2.1 ACTING UNDER UNCERTAINTY

Uncertainty is a situation which involves imperfect and/or unknown information. However, "uncertainty is an unintelligible expression without a straightforward description.

Uncertainty can also arise because of incompleteness, incorrectness in agents understanding the properties of environment.

So to represent uncertain knowledge, where we are not sure about the predicates, we need uncertain reasoning or probabilistic reasoning.

Causes of uncertainty:

Following are some leading causes of uncertainty to occur in the real world.

1. Information occurred from unreliable sources.
2. Experimental Errors
3. Equipment fault
4. Temperature variation
5. Climate change.

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.

We use probability in probabilistic reasoning because it provides a way to handle the uncertainty that is the result of someone's laziness and ignorance.

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 here we use probabilistic reasoning.

A **certainty factor (CF)** is a numerical value that expresses a degree of subjective belief that a particular item is true. The item may be a fact or rule.

Suppose that a certainty is defined to be a real number between -1.0 and +1.0, where 1.0 represents complete certainty that an item is true and -1.0 represents complete certainty that an item is false. Here a CF of 0.0 indicates that no information is available about either the truth or the falsity of an item. Hence positive values indicate a degree of belief or evidence that an item is true, and negative values indicate the opposite belief.

Ad-hoc methods of dealing with uncertainty are methods which have no formal theoretical basis.

Ad-hoc procedures are implemented in the MYCIN system. It is the expert systems developed to diagnose meningitis and infectious blood disease.

MYCIN

Its knowledge base is composed of if….then rules which are used to assess various forms of patient evidence with the ultimate goal being the formulation of a correct diagnosis and recommendation for a suitable therapy.

A typical rule has the form:

IF: The stain of the organism is gram positive, and

The morphology of the organism is Coccus, and

The growth conformation of the organism is chains

THEN: There is suggestive evidence (0.7) that the identity of the organism is streptococcus.

Rule-Implication;

- The rule is used by the inference mechanism to help identify the offending organism.
- The 3 conditions given in the IF part of the rule refer to attributes that help to characterize and identify organisms.

- When identification is certain, an appropriate therapy is recommended.
- The numeric value (0.7) given in the THEN part of the above rule corresponds to an experts estimate of degree of belief one can place in the rule conclusion when the three conditions if the IF part have been satisfied.
- The Belief associated with the rule may be thought of as a conditional probability.

$$P(H/E1,E2,E3) = 0.7$$

Where H is the hypothesis that the organism is streptococcus, and E1, E2 and E3 correspond to the three pieces of joint evidence given in the IF part, respectively.

MYCIN – Belief and Disbelief :

- MYCIN uses measures of both belief and disbelief to represent degrees of confirmation and disconfirmations respectively in a given hypothesis.
- The measure of belief denoted by MB (H,E) is a measure of the increased belief in hypothesis H due to the evidence E.
- Its equivalent to the estimated increase in probability of P(H/E) over P(H) given by an expert as a result of the knowledge gained by E.
- A value of 0 corresponds to no increase in belief and 1 corresponds to maximum increase or absolute belief.
- MD(H,E) is a measure of the increased disbelief in hypothesis H due to evidence E.
- MD ranges from 0 to +1, with +1 representing maximum increase in disbelief 0 representing no increase.
- In both measures, the evidence E may be absent or may be placed with another hypothesis, MB(H1,H2). This represents the increased belief in H1 given H2 is true.

Uncertainty measure – MYCIN:

MB and MD given in terms of prior and conditional probabilities.

The actual values are subjective probability estimates provided by a physician.

$$MB(H,E) = \begin{cases} 1 & if\ P(H) = 1 \\ \dfrac{max[P(H|E), P(H)] - P(H)}{max[1,0] - P(H)} & Otherwise \end{cases}$$

$$MD(H,E) = \begin{cases} 1 & if\ P(H) = 0 \\ \dfrac{min[P(H|E), P(H)] - P(H)}{max[1,0] - P(H)} & Otherwise \end{cases}$$

Note that when 0 < P(H) <1, and E and H are independent (so P(H|E)=P(H), then MB =MD=0.This would be the case if E provided no useful information.

The two measures MB and MD are combined into a single measure called the certainty factor (CF), denoted by

CF(H,E) =MB(H,E) – MD(H,E)

Note that the value of CF ranges from -1 (certain disbelief) to +1(certain belief). Furthermore, a value of CF=0 will result if E neither confirm nor unconfirms H(E and H are independent).

In MYCIN, each rule hypothesis Hi has an associated MB and MD initially set to 0,

- Evidence are accumulated, they are updated using intermediate combining functions, and when all applicable rules have been executed, a final CF is calculated for each H_i
- These are then compared and the largest cumulative confirmations or disconfirmations are used to determine the appropriate therapy.
- A threshold value of |CF| > 0.2 is used to prevent the acceptance of a weakly supported hypothesis.
- In the initial assignment of belief values an expert will consider all available confirming and disconfirming evidences E1, E2,……EK and assign properties, consistent values to both.

2.2 BAYESIAN INFERENCE

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.

We use probability in probabilistic reasoning because it provides a way to handle the uncertainty that is the result of someone's laziness and ignorance.

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 here we use probabilistic reasoning.

Need of probabilistic reasoning in AI:

- When there are unpredictable outcomes.
- When specifications or possibilities of predicates becomes too large to handle.
- When an unknown error occurs during an experiment.

In probabilistic reasoning, there are two ways to solve problems with uncertain knowledge:

- **Bayes' rule**
- **Bayesian Statistics**

As probabilistic reasoning uses probability and related terms, so before understanding probabilistic reasoning, let's understand some common terms:

Probability: Probability can be defined as a chance that an uncertain event will occur. It is the numerical measure of the likelihood that an event will occur. The value of probability always remains between 0 and 1 that represent ideal uncertainties.

- $0 \leq P(A) \leq 1$, where P(A) is the probability of an event A.

- P(A) = 0, indicates total uncertainty in an event A.
- P(A) =1, indicates total certainty in an event A.

We can find the probability of an uncertain event by using the below formula.

$$\textbf{Probability of occurrence} = \frac{\text{Number of desired outcomes}}{\text{Total number of outcomes}}$$

- ○ P(¬A) = probability of a not happening event. ○ P(¬A) + P(A) = 1.

Conditional probability:

Conditional probability is a probability of occurring an event when another event has already happened.

Let's suppose, we want to calculate the event A when event B has already occurred, "the probability of A under the conditions of B", it can be written as:

$$P(A|B) = \frac{P(A \wedge B)}{P(B)}$$

Where P($A \wedge B$) = Joint probability of a and B

P(B)= Marginal probability of B.

If the probability of A is given and we need to find the probability of B, then it will be given as:

$$P(B|A) = \frac{P(A \wedge B)}{P(A)}$$

Example:

In a class, there are 70% of the students who like English and 40% of the students who likes English and mathematics, and then what is the percent of students those who like English also like mathematics?

Solution:

Let, A is an event that a student likes Mathematics B

is an event that a student likes English.

$$P(A|B) = \frac{P(A \wedge B)}{P(B)} = \frac{0.4}{0.7} = 57\%$$

Hence, 57% are the students who like English also like Mathematics.

BAYESIAN REASONING (BAYESIAN INFERENCE)

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

- In probability theory, it relates the conditional probability and marginal probabilities of two random events.
- Bayes' theorem was named after the British mathematician **Thomas Bayes**. The **Bayesian inference** is an application of Bayes' theorem, which is fundamental to Bayesian statistics.
  - It is a way to calculate the value of P(B|A) with the knowledge of P(A|B).
- Bayes' theorem allows updating the probability prediction of an event by observing new information of the real world.

Example: If cancer corresponds to one's age then by using Bayes' theorem, we can determine the probability of cancer more accurately with the help of age.

Bayes' theorem can be derived using product rule and conditional probability of event A with known event B:

As from product rule we can write:

1. $P(A \wedge B) = P(A|B) P(B)$ or

Similarly, the probability of event B with known event A:

1. P(A ∧ 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**.

It shows the simple relationship between joint and conditional probabilities. Here,

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.

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.

Applying Bayes' rule:

Bayes' rule allows us to compute the single term P(B|A) in terms of P(A|B), P(*B*), and P(A). This is very useful in cases where we have a good probability of these three terms and want to determine the fourth one. Suppose we want to perceive the effect of some unknown cause, and want to compute that cause, then the Bayes' rule becomes:

$$\textbf{P(cause | effect)} = \frac{P(\text{effect}|\text{cause})\ P(\text{cause})}{P(\text{effect})}$$

Example-1:

Question: what is the probability that a patient has diseases meningitis with a stiff neck?

Given Data:

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%.

Let a be the proposition that patient has stiff neck and b be the proposition that patient has meningitis. , so we can calculate the following as:

P(a|b) = 0.8

P(b) = 1/30000

P(a)= .02

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

Hence, we can assume that 1 patient out of 750 patients has meningitis disease with a stiff neck.

Application of Bayes' theorem in Artificial intelligence:

Following are some applications 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.