SNS COLLEGE OF ENGINEERING

Kurumbapalayam(Po), Coimbatore – 641 107
Accredited by NAAC-UGC with ‘A’ Grade
Approved by AICTE, Recognized by UGC & Affiliated to Anna University, Chennai

Department of Information Technology
Course Name – Software Engineering

II Year / III Semester

Unit-3 Reasoning Under Uncertainity

Beliefs (Bayesian Probabilities)

• We use probability to describe uncertainty due to:
– Laziness: failure to enumerate exceptions, qualifications etc.
– Ignorance: lack of relevant facts, initial conditions etc.
– True randomness? Quantum effects? ...
• Beliefs (Bayesian or subjective probabilities) relate propositions to one's current state of knowledge
– E.g. P(A(25)| no reported accident) = 0.1 These are not assertions about the world / absolute truth
23.Oct.2022
• Beliefs change with new evidence:
– E.g. P(A(25)| no reported accident, 5am) = 0.2
• This is analogous to logical entailment:
KB # α means that α is true given the KB, but may not be true in general.

Making Decisions Under Uncertainty

• Suppose I believe the following:

P(A(25) gets me there on time | ...) = 0.04 P(A(90) gets me there on time | ...) = 0.70

P(A(120) gets me there on time | ...) = 0.95

P(A(1440) gets me there on time |...) = 0.9999

• Which action should I choose?

Making Decisions Under Uncertainty

• Suppose I believe the following:

P(A(25) gets me there on time | ...) = 0.04

P(A(90) gets me there on time | ...) = 0.70

P(A(120) gets me there on time | ...) = 0.95

P(A(1440) gets me there on time | ...) = 0.9999

• Which action should I choose?

– Depends on my preferences for missing flight vs. airport cuisine, etc.

– Utility theory is used to represent and infer preferences.

– Decision theory = utility theory + probability theory

Bayes Rule

• Bayes rule is another alternative formulation of the product rule:

P(A|B) = P(B|A)P(A) P(B)

• The complete probability formula states that:

P(A) = P(A|B)P(B) + P(A|¬B)P(¬B) or more generally,

P(A) = ! i P(A|bi)P(bi), where bi form a set of exhaustive and mutually exclusive events.

Conditional vs unconditonal probability

• Bertrand’s coin box problem

• What are we supposed to compute?

Using Bayes Rule for Inference

- Suppose we want to form a hypothesis about the world based on observable variables.
- Bayes rule tells us how to calculate the belief in a hypothesis H given evidence

eg: $P(H|e)=(P(e|H)P(H))/P(e)$

– $P(H|e)$ is the posterior probability

– $P(H)$ is the prior probability

– $P(e|H)$ is the likelihood

– $P(e)$ is a normalizing constant, which can be computed as:

$$P(e) = P(e|H)P(H) + P(e|\neg H)P(\neg H)$$

Sometimes we write $P(H|e) \propto P(e|H)P(H)$

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:

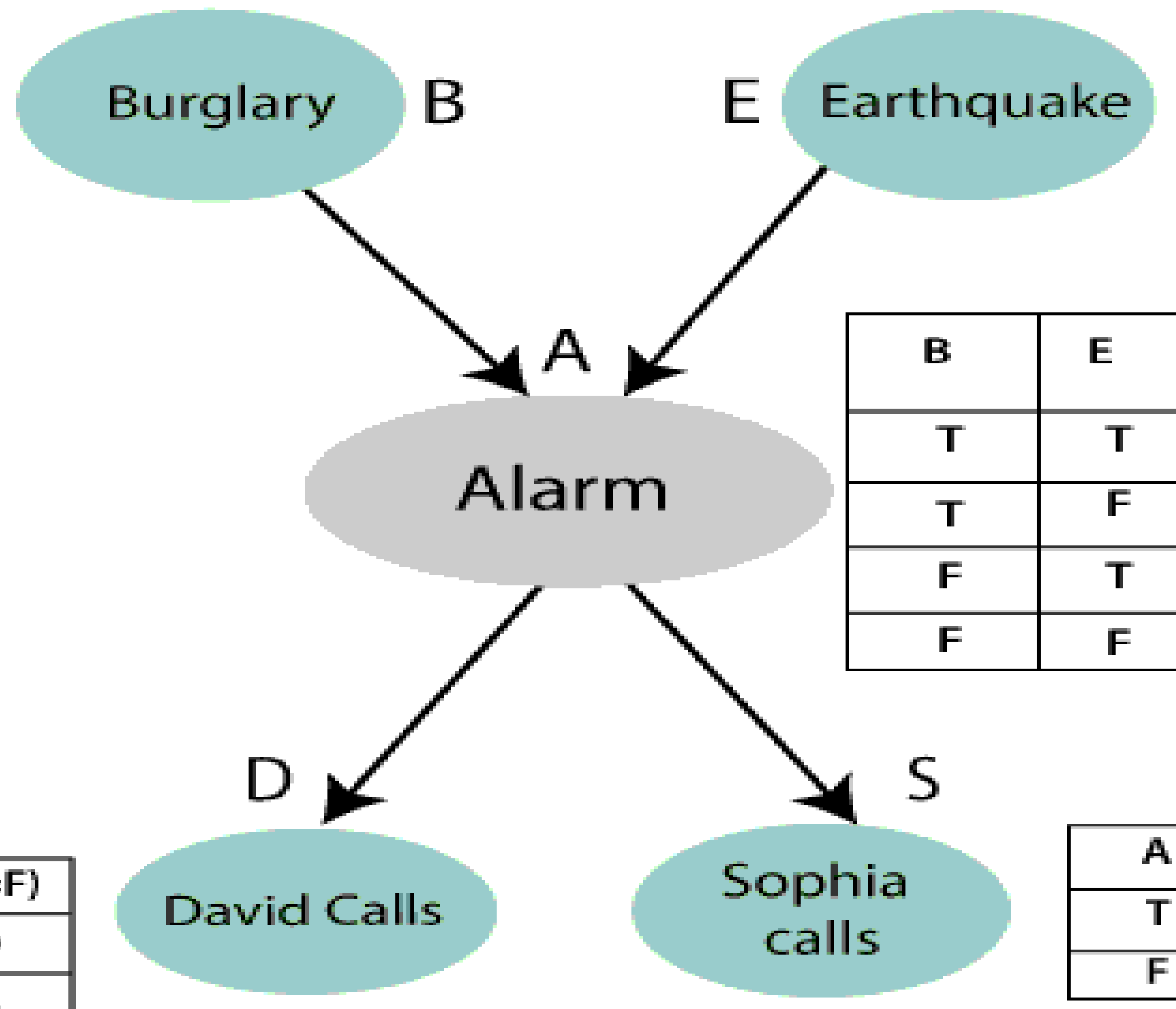
P[D, S, A, B, E]= P[D | S, A, B, E]. P[S, A, B, E]

=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	0.001
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 ^ ¬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.

It is helpful in designing inference procedure.