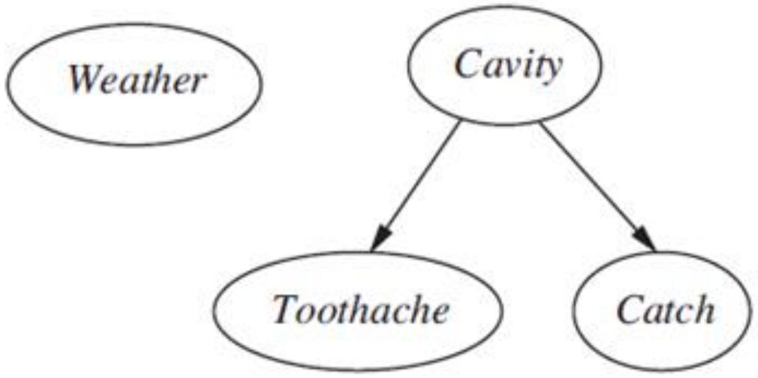
Chapter 14 (AIAMA)

Probabilistic Reasoning

Sukarna Barua
Assistant Professor, CSE, BUET

Bayesian Network

- A Bayesian network (BN) is a directed graph where each node is a random variable and edges show relationship (cause and effect) between the random variables.



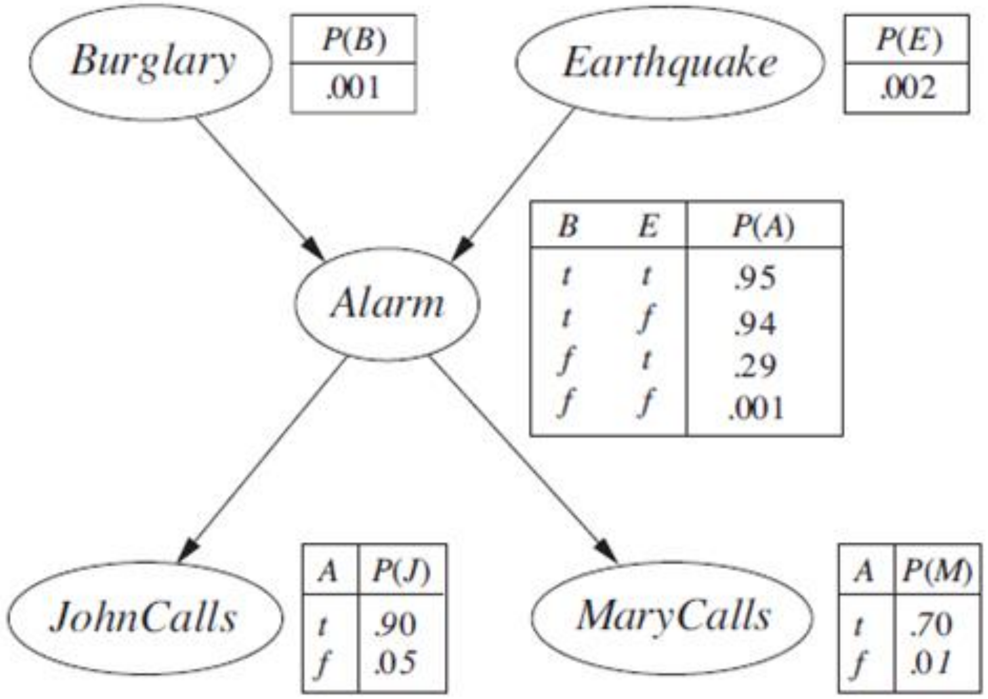
A simple Bayesian network in which *Weather* is independent of the other three variables and *Toothache* and *Catch* are conditionally independent, given *Cavity*

Bayesian Network: Example

- You have a new burglar alarm installed at home. It is fairly 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 nearly 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 rather loud music and often misses the alarm altogether. Given the evidence of who has or has not called, we would like to estimate the probability of a burglary.

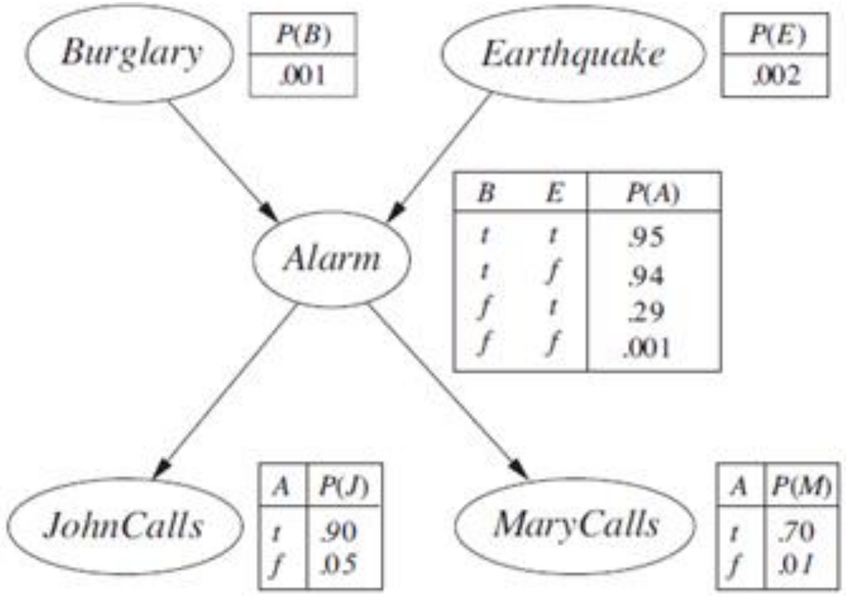
Bayesian Network: Example

- Bayesian network for burglary alarm Example



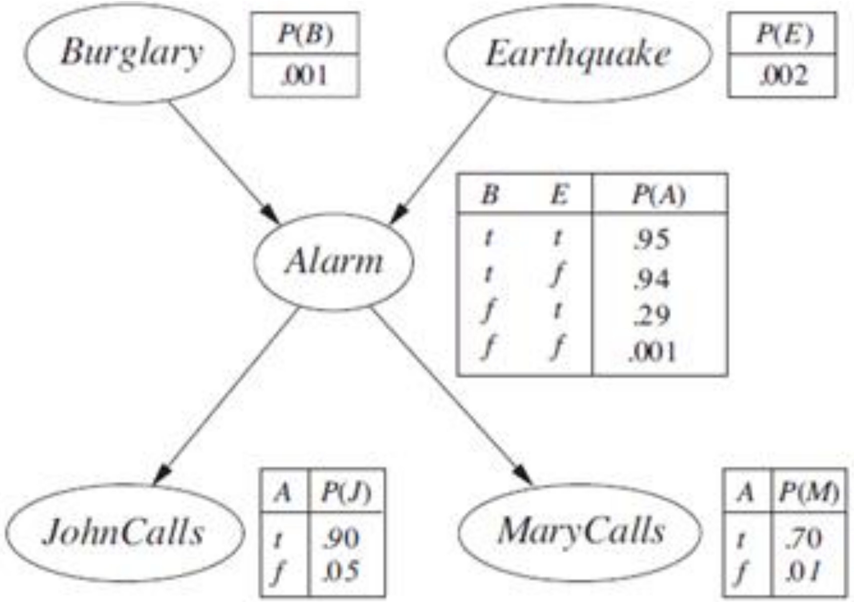
Bayesian Network: Some Observations

- Burglay and Earthquake directly affect the probability of Alarm going off
- John and Marry calls depending only on the Alarm (They know nothing about Burglary or Earthquake)



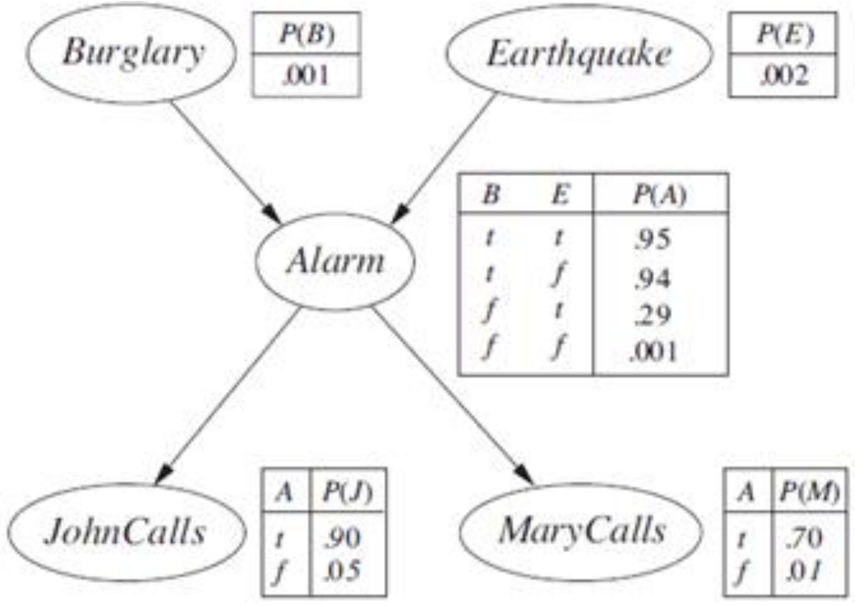
Bayesian Network: Example

- In a Bayesian network, each node is also annotated with **quantitative probability information** (conditional probability distribution).



Bayesian Network: Example

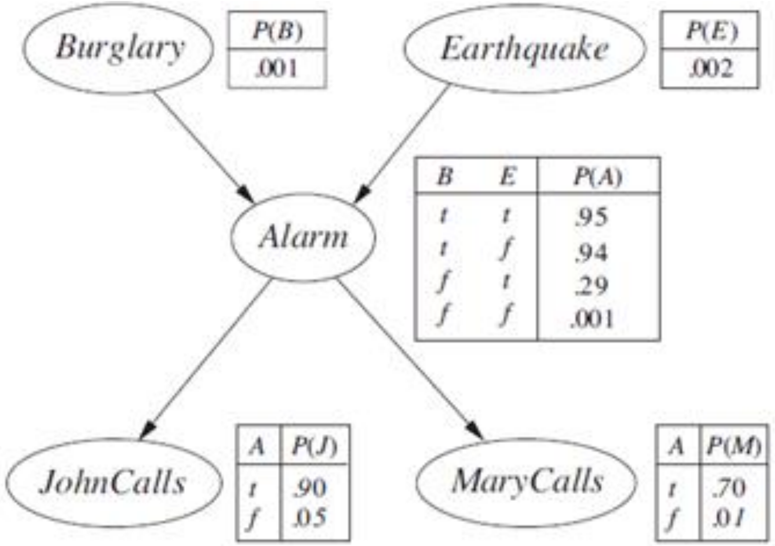
- Conditional probability distribution specified by a conditional probability table (CPT)
  - Each row specifies the probability of all values (of the node) given a combination of values of the parent nodes
  - Each row must sum to 1
  - For binary valued nodes, only one (1) probability value may be specified for each row. [as probabilities sum to 1]



Bayesian Network: Size of CPTs

- A Boolean node have k Boolean parents. What will be the size of CPT for this node?

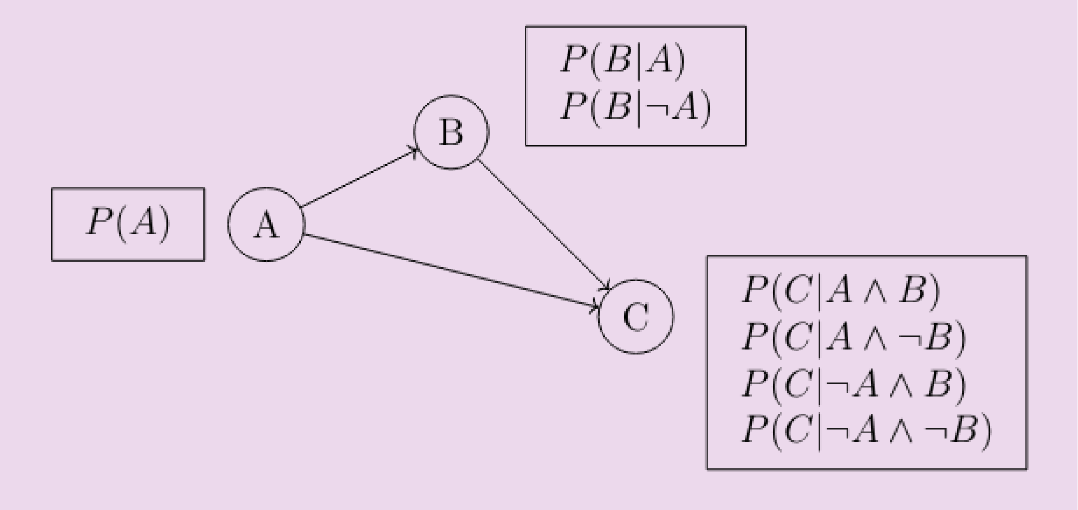
  Number of rows
  = possible combination of parent values
  = 2^k
  Number of entries
  = $(2^k)*1$ [Only one value need to be stored by row]



- How many probabilities are needed to represent the entire Bayesian Network? [Answer: 10]

Bayesian Network: Size of CPTs

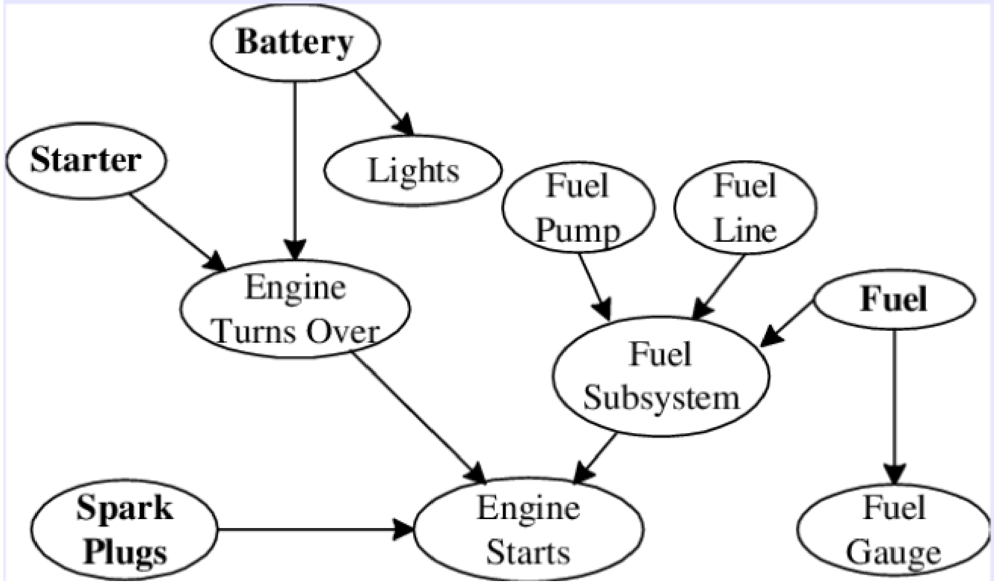
- How many probabilities are needed to represent the Bayesian Network? Assume all are Boolean variables. [Answer: 7]



- How many will be needed if B and C are conditionally independent given A?

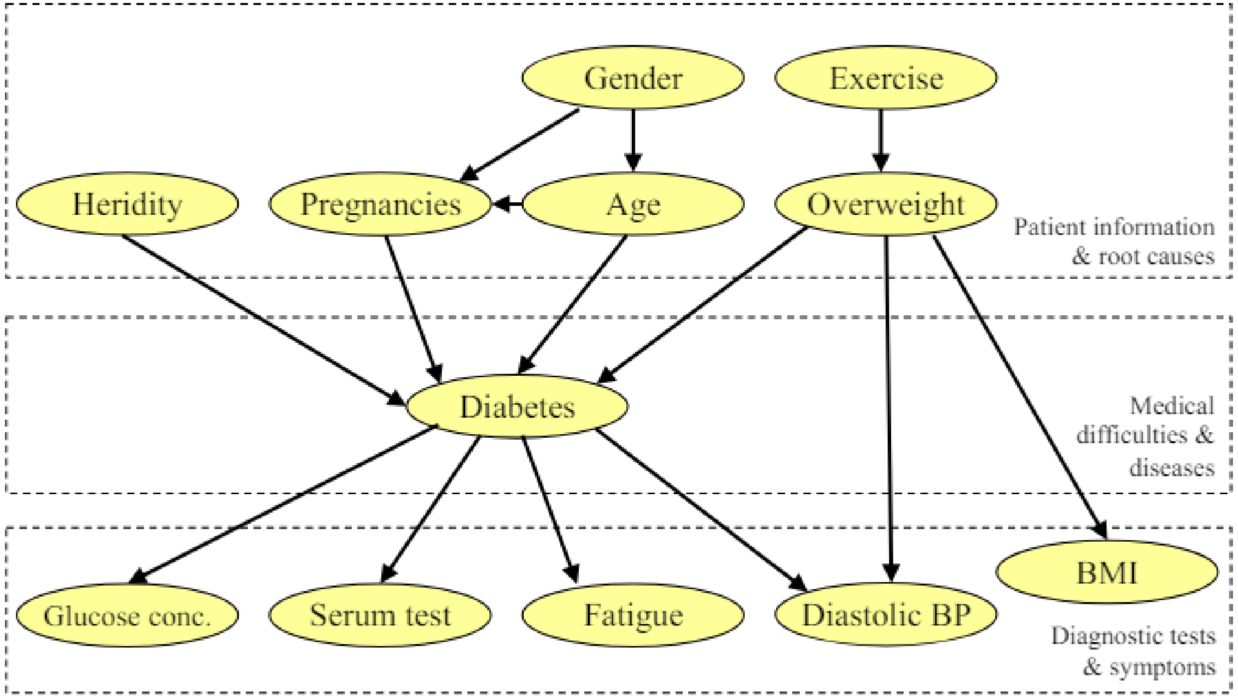
Bayesian Network: More Example

- A Bayesian Network for car diagnostic system.
  - we may have problems in a particular part of the car, and that may be caused by many different reasons.
  - The battery might cause the lights to have different problems or it may cause the engine to have problems, and that in turn may affect whether or not the engine can start.



Bayesian Network: More Example

- A Bayesian Network for diabetes diagnosis.
  - Symptoms and tests
  - Patient information: the patient's age, weight, medical history, genetics, and so on.



Bayesian Network: Full Specification

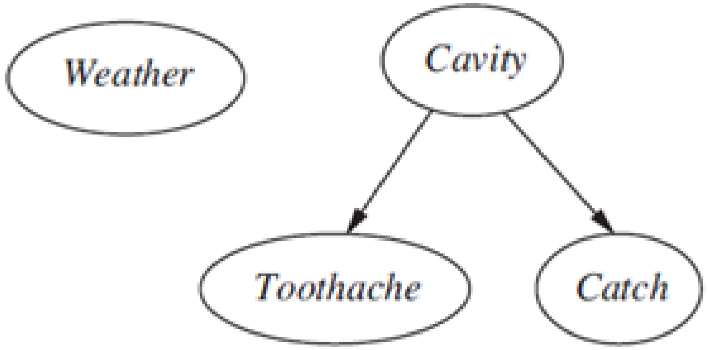
1. Each node corresponds to a random variable, which may be discrete or continuous.
2. Set of directed links or arrows connects pairs of nodes.
   - If there is an arrow from node X to node Y , X is said to be a parent of Y. The graph has no directed cycles and hence is a directed acyclic graph, or DAG.
3. Each node X_i has a conditional probability distribution $P(X_i \,|Parents(X_i))$ that quantifies the effect of the parents on the node.

Bayesian Network: Conditional Independence

- Implicitly defines conditional independence of random variables

  - *Weather* is Independent of other variables

  - *Toothache* and *Catch* are conditionally independent given *Cavity*



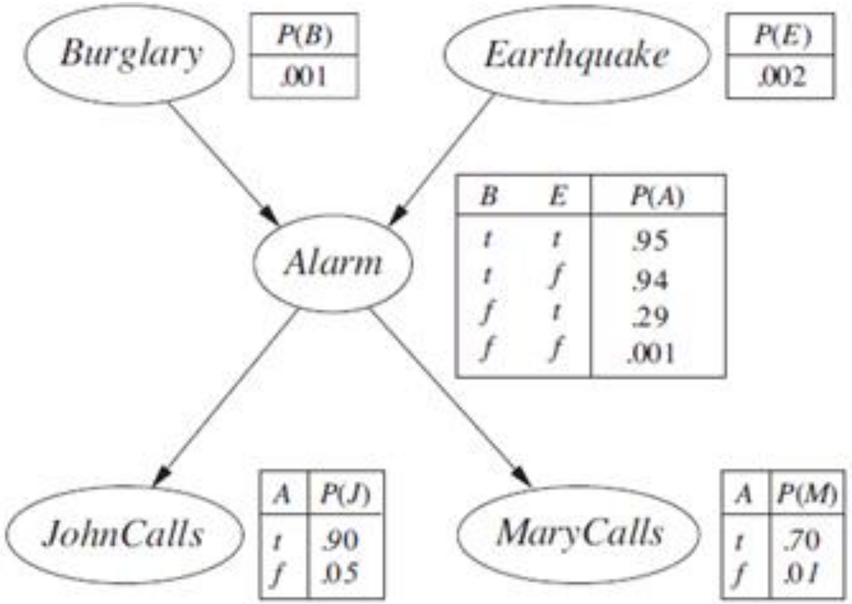
Bayesian Network: Conditional Independence

- Implicitly defines conditional independence of random variables

  - JohnCalls is Independent of MarryCalls given Alarm

  - MarryCalls is independent of EarthQuake given Alarm

  - Is Burglary independent of Earthquake? [Yes, Why?]

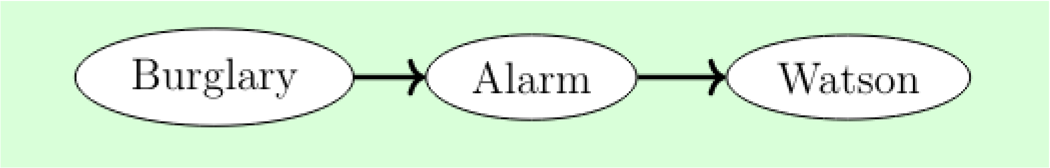


Bayesian Network: Conditional Independence

- Implicitly defines conditional independence of random variables

  - Are Burglary and Watson independent? [No]

  - Are Burglary and Watson independent given Alarm? [Yes]

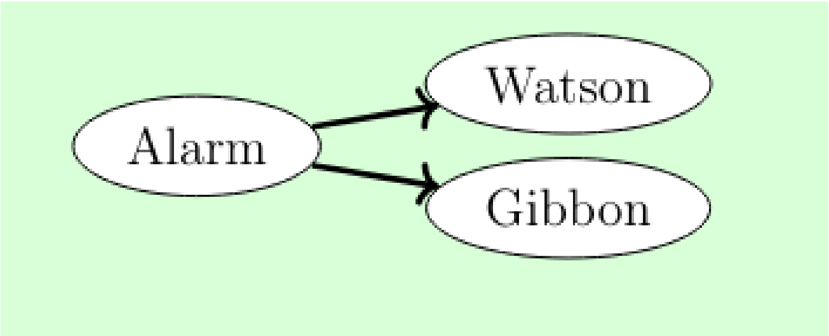


Bayesian Network: Conditional Independence

- Implicitly defines conditional independence of random variables

  - Are Watson and Gibbon independent? [No]

  - If Watson is calling, it is more likely that Alarm is off, which means that Gibbon is more likely to call



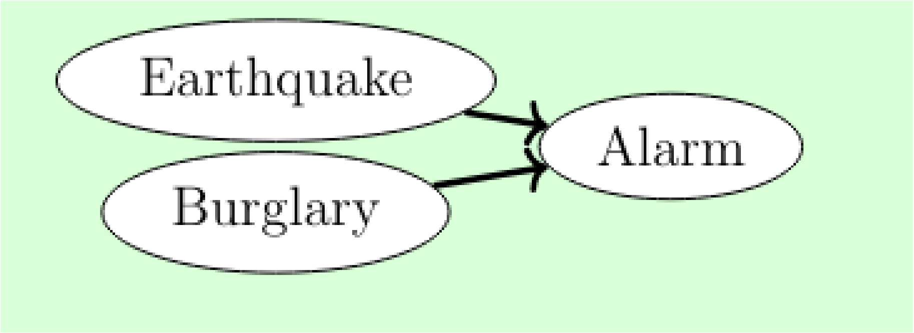
Bayesian Network: Conditional Independence

- Implicitly defines conditional independence of random variables

    - Are Earthquake and Burglary independent [Yes]

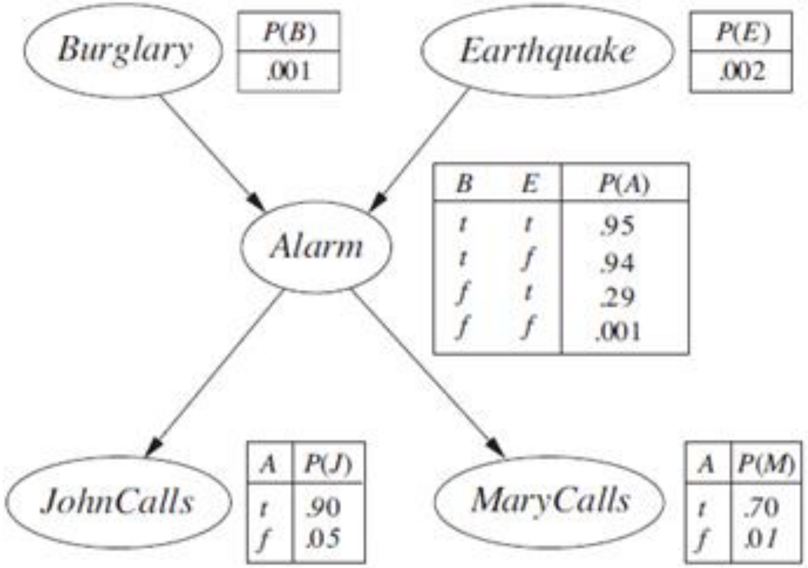
    - Are Earthquake and Burglary independent given Alarm [No]

    [If Earthquake is happening, then it is less likely that the Alarm is caused by Burglary.]



Bayesian Network: Full Joint Distribution

- *Bayesian network + conditional probability distribution* for each variable, given its parents specify full joint distribution for all the variables.



Bayesian Network: Full Joint Distribution

- Given a Bayesian Network with variables/nodes $x_1, x_2, \dots, x_n$
- Find the joint probability $P(x_1, x_2, \dots, x_n)$
- We can show that:

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

Bayesian Network: Full Joint Distribution

- If nodes are ordered such that all parent nodes of $x_i \in \{x_{i-1} \dots, x_1\}$, then the full joint distribution can be written as:

$$P(x_1, x_2, \dots, x_n) = \prod_{i=1}^{n} P(x_i | Parent(X_i)$$

- We just need to find a topological sorting of the nodes for the above equation to become correct!

Bayesian Network: Compute Full Joint Distribution

- Calculate the probability that the alarm has sounded, but neither a burglary nor an earthquake has occurred, and both John and Mary call.

We have following values:

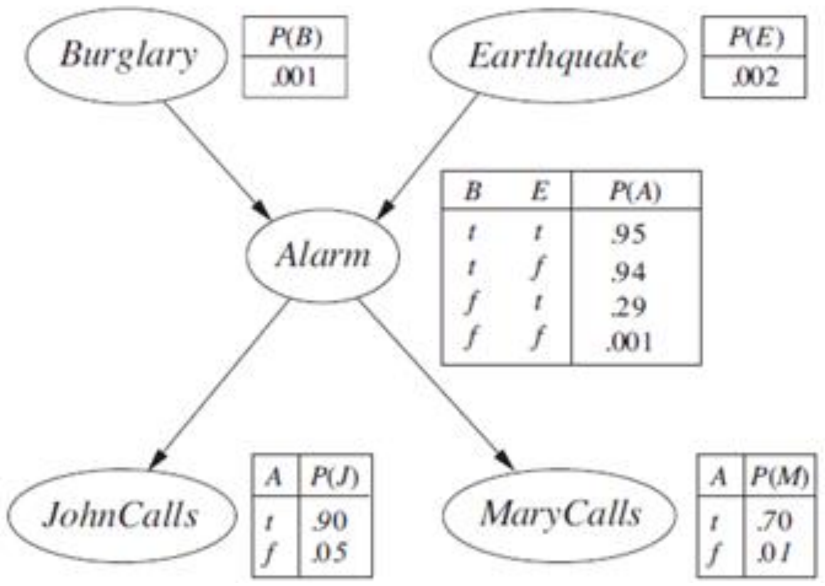
a = alarm sounds

$\neg b$ = no burglary

$\neg e$ = no earthquake

j = John calls

m = Marry calls

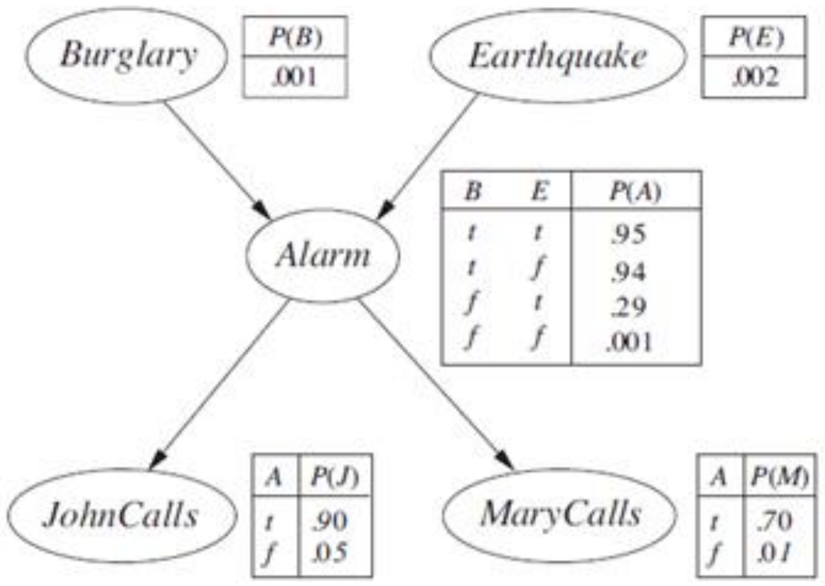


Bayesian Network: Compute Full Joint Distribution

- Calculate the probability that the alarm has sounded, but neither a burglary nor an earthquake has occurred, and both John and Mary call.

- We need to find: $P(a, \neg b, \neg e, j, m)$

- We can compute the full joint distribution using CPTs.



Bayesian Network: Compute Full Joint Distribution

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

$= P(j, m, a, \neg b, \neg e)$ [rearrange as per topological sorting]

$= P(j|m, a, \neg b, \neg e)P(m, a, \neg b, \neg e)$ $[P(a, b) = P(a|b)P(b)]$

$= P(j|a)P(m, a, \neg b, \neg e)$ [j is independent of other variables given a]

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

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

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

[b and e are independent of each other]

$= 0.9 * 0.7 * 0.001 * 0.999 * 0.998$

$= 0.000628$