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

Dempster-Shafer Theory

- Probability theory limitation
  - Assign a single number to measure any situation, no matter how it is complex
  - Cannot deal with missing evidence, heuristics, and limited knowledge
- Dempster-Shafer theory
  - Extend probability theory
  - Consider a set of propositions as a whole
  - Assign a set of propositions an interval [believe, plausibility] to constraint the degree of belief for each individual propositions in the set
  - The belief measure bel is in [0,1]
    - 0 – no support evidence for a set of propositions
    - 1 – full support evidence for a set of propositions
  - The plausibility of p,
    - pl(p) = 1 – bel(not(p))
    - Reflect how evidence of not(p) relates to the possibility for belief in p
    - Bel(not(p))=1: full support for not(p), no possibility for p
    - Bel(not(p))=0: no support for not(p), full possibility for p
    - Range is also in [0,1]

23.Oct.2022

Properties of Dempster-Shafer

- Initially, no support evidence for either competing hypotheses, say h1 and h2
  - Dempster-Shafer: [bel, pl] = [0, 1]
  - Probability theory: p(h1)=p(h2)=0.5
- Dempster-Shafer belief functions satisfy weaker axioms than probability function
- Two fundamental ideas:
  - Obtaining belief degrees for one question from subjective probabilities for related questions
  - Using Dempster rule to combine these belief degrees when they are based on independent evidence

An Example

Two persons M and B with reliabilities detect a computer and claim the result independently. How you believe their claims?

Question (Q): detection claim

Related question (RQ): detectors' reliability

Dempster-Shafer approach

- Obtain belief degrees for Q from subjective (prior) probabilities for RQ for each person
- Combine belief degrees from two persons

Person M:

- reliability 0.9, unreliability 0.1
- Claim h1
- Belief degree of h1 is bel(h1)=0.9
- Belief degree of not(h1) is bel(not(h1))=0.0, different from probability theory, since no evidence supporting not(h1)
- pl(h1) = 1 – bel(not(h1)) = 1-0 =1
- Thus belief measure for M claim h1 is [0.9, 1]

Person B:

- Reliability 0.8, unreliability 0.2
- Claim h2
- bel(h2) =0.8, bel(not(h2))=0, pl(h2)=1-bel(not(h2))=1-0
- Belief measure for B claim h2 is [0.8,1]

Combining Belief Measure

Set of propositions: M claim h1 and B claim h2

- Case 1: h1 = h2
  - Reliability M and B: 09x0.8=0.72
  - Unreliability M and B: 0.1x0.2=0.02
  - The probability that at least one of two is reliable: 1-0.02=0.98
  - Belief measure for h1=h2 is [0.98,1]
- Case 2: h1 = not(h2)
  - Cannot be both correct and reliable
  - At least one is unreliable
    - Reliable M and unreliable B: 0.9x(1-0.8)=0.18
    - Reliable B and unreliable M: 0.8x(1-0.1)=0.08
    - Unreliable M and B: (1-0.9)x(1-0.8)=0.02
    - At least one is unreliable: 0.18+0.08+0.02=0.28
  - Given at least one is unreliable, posterior probabilities
    - Reliable M and unreliable B: 0.18/0.28=0.643
    - Reliable B and unreliable M: 0.08/0.28=0.286
  - Belief measure for h1
    - Bel(h1)=0.643, bel(not(h1))=bel(h2)=0.286
    - Pl(h1)=1-bel(not(h1))=1-0.286=0.714
    - Belief measure: [0.643, 0.714]
  - Belief measure for h2
    - Bel(h2)=0.286, bel(not(h2))=bel(h1)=0.683
    - Pl(h2)=1-bel(not(h2))=1-0.683=0.317
    - Belief measure: [0.286, 0.317]

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:

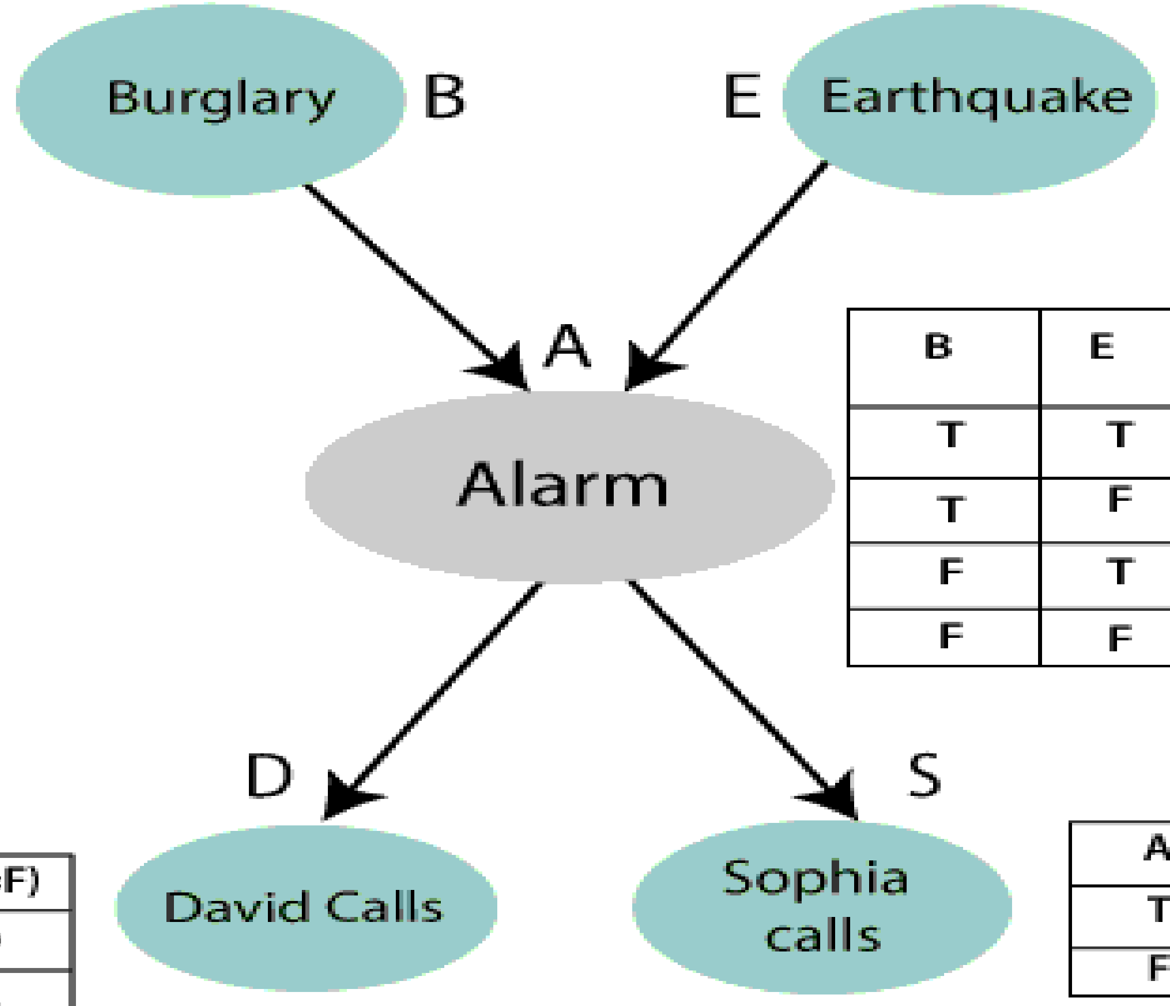
P[D, S, A, B, E]= P[D | S, A, B, E]. P[S, A, B, E]

=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.