Copyright © 1979, by the author(s).
All rights reserved.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission.

APPROXIMATE REASONING BASED ON FUZZY LOGIC

by

L. A. Zadeh

Memorandum No. UCB/ERL M79/32

May 1979

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720

February 1, 1979

APPROXIMATE REASONING BASED ON FUZZY LOGIC

L.A. Zadeh*

Abstract

During the past several years, the emergence of expert systems as a field of considerable practical as well as theoretical importance within AI has provided a strong impetus for the development of theories of approximate reasoning and credibility assessment of inference processes in knowledge-based systems.

The approach to approximate reasoning described in this paper is based on a fuzzy logic, FL, in which the truth-values and quantifiers are defined as possibility distributions which carry linguistic labels such as true, quite true, not very true, many, not very many, several, almost all, etc. Based on the concept of a possibility distribution, a set of translation and inference rules is developed and their application to inference from imprecise premises is illustrated by examples.

Key words: approximate reasoning, fuzzy logic, expert systems.

*Computer Science Division, Department of Electrical Engineering and Computer Sciences and the Electronics Research Laboratory, University of California, Berkeley, CA 94720. Research supported by Naval Electronic Systems Command Contract N00039-78-C0013 and National Science Foundation Grant MCS77-07568.

APPROXIMATE REASONING BASED ON FUZZY LOGIC

L.A. Zadeh*

1. Introduction

It has long been recognized that much, perhaps most, of human reasoning is approximate rather than exact in nature.[1] And yet, logicians and cognitive scientists have devoted scant attention to the development of theories of approximate reasoning, partly because of the deeply entrenched tradition of respect for what is precise and disdain for what is fuzzy, and partly because there was no compelling need for such theories before the advent of artificial intelligence.

In recent years, however, the rapid growth of interest in natural language processing and the emergence of expert systems as an important application area within AI have made it increasingly clear that a better understanding of the processes of approximate reasoning is a prerequisite to the development of knowledge-based systems which can manipulate information that is imprecise, incomplete or not totally reliable.

In a series of papers starting in 1973 [2]-[7], we have advanced the view that conventional logical systems do not provide an appropriate basis for approximate reasoning and have suggested a fuzzy logic, FL, for this purpose. In contrast to two-valued and multi-valued logics, the truth-values of FL are linguistic rather than numerical, as are quantifiers exemplified by

---

*Computer Science Division, Department of Electrical Engineering and Computer Sciences and the Electronics Research Laboratory, University of California, Berkeley, CA 94720. Research supported by Naval Electronic Systems Command Contract N00039-78-C0013 and National Science Foundation Grant MCS77-07568.

[1] The terms "approximate reasoning" and "fuzzy reasoning" are frequently used interchangeably in the literature. A comprehensive exposition of the foundations of fuzzy reasoning may be found in the paper by B.R. Gaines [1].

many, most, not very many, few, several, almost all, etc. Furthermore, the rules of inference in FL are approximate rather than exact. Thus, in spirit as well as in substance, FL represents a sharp break with the long-standing tradition of extreme precision in logic and a move toward an accommodation with the pervasive imprecision of human thought, perception, language and decision-making.

Another important difference between FL and classical logical systems centers on the concept of truth. Thus, whereas in the latter systems [8]-[9] the truth of a proposition serves as a starting point for the definition of various basic concepts, the corresponding role in FL is played not by truth but by the concept of a possibility distribution, by which is meant a mapping associated with a variable X, $\pi_X: U \rightarrow [0,1]$, which assigns to each point u in the domain of X, U, the possibility that X could take u as a value. The concept of truth, then, is employed to express the compatibility of the possibility distribution induced by a proposition p with that of a reference proposition r. In this sense, the truth of a proposition is a relative concept and is a measure of its compatibility or consistency with a collection of propositions which constitute a database.

Our main concern in the present paper is with (a) the establishment of translation rules for various types of imprecise propositions; (b) inference from such propositions; and (c) the development of a conceptual framework for dealing with the issues of belief and credibility. The latter issues play a particularly important role in expert systems [10]-[17] as well as in the theories of evidence and decision analysis [18]-[28].

2. The Concept of a Possibility Distribution

As a preliminary to the consideration of translation rules for various types of propositions, it will be helpful to establish some of the properties of possibility distributions which will be needed in later sections.[2]

Informally, if X is a variable taking values in U, then a possibility distribution, Π_X, associated with X may be viewed as a fuzzy (or elastic) constraint on the values that may be assigned to X. Such a distribution is characterized by a possibility distribution function $\pi_X: U \to [0,1]$ which associates with each $u \in U$ the "degree of ease" or the possibility that X may take u as a value.

In some cases, the constraint on the values of X is physical in origin. For example, if X is the number of students in a classroom which has, say, 50 seats, then $\pi_X(u) = 1$ for u up to 50, with $\pi_X(u)$ gradually declining to 0 at, say, $u = 75$. In this case, an intermediate value of π_X, say $\pi_X(65) = 0.3$, would signify that, by some explicit or implicit criterion, the degree of ease with which 65 students could be crowded into the classroom is 0.3 (on the scale from 0 to 1).

In most cases, however, the possibility distribution which is associated with a variable is epistemic rather than physical in origin. A basic assumption in fuzzy logic is that such epistemic possibility distributions are induced by propositions expressed in a natural language. In more concrete terms, this assumption may be stated as the

Possibility Postulate. If F is a fuzzy subset of U characterized by its membership function $\mu_F: U \to [0,1]$, then the proposition "X is F" induces a possibility distribution Π_X which is equal to F. Equivalently, "X is F"

[2] An introductory exposition of possibility theory may be found in [29]. A historical account of the concepts of possibility and probability is contained in [30].

translates into the <u>possibility assignment equation</u> $\Pi_X = F$, i.e.,

$$X \text{ is } F \rightarrow \Pi_X = F \tag{1}$$

which signifies that the proposition "X is F" has the effect of constraining the values that may be assumed by X, with the possibility distribution Π_X identified with F.

As a simple illustration of (1), if in the proposition "X is small" small is regarded as a label of a fuzzy subset of $U = \{0,1,2,\ldots\}$ which is defined by[3]

$$\text{SMALL} = 1/0 + 1/1 + 0.8/2 + 0.6/3 + 0.4/4 + 0.2/5 \tag{2}$$

then

$$\begin{aligned} &\text{Poss}\{X=0\} = 1 \\ &\text{Poss}\{X=1\} = 1 \\ &\text{Poss}\{X=2\} = 0.8 \\ &\text{- - - - - - - - - -} \\ &\text{Poss}\{X=5\} = 0.2 \end{aligned} \tag{3}$$

where $\text{Poss}\{X=u\} = \pi_X(u)$ is the possibility that X may take u as a value.

An important aspect of the concept of a possibility distribution is that it is nonstatistical in nature. As a consequence, if P_X is a probability distribution associated with X then the only connection between Π_X and P_X is that impossibility (i.e., zero possibility) implies improbability but not vice-versa. Thus, Π_X cannot be inferred from P_X nor can P_X be inferred from Π_X.

[3]The notation $F = \mu_1/x_1 + \cdots + \mu_n/x_n$ signifies that F is a fuzzy subset of the set $\{x_1,\ldots,x_n\}$, with μ_i, $i = 1,\ldots,n$, being the grade of membership of x_i in F. Uppercase symbols (e.g., F, SMALL) are employed to denote names of sets (fuzzy or nonfuzzy).

As in the case of probabilities, one can define joint and conditional possibilities. Thus, if X and Y are variables taking values in U and V, respectively, then we can define the joint and conditional possibility distributions through their respective distribution functions:

$$\pi_{(X,Y)}(u,v) = \text{Poss}\{X=u,\ Y=v\}\ ,\quad u\in U,\ v\in V \tag{4}$$

and

$$\pi_{(X|Y)}(u|v) = \text{Poss}\{X=u|Y=v\} \tag{5}$$

where (5) represents the conditional distribution function of X given Y.

If we know the distribution function of X and the conditional distribution function of Y given X, then we can construct the joint distribution function of X and Y by forming the conjunction ($\wedge \triangleq \min$)

$$\pi_{(X,Y)}(u,v) = \pi_X(u) \wedge \pi_{(Y|X)}(v|u)\ . \tag{6}$$

However, unlike the identity that holds in the case of probabilities, we can also obtain $\pi_{(X,Y)}(u,v)$ by forming the conjunction of $\pi_{(X|Y)}(u|v)$ and $\pi_{(Y|X)}(v|u)$:

$$\pi_{(X,Y)}(u,v) = \pi_{(X|Y)}(u|v) \wedge \pi_{(Y|X)}(v|u)\ . \tag{7}$$

In yet another deviation from parallelism with probabilities, the marginal possibility distribution function of X may be expressed in more than one way in terms of the joint and conditional possibility distribution functions. More specifically, we may have

(a) $$\pi_X(u) = \vee_v \pi_{(X,Y)}(u,v) \tag{8}$$

where $\vee_v$ denotes the supremum over $v \in V$;

(b) $$\pi_X(u) = \vee_v \pi_{(X|Y)}(u|v) \tag{9}$$

and

(c) $$\pi_X(u) = \pi_{(X|Y)}(u,\tilde{v}(u)) \tag{10}$$

where, for a given u, $\tilde{v}(u)$ is the value of v at which $\pi_{(X|Y)}(u,v) = 1$, if $\tilde{v}(u)$ is defined for every $u \in U$.

Intuitively, (a) represents the possibility of assigning a value to X as perceived by an observer ((X,Y) observer) who observes the joint possibility distribution $\Pi_{(X,Y)}$. Similarly, (b) represents the perception of an observer ($(X|Y)$ observer) who observes only the conditional possibility distribution $\Pi_{(X|Y)}$ and is unconcerned with or unaware of $\Pi_{(Y|X)}$. And (c) expresses the perception of an observer who assumes that v is assigned that value, if it exists, which makes $\pi_{(X|Y)}(u,v)$ equal to unity.

In relating π_X to $\pi_{(X|Y)}$ through the operator $\vee$ (supremum), we are tacitly invoking the principle of maximal restriction [4], which asserts that, in the absence of complete information about Π_X, we should equate Π_X to the maximal (i.e., least restrictive) possibility distribution which is consistent with the partial information about Π_X. In the case of (a), for example, this would be the supremum of $\pi_{(X,Y)}(u,v)$ over $v \in V$.

As will be seen in Section 2, the concept of a conditional possibility distribution plays a basic role in the formulation of a generalized form of <u>modus ponens</u> and in defining a measure of belief. What is as yet an unsettled issue revolves around the question of how to derive $\pi_{(X|Y)}$ and $\pi_{(Y|X)}$ from $\pi_{(X,Y)}$. Somewhat different answers to this question are presented in [27], [31] and [32]. It may well turn out to be the case that, in contrast to probabilities, there does not exist a unique solution to the problem and that, in general, the answer depends on the perspective of the observer.

3. Possibility Theory and Fuzzy Logic

As was alluded to already, the concept of a possibility distribution and the possibility theory which is based on it, play a central role in fuzzy logic by providing a means for the representation of the meaning of imprecise premises and the generation of propositions which follow logically from them.

Typically, a variable in fuzzy logic is treated as a *linguistic variable* [4], that is, a variable whose values are represented as words or sentences in a natural or synthetic language, with each such value defining a possibility distribution in the domain of the variable. In effect, what this implies is that a linguistic variable is a microlanguage with its own syntax and semantics. For example, in the case of FL, the linguistic values of the variable *Truth* may be generated by a context-free grammar and interpreted by an attributed grammar [33]. Thus, starting with (a) the primary term *true* and its antonym *false*; and (b) a finite set of modifiers and connectives such as *and*, *or*, *not*, *very*, *more or less*, *quite*, *extremely*, etc., the linguistic values of *Truth* may be represented as:

true	false
not true	not false
very true	very false
not very true	not very false
more or less true	more or less false
quite true	quite false
not quite true	not quite false
- - - - -	- - - - -

not true and not false
not very true and not very false
- - - - -

In FL, a linguistic truth-value is regarded as a composite label of a possibility distribution in the interval [0,1], which is the set of truth-values in Lukasiewicz's L_{Aleph_1} logic. What this means is that in FL we generally deal not with numerically-valued truth-values but with their possibility distributions.

An important distinction between FL and L_{Aleph_1} is that in L_{Aleph_1} -- as in all other multi-valued logics -- there are only two quantifiers <u>all</u> and <u>some</u>, whereas in FL we can employ a large variety of fuzzy quantifiers exemplified by <u>few</u>, <u>several</u>, <u>many</u>, <u>most</u>, <u>almost all</u>, <u>very many</u>, <u>not very many</u>, etc. This feature of FL makes it possible to translate, and infer from, imprecise premises exemplified by

Most tall women are not very fat.

Among the many men who are heavy smokers, quite a few are overweight and some are heavy drinkers.

Carol has several close friends who have many children.

The meaning of a quantifier in FL is based on the concept of the cardinality of a fuzzy set. Thus, if F is a fuzzy subset of a finite set $U = \{u_1,\ldots,u_n\}$ characterized as

$$F = \mu_1/u_1 + \cdots + \mu_N/u_N$$

where μ_i is the grade of membership of u_i in F, $i = 1,\ldots,N$, then the <u>power</u> of F -- which is roughly a measure of its cardinality [34] -- may be defined as

$$|F| = \sum_{i=1}^{n} \mu_i . \tag{11}$$

Alternatively, and perhaps more appropriately, the cardinality of F may be defined as a fuzzy number, as is done in [6]. Thus, if the elements of F are sorted in descending order, so that $\mu_n \leq \mu_m$ if $n \geq m$, then the truth-value of the proposition

$$p \triangleq F \text{ has at least n elements} \tag{12}$$

is defined to be equal to μ_n, while that of q,

$$q \triangleq F \text{ has at most n elements} , \tag{13}$$

is taken to be $1 - \mu_{n+1}$. From this it follows that the truth-value of the proposition r,

$$r \triangleq F \text{ has exactly n elements} , \tag{14}$$

is given by $\mu_n \wedge (1-\mu_{n+1})$. An illustration of these alternative definitions of cardinality is provided by Example 5 in Section 4.

The translation rules for propositions expressed in a natural language form an important part of PRUF -- a meaning representation language based on possibility theory [7].[4] In what follows, we shall present in a summarized form a subset of such rules which play a basic role in FL.

Translation Rules

1. Modifier rule

If

$$X \text{ is } F \rightarrow \Pi_X = F \tag{15}$$

then

$$X \text{ is } mF \rightarrow \Pi_X = F^+$$

where m is a modifier such as not, very, more or less, etc., and F^+ is a modification of F induced by m. More specifically: If m = not, then

[4]There are several implemented languages which have extensive facilities for the manipulation of fuzzy propositions and execution of fuzzy instructions. Prominent among these are FUZZY [35] and FSTDS [36], [66].

$F^+ = F' =$ complement of F, i.e.,

$$\mu_{F^+}(u) = 1 - \mu_F(u) \ , \quad u \in U \ . \tag{16}$$

If $m =$ very, then $F^+ = F^2$, i.e.,

$$\mu_{F^+}(u) = \mu_F^{\,2}(u) \ , \quad u \in U \ . \tag{17}$$

If $m =$ more or less, then $F^+ = \sqrt{F}$, i.e.,

$$\mu_{F^+}(u) = \sqrt{\mu_F(u)} \ , \quad u \in U \ . \tag{18}$$

As a simple illustration of (15), if SMALL is defined as in (2), then

$$X \text{ is very small} \rightarrow \Pi_X = F^2 \tag{19}$$

where

$$F^2 = 1/0 + 1/1 + 0.64/2 + 0.36/3 + 0.16/4 + 0.04/5 \quad . \tag{20}$$

It should be noted that (16), (17) and (18) should be viewed as default rules which may be replaced by other translation rules in cases in which some alternative interpretations of the modifiers <u>very</u> and <u>more or less</u> are more appropriate.

2. <u>Conjunctive, disjunctive and implicational rules</u>

If

$$X \text{ is } F \rightarrow \Pi_X = F \quad \text{and} \quad Y \text{ is } G \rightarrow \Pi_Y = G \tag{21}$$

where F and G are fuzzy subsets of U and V, respectively, then

(a)
$$X \text{ is } F \text{ and } Y \text{ is } G \rightarrow \Pi_{(X,Y)} = F \times G \tag{22}$$

where

$$\mu_{F\times G}(u,v) \triangleq \mu_F(u) \wedge \mu_G(v) \qquad (\wedge = \min) \ . \tag{23}$$

(b) $$X \text{ is } F \text{ or } Y \text{ is } G \rightarrow \Pi_{(X,Y)} = \bar{F} \cup \bar{G} \tag{24}$$

where

$$\bar{F} \triangleq F \times V \;, \quad \bar{G} \triangleq U \times G \tag{25}$$

and

$$\mu_{\bar{F} \cup \bar{G}}(u,v) = \mu_F(u) \vee \mu_G(v) \;. \tag{26}$$

(c) $$\text{If } X \text{ is } F \text{ then } Y \text{ is } G \rightarrow \Pi_{(Y|X)} = \bar{F}' \oplus \bar{G} \tag{27}$$

where $\Pi_{(Y|X)}$ denotes the conditional possibility distribution of Y given X, and the bounded sum $\oplus$ is defined by

$$\mu_{\bar{F}' \oplus \bar{G}}(u,v) = 1 \wedge (1 - \mu_F(u) + \mu_G(v)) \;. \tag{28}$$

As simple illustrations of (22), (24) and (27), if

$$F \triangleq SMALL = 1/1 + 0.6/2 + 0.1/3 \tag{29}$$

$$G \triangleq LARGE = 0.1/1 + 0.6/2 + 1/3 \tag{30}$$

then

$$\begin{aligned} &X \text{ is small and } Y \text{ is large} \rightarrow \Pi_{(X,Y)} \\ &\quad = 0.1/(1,1) + 0.6/(1,2) + 1/(1,3) + 0.1/(2,1) + 0.6/(2,2) \\ &\qquad + 0.6/(2,3) + 0.1/(3,1) + 0.1/(3,2) + 0.1/(3,3) \end{aligned} \tag{31}$$

$$\begin{aligned} &X \text{ is small or } Y \text{ is large} \rightarrow \Pi_{(X,Y)} \\ &\quad = 1/(1,1) + 1/(1,2) + 1/(1,3) + 0.6/(2,1) + 0.6/(2,2) + 1/(2,3) \\ &\qquad + 0.1/(3,1) + 0.6/(3,2) + 1/(3,3) \end{aligned} \tag{32}$$

and

If X is small then Y is large $\rightarrow \Pi_{(Y|X)}$ (33)

$$= 0.1/(1,1) + 0.6/(1,2) + 1/(1,3) + 0.5/(2,1) + 1/(2,2) + 1/(2,3) + 1/(3,1) + 1/(3,2) + 1/(3,3) \; .$$

3. Quantification rule

If $U = \{u_1, \ldots, u_N\}$, Q is a quantifier such as many, few, several, all, some, most, etc., and

$$X \text{ is } F \rightarrow \Pi_X = F \tag{34}$$

then the proposition "QX are F" (e.g., "many X's are large") translates into

$$\Pi_{\text{Count}(F)} = Q \tag{35}$$

where Count(F) denotes the number (or the proportion) of elements of U which are in F. By definition (11), if the fuzzy set F is expressed as

$$F = \mu_1/u_1 + \mu_2/u_2 + \cdots + \mu_N/u_N \tag{36}$$

then

$$\text{Count}(F) = \sum_{i=1}^{N} \mu_i \; . \tag{37}$$

As a simple illustration of (35), if the quantifier several is defined as

$$\text{SEVERAL} \triangleq 0/1 + 0.4/2 + 0.6/3 + 1/4 + 1/5 + 1/6 + 0.6/7 + 0.2/8 \tag{38}$$

then

$$\text{Several X's are large} \rightarrow \Pi_{\sum_{i=1}^{N} \mu_{\text{LARGE}}(u_i)} \tag{39}$$

$$= 0/1 + 0.4/2 + 0.6/3 + 1/4 + 1/5 + 1/6 + 0.6/7 + 0.2/8$$

where $\mu_{\text{LARGE}}(u_i)$ is the grade of membership of the i^{th} value of X in the fuzzy set LARGE.

4. Truth qualification rule

Let τ be a fuzzy truth-value, e.g., very true, quite true, more or less true, etc. Such a truth-value may be regarded as a fuzzy subset of the unit interval which is characterized by a membership function $\mu_\tau: [0,1] \to [0,1]$.

A truth-qualified proposition, e.g., "It is τ that X is F," is expressed as "X is F is τ." As shown in [5], the translation rule for such propositions is given by

$$\text{X is F is } \tau \longrightarrow \Pi_X = F^+ \tag{40}$$

where

$$\mu_{F^+}(u) = \mu_\tau(\mu_F(u)) \ . \tag{41}$$

As an illustration, consider the truth-qualified proposition

Bob is young is very true

which by (40), (41) and (17) translates into

$$\Pi_{Age(Bob)} = \mu_{TRUE^2}(\mu_{YOUNG}(u)) \ . \tag{42}$$

Now, if we assume that

$$\mu_{YOUNG}(u) = (1+(\frac{u}{25})^2)^{-1} \ , \quad u \in [0,100] \tag{43}$$

and

$$\mu_{TRUE}(v) = v^2 \ , \quad v \in [0,1] \tag{44}$$

then (41) yields

$$\Pi_{Age(Bob)} = (1+(\frac{u}{25})^2)^{-4} \tag{45}$$

as the possibility distribution of the age of Bob.

Used in combination, the translation rules stated above provide a system for the determination of the possibility distribution induced by a fairly complex composite proposition. For example, the proposition

If X is not very large and Y is more or less small then

Z is very very large.

induces the conditional possibility distribution described by

$$\pi_{(Z|X,Y)}(w|u,v) = 1 \wedge (1 - (1 - \mu^{2}_{LARGE}(u)) \wedge \mu^{0.5}_{SMALL}(v) + \mu^{4}_{LARGE}(w)) \quad . \tag{46}$$

It should be noted that rules of this type have found practical applications in the design of fuzzy logic controllers in steel plants, cement kilns and other types of industrial process control applications in which instructions expressed in a natural language are transformed into control signals [61]-[65].

Rules of Inference

In our approach to approximate reasoning, the translation rules for imprecise propositions serve as a preliminary to the application of various rules of inference to the possibility distributions which are induced by such propositions, leading to other possibility distributions which upon retranslation yields the conclusions which may be drawn from the premises.

More specifically, the basic rules of inference in FL are the following.

1. Projection rule

Consider a fuzzy proposition whose translation is expressed as

$$p \rightarrow \Pi_{(X_1,\ldots,X_n)} = F \tag{47}$$

and let $X_{(s)}$ denote a subvariable of the variable $X \triangleq (X_1,\ldots,X_n)$, i.e.,

$$X_{(s)} = (X_{i_1},\ldots,X_{i_k}) \tag{48}$$

where the index sequence $s \triangleq (i_1,\ldots,i_k)$ is a subsequence of the sequence $(1,\ldots,n)$.

Furthermore, let $\Pi_{X_{(s)}}$ denote the marginal possibility distribution of $X_{(s)}$; that is,

$$\Pi_{X_{(s)}} = \mathrm{Proj}_{U_{(s)}} F \tag{49}$$

where U_i, $i = 1,\ldots,n$, is the universe of discourse associated with X_i;

$$U_{(s)} = U_{i_1} \times \cdots \times U_{i_k} \tag{50}$$

and the projection of F on $U_{(s)}$ is defined by the possibility distribution function

$$\pi_{X_s}(u_{i_1},\ldots,u_{i_k}) = \mathrm{Sup}_{u_{j_1},\ldots,u_{j_m}} \mu_F(u_1,\ldots,u_n) \tag{51}$$

where $s' \triangleq (j_1,\ldots,j_m)$ is the index subsequence which is complementary to s, and μ_F is the membership function of F.

Now let q be a retranslation of the possibility assignment equation

$$\Pi_{X_{(s)}} = \mathrm{Proj}_{U_{(s)}} F \,. \tag{52}$$

Then, the projection rule asserts that q may be inferred from p. In a schematic form, this assertion may be expressed more transparently as

$$\begin{array}{c} p \longrightarrow \Pi_{(X_1,\ldots,X_n)} = F \\ \downarrow \\ q \longleftarrow \Pi_{X_{(s)}} = \mathrm{Proj}_{U_{(s)}} F \end{array} \tag{53}$$

The operation of projection is easy to perform when Π_X is expressed as a linear form. As an illustration, assume that $U_1 = U_2 = \{a,b\}$, and

$$\Pi_{(X_1,X_2)} = 0.8aa + 0.6ab + 0.4ba + 0.2bb \tag{54}$$

in which a term of the form 0.6ab signifies that

$$\mathrm{Poss}\{X_1 = a,\ X_2 = b\} = 0.6\ . \tag{55}$$

To obtain the projection of Π_X on, say, U_2 it is sufficient to replace the value of X_1 in each term by the null string Λ. Thus

$$\mathrm{Proj}_{U_2}\, \Pi_{(X_1,X_2)} = 0.8a + 0.6b + 0.4a + 0.2b = 0.8a + 0.6b \tag{56}$$

and hence from the proposition

$$(X_1,X_2) \ \text{ is } \ 0.8aa + 0.6ab + 0.4ba + 0.2bb \tag{57}$$

we can infer by (53) that

$$X_1 \ \text{ is } \ 0.8a + 0.6b\ . \tag{58}$$

2. <u>Conjunction rule</u>

Consider a proposition p which is an assertion concerning the possible values of, say, two variables X and Y which take values in U and V, respectively. Similarly, let q be an assertion concerning the possible values of the variables Y and Z, taking values in V and W. With these assumptions, the translations of p and q may be expressed as

$$\begin{aligned} p &\rightarrow \Pi^p_{(X,Y)} = F \\ q &\rightarrow \Pi^q_{(Y,Z)} = G \end{aligned}\ . \tag{59}$$

Let $\bar{F}$ and $\bar{G}$ be, respectively, the cylindrical extensions of F and G in $U \times V \times W$. Thus,

$$\bar{F} = F \times W \tag{60}$$

and

$$\bar{G} = U \times G\ . \tag{61}$$

Using the conjunction rule, we can infer from p and q a proposition which is defined by the following scheme (the reverse arrow $\leftarrow$ denotes retranslation, i.e., reverse translation):

$$r \rightarrow \Pi^{p}_{(X,Y)} = F \tag{62}$$

$$q \rightarrow \Pi^{q}_{(Y,Z)} = G \tag{63}$$

$$r \leftarrow \Pi_{(X,Y,Z)} = \bar{F} \cap \bar{G} \tag{64}$$

On combining the projection and conjunction rules, we obtain the <u>compositional rule of inference</u> (67) which includes the classical <u>modus ponens</u> as a special case.

More specifically, on applying the projection rule to (64), we obtain the following inference scheme

$$\begin{array}{l} p \rightarrow \Pi^{p}_{(X,Y)} = F \\ q \rightarrow \Pi^{q}_{(Y,Z)} = G \\ \hline r \leftarrow \Pi^{r}_{(X,Z)} = F \circ G \end{array} \tag{65}$$

where the composition of F and G is defined by

$$\mu_{F \circ G}(u,w) = \mathrm{Sup}_{v}(\mu_F(u,v) \wedge \mu_G(v,w)) \ . \tag{66}$$

In particular, if p is a proposition of the form "X is F" and q is a proposition of the form "If X is G then Y is H," then (65) becomes

$$\begin{array}{l} p \rightarrow \Pi_{X} = F \\ q \rightarrow \Pi_{(Y|X)} = \bar{G}' \oplus \bar{H} \\ \hline r \leftarrow \Pi_{(Y)} = F \circ (\bar{G}' \oplus \bar{H}) \end{array} \tag{67}$$

The rule expressed by (67) may be viewed as a generalized form of modus ponens which reduces to the classical modus ponens when $F = G$ and F, G, H are nonfuzzy sets.

Stated in terms of possibility distributions, the generalized modus ponens places in evidence the analogy between probabilistic and possibilistic inference. Thus, in the case of probabilities, we can deduce the probability distribution of Y from the knowledge of the probability distribution of X and the conditional probability distribution of Y given X. Similarly, in the case of possibility distributions, we can infer the possibility distribution of Y from the knowledge of the possibility distribution of X and the conditional possibility distribution of Y given X.

It is important to note that the generalized modus ponens as expressed by (67) may be used to enlarge significantly the area of applicability of rule-based systems of the type employed in MYCIN and other expert systems. This is due primarily to two aspects of (67) which are not present in conventional rule-based systems: (a) in the propositions "X is F" and "If X is G then Y is H," F, G and H may be fuzzy sets; and (b) F and G need not be identical. Thus, as a result of (a) and (b), a rule-based system employing (67) may be designed to have an interpolative capability [6].

4. Probability/Possibility and Credibility Analysis

An important issue in approximate reasoning relates to the assessment of the credibility of a conclusion which is inferred from a collection of premises which are imprecise and/or have less than complete credibility. This issue plays a particularly important role in the case of expert systems, e.g., MYCIN, in which it is essential for the user to have an indication of the degree of confidence with which a conclusion supplied by the system may be used as a basis for a decision [12], [13].

Viewed from the perspective of fuzzy logic, the problem of credibility assessment bears an intimate relation to inference from propositions in which the imprecision is partly possibilistic and partly probabilistic. A simple example of a proposition of this type is the probability-qualified proposition

$$p \triangleq X \text{ is small is likely} \tag{68}$$

in which likely is interpreted as a linguistic probability [4], i.e., a possibility distribution of a numerical probability which takes values in the interval $[0,1]$. Thus, likely is characterized by a possibility distribution function $\pi_{LIKELY}(v)$, $0 \le v \le 1$, which expresses one's subjective perception of the meaning of likely.

As shown in [5], p translates into a possibility distribution of probability distributions. More specifically, if $\rho_X(u)du$ is the probability that X lies in the interval $[u,u+du]$, then the possibility distribution function of ρ_X which is induced by p is expressed by

$$\pi(\rho_X) = \pi_{LIKELY}\left(\int_U \mu_{SMALL}(u)\rho_X(u)du\right) \tag{69}$$

where μ_{SMALL} denotes the membership function of small in (68).

If p in (68) is viewed as a piece of evidence, then its translation (69) shows that the evidence in question is a possibility distribution of probability distributions. A dual case which plays an important role in Shafer's theory of evidence and in the calculation of upper and lower probabilities [37], [38], [21], is one in which the evidence may be viewed as a probability distribution of possibility distributions. More specifically, let X be a random variable which is associated with a probability distribution P_X and let $\Pi_{(Y|X)}$ be the conditional possibility distribution of a variable Y given X. Then, viewing P_X and $\Pi_{(Y|X)}$ as the available evidence about Y, the question is: What is the degree of credibility which can be associated with the proposition $q \triangleq Y$ is in A, where A is a subset of the domain of Y?

For our purposes, it will be convenient to formulate an answer to this question in terms of the concepts of conditional certainty:

$$\mathrm{Cert}\{Y \in A | X\} \tag{70}$$

and conditional possibility

$$\mathrm{Poss}\{Y \in A | X\} \quad . \tag{71}$$

Thus, we shall associate with Y two measures of credibility: (a) the expectation of conditional certainty

$$C(A) \triangleq \mathrm{Cert}\{Y \in A\} = E_X \mathrm{Cert}\{Y \in A | X\} \tag{72}$$

and (b), the expectation of conditional possibility

$$\Pi(A) = \mathrm{Poss}\{Y \in A\} = E_X \mathrm{Poss}\{Y \in A | X\} \quad . \tag{73}$$

For the case where $\Pi_{(Y|X)}$ can take only two values, 0 or 1, which is equivalent to the case considered by Shafer,

$$\begin{aligned} \mathrm{Cert}\{Y \in A|X\} &= 1 \quad \text{if} \quad \Pi(Y|X) \in A \\ &= 0 \quad \text{otherwise} \end{aligned} \tag{74}$$

while

$$\begin{aligned} \mathrm{Poss}\{Y \in A|X\} &= 1 \quad \text{if} \quad \Pi(Y|X) \notin A' \\ &= 0 \quad \text{otherwise} \end{aligned} \tag{75}$$

where A' is the complement of A in the domain of Y. In this case, $C(A)$ as defined by (72) reduces to Shafer's belief function $B(A)$, while $\Pi(A)$ reduces to Shafer's plausibility function.

A basic problem in credibility analysis is that of defining a rule or rules for various combinations of evidence and hypotheses [12]. The so-called Dempster rule [37] relates, in essence, to the case where the evidence consists of $E_1 = \{P_{X_1}, \Pi_{(Y|X_1)}\}$ and $E_2 = \{P_{X_2}, \Pi_{(Y|X_2)}\}$, where X_1 and X_2 are independent random variables and the conditional possibility distributions are non-fuzzy. In this case, we have

$$C_1(A) = E_{X_1} \mathrm{Cert}\{Y \in A|X_1\} \tag{76}$$

$$C_2(A) = E_{X_2} \mathrm{Cert}\{Y \in A|X_2\} \tag{77}$$

and the rule in question yields the combined measure of certainty that $Y \in A$, i.e.,

$$C_{1,2}(A) = E_{(X_1,X_2)} \mathrm{Cert}\{Y \in A|(X_1,X_2)\} \quad . \tag{78}$$

It should be noted that, in the form stated in (78), the combination rule can readily be generalized to the case where $\pi_{(Y|X_1)}$ and $\pi_{(Y|X_2)}$ take values in the interval $[0,1]$, rather than just 0 or 1 -- which is the case considered by Dempster and Shafer.[5]

The above results apply only to the special case where the evidence consists of P_X and $\Pi_{(Y|X)}$. At this juncture, there does not exist a more general and widely applicable theory of evidence which could serve as a basis for computing the credibility indices of conclusions inferred from a collection of imprecise premises. A difficulty that stands in the way of developing such a theory is that the relation between a hypothesis and evidence may involve a number of variables which are interrelated by a mixture of probability and possibility distributions. The rules for combining a mixture of such distributions tend to be cumbersome and complex, suggesting that it may be necessary to reconcile ourselves, at least for the foreseeable future, to the employment of techniques of credibility analysis which are ad hoc, approximate and heuristic, rather than formal and exact.

[5] It should be noted that the expression for $C_{1,2}$ as defined by (78) differs from that yielded by the Dempster/Shafer rule in that $C_{1,2}$ is not normalized. As shown in [67], the normalization employed by Dempster and Shafer is of questionable validity.

5. Examples of Approximate Reasoning

A concomitant of the imprecision which is inherent in approximate reasoning is that the chains of reasoning based on fuzzy logic are generally quite short in comparison with the long chains of reasoning which are possible when precise reasoning is employed. In what follows, we shall consider several simple examples which illustrate some of the basic aspects of the application of fuzzy logic to inference from fuzzy propositions. In some of these examples, we shall make use of concepts which form a part of the fuzzy-logic-based meaning representation language PRUF. A detailed exposition of this language may be found in [7].

Example 1. Suppose that we have the following information concerning three real numbers X, Y, Z.

X is small.

Y is approximately equal to X.

Z is much larger than both X and Y.

Question. How large is Z?

Let SMALL, APPROXIMATELY EQUAL, and MUCH LARGER THAN be the fuzzy subsets of R and R^2 ($R \triangleq$ real line) representing the denotations of small, approximately equal and much larger than, respectively. Then, on applying the compositional rule of inference (65) and the conjunctive rule (21), we obtain the following expression for the possibility distribution of Z:

$$\Pi_Z = (\text{MUCH LARGER THAN} \circ \text{APPROXIMATELY EQUAL} \circ \text{SMALL}) \cap \text{MUCH LARGER THAN} \circ \text{SMALL} \tag{79}$$

Example 2.

Premise: $p \triangleq$ Naomi is not very tall is true.

Question: How true is it that Naomi is tall?

Suppose that the answer to the question is expressed as a proposition q:

$$q \triangleq \text{Naomi is tall is } \tau \tag{80}$$

where τ is a linguistic truth-value, e.g., very true, more or less true, etc.

To determine τ, we set q semantically equal to p [7], i.e., we assert that the possibility distributions induced by p and q are equal. Now, by (17) and (40), we have

$$\text{Naomi is not very tall is true} \rightarrow \Pi_{\text{Height(Naomi)}} = F \tag{81}$$

where

$$\mu_F(u) = \mu_{TRUE}(1 - \mu^2_{TALL}(u)) \tag{82}$$

and

$$\text{Naomi is tall is } \tau \rightarrow \mu_\tau(\mu_{TALL}(u)) \tag{83}$$

where μ_{TALL} and μ_{TRUE} are the membership functions of TALL and TRUE, respectively. Consequently, for all u in the domain of the variable Height(Naomi), we have

$$\mu_{TRUE}(1 - \mu^2_{TALL}(u)) = \mu_\tau(\mu_{TALL}(u)) \tag{84}$$

from which it follows that the membership function of τ is given by

$$\mu_\tau(v) = 1 - v^2\,, \qquad v \in [0,1]\,. \tag{85}$$

Thus, if μ_{TRUE} is defined by

$$\mu_{TRUE}(v) = v^2 \tag{86}$$

then

$$\mu_\tau(v) = 1 - \mu_{TRUE}(v) \tag{87}$$

and hence

$$\tau = \text{not true} . \tag{88}$$

On the other hand, if

$$\mu_{TRUE}(u) = v \tag{89}$$

then

$$\mu_\tau(u) = 1 - \mu^2_{TRUE}(v) \tag{90}$$

and

$$\tau = \text{not very true} . \tag{91}$$

Example 3.

Premises: $p \triangleq$ Marvin lives near MIT.

$q \triangleq$ Lucia lives near MIT.

Question: What is the distance between the residences of Marvin and Lucia?

Let (X_M,Y_M) and $X_L,Y_L)$ be the coordinates of the residences of Marvin and Lucia, respectively. Furthermore, let $\Pi_{(X_M,Y_M)}$ and $\Pi_{(X_L,Y_L)}$ be the possibility distributions induced by p and q, that is, derived from the definition of the binary fuzzy relation NEAR.

Now, the distance between the residences of Marvin and Lucia is expressed by

$$d = \sqrt{(X_M-X_L)^2 + (Y_M-Y_L)^2} . \tag{92}$$

Using (92) and applying the extension principle [4], the possibility distribution function of d is found to be given by

$$\pi_d(w) = \mathrm{Sup}_{u_1,v_1,u_2,v_2}\left(\pi_{(X_M,Y_M)}(u_1,v_1) \wedge \pi_{(X_L,Y_L)}(u_2,v_2)\right) \tag{93}$$

subject to

$$w = \sqrt{(u_1-u_2)^2 + (v_1-v_2)^2} \tag{94}$$

where the supremum is taken over all possible values of X_M, Y_M, X_L and Y_L subject to the constraint (94). Generally, π_d as defined by (93) will be a monotone decreasing function of w, with $\pi_d(w) = 1$ for sufficiently small values of w.

Example 4.

Premise: $p \triangleq$ Marian is much taller than her close friends.

Question: How tall is Marian?

As a first step it is convenient, but not essential, to replace p with a proposition q which is semantically equivalent to p, namely:

$q \triangleq$ Not(Marian is not much taller than some of her close friends) . (95)

Next, assume that the database consists of three relations whose frames (i.e., names of relations and their attributes) are the following. (To rule out a trivial solution, we assume that Marian is not listed under Name in POPULATION.)

POPULATION ‖ Name | Height |

FRIEND ‖ Name1 | Name2 | μ |

MUCH TALLER ‖ Height1 | Height2 | μ |

In the second of these relations, μ is the degree to which the individual whose name appears under Name2 is a friend of the individual whose name appears in the same row under Name1. Similarly, in the third relation, μ is the degree to which an individual whose height is Height1 is much taller than one whose height is Height2.

Now in the notation of PRUF, the fuzzy set of Marian's close friends may be expressed as

$$F = {}_{\mu\times \mathrm{Name2}}\mathrm{FRIEND}^2[\mathrm{Name1} = \mathrm{Marian}] \tag{96}$$

which signifies that F is the projection of a particularized and intensified form of the relation FRIEND on the attributes μ and Name2. Typically, F is of the form

$$F = 0.8\mathrm{Jean} + 0.6\mathrm{Vicki} + 0.3\mathrm{Edie} + \cdots$$

which means that Jean, Vicki and Edie are Marian's close friends to the degree 0.8, 0.6 and 0.3, respectively.

Having F, we can obtain the fuzzy set of heights of Marian's close friends by forming the expression

$$H = {}_{\mathrm{Height}}\mathrm{POPULATION}[\mathrm{Name} = F] \tag{97}$$

which in conjunction with (95) implies that the possibility distribution of Marian's height is expressed by

$$\pi_{\mathrm{Height(Marian)}} = ({}_{\mu\times \mathrm{Height1}}\mathrm{MUCH\ TALLER}'[\Pi_{\mathrm{Height2}} = H])' \tag{98}$$

where the primes denote the complements (to account for the negations in (95)). From (98), then, it follows that the possibility distribution function of Marian's height is given by

$$\pi_{\text{Height(Marian)}}(u) = 1 - \text{Sup}_v(\mu_H(v) \wedge (1 - \mu_{\text{MUCH TALLER}}(u,v))) \ , \quad u \in U \quad (99)$$

where the supremum is taken over U, the domain of the variable Height.

It should be noted that the same result could be obtained by replacing p with the semantically equivalent proposition

$r \triangleq (\forall x)$(If x is a close friend of Marian then Marian is much taller than x)

provided that the implicational rule is taken to be

$$\text{If } X \text{ is } F \text{ then } Y \text{ is } G \longrightarrow \Pi_{(Y|X)} = \bar{F}' + \bar{G} \quad (100)$$

rather than (27).

This leads to the expression

$$\pi_{\text{Height(Marian)}}(u) = \text{Inf}_v((1 - \mu_H(v)) \vee \mu_{\text{MUCH TALLER}}(u,v)) \quad (101)$$

where Inf_v denotes the infimum over $v \in U$.

Example 5.

Premise: There are many more female students at Berkeley than rich students of both sexes.

Question: How many female students are there at Berkeley?

Let X be the number of female students at Berkeley. Our aim is to infer the possibility distribution of X from the given premise. To this end, assume that the frames of the relations in the database are of the form

STUDENT	Name	μ_{rich}

MANY MORE	ρ	γ	μ

In the first relation the names of all students are listed and μ_{rich} is the degree to which an individual is a rich student. In the second relation, μ is the degree to which the proportion ρ fits the description "many more" in relation to the proportion γ.

If the cardinality of a fuzzy set is identified with its power (11), then the proportion of rich students is expressed by

$$\frac{\mathrm{Count}({}_{\mathrm{Name}\times\mu_{rich}}\mathrm{STUDENT})}{\mathrm{Count}({}_{\mathrm{Name}}\mathrm{STUDENT})} \tag{102}$$

in which the numerator is the count of rich students and the denominator is the total count of students ($\triangleq N$) of both sexes. From the initial premise, then, it follows that the possibility distribution of the proportion of female students is given by

$$\Pi_\rho = {}_{\mu\times\rho}\mathrm{MANY\ MORE}[\gamma = \tfrac{1}{N}\mathrm{Count}({}_{\mathrm{Name}\times\mu_{rich}}\mathrm{STUDENT})] \ . \tag{103}$$

Alternatively, if the cardinality of the fuzzy set of rich students is assumed to be a fuzzy set, then the possibility distribution of ρ may be obtained as follows.

First, the fuzzy set STUDENT is sorted in descending order according to the values of μ_{rich}. The result, then, is of the form

$$\mathrm{STUDENT} = \mu_1\mathrm{Name}_1 + \mu_2\mathrm{Name}_2 + \cdots + \mu_N\mathrm{Name}_N \tag{104}$$

where μ_i is the degree to which Name_i is rich and $\mu_i \leq \mu_j$ for $j > i$.

The fuzzy cardinality of this set may be expressed as the fuzzy set

$$\mathrm{FCount(STUDENT)} = \mu_1 1 + \mu_2 2 + \cdots + \mu_N N \tag{105}$$

in which μ_n, $n = 1,\ldots,N$, represents the degree to which the fuzzy set

STUDENT has $\geq n$ elements.

The original premise may now be expressed compactly in the form

$$p \triangleq \gamma \text{ is many more than } \frac{1}{N} \text{FCount(STUDENT)} \ . \tag{106}$$

On applying to (106) the approach employed in Example 4, we are led to the following explicit expression for the possibility distribution function of ρ

$$\pi_\rho(u) = \text{Inf}_n((1-\mu_n) \vee \mu_{\text{MANY MORE}}(u,\tfrac{n}{N})) \ , \quad u \in [0,1] \ . \tag{107}$$

This expression represents the desired possibility distribution function of the proportion of female students at Berkeley given the premise: There are many more female students at Berkeley than rich students.

6. Concluding Remarks

The approach described in this paper may be viewed as an attempt at constructing a conceptual framework for inference from propositions whose meaning is not sharply defined. Through the use of fuzzy logic, the answer to a query is usually expressed in the form of a possibility distribution of one or more variables. In contrast to the conventional techniques of inference, the standards of precision in fuzzy logic are generally not high. More importantly, through the use of linguistic variables and linguistic approximation, these standards can be adjusted to fit the imprecision and unreliability of the information which is resident in the database.

References and Bibliography

[1] B.R. Gaines, "Foundations of fuzzy reasoning," Int. J. of Man-Machine Studies 6: 623-668, 1975.

[2] L.A. Zadeh, "Outline of a new approach to the analysis of complex systems and decision processes," IEEE Trans. Systems, Man and Cybernetics SMC-3: 28-44, 1973.

[3] L.A. Zadeh, "Fuzzy logic and approximate reasoning (in memory of Grigore Moisil)," Synthese 30: 407-428, 1975.

[4] L.A. Zadeh, "The concept of a linguistic variable and its application to approximate reasoning. Part I," Information Sciences 8: 119-249; Part II, Information Sciences 8: 301-357; Part III, Information Sciences 9: 43-80, 1975.

[5] L.A. Zadeh, "Local and fuzzy logics" (with R.E. Bellman), in Modern Uses of Multiple-Valued Logic, J.M. Dunn and G. Epstein, eds., D. Reidel, Dordrecht, 1977, pp. 103-165.

[6] L.A. Zadeh, "A theory of approximate reasoning," Memorandum M77/58, Electronics Research Laboratory, University of California, Berkeley, 1977. To appear in Machine Intelligence 9.

[7] L.A. Zadeh, "PRUF--a meaning representation language for natural languages," Int. J. Man-Machine Studies 10: 395-460, 1978.

[8] A. Tarski, Logics, Semantics, Metamathematics, Clarendon Press, Oxford, 1956.

[9] S. Haack, Philosophy of Logics, Cambridge University Press, Cambridge, 1978.

[10] E. Feigenbaum, B.G. Buchanan and J. Lederberg, "On generality and problem solving: a case study using the DENDRAL program," Machine Intelligence 6, Edinburgh University Press, 1971.

[11] G.A. Gorry, J.P. Kassirer, A. Essig and W.B. Schwartz, "Decision analysis as the basis for computer-aided management of acute renal failure," Amer. J. Medicine 55: 473-484, 1973.

[12] E.H. Shortliffe and B.G. Buchanan, "A model of inexact reasoning in medicine," Mathematical Biosciences 23: 351-379, 1975.

[13] E. Shortliffe, MYCIN: Computer-Based Medical Consultations, American Elsevier, New York, 1976.

[14] R. Duda et al., "Development of a computer based consultant for mineral exploration (PROSPECTOR)," Annual Report SRI International, Nos. 5821 and 6415, 1977.

[15] P. Szolovits, L.B. Hawkinson and W.A. Martin, "An overview of OWL, a language for knowledge representation," LCS Memorandum, M.I.T., 1977.

[16] J. McCarthy, "Epistomological problems of artificial intelligence," Proc. Fifth IJCAI, M.I.T., 1977, pp. 1038-1044.

[17] D.G. Bobrow and T. Winograd, "An overview of KRL, a knowledge representation language," Cognitive Science 1: 3-46, 1977.

[18] R.G. Swinburne, An Introduction to Confirmation Theory, Methuen, London, 1973.

[19] P. Suppes, "The measurement of belief," J. Roy. Statist. Soc. B 36: 160-175, 1974.

[20] K.M. Colby, "Simulations of belief systems," in Computer Models of Thought and Language, R.C. Schank and K.M. Colby, eds., W. Freeman, San Francisco, 1973.

[21] G. Shafer, A Mathematical Theory of Evidence, Princeton University Press, Princeton, 1976.

[22] B.C. Bruce, "Belief systems and language understanding," N.I.H. Report CBM-TR-41, Rutgers University, 1975.

[23] R.P. Abelson, "The structure of belief systems," in Computer Models of Thought and Language, R.C. Schank and K.M. Colby, eds., W. Freeman, San Francisco, 1973.

[24] R.D. Rosenkrantz, Inference, Method and Decision, D. Reidel, Dordrecht, 1977.

[25] N. Rescher, Plausible Reasoning, Van Gorcum, Amsterdam, 1976.

[26] A. Tversky and D. Kahneman, "Judgment under uncertainty: heuristics and biases," Science 185: 1124-1131, 1974.

[27] R. Buxton, "The interpretation and justification of the subjective Bayesian approach to statistical inference," Brit. J. Phil. Sci. 29: 25-38, 1978.

[28] H.E. Kyburg, "Decisions, conclusions and utilities," Synthese 36: 87-96, 1977.

[29] L.A. Zadeh, "Fuzzy sets as a basis for a theory of possibility," Fuzzy Sets and Systems 1: 3-28, 1978.

[30] I. Hacking, The Emergence of Probability, Cambridge University Press, Cambridge, 1975.

[31] H.T. Nguyen, "On conditional possibility distributions," Fuzzy Sets and Systems 1: 299-309, 1978.

[32] E. Hisdal, "Conditional possibilities: independence and noninteraction," Fuzzy Sets and Systems 1: 283-297, 1978.

[33] P.M. Lewis, D.J. Rosenkrantz and R.E. Stearns, "Attributed translations," J. Computer and System Sciences 9: 279-307, 1974.

[34] A. DeLuca and S. Termini, "A definition of a nonprobabilistic entropy in the setting of fuzzy sets theory," Information and Control 20: 301-312, 1972.

[35] R.A. LeFaivre, "FUZZY: a programming language for fuzzy problem solving," Technical Report 202, Department of Computer Science, University of Wisconsin, Madison, 1974.

[36] K. Noguchi, M. Umano, M. Mizumoto and K. Tanaka, "Implementation of fuzzy artificial intelligence language FLOU," Technical Report on Automation and Language of IECE, 1976.

[37] A.P. Dempster, "Upper and lower probabilities induced by a multi-valued mapping," Ann. Math. Statist. 38: 325-329, 1967.

[38] P. Suppes and M. Zanotti, "On using random relations to generate upper and lower probabilities," Synthese 36: 427-440, 1977.

[39] J.F. Baldwin and B.W. Pilsworth, "Axiomatic approach to implication for approximate reasoning using a fuzzy logic," Report EM/FS 9, University of Bristol, 1978.

[40] M. Minsky, "A framework for representing knowledge," in The Psychology of Computer Vision, P. Winston, ed., McGraw-Hill, New York, 1975, pp. 211-380.

[41] A. Newell and H.A. Simon, Human Problem Solving, Prentice-Hall, Englewood Cliffs, 1972.

[42] R.C. Schank, Conceptual Information Processing, American Elsevier, New York, 1975.

[43] D. Bobrow and A. Collins, Representation and Understanding, Academic Press, New York, 1975.

[44] A. Collins, E. Warnock, N. Aiello and M. Miller, "Reasoning from an incomplete knowledge," in Representation and Understanding, D. Bobrow and A. Collins, eds., Academic Press, New York, 1975.

[45] J.S. Brown and R. Burton, "Multiple representations of knowledge for tutorial reasoning," in Representation and Understanding, D. Bobrow and A. Collins, eds., Academic Press, New York, 1975.

[46] D.A. Norman, D.E. Rumelhart and the LNR Research Group, Explorations in Cognition, W. Freeman, San Francisco, 1975.

[47] R.F. Simmons, "Semantic networks: their computation and use for understanding English sentences," in Computer Models of Thought and Language, R.C. Schank and K.M. Colby, eds., W. Freeman, San Francisco, 1973.

[48] W.A. Woods, R.M. Kaplan and B. Nash-Webber, "The lunar sciences natural language information system," BBN Report 2378, Cambridge, Mass., 1972.

[49] S. Fahlman, "A system for representing and using real-world knowledge," A.I. Memo 331, M.I.T., 1975.

[50] C. Rieger, "The representation and selection of commonsense knowledge for natural language comprehension," Proc. Georgetown University Linguistics Roundtable, 1976.

[51] G.A. Miller and P.N. Johnson-Laird, Language and Perception, Harvard University Press, Cambridge, 1976.

[52] I. Goldstein and S. Papert, "Artificial intelligence, language, and the study of knowledge," Cognitive Science 1: 84-123, 1977.

[53] B. Adams, "A probability model of medical reasoning and the MYCIN model," Mathematical Biosciences 32: 177-186, 1976.

[54] A. Joshi, "Inferencing in partial information," in Pattern Directed Inference, F. Hays-Roth and D. Waterman, eds., Academic Press, 1978.

[55] R. Reiter, "On reasoning by default," Proc. TINLAP-2, Urbana, 1978, pp. 210-218.

[56] A. Collins, "Fragments of a theory of human plausible reasoning," Proc. TINLAP-2, Urbana, 1978, pp. 194-201.

[57] J.A. Goguen, "Concept representation in natural and artificial languages: axioms, extension and applications for fuzzy sets," Int. J. Man-Machine Studies 6: 513-561, 1974.

[58] H. Wechsler, "Applications of fuzzy logic to medical diagnosis," Proc. Int. Symp. on Multiple-Valued Logic, University of Indiana, Bloomington, 1975, pp. 162-174.

[59] M. Mizumoto, S. Fukame and K. Tanaka, "Fuzzy reasoning methods by Zadeh and Mamdani, and improved methods," Proc. Third Workshop on Fuzzy Reasoning, Queen Mary College, London, 1978.

[60] J.F. Baldwin, "A model of fuzzy reasoning and fuzzy logic," Report EM/FS 10, University of Bristol, 1978.

[61] J.F. Baldwin, N.C.F. Guild and B.W. Pilsworth, "Improved logic for fuzzy controllers," Report EM/FS 5, University of Bristol, 1978.

[62] W.J.M. Kickert and H.R. van Nauta Lemke, "Application of a fuzzy controller in a warm water plant," Automatica 12: 301-308, 1976.

[63] E.H. Mamdani and S. Assilian, "An experiment in linguistic synthesis with a fuzzy logic controller," Int. J. Man-Machine Studies 7: 1-13, 1975.

[64] R.M. Tong, "Synthesis of fuzzy models for industrial processes--some recent results," Int. J. General Systems 4: 143-162, 1978.

[65] M. Mamdani, "Applications of fuzzy set theory to control systems: a survey," in Fuzzy Automata and Decision Processes, M.M. Gupta, G.N. Saridis and B.R. Gaines, eds., Elsevier North-Holland, New York, 1977.

[66] M. Uragami, M. Mizumoto and K. Tanaka, "Fuzzy robot controls," J. Cybernetics 6: 39-64, 1976.

[67] L.A. Zadeh, "On the validity of Dempster's rule of combination of evidence," Memorandum M79/24, Electronics Research Laboratory, University of California, Berkeley, 1979.