CHAPTER – 7
STATISTICAL REASONING

Subject : AI **Code : 2180703**	**Prepared By:** **Asst. Prof. Twinkal Panchal** **(CSE Department, ACET)**	AMIRAJ COLLEGE OF ENGINEERING & TECHNOLOGY

Bayes' Theorem

- Bayes' theorem is also known as **Bayes' rule, Bayes' law**, or **Bayesian reasoning**, which determines the probability of an event with uncertain knowledge.

- It is a way to calculate the value of P(B|A) with the knowledge of P(A|B).
- Bayes' theorem can be derived using product rule and conditional probability of event A with known event B:

  $\mathbf{P(A \wedge B) = P(A|B)\ P(B)}$ or
- Similarly, the probability of event B with known event A:

  $\mathbf{P(A \wedge B) = P(B|A)\ P(A)}$
- Equating right hand side of both the equations, we will get:

$$P(A|B) = \frac{P(B|A)\,P(A)}{P(B)} \quad \text{....(a)}$$

- The above equation (a) is called as **Bayes' rule** or **Bayes' theorem**. This equation is basic of most modern AI systems for **probabilistic inference**.

- P(A|B) is known as **posterior**, which we need to calculate, and it will be read as Probability of hypothesis A when we have occurred an evidence B.

- P(B|A) is called the likelihood, in which we consider that hypothesis is true, then we calculate the probability of evidence.

AMIRAJ
COLLEGE OF ENGINEERING & TECHNOLOGY

- P(A) is called the **prior probability**, probability of hypothesis before considering the evidence

- P(B) is called **marginal probability**, pure probability of an evidence.

- In the equation (a), in general, we can write P (B) = P(A)*P(B|Ai), hence the Bayes' rule can be written as:

$$P(A_i|B) = \frac{P(A_i)*P(B|A_i)}{\sum_{i=1}^{k} P(A_i)*P(B|A_i)}$$

Where A_1, A_2, A_3,........, A_n is a set of mutually exclusive and exhaustive events.

AMIRAJ COLLEGE OF ENGINEERING & TECHNOLOGY

Ex-

- A doctor is aware that disease meningitis causes a patient to have a stiff neck, and it occurs 80% of the time. He is also aware of some more facts, which are given as follows:

- The Known probability that a patient has meningitis disease is 1/30,000.The Known probability that a patient has a stiff neck is 2%.

$$P(a|b) = 0.8$$

$$P(b) = 1/30000$$

$$P(a) = .02$$

$$P(b|a) = \frac{P(a|b)P(b)}{P(a)} = \frac{0.8*(\frac{1}{30000})}{0.02} = 0.001333333.$$

AMIRAJ COLLEGE OF ENGINEERING & TECHNOLOGY

Application of Bayes' theorem

- It is used to calculate the next step of the robot when the already executed step is given.
- Bayes' theorem is helpful in weather forecasting.
- It can solve the Monty Hall problem.

Certainty factors and rule based system

A certainty factor (CF [h, e]) is defined in terms of two components:

1. MB [h, e]

- A measure (between 0 and 1) of belief in hypothesis h given the evidence e. MB measures the extent to which the evidence supports the hypothesis. It is zero if the evidence fails to support the hypothesis.

2. **MD [h, e]**

- A measure (between 0 and 1) of disbelief in hypothesis h given the evidence e. MD measures the extent to which the evidence supports the negation of the hypothesis. It is zero if the evidence supports the hypothesis.

CF [h, e] = MB[h, e] – MD[h, e]

AMIRAJ COLLEGE OF ENGINEERING & TECHNOLOGY

- The measures of belief and disbelief of a hypothesis given, two observations S_1 and S_2 are computed from:

$$\text{If } MD[h, s_1 \wedge s_2] = 1 \text{ otherwise}$$

$$MB[h_1, s_1 \wedge s_2] = \begin{cases} 0 \\ MB[h, s_1] + MD[h, s_2] \cdot (1 - MB(h, s_1) \end{cases} \quad \ldots (7.5)$$

$$\text{if } MB[h, s_1 \wedge s_2] = 1$$

$$MD[h_1, s_1 \wedge s_2] = \begin{cases} 0 \\ MD[h, s_1] + MD[h, s_2] \cdot (1 - MD(h, s_1) \end{cases} \text{otherwise}$$

AMIRAJ
COLLEGE OF ENGINEERING & TECHNOLOGY

Ex -

- Suppose we make an initial observation corresponding to Fig. 7.4 (a) which confirms our belief in h with MB = 0.3. Then MD [h, s_1] = 0 and CF [h, s_1] = 0.3 Now we make a second observation, which also confirms h, with MB [h, s_2] = 0.2.

$$MH[h, s_1 \wedge s_2] = 0.3 + 0.2 \times 0.7$$
$$= 0.44$$

$$MD[h, s_1 \wedge s_2] = 0.0$$
$$CF[h, s_1 \wedge s_2] = 0.44$$

AMIRAJ COLLEGE OF ENGINEERING & TECHNOLOGY

- MYCIN uses for the MB the conjunction and the disjunction of two hypotheses as given below:

$$MB[h_1 \wedge h_2, e] = \min(MB[h_1, e], MB[h_2, e])$$
$$MB[h_1 \vee h_2, e] = \max(MB[h_1, e], MB[h_2, e]) \quad ...(7.6)$$

AMIRAJ
COLLEGE OF ENGINEERING & TECHNOLOGY

Bayesian Network

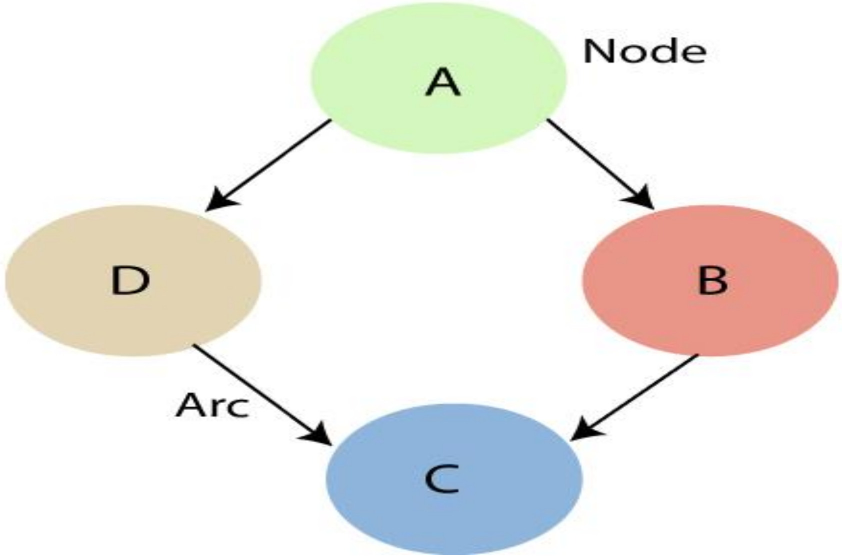
- "A Bayesian network is a probabilistic graphical model which represents a set of variables and their conditional dependencies using a directed acyclic graph."

- It is also called a **Bayes network, belief network, decision network**, or **Bayesian model**.

- Bayesian Network consists of two parts:

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

- A Bayesian network graph is made up of nodes and Arcs (directed links), where:



- Each node corresponds to the random variables, and a variable can be continuous or discrete.
- Arc or directed arrows represent the causal relationship or conditional probabilities between random variables

- In the above diagram, A, B, C, and D are random variables represented by the nodes of the network graph.
- If we are considering node B, which is connected with node A by a directed arrow, then node A is called the parent of Node B.
- Node C is independent of node A.

Dempster – Shafer Theory

- The Dempster-Shafer theory was designed to mathematically model and validate the uncertainty involved in statistical inferences. This theory allows one to combine evidence from different sources and arrive at a degree of belief, which is mathematically defined by belief functions (Bel).

Belief Functions

- Belief functions were initially proposed as a way to achieve generalized Bayesian inference without priors and mathematically related to random sets. Belief functions combine a set of representations and model data when there is a lack of information.

AMIRAJ
COLLEGE OF ENGINEERING & TECHNOLOGY

Belief And Plausibility

- Dempster's rule of combination is sometimes interpreted as an approximate generalisation of Bayes' rule.

- Belief in a hypothesis is calculated as the summation of all these masses of all subsets of the hypothesis-set. The value of belief gives a lower bound on its probability.

- Belief(Bel) ranges from 0 to 1 where 0 indicates no evidence and 1 indicates certainty.

Plausibility(Pl) = 1-Bel.

Fuzzy Logic

- Fuzzy Logic Systems (FLS) produce acceptable but definite output in response to incomplete, ambiguous, distorted, or inaccurate (fuzzy) input.

- Fuzzy Logic (FL) is a method of reasoning that resembles human reasoning. The approach of FL imitates the way of decision making in humans that involves all intermediate possibilities between digital values YES and NO.

- The conventional logic block that a computer can understand takes precise input and produces a definite output as TRUE or FALSE, which is equivalent to human's YES or NO.

Application Areas of Fuzzy Logic

1. **Automotive Systems -** Automatic Gearboxes
   Four-Wheel Steering
   Vehicle environment control
2. **Consumer Electronic Goods -** Hi-Fi Systems
   Photocopiers
   Television
3. **Domestic Goods -** Microwave Ovens
   Refrigerators
   Toasters
   Vacuum Cleaners
   Washing Machines

AMIRAJ
COLLEGE OF ENGINEERING & TECHNOLOGY

Advantages

- Mathematical concepts within fuzzy reasoning are very simple.
- Fuzzy logic Systems can take imprecise, distorted, noisy input information.
- FLSs are easy to construct and understand.

Disadvantages

- There is no systematic approach to fuzzy system designing.
- They are understandable only when simple.
- They are suitable for the problems which do not need high accuracy.

AMIRAJ
COLLEGE OF ENGINEERING & TECHNOLOGY