5

CHAPTER

REASONING UNCERTAINTY

5.0 INTRODUCTION

Reasoning means deriving conclusions from the available set of data and information. We deal with many types of situations in this real world that are full of uncertainties. Our AI systems should be such that they should be able to handle these uncertainties. These uncertainties occur because the knowledge that we acquire is from beliefs and hypotheses. When we get some new information or when our beliefs change, we face uncertainty. Experimental errors, Randomness, lack of evidences are all the causes of uncertainties. **The knowledge stored in KB is incomplete. So, an intelligent AI system should be able to handle these uncertainties. Note that the normal logical problem solving techniques-propositional and predicate logic cannot handle such uncertain situations.** To handle uncertain situations, the uncertain knowledge needs to be represented and thus reasoned.

5.1 UNCERTAIN KR METHOD

Certain events are "**certain events**" like "New Delhi is the Capital of India" but some events like "It may rain today" are "**uncertain events**" These uncertain events need to be stored in KB now. To achieve this, the uncertain situation is represented by attaching a **degree of belief factor**. It is related to the probability theory. So, representation of uncertain events uses the following terminology—

(*a*) **Belief:** It is any meaningful and coherent expression that can be represented. It may be true or false.

(*b*) **Belief factor:** It indicates the correctness of knowledge. The value of this factor is between 0 and 1.

(*c*) **Evidence:** It is the observation (s) obtained in the real world.

(*d*) **Hypothesis:** It is a justified belief that is supported by some evidence. It is not known to be true.

(*e*) **Knowledge:** It is a true justified belief.

(*f*) **Meta knowledge:** It is the knowledge about the knowledge.

How to reason this uncertain knowledge?

AI aims at building intelligent systems to solve those real world problems that involve uncertainties. Reasoning is done on uncertain information contained in the KB. **Please understand that without this ability an AI system is just like other programming languages or like a search table doing sequential search.** So, we can say,

AI $\longrightarrow$ used to build intelligent systems $\longrightarrow$ Use facts and rules stored in KB

Three types of reasoning are used here—

1. Inductive Reasoning

It is based on the generalizations of the previous experiences about the problem. For e.g.: Consider the situation—

(*a*) Rajiv is an Associate Professor.

(*b*) He teaches in a college.

Now, if some one asks you to guess about college environment then from your experience, you can infer that college has — Professors, Lecture Rooms, Conference Rooms, Labs, play grounds etc.

This is known as inductive **reasoning**. In AI field, this method is used in **machine learning**.

2. Abductive Reasoning

It is an unsound rule of inferencing meaning that the conclusion is not necessarily true for every interpretation made out of the environment. It is not very reliable reasoning method and is preferred in cases where knowledge is incomplete. It looks back through the chain of events to do reasoning.

For e.g. If a student fails in his/her exam although the paper is easy. To find the reason as to why he/she failed, we need to look back in the chain of events that paper is easy, student studied, student failed, teacher is not strict. This type of reasoning is **abductive reasoning.**

3. Deductive Reasoning

It is a kind of explicit, non-reversible reasoning that works on standard logic.

For e.g., "If mobile is OFF", it is certain that no calls can be made. This is deductive reasoning. But the reverse is NOT true.

5.1.1 Monotonic versus Non-Monotonic Reasoning

Monotonic Reasoning	Non-Monotonic Reasoning
1. All new knowledge that is added to the KB must be consistent with the previous knowledge.	1. In uncertain situations addition of new axioms may contradict with earlier and might be required to be removed from the KB.
2. The size of the KB always increases monotonically.	2. The size of the KB increases-non monotonically.
3. No retractions of rules (removals) are allowed.	3. Allow retractions also i.e., removal of contradictory facts.
4. Inference methods available are insufficient.	4. Inference methods are complete and justified.
5. Traditional KR techniques like PL, FOPL, Propositional logic cannot accommodate the real world's changing situations. **For e.g.:** PL, FOPL, Propositional logic, Theorem Proving etc.	5. These systems like TMS allows the revision in belief and accommodate the real world changing situations. **For e.g.:** Truth Maintenance Systems (TMS).

5.1.2 TMS-Truth Maintenance Systems (or Belief or Revision Maintenance Systems)

1. TMS are "truth-maintenance systems" which have been implemented to permit a form of non monotonic reasoning by permitting the addition of changing (even contradictory) statements to a knowledge base.
2. Also known as "belief-revision" and "revision-maintenance-systems" are companion-components to inference system, (of our basic KBS).

3. Main job– of TMS is to maintain "**consistency of knowledge**" being used by the problem solver and not to perform any inference functions.

   As such, it **frees** the problem solver from any concerns of consistency and allows it to concentrate on the problem solution aspects.

4. TMS also gives the inference–component, the latitude to perform non-monotonic inferences.

5. When new discoveries are made, this more recent information can displace previous conclusions that are no longer valid. In this way, set of beliefs available to the problem solver will continue to be current and consistent.

   For e.g., Shows role of TMS

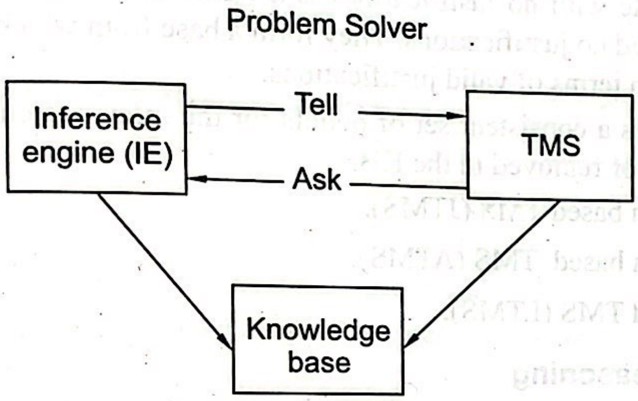


Fig. 5.1 Architecture of the problem solver with a TMS.

Working: The inference engine (IE) solves domain problems based on its current belief set, while the TMS maintains the currently active belief set. The updating process is incremental. After each inference, information is exchanged between the two components. The IE tells the TMS what deductions it has made. The TMS, in turn, **asks** question about current beliefs and reasons for failure. If maintains a consistent set of beliefs for the IE to work with even if now knowledge is added and removed.

For e.g., S1: Say, KB contained the propositions P, P → Q and modus ponens.

S2: From this, the IE would right fully conclude Q and add this conclusion to the KB.

S3: Later, if it was learned that ~P was appropriate, if would be added to the KB resulting in a contradiction.

S4: Consequently, it would be necessary to remove P to eliminate the inconsistency. But, with P now removed, Q is no longer a justified belief. It too should be removed.

This type of belief revision is the job of the TMS.

7. Actually, the TMS does not discard conclusions like Q as suggested. That could be wasteful, since P may again become valid, which would require that Q and facts justified by Q be received. Instead, the TMS maintains **dependency records** for all such conclusion. These records determine which set of beliefs are current (which are to be used by the IE). Thus, Q would be removed from the current belief set by making appropriate updates to the records and not by erasing Q. Since Q would not be lost, its rederivation would not be necessary if P became valid once again.

8. TMS maintains complete records of reasons of justifications for beliefs. Each proposition or statement having at least one valid justification is made a part of the current belief set. Statements lacking acceptable justifications are excluded from this set. When a contradiction is discovered, the statements responsible for the contradiction are identified and an

appropriate one is retracted. This in turn may result in other retractions and additions. The procedure used to perform this process is called **"dependency-directed back tracking"**.

9. TMS maintains records to reflect retractions and additions so that the IE will always know its current belief set. The records are maintained in the form of a "dependency- –network".
   (*a*) The **nodes** in the network represent KB entries such as premises, conclusions, inference rules and the like.
   (*b*) **Attached** to the nodes are justifications which represent the inference steps from which the node was derived.

Nodes in the belief set must have valid justifications.

A premise (or a node with no justification) is a fundamental belief which is assumed to be always true. Premises need no justifications. They form a base from which all other currently active nodes can be explained in terms of valid justifications.

Thus, TMS maintains a consistent set of beliefs for the inference engine to rely upon even if new knowledge is added or removed in the KB.

For e.g., (*a*) Justification based TMS (JTMS).
(*b*) Assumption based TMS (ATMS).
(*c*) Logic based TMS (LTMS).

5.1.3 Probabilistic Reasoning

It is a statistical method of reasoning based on some numeric value assigned to them. It involves 4 methods—

(*a*) Bayesian belief networks.
(*b*) Dempster Shaffer theory.
(*c*) Reasoning with certainty factors.
(*d*) Fuzzy reasoning.

Let us discuss these reasoning one by one.

5.1.3.1 Bayesian belief networks

Bayesian networks are the graphical representation of dependent events. These networks are an easy graphical representation of real world dependent events. These are directed acyclic graphs in which **nodes represent the events** or evidence **and the arcs represent their dependence.**

Basic Properties of Bayesian networks

1. They have a set of random, discrete or continuous variables generating the nodes of the network.
2. They have directed links or arrows connecting a pair of nodes.
   For e.g. a → b
   Here, 'a' is parent of 'b'.
3. Each nodes (n_i) has a conditional probability distribution i.e.,
   P (n_i|parent (n_i))

This quantifies the effect of parent on the node.

4. **The graph has no directed cycles** i.e., these networks are directed acyclic graphs (DAG).
5. **The topology of graph** indicates that node specifies the conditional independence relationships of the problem.

6. Once belief network is constructed, an inference engine of an intelligent system can use it to maintain and propagate beliefs. If some new information is received, the effects of same can be propagated throughout the network until equilibrium probabilities are reached.
7. **At equilibrium all propositions will have consistent probabilities.**

For example : Medical diagnosis systems like CASNET uses this reasoning.

5.1.3.2 Dempster Shafer Theory

There are various theories available for example Bayesian networks that consider or find out the belief from some premises from a single permit of observation.

Dempster Shafer theory considers the fact that the observations or correct belief can be made out by not one degree of final belief but it actually considers an interval defined as,

(Belief, Plausibility)

1. The degree of correct observation lies between 0 to 1 according to the range of above interval.
2. Belief is made on certain facts that are considered to be true and plausibility is based on the value of belief.

$$\boxed{\text{Plausibility} = 1 - \text{Belief}\ (\neg S)}$$

3. The value of the interval defined by the Dempster shafer ranges from 0 (uncertainty) to 1 (certain).
4. When the value of belief on the some observation is, 's 'that belief $(\neg S)$ is 1, the plausibility is 0 and its 1 otherwise.

EXAMPLE:

Let us consider a set that considers a diagnosis problem from bacterial in infection. It is {All, Flu, Cold, Pneu}, where

All	:	Allergy
Flu	:	Flu
Cold	:	Cold
Pneu	:	Pneumonia

5. The Dempster Shafer theory applies not only to the members of sets but also to its subsets including the Singleton set.

We will call the set as θ.

θ: {All, Flu, Cold, Pneu}

If there are n members of the set, then total no. of subsets that are there can be 2^n.

6. The degree of functionality, 'm' can then be found out from the set θ.
7. The degree of functionality 'm' can apply not to all members of sets.

For example, consider that a value fever that may not happen in the case of allergy.

Therefore, only following members of set are

θ': = { Flu, Cold Pneu}

8. Giving values to these subsets, the value of m i.e. degree of functionality can be found out.

5.1.3.3 Reasoning with Certainty Factors (CF)

The certainty factors are human estimates of symton/cause relationship and their probability measures. In this type of reasoning, we need to combine the effects of all of the facts and assign a single value to the rule known as CF i.e., here the problem is represented as a simple rule as follows—

IF p and q and r THEN S(CF).

Please note that CF-certainty factor will hold good if we are certain about facts. Also note that a new CF would be calculated whenever there is any change in the measure of belief of any or all of the facts.

Certainty theory does not produce an algebra for correct reasoning. It **is adhoc** meaning that the confidence of human expert in his result is approximate, heuristic and informal.

For e.g.: MYCIN expert system uses this type of reasoning.

The certainty theory makes simple assumption for creating and assigning confidence measures 'for' and 'against' any given situation.

5.1.3.4 *Fuzzy Logic Reasoning*

According to George Boole, human thinking and decision are based on Yes or No reasoning or logic 1 and logic 0. Accordingly, Boolean algebra was developed. Even expert systems founded this logic. But it has been argued that **human thinking does not always follows crisp** YES or NO logic rather it is quite **vague, uncertain, imprecise or fuzzy in nature.** Based on the nature of fuzzy human thinking, **Lofti Zadeh**, a computer scientist at the University of California, Berkeley, originated the fuzzy logic or fuzzy set theory in 1965. Classical set theory/crisp set theory (given by George Cantor) is fundamental to the study of fuzzy sets. **Please understand that as Boolean logic had its roots in the theory of crisp sets, fuzzy logic has its roots in the theory of fuzzy sets.**

Also note that this process of making a crisp quantity fuzzy is known as fuzzification. We do this by recognizing that many of the quantities we consider to be crisp and deterministic are actually non-deterministic. They carry considerable uncertainty. If the form of uncertainty happens to arise because of imprecision, ambiguity or vagueness then the variable is probable fuzzy and can be represented by a membership function.

> **NOTE:** Reasoning about imprecise propositions is an ultimate goal of fuzzy logic and is referred to as approximate reasoning.

Features of membership characteristics functions

1. It describes vague natural real world concepts.
2. A fuzzy set admits the possibility of partial members with associated membership characteristic function value.
3. The degree of an object belonging to a fuzzy set is denoted by membership value between 0 and 1.
4. A membership function associated with given fuzzy set maps input value to its appropriate membership value.
5. Uncertainties are represented with membership functions and then this function is manipulated in a method defined in fuzzy theory.

Laws of operations on fuzzy sets

1. **Law of Equality:** "If A and B are two fuzzy sets such that A, B $\subseteq$ U then A = B iff mA(x) = mB(x) for all x $\in$ U".
2. **Law of containment:** "A $\subseteq$ B iff mA(x) $\leq$ mB(x) for all x $\in$ U.
3. **Union law:** " A $\cup$ B is the smallest fuzzy subset of which both A and B are subsets."
4. **Intersection law:** " A $\cap$ B is the largest fuzzy subset which is a subset of both A and B".
5. **Distributive law**

   (*a*) $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$

   (*b*) $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$

6. **Associative law**
   (*a*) $(A \cup B) \cup C = A \cup (B \cup C)$
   (*b*) $(A \cap B) \cap C = A \cap (B \cap C)$
7. **Commutative law**
   (*a*) $A \cup B = B \cup A$
   (*b*) $A \cap B = B \cap A$
8. **Idempotency law**
   (*a*) $A \cup A = A$
   (*b*) $A \cap A = A$

Certain other laws used on membership characteristics functions are as follows–

(*a*) **Union law**

$m(A(x) \cup B(x))$ or $mA(x)$ or $B(x)$

$= MAX\{mA(x), mB(x)\}$

(*b*) **Intersection law**

$m(A(x) \cap B(x))$ or $mA(x)$ or $B(x)$

$= MIN\{mA(x), mB(x)\}$

(c) **Negation law**

Negation of $mA(x) = 1 - mA(x)$

(*d*) **Dilation**

If A is a fuzzy set such that $A \subseteq U$, for all $x \in A$, the dilation of A is defined as–

$DIL(A) = [mA(x)]^{1/2}$ for $x \in U$

(*e*) **Concentration**

The concentration of A is defined as–

$CON(A) = [mA(x)]^2$ for all $x \in U$

(*f*) **Normalization**

The normalization of A is defined as–

$NORM(A) = mA(x)/MAX\{mA(x)\}$

NOTE: Concentration is opposite of dilation.

Normalization is the process of normalizing all characteristic functions to the same base. These operations are used to manipulate the membership characteristic function according to the variation in fuzziness of a function.

To summarize, we can say in nutshell that because real world situations are very vague and cannot be crisply defined, so they are modeled using fuzzy theory only in AI.

SUMMARY

We encounter various situations in our real life where the outcome is uncertain. This chapter focuses on how this type of knowledge is represented and reasoned. Various theories have been given to accomplish the same.

MULTIPLE CHOICE QUESTIONS [MCQS]

1. The knowledge about the knowledge is known as
   (*a*) knowledge (*b*) Meta knowledge
   (*c*) Knowledge base (d) None of the above.
2. A type of reasoning which is based on generalizations of previous experience about t problem is known as
   (*a*) Inductive reasoning (*b*) Aductive reasoning
   (*c*) Deductive reasoning (*d*) None of the above.
3. PL and FOPL are examples of
   (*a*) Monotonic reasoning (*b*) Non-monotonic reasoning
   (*c*) TMS (*d*) None of the above.
4. TMS stands for
   (*a*) True management system (*b*) Teller management system
   (*c*) Truth maintenance system (*d*) None of the above.
5. Which unit makes inferences based on current set of beliefs
   (*a*) Inference engine (*b*) TMS
   (*c*) KB (*d*) None of the above.
6. CASNET uses
   (*a*) Fuzzy logic reasoning (*b*) Bayesian belief networks
   (*c*) Dempster shaffer theory (*d*) None of the above.
7. MYCIN uses
   (*a*) Reasoning with certainity factor (*b*) Fuzzy logic
   (*c*) Bayesian networks (*d*) None of the above.
8. Fuzzy logic has its roots in the theory of
   (*a*) Crisp sets (*b*) Fuzzy sets
   (*c*) Boolean algebra (*d*) None of the above.
9. The process of making a crisp quantity fuzzy is known as
   (*a*) fuzzification (*b*) de-fuzzification
   (*c*) assumption (*d*) None of the above.
10. Concentration is opposite of
    (*a*) Fuzzification (*b*) Dilation
    (*c*) Normalization (*d*) None of the above.

ANSWERS

1. (*b*)	**2.** (*a*)	**3.** (*a*)	**4.** (c)	**5.** (*a*)
6. (*b*)	**7.** (*a*)	**8.** (*b*)	**9.** (*a*)	**10.** (*b*)

CONCEPTUAL SHORT QUESTIONS WITH ANSWERS

Q. 1. In a database, a rule is defined as (P1 and P2) or P3? R1 (0.8) and R2 (0.3) where P1, P2, P3 are premises and R1, R2 are the conclusions of the rule with CFs 0.8, 0.3 respectively. If any running program has produced P1, P2 and P3 with CFs as 0.5, 0.8, 0.2 respectively. Find CF of results on the basis of premises.

Ans. The CF of R1 is 0.8 (given)

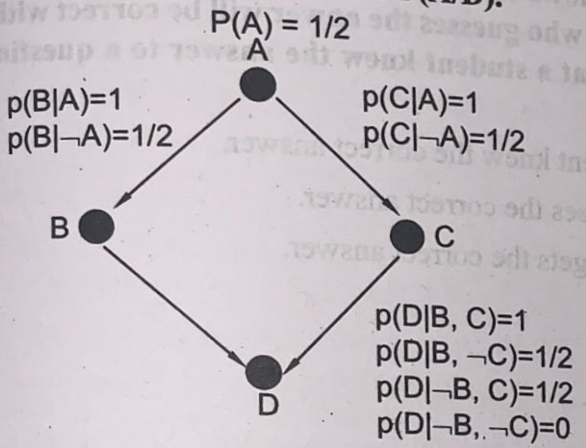
∴ CF of R1 based on premises will be

$$= 0.8 * 0.5$$
$$= 0.40.$$

Similarly, CF of R2 based on premises will be $(0.3 * 0.5) = 0.15$.

Q. 2. An admissions committee for a college is trying to determine the probability that an admitted candidate is really qualified. The relevant probabilities are given in the Bayes network shown here. Calculate P(A/D).

Ans. 1.



A = application is qualified

B = applicant has high grade point average

C = applicant has excellent recommendations

D = applicant is admitted.

Find $p(A/D) = ?$

Ans. 2.

$$p(A|D) = p(A, D)/p(D)$$

$$\begin{aligned} p(A,D) = {} & p(D|B, C)p(B|A)p(C|A)p(A) \\ & + p(D|B, \neg C)p(B|A)p(\neg C|A)p(A) \\ & + p(D|\neg B, C)p(\neg B|A)p(C|A)p(A) \\ & + p(D|\neg B, \neg C)p(\neg B|A)p(\neg C|A)p(A) \\ = {} & 1/2 \end{aligned}$$

$$\begin{aligned} p(D) = {} & p(D/B, C)p(B/A)p(C/A)p(A) \\ & + p(D|B, \neg C)p(B|A)p(\neg C|A)p(A) \\ & + p(D|\neg B, C)p(\neg B|A)p(C|A)p(A) \\ & + p(D|\neg B,\neg C)p(\neg B/A)p(\neg C|A)p(A) \\ & + p(D|B,C)p(B|\neg A)p(C|\neg A)p(\neg A) \\ & + p(D|B,\neg C)p(B|\neg A)p(\neg C|\neg A)p(\neg A) \end{aligned}$$

$$+p(D|\neg B,C)p(\neg B|\neg A)p(C|\neg A)p(\neg A)$$

$$+p(D|\neg B,C)p(\neg B|\neg A)p(\neg C|\neg A)p(\neg A)$$

$$= 3/4$$

Therefore, $p(A|D) = 2/3$.

Q. 3. What is defuzzification?

Ans. It is the process of **Rounding fuzzy set off** from its location in the unit hypercube to the nearest vertex in a geometric sense. If a fuzzy set is a collection of membership values or a vector of values on the unit interval then defuzzification reduces this vector to a single scalar quantity. Aggregating two or more fuzzy output sets yields a new fuzzy set in the fuzzy inference algorithm. A result in form of a fuzzy set is converted into a crisp result by this process of defuzzification.

Q. 4. In answering a question in TNPCE, a student either knows the answer or guesses. Let p be the probability that he knows the answer and (1 – *p*) be the probability that he guesses. Assume that a student who guesses the anwer will be correct with probability 1/5. What is the probability that a student knew the answer to a question given that he answered it correctly?

Ans. Let E_1 be the event that the student knew the correct answer.

E_2 be the event that student guesses the correct answer.

A be the student that the student gets the correct answer.

We need to find $P(E_1|A)$.

From the data,

$$P(E_1) = p$$

and $$P(E_2) = 1 - p$$

and $$P(A|F_2) = \frac{1}{5}$$

$$P(A|E_1) = \text{P(the student gets the correct answer given that he knew the correct answer)} = 1$$

$$\therefore \quad P(E_1|A) = \frac{P(E_1)\,P(A\,|\,E_1)}{P(E_1)\cdot P(A\,|\,E_1) + P(E_2)\cdot P(A\,|\,E_2)}$$

$$= \frac{p(1)}{p(1)+(1-p)\frac{1}{5}}$$

$$= \frac{5p}{4p+1}.$$

Q. 5. Compare JTMS and LTMS in a tabular form.

[GGIPU, B. Tech (CS)-8th Sem., 2nd Internal Test

Ans.

JTMS (by Jon Doyle)	LTMS (by Mc Allester)
1. It is a justified TMS.	1. It is a logic based TMS.
2. It is used for Non-monotonic reasoning where information is incomplete.	2. It uses default logic and non-monotonic reasoning, abductions and inheritance.

JTMS (by Jon Doyle)	LTMS (by Mc Allester)
3. Representation is in a manner where rules are used.	3. Representation is in a manner where logic is used.
4. Both logic and its contradiction are not asserted.	4. Both logic and contradictions are asserted.

EXERCISE QUESTIONS

Q. 1. (a) What are the different types of fuzzy membership function? What is the criteria to define a fuzzy membership function?

(b) How uncertainty is captured using probabilistic reasoning? How is it different from handling uncertainty using fuzzy logic?

[GGIPU, B. Tech (CSE)-8th Sem., May 2011]