4. Possibility, Probability and Fuzzy Set Theory

Since L. Zadeh proposed the concept of fuzzy set in 1965 the relationships between probability theory and fuzzy set theory have been discussed. Both theories seem to be similar in the sense that both are concerned with some type of uncertainty and both use the interval [0, 1] for their measures as the range of their respective functions. (At least as long as one considers normalized fuzzy sets only!)

The comparison between probability theory and fuzzy set theory is difficult primarily for two reasons:

1. The comparison could be made on very different levels, that is, mathematically, semantically, linguistically, and so on.

2. Fuzzy set theory is not or is no longer uniquely defined mathematical structure, such as Boolean algebra or dual logic. It is rather a very general family of theories (consider, for instance, all the possible operations have been discussed in the previous 3 chapters, or the different types of membership functions). In this respect, fuzzy set theory could rather be compared with the different existing theories of multivalued logic. Further, there does not yet exist and probably never will exist, a unique context-independent definition of what fuzziness means. On the other hand, neither is probability theory uniquelly defined. There are different definitions and different linguistic appearances of "probability".

In recent years some specific interpretations of fuzzy set theory have been suggested. One of them, possibility theory, used to correspond, roughly speaking, to the min-max version of fuzzy set theory, that is, to fuzzy set theory in which the intersection is modeled by the min-operator and the union by the max-operator. This interpretation of possibility theory, however, is no longer correct. Rather it has been developed into a well-founded and comprehensive theory.

We shall first describe the essentials of possibility theory and then compare with other theories of uncertainty.

4.1. Possibility Theory

4.1.1. Fuzzy Sets and Possibility Distributions

Possibility theory focuses primarily on imprecision, which is intrinsic in natural languages and is assumed to be rather "possibilistic" than probabilistic. Therefore the term variable is very often used in a more linguistic sense than in strictly mathematical one. This is one reason why the terminology and the symbolism of possibility theory differs in some respects from that of fuzzy set theory. In order to facilitate the study of possibility theory, we will therefore use the common possibilistic terminology but always show the correspondence to fuzzy set theory.

Suppose, for instance, we want to consider the proposition "*X* is *F*", where *X* is the name of an object, a variable, or a proposition, and *F* is a fuzzy set. For instance, in "*X* is a small integer", *X* is the name of a variable. In "Peter is young", Peter is the name of an object. F (i.e., "small integer" or "young") is a fuzzy set characterized by its membership function μ_F.

One of the central concepts of possibility theory is that of a possibility distribution (as opposed to a probability distribution). In order to define a possibility distribution, it is convenient first to introduce the notation of fuzzy restriction. To visualize a fuzzy restriction the reader should imagine an elastic suitcase which acts on the possible volume of its contents as a constraint. For a hardcover suitcase, the volume is a crisp number. For a soft valise, the volume of its contents depends to a certain degree on the strength that is used to stretch it. The variable in this case would be the volume of the valise; the values this variable (*X*) can assume different values of *u* is expressed by $\mu_F(u)$. Zadeh [] defines these relationships as follows.

Definition 4-1 (fuzzy restriction): Let *F* be a fuzzy set of the universe *U* characterized by a membership function $\mu_F(u)$. *F* is a fuzzy restriction on the variable *X* if *F* acts as an elastic constraint on the values that may be assigned to *X*, in the sense that the assignment of the values *u* to *X* has the form

$$X = u: \mu_F(u).$$

$\mu_F(u)$ is the degree to which the constraint represented by *F* is satisfied when *u* is assigned to *X*. Equivalently, this implies that $1 - \mu_F(u)$ is the degree to which the constraint has to be stretched in order to allow the assignment of the values *u* to the variable *X*.

Whether a fuzzy set can be considered as a fuzzy restriction or not obviously depends on its interpretation: This is only the case if it acts as a constraint on the values of a variable, which might take the form of a linguistic term or a classical variable.

Definition 4-2 (relational assignment equation): Let *R*(*X*) be a fuzzy restriction associated with *X* such as defined in definition 4-1. Then *R*(*X*) = *F* is called a relational assignment equation which assigns the fuzzy set *F* to the fuzzy restriction *R*(*X*).

Let us now assume that *A*(*X*) is an implied attribute of the variable *X*. For instance, *A*(*X*) = "age of Jim" and *F* is the fuzzy set "young". The proposition "Jim is young" (or better "the age of Jim is young") can then be expressed as *R*(*A*(*X*)) = *F*.

Example 4-1

Let *p* be the proposition "Peter is young" in which young is a fuzzy set of the universe *U* = [0, 100] chracterized by the membership function

$$\mu_{young}(u) = S(u; 20, 30, 40)$$

where *u* is the numerical age and the *S*-function is defined by

$$S(u;\ \alpha,\ \beta,\ \gamma) = \begin{cases} 0 & \text{for } u \le \alpha \\ 2\left(\frac{u-\alpha}{\gamma-\alpha}\right)^2 & \text{for } \alpha \le u \le \beta \\ 1 - 2\left(\frac{u-\alpha}{\gamma-\alpha}\right)^2 & \text{for } \beta \le u \le \gamma \\ 1 & \text{for } u \ge \gamma \end{cases}$$

In this case, the implied attribute $A(X)$ is Age(Peter) and the translation of "Peter is young" has the form

$$\text{Peter is young} \rightarrow R(\text{Age (Peter)}) = \text{young}$$

Zadeh related the concept of a fuzzy restriction to that of a possibility distribution as follows.

> {"Consider a numerical age, say $u = 28$, whose grade of membership in the fuzzy set "young" is approximately 0.7. First we interpret 0.7 as the degree of compatibility of 28 with the concept labelled young. Then we postulate that the proposition "Peter is young" converts the meaning of 0.7 from the degree of compatibility of 28 with young to the degree of possibility that Peter is 28 given the proposition "Peter is young". In short, the compatibility of a value of u given "Peter is young".}[]

The concept of possibility distribution can now be defined as follows:

Definition 4-3 (possibility distribution): Let F be a fuzzy set in a universe of discourse U which is characterized by its membership function $\mu_F(u)$, which is interpreted as the compatibility of $u \in U$ with the concept labelled F.

Let X be a variable taking values in U and F act as a fuzzy restriction, $R(X)$, associated with X. Then the proposition "X is F", which translates into $R(X) = F$ associates a possibility distribution, π_x, with X which is postulated to be equal to $R(X)$.

The possibility distribution function, $\pi_x(u)$, characterizing the possibility distribution π_x is defined to be numerically equal to the membership function $\mu_F(u)$ of F, that is,

$$\pi_x := \mu_F$$

Where the symbol := stands for "denotes" or "is defined to be". In order to stay in line with the common symbol of possibility theory we will denote a possibility distribution with π_x.

Example 4-2

Let U be the universe of integers and F be the fuzzy set of small integers defined by

$$F = \{(1, 1), (2, 1), (3, 0.8), (4, 0.6), (5, 0.4), (6, 0.2)\}$$

Then the proposition "X is a small integer" associates with X the possibility distribution

$$\pi_x = F$$

in which a term such as (3, 0.8) signifies that the possibility that x is 3, given that x is a small integer, is 0.8.

Even though definition 4-3 does not assert that our intuition of what we mean by possibility agrees with min-max fuzzy set theory, it might help to realize their common origin. It might also make more obvious the difference betweeen possibility distribution and probability distribution.

Zadeh[] illustrates this difference by a simple but impressive example:

Example 4-3:

Consider the statement ″Billy ate X eggs for breakfast". X = {1, 2, ...}. A possibility distribution as well as a probability distribution may be associated with X. The possibility distribution $\pi_x(u)$ can be interpreted as the degree of ease with which Billy can eat u eggs while the probability distribution might have been determined by observing at breakfast for 100 days. The values of $\pi_x(u)$ and $P_x(u)$ might be as shown in the following table:

u	1	2	3	4	5	6	7	8
$\pi_x(u)$	1	1	1	1	0.8	0.6	0.4	0.2
$P_x(u)$	0.1	0.8	0.1	0	0	0	0	0

We observe that a high degree of possibility does not imply a high degree of probability. If however, an event is not possible, it is also improbable. Thus, in way the possibility is an upper bound for the probability.

This principle is not intended as a crisp principle, from which exact probabilities or possibilities can be computed but rather as a heuristic principle, expressing the principle relationship between possibilities and probabilities.

4.1.2. Possibility and Necessity Measures

In the previous chapter a possibility measure was alredy defined (see definition 3-5) for the case that A is a crisp set. If A is a fuzzy set a more general definition of possibility measure has to be given[].

Definition 4-4 (possibility measure on a fuzzy set): Let A be a fuzzy set in the universe of discourse U and π_x a possibility distribution associated with a variable X which takes values in U. The possibility measure, $\pi(A)$, of A is defined by

$$\text{poss}(X \text{ is } A) = \pi(A) := \sup_{u \in U} \min(\mu_A(u), \pi_x(u)) \quad \text{Kapcsos zarojelek!!!)}$$

Example 4-4:

Let us consider the possibility distribution induced by the proposition "X is a small integer" (see examle 4-2)

$$\pi_x = \{(1, 1), (2, 1), (3, 0.8), (4, 0.6), (5, 0.4), (6, 0.2)\}$$

and the crisp set $A = \{3, 4, 5\}$.

The possibility measure $\pi(A)$ is then

$$\pi(A) = \max \{0.8, 0.6, 0.4\} = 0.8.$$

If A, on the other hand, is assumed to be fuzzy set "integers which are not small", defined as

$$A = \{(3, 0.2), (4, 0.4), (5, 0.6), (6, 0.8), (7, 1),...\},$$

then the possibility measure of "X is not a small integer" is

$$\text{poss}(X \text{ is not a small integer}) = \max \{0.2, 0.4, 0.4, 0.2\} = 0.4.$$

Fuzzy measures as defined in definition 3-1 express the degree to which a certain subset of a universe, Ω, or an event is possible. Hence, we have

$$g(0) = 0 \quad \text{and} \quad g(\Omega) = 1.$$

As a consequence of condition m2 of definition 3-1, that is,

$$A \subseteq B \Rightarrow g(A) \leq g(B)$$

we have

$$g(A \cup B) \geq \max (g(A), g(B)) \quad \text{and}$$

$$g(A \cap B) \leq \min (g(A), g(B)) \quad \text{for } A, B \subseteq \Omega.$$

Possibility measures are defined for the limiting cases:

$$\pi(A \cup B) = \max (\pi(A), \pi(B))$$

$$\pi(A \cap B) = \min (\pi(A), \pi(B)).$$

If $\overline{A}$ is the complement of A in Ω, then

$$\pi(A \cup \overline{A}) = \max (\pi(A), \pi(\overline{A})) = 1$$

which expresses the fact that either A or $\overline{A}$ is completely possible.

In possibility theory another additional measure is defined, which uses the conjunctive relationship and, in a sense, is dual to the possibility measure:

$$N(A \cap B) = \min (N(A), N(B)).$$

N is called the *necessity measure*. $N(A) = 1$ indicates that A is necessarily true (A is sure). The dual relationship of possibility and necessity requires that

$$\pi(A) = 1 - N(\overline{A})) \qquad \forall A \subseteq \Omega.$$

Necessity measures satisfy the condition

$$\min (N(A), N(\overline{A})) = 0$$

The reationships between possibility measures and necessity measures satisfy also the following conditions[]:

$$\pi(A) \geq N(\overline{A}), \qquad \forall A \subseteq \Omega$$

$$N(A) \geq 0 \Rightarrow \pi(A) = 1$$

$$\pi(A) \leq 1 \Rightarrow N(A) = 0.$$

Here Ω is always assumed to be finite.

4.2. Probability of Fuzzy Events

By now it should have become clear that possibility is not substitute for probability but rather another kind of uncertainty. Let us now assume that an event is not crisply defined expect by a possibility distribution (a fuzzy set) and that we are in classical situation of stochastic uncertainty, that is, that the happening of this (fuzzily described) event is not certain and that we want to express the probability of its happening. Two views on this can be adopted: Either this probability should be a scalar (measure) or this probability can be considered as a fuzzy set also. We shall consider both views briefly.

4.2.1. Probability of a Fuzzy Event as a Scalar

In classical probability theory an event, A, is a member of an α-field a, of subsets of a sample space Ω. A probability measure P is a normalized measure over a measurable space (Ω, a) that is, P is real-valued function which assigns to every A in a a probability, $P(A)$ such that

p1: $P(A) = 0 \qquad A \in a$

p2: $P(\Omega) = 1$

p3: If $A_i \in a, \quad i \in I \subseteq \mathrm{N}, \quad$ pairwise disjoint then

$$P(\bigcup_{i \in I} \mathrm{A}_i) = \sum_{i \in I} P(\mathrm{A}_i)$$

As well as, obviously, the above expression can be interpreted for continuous case. Let Ω is, for instance a euclidean n-space and a the σ-field of Borel-sets[1] in R^n then

the probability of A can be expressed as

[1] The Borel-set $\mathbf{B} \subset P(U)$ is a family of subsets of U such that:

b1. $0 \in \mathbf{B}$ and $U \in \mathbf{B}$;

b2. iIf $A \in$ B, then $\overline{A} \in$ B;

b3. B is closed under the operation of set union, that is, if $A \in$ B and $B \in$ B, then also $A \cup \mathrm{B} \in \mathbf{B}$

$$P(A) = \int_A dP$$

If $\mu_A(x)$ denotes the characteristic function of a crisp set of A and $E_p(\mu_A)$ the expectation of $\mu_A(x)$ then

$$P(A) = \int_{R^n}^{\mu_A} (x)dP = E_P(\mu_A)$$

If $\mu_A(x)$ does not denote the characteristic function of a crisp set but rather the membership function of a fuzzy set the basic definition of the probability af A should not change.

Definition 4-5 (probability of a fuzzy event I.): Let ($\mathbf{R}^n$, a, P) be a probability space in which a is the σ-field of Borel-sets in $\mathbf{R}^n$ and P is a probability measure over $\mathbf{R}^n$. Then a *fuzzy event* in $\mathbf{R}^n$ is a fuzzy set F in $\mathbf{R}^n$ whose membership function $\mu_F(x)$ is Borel-measurable.

The probability of a fuzzy event F is then defined by the integral:

$$P(F) = \int_{R^n}^{\mu_F} (x)dP = E_P(\mu_F).$$

4.2.2. Probability of a Fuzzy Event as a Fuzzy Set

In the following we shall consider sets with a finite number of elements. Let us assume that there exists a probability measure P defined on the set of all crisp subsets of (the universe) X the Borel set. $P(x_i)$ shall denote the probability of element $x_i \in X$.

Let $A(x) = \{(x, \mu_A(x)| \ x \in X)\}$ be a fuzzy set representing a fuzzy event. The degree of membership of element $x_i \in A$ is denoted by $\mu_A(x_i)$. α-level sets or α-cuts such as alredy defined in definition 1-4 shall be denoted by A_α.

Yager[] suggests that it is quite natural to define the probability of an α-level set as

$$P(A_\alpha) = \sum_{x \in A_\alpha} \mathrm{P}(x)$$

On the basis of this the probability of a fuzzy event can be defined as follows.

Definition 4-6 (probability of fuzzy event II.): Let A_α be the α-level set of a fuzzy set A representing a fuzzy event. Then the probability of a fuzzy event can be defined as

$$P_Y(A) = \{(P(A_\alpha), \alpha)| \ \alpha \in [0, 1])\}$$

with the interpretation "the probability of at least an α degree of satisfaction to the condition A."

The subscript Y of P_Y indicates that P_Y is a definition of probability due to Yager which differs from Zadeh's definition which is denoted by P. It should be very clear that Yager considers α, which is used as the degree of membership of the probabilities $P(A_\alpha)$ in the fuzzy set $P_Y(A)$, as a kind of significance level for the probability of a fuzzy event.

Yager also suggests another definition for the probability of a fuzzy event, which is derived as follows:

Definition 4-7 (the truth of the proposition): The truth of the proposition "the probability A is at leas w" is defined as the fuzzy set $P^*(A)$ with the membership function

$$P^*(A)(w) = \sup_{\alpha}\{\alpha |\ P(A_\alpha) \geq w\}, \qquad w \in [0, 1]$$

We should realize, that now the "indicator" of significance of the probability measure is w and no longer α. We should also be aware of the fact that we have used Yager's terminology denoting the values of the membership function by $P^*(A)(w)$. This will facilitate reading Yager`s work[].

If we denote the complement of A by $\overline{A} = \{(x, 1-\mu_A(x))|\ x \in X\}$ and the α-level sets of $\overline{A}$ by $\overline{A}_\alpha$ then $P^*(\overline{A})(w) = \sup_{\alpha}\{\alpha |\ P(\overline{A}) \geq w\}, \quad w \in [0, 1]$ can be interpreted as the truth of the proposition the "probability of A is at least w".

Let us define $\overline{P}^*(A) = 1\text{-}P^*(\overline{A})$. If $\overline{P}^*(A)(w)$ is interpreted as the truth of the proposition "probability of A is at most w", then we can argue as follows: The "and" combination of the "probability of A is at least w" and the "probability of A is at most w" might be considered as "probability of A is exactly w". If $P^*(A)$ and $\overline{P}^*(A)$ are considered as possibility distributions then their conjunction is their intersection (modeled by applying the min-operator to the respective membership functions). Hence the following definition:

Definition 4-9 ("the probability of A is exactly w"): Let $P^*(\overline{A})$ and $\overline{P}^*(A)$ be defined as above. The possibility distribution associated with the proposition "probability of A is exactly w" can be defined as

$$\overline{P}(A)(w) = \min\{P^*(A)(w),\ \overline{P}^*(A)(w)\}$$

Example 4-5

Let $A\{(x_1, 1), (x_2, 0.7), (x_3, 0.6), (x_4, 0.2)\}$ be a fuzzy event with the probability defined for the generic elements: $P_1 = 0.1$, $P_2 = 0.4$, $P_3 = 0.3$, $P_4 = 0.2$; $p\{x_2\}$ is 0.4, where the element x_2 belongs to the fuzzy event A with a degree of 0.7.

First we compute $P^*(A)$. We start by determining the α-level sets A_α for all $\alpha \in [0, 1]$. Then we compute the probability of the crisp events A_α and give the intervals of w for which $P(A_\alpha) \geq w$. We finally obtain $P^*(A)$ as the respective supremum of α.

The computing is summarized in the following table:

α	A_α	$P(A)$	w	$P^*(A)$

[0, 0.2]	$\{x_1, x_2, x_3, x_4\}$	1	[0.8, 1]	0.2
[0.2 0.6]	$\{x_1, x_2, x_3\}$	0.8	[0.5 0.8]	0.6
[0.6, 0.7]	$\{x_1, x_2\}$	0.5	[0.1, 0.5]	0.7
[0.7, 1]	$\{x_1\}$	0.1	[0, 0.1]	1.0

Analogously we obtain for $\overline{P}^*(A) = 1\text{-}P^*(\overline{A})$:

α	$\overline{A}_\alpha$	$P(\overline{A})_\alpha$	w	$\overline{P}^*(\overline{A})$	$\overline{P}^*(A)$
0	$\{x_1, x_2, x_3, x_4\}$	1	[0.9, 1]	0	1
[0, 0.3]	$\{x_1, x_2, x_3\}$	0.9	[0.5 0.9]	0.3	0.7
[0.3, 0.4]	$\{x_1, x_2\}$	0.5	[0.2, 0.5]	0.4	0.6
[0.4, 8]	$\{x_1\}$	0.2	[0, 0.2]	0.8	0.2
[0.8, 1]	0	0	0	1.0	0

The probability $\overline{P}(A)$ of the fuzzy event A is now determined by the intersection of the fuzzy sets $P^*(A)$ and $\overline{P}^*(A)$ modeled by the min-operator as in definition 4-7:

$$\overline{P}(A)(w) = \begin{cases} 0.0, & w = 0 \\ 0.2, & w \in [0.0,\ 0.2] \\ 0,6, & w \in [0.2,\ 0.8] \\ 0.2, & w \in [0.8,\ 1.0] \end{cases}$$

5. Fuzzy Logic and Approximate Reasoning

5.1. Linguistic Variables

"In retraining from precision in the face of overpowering complexity, it is natural to explore the use of what might be called *linguistic* variables, that is, variables whose values are not numbers but words or sentences in a natural or artifciial language.

The motivation for the use of words and sentences rather than numbers is that linguistic characterizations are, in gerneral, less specific than numerical ones." []

This quotion presents in a nutshell the motivation and justification for fuzzy logic and approximate reasoning. Another quotation might be added, which is much older. The philosopher B. Russel noted in his 1923 study:

"All traditional logic habitually assumes that precise symbols are being employed. It is therefore not applicable to this terrestrial life but only to an imagined celestal existece." []

One of the basic tools for fuzzy logic and approximate reasonong is the notion of a linguistic variable which in 1973 was was called a *variable of higher order* rather than a fuzzy variable and defined as follows []:

Definition 5-1 (linguistic variable): A linguistic variable is characterized by a quintuple

$$(x, T(x), U, G, M)$$

in which:

- x is the name of the variable;
- $T(x)$ (or simlpy T) denotes the *term set of* x, that is, the set of names of *linguistic values* of x, with each value being a fuzzy variable denoted generically by x and ranging over a
- *universe of discourse* U which is associated with the base variable u;
- G is a *syntactic rule* (which usually has the form of a grammar) for generating the name, X, of values of x;
- M is a *semantic rule* for associating with each X its meaning, $M(X)$ which is a fuzzy subset of the universe of discourse.

A particular X, that is a name generated by G, is called *term*. It should be noted that the variable u can also be vector-valued.

In order to facilitate the symbolism in what follows, some symbols will have two meanings wherever clarity allows: x will denote the name of the variable ("the label") and the generic name of its values. The same will be true for X, and $M(X)$.

Example 5-1

Let X be a linguistic variable with the label "Age" (i.e., the label of this variable is "Age" and the values of it will also be called "Age") with U = [0, 100]. Terms of this linguistic variable, which are again fuzzy sets, could be called "old", "young", "very old", and so on. The base-variable u is the age in years of life, $M(X)$ is the rule that assigns a meaning, that is, a fuzzy set, to the terms.

$$M(\text{old}) = \{(u, \mu_{\text{old}}(u)) \mid u \in [0, 100]\}$$

where

$$\mu_{\text{old}}(u) = \begin{cases} 1 & u \in [0,\ 50] \\ \left[1 + \left(\frac{u-50}{5}\right)^2\right]^{-1} & u \in [50,\ 100] \end{cases}$$

T(Age) will define the term set of the variable x, for instance, in the case,

$$T(\text{Age}) = \{\text{old, very old, not so old, more or less young, quite young, wery young}\}$$

where $G(X)$ is a rule which generates the (labels of) terms in the term set. Figure 5-1 sketches the above-mentioned relationships.

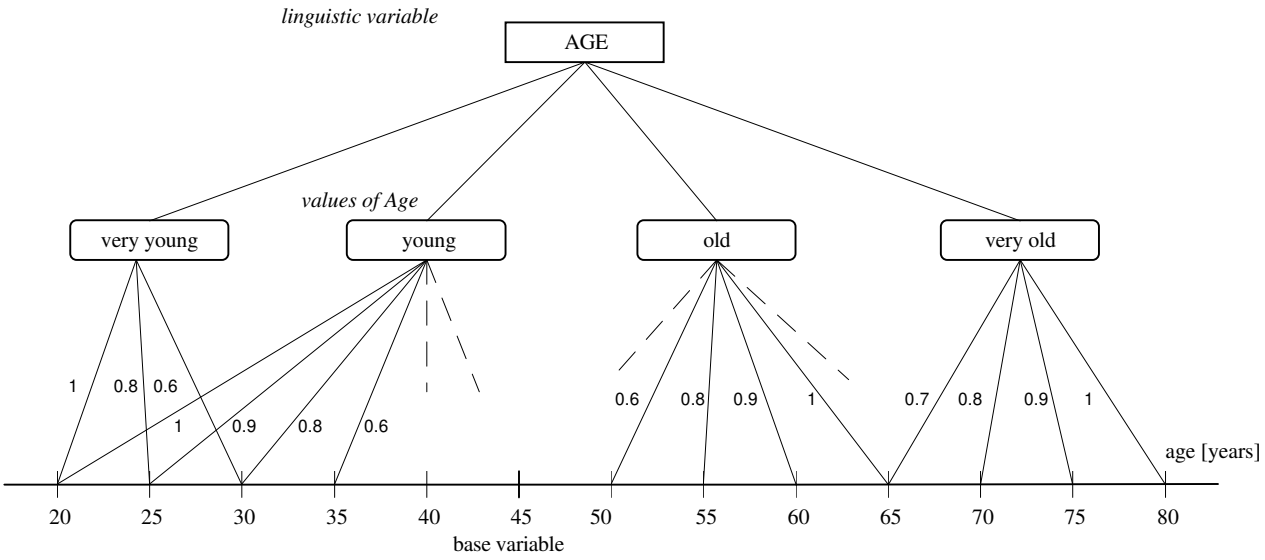

Figure 5-1. Linguistic variable "Age"

Two linguistic variables of particular interest in fuzzy logic and in (fuzzy) probability theory are two linguistic variables "Truth" and "Probability". The linguistic variable "Probability" is depicted in figure 5-2.

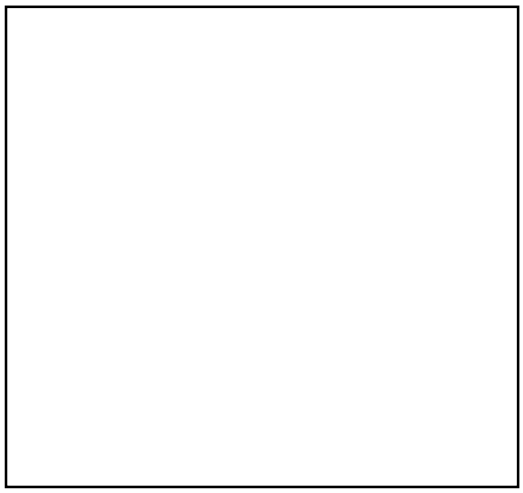

Figure 5-2. Linguistic variable "Probability".

The term set of linguistic variable "Truth" might be defined by different ways. Boldwin [] defines, for instance, some of the terms follows: (see figure 5-3)

$$\mu_{true}(v) = v$$

$$\mu_{very\ true}(v) = (\mu_{true}(v))^2$$

$$\mu_{fairly\ true}(v) = (\mu_{true}(v))^{1/2}$$

$$\mu_{very\ false}(v) = 1\text{-}\mu_{very\ true}(v)$$

$$\mu_{fairly\ false}(v) = 1\text{-}\mu_{fairly\ true}(v)$$

$$\mu_{false}(v) = 1\text{-}\mu_{true}(v)$$

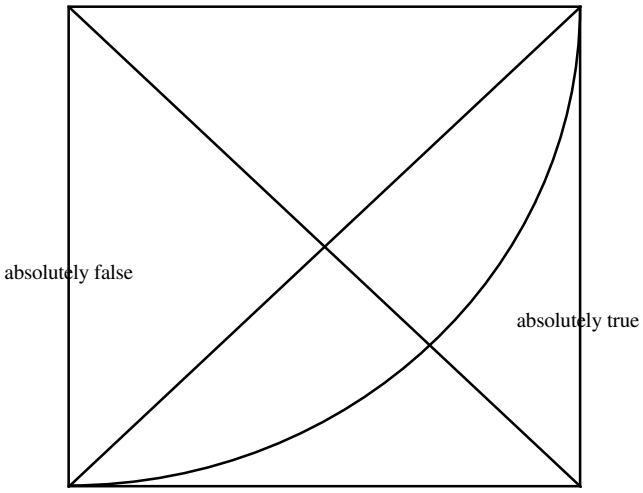

Figure 5-3. Linguistic variable "Truth"

Zadeh suggests for the term *true* the membership function

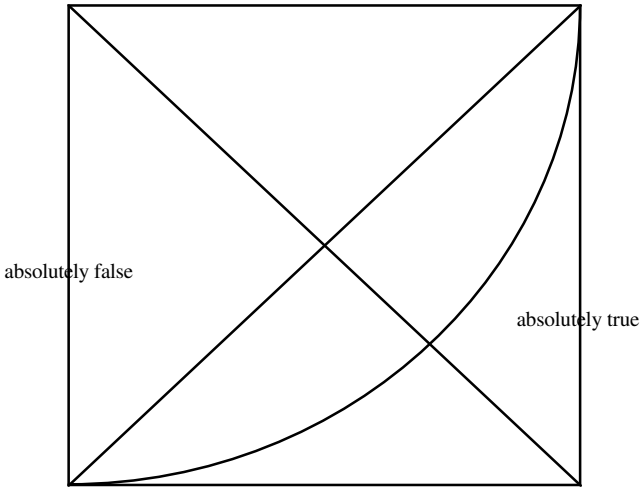

Figure 5-4. Terms "True" and "False"

$$\mu_{true}(v) = \begin{cases} 0 & \text{for } 0 \le v \le a \\ 2\left(\dfrac{v-a}{1-a}\right)^2 & \text{for } a \le v \le \dfrac{a+1}{2} \\ 1-2\left(\dfrac{v-1}{1-a}\right)^2 & \text{for } \dfrac{a+1}{2} \le v \le 1 \end{cases}$$

where $v = (1+a)/2$ a crossover point, and $a \in [0, 1]$ is a parameter that indicates the subjective judgment abaut the minimum value of v in order to consider a statement as "true" at all.

The membership function of "false" is considered as the mirror image of "true", that is

$$\mu_{false}(v) = \mu_{true}(1\text{-}v) \qquad 0 \le v \le 1$$

Figure 9-4 shows the above *true* and *false* terms.

Of course the membership functions of true and false, respectively, can also be chosen from the finite universe of truth values. The term set of the linguistic variable "Truth" is then defined as

T(Truth) = {true, not true, very true, not very true,...,false, not false, very false,...,not very true and not very false,...}

The fuzzy sets (possibility distribution) of those terms can essentially be determined from the term *true* or the term *false* by applying appropriately the above-mentioned modifiers (hedges).

Definition 5-2 (stuctured linguistic variable): A linguistic variable x is called structured if the term set $T(x)$ and the meaning $M(x)$ can be characterized algorithmically. For structured linguistic variable, $M(x)$ and $T(x)$ can be regarded as algorithms which generate the terms of the term set and associate meanings with them.

Before we illustrate this by an example we need to define what we mean by "hedge" or "modifier".

Definition 5-3 (linguistic hedge or a modifier): A linguistic hedge or a modifier is an operation that modifies the meaning of a term, more generally, of a fuzzy set. If A is a fuzzy set then the modifier m generates the (composite) term $B = m(A)$.

Matehematical models frequently used for modifiers are:

Comcentration: $\mu_{con(A)}(u) = (\mu_A(u))^2$

Dilation: $\mu_{dil(A)}(u) = (\mu_A(u))^{1/2}$

contrast intensification:
$$\mu_{int(A)}(u) = \begin{cases} 2(\mu_A(u))^2 & \text{for } \mu_A(u) \in [0,\ 0{,}5] \\ 1 - 2(1-\mu_A(u))^2 & \text{otherwise} \end{cases}$$

Generally the following linguistic hedges (modifiers) are associated with above-mentioned mathematical operators.

If A is a term (a fuzzy set) then

very	$A = \text{con}(A)$
more or less	$A = \text{dil}(A)$
plus	$A = A^{1.25}$
slightly	$A = \text{int}[\text{plus } A \text{ and not (very } A)]$

were "and" is interpreted possibilistically.

Example 5-2:

Let us reconsider from example 4-1 the linguistic variable "Age". The term set shall be assumed to be

$$T(\text{Age}) = \{\text{old, very old, very very old...}\}$$

The term set can now be generated recursively by using the following rule (algorithm):

$$T^{i+1} = \{\text{old}\} \cup \{\text{very } T^i\}$$

that is,

$T^0 = 0$

$T^1 = \{\text{old}\}$

$T^2 = \{\text{old, very old}\}$

$T^3 = \{\text{old, very old, very very old}\}$

For the semantic rule we only need to know the meaning of "old" and the meaning of the modifier "very" in order to determine the meaning of an arbitrary term of the term set. If one defines "very" as the concentration, then terms of the term set of the structured linguistic variable "Age" can be determined, given that the membership function of the term "old" is known.

Definition 5-4 (Boolean linguistic variable): A Boolean linguistic variable is a linguistic variable whose terms, X, are Boolean expressions in variables of the form X_p, $m(X_p)$ where X_p is a primari term and m is a modifier. $m(X_p)$ is a fuzzy set resulting from acting with m on X_p.

Example 5-3

Let "Age" be a Boolean linguistic variable with the term set

T(Age) = {young, not young, old, not old, very young, not young and not old young or old...}

Identifying "and" with intersection, "or" with the union, "not" with the complementation, and "very" with the contcentration we can derive the meaning of different terms of the term set as follows:

M(not young) = ¬young
M(not very young) = ¬(young)^2
M(young or old) = young ∪ old
...

Given the two Fuzzy sets (primary terms)

$$M(\text{young}) = \{(u, \mu_{\text{young}}(u)) \mid u \in [0, 100]\}$$

where

$$\mu_{\text{young}}(u) = \begin{cases} 1 & u \in [0,\ 25] \\ \left[1 + \left(\frac{u-25}{5}\right)^2\right]^{-1} & u \in [25,\ 100] \end{cases}$$

and

$$M(\text{old}) = \{(u, \mu_{\text{old}}(u)) \mid u \in [0, 100]\}$$

where

$$\mu_{\text{old}}(u) = \begin{cases} 0 & u \in [0,\ 50] \\ \left[1 + \left(\frac{u-50}{5}\right)^{-2}\right]^{-1} & u \in [50,\ 100] \end{cases}.$$

Then the membership function of the term "young or old" would, for instance, be

$$\mu_{\text{true}}(v) = \begin{cases} 0 & \text{for } 0 \le v \le a \\ 2\left(\frac{v-a}{1-a}\right)^2 & \text{for } a \le v \le \frac{a+1}{2} \\ 1 - 2\left(\frac{v-a}{1-a}\right)^2 & \text{for } \frac{a+1}{2} \le v \le 1 \end{cases} \qquad \text{z139!!!!!}$$

5.2. Fuzzy Logic

5.2.1. Classical Logics Revisited

Logics a bases for reasoning can be distingushed essentially by their three topic-neutral (context-independent) items: truth values, vocubulary (operators), and reasoning procedure (tautologies, syllogisms).

In Boolean Logic, truth values can be 0 (false) or (1) (true) and by means of these truth values (operators) is defined via truth tables.

Let us consider two statements, *A*, and *B*, either of which can be true or false, that is, have the truth value 1 or 0. We can construct the following truth tables (where one column is one truth table of a Boolean logical operation):

A	*B*	$\wedge$	$\vee$	$x\vee$	$\Rightarrow$	$\Leftrightarrow$	?
1	1	1	1	0	1	1	1
1	0	0	1	1	0	0	1
0	1	0	1	1	1	0	0
0	0	0	0	0	1	1	0

Table 5.1. Truth tables for two statements

There are 16 truth tables, each defining an operator. Assigning meanings (words) to these operators is not difficult for the first 4 or 5 columns: the first obviously characterizes the "and", the second the "inclusive or", the third the "exclusive or", and the fourth and fifth the implication and the equivalence. We will have difficultie, hawever, interpreting the remaining nine columns in terms of our language. If we have three statements rather than two, this task of assigning meanings to truth tables becomes even more difficult.

So far it has been assumed that each statement, *A* and *B*, could clearly be classified as true or false. If this is no longer true then additional truth values, such as "undecided" or similar, can and have to be introduced, which leads to the many existing systems of multivalued logic. It is not difficult to see haw the above-mentioned problems of two-valued logic in "calling" truth tables or operators increase as we move to multivalued logic. For only two statements and three possible truth values there are alredy $3^{3^2} = 729$ truth tables! The uniqueness of interpretation of truth tables, which is so convenient in Boolean logic, disappears immediately because many truth tables in three-valued logic look very much alike.

The third topic-neutral item of logical systems is the reasoning procedure itself which generally bases on tautologies such as

modus ponens: $(A \wedge (A \Rightarrow B)) \Rightarrow B$

modus tollens: $((A \Rightarrow B) \wedge \neg B) \Rightarrow \neg A$

syllogism: $((A \Rightarrow B) \wedge (B \Rightarrow C)) \Rightarrow (B \Rightarrow C)$

contraposition: $(A \Rightarrow B) \Rightarrow (\neg B \Rightarrow \neg \mathrm{A})$

Let us consider the modus ponens which could be interpreted as: "If *A* is true and if the statement "If *A* is true then *B* is true" is also true then *B* is true.

The term *true* is used at different places and in two defferent senses: All but the last "true's" are material true`s, that is, they are taken as a matter of fact, while the last "true" is a topic-neutral true.

In Boolean logic, however, these "true's" are all terated the same way []. A distinction between a material and logical (necessary) truth is made in so called *extended logics*: Modal logic [] distuinguishes between necessery and possible truth, tense logic between statemens that were true in the past and those that will be true in the future. Epistemic logic deals with knowledge and belief and deontic logic with what ought to be done and what is permitted to be true. Modal logic, in particular, might be a very good basis for applying different measures and theories of uncertainty.

Another extension of Boolean logic is predicate calculus, which is a set theoretic logic using quantifiers and predicates in addition to the operators of Boolean logic.

Fuzzy logic is an extension of set theoretic multivalued logic in which the truth values are linguistic variables (or terms of the linguistic variable truth).

Since operators $\wedge$, $\vee$, $\neg$ $\Rightarrow$ in fuzzy logic are also defined by using truth tables, the extension principle can be applied to derive definitions of the operators. So far, possibility theory has primary been used in order to define operators in fuzzy logic, even though other operators have also been investigated [], and could also be used. In this book we will limit considerations to possibilistic interpretations of linguistic variables and we will also stick to the original proposals of Zadeh []. To the interested reader several supplemental study of alternative approaches can be found in the [].

Iv $v(\text{not}A)$ is a point in $V = [0, 1]$, representing the truth value of the proposition "u is A, then the truth value of not A is given by

$$v(\text{not}A) = 1 - v(A)$$

Definition 5-5 (the truth value of v*(notA)): If $v^*(A)$ is a normalized fuzzy set, $v^*(A) = \{(v_i, \mu_i) | \ i=1, ..., n, v_i \in [0, 1]\}$ then by applying the extension principle, the truth value of $v^*(\text{not}A)$ is defined as

$$v^*(\text{notA}) = \{(1 - v_i, \mu_i) | \ i=1, ..., n, v_i \in [0, 1]\}$$

In particular "false" is interpreted as "not true", that is

$$v^*(\text{false}) = \{(1 - v_i, \mu_i) | \ i=1, ..., n, v_i \in [0, 1]\}$$

Examle 5-4

Let us consider the terms *true* and *false*, respectively, defined as the following possibility distributions:

$v^*(true) = \{(0.5, 0.6), (0.6, 0.7), (0.7, 0.8), (0.8, 0.9), (0.9, 1), (1, 1)\}$

$v^*(false) = v(\text{not } true) = \{(0.5, 0.6), (0.4, 0.7), (0.3, 0.8), (0.2, 0.9), (0.1, 1), (0, 1)\}$

then

$v^*(\text{very } true) = \{(0.5, 0.36), (0.6, 0.49), (0.7, 0.64), (0.8, 0.81), (0.9, 1), (1, 1)\}$

$v^*(\text{very } false) = \{(0.5, 0.36), (0.4, 0.49), (0.3, 0.64), (0.2, 0.81), (0.1, 1), (0, 1)\}$.

It has alredy been mentioned that fuzzy logic is essentially considered as an application of possibility theory to logic. Hence the logical operators "and", and "or", and "not" are defined accordingly.

Definition 5-5 (the four logical basic operations): For numerical truth values $v(A)$ and $v(B)$ the logical operations *and*, and *or*, *not* and *implied* are defined as

$$v(A) \wedge v(B) = v(A \text{ and } B) = \{(v, \min\{\mu_A(v), \mu_B(v)\}\}$$
$$v(A) \vee v(B) = v(A \text{ or } B) = \{(v, \max\{\mu_A(v), \mu_B(v)\}\}$$
$$\neg vA = \{v, 1 - \mu_A(v)\}$$
$$v(A) \Rightarrow v(B) = v(A \Rightarrow B) = \neg v(A) \vee v(B) = \{(v, \max\{1 - \mu_A(v), \mu_B(v)\}\}$$

If

$$v^*(A) = \{(v_i, \alpha_i), \quad v_i, \alpha \in [0, 1] \quad i = 1,...,n$$
$$v^*(B) = \{(v_j, \beta_j), \quad v_j, \beta_i \in [0, 1] \quad j = 1,...,m$$

then

$$v^*(A) \wedge v^*(B) = v^*(A \text{ and } B) = \{(u = \min\{v_i, v_j\}, \max_u \min\{\alpha_i, \beta_j\} | \ i = 1, ..., n; \ j = 1, ..., m.$$

The other operators defined accordingly.

Example 5-5:

Let $v^*(A)$ = true {(0.5, 0.6), (0.6, 0.7), (0.7, 0.8), (0.8, 0.9), (0.9, 1), (1, 1}

then

$\neg v^*A$ = {(0, 1), (0.1, 1), (0.2, 1), (0.3, 1), (0.4, 1), (0.5, 0.4), (0.6, 0.3), (0.7, 0.2), (0.8, 0.1)}.

5.2.2. Truth Tables and Linguistic Approximation

As mentioned at the beginning of this section binary connectives (operators) in classical two- and many-valued logics are normally defined by the tabulation of truth values in truth tables. In fuzzy logic the number of truth values, in general, infinite. Hence tabulation of the truth values for operators is not possible. We can, however, tabulate truth values, that is, terms of the linguistic variable "Truth" for a finite number of terms, such as true, not true, very true, etc.

Zadeh [] suggests truth tables for the determination of truth values for operators using a four-valued logic including the truth values true, false, undecided, and unknown. "Unknown" is then interpreted as "true or falese" (T+F).

Extending the normal Boolean logic with truth values true (1) and false (0) to a (fuzzy) three-valued logic with a universe of truth values being two-valued (true an false) we obtain the following truth tables where the first columns contain the truth values for a statement *A* and the first rows those for a statement *B* [].

$\wedge$	T	F	T+F
T	T	F	T+F
F	F	F	F
T+F	T+F	F	T+F

"and "

$\vee$	T	F	T+F

T	T	T	T
F	T	F	T+F
T+F	T	T+F	T+F

"or"

	¬
T	F

F	T
T+F	T+F

"not"

Table 5-2 Truth tables for three-valued logic

If the number of truth values (terms of the linguistic variable truth) increases one can still "tabulate" the truth table for operators by using definition 4-6 as follows:

Let us assume, that *i*th row of the table represents "not true" and the *j*th column "more or less true". The (*i*, *j*)th entry in the truth table for "and" would than contain the entry for "not true $\wedge$ more or less true". The resulting fuzzy set would, however, most likely not correspond to any fuzzy set assigned to the terms of the term set of "truth". In this case one could try to find the fuzzy set of the term which is most similar to the fuzzy set resulting from the computations. Such a term would then be called *linguistic approximation*. This is an analogy to statistics, where empirical distribution functions are often approximated by well-known standard distribution functions.

Example 5-6:

Let $V = \{0, 0.1, 0.2, ..., 1\}$ the universe,
true = {(0.8, 0.9), (0.9, 1), (1, 1)},
more or less true = {(0.6, 0.2), (0.7, 0.4), (0.8, 0.7), (0.9, 1), (1, 1)}, and
almost true = {(0.8, 0.9), (0.9, 1), (1, 0.8)}.

Let "more or less true" be the *i*th row and "almost true" the *j*th column of the truth table "or". Then "more or less true $\vee$ almost true" is the (*i*, *j*)th entry in the table:

$$\text{more or less true} \vee \text{almost true}$$
$$=$$
$$\{(0.6, 0.2), (0.7, 0.4), (0.8, 0.7), (0.9, 1), (1, 1)\}$$
$$\vee$$
$$\{(0.8, 0.9), (0.9, 1), (1, 0.8)\}.$$

Now we can approximate the right-hand side of this equation by

$$\text{true} = \{(0.8, 0.9), (0.9, 1), (1, 1)\}$$

This yields

$$\text{"more or less true} \vee \text{almost true"} \approx \text{"true"}.$$

Baldwin [] suggests another version of fuzzy logic – fuzzy truth tables, and their determination:

The truth values on which he bases his suggestions were shown graphically in figure 9-3. They were defined as

$$\text{true} = \{(v, \mu_{true}(v) = v) | \; v \in [0, 1]\}$$
$$\text{false} = \{(v, \mu_{false}(v) = 1\text{-}\mu_{true}(v)) | \; v \in [0, 1]\}$$

very true $= \{(v, \mu_{true}(v)^2 | \ v \in [0, 1]\}$

fairly true $= \{(v, \mu_{true}(v)^{1/2} | \ v \in [0, 1]\}$

undecided $= \{(v, 1 | \ v \in [0, 1]\}$

Very false and fairly false were defined correspondingly, and

absolutely true $= \{(v, \mu_{at}(v)) | \ v \in [0, 1]\}$ where $\mu_{at}(v) = \begin{cases} 1 & \text{for } v = 1 \\ 0 & \text{otherwise} \end{cases}$

absolutely false $= \{(v, \mu_{at}(v)) | \ v \in [0, 1]\}$ where $\mu_{af}(v) = \begin{cases} 1 & \text{for } v = 0 \\ 0 & \text{otherwise} \end{cases}$.

Hence

(very)^{k}true $\rightarrow$ absolutely true as $k \rightarrow \infty$

(very)^{k}false $\rightarrow$ absolutely false as $k \rightarrow \infty$

(fairly)^{k}true $\rightarrow$ undecided as $k \rightarrow \infty$

(fairly)^{k}false $\rightarrow$ undecided as $k \rightarrow \infty$.

Using figure 5-3 and the interpretations of "and" and "or" as minimum and maximum, respectively, the following truth table results:

$v(P)$	$v(Q)$	$v(P$ and $Q)$	$v(P$ or $Q)$
false	false	false	false
true	false	false	true
true	true	true	true
undecided	false	false	undecided
undecided	true	undecided	true
undecided	undecided	undecided	undecided
true	very true	true	very true
true	fairly true	fairly true	true

Some more considerations and assumptions are needed to derive the truth table for the implication. Baldwin considers his fuzzy logic to rest on two pillars: the denumerably infinite multivalued logic system of Lukasewicz logic and fuzzy theory.

"Implication statements are treated by composition of fuzzy truth value restrictions with a Lukasewicz logic implication relation on a fuzzy truth space. Set theoretic considerations are used to obtain fuzzy truth value restrictions from conditional fuzzy linguistic statements using an inverse truth functional modification procedure. Finally true functions modification is used to obtain the final conclusion" [].

5.3. Approximate Reasoning

We alredy mentioned, that in traditional logic the main tools of reasonong are tautologies such as, for instance, the modus ponens, that is $(A \wedge (A \Rightarrow B)) \Rightarrow B$ or

Premise	A is true
Implication	If A then B
Conclusion	B is true.

A and B are statements or propositions (crisply defined) and the B in the conditional statement is identical to the B of the conclusion.

On the basis of what has been said in section 4.1. and 4.2., two quite obvios generalizations of the modus ponens are:

1. To allow statements that are characterized by fuzzy sets.
2. To relax (slightly) the identity of the "B's" in the implication and the conclusions.

This version of modus ponens is then called "generalized modus ponens" [].

Example 5-7

Let A, A', B, B' be fuzzy statements, then the generalized modus ponens reads

Premise:	x is A'
Implication:	If x is A then y is B
Conclusion:	y is B'.

For instance []:

Premise:	This apple is very red
Implication:	If an apple is red then the apple is ripe.
Conclusion:	This apple is very ripe.

Zadeh suggested the compositional rule of inference for the above-mentioned of fuzzy conditional inference []. In the meantime other other authors have suggested different methods and investigated also the modus tollens, syllogism, and contraposition (see []). In the frame of this textbook, however, we shall restrict considerations to Zadeh's compositional rule of inference.

Definition 5-7 (compositional rule of inference): Let $R(x)$, $R(x, y)$ and $R(y)$, $x \in X$, $y \in Y$, be fuzzy relations in X, $X \times Y$, and Y respectively, (where X and Y are crisp sets,) which act as fuzzy restrictions on x, (x, y), and y, respectively. Let A and B denote particular fuzzy sets in X and $X \times Y$. Then the compositional rule of inference asserts, that the solution of the relational assigment equations (see definition 4-1) $R(x) = A$ and $R(x, y) = B$ is given by $R(y) = A \circ B$, where $A \circ B$ is the composition of A and B.

Example 5-8

Let the universe be $X = \{1, 2, 3, 4\}$.

A = little = {(1, 1), (2, 0.6), (3, 0.2), (4, 0)}

R = "approximately equal" be a fuzzy relation defined by

R:	1	2	3	4
1	1	0.5	0	0
2	0.5	1	0.5	0
3	0	0.5	1	0.5
4	0	0	0.5	1

For the formal inference denote

$$R(x) = A,\ R(x, y) = B, \text{ and } R(y) = A \lozenge B$$

Applying the max-min composition for computing $R(y) = A \circ B$ yields

$$R(y) = \max_x \min \{\mu_A(x), \mu_R(x, y)\} = \{(1, 1), (2, 0.6), (3, 0.5), (4, 0.2)\}$$

A possible interpretation of the inference may be the following:

Premise: x is little

Implication: x and y are approximately equal

Conclusion: y is more or less little.

There are several direct applications of the approximate reasoning, for instance, the fuzzy algorithm and fuzzy languages, however, it is not our aim to deal with them. For more information of the above-mentioned fields see [].

5. 5. Selected Methods of Determination Memembership Functions

5.5.1. Some comments

The problem of obtaining the values of membership function, or at least their estimation, is of interest in the further alpplication of fuzzy set techniques. Following a stream of fuzzy sets as formulae describing vague notions, it is difficult to see how this problem could have a straightforward solution. Firstly, fuzzy sets model a subjective category; therefore, their membership functions can be evalueted in a subjective fashion. We should also bear in mind that notations or categories modelled by fuzzy sets have a local character, that is, to say the meaning of certain category relies upon the context (situation) in which its application is planned. For instance, when talking abaut a concept *large steady state error* in a certain community (e.g. in control of a certain

industrial system) and after establishing the relevant membership function, it is not possible to play with the same membership function in a completely different community; usually at least some scaling will be necessary. From the measurement theory point of view, it is not clear which type of scale shoul be used for estimation of the membership function. Thole an Zimmerman [], for instance, used an absolute scale, in Saaty's approach [], a ratio scale is suggested, while Goguen [] argues that no stronger scale than an ordinal one may be obtained. Leaving this questions open we focus our attention on the discussion of some methods that may be used in engineering practice when the membership function has to be estimated.

A straightforward method for estimating the values of the membership function, whose roots form the example cited by Borel, can be compactly stated as follows.

Consider a group of researchers (experts) involved with the same area of investigation. They are asked to answer a question having a format:

Can x_0 be viewed as compatible with the concept represented by the fuzzy set A?

where x_0 is a fixed element of the universe of discourse. The answer is "yes" or "no". Then, counting the fraction of positive ("yes") response $n(x_0)$ found in the total number of responses, we get the value of the membership funcion of this element of the universe of discourse x_0

$$\mu_A(x_0) = \frac{n(x_0)}{N}$$

with N being the total number of responses dealt with x_0. Moreover, following this statistical approach, one can also determine a confidence interval at a prespecified level of probability. Denote the obtained bounds by x_0 $\mu_A^-(x_0)$ and $\mu_A^+(x_0)$, respectively, Thus, in fact an appropriate interpretation of the result enables us to evaluate the precision of the membership obtained. This method for determining the membership function is mentioned in [] and [], however, no method for expressing its precision is provided.

5.5.2. "Pairwise comparison" method

The membersip function expressed on a ratio scale can be conveniently estimated by a pairwise comparison as proposed in the []. As usual, μ_i denotes the degree to which the ith element of the universe of discourse fulfils the fuzzy notion **A**. Take, now, the ratios μ_i/μ_j for all $i, j = 1, 2, ..., n$ and arrange them in the form of square matrix $[\mu_i/\mu_j]$. Multiplying it by a vector $\mu_i = [\mu_1, \mu_2, ..., \mu_n]$, we get a system of equations

$$[\mu_i/\mu_j]\mu = n\mu$$

i.e.

$$([\mu_i/\mu_j]\mu - n)\mu = 0.$$

Thus, 'n' forms the largest eigenvalue of the above eigenvalue problem (the remainder are equal to zero), and μ is the corresponding eigenvector.

An experiment performed using the idea that stems from the above finding relies on obtaining estimates of these grades by making pairwise comparison of the elements of the universe of

discourse. Assuming a certain scale in which this comparison is realized (usually consisting of abaut 7 grades), a researcher is asked to evaluate the *i*th element of the universe of discourse with respect to the *j*th one. The more preferable the *i*th element is with respect to the *j*th one, the closer the value of the (*i*, *j*)th element of the matrix **A** to the highest value in the scale established. Conversely, if the *i*th element is completely rejected with respect to the *j*th one, the estimated value of **A** is equal to the reciprocal of the highest value of the scale. Also, for each element lying onthe diagonal of **A**, we write 1.0. Summarising after performing all pairwise comparisons, a matrix **A** is given. Obviously, now, only an approximate equality might be expected, namely

$$a_{ij} = \mu_i/\mu_j$$

and the transitivity property is no longer preserved (i.e. $a_{ik}a_{kj} \neq a_{ij}$).

There are some numerical schemes useful for the calculation of the membership function on the basis of matrix **A**. One of them, [], takes the problem of minimisation of a sum of squares

$$\min_{\mu} \sum_{i=1}^{n}\sum_{j=1}^{n}(a_{ij}\cdot\mu_j - \mu_i)^2$$

subject to constraints

$$\sum_{i=1}^{n}\mu_i = 1.$$

Then, renormalising μ_i's in such a fashion that the maximal value of μ is equal ti 1, we get the membership function of a normal fuzzy set.

5.5.3. "Probabilistic characteristics" method

In the methods described so far, the membership function was calculated with the aid of a probability function obtained by the previous experiment. In [], a bijective mapping, turning a probability function into a membership function and vica versa, has been introduced. For the discrete probabilities $p_1, p_2, ..., p_n$, $\Sigma p_i = 1$, arranged in descending order $p_1 \geq p_2 \geq ... \geq p_n$, the values of the membership function $\mu_1, \mu_2, ..., \mu_n$ are computed by the formula

$$\mu_i = 1 - \sum_{j=1}^{i-1}(p_j - p_i), \quad i > 1. \tag{1}$$

Setting $\mu_1 = 1$, or noticing the normalisation condition, one also has

$$\mu_i = ip_i + \sum_{j=i+1}^{n} p_j \qquad \text{with the convention } \sum_{j=n+1}^{n} p_j = 0.$$

Not assuming any arrangement of p_i's, one has

$$\mu_i = \sum_{j=1}^{n}\min(p_i, p_j)$$

By inspection, the membership function and the corresponding probability function have the same shape. Therefore, $\mu_i = \mu_j$ implies $p_i = p_j$ and, from $\mu_i \geq \mu_j$, we know that $p_i \geq p_j$. Solving (1) with respect to p_j for μ_j known, we get

$$p_i = \sum_{j=1}^{n}\frac{1}{j}(\mu_j - \mu_{j+1}), \quad i = 1, 2, ..., n \text{ and } \mu_{n+1} = 0.$$

In [], a continuous case has been investigatid. Given a probability density function (PDF) defined in **R**, the corresponding membership function is calculated to minimise a performance index

$$\min_{A} \int A^2(x)\mathrm{d}x$$

subject to the following constraints:

(i) $E(A(x)) \mid x$ is distributed according to the PDF $\geq c$

Where $E(\cdot)$ stands for the expected value of the membership function of A while 'c' denotes a confidence level put close to 1.

(ii) $0 \leq A(x) \leq 1$.

An integral in the above formulation, which is minimised, visualises the fact that the obtained membership function, A, is 'sharp' (i.e. selective) enough (in the sense of the energy measure of fuzziness). Between the two constraints, the second one is obvious, while the first states the fact that the elements which are most likely (in the sense of probability) should have high membership values.

By the use of constrained optimisation techniques for infinite dimensional space, the optimal membership function is equal to

$$A(x) = \begin{cases} \lambda p(x) & \text{if } \lambda p(x) < 1 \\ 1 & \text{otherwise} \end{cases}$$

with λ resulting from the following equation

$$\lambda \int_{\lambda p(x)<1} p(x)\mathrm{d}x + \int_{\lambda p(x)\geq 1} p(x)\mathrm{d}x - c = 0$$

Treating a fuzzy set as a projectable random set can be fruitful in determining a membership function following such a contsruction. When describing vague notions, it is evident that the most difficult situation is to express a grade of membership for intermediate elements, not for those which completely belong or do not belong to the concept (satisfy it).

Let us consider regions between two such fuzzy notions (for example, modelling a concept of *high* and *medium*) where uncertainty of classification of these objects to one of the categories is significant. Suppose it can be characterised by means of the probability density function $p(x)$. This function can be estimated, for instance, by making use of histogram of responses 'don't know'. Assume $p(x)$ takes non-zero values in a certain closed interval of the space in which the fuzzy sets are defined. Then the membership functions of A and B, say *medium* and *high*, are constructed accordingly (see Figure 5-1):

$$A(x) = \begin{cases} 1 & \text{if } x < a_1 \\ \int_x^{a_2} p(u)\mathrm{d}u & \text{if } a_1 \leq x \leq a_2 \\ 0 & \text{otherwise} \end{cases}$$

and

$$B(x) = \begin{cases} 0 & \text{if } x < a_1 \\ \int_{a_1}^{x} p(u)\,du & \text{if } a_1 \le x \le a_2 \\ 1 & \text{otherwise} \end{cases}$$

Figure 5-1. Construction of membership function with the aid of probability density function.

Some relations of the problem of membership function estimation to psychometric scaling techniques discussed by Thurstone [] has been discussed in [].

5. 6. Concluding remarks

The first chapter has contained the main ideas of fuzzy sets, giving the reader a concise presentation of fundamentals of their theory. The reader can get an overall view of some techniques that are characteristic for fuzzy set, has inspected some fuzzy set theoretic operations, some techniques of fuzzy measures and measures of fuzzyness, the possibility and probability theory. After this the progressively growing field of fuzzy logic and approximate reasoning have been introduced.

A reader who wishes to become involved in greater detail concerning fuzzy sets can refer to existing literature.

References

[1] Zadeh, L. A. 1965. Fuzzy sets. Information & Control, vol 8, 338-353.

[2] Zadeh, L. A. 1965. Fuzzy sets and systems Proc. Symp. on System Theory, Polytech. Inst. Brooklyn, 29-37.

[3] Godal, R.C., and T.J. Goodman 1980. Fuzzy sets and borel. IEEE Trans. on System Man, and Cybernetics, vol 10, 637.

[4] Bellman, R. and M. Giertz 1973. On the analitic formalism of the theory of fuzzy sets. Information Sciences vol 5, 149-156.

[5] Bonissone, P.P., and K.S. Decker 1986. Selecting uncertainty calculi and granularity: An experiment in trading-off precision and complexity. In Kanal and Lemmer. 217-247.

[6] Dubois, D., and H. Prade 1985. A review of fuzzy set aggregation connectives. Information Science 36, 85-121.

[7] Mizumot, M. 1989. Pictorial representations of fuzzy connectives, Part I: cases of T-norms, T-conorms and averaging operators. *FSS* 31, 217-242.

[8] Dubois, D., and H. Prade 1980. New results abaut properties and semantics of fuzzy set-theoretic operators. In Wang and Chang, 59-75.

[9] Dubois, D., and H. Prade 1982. A class of fuzzy measures based on triangular norms. Inter. J. Gen. Syst. 8, 43-61.

[10] Thole, U., H.J. Zimmermann, and P. Zysno, 1979. On the suitability of minimum and product operators for the intersection of fuzzy sets. *FSS* 2, 167-180.

[11] Werners, B. 1984. Interaktive Entscheidungsunterstutzung durch ein flexibles mathematisches Programmierungsytem. Munchen.

[12] H.J. Zimmermann, H.J., and P. Zysno, 1983. Decision and evaluations by hierarchical aggregation of information. *FSS* 10, 243-266.

[13] Werners, B. 1988. Aggregation models in mathematical programming. In Mitra, 259-319.

[14] H.J. Zimmermann, H.J., and P. Zysno, 1980. Latent connectivities in human decision making. *FSS* 4, 37-51.

[15] Klir, G.J., and T.A. Folger 1988. Fuzzy Sets, Uncertainty and Information. Englewood Cliffs.

[16] Sugeno, m. 1977. Fuzzy measures and fuzzy integrals–A survey. In Gupta, Saridis, and Gaines. 89-102.

[17] Murofushi, T., and M. Sugeno, 1989. An interpretation of fuzzy measures and the choquet integral as an integral with respect to a fuzzy measure. *FSS* 29, 201-227.

[18] Dubois, D., and H. Prade, 1988. Possibility Theory. Plenum Press, New York.

[19]

[20]

[21]

[22]

[]

[]