Chapter

6

Dealing with Uncertainty

One of the most important capabilities of a human expert, and one of the most difficult to faithfully replicate in an ES, is the ability to deal effectively with imprecise, incomplete, and sometimes uncertain information.

There are many different types of uncertainty that are common in expert domains:

Uncertain knowledge. Frequently the expert will have only heuristic knowledge regarding some aspect of the domain. For example, the expert may know only that a certain set of evidence probably implies a certain conclusion.

Uncertain data. Even when we are certain of the domain knowledge, there may still be uncertainty in the data that describes the external environment. For example, when attempting to infer a specific cause from an observed effect, we may have to rely on questionable test results.

Incomplete information. It is frequently necessary to make decisions based on incomplete information. This can occur for several reasons. For example, we must make such decisions in the course of processing incrementally acquired information.

Principles of AI, and Intelligence and Expert Systems Development
- David W. Rolston. McGraw Hill International Edition, NY,
Computer Science Series

fact
9 April

Randomness. Some domains are inherently random; even though the available knowledge and information is complete and the knowledge is certain, the domain still has stochastic properties.

ESs using current technology are not capable of dealing with uncertainty as effectively as their human counterparts, and this subject remains an important research topic. The following sections describe several methods for handling uncertainty. Although, as noted above, these methods will not completely duplicate human capabilities, each has proved to be useful in the development of actual systems.

6.1 Reasoning Based on Partial Information

Reasoning systems based on predicate logic (as described in Chap. 3) are conceptually elegant and intellectually appealing because they are precise and rigorous. By using formal logic, truth can be given or derived with equal assurance. Once established, truth is always true. Moreover, derived truth will never produce a contradiction, given that no contradictions exist within the axioms.

Because of these characteristics, predicate logic is a *monotonic* reasoning system. "Monotonic," which means to "move in one direction only," is used in describing a predicate logic system to convey the idea of a reasoning process that moves in one direction only—that of continuously adding additional truth.

PRINCIPLE 6.1: In a monotonic reasoning system the number of facts known to be true at any specified time is always increasing, never decreasing.

Unfortunately, although they provide a basis for consistent, reliable inference, these characteristics also limit the extent to which pure logic systems can be applied in the real world. As introduced in Sec. 6.1, reasoning processes that are to be applied to practical, unstructured problems must recognize at least the following:

- Available information is frequently incomplete, at least at any given decision point.
- Conditions change over time.
- There is frequently a need to make an efficient, but possibly incorrect, guess when reasoning reaches a dead end.

6.2 Nonmonotonic Reasoning

In dealing with these difficulties, human problem solvers often augment absolute truth with beliefs that are subject to change given fur-

class. A f

ther information. These *tentative beliefs* are generally based *assumptions* that are made in light of the lack of evidence to the c trary. For example, even though the decision to board a commer airliner is very important, when we actually make the decision, w will probably assume that the pilot is competent and that the aircraft is airworthy, unless there is something that indicates otherwise. A *nonmonotonic reasoning system* (NMRS) is based on this concept.

PRINCIPLE 6.2: A nonmonotonic reasoning system tracks a set of tentative beliefs and revises those beliefs when new knowledge is observed or derived.

An NMRS typically includes a set of *premises* that are held to be immutably true (analogously to the axioms in a predicate logic system). In addition to premises, the system keeps a collection of tentative beliefs: pieces of knowledge that are explicitly recognized to be potentially incorrect because they are either assumptions or beliefs inferred from assumptions. For each tentative belief the system maintains a *dependency record* that tracks the belief vs. its *justification:* the facts, beliefs, and inferences that were used to generate the tentative belief.

Example 6.1 Consider the following rather contrived situation: A woman who has never been out of the United States wakes from unconsciousness on a deserted bank of a very large river in unfamiliar territory. She can remember (somehow) only that she is near a large city and that the city is at the mouth of the river. Although it is dark, she can sense that the area surrounding her is warm, humid, and tropical.

Given that she can recall the facts listed below (from past study of world geography) and that she assumes she is in the United States (given the lack of evidence to the contrary), describe the set of beliefs and associated justifications that she would probably develop.

```
fact 1    The sky is blue.
              .
              .
              .
fact 37   The Amazon river—
          A   flows from west to east
          B   empties into the South Atlantic Ocean
          C   ends at Almirim, Brazil
fact 38   Most of the area near the mouth of the Amazon is warm, hu-
          mid, and tropical.
fact 39   The mouth of the Nile is in Egypt.
fact 40   The Columbia River—
          A   flows from east to west
          B   empties into the Pacific Ocean
          C   ends at Astoria, Oregon
fact 41   The Nile river—
          A   flows from south to north
          B   empties into the Mediterranean Sea
          C   ends at Cairo, Egypt
fact 42   Most of the area near the mouth of the Nile is hot and dry
          with sparse vegetation.
```

```
fact 43    The predominate language in the United States is English.
fact 44    The Mississippi river-
           A    flows from north to south
           B    empties into the Gulf of Mexico
           C    ends at New Orleans, Louisiana
fact 45    The predominant language in Brazil is Portuguese.
fact 46    The Columbia river is in the United States.
fact 47    The Mississippi river is in the United States.
fact 48    The predominant language in Egypt is Arabic.
fact 49    Most of the area surrounding the mouth of the Columbia river
           is cool, moist, and includes large conifers.
fact 50    The Amazon river is in Brazil.
fact 51    Most of the area surrounding the mouth of the Mississippi
           river is warm, humid, and tropical.
           .
           .
fact N     It is dark after sunset.
```

Given the above information, the woman could, using the following beliefs, quite reasonably conclude that she was traveling south, toward New Orleans and the Gulf of Mexico:

Belief number	Belief	Justification
6	Traveling south	Fact 44A and belief 3
5	Traveling toward Gulf of Mexico	Fact 44B and belief 3
4	Traveling toward New Orleans	Fact 44C and belief 3
3	The river is the Mississippi	Fact 47, fact 51, belief 2, and belief 1
2	The surrounding area is warm, humid, and tropical	External observation
1	The area is in the United States	Assumption based on lack of contrary information

Because an NMRS includes tentative knowledge, it is possible to add new piece of knowledge that will cause a previously believed tentative uth to become false.

INCIPLE 6.3: When a belief in an NMRS is revised, then any beliefs at rest on it, directly or indirectly, must also be revised.

The *belief revision* portion of an NMRS propagates the effect of any inge in belief through the use of *dependency-directed backtracking*.

xample 6.2: Given the situation from Example 6.1, what belief revisions would cur as a result of the woman's finding several public notices written in Portuguese?

Having observed such a sign, the woman would be forced to replace the assumption that she is in the United States with the belief that she is in Brazil. The cascading effect of this change is illustrated by the following set of revised beliefs:

Belief number	Belief	Justification
6	Traveling east	Fact 37A and belief 3
5	Traveling toward South Atlantic	Fact 37B and belief 3
4	Traveling toward Almirim, Brazil	Fact 37C and belief 3
3	The river is the Amazon	Fact 38, fact 50, belief 2, and belief 1
2	The surrounding area is warm, humid, and tropical	External observation
1	The area is in Brazil	Fact 45

Her conclusion, which changed radically as the result of adding a small piece of knowledge, is now that she is heading east on the Amazon toward Almirim and the South Atlantic Ocean.

PRINCIPLE 6.4: The addition of a small piece of new knowledge can result in a great deal of belief revision.

The increased power and flexibility of default reasoning makes an NMRS especially useful for problem-solving domains, such as planning and design, that require a large number of tentative assumptions based on partial information. This increased usefulness is not, however, without cost. Specifically, an NMRS can require a large amount of memory to store the dependency information and a large amount of processing time to propagate changes in beliefs.

For additional information on nonmonotonic reasoning, See D. McDermott, 1980, Winograd, 1980, and Davis, 1980.

6.3 Truth Maintenance System

The Truth Maintenance System (TMS) is an implementation of an NMRS [Doyle, 1979*a* and 1979*b*]. TMS operates as a knowledge base management system and is called every time the reasoning system generates a new truth value. TMS, using belief revision, takes any action required to modify dependent beliefs to maintain consistency in the knowledge base. TMS's role is purely passive; it never initiates the generation of inferences.

In TMS a *node* represents one unit of knowledge: a fact, rule, assertion, etc. At any point in execution, every node is in one of two conditions:

IN Currently believed to be true.

OUT Currently believed to be false. A node can be OUT because there is no possible condition that would make it true or because the conditions required to make it true are not currently IN.

Support list justifications

Associated with each node are *justifications* for the node's truth value. (We will assume one justification per node for simplicity's sake.) For each node that is IN, TMS records a *well-founded support:* proof of the validity of the node, starting from the system's facts and justifications.

The simplest form of justification is a *support list* (SL), which has the following form:

[SL (in_nodes)(out_nodes)]

where (in_nodes) ::= list of all nodes that must be IN for this node to be true

(out_nodes) ::= list of all nodes that must be OUT for this node to be true

Example 6.3 Develop a set of nodes, and associated SL justifications, for the following:

1. It is sunny.
2. It is daytime.
3. It is raining.
4. It is warm.

Node number	Knowledge	Justification
1	It is sunny.	[SL (2) (3)]
2	It is daytime.	[SL ( ) ( )]
3	It is raining.	[SL ( ) (1)]
4	It is warm.	[SL(1)(3)]

The content of the IN and OUT lists at any time describe a snapshot of the system's belief. As system execution proceeds, the content of each of these lists is changed by TMS to ensure consistency.

Example 6.4 Given the nodes and justifications from Example 6.3 and the following IN and OUT lists:

IN 1, 2, 4
OUT 3

What is the effect of the reasoning system passing to TMS the following: "It is raining."?

The content of the IN and OUT lists will become:

IN 3, 2
OUT 1, 4

Several different types of nodes can be justified by a support list.

Premise. A premise is a fact that is always valid. The in_list and out_list portions of the SL for a premise will always be empty. Given that we ignore time of day, node 2 in Example 6.3 is a premise.

Normal deduction. A normal deduction is an inference that is formed in the normal sense of a monotonic system. The out_list portion of the SL for a normal deduction will always be empty; the belief in the deduction follows from the belief in the nodes listed in the in_list portion of the SL. Example 6.3 could be extended to include a belief, "It is wet," with an SL of [SL (3) ( )] that would be an example of a normal deduction.

Assumption. An assumption is a belief that is supported by the lack of contrary information. The out_list portion of the SL for an assumption will never be empty. The nodes that are on in_list can be viewed as the *reasons* for making the assumptions, and the nodes on the out_list are the nodes whose presence would provide contrary information that would invalidate the assumption.

Assumptions provide the basis for default reasoning in TMS. Node 1 in Example 6.3 is an example of an assumption. Its justification can be interpreted as, "Assume that it is sunny given that it is daytime and that nothing suggests that it is raining."

Conditional proof justifications

A *conditional proof* (CP) is a second type of justification that is used to support hypothetical reasoning. The format of a CP is:

[CP(consequent) (in_hypothesis)]

(A CP also includes an out_hypothesis entry, but it is almost alwa[ys] empty and is ignored in this discussion.) A CP justification is valid [if] and only if the consequent node is IN whenever all of the nodes [in] in_hypothesis are IN. The usefulness of such justification is illustra[ted] in the following section.

Dependency-directed backtracking in TMS

When TMS discovers an inconsistency in the current set of beliefs, [as] a result of a newly added justification, it invokes dependency-direc[ted] backtracking to restore consistency. This backtracking activity is based on the following principle:

PRINCIPLE 6.5: Contradictions in a set of beliefs occur as a result of incorrect assumptions.

To restore consistency, TMS must retract belief in one or more assumptions. This process consists of the following:

1. A node is marked as a *contradiction* when belief in it causes inconsistencies. Well-founded support in a contradiction is unacceptable, and, therefore, backtracking must continue until the contradiction becomes OUT.

2. The backtracker traces backward through the well-founded support for the contradiction, attempting to find a possible cause. Because all normal deductions are correct, the cause of the inconsistency must be a bad assumption, and, therefore, the backtracker looks only for such assumptions. The result of this activity is the accumulation of the set of suspects

$$S_a = \{A_1, A_2, A_3, \ldots, A_n\}$$

the assumptions in the well-founded support for the contradiction.

3. The backtracker creates a new node, *nogood,* that indicates that S_a is inconsistent [e.g., representing the fact

$$\sim(A_1 \wedge A_2 \wedge A_3 \wedge \ldots \wedge A_n)]$$

S_a is then called the *nogood-set.*

The justification for the nogood node is generally represented by using the CP form:

node# nogood [CP (contradiction) (S_a)]

We know that "nogood" should be IN because, if all of the nodes in the nogood-set S_a are simultaneously IN, then the contradiction node will also be IN. This stores the relationship between the assumptions and the contradictions. (The use of the term "condition proof" for this form of justification corresponds to this usage. The truth of the nogood node is implied as a result of a specific type of proof technique—called conditional proof—of the contradiction from the nogood-set [Klenk, 1983].)

4. The backtracker uses the information from the nogood-set to identify an assumption that must be retracted to resolve the inconsistency and so remove the contradiction.

Example 6.5 Assuming that each fact from Example 6.1 is now represented as a premise node (where node numbers for nodes shown in this list correspond to the fact numbers from Example 6.2), the following set of beliefs are IN—in addition to nodes 37 to 51 (which are premises):

Node number	Belief	Justification
6	Traveling south	[SL (3 , 44A) ( )]
5	Traveling toward Gulf of Mexico	[SL (3 , 44B) ( )]
4	Traveling toward New Orleans	[SL (3 , 44C) ( )]
3	The river is the Mississippi	[SL (26,2,47,51) ( )]
2	The surrounding area is warm, humid, and tropical	[SL ( ) ( )]
26	The area is in the United States	[SL (21) ( )]
21	The predominant language is English	[SL ( ) (20, 22)]

The following nodes are OUT:

24 The area is in Brazil.
25 The area is in Egypt.
20 The predominant language is Portuguese.
21 The predominant language is Arabic.

After the observation of the Portuguese signs occurs, a contradiction is eventually observed and a contradiction node is added:

50 contradiction-1 [SL (26, 24) ( )]

Backtracking is invoked as a result of the recognition of the contradiction. The backtracker looks backward through the well-founded support for node 50 in search of an assumption that could have caused the inconsistency. The backtracker creates a node, nogood-1, by using a CP justification to represent the inconsistent assumption:

56 nogood-1 [CP 55 (21,20)]

This node can be interpreted as, "The contradiction called contradiction-1 results from assuming node 21." To remove the assumption, the backtracker must move a node (from the out_list of the assumption) from OUT to IN. Once this move occurs, the faulty assumption will be forced out. Any nodes that depended on the faulty assumption will also be forced OUT. In this case all of the IN list will be removed (except for nodes 20 and 2 and the premises), and a new IN list will be generated.

Several modifications of Doyle's TMS have been developed (e.g., Charniak, 1979, Martins, 1984, and Petrie, 1985.) For a history of such development see Doyle, 1980.

6.4 Reasoning Based on Probability

The techniques of *probability theory* have been widely used in many different disciplines in attempts to quantify uncertainty. The appeal of probability is based in part on the fact that it has a solidly established mathematical basis; techniques for using probabilities are widely published.

$P(E)$ is the probability that an event E will occur; it represents a quantification of the likelihood of this occurrence. In most cases the value of $P(E)$ is established by statistical analysis (e.g., measuring the frequency of occurrence of E in a random series of tests).

Probabilities have a value from 0 to 1, where 1 represents absolute knowledge of E and 0 represents absolute knowledge that E will not occur. When an event being considered has several possible outcomes (e.g., rolling a pair of dice) a probability is associated with each outcome. The sum of the probabilities for all possible outcomes for an event must equal 1.

An *objective probability* is a probability that is measured by using frequency ratio techniques described above. Unfortunately, in most ES domains it is impossible to preform such measurements. It is much more common to attempt to collect estimates of the probability values, called *subjective probabilities,* by interviewing experts.

There are several problems that make it difficult to use probability for dealing with uncertainty in ES. For example, even though a person is an expert in a domain, it can still be very difficult for that person to accurately estimate probabilities.

Because the actions of an expert system are typically the result of piecing together many different fragments of knowledge, each with different probability characteristics, we must be able to combine probabilities. Probability values can be combined by using many established techniques. For example, a commonly used formula

$$P(E_1 \text{ and } E_2) = P(E_1) * P(E_2)$$

can be used to find the probability that both E_1 and E_2 will occur, given the individual probabilities of E_1 and E_2. For example, if the probability of drawing an ace of spades from a deck of cards is $1/52$, and the probability of drawing any diamond from a separate deck is $13/52$, then the probability of drawing both the ace of spades and a diamond is $1/52 * 13/52$.

Bayes's rule is used for more complicated situations. It employs the following terms:

B_i ::= a specific belief
$P(B_i)$ ::= the a priori probability that B_i is true; this is the probability that B_i is true, given no specific evidence
$P(B_i|E)$ ::= the *conditional probability* that B_i is true, given evidence E; this indicates our revised belief in B_i upon finding that E is true.
$P(E|B_i)$ ::= the probability of of observing E, given that B_i is true
k ::= total number of possible beliefs

Bayes's rule states:

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

A typical use of Bayes's rule is in diagnostic domains. In such cases, B_i represents a cause of the effect E. $P(B_i|E)$ is, in effect, inferred belief that B_i is the cause of effect E. We can use this relation to quantify the extent to which we should believe B_i given the available evidence E. For this reason, Bayes's rule is often referred to as the *probability of causes theorem* [Spiegel, 1975].

Probability theory has been successfully applied in several ESs, for example, Page-1 [Strandberg, 1985] and most notably the Prospector system [Duda, 1980].

6.5 Certainty Factors

A *certainty factor* (CF) is a relatively informal mechanism for quantifying the degree to which, based on the presence of a given set of evidence, we believe (or disbelieve) a given conclusion. Certainty factors have been most widely applied to domains that use incrementally acquired evidence.

The concept of CF was developed for Mycin [Buchanan, 1984*a*], and it has been used successfully in many other systems [Buchanan, 1983]. For Mycin it was decided, after much discussion, to use a new technique for certainty quantification rather than the more traditional probabilistic methods, primarily because of the difficulty of accurately estimating the a priori and conditional probabilities required for the application of Bayes's rule. (See Adams, 1984, for a discussion of the use of CF vs. formal probability.)

Description of certainty factors

A CF is a numerical value that expresses the extent to which, based on a given set of evidence, we should accept a given conclusion. A CF with a value of 1 indicates total belief, whereas a CF with a value of −1 indicates total disbelief.

Component certainty factors

For each rule in the system, a CF is assigned by the domain expert. It is based on the expert's knowledge and experience.

PRINCIPLE 6.6: A CF is a subjective quantification of an expert's judgment and intuition.

Example 6.6 The following simple rule for diagnosing the failure of a table lamp illustrates the use of a CF (in this case, 0.8).

IF a table lamp is plugged into a receptacle,
the receptacle has current, and
the lamp's switch is on,
THEN there is suggestive evidence (0.8) that the light bulb is faulty.

The CF that is included in a rule is a *component certainty factor* (CF_{comp}), and it describes the credibility of the conclusion, given only the evidence represented by the preconditions of the rule.

Typically, in a large rule-based system, many different rules will relate to the same conclusion. The following principle relates an important constraint regarding these multiple rules:

PRINCIPLE 6.7: In a system that uses CFs, the rules must be so structured that any given rule either adds to belief in a given conclusion or adds to disbelief.

A *measure of belief* $MB[c,e]$ is a number that indicates the degree to which our belief in conclusion c is increased, based on the presence of evidence e. By definition: $0 \le MB[c,e] \le 1$.

Similarly, a *measure of disbelief*, $MD[c,e]$, is a number that indicates the degree to which our disbelief in c is increased, based on the presence of e.

Because of the restriction described in Principle 6.7, for any given rule if $MB[c,e] > 0$, then $MD[c,e] = 0$, and if $MD[c,e] > 0$, then $MB[c,e] = 0$. The component certainty factor can now be more formally described as:

$$CF_{comp}[c,e] = MB_{comp}[c,e] - MD_{comp}[c,e] \tag{6.1}$$

where: c ::= the conclusion under consideration
e ::= the evidence relating to c

Note that either the MB term or the MD term must equal zero.

Because there are many rules that relate to any given conclusion, each of which can add to our *overall* belief or disbelief in a conclusion, a *cumulative certainty factor* is used to express the certainty of the conclusion, at a given point in execution, in light of *all* of the evidence that has been considered up to that point.

Calculation of certainty factors

The cumulative certainty factor, which provides a means of assessing the certainty of a conclusion from a global viewpoint, is formed by combining global degrees of belief and disbelief represented by the *cumulative measure of belief* and the *cumulative measure of disbelief* for the conclusion.

Specifically, a cumulative certainty factor is defined, for a specified point during system execution, as follows:

$$CF[c,e_c] = MB[c,e_f] - MD[c,e_a] \tag{6.2}$$

where c ::= the conclusion under consideration
e_c ::= *all* of the evidence that relates to c that has been considered up to the specified point of execution
$CF[c,e_c]$ ::= the cumulative certainty factor for c given e_c (the net belief in the conclusion, given the current evidence)
e_f ::= all of the evidence *for* c that has been considered
e_a ::= all of the evidence *against* c that has been considered
$MB[c,e_f]$ ::= cumulative measure of belief for c given e_f
$MD[c,e_a]$ = cumulative measure of disbelief for c given e_a

The above definition implies the need for calculating MB and MD for each possible conclusion in the system. This calculation is performed by first initializing both terms to zero and then incrementally including the effect of each applicable rule. Every time an additional rule is considered, a new MB and MD is calculated on the basis of the effect of the new rule combined with the existing MB and MD.

Combining functions for performing this activity are based on the constraint that the collections of evidence being considered are independent. For example, in an automotive diagnostic system, "a low battery" and "dim lights" do not individually contribute new evidence because they routinely occur together. All pieces of *related* evidence must occur in the same rule.

The measure of belief that results from considering two sources of evidence can be calculated by using the following formula:

$$\text{MB}[c,s_1\&s_2] = \qquad (6.3)$$

$$\begin{array}{ll} \text{IF} & \text{MD}[c,s_1\&s_2] = 1, \text{ THEN } 0 \\ \text{ELSE} & \text{MB}[c,s_1] + \text{MB}[c,s_2](1 - \text{MB}[c,s_1]) \end{array}$$

where $\text{MB}[c,s_1\&s_2]$ ::= the measure of belief based on a pair of sources

In the elementary case, s_1 and s_2 are simply two individual rules r_1 and r_2. In general, s_1 may represent a set of rules whose cumulative effects have previously been considered and s_2 represents a new rule whose effect is to be added to the previously existing cumulative belief. ($\text{MD}[c,s_1\&s_2]$ is the measure of disbelief for the same pair of sources and is equal to 1 if and only if the conclusion is known to be false with absolute assurance).

Similarly, MD is defined by using

$$\text{MD}[c,s_1\&s_2] = \qquad (6.4)$$

$$\begin{array}{ll} \text{IF} & \text{MB}[c,s_1\&s_2] = 1, \text{ THEN } 0 \\ \text{ELSE} & \text{MD}[c,s_1] + \text{MD}[c,s_2](1 - \text{MD}[c,s_1]) \end{array}$$

The reasonableness of this function is clear when we recognize that the addition of new evidence that supports belief in a conclusion increases, but does not absolutely establish, the credibility of the conclusion. As a large number of elements of supporting evidence are combined, the overall MB grows asymptotically toward unity.

The factor

$$\text{MB}[c,s_2](1 - \text{MB}[c,s_1])$$

in Eq. (6.3) describes the contribution to MB provided by a new piece of evidence. This factor can be viewed as a measure of the extent to which the new evidence mitigates the doubt that remained after the previous evidence had been considered. The degree of mitigation is, quite reasonably, proportional to the strength of the new evidence. A similar argument holds for MD.

Example 6.7 Given that there are only four rules that suggest conclusion c, find the cumulative certainty factor for c, given the following component certainty factors:

Rule	CF
1	0.8
2	0.3
3	−0.2
4	0.7

For rule 1, using Eqs. (6.1) and (6.2):

$$\text{MB} = \text{MB}_{\text{comp}} = 0.8 \qquad \text{MD} = \text{MD}_{\text{comp}} = 0$$

Equation (6.2) is then used to include the effect of rule 2:

$$\text{MB} = 0.8 + 0.3(1 - 0.8) = 0.86 \qquad \text{MD} = 0$$

After considering rule 3,

$$\text{MB} = 0.86 \qquad \text{MD} = -0.2$$

Finally, the effect of rule 4 is included:

$$\text{MB} = 0.86 + 0.7(1 - 0.86) = 0.96 \qquad \text{MD} = -0.2$$

and the final certainty factor is developed:

$$\text{CF} = 0.96 - 0.2 = 0.76.$$

The following formulas are used to calculate MB and MD on the basis of the combination of separate conclusions rather than separate evidence:

Conjunction of conclusions:

$$\text{MB}[c_1 \text{ or } c_2,e] = \overset{\text{min}}{\text{max}}(\text{MB}[c_1,e], \text{MB}[c_2,e]) \qquad (6.5)$$

$$\text{MD}[c_1 \text{ or } c_2,e] = \min(\text{MD}[c_1,e], \text{MD}[c_2,e]) \qquad (6.6)$$

where c_1 = conclusion 1
c_2 = conclusion 2
e = all available evidence

Disjunctions of conclusions:

$$\text{MB}[c_1 \text{ or } c_2,e] = \max(\text{MB}[c_1,e], \text{MB}[c_2,e]) \qquad (6.7)$$

$$\text{MD}[c_1 \text{ or } c_2,e] = \max(\text{MD}[c_1,e], \text{MD}[c_2,e]) \qquad (6.8)$$

Example 6.8 Given that with TR, the results of an automotive diagnostic test, the following conclusions can be drawn, each with the associated CF:

C1 The problem requires immediate attention (0.8).
C2 There is a problem in the electrical system (0.6).
C3 There is a short in the electrical system (0.4).
C4 There is a fault in the flow control computer (0.2).

Find the measure of belief that there is a problem in the electrical system that requires immediate attention and that the problem is either a short or a computer fault.

Using Eqs. (6.5) and (6.7):

$$\begin{aligned}\text{MB[C1\&C2\&(C3 or C4),TR]} &= \min(\text{MB[C1,TR], MB[C2,TR],}\\ &\qquad \text{MB[C3 or C4,TR]})\\ &= \min(\text{MB[C1,TR], MB[C2,TR],}\\ &\qquad \max(\text{MB[C3,TR], MB[C4,TR]}))\\ &= \min(0.8,0.6,\max(0.4,0.2))\\ &= \min(0.8,0.6,0.4)\\ &= 0.4\end{aligned}$$

The domain expert assigns a CF to a rule based on the assumption that all conditions in the premise are known with certainty. If there is uncertainty regarding the conditions, then the CF normally associated with the rule is not fully applicable. In this case, the following formulas can be used to calculate the revised CF for the rule. The calculation is based on a CF that describes our degree of belief in the required condition (i.e., the evidence for the rule).

$$\text{MB}[c,s] = \text{MB}'[c,s] * \max(0,\text{CF}[s,e]) \tag{6.9}$$

$$\text{MD}[c,s] = \text{MD}'[c,s] * \max(0,\text{CF}[s,e]) \tag{6.10}$$

where c ::= conclusion of the rule
s ::= evidence required for the rule
MB′[c,s] ::= MB in the conclusion, given complete confidence in s
CF[s,e] ::= actual degree of belief in s (established by some prior evidence e)

The situation described above typically occurs in one of two ways:

1. There is less than total confidence in the evidence when it is provided to the system (e.g., the evidence represents the conclusions of a test that had mixed results).
2. The evidence is a conclusion that resulted from previous execution of a different rule with which a CF was associated.

The second case frequently occurs during the process of chaining rules. To calculate the CF for the final conclusion in a chaining process, it is generally necessary to use both the combining functions for strength of evidence [Eqs. (6.9) and (6.10) and the combining functions for conjunction and disjunction of hypothesis [Eqs. (6.7) and (6.8).].

Example 6.9 Given the following rules from an ES for diagnosis of electrical problems:

1. IF a system fault has been reported and the WPLMI board is reporting low voltage,
 THEN there is suggestive evidence (0.7) that there is a problem in the power supply in the WPLMI board.
2. IF there is a problem in the power supply to the WPLMI board, the CPU port has been closed, and the voltage at the input to the CPU is less than 4.5 V,
 THEN there is suggestive evidence (0.9) that the CPU power supply has failed.
3. IF the CPU is not responding to port commands and a system fault has been reported,
 THEN there is suggestive evidence (0.4) that the CPU port has been closed.
4. IF a system fault has been reported and the WPLMI board is the source of the system fault,
 THEN there is suggestive evidence (0.6) that there is a problem in the power supply to the WPLMI.

Given also the following observations, each of which is known with complete certainty:

E1 A system fault has been reported.
E2 The CPU is not responding to port commands.
E3 The voltage at the input to the CPU is 3.8 V.
E4 The WPLMI board is the fault source.

find the certainty with which we can conclude that the CPU power supply has failed.

From rule 1, we conclude C1—there is a problem in the power supply to the WPLMI board—with a CF of 0.7. Rule 4 is then applied to C1:

$$\text{CF}_{c1} = 0.7 + 0.6(1 - 0.7) = 7.2$$

Using rule 3, we conclude C2—the CPU port has been closed—with a CF of 0.4.

We can conclude C3—the voltage at the input to the CPU is less than 4.5 V—with absolute certainty, based on E3. By using these conclusions as evidence for rule 2, the rule, including the CFs for each condition, can now be written as:

IF there is a problem in the power supply to the WPLMI board (7.2),
 the CPU port has been closed (0.4), and
 the voltage at the input to the CPU is less than 4.5 V (1.0),
THEN there is suggestive evidence (0.9) that the CPU power supply has failed.

The measure of belief in the desired conclusion can now be characterized as:

$$MB[C_{final},C1\&C2\&C3]$$

$$= MB'[C_{final},C1\&C2\&C3] * \max(0,CF[C1\&C2\&C3,E_p])$$

where E_p ::= previous evidence (in this case, E1, E2, E3, and E4)

$$MB[C_{final},C1\&C2\&C3] = 0.9 * \max(0,CF[C1\&C2\&C3,E_p])$$

Using Eq. (6.5):

$$CF[C1\&C2\&C3,E_p] = \min(CF[C1,E_p], CF[C2,_p], CF[C3,E_p])$$

$$= \min(7.2,0.4,1.0)$$

$$= 0.4$$

$$MB[c_{final},C1\&C2\&C3] = 0.9 * \max(0,0.4)$$

$$= 0.36$$

$$CF[c_{final},C1\&C2\&C3] = 0.36 - 0 = 0.36$$

The use of CFs as described in this section is essentially a technique for supplementing existing reasoning processes with information regarding uncertainty.

6.6 Fuzzy Reasoning

Fuzzy reasoning is specifically designed to deal with the inexactness (or fuzziness) that is present in the knowledge used by human experts [Zadeh, 1965].

In conventional mathematics, a set has clearly defined boundaries that specifically identify an exact, although potentially infinite, group of elements (e.g., the set of integers larger than 2 and smaller than 10). Given a correct understanding of a set's definition, it is possible to determine whether a given candidate is or is not a member of the set. Unfortunately, it is very difficult, if not impossible, to develop exact set definitions for many of the concepts and classification mechanisms that are used by humans:

PRINCIPLE 6.8: Humans routinely and subconsciously place things into classes whose meaning and significance are well understood but whose boundaries are not well defined.

In fuzzy reasoning the concept of a *fuzzy set* corresponds to such a class. A fuzzy set is a class of elements with loosely defined boundaries.

Examples of fuzzy sets include the set of "large cars," "fast horses," and "rich people." To identify the members of a fuzzy set, we associate a *grade of membership* with each element that could potentially be a member. The grade of membership is a number, between 0 and 1, that indicates the extent to which an element is a member of the set. A grade of 1 indicates that the candidate is definitely a member, whereas a grade of 0 indicates that the candidate is definitely not a member. The transition between these extremes is gradual rather than distinct.

PRINCIPLE 6.9: Grades of membership for fuzzy sets are subjectively assigned on the basis of context.

For example, we might assign a value of 0.2 as the grade of membership for a 28-year-old college freshman in the set of "young college freshmen." Note that there is no deterministic procedure for establishing the validity of this value; rather it can be viewed as an intuitive statement, expressed on a scale of 0 to 1, of the extent to which the label "young college freshman" applies to the subject person. Overall, a fuzzy set is defined by a *membership function* that associates a grade of membership with each candidate element. A membership function for the example mentioned above is shown in Fig. 6.1.

In some cases the grade of membership may itself be inexactly represented as a *fuzzy number* (e.g., a membership grade that is "near 0.2" or "about 0.5"). A fuzzy set that uses fuzzy values in its membership function is called an *ultrafuzzy set*. An example of a membership for an ultrafuzzy set is shown in Fig. 6.2.

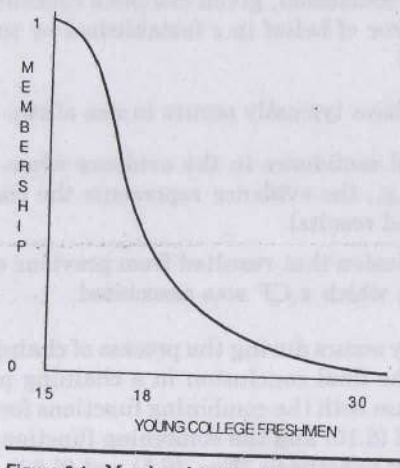


Figure 6.1 Membership function for a fuzzy set.

ist horses,"
ve associate
ntially be a
and 1, that
f the set. A
er, whereas
a member.
ian distinct.

ojectively as-

ade of mem-
oung college
re for estab-
s an intuitive
to which the
son. Overall,
iates a grade
ship function

nexactly rep-
that is "near
s membership
embership for

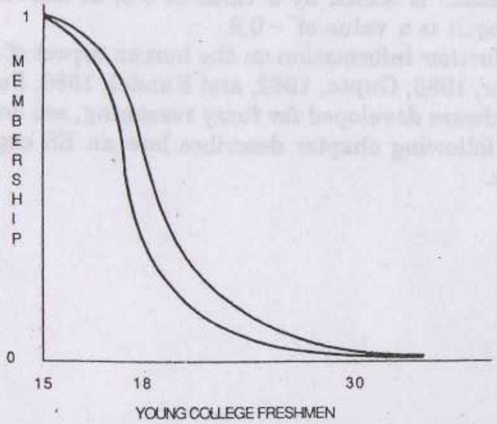


Figure 6.2 Membership function for an ultrafuzzy set.

Conceptually, membership grades are not intended to correspond to probabilities. Rather than representing the probability of an element being a member of a given set—based on an analysis of randomness in the data—a grade of membership is intended to quantify an intuitive understanding of the extent to which a given element is compatible with a given concept.

Manipulation of fuzzy sets

Fuzzy logic is a well-defined reasoning system that is based on the use of fuzzy sets rather than on the binary values associated with traditional bivalent logic.

PRINCIPLE 6.10: In a fuzzy logic system only the elements being manipulated are fuzzy; the rules of logic are well defined.

The task of translating human expressions in fuzzy logic systems is relatively simple because humans tend to communicate ideas and quantifications by using verbal rather than numeric descriptions. These verbal expressions are based on the use of *linguistic variables* such as big, old, and fast. To translate these values in a fuzzy logic system, we simply describe a membership function that represents them. Once the basic membership functions have been established, *fuzzy modifiers* can be used to modify a membership function (e.g., very old). Figure 6.3 illustrates the use of the modifiers "very" and "not."

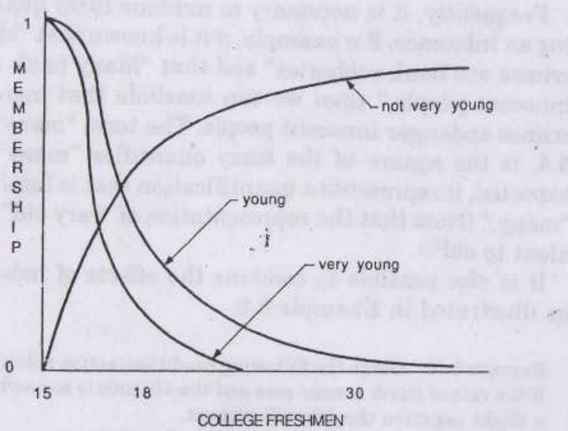


Figure 6.3 Fuzzy modifiers.

Fuzzy quantifiers are used to represent approximate quantification in fuzzy logic. Fuzzy quantifiers are frequently used for *dispositions*, statements that include implied fuzzy quantification. For example, the statement, "Horses have tails," actually implies the use of the fuzzy quantifier "most" (e.g., "Most horses have tails").

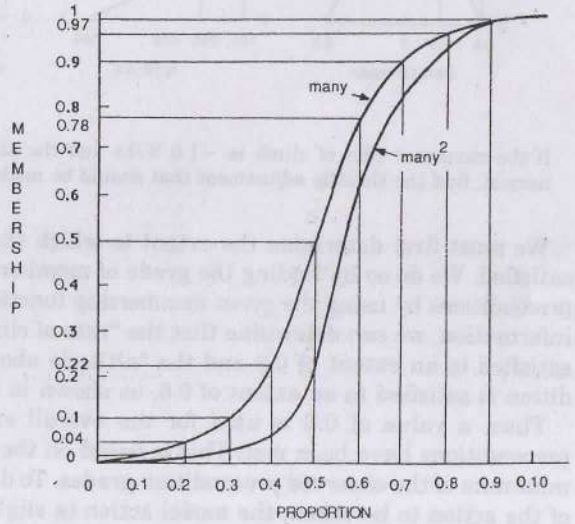
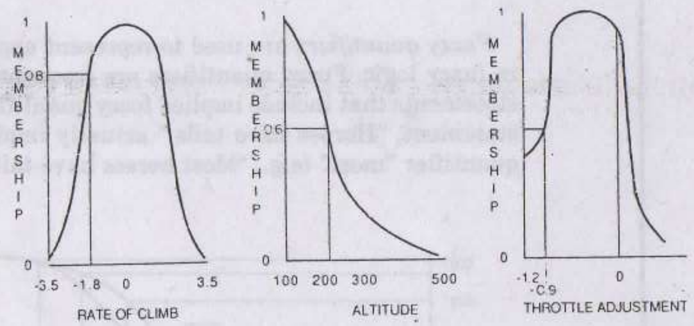


Figure 6.4 Manipulation of fuzzy quantifiers.

Frequently, it is necessary to combine fuzzy quantifiers when forming an inference. For example, if it is known that "Many of A. B. Jones's crimes are bank robberies" and that "Many bank robberies endanger innocent people," then we can conclude that many^2 of A. B. Jones's crimes endanger innocent people. The term "many^2", as shown in Fig. 6.4, is the square of the fuzzy quantifier "many" and, as would be expected, it represents a quantification that is fuzzier than the original "many." (Note that the representation of "very old" in Fig. 6.3 is equivalent to old^2.)

It is also possible to combine the effects of independent fuzzy sets, as illustrated in Example 6.9.

Example 6.10 Given the following condition/action rule (for an aviation domain): If the rate of climb is near zero and the altitude is somewhat too high, then make a slight negative throttle adjustment.
Given also the following membership functions:



If the measured rate of climb is −1.8 %/hr and the altitude is 200 ft above normal, find the throttle adjustment that should be made.

We must first determine the extent to which the preconditions are satisfied. We do so by finding the grade of membership for each of the preconditions by using the given membership functions. From the given information, we can determine that the "rate of climb" precondition is satisfied to an extent of 0.8 and the "altitude above normal" precondition is satisfied to an extent of 0.6, as shown in the above figure.

Thus, a value of 0.6 is used for the overall extent to which the preconditions have been met. This is based on the fact that 0.6 is the minimum of the observed precondition grades. To determine the extent of the action to be taken, the model action (a slight negative throttle

adjustment) is scaled by a value of 0.6, as shown in the figure. The final result is a value of −0.9.

For further information on the human aspect of fuzzy reasoning see Bandler, 1983, Gupta, 1982, and Kandel, 1986. For more information on hardware developed for fuzzy reasoning, see Johnson, 1985.

The following chapter describes how an ES explains its reasoning process.