BCS-404 ARTIFICIAL INTELLIGENCE (3-1-0) Cr.-04

Formalized symbolic logic: Propositional logic-first order predicate logic, wff conversion to clausal form, inference rules, the resolution principle, Dealing with inconsistencies and uncertainties, fuzzy logic.

Module - II

Probabilistic Reasoning Structured knowledge, graphs, frames and related structures, Knowledge organization and manipulation.

Module – III

Matching Techniques, Knowledge organizations, Management.

Module - IV

Natural Language processing, Pattern recognition, expert systems.

- This starting position has been generated by placing the queens such that there are no conflicts on rows or columns. The only conflict here is that the queen in column 3 (on c7) is on a diagonal with the queen in column h (on h2).
- To move toward a solution, we choose to move the queen that is on column h.
- We will only ever apply a move that keeps a queen on the same column because we already know that we need to have one queen on each column.
- Each square in column h has been marked with a number to show how many other queens that square conflicts with. Our first move will be to move the queen on column h up to row 6, where it will conflict only with one queen. Then we arrive at the position shown in below Fig
- Because we have created a new conflict with the queen on row 6 (on f6), our next move must be to move this queen. In fact, we can move it to a square where it has zero conflicts. This means the problem has been solved, and there are no remaining conflicts.
- This method can be used not only to solve the eight-queens problem but also has been successfully applied to the n-queens problem for extremely large values of n. It has been shown that, using this method, the 1,000,000 queens problem can be solved in an average of around 50 steps.
- Solving the 1,000,000-queens problem using traditional search techniques would be impossible because it would involve searching a tree with a branching factor of 10^{12}.

Fuzzy Set Theory

What is Fuzzy Set ?

- The word "fuzzy" means "vagueness". Fuzziness occurs when the boundary of a piece of information is not clear-cut.

- Fuzzy sets have been introduced by Lotfi A. Zadeh (1965) as an extension of the classical notion of set.

- Classical set theory allows the membership of the elements in the set in binary terms, a bivalent condition - an element either belongs or does not belong to the set.

  Fuzzy set theory permits the gradual assessment of the membership of elements in a set, described with the aid of a membership function valued in the real unit interval [0, 1].

- **Example:**

  Words like young, tall, good, or high are fuzzy.

  - There is no single quantitative value which defines the term young.
  - For some people, age 25 is young, and for others, age 35 is young.
  - The concept young has no clean boundary.
  - Age 1 is definitely young and age 100 is definitely not young;
  - Age 35 has some possibility of being young and usually depends on the context in which it is being considered.

Introduction

In real world, there exists much fuzzy knowledge;
Knowledge that is vague, imprecise, uncertain, ambiguous, inexact, or probabilistic in nature.

Human thinking and reasoning frequently involve fuzzy information, originating from inherently inexact human concepts. Humans, can give satisfactory answers, which are probably true.

However, our systems are unable to answer many questions. The reason is, most systems are designed based upon classical set theory and two-valued logic which is unable to cope with unreliable and incomplete information and give expert opinions.

Fuzzy Intersection A x B is defined as :
for all x in the set X,
$(A \cap B)(x) = \min [A(x), B(x)]$,

***A x B* =**

A \ B	1	2	3	4
a	0	0	0	0
b	0.2	0.8	0.8	0
c	0.2	0.6	0.6	0
d	0.2	1	0.8	0

Fuzzy Intersection ¬*A* x Y is defined as :
for all x in the set X
$(\neg A \cap Y)(x) = \min [A(x), Y(x)]$,

¬*A x Y* =

A \ y	1	2	3	4
a	1	1	1	1
b	0.2	0.2	0.2	0.2
c	0.4	0.4	0.4	0.4
d	0	0	0	0

Fuzzy Union is defined as **$(A \cup B)(x) = \max [A(x), B(x)]$ for all $x \in X$**

Therefore **R = (A x B) U (¬ A x Y)** gives

***R* =**

x \ y	1	2	3	4
a	1	1	1	1
b	0.2	0.8	0.8	0
c	0.4	0.6	0.6	0.4
d	0.2	1	0.8	0

This represents **If x is A THEN y is B** ie **$T(A \Rightarrow B) = \max (1- T(A), T(B))$**

To determine implication relations (ii) compute : (Ref : Previous slide)

Given **X = {a, b, c, d}** ,

A = {(a, 0) (b, 0.8) (c, 0.6) (d, 1)}
B = {(1, 0.2) (2, 1) (3, 0.8) (4, 0)}
C = {(1, 0) (2, 0.4) (3, 1) (4, 0.8)}

Here, the operator ⇒ represents **IF-THEN-ELSE** statement like,

IF x is A THEN y is B Else y is C, is equivalent to

R = (A x B) U (¬ A x C) and

the membership function of **R** is given by

$$\mu_R (x , y) = \max [\min (\mu_A (x) , \mu_B (y)) , \min(1 - \mu_A (x), \mu_C (y)]$$

- When a knowledge base is too large to be held in main memory, it must be stored as a file in secondary storage (disk, drum or tape).
- Storage and retrieval of information in secondary memory is then performed through the transfer of equal-size physical blocks consisting of between 256 and 4096 bytes.
- When an item of information is retrieved or stored, at least one complete block must be transferred between main and secondary memory.
- The time required to transfer a block typically ranges between 10ms and 100ms, about the same amount of time required to sequentially searching the whole block for an item.
- Grouping related knowledge together as a unit can help to reduce the number of block transfers, hence the total access time
- An example of effective grouping can be found in some expert system KB organizations
- Grouping together rules which share some of the same conditions and conclusions can reduce block transfer times since such rules are likely to be needed during the same problem solving session
- Collecting rules together by similar conditions or content can help to reduce the number of block transfers required
- **Indexed Organization**
  - While organization by content can help to reduce block transfers, an indexed organization scheme can greatly reduce the time to determine the storage location of an item
  - Indexing is accomplished by organizing the information in some way for easy access
  - One way to index is by segregating knowledge into two or more groups and storing the locations of the knowledge for each group in a smaller index file
  - To build an indexed file, knowledge stored as units is first arranged sequentially by some key value
  - The key can be any chosen fields that uniquely identify the record
  - A second file containing indices for the record locations is created while the sequential knowledge file is being loaded
  - Each physical block in this main file results in one entry in the index file
  - The index file entries are pairs of record key values and block addresses

John gave the dog the sandwich. It wagged its tail.

- In this case, a human listener would know very well that it was the dog that wagged its tail, and not the sandwich. Without specific world knowledge, the natural language processing system might not find it so obvious.
- A local ambiguity occurs when a part of a sentence is ambiguous; however, when the whole sentence is examined, the ambiguity is resolved. For example, in the sentence There are longer rivers than the Thames, the phrase longer rivers is ambiguous until we read the rest of the sentence, than the Thames.
- Another cause of ambiguity in human language is vagueness. we examined fuzzy logic, words such as tall, high, and fast are vague and do not have precise numeric meanings.
- The process by which a natural language processing system determines which meaning is intended by an ambiguous utterance is known as disambiguation.
- Disambiguation can be done in a number of ways. One of the most effective ways to overcome many forms of ambiguity is to use probability.
- This can be done using prior probabilities or conditional probabilities. Prior probability might be used to tell the system that the word bat nearly always means a piece of sporting equipment.
- Conditional probability would tell it that when the word bat is used by a sports fan, this is likely to be the case, but that when it is spoken by a naturalist it is more likely to be a winged mammal.
- Context is also an extremely important tool in disambiguation. Consider the following sentences:

    I went into the cave. It was full of bats.

    I looked in the locker. It was full of bats.

- In each case, the second sentence is the same, but the context provided by the first sentence helps us to choose the correct meaning of the word “bat” in each case.
- Disambiguation thus requires a good world model, which contains knowledge about the world that can be used to determine the most likely

- When the *(run)* command is issued to the system, it inserts a new fact into the database, which is a command to the elevator to go to floor 3.

Backward Chaining in Rule-Based Expert Systems

- A common method for building expert systems is to use a rule-based system with backward chaining. Typically, a user enters a set of facts into the system, and the system tries to see if it can prove any of the possible hypotheses using these facts.
- In some cases, it will need additional facts, in which case the expert system will often ask the user questions, to ascertain facts that could enable further rules to fire.
- The algorithm is applied as follows:
- To prove a conclusion, we must prove a set of hypotheses, one of which is the conclusion. For each hypothesis, H:

  If H is in the facts database, it is proved.

  Otherwise, if H can be determined by asking a question, then enter

  the user's answer in the facts database. Hence, it can be determined

  whether H is true or false, according to the user's answer.

  Otherwise, find a rule whose conclusion is H. Now apply this algorithm

  to try to prove this rule's antecedents.

  If none of the above applies, we have failed to prove H.

- Typically, backward chaining is used in combination with forward chaining. Whenever a new fact is added to the database, forward chaining is applied to see if any further facts can be derived.
- Backward chaining is then used to try to prove each possible hypothesis.
- Let us imagine a simple medical expert system that is designed to diagnose and treat patients' diseases. The rules might be as follows:

  Rule 1: If headache then prescribe painkiller

  Rule 2: If headache and sore throat and coughing then diagnose flu

  Rule 3: If tired and headache then diagnose glandular fever