Fuzzy Logic

Lofti A. Zadeh
University of California, Berkeley

Fuzzy logic — the logic underlying approximate, rather than exact, modes of reasoning — is finding applications that range from process control to medical diagnosis.

Logic, according to Webster's dictionary, is the science of the normative formal principles of reasoning. In this sense, fuzzy logic is concerned with the formal principles of approximate reasoning, with precise reasoning viewed as a limiting case.

In more specific terms, what is central about fuzzy logic is that, unlike classical logical systems, it aims at modeling the imprecise modes of reasoning that play an essential role in the remarkable human ability to make rational decisions in an environment of uncertainty and imprecision. This ability depends, in turn, on our ability to infer an approximate answer to a question based on a store of knowledge that is inexact, incomplete, or not totally reliable. For example:

(1) Usually it takes about an hour to drive from Berkeley to Stanford and about half an hour to drive from Stanford to San Jose. How long would it take to drive from Berkeley to San Jose via Stanford?

(2) Most of those who live in Belvedere have high incomes. It is probable that Mary lives in Belvedere. What can be said about Mary's income?

(3) Slimness is attractive. Carol is slim. Is Carol attractive?

(4) Brian is much taller than most of his close friends. How tall is Brian?

There are two main reasons why classical logical systems cannot cope with problems of this type. First, they do not provide a system for representing the meaning of propositions expressed in a natural language when the meaning is imprecise; and second, in those cases in which the meaning can be represented symbolically in a meaning representation language, for example, a semantic network or a conceptual-dependency graph, there is no mechanism for inference.

As will be seen, fuzzy logic addresses these problems in the following ways. First, the meaning of a lexically imprecise proposition is represented as an elastic constraint on a variable; and second, the answer to a query is deduced through a propagation of elastic constraints.

During the past several years, fuzzy logic has found numerous applications in fields ranging from finance to earthquake engineering. But what is striking is that its most important and visible application today is in a realm not anticipated when fuzzy logic was conceived, namely, the realm of fuzzy-logic-based process control. The basic idea underlying fuzzy logic control was suggested in notes published in 1968 and 1972[1,2] and described in greater detail in 1973.[3] The first implementation was pioneered by Mamdani and Assilian in 1974[4] in connection with the regulation of a steam engine. In the ensuing years, once the basic idea underlying fuzzy logic control became well understood, many applications followed. In Japan, in particular, the use of fuzzy logic in control processes is being pursued in many application areas, among them automatic train operation (Hitachi),[5] vehicle control (Sugeno Laboratory at Tokyo Institute of Technology),[5] robot control (Hirota Laboratory at Hosei University),[5] speech recognition (Ricoh),[5] universal controller (Fuji),[5] and stabilization control (Yamakawa Laboratory at Kumamoto University).[5] More about

0018-9162/88/0400-0083$01.00 © 1988 IEEE

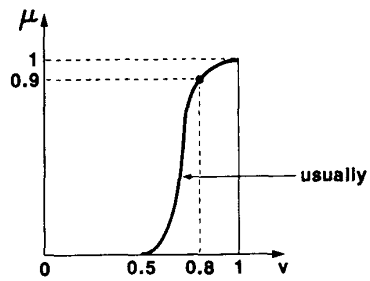

Figure 1. Representation of "usually" as a fuzzy proportion.

some of these projects will be said in the section dealing with applications.

In most of the current applications of fuzzy logic, software is employed as a medium for the implementation of fuzzy algorithms and control rules. What is clear, however, is that it would be cheaper and more effective to use fuzzy logic chips and, eventually, fuzzy computers. The first logic chip was developed by Togai and Watanabe at Bell Telephone Laboratories in 1985, and it is likely to become available for commercial use in 1988 or 1989. On the heels of this important development came the announcement of a fuzzy computer designed by Yamakawa at Kumamoto University. These developments on the hardware front may lead to an expanded use of fuzzy logic not only in industrial applications but, more generally, in knowledge-based systems in which the deduction of an answer to a query requires the inference machinery of fuzzy logic.

One important branch of fuzzy logic may be called *dispositional logic*. This logic, as its name implies, deals with *dispositions*, that is, propositions that are preponderantly but not necessarily always true. For example, "snow is white" is a disposition, as are the propositions "Swedes are blond" and "high quality is expensive." A disposition may be viewed as a usuality-qualified proposition in which the qualifying quantifier "usually" is implicit rather than explicit. In this sense, the disposition "snow is white" may be viewed as the result of suppressing the fuzzy quantifier "usually" in the usuality-qualified proposition

usually (snow is white)

In this proposition, "usually" plays the role of a fuzzy proportion of the form shown in Figure 1.

The importance of dispositional logic stems from the fact that most of what is usually referred to as common sense knowledge may be viewed as a collection of dispositions. Thus, the main concern of dispositional logic lies in the development of rules of inference from common sense knowledge.

In what follows, I present a condensed exposition of some basic ideas underlying fuzzy logic and describe some representative applications. More detailed information regarding fuzzy logic and its applications may be found in the cited literature.

Basic principles

Fuzzy logic may be viewed as an extension of multivalued logic. Its uses and objectives, however, are quite different. Thus, the fact that fuzzy logic deals with approximate rather than precise modes of reasoning implies that, in general, the chains of reasoning in fuzzy logic are short in length, and rigor does not play as important a role as it does in classical logical systems. In a nutshell, in fuzzy logic everything, including truth, is a matter of degree.

The greater expressive power of fuzzy logic derives from the fact that it contains as special cases not only the classical two-valued and multivalued logical systems but also probability theory and probabilistic logic. The main features of fuzzy logic that differentiate it from traditional logical systems are the following:

(1) In two-valued logical systems, a proposition p is either true or false. In multivalued logical systems, a proposition may be true or false or have an intermediate truth value, which may be an element of a finite or infinite truth value set T. In fuzzy logic, the truth values are allowed to range over the fuzzy subsets of T. For example, if T is the unit interval, then a truth value in fuzzy logic, for example, "very true," may be interpreted as a fuzzy subset of the unit interval. In this sense, a fuzzy truth value may be viewed as an imprecise characterization of a numerical truth value.

(2) The predicates in two-valued logic are constrained to be crisp in the sense that the denotation of a predicate must be a nonfuzzy subset of the universe of discourse. In fuzzy logic, the predicates may be crisp—for example, "mortal," "even," and "father of"—or, more generally, fuzzy—for example, "ill," "tired," "large," "tall," "much heavier," and "friend of."

(3) Two-valued as well as multivalued logics allow only two quantifiers: "all" and "some." By contrast, fuzzy logic allows, in addition, the use of fuzzy quantifiers exemplified by "most," "many," "several," "few," "much of," "frequently," "occasionally," "about ten," and so on. Such quantifiers may be interpreted as fuzzy numbers that provide an imprecise characterization of the cardinality of one or more fuzzy or nonfuzzy sets. In this perspective, a fuzzy quantifier may be viewed as a second-order fuzzy predicate. Based on this view, fuzzy quantifiers may be used to represent the meaning of propositions containing fuzzy probabilities and thereby make it possible to manipulate probabilities within fuzzy logic.

(4) Fuzzy logic provides a method for representing the meaning of both nonfuzzy and fuzzy predicate-modifiers exemplified by "not," "very," "more or less," "extremely," "slightly," "much," "a little," and so on. This, in turn, leads to a system for computing with *linguistic variables*,[3] that is, variables whose values are words or sentences in a natural or synthetic language. For example, "Age" is a linguistic variable when its values are assumed to be "young," "old," "very young," "not very old," and so forth. More about linguistic variables will be said at a later point.

(5) In two-valued logical systems, a proposition p may be qualified, principally by associating with p a truth value, "true" or "false"; a modal operator such as "possible" or "necessary"; and an intensional operator such as "know" or "believe." Fuzzy logic has three principal modes of qualification:

- *truth-qualification*, as in

  (Mary is young) is not quite true,

  in which the qualified proposition is (Mary is young) and the qualifying truth value is "not quite true";
- *probability-qualification*, as in

  (Mary is young) is unlikely,

  in which the qualifying fuzzy probability is "unlikely"; and
- *possibility-qualification*, as in

  (Mary is young) is almost impossible,

  in which the qualifying fuzzy possibility is "almost impossible."

An important issue in fuzzy logic relates to inference from qualified propositions, especially from probability-qualified propositions. This issue is of central importance in the management of uncertainty in expert systems and in the formalization of common sense reasoning. In the latter, it's important to note the close connection between probability-qualification and usuality-qualification and the role played by fuzzy quantifiers. For example, the disposition

Swedes are blond

may be interpreted as

most Swedes are blond;

or, equivalently, as

(Swede is blond) is likely,

where "likely" is a fuzzy probability that is numerically equal to the fuzzy quantifier "most"; or, equivalently, as

usually (a Swede is blond),

where "usually" qualifies the proposition "a Swede is blond."

As alluded earlier, inference from propositions of this type is a main concern of dispositional logic. More about this logic will be said at a later point.

Meaning representation and inference

A basic idea serving as a point of departure in fuzzy logic is that a proposition p in a natural or synthetic language may be viewed as a collection of elastic constraints, $C_1, \ldots, C_k$, which restrict the values of a collection of variables $X = (X_1, \ldots, X_n)$.[6] In general, the constraints as well as the variables they constrain are implicit rather than explicit in p. Viewed in this perspective, representation of the meaning of p is, in essence, a process by which the implicit constraints and variables in p are made explicit. In fuzzy logic, this is accomplished by representing p in the so-called *canonical form*

$$p \rightarrow X \text{ is } A$$

in which A is a fuzzy predicate or, equivalently, an n-ary fuzzy relation in U, where $U = U_1 \times U_2 \times \ldots \times U_n$, and U_i, $i = 1, \ldots, n$, is the domain of X_i. Representation of p in its canonical form requires, in general, the construction of an explanatory database and a test procedure that tests and aggregates the test scores associated with the elastic constraints $C_1, \ldots, C_k$.[6]

In more concrete terms, the canonical form of p implies that the possibility distribution[6] of X is equal to A—that is,

$$\Pi_X = A \tag{1}$$

which in turn implies that

$$\text{Poss}\{X=u\} = \mu_A(u),\ u \in U$$

where μ_A is the membership function of A and Poss$\{X=u\}$ is the possibility that X may take u as its value. Thus, when the meaning of p is represented in the form of Equation 1, it signifies that p induces a possibility distribution Π_X that is equal to A, with A playing the role of an elastic constraint on a variable X that is implicit in p. In effect, the possibility distribution of X, Π_X, is the set of possible values of X, with the understanding that possibility is a matter of degree. Viewed in this perspective, a proposition p constrains the possible values that X can take and thus defines its possibility distribution. This implies that the meaning of p is defined by (1) identifying the variable that is constrained and (2) characterizing the constraint to which the variable is subjected through its possibility distribution. Note that Equation 1 asserts that the possibility that X can take u as its value is numerically equal to the grade of membership, $\mu_A(u)$, of u in A.

As an illustration, consider the proposition

$$p \triangleq \text{John is tall}$$

in which the symbol $\triangleq$ should be read as "denotes" or "is equal to by definition." In this case, X = Height(John), A = TALL, and the canonical form of p reads

Height(John) is TALL

where the fuzzy relation TALL is in uppercase letters to underscore that it plays the role of a constraint in the canonical form. From the canonical form, it follows that

$$\text{Poss}\{\text{Height (John)} = u\} = \mu_{\text{TALL}}(u)$$

where μ_{TALL} is the membership function of TALL and $\mu_{\text{TALL}}(u)$ is the grade of membership of u in TALL or, equivalently, the degree to which a numerical height u satisfies the constraint induced by the relation TALL.

When p is a conditional proposition, its canonical form may be expressed as "Y is B if X is A," implying that p induces a conditional possibility distribution of Y given X, written as $\Pi_{(Y|X)}$. In fuzzy logic, $\Pi_{(Y|X)}$ may be defined in a variety of ways,[7] among which is a definition consistent with the definition of implication in Lakasiewicz's L_{Aleph_0} logic. In this case, the conditional possibility distribution function, $\pi_{(Y|X)}$, which defines $\Pi_{(Y|X)}$, may be expressed as

$$\pi_{(Y|X)}(u, v) = 1 \wedge (1 - \mu_A(u) + \mu_B(v)),\quad u \in U,\ v \in V, \tag{2}$$

where

$$\pi_{(Y|X)}(u, v) \triangleq \text{Poss}\{X=u,\ Y=v\}$$

μ_A and μ_B denote the membership functions of A and B, respectively; and $\wedge$ denotes the operator min.

When p is a quantified proposition of the form

$$p \triangleq Q\ A\text{'s are } B\text{'s}$$

for example,

$$p \triangleq \text{most tall men are not very fat}$$

where Q is a fuzzy quantifier and A and B are fuzzy predicates, the constrained variable, X, is the proportion of B's in A's, with Q representing an elastic constraint on X. More specifically, if U is a finite set $\{u_1, \ldots, u_m\}$, the proportion of B's in A's is defined as the *relative sigma-count*

$$\Sigma\text{Count}(B/A) = \frac{\sum_j \mu_A(u_j) \wedge \mu_B(u_j)}{\sum_j \mu_A(u_j)} \quad j = 1, \ldots, m \tag{3}$$

where $\mu_A(u_j)$ and $\mu_B(u_j)$ denote the grades of membership of u_j in A and B, respectively. Thus, expressed in its canonical form, Equation 3 may be written as

$$\Sigma Count(B/A) \text{ is } Q$$

which places in evidence the constrained variable, X, in p and the elastic constraint, Q, to which X is subjected. Note that X is the relative sigma-count of B in A.

The concept of a canonical form pro-

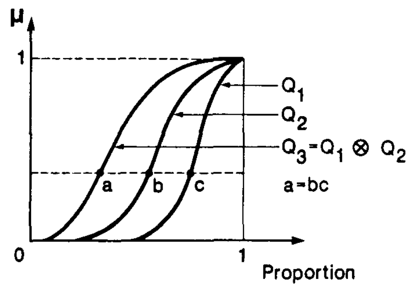

Figure 2. Representation of fuzzy quantifiers in the intersection/product syllogism.

vides an effective framework for formulating the problem of inference in expert systems. Specifically, consider a knowledge base, *KB*, which consists of a collection of propositions $\{p_1, \ldots, p_N\}$. Typically, a constituent proposition, p_i, $i = 1, \ldots, N$, may be (1) a fact that may be expressed in a canonical form as "X is A" or (2) a rule that may be expressed in a canonical form as "Y is B_i" if "X is A_i." More generally, both facts and rules may be probability-qualified or, equivalently, expressed as quantified propositions. For example, a rule of the general form "Q A's are B's" may be interpreted as the probability-qualified proposition (X is B if X is A) is λ, where λ is a fuzzy probability whose denotation as a fuzzy subset of the unit interval is the same as that of the fuzzy quantifier Q and X is chosen at random in U.

Now if p_i induces a possibility distribution $\Pi^i_{(X_1, \ldots, X_n)}$, where $X_1, \ldots, X_n$ are the variables constrained by p_i, then the possibility distribution $\Pi_{(X_1, \ldots, X_n)}$, which is induced by the totality of propositions in *KB* is given by the intersection[3] of the $\Pi^i_{(X_1, \ldots, X_n)}$. That is,

$$\Pi_{(X_1, \ldots, X_n)} = \Pi^1_{(X_1, \ldots, X_n)} \cap \ldots \cap \Pi^N_{(X_1, \ldots, X_n)}$$

or, equivalently,

$$\pi_{(X_1, \ldots, X_n)} = \pi^1_{(X_1, \ldots, X_n)} \wedge \ldots \wedge \pi^N_{(X_1, \ldots, X_n)}$$

$\pi_{(X_1, \ldots, X_n)}$ is the possibility distribution function of $\Pi_{(X_1, \ldots, X_n)}$. Note that there is no loss of generality in assuming that the constrained variables $X_1, \ldots, X_n$ are the same for all propositions in *KB* since the set $\{X_1, \ldots, X_n\}$ may be taken to be the union of the constrained variables for each proposition.

Now suppose that we are interested in inferring the value of a specified function $f(X_1, \ldots, X_n)$, $f: U \rightarrow V$, of the variables constrained by the knowledge base. Because of the incompleteness and imprecision of the information resident in *KB*, what we can deduce, in general, is not the value of $f(X_1, \ldots, X_n)$ but its possibility distribution, Π_f. By employing the extension principle,[8] it can be shown that the possibility distribution function of f is given by the solution of the nonlinear program

$$\pi_f(v) = \max_{u_1, \ldots, u_n} [\pi^1_{(X_1, \ldots, X_n)}(u_1, \ldots, u_n) \wedge \ldots \wedge \pi^N_{(X_1, \ldots, X_n)}(u_1, \ldots, u_n)] \tag{4}$$

subject to the constraint

$$v = f(u_1, \ldots, u_n)$$

where $u_j \in U_j$, $i = 1, \ldots, n$, and $v \in V$. The reduction to the solution of a nonlinear program constitutes the principal tool for inference in fuzzy logic.

Fuzzy syllogisms. A basic fuzzy syllogism in fuzzy logic that is of considerable relevance to the rules of combination of evidence in expert systems is the *intersection/product syllogism*—a syllogism that serves as a rule of inference for quantified propositions.[9] This syllogism may be expressed as the inference rule

$$\begin{array}{l} Q_1\ A\text{'s are } B\text{'s} \\ \underline{Q_2\ (A \text{ and } B)\text{'s are } C\text{'s}} \\ (Q_1 \otimes Q_2)\ A\text{'s are } (B \text{ and } C)\text{'s} \end{array} \tag{5}$$

in which Q_1 and Q_2 are fuzzy quantifiers, A, B, and C are fuzzy predicates, and $Q_1 \otimes Q_2$ is the product of the fuzzy numbers Q_1 and Q_2 in fuzzy arithmetic.[10] (See Figure 2). For example, as a special case of Equation 5, we may write

most students are single
a little more than a half of single students are male

(most ⊗ a little more than a half) of students are single and male

Since the intersection of B and C is contained in C, the following corollary of Equation 5 is its immediate consequence.

$$\begin{array}{l} Q_1\ A\text{'s are } B\text{'s} \\ \underline{Q_2\ (A \text{ and } B)\text{'s are } C\text{'s}} \\ \geq (Q_1 \otimes Q_2)\ A\text{'s are } C\text{'s} \end{array} \tag{6}$$

where the fuzzy number $\geq (Q_1 \otimes Q_2)$ should be read as "at least $(Q_1 \otimes Q_2)$." In particular, if the fuzzy quantifiers Q_1 and Q_2 are monotone increasing (for example, when "$Q_1 = Q_2 \triangleq$ most"), then

$$\geq (Q_1 \otimes Q_2) = Q_1 \otimes Q_2$$

and Equation 6 becomes

$$\begin{array}{l} Q_1\ A\text{'s are } B\text{'s} \\ \underline{Q_2\ (A \text{ and } B)\text{'s are } C\text{'s}} \\ (Q_1 \otimes Q_2)\ A\text{'s are } C\text{'s} \end{array} \tag{7}$$

Furthermore, if B is a subset of A, then A and $B = B$, and Equation 7 reduces to the *chaining rule*

$$\begin{array}{l} Q_1\ A\text{'s are } B\text{'s} \\ \underline{Q_2\ B\text{'s are } C\text{'s}} \\ (Q_1 \otimes Q_2)\ A\text{'s are } C\text{'s} \end{array} \tag{8}$$

For example,

most students are undergraduates
most undergraduates are young
---
most^2 students are young

where "most^2" represents the product of the fuzzy number "most" with itself (see Figure 3).

What is important to observe is that the chaining rule expressed by Equation 8 serves the same purpose as the chaining rules in Mycin, Prospector, and other probability-based expert systems. However, Equation 8 is formulated in terms of fuzzy quantifiers rather than numerical probabilities or certainty factors, and it is a logical consequence of the concept of a relative sigma-count in fuzzy logic. Furthermore, the chaining rule (Equation 8) is robust in the sense that if Q_1 and Q_2 are close to unity, so is their product $Q_1 \otimes Q_2$. More specifically, if Q_1 and Q_2 are expressed as

$$Q_1 = 1 \ominus \varepsilon_1$$
$$Q_2 = 1 \ominus \varepsilon_2$$

where ε_1 and ε_2 are small fuzzy numbers, then, to a first approximation, Q may be expressed as

$$Q = 1 \ominus \varepsilon_1 \ominus \varepsilon_2$$

An important issue concerns the general properties Q_1, Q_2, A, B, and C must have to ensure robustness. As shown above, the containment of B in A and the monotonicity of Q_1 and Q_2 are conditions for robustness in the case of the intersection/product syllogism.

Another basic syllogism is the *consequent conjunction syllogism*

$$\frac{\begin{array}{l} Q_1\ A\text{'s are } B\text{'s} \\ Q_2\ A\text{'s are } C\text{'s} \end{array}}{Q\ A\text{'s are } (B \text{ and } C)\text{'s}} \qquad (9)$$

where

$$0 \circledcirc (Q_1 \oplus Q_2 \ominus 1) \leq Q \leq Q_1 \circledcirc Q_2$$

in which the operators $\circledcirc$, $\circledcirc$, $\oplus$, $\ominus$, and the inequality $\leq$ are the extensions of $\wedge$, $\vee$, $+$, $-$, and $\leq$, respectively, to fuzzy numbers.

The consequent conjunction syllogism plays the same role in fuzzy logic as the rule of combination of evidence for conjunctive hypotheses does in Mycin and Prospector.[11] However, whereas in Mycin and Prospector the qualifying probabilities and certainty factors are real numbers, in the consequent conjunction syllogism the fuzzy quantifiers are fuzzy numbers. As can be seen from the result expressed by Equation 9, the conclusion yielded by the

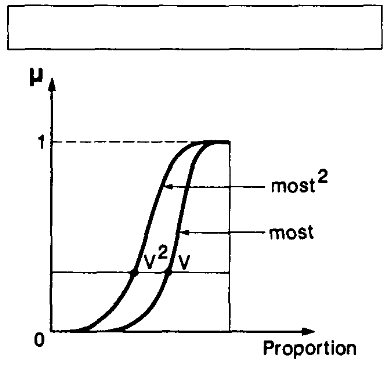

Figure 3. Representation of "most" and "most²."

Inference with fuzzy probabilities

An example of an important problem to which the reduction to a nonlinear program may be applied is the following. Assume that from a knowledge base $KB = \{p_1, \ldots, p_N\}$ in which the constituent propositions are true with probability one, we can infer a proposition q which, like the premises, is true with probability one. Now suppose that each p_i in KB is replaced with a probability-qualified proposition "$p_i^* \triangleq p_i$ is λ_i," in which λ_i is a fuzzy probability. For example

$p_i \triangleq X$ is small

and

$p_i^* \triangleq X$ is small is very likely

As a result of the qualification of the p_i, the conclusion, q, will also be a probability-qualified proposition that may be expressed as

$q^* = q$ is λ

in which λ is a fuzzy probability. The problem is to determine λ as a function of the λ_i, if such a function exists. A special case of this problem, which is of particular relevance to the management of uncertainty in expert systems, is one in which the fuzzy probabilities λ_i are close to unity. We shall say that the inference process is *compositional* if λ can be expressed as a function of the λ_i; it is *robust* if whenever the λ_i are close to unity, so is λ.

By reducing the determination of λ to the solution of a nonlinear program, it can be shown that, in general, the inference process is not compositional if the λ_i and λ are numerical probabilities. This result calls into question the validity of the rules of combination of evidence in those expert systems in which the certainty factor of the conclusion is expressed as a function of the certainty factors of the premises. However, compositionality does hold, in general, if the λ_i and λ are assumed to be fuzzy probabilities, for this allows the probability of q to be interval-valued when the λ_i are numerical probabilities, which is consistent with known results in inductive logic.

Another important conclusion relating to the robustness of the inference process is that, in general, robustness does not hold without some restrictive assumptions on the premises. For example, the brittleness of the transitivity of implication is an instance of the lack of robustness when no assumptions are made regarding the fuzzy predicates A, B, and C. On the other hand, if in the inference schema

$$\frac{\begin{array}{l} X \text{ is } A \\ Y \text{ is } B \text{ if } X \text{ is } A \end{array}}{Y \text{ is } B}$$

the major premise is replaced by "X is A is probable," where "probable" is a fuzzy probability close to unity, then it can be shown that, under mildly restrictive assumptions on A, the resulting conclusion may be expressed as "Y is B is $\geq$ probable," where "$\geq$ probable" is a fuzzy probability that, as a fuzzy number, is greater than or equal to the fuzzy number "probable." In this case, then, robustness does hold, for if "probable" is close to unity, so is "$\geq$ probable."

application of fuzzy logic to the premises in question is both robust and compositional.

A more complex problem is presented by what in Mycin and Prospector corresponds to the conjunctive combination of evidence. Stated in terms of quantified premises, the inference rule in question may be expressed as

$$\begin{array}{l} Q_1\ A\text{'s are } C\text{'s} \\ \underline{Q_2\ B\text{'s are } C\text{'s}} \\ Q\ (A \text{ and } B)\text{'s are } C\text{'s} \end{array} \qquad (10)$$

where the value of Q is to be determined. To place in evidence the symmetry between Equation 9 and Equation 10, we shall refer to the rule in question as the *antecedent conjunction syllogism*.

It can readily be shown that, without any restrictive assumptions on Q_1, Q_2, A, B, and C, there is nothing that can be said about Q, which is equivalent to saying that "Q = none to all." A basic assumption in Mycin, Prospector, and related systems is that the items of evidence are conditionally independent, given the hypothesis (and its complement). That is,

$$P(E_1, E_2 \mid H) = P(E_1 \mid H)\, P(E_2 \mid H)$$

where $P(E_1, E_2 \mid H)$ is the joint probability of E_1 and E_2, given the hypothesis H; and $P(E_1 \mid H)$ and $P(E_2 \mid H)$ are the conditional probabilities of E_1 given H and E_2 given H, respectively. Expressed in terms of the relative sigma-counts, this assumption may be written as

$$\Sigma\text{Count}\,(A \cap B \,/\, C) = \Sigma\text{Count}\,(A \,/\, C)\, \Sigma\text{Count}\,(B \,/\, C) \qquad (11)$$

where $\cap$ denotes the intersection of fuzzy sets.[3]

To determine the value of Q in Equation 10 we have to compute the relative sigma-count of C in $A \cap B$. It can be verified that, under the assumption (Equation 11), the sigma-count in question is given by

$$\Sigma\text{Count}\,(C \,/\, A \cap B) = \Sigma\text{Count}\,(C \,/\, A)\, \Sigma\text{Count}\,(C \,/\, B)\, \delta$$

where the factor δ is expressed by

$$\delta = \frac{\Sigma\text{Count}\,(A)\, \Sigma\text{Count}\,(B)}{\Sigma\text{Count}\,(A \cap B)\, \Sigma\text{Count}\,(C)} \qquad (12)$$

Inspection of Equation 12 shows that the assumption expressed by Eqaution 11 does not ensure the compositionality of Q. However, it can be shown that compositionality can be achieved through the use of the concept of a relative *ϱsigma-count*, which is defined as

$$\varrho\Sigma\text{Count}\,(B/A) = \frac{\Sigma\text{Count}(B/A)}{\Sigma\text{Count}\,(\neg B/A)}$$

where $\neg B$ denotes the negation of B. The use of ϱsigma-counts in place of sigma-counts is analogous to the use of odds instead of probabilities in Prospector, and it serves the same purpose.

Interpolation

An important problem that arises in the operation of any rule-based system is the following. Suppose the user supplies a fact that, in its canonical form, may be expressed as "X is A," where A is a fuzzy or nonfuzzy predicate. Furthermore, suppose that there is no conditional rule in *KB* whose antecedent matches A exactly. The question arises: Which rules should be executed and how should their results be combined?

An approach to this problem, sketched in Reference 8, involves the use of an interpolation technique in fuzzy logic which requires a computation of the degree of partial match between the user-supplied fact and the rows of a decision table. More specifically, suppose that upon translation into their canonical forms, a group of propositions in *KB* may be expressed as a fuzzy relation of the form

R	X_1	X_2	.	X_n	X_{n+1}
	R_{11}	R_{12}	.	R_{1n}	Z_1
	.	.	.	.	.
	R_{m1}	R_{m2}	.	R_{mn}	Z_m

in which the entries are fuzzy sets; the input variables are $X_1, \ldots, X_n$, with domains $U_1, \ldots, U_n$; and the output variable is X_{n+1}, with domain U_{n+1}. The problem is: Given an input n-tuple $(R_1, \ldots, R_n)$, in which R_j, $j = 1, \ldots, n$, is a fuzzy subset of U_j, what is the value of X_{n+1} expressed as a fuzzy subset of U_{n+1}?

A possible approach to the problem is to compute for each pair (R_{ij}, R_j) the degree of consistency of the input R_j with the R_{ij} element of R, $i = 1, \ldots, m$, $j = 1, \ldots, n$. The degree of consistency, γ_{ij}, is defined as

$$\begin{aligned} \gamma_{ij} &\triangleq \sup\,(R_{ij} \cap R_j) \\ &= \sup_{u_j}\,(\mu_{R_{ij}}(u_j) \wedge \mu_{R_j}(u_j)) \end{aligned}$$

in which $\mu_{R_{ij}}$ and μ_{R_j} are the membership functions of R_{ij} and R_j, respectively; u_j is a generic element of U_j; and the supremum is taken over u_j.

Next, we compute the overall degree of consistency, γ_i, of the input n-tuple $(R_1, \ldots, R_n)$ with the ith row of R, $i = 1, \ldots, m$, by employing $\wedge$ (min) as the aggregation operator. Thus,

$$\gamma_i = \gamma_{i1} \wedge \gamma_{i2} \wedge \ldots \wedge \gamma_{in}$$

which implies that γ_i may be interpreted as a conservative measure of agreement between the input n-tuple $(R_1, \ldots, R_n)$ and the ith-row n-tuple $(R_{i1}, \ldots, R_{in})$. Then, employing γ_i as a weighting coefficient, the desired expression for X_{n+1} may be written as a "linear" combination

$$X_{n+1} = \gamma_1 \wedge Z_1 + \ldots + \gamma_m \wedge Z_m$$

in which $+$ denotes the union, and $\gamma_i \wedge z_i$ is a fuzzy set defined by

$$\mu_{\gamma_i \wedge z_i}(u_{i+1}) = \gamma_i \wedge \mu_{z_i}(u_{i+1}),\ i = 1, \ldots, m$$

The above approach ceases to be effective, however, when R is a sparse relation in the sense that no row of R has a high degree of consistency with the input n-tuple. For such cases, a more general interpolation technique has to be employed.

Basic rules of inference

One distinguishing characteristic of fuzzy logic is that premises and conclusions in an inference rule are generally expressed in canonical form. This representation places in evidence the fact that each premise is a constraint on a variable and that the conclusion is an induced constraint computed through a process of constraint propagation — a process that, in general, reduces to the solution of a nonlinear program. The following briefly presents — without derivation — some of the basic inference rules in fuzzy logic. Most of these rules can be deduced from the basic inference rule expressed by Equation 4.

The rules of inference in fuzzy logic may be classified in a variety of ways. One basic class is *categorical rules*, that is, rules that do not contain fuzzy quantifiers. A more general class is *dispositional rules*, rules in which one or more premises may contain, explicitly or implicitly, the fuzzy quantifier "usually." For example, the inference rule known as the *entailment principle*:

$$\begin{array}{l} X \text{ is } A \\ \underline{A \subset B} \\ X \text{ is } B \end{array} \qquad (13)$$

where X is a variable taking values in a universe of discourse U, and A and B are fuzzy subsets of U, is a categorical rule. On the other hand, the *dispositional entailment principle* is an inference rule of the form

$$\begin{array}{l} \text{usually } (X \text{ is } A) \\ \underline{A \subset B \qquad\qquad} \\ \text{usually } (X \text{ is } B) \end{array} \qquad (14)$$

In the limiting case where "usually" becomes "always," Equation 14 reduces to Equation 13.

In essence, the *entailment principle* asserts that from the proposition "X is A" we can always infer a less specific proposition "X is B." For example, from the proposition "Mary is young," which in its canonical form reads

Age(Mary) is YOUNG

where YOUNG is interpreted as a fuzzy set or, equivalently, as a fuzzy predicate, we can infer "Mary is not old," provided YOUNG is a subset of the complement of OLD. That is

$$\mu_{\text{YOUNG}}(u) \subset 1 - \mu_{\text{OLD}}(u), \; u \in [0, 100]$$

where μ_{YOUNG} and μ_{OLD} are, respectively, the membership functions of YOUNG and OLD, and the universe of discourse is the interval [0, 100].

Viewed in a different perspective, the entailment principle in fuzzy logic may be regarded as a generalization to fuzzy sets of the inheritance principle widely used in knowledge representation systems. More specifically, if the proposition "X is A" is interpreted as "X has property A," then the conclusion "X is B" may be interpreted as "X has property B," where B is any superset of A. In other words, X inherits property B if B is a superset of A.

Among other categorical rules that play a basic role in fuzzy logic are the following. In all of these rules, $X, Y, Z, \ldots$ are variables ranging over specified universes of discourse, and $A, B, C, \ldots$ are fuzzy predicates or, equivalently, fuzzy relations.

Conjunctive rule.

$$\begin{array}{l} X \text{ is } A \\ \underline{X \text{ is } B \qquad} \\ X \text{ is } A \cap B \end{array}$$

where $A \cap B$ is the intersection of A and B defined by

$$\mu_{A \cap B}(u) = \mu_A(u) \wedge \mu_B(u), \quad u \in U$$

Cartesian product.

$$\begin{array}{l} X \text{ is } A \\ \underline{Y \text{ is } B \qquad\quad} \\ (X, Y) \text{ is } A \times B \end{array}$$

where (X, Y) is a binary variable and $A \times B$ is defined by

$$\mu_{A \times B}(u, v) = \mu_A(u) \wedge \mu_B(v), \quad u \in U, v \in V$$

Projection rule.

$$\begin{array}{l} \underline{(X, Y) \text{ is } R} \\ X \text{ is } {}_X R \end{array}$$

where ${}_X R$, the projection of the binary relation R on the domain of X, is defined by

$$\mu_{{}_X R}(u) = \sup_v \mu_R(u,v), \quad u \in U, v \in V$$

where $\mu_R(u,v)$ is the membership function of R and the supremum is taken over $v \in V$.

Compositional rule.

$$\begin{array}{l} X \text{ is } A \\ \underline{(X, Y) \text{ is } R} \\ Y \text{ is } A \circ R \end{array}$$

where $A \circ R$, the composition of the unary relation A with the binary relation R, is defined by

$$\mu_{A \circ R}(v) = \sup_u (\mu_A(u) \wedge \mu_R(u,v))$$

The compositional rule of inference may be viewed as a combination of the conjunctive and projection rules.

Generalized modus ponens.

$$\begin{array}{l} X \text{ is } A \\ \underline{Y \text{ is } C \text{ if } X \text{ is } B} \\ Y \text{ is } A \circ (\neg B \oplus C) \end{array}$$

where $\neg B$ denotes the negation of B and the bounded sum is defined by

$$\mu_{\neg B \oplus C}(u,v) = 1 \wedge (1 - \mu_B(u) + \mu_C(v))$$

An important feature of the generalized modus ponens, which is not possessed by the modus ponens in binary logical systems, is that the antecedent "X is B" need not be identical with the premise "X is A." It should be noted that the generalized modus ponens is related to the interpolation rule which was described earlier. An additional point that should be noted is that the generalized modus ponens may be regarded as an instance of the compositional rule of inference.

Dispositional modus ponens. In many applications involving common sense reasoning, the premises in the generalized modus ponens are usuality-qualified. In such cases, one may employ a dispositional version of the modus ponens. It may be expressed as

$$\begin{array}{l} \text{usually } (X \text{ is } A) \\ \underline{\text{usually } (Y \text{ is } B \text{ if } X \text{ is } A)} \\ \text{usually}^2 \, (Y \text{ is } B) \end{array}$$

where "usually^2" is the square of "usually" (see Figure 4). For simplicity, it's assumed that the premise "X is A" matches the antecedent in the conditional proposition; also, the conditional proposition is interpreted as the statement, "The

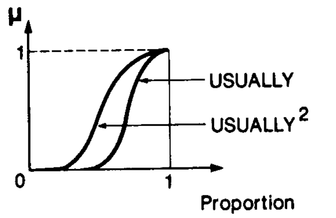

Figure 4. Representation of "usually" and "usually^2."

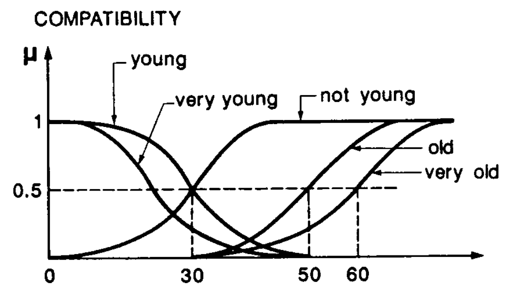

Figure 5. The linguistic values of "Age."

value of the fuzzy conditional probability of B given A is the fuzzy number USUALLY."

Extension principle. The extension principle plays an important role in fuzzy logic by providing a mechanism for computing induced constraints. More specifically, assume that a variable X taking values in a universe of discourse U is constrained by the proposition "X is A." Furthermore, assume that f is a mapping from U to V so that X is mapped into $f(X)$. The question is: What is the constraint on $f(X)$ which is induced by the constraint on X?

The answer provided by the extension principle may be expressed as the inference rule

$$\frac{X \text{ is } A}{f(X) \text{ is } f(A)}$$

where the membership function of $f(A)$ is defined by

$$\mu_{f(A)}(v) = \sup_u \mu_A(u) \tag{15}$$

subject to the condition

$$v = f(u),\ u \in U,\ v \in V$$

In particular, if the function f is 1:1, then Equation 15 simplifies to

$$\mu_{f(A)}(v) = \mu_A(v^{-1}),\ v \in V$$

where v^{-1} is the inverse of v. For example,

$$\frac{X \text{ is small}}{X^2 \text{ is small}^2}$$

and

$$\mu_{\text{SMALL}^2}(v) = \mu_{\text{SMALL}}(\sqrt{v})$$

As in the case of the entailment rule, the dispositional version of the extension principle has the simple form

$$\frac{\text{usually } (X \text{ is } A)}{\text{usually } (f(X) \text{ is } f(A))}$$

The dispositional extension principle plays an important role in inference from common sense knowledge. In particular, it is one of the inference rules that play an essential role in answering the questions posed in the introduction.

The linguistic variable and its application to fuzzy control

A basic concept in fuzzy logic that plays a key role in many of its applications, especially in the realm of fuzzy control and fuzzy expert systems, is a *linguistic variable*.

A linguistic variable, as its name suggests, is a variable whose values are words or sentences in a natural or synthetic language. For example, "Age" is a linguistic variable if its values are "young," "not young," "very young," "old," "not old," "very old," and so on.

In general, the values of a linguistic variable can be generated from a *primary term* (for example, "young") its antonym ("old"), a collection of modifiers ("not," "very," "more or less," "quite," "not very," etc.), and the connectives "and" and "or." For example, one value of "Age" may be "not very young and not very old." Such values can be generated by a context-free grammar. Furthermore, each value of a linguistic variable represents a possibility distribution, as shown in Figure 5 for the variable "Age." These possibility distributions may be computed from the given possibility distributions of the primary term and its antonym through the use of attributed grammar techniques.

An interesting application of the linguistic variable is embodied in the fuzzy car conceived and designed by Sugeno of the Tokyo Institute of Technology.[5] The car's fuzzy-logic-based control system lets it move autonomously along a track with rectangular turns and park in a designated space (see Figure 6). An important feature is the car's ability to learn from examples.

The basic idea behind the Sugeno fuzzy car is the following. The controlled variable Y, which is the steering angle, is assumed to be a function of the state variables $X_1, X_2, X_3, \ldots, X_n$, which represent the distances of the car from the boundaries of the track at a corner (see Figure 7). These values are treated as linguistic variables, with the primary terms represented as triangular possibility distributions (see Figure 8).

The control policy is represented as a finite collection of rules of the form

R^i: if (X_1 is $A^i{}_1$) and . . . (X_n is A^i_n), then
$$Y^i = a^i_0 + a^i{}_1 X_1 + \ \ldots \ + a^i_n X_n$$

where R^i is the ith rule; A^i_j is a linguistic value of X_j in R^i; Y^i is the value of the control variable suggested by R^i; and a^i_0, . . . , a^i_n are adjustable parameters, which define Y^i as a linear combination of the state variables.

In a given state (X_1, . . . , X_n), the truth value of the antecedent of R^i may be expressed as

$$W^i = A^i_1(X_1) \wedge \ldots \wedge A^i_n(X_n)$$

where $A^i_j(X_j)$ is the grade of membership of X_j in A^i_j. The aggregated value of the controlled variable Y is computed as the normalized linear combination

$$Y = \frac{W_1 Y^1 + \ldots + W_n Y^n}{W_1 + \ldots + W_n} \qquad (16)$$

Thus, Equation 16 may be interpreted as the result of a weighted vote in which the value suggested by R^i is given the weight $W_i/(W_1 + \ldots + W_n)$.

The values of the coefficients a^i_1, . . . , $a^i{}_n$ are determined through training. Training consists of an operator guiding a model car along the track a few times until an identification algorithm converges on parameter values consistent with the control rules. By its nature, the training process cannot guarantee that the identification algorithm will always converge on the correct values of the coefficients. The justification is pragmatic: the system works in most cases.

Variations on this idea are embodied in most of the fuzzy-logic-based control systems developed so far. Many of these systems have proven to be highly reliable and superior in performance to conventional systems.[5]

Since most rules in expert systems have fuzzy antecedents and consequents, expert systems provide potentially important applications for fuzzy logic. For example[11]:

IF the search "space" is moderately small
THEN exhaustive search is feasible

IF a piece of code is called frequently
THEN it is worth optimizing

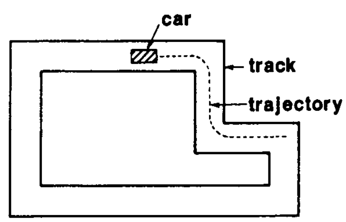

Figure 6. The Sugeno fuzzy car.

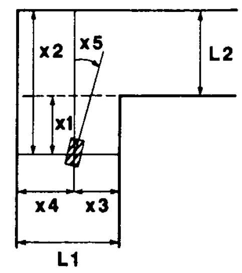

Figure 7. The state variables in Sugeno's car.

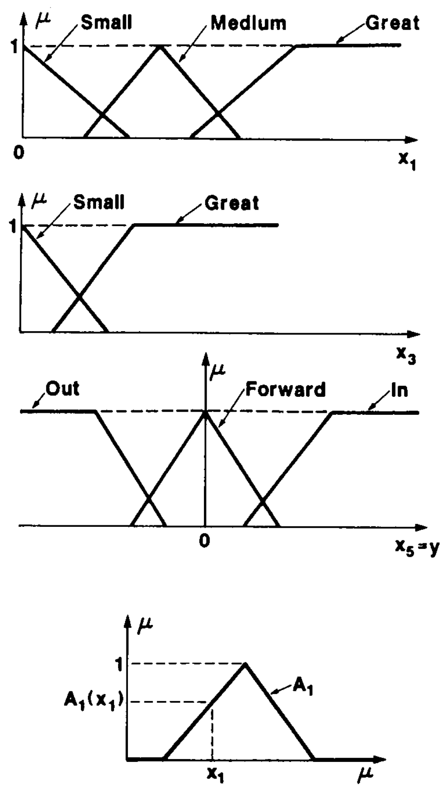

Figure 8. The linguistic values of state variables.

IF large oil spill or strong acid spill
THEN emergency is strongly suggested

The fuzziness of such rules is a consequence of the fact that a rule is a summary, and summaries, in general, are fuzzy. However, in the context of expert systems and fuzzy logic control, fuzziness has the positive effect of reducing the number of rules needed to approximately characterize a functional dependence between two or more variables.

Fuzzy hardware. Several expert system shells based on fuzzy logic are now commercially available, among them Reveal and Flops.[5] The seminal work of Togai and Watanabe at Bell Telephone Laboratories, which resulted in the development of a fuzzy logic chip, set the stage for using such chips in fuzzy-logic-based expert systems and, more generally, in rule-based systems not requiring a high degree of precision.[12] More recently, the fuzzy computer developed by Yamakawa of Kumamoto University has shown great promise as a general-purpose tool for processing linguistic data at high speed and with remarkable robustness.[5]

Togai and Watanabe's fuzzy inference chip consists of four major components: a rule set memory, an inference processor, a controller, and I/O circuitry. In a recent implementation, a rule set memory is realized by a random-access memory. In the inference processor, there are 16 data paths; one data path is laid out for each rule. All 16 rules on the chip are executed in parallel. The chip requires 64 clock cycles to produce an output. This translates to an execution speed of approximately 250,000 fuzzy logical inferences per second (FLIPS) at 16 megahertz clock. A fuzzy inference accelerator, which is a coprocessor board for a designated computer, is currently being designed. This board accommodates the new chips.

In the current implementation, the control variables are assumed to range over a finite set having no more than 31 elements. The membership function is quantized at 16 levels, with 15 representing full membership. Once the Togai/Watanabe chip becomes available commercially in 1988 or 1989, it should find many uses in both fuzzy-logic-based intelligent controllers and expert systems.

Yamakawa's fuzzy computer, whose hardware was built by OMRON Tateise Electronics Corporation, is capable of performing fuzzy inference at the very high speed of 10 megaFLIPS. Yamakawa's computer employs a parallel architecture. Basically, it has a fuzzy memory, a set of inference engines, a MAX block, and a defuzzifier. The computer is designed to process linguistic inputs, for example, "more or less small" and "very large," which are represented by analog voltages on data buses. A binary RAM, an array of registers and a membership function generator form the computer's fuzzy memory.

The linguistic inputs are fed to inference engines in parallel, with each rule yielding an output. The outputs are aggregated in the MAX block, yielding an overall fuzzy output that appears in the output data bus as a set of distributed analog voltages. In intelligent fuzzy control and other applications requiring nonfuzzy commands, the fuzzy output is fed to a defuzzifier for transformation into crisp output.

Yamakawa's fuzzy computer may be an important step toward a sixth-generation computer capable of processing common sense knowledge. This capability is a prerequisite to solving many AI problems — for example, handwritten text recognition, speech recognition, machine translation, summarization, and image understanding — that do not lend themselves to cost-effective solution within the bounds of conventional technology. □

Acknowledgment

The research reported in this article has been supported in part by NASA Grant NCC-2-275 and NSF Grant DCR-8513139. The article was written while the author was a visiting scholar at the Center for the Study of Language and Information at Stanford University. It is dedicated to the Japanese scientists and engineers who have contributed so importantly to the development of fuzzy logic and its applications.

References

1. L.A. Zadeh, "Fuzzy Algorithms," *Information and Control*, Vol. 12, 1968, pp. 94-102.
2. L.A. Zadeh, "A Rationale for Fuzzy Control," *J.Dynamic Systems, Measurement and Control*, Vol. 94, Series G, 1972, pp. 3-4.
3. L.A. Zadeh, "Outline of a New Approach to the Analysis of Complex Systems and Decision Processes," *IEEE Trans. Systems, Man, and Cybernetics*, Vol. SMC-3, 1973, pp. 28-44.
4. E.H. Mamdani and S. Assilian, "A Case Study on the Application of Fuzzy Set Theory to Automatic Control," *Proc. IFAC Stochastic Control Symp.*, Budapest, 1974.
5. *Preprints of the Second Congress of the International Fuzzy Systems Association*, Tokyo, Japan, 1987.
6. L.A. Zadeh, "Test-Score Semantics as a Basic for a Computational Approach to the Representation of Meaning," *Literary and Linguistic Computing*, Vol. 1, 1986, pp. 24-35.
7. D. Dubois and H. Prade, *Theorie des Possibilites*, Masson, Paris, 1985.
8. L.A. Zadeh, "The Role of Fuzzy Logic in the Management of Uncertainty in Expert Systems," *Fuzzy Sets and Systems*, Vol. 11, 1983, pp. 199-227.
9. L.A. Zadeh, "Syllogistic Reasoning in Fuzzy Logic and its Application to Usuality and Reasoning with Dispositions," *IEEE Trans. Systems, Man, and Cybernetics*, Vol. SMC-15, 1985, pp. 754-763.
10. A. Kaufmann and M. M. Gupta, *Introduction to Fuzzy Arithmetic*, Van Nostrand, New York, 1985.
11. B. Buchanan and E. K. Shortliffe, *Rule-Based Expert Systems*, Addison-Wesley, Reading, Mass., 1984.
12. M. Togai and H. Watanabe, "Expert Systems on a Chip: An Engine for Real-Time Approximate Reasoning," *IEEE Expert*, Vol. 1, 1986, pp. 55-62.

Related Publications

Adams, E.W., and H.F. Levine, "On the Uncertainties Transmitted from Premises to Conclusions in Deductive Inferences," *Synthese*, Vol. 30, 1975, pp. 429-460.

Baldwin, J.F., and S.Q. Zhou, "A New Approach to Approximate Reasoning Using a Fuzzy Logic," *Fuzzy Sets and Systems*, Vol. 2, 1979, pp. 302-325.

Bellman, R.E., and L.A. Zadeh, "Local and Fuzzy Logics," in G. Epstein, ed., *Modern Uses of Multiple-Valued Logic*, Reidel, Dordrecht, 1977, pp. 103-165.

Dubois, D., and H. Prade, *Fuzzy Sets and Systems: Theory and Applications*, Academic Press, New York, 1980.

Goguen, J.A., "The Logic of Inexact Concepts," *Synthese*, Vol 19, 1969, pp. 325-373.

Klir, G.J., and T.A. Folger, *Fuzzy Sets, Uncertainty, and Information*, Prentice Hall, Englewood Cliffs, N.J., 1988.

Moisil, G.C., *Lectures on the Logic of Fuzzy Reasoning*, Scientific Editions, Bucharest, 1975.

Prade, H., and C.V. Negoita, *Fuzzy Logic in Knowledge Engineering*, Verlag TÜV Rheinland, Köln, 1986.

Sugeno, M., ed., *Industrial Applications of Fuzzy Control*, Elsevier Science Publishers BV, The Netherlands, 1985.

Watanabe, H., and W. Dettloff, "Fuzzy Logic Inference Processor for Real Time Control: A Second Generation Full Custom Design," *Proc. 21st Asilomar Conference on Signals, Systems, and Computers*, Asilomar, Calif., 1987.

Lofti A. Zadeh joined the Department of Electrical Engineering at the University of California, Berkeley, in 1959, serving as its chairman from 1963 to 1968. Earlier he was a member of the electrical engineering faculty at Columbia University. He has held a number of visiting appointments, among them a visiting scientist appointment at the IBM Research Laboratory, San Jose, and a visiting scholar appointment at the AI Center, SRI International.

Zadeh's work centered on system theory and decision analysis until 1965 when his interests shifted to the theory of fuzzy sets and its applications.

An alumnus of the University of Teheran in Iran, MIT, and Columbia University, Zadeh is a fellow of the IEEE and AAAS and a member of the National Academy of Engineering. He holds a doctorate honoris causa from Paul Sabatier University, France, in recognition of his development of the theory of fuzzy sets.

Zadeh's address is Computer Science Division, Dept. of EECS, University of California, Berkeley, CA 94720.