Uncertain Knowledge Representation in AI

Dr. Singara Singh Kasana

Associate Professor

Computer Science and Engineering Department

Thapar Institute of Engineering and Technology

Patiala, Punjab

Probabilistic Reasoning

Probabilistic reasoning is a way of knowledge representation where we apply the concept of probability to indicate the uncertainty in knowledge. In probabilistic reasoning, we combine probability theory with logic to handle the uncertainty.

In the real world, there are lots of scenarios, where the certainty of something is not confirmed, such as

- It will rain today.
- Behavior of someone for some situations
- A match between two teams or two players

These are probable sentences for which we can assume that it will happen but not sure about it, so there is the need of probabilistic reasoning.

Independent Events

Each event is **not affected** by any other events.

Example: Tossing a coin

Each toss of a coin is a perfect isolated thing.

What it did in the past will not affect the current toss.

The chance is simply 1-in-2, or 50%, just like ANY toss of the coin.

So each toss is an **Independent Event**.

Dependent Events

Events can also be "dependent" ... which means they **can be affected by previous events**

Example: 2 blue and 3 red balls are in a bag. What are the chances of getting a red ball?

The chance is **3 in 5**
But after taking one ball out, the chances change!
So the next time:

- if we got a **red** ball before, then the chance of a **blue** ball next is **2 in 4**

- if we got a **blue** ball before, then the chance of a **red** ball next is **1 in 4**

Replacement

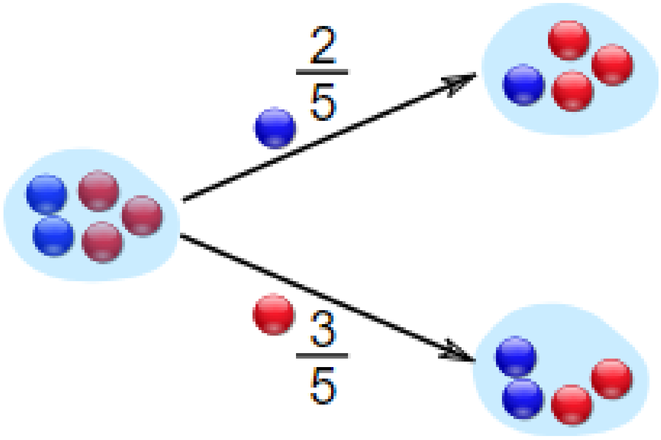
Note: if we **replace** the ball in the bag each time, then the chances do **not** change and the events are independent:

With Replacement: the events are **Independent** (the chances don't change)

Without Replacement: the events are **Dependent** (the chances change)

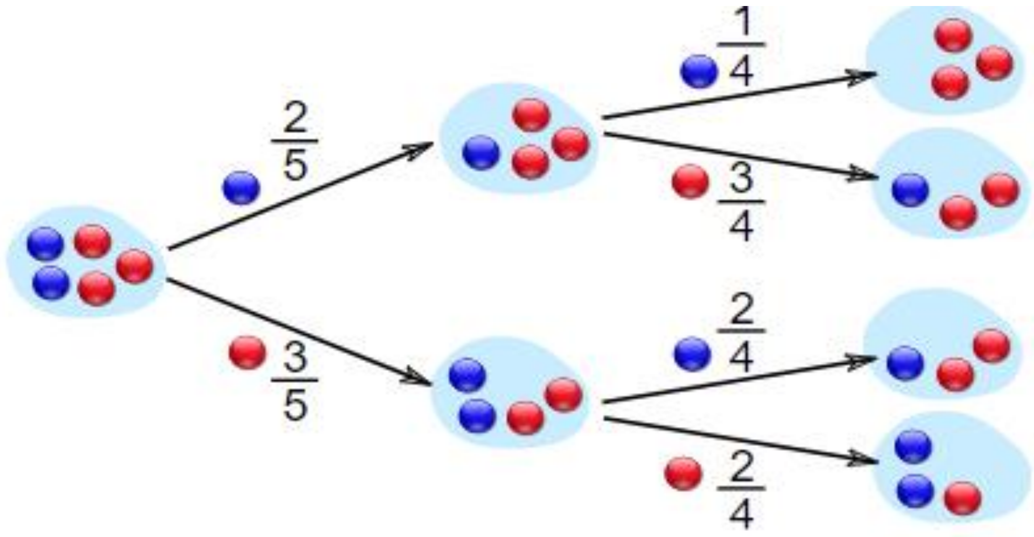
Tree Diagram

There is a 2/5 chance of pulling out a blue ball, and a 3/5 chance for a red ball:



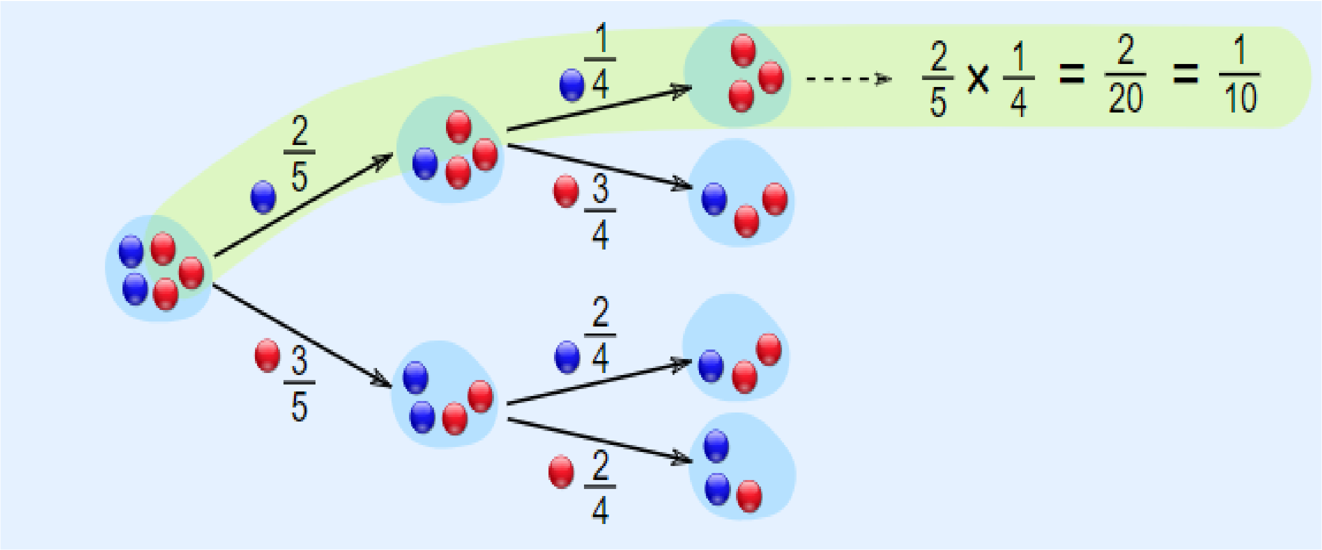
Tree Diagram- Continued

We can go one step further and see what happens when we pick a second ball:



Now we can answer questions like **"What are the chances of drawing 2 blue balls?"**

Answer: it is a **2/5 chance** followed by a **1/4 chance**:



Event A is "get a Blue ball first" with a probability of 2/5:

P(A) = 2/5

Event B is "get a Blue ball second" ... but for that we have 2 choices:

If we got a **Blue ball first** the chance is now **1/4**

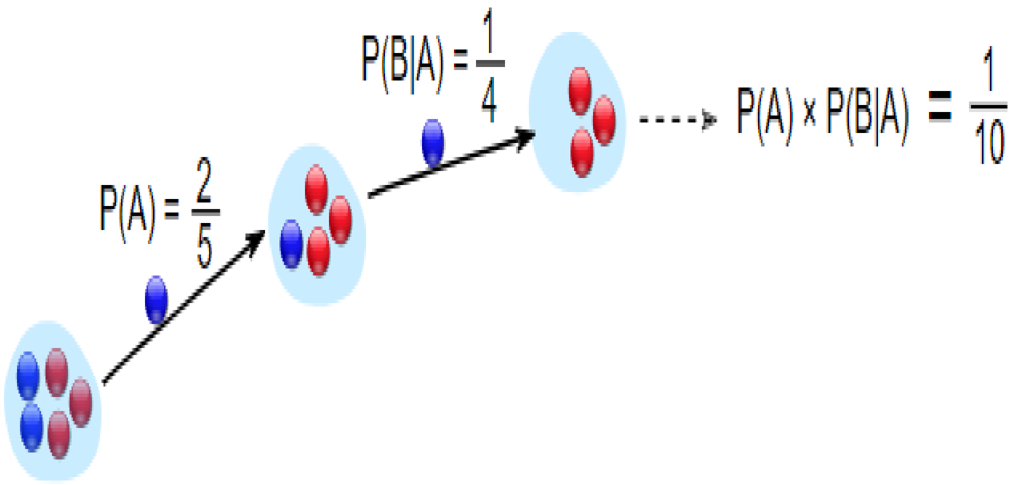
If we got a **Red ball first** the chance is now **2/4**

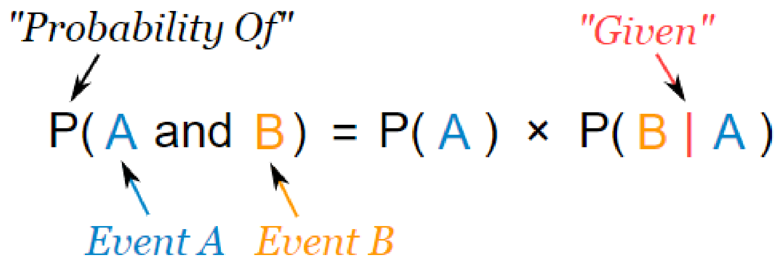
So we have to say **which one we want**, and use the symbol "|" to mean "given":

P(B|A) means "Event B **given** Event A"

In other words, event A has already happened, now what is the chance of event B?

P(B|A) is also called the "Conditional Probability" of B given A.





Question: Drawing 2 Queens from a Deck

Event A is drawing a Queen first, and **Event B** is drawing a second Queen.

P(A and B) = P(A) x P(B|A) = (4/52) x (3/51) = 12/2652 = **1/221**

Conditional Probability

- **Conditional probability** is defined as the likelihood of occurrence of an event, based on the occurrence of some other event.
- Conditional **Probability** of event B **given** A is represented as P(B|A).

- P(B|A) means "Probability of event B **given** event A" In other words, event A has already happened, now what is the chance of event B?

- $P(B|A) = P(A \cap B) / P(A)$ when $P(A)>0$

$$P(A \cap B) = P(A) \times P(B|A)$$

Probability of event A and event B equals to probability of event A times probability of event B given event A.

$$P(B|A) = P(A \cap B) / P(A)$$

If A and B are independent events then $P(A \cap B) = P(A) \times P(B)$ then

$$P(B|A) = P(A) \times P(B) / P(A)$$
$$= P(B)$$

Example: If 70% of your friends like Pizza, and 35% like Pizza AND like Cold Drink. What percent of those who like Pizza also like Cold Drink?

P(Pizza and Cold Drink) = 35/100 = 0.35
P(Pizza) = 70/100 = 0.70

P(Cold Drink |Pizza) = P(Pizza and Cold Drink) / P(Pizza)
P(Cold Drink |Pizza) = 0.35 / 0.7 = 0.5

To calculate the probability of the intersection of more than two events, the conditional probabilities of *all* of the preceding events must be considered. In the case of three events, *A*, *B*, and *C*, the probability of the intersection

$$P(A \cap B \cap C) = P(A) \times P(B|A) \times P(C|A \cap B).$$

Chain rule for conditional probability:

$$P(A_1 \cap A_2 \cap \cdots \cap A_n) = P(A_1)P(A_2|A_1)P(A_3|A_2 \cap A_1) \cdots P(A_n|A_{n-1} \cap A_{n-2} \cdots \cap A_1)$$

Conditional probability itself is a probability measure, so it satisfies probability axioms.

Axiom 1: For any event A, $P(B|A) \geq 0$.

Axiom 2: Conditional probability of B given B is 1, i.e., $P(B|B)=1$.

Axiom3: If $A_1, A_2, A_3, \cdots$ are disjoint events, then

$P(B_1 \cup B_2 \cup B_3 \cdots |A)=P(B_1|A)+P(B_2|A)+P(B_3|A)+\cdots$.

Special cases of conditional probability:

Case1: When A and B are disjoint. In this case, $A \cap B = \emptyset$, so

$$P(B|A) = P(A \cap B)/P(A) = P(\emptyset)/P(A) = 0.$$

Case 2: When B is a subset of A,

$$P(B|A) = P(A \cap B)/P(A) = P(B)/P(A)$$

Case 3: When A is a subset of B: In this case $A \cap B = A$, so

$$P(B|A) = P(A \cap B)/P(A) = P(A)/P(A) = 1$$

Bayes Theorem

$$\mathbf{P(B|A) = P(A|B)\,P(B) \,/\, P(A)}$$

P(B|A) is the likelihood of B given A
P(A) is the prior probability of A
P(B) is the marginal probability of B

Example

A desk lamp produced by The Luminar Company was found to be defective (*D*). There are three factories (*A*, *B*, *C*) where such desk lamps are manufactured. A Quality Control Manager (QCM) is responsible for investigating the source of defects. This is what the QCM knows about the company's desk lamp production and the possible source of defects:

Factory	% of total production	Probability of defective lamps
A	$0.35 = P(A)$	$0.015 = P(D \mid A)$
B	$0.35 = P(B)$	$0.010 = P(D \mid B)$
C	$0.30 = P(C)$	$0.020 = P(D \mid C)$

The QCM would like to answer the following question: If a randomly selected lamp is defective, what is the probability that the lamp was manufactured in factory C, when **P(D), the probability that a lamp manufactured by The Luminar Company is defective, is 0.01475.**

Solution:

The probability that a lamp was manufactured in factory A given that it is defective is:

$$P(A|D)=P(A\cap D)/P(D)=P(D|A)\times P(A)/P(D)=(0.015)\ \times(0.35)/0.01475=0.356$$

And, the probability that a lamp was manufactured in factory B given that it is defective is:

$$P(B|D)=P(B\cap D)/P(D)=P(D|B)\times P(B)/P(D)=(0.01)\ \times(0.35)/0.01475=0.237$$

$$P(C|D) = 0.407$$
$$P(B|D) = 0.237$$
$$P(A|D) = 0.356$$

Generalization of Bayes Rule

$$P(B_i \mid A) = \frac{P(A \mid B_i)P(B_i)}{P(A \mid B_1)P(B_1) + P(A \mid B_2)P(B_2) + \cdots + P(A \mid B_k)P(B_k)}$$

- where:

  B_i = i^{th} event of k mutually exclusive and collectively exhaustive events

  A = new event that might impact $P(B_i)$

Bayes Theorem

P(B|A) = (P(A|B) *P(B)) / P(A)

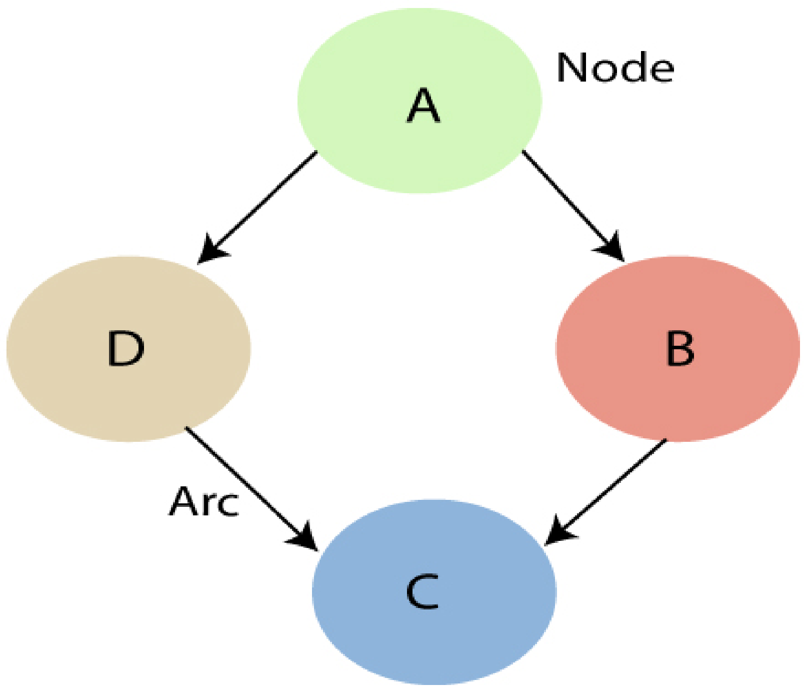
- Basically, we need to find probability of event B, given the event A is true. Event A is also termed as **evidence**.
- P(B) is the **priori** of B (the prior probability, i.e. Probability of event before evidence is seen). The evidence is an attribute value of an unknown instance(here, it is event B).
- P(B|A) is a posteriori probability of B, i.e. probability of event after evidence is seen.

Bayesian Network

- A Bayesian network is a probabilistic graphical model which use a directed acyclic graph to represent a set of variables and their conditional dependencies.
- These networks are probabilistic in nature, because they use concepts of probability theory.
- These are very useful in understanding the dependency among events and assigning probabilities to them thus ascertaining how probable or what is the chance of occurrence of one event given the other.

Bayesian Network consists of two parts:

1. **Directed Acyclic Graph**
2. **Table of conditional probabilities.**



- Each node corresponds to the random variables, and a variable can be continuous or discrete.
- Arc or directed arrows represent the relationship or conditional probabilities between random variables. These directed links or arrows connect the pair of nodes in the graph.
- These links represent that one node directly influence the other node, and if there is no directed link that means that nodes are independent with each other

- In the diagram(shown in previous slide), A, B, C, and D are random variables, represented by the nodes of the network graph.
- If we consider node B, which is connected with node A by a directed arrow, then node A is the parent of Node B.
- Node C is independent of node A.

Joint probability distribution

Joint probability is the likelihood of more than one event occurring at the same time.

Joint probability for two events, A and B, is expressed mathematically as P(A,B)and is calculated by multiplying the probability, P(A), of event A by the probability, P(B), of event B.

For example, suppose a statistician wishes to know the probability that the number five will occur on two dices when they are rolled at the same time. Since each die has six possible outcomes, the probability of a five occurring on each die is 1/6 or 0.1666.

$P(A)=0.1666$

$P(B)=0.1666$

$P(A,B)=0.1666 \times 0.1666=0.02777$

This means the joint probability that a five will be rolled on both dice at the same time is 0.02777.

Conditions for Joint Probability

i. One condition is that events X and Y must happen at the same time. Example: Throwing two dice simultaneously.
ii. The other is that events X and Y must be independent of each other. That means the outcome of event X does not influence the outcome of event Y.

Example: Rolling two Dice.

Conditional Probability

The conditional probability of an event B is the probability that the event will occur given the knowledge that an event A has already occurred. It is denoted by $P(B|A)$.

The joint probability of two dependent events becomes $P(A \text{ and } B) = P(A)P(B|A)$

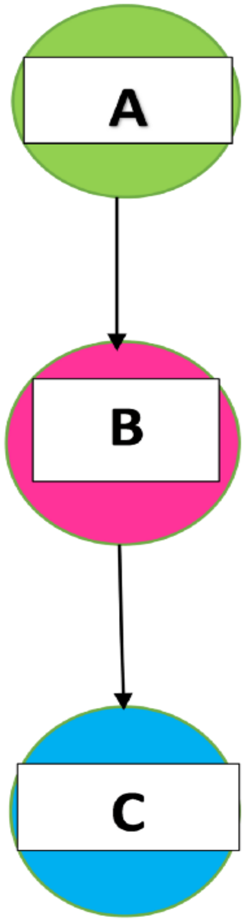
Joint probability distribution

If we have n variables $x1, x2, x3, \ldots, xn$.

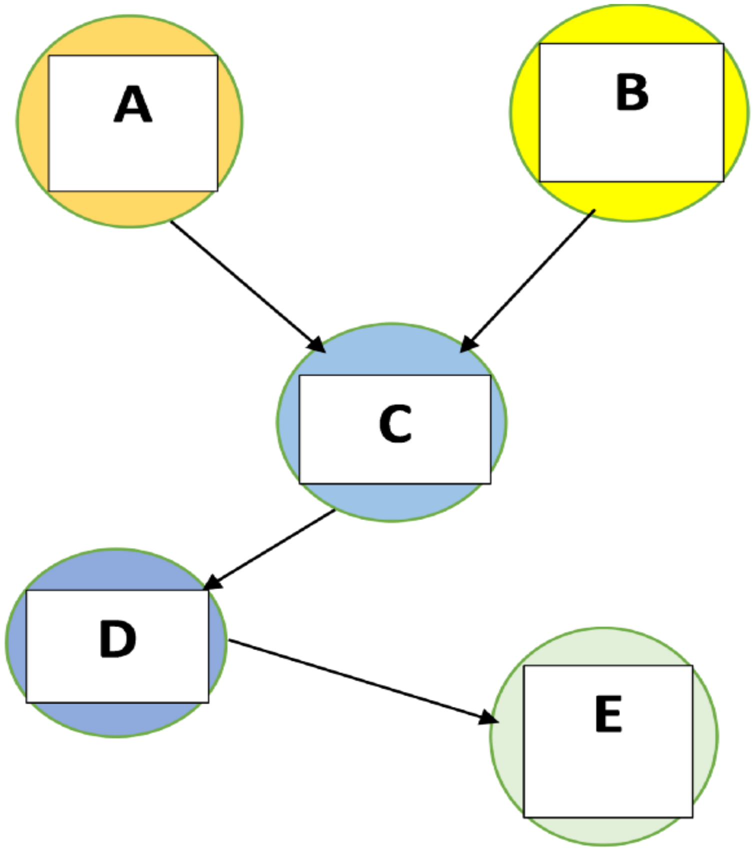
then $P[x1, x2, x3, \ldots, xn]$ can be written as the following way in terms of the joint probability distribution.

$$= P[x1 \mid x2, x3, \ldots, xn]P[x2, x3, \ldots, xn]$$

$$= P[x1 \mid x2, x3, \ldots, xn]P[x2 \mid x3, \ldots, xn] \ldots P[xn-1 \mid xn]P[xn].$$



$$P(A,B,C) = P(C|B).P(B|A).P(A)$$



$$P(A,B,C,D,E) = P(E|D).P(D|C).P(C|A,B).P(B).P(A)$$

Example: Ram installed a alarm system at his home to detect burglary. The alarm reliably responds at detecting a burglary but also responds for fire. Ram has two neighbors- Krishan and Sonia, who have taken a responsibility to inform Ram at work when they hear the alarm. Krishan always calls Ram when he hears the alarm, but sometimes he got confused with the phone ringing and calls at that time too. On the other hand, Sonia likes to listen to high music, so sometimes she misses to hear the alarm. **Calculate the probability that alarm has sounded, but there is neither a burglary, nor an fire catching occurred, and Krishan and Sonia both called the Ram.**

List of all events occurring in this network:

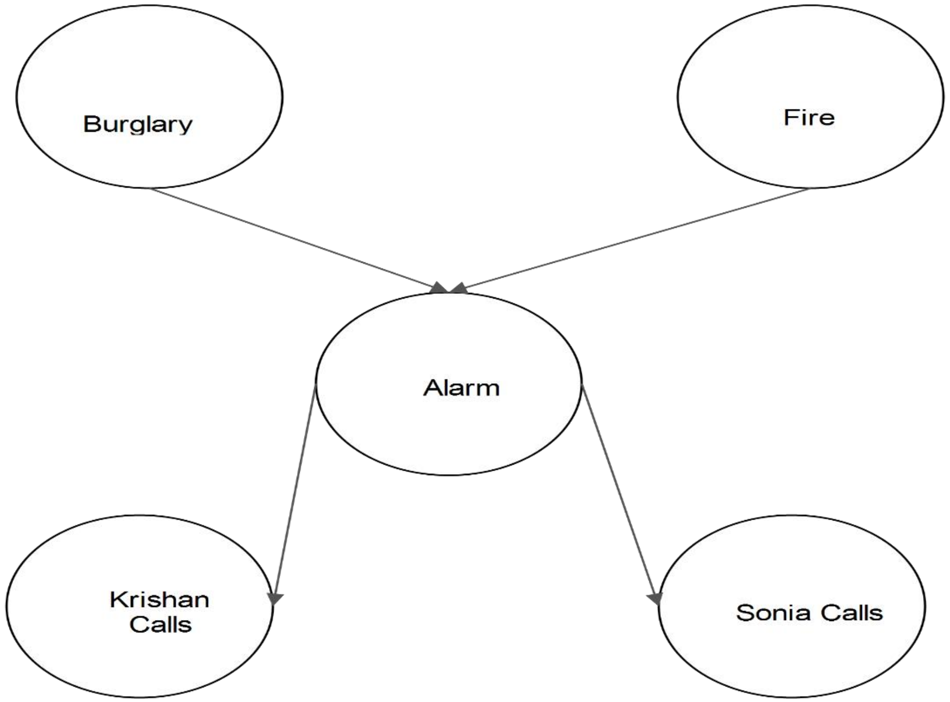
- Burglary (B)
- Fire Catching(F)
- Alarm(A)
- Krishan Calls(K)
- Sonia calls(S)

Burglary Probabilities

T	0.002
F	0.998

Fire Probabilities

T	0.001
F	0.999



Burglary	Fire	P(Alarm=T)	P(Alarm =F)
T	T	0.94	0.06
T	F	0.95	0.05
F	T	0.69	0.31
F	F	0.001	0.999

Alarm	P(Krishna calls)	P(Krishna does not call)
T	0.91	0.09
F	0.05	0.95

Alarm	P(Sonia calls)	P(Sonia does not call)
T	0.75	0.25
F	0.02	0.98

Let's take the observed probability for the Burglary and earthquake component:

P(B= True) = 0.002, which is the probability of burglary.

P(B= False)= 0.998, which is the probability of no burglary.

P(F= True)= 0.001, which is the probability of a catching fire

P(F= False)= 0.999, Which is the probability of not catching.

Conditional probability table for Alarm A:

The Conditional probability of Alarm A depends on Burglar and fire catching:

B	F	P(A= True)	P(A= False)
True	True	0.94	0.06
True	False	0.95	0.05
False	True	0.69	0.31
False	False	0.001	0.999

Conditional probability table for Krishan Calls:
The Conditional probability of Krishan that he will call depends on the probability of Alarm.

A	P(K= True)	P(K= False)
True	0.91	0.09
False	0.05	0.95

Conditional probability table for Sonia Calls:
The Conditional probability of Sonia that she calls is depending on its Parent Node "Alarm."

A	P(S= True)	P(S= False)
True	0.75	0.25
False	0.02	0.98

From the formula of joint distribution, we can write the problem statement in the form of probability distribution:

$P(S, K, A, \neg B, \neg F) = P(S|A) * P(K|A) * P(A|\neg B \wedge \neg F) * P(\neg B) * P(\neg F)$.

$= 0.75 * 0.91 * 0.001 * 0.998 * 0.999$

$= 0.00068045$.

Thanks