2.3 BAYESIAN NETWORKS

Network representation of knowledge is used to exhibit the interdependencies which exist between related pieces of knowledge.

Bayesian network is a probabilistic graphical model that represents a set of random variables and their conditional dependencies via a directed acyclic graph (DAG).

For example, a bayesian network could represent the probabilistic relationships between diseases and symptoms. Given symptoms, the networks can be used to compute the probabilities of the presence of various diseases.

- Bayesian networks are DAG's whose node represents random variables, an observable quantities. Edges represents conditional dependencies.
- Nodes that are not connected indicate the variables that are conditionally independent.
- Each node is associated with a probability function, the input to the probability function is the set of values for the nodes parent variable. Output is the probability of the variable represented by the node.
- Network representations which depicts the degrees of belief of proportions and the casual dependencies that exist between them.
- Inference in a network leads to propagating the probabilities of given and related information through the network to one or more conclusion nodes.

Bayesian belief networks :

Inference process, problem domain is with a network of nodes which represent propositional variables x_i, connected by arcs represents casual influences or dependencies among the nodes.

The strength of the influences are quantified by conditional probabilities of each variable.

Example :

To represent casual relationships between the proportional variables x1, x2,..., x6 as shown in the below figure, Joint probability P(x1,x2,...x6) is the product of conditional probabilities.

$$P(x_1,x_2,....,x_6) = P(x_6/x_5).P(x_5/x_2,x_3).P(x_4/x_1,x_2).P(x_3/x_1).P(x_2/x_1).P(x_1).$$

In general for each variable Xi, we can write the equation as:

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

An inference engine can use network to maintain and propagate beliefs.

When new information is received, the effects can be propagated throughout the network.

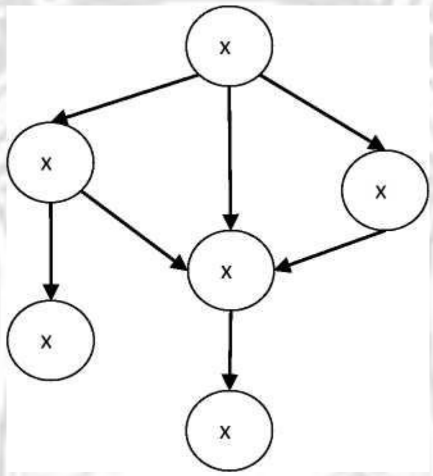


Fig 3.14 Example of Bayesian belief network

Example: Harry installed a new burglar alarm at his home to detect burglary. The alarm reliably responds at detecting a burglary but also responds for minor earthquakes. Harry has two neighbors David and Sophia, who have taken a responsibility to inform Harry at work when they hear the alarm. David always calls Harry when he hears the alarm, but sometimes he got confused with the phone ringing and calls at that time too. On the other hand, Sophia likes to listen to high music, so sometimes she misses to hear the alarm. Here we would like to compute the probability of Burglary Alarm.

Problem:

Calculate the probability that alarm has sounded, but there is neither a burglary, nor an earthquake occurred, and David and Sophia both called the Harry.

Solution:****

- The Bayesian network for the above problem is given below. The network structure is showing that burglary and earthquake is the parent node of the alarm and directly affecting the probability of alarm's going off, but David and Sophia's calls depend on alarm probability.

- The network is representing that our assumptions do not directly perceive the burglary and also do not notice the minor earthquake, and they also not confer before calling.
- The conditional distributions for each node are given as conditional probabilities table or CPT.
- Each row in the CPT must be sum to 1 because all the entries in the table represent an exhaustive set of cases for the variable.
- In CPT, a boolean variable with k boolean parents contains 2^K probabilities. Hence, if there are two parents, then CPT will contain 4 probability values

List of all events occurring in this network:

- **Burglary (B)**
- **Earthquake(E)**
- **Alarm(A)**
- **David Calls(D)**
- **Sophia calls(S)**

We can write the events of problem statement in the form of probability: **P[D, S, A, B, E]**, can rewrite the above probability statement using joint probability distribution:

P[D, S, A, B, E]= P[D | S, A, B, E]. P[S, A, B, E]

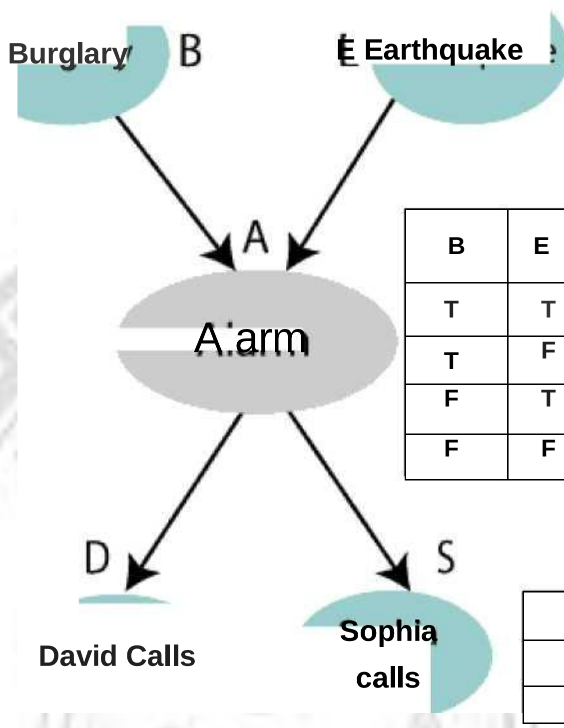
=P[D | S, A, B, E]. P[S | A, B, E]. P[A, B, E]

= P [D| A]. P [S| A, B, E]. P[A, B, E]

= P[D | A]. P[S | A]. P[A| B, E]. P[B, E]

= P[D | A]. P[S | A]. P[A| B, E]. P[B |E]. P[E]

T	0,002
F	0.998



T	o.oOi
F	0.999

B	E	P(A=T)	P(A=F)
T	T	0.94	0.06
T	F	0.95	0.04
F	T	0.69	0.69
F	F	0.999	0.999

A	P{D=T)	P(D=F)
T	0.91	0.09
F	0.05	0.95

A	P<S=T)	P(S=F)
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(E= True)= 0.001, which is the probability of a minor earthquake

P(E= False)= 0.999, Which is the probability that an earthquake not occurred.

We can provide the conditional probabilities as per the below tables:

Conditional probability table for Alarm A:

The Conditional probability of Alarm A depends on Burglar and earthquake:

B	E	P(A= True)	P(A= False)

True	True	0.94	0.06
True	False	0.95	0.04
False	True	0.31	0.69
False	False	0.001	0.999

Conditional probability table for David Calls:

The Conditional probability of David that he will call depends on the probability of Alarm.

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

Conditional probability table for Sophia Calls:

The Conditional probability of Sophia 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, D, A, -B, -E) = P (S|A) *P (D|A)*P (A|-B ^{A} -E) *P (-B) *P (-E).

= 0.75* 0.91* 0.001* 0.998*0.999

= 0.00068045.

Hence, a Bayesian network can answer any query about the domain by using Joint distribution.

The semantics of Bayesian Network:

There are two ways to understand the semantics of the Bayesian network, which is given below:

1. **To understand the network as the representation of the Joint probability distribution.**

It is helpful to understand how to construct the network.

2. **To understand the network as an encoding of a collection of conditional independence statements.**