Chapter 1

Introduction

“The key elements in human thinking are not numbers but labels of fuzzy sets.”

—L. A. Zadeh

1.1 Historical Background

The concept of a fuzzy set emanates from the observation made by Zadeh (1965) that “more often than not, the classes of objects encountered in the real physical world do not have precisely defined criteria of membership”. This observation gives emphasis to the gap existing between intellectual representations of reality and usual mathematical representations, which supports binary logic, precise numbers, differential equations etc.

The specificity of fuzzy sets is to capture the conception of partial membership. The characteristic function of a fuzzy set, often called membership function is a function whose range is an authoritatively ordered membership set containing more than two (often a continuum of) values (typically, the unit interval).
Therefore, a fuzzy set is often understood as a function. This has been a source of reproval from mathematicians [Arbib (1977)] as functions are already well-known and a theory of functions already exists. The novelty of fuzzy set theory was first proposed by Zadeh (1965), it treat the functions as if they were subsets of their domains, such functions are used to represent steady classes. It follows that classical set-theoretic notions like intersection, union, complement, inclusion, etc. are extended so as to combine functions ranging on an ordered membership set. In elementary fuzzy set theory, the set-union of functions is performed by taking their point-wise maximum, their intersection by their point-wise minimum, their complementation by means of an order-reversing automorphism of the membership scale and set-inclusion by the point-wise inequality between functions.

Let us consider an example of image processing, a machine has to read typewritten characters which computes correlations with various pattern prototypes and extracts certain “features.” Actual samples of the letter “A” may produce a variety of values of these criteria and some of the criteria intended to detect “A’s” may denounce, which produce an ambiguity. There may not have any way of determining whether or not some character is an “A”. Thus, the set of characters intended to be apprehended as “A’s” is a fuzzy set, a set without a well defined boundary. The fuzziness appears to be an essential aspect of this problem.

Fuzzy ambiguities and vagueness are eccentricity of many problems. Such problems may be ill-posed as they do not contain unique solutions or in other words, they may not have solutions at all in the usual sense. The theory of fuzzy sets studies formal properties of such ill-posed problems and ill defined sets, in the same manner as ordinary set theory does for ordinary sets. As “hard sciences”, such as physics, find crisp (as opposed to fuzzy) relations between their observable, the so-called “soft sciences” (biology, psychology, etc.) may involve fuzzy relations between variables.

The aim of fuzzy set theory is to provide a formal setting for incomplete
and gradual information as expressed by people in natural language. There is
a very long tradition of philosophical interest in ambiguity and imprecision of
knowledge. The concepts like inexactness, vagueness, uncertainty, etc. are de-
tained in different way by scholars. There is however a quite general agreement
in considering all such notions as relevant for representational systems, of which
a language is a typical example.

1.2 Crisp and Fuzzy Sets

The foundation of crisp set was laid down by cantor, Russell, Frege and others.
The theory of sets influenced and enriched almost every branch of mathematics
and has also helped to elucidate the relation between mathematics and philos-
ophy. The fundamental conception of a set is the origin point in the study of
modern algebra.

"Collection of well-defined and distinct objects is called a set".

The words aggregate, class or collection are also used in place of the word “set”.
But the use of the word set is common. In general capital letters like A, B, C, ...
etc. are used to denote the sets and lower letters a, b, c, ... to denote the objects or
elements belonging to these sets. We express the relation between an object and
a set to which it belongs by writing, \( a \in A \). Characteristic (Indicator) function
\( \chi_A \) that declares which element of \( X \) are members of the set and which are not,
as follows

\[
\chi_A = \begin{cases} 
1, & \text{when } x \in A, \\
0, & \text{when } x \notin A,
\end{cases}
\]

which is formally expressed by \( \chi_A : X \to \{0, 1\} \).
Operations on Crisp Set

**Union**- If A and B are two non empty sets, then $A \cup B$ defined as

$$A \cup B = \{x| x \in A \text{ or } x \in B\} ,$$

For a family of sets $\{A_i | i \in N\}$, we have

$$\bigcup A_i = \{x| x \in A_i \text{ for some } i \in N\} .$$

**Intersection**- If A and B are two non empty sets, then $A \cap B$ defined as

$$A \cap B = \{x| x \in A \text{ and } x \in B\} ,$$

For a family of sets $\{A_i | i \in N\}$, we obtain

$$\bigcap A_i = \{x| x \in A_i \text{ for all } i \in N\} .$$

**Complement of a set**- complement of a set $A$ is denoted by $A^c$ and defined as

$$A^c = \{x \in X| x \notin A\} .$$

**De-Morgan’s law**

$$(A \cap B)^c = A^c \cup B^c ;$$
and

$$(A \cup B)^c = A^c \cap B^c.$$  

In general, De Morgan’s principles can be stated for $n$ sets as provided here for events, $A_i$:

$$
(A_1 \cup A_2 \cup \ldots \cup A_n) = \overline{A_1} \cap \overline{A_2} \cap \ldots \cap \overline{A_n};
$$

$$
(A_1 \cap A_2 \cap \ldots \cap A_n) = \overline{A_1} \cup \overline{A_2} \cup \ldots \cup \overline{A_n}.
$$

**Fuzzy Set**

In classical set theory, the membership of elements in a set is evaluated in binary digits corresponding to a bivalent condition — an element either belongs or does not belong to the set. But in real life situations, many state of affairs take place where inclusion and non-inclusion in a set are not obviously defined; for instance, the classes of tall girl, beautiful paintings, expensive cars, sunny days, etc. The margins of such sets are vague and the transition from member to non-member appears rather than sudden. Fuzzy set theory allows the steady assessment of the membership of elements in a set; this is described with the aid of a membership function valued in the real unit interval $[0, 1]$. This membership was extensive to possess various “degree of membership” on the real continuous interval $[0, 1]$.

Fuzzy sets take a broad view of classical sets, since the indicator functions of classical sets are special cases of the membership functions of fuzzy sets, if the latter only take values 0 or 1. In fuzzy set theory, classical bivalent sets are usually called crisp sets. Fuzzy sets were introduced by Lotfi A. Zadeh (1965) and Dieter Klaua (1965) in 1965 as an extension of the classical idea of set.

Zadeh (1965) defined the *fuzzy sets* as the sets on the universe $X$ which can accommodate “degrees of membership.” The concept of a fuzzy set contrasts with a classical concept of a bivalent set (crisp set), whose boundary is required to be
precise, i.e., a crisp set is a collection of things for which it is known whether any given object lies inside it or not. Zadeh (1965) generalized the idea of a crisp set by extending a valuation set \( \{0, 1\} \) (definitely in/definitely out) to the interval of real values (degrees of membership) between 1 and 0 denoted as \([0, 1]\). We can say that the degree of membership of any particular element of a fuzzy set express the degree of compatibility of the element with a concept represented by fuzzy set. It means that a fuzzy set \( A \) contains an object \( x \) to degree \( \mu(x) \), i.e., \( \mu(x) = \text{Degree}(x \in A) \), and the map \( \mu : X \rightarrow \{\text{Membership Degrees}\} \) is called a set function or membership function.

**Definition 1.2.1 [Zadeh(1965)]:** Let \( X = \{x_1, x_2, ..., x_n\} \) be a finite universe of discourse and let \( A \subset X \), then \( A \) is a fuzzy set is defined by

\[
A = \{(x_i, \mu_A(x_i)) : \mu_A(x_i) \in [0, 1] ; \forall x_i \in X\},
\]

where \( \mu_A : X \rightarrow [0, 1] \) is membership function of \( A \). The number \( \mu_A(x_i) \) shows the degree of membership of \( x_i \in X \) to \( A \).

From Figure 1.2, it can be noted that \( a \) is clearly a member of fuzzy set \( A \), \( c \) is clearly not a member of fuzzy set \( A \), but the membership of \( b \) is found to be vague. Hence \( a \) can take membership value 1, \( c \) can take membership value 0 and \( b \) can take membership value between 0 and 1 \([0 \text{ to } 1]\), say 0.4, 0.7, etc. This is set to be a partial membership of fuzzy set \( A \).

**Representation of Fuzzy Sets**

There are various notation conventions in fuzzy sets- When the universe of discourse \( X \) is finite and discrete:

**Vector form:** \( A = \left\{ \frac{\mu_A(x_i)}{x_i} : x_i \in X, \ i = 1(1)n \right\} \).

**Summation form:** \( A = \frac{\mu_A(x_1)}{x_1} + \frac{\mu_A(x_2)}{x_2} + ... + \frac{\mu_A(x_n)}{x_n} = \sum_i \frac{\mu_A(x_i)}{x_i} \).

**Ordered pairs:** \( A = \{(x_i, \mu_A(x_i)) : x_i \in X\} \).
When the universe of discourse $X$ is continuous:

**Integral form:** $A = \left\{ \int \frac{\mu_A(x)}{x} \right\}$.

**Function or graphical form:** By means of a graph as in Figure-1.3 (a) and (b).

Here, in these notations, the horizontal bar is not a quotient but rather a delimiter. The summation is not the algebraic summation, but a theoretical aggregation operator or collection operator or similarly, the integral sign.

**Example 1.1:** Let $X$ be a set of positive real number and $A$ be a fuzzy set “about 50 years old” such that $A = \{(x, \mu_A(x)) : x \in X\}$, where

$$\mu_A(x) = \frac{1}{1 + \left(\frac{x-50}{10}\right)^2}.$$ 

**Example 1.2:** $A = \text{“real numbers close to 10”}$

$$A = \left\{ (x, \mu_A(x)) : \mu_A(x) = (1 + (x - 10)^2)^{-1} \right\}.$$
Operations and Properties on Fuzzy Sets

Let $A$, $B$, $C$ be three fuzzy sets defined on universe of discourse $X$. For a given arbitrary element $x$ of the universe, the following function theoretic operations are defined for $A$, $B$, $C$ on $X$:

**Union (OR operation):** It is denoted by $\mu_{A \cup B}(x) = \mu_A(x) \lor \mu_B(x)$. This is also known as $t$-conorm operators and are defined as follows:

- **Maximum:** $\max \{ \mu_A(x), \mu_B(x) \}$
- **Probabilistic Sum:** $\mu_A(x) + \mu_B(x) - \mu_A(x).\mu_B(x)$
- **Bounded Sum:** $\min \{1, \mu_A(x) + \mu_B(x)\}$.

**Intersection (AND operation):** It is denoted by $\mu_{A \cap B}(x) = \mu_A(x) \land \mu_B(x)$. This is also known as $t$-norm operators and are defined as follows:

- **Minimum:** $\min \{ \mu_A(x), \mu_B(x) \}$
- **Product:** $\mu_A(x) \cdot \mu_B(x)$
- **Bounded Product:** $\max \{0, \mu_A(x) + \mu_B(x) - 1\}$.
Complement (NOT operation): It is denoted by $\mu_T(x) = \mu_{\sim A}(x)$. This is also known as fuzzy complement operator which is almost universally used in fuzzy inference systems and is defined as

**Fuzzy Complement:** $1 - \mu_A(x)$.

Magnitude or Scalar Cardinality of Fuzzy Set (Sigma Count): Let $A$ be a fuzzy set defined on $X$, then scalar cardinality (or sigma count) of $A$ is denoted by $|A|$ and defined as

$$|A| = \sum_{x \in X} \mu_A(x) = \sum_{\text{count}} (A).$$

It is observed that $|A| \geq 0$.

Secondly, comparing the magnitude of fuzzy set $A$ with that of universal set $X$ can be an idea.

$$\|A\| = \frac{|A|}{|X|}.$$  

This is called “relative cardinality”.

**Subsethood:** Let $A$ and $B$ be any two fuzzy sets defined on universe of discourse $X$, then the degree of subsethood of $A$ in $B$ is denoted by $S(A, B)$ and is defined as

$$S(A, B) = \frac{1}{|A|} \left( |A| - \sum_{x \in X} \max \{0, \mu_A(x) - \mu_B(x)\} \right)$$

$$= 1 - \frac{\sum_{x \in X} \max \{0, \mu_A(x) - \mu_B(x)\}}{\sum_{\text{count}} (A)}.$$  

Kosko (1986) examined a fuzzy set as a fuzzy message. He extended the bit value representation of an ordinary set to fuzzy set. If $X = \{a, b, c, d\}$, then the subset $A = \{b, d\}$ can be represented by a bit vector $A = \{0, 1, 0, 1\}$. Similarly, a fuzzy subset of $X$ can be expressed as a fit vector or fuzzy message, e. g., $A = \{0.5, 0.1, 0.4, 0.2\}$. The containment of a fuzzy message in another one has been expressed through a fuzzy message conditioning measure called subsethood [Kosko (1986, 1992)] of a fuzzy set in another one.
In other hand, Subsethood is defined as
\[
S(A, B) = \frac{|A \cap B|}{|A|} = \frac{\sum_{A \cap B} \text{count}(A \cap B)}{\sum_{A} \text{count}(A)},
\]
where intersection is the fuzzy intersection. Thus, it can be observed that two fuzzy sets \(A\) and \(B\) are equal if and only if \(S(A, B) = 1\).
\(S(A, B) = 0\) if and only if \(A\) and \(B\) have no points in common, i.e., have no points with fuzzy membership is greater than 0 (zero).
The sum term in the formula describes the sum of the degree to which the subset inequality \(\mu_A(x) \leq \mu_B(x)\) violated. \(|A|\) in the denominator is a normalizing factor for getting the range \(0 \leq S(A, B) \leq 1\).

**Properties of Fuzzy Sets**

The properties of the classical set also suits for the properties of the fuzzy sets.
The important properties of fuzzy set include:

**Commutativity:** \(A \cup B = B \cup A\) and \(A \cap B = B \cap A\).

**Associativity:** \(A \cup (B \cup C) = (A \cup B) \cup C\) and \(A \cap (B \cap C) = (A \cap B) \cap C\).

**Distributivity:** \(A \cup (B \cap C) = (A \cup B) \cap (A \cup C)\) and \(A \cap (B \cup C) = (A \cap B) \cup (A \cap C)\).

**De Morgan’s Law:** \((A \cap B) = \overline{A} \cup \overline{B}\) and \((A \cup B) = \overline{A} \cap \overline{B}\).

**Absorption:** \(A \cup (A \cap B) = A\), \(A \cap (A \cup B) = A\).

**Idempotency:** \(A \cup A = A\) and \(A \cap A = A\).

**Identity:** \(A \cup \varnothing = A\) and \(A \cap X = A\).

**Zero Law:** \(A \cap \varnothing = \varnothing\) and \(A \cup X = X\).

**Transitivity:** If \(A \subset B \subset C\) then \(A \subset C\).

**Involution:** \(\overline{\overline{A}} = A\).

Some important concepts related to fuzzy set theory are defined as follows:
Equality of Two Fuzzy Sets

Two fuzzy sets $A$ and $B$ are said to be equal iff $\mu_A(x_i) = \mu_B(x_i), \forall x_i \in X$.

Standard Fuzzy Sets

A standard fuzzy set is that member of the class of fuzzy equivalent sets whose all membership values are less than or equal to 0.5.

Support of a Fuzzy Set

Let $A$ be fuzzy set defined on $X$, then the support of fuzzy set $A$ is a crisp set. It is denoted by $Supp(A)$ and defined as

$$Supp(A) = \{ x \in X : \mu_A(x) > 0 \}.$$ 

Core

The core of a membership function for some fuzzy set $A$ is defined as that region of the universe that is characterized by complete and full membership in the set $A$. That is, the core comprises those elements $x$ of the universe such that $\mu_A(x) = 1$.

Boundary

If the region of universe has a non zero membership but not full membership, this defines the boundary of a membership function for fuzzy set $A$; The boundary has the elements whose membership lies between 0 and 1, i.e., $0 < \mu_A(x) < 1$.

Crossover Point

The crossover point of a membership function is the elements in universe, whose membership value is equal to 0.5, i.e., $\mu_A(x) = 0.5$.

Height

The height of the fuzzy set $A$ is the maximum value of the membership function, $\max(\mu_A(x))$. 
Normal Fuzzy Set

If the membership function has at least one element in the universe whose value is equal to 1, then that set is called as normal fuzzy set.

Subnormal Fuzzy Set

If the membership function has the membership values less than 1, then that set is called as subnormal fuzzy set.

α- Cut Set

Let A be fuzzy set defined on universe of discourse and $\alpha \in [0, 1]$ be any number, then $\alpha$- cut set of A is a crisp set. It is denoted by $^{\alpha}A$ and defined as

$$^{\alpha}A = \{ x \in X : \mu_A(x) \geq \alpha \}.$$  

And strong $\alpha$- cut set of fuzzy set A is denoted by $^{\alpha+}A$ and defined as

$$^{\alpha+}A = \{ x \in X : \mu_A(x) > \alpha \}.$$  

Special Note: For any fuzzy set A and pair $\alpha_1, \alpha_2 \in [0, 1]$ of distinct values
such that $\alpha_1 < \alpha_2$, we have $\alpha_1 A \supseteq \alpha_2 A$ and $\alpha_1^+ A \supseteq \alpha_2^+ A$.

**Convex Fuzzy Set:**

If the membership function has membership values those are monotonically increasing or monotonically decreasing or they are monotonically increasing and decreasing with the increasing values for elements in the universe, those fuzzy set $A$ is called convex fuzzy set. In other hand, for any elements $x$, $y$ and $z$ in a fuzzy set $A$, the relation $x < y < z$ implies that

$$\mu_A(y) \geq \min [\mu_A(x), \mu_A(z)].$$

**Non Convex Fuzzy Set**

If the membership function has membership values which are not strictly monotonically increasing or monotonically decreasing or both monotonically increasing and decreasing with increasing values for elements in the universe of discourse, then this is called as non convex fuzzy set.

**Fuzzy Number**

A fuzzy number is a quantity whose value is imprecise, rather than exact as is the case with “ordinary” (single-valued) numbers. Any fuzzy number can be thought of as a function whose domain is a specified set. Fuzzy number is an important concept in Fuzzy set theory. It has helped in providing a language which integrates different branch of the topic into one unified whole.

Among the various types of fuzzy sets, of special significance are fuzzy sets that are defined on the set $R$ of real numbers. Membership functions of these sets, which have the form

$$A : \mathbb{R} \rightarrow [0, \ 1],$$

clearly have a quantitative meaning and may under certain conditions be viewed
as fuzzy numbers or fuzzy intervals. To view them in this way, they should capture our intuitive conceptions of approximate numbers or intervals, such as “numbers that are close to a given real number” or “numbers that are around a given interval of real numbers.” Such concepts are essential for characterizing states of fuzzy variables and consequently, play an important role in many applications, including fuzzy control, decision making, approximate reasoning, optimization, and statistics with imprecise probabilities.

**Definition 1.2.2:** To qualify as a fuzzy number, a fuzzy set \( A \) on \( \mathbb{R} \) must possess at least the following three properties:

(a). \( A \) must be normal fuzzy set;

(b). \( \alpha \)-cut must be closed interval for every \( \alpha \in [0, 1] \);

(c). support of \( A, \, 0^+A \), must be bounded.

Since \( \alpha \)-cuts of any fuzzy number are required to be closed intervals for all \( \alpha \in [0, 1] \), every fuzzy number is a convex fuzzy set. The inverse, however, is not necessarily true, since \( \alpha \)-cuts of some convex fuzzy sets may be open or half-open intervals.

**Linguistic Variables**

In retreating 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 artificial language.

The motivation for the use of words or sentences rather than numbers is that linguistic characterizations are, in general, less specific than numerical ones [Zadeh (1973)].
Definition 1.2.3 [Zadeh (1973)]: A linguistic variable is fully characterized by a quintuple \((x, T(x), X, f, m)\) in which \(x\) is the name of the variable, \(T\) is the set of linguistic terms of \(x\) that refer to a base variable whose values range over a universal set \(X\), \(g\) is a syntactic rule (a grammar) for generating linguistic terms and \(m\) is a semantic rule that assigns to each linguistic term \(t \in T\) its meaning, \(m(t)\), which is a fuzzy set on \(X\) (i.e., \(m : T \rightarrow F(X)\)).

An example of a linguistic variable is shown in Figure 1.5. Its name is performance. This variable expresses the performance (which is the base variable in this example) of a goal-oriented entity (a person, machine, organization, method, etc.) in a given context by five basic linguistic terms—very small, small, medium, large, very large—as well as other linguistic terms generated by a syntactic rule (not explicitly shown in Figure 1.5), such as not very small, large or very large, very very small, and so forth. Each of the basic linguistic terms is assigned one of five fuzzy numbers by a semantic rule, as shown in the figure. The fuzzy num-
bers, whose membership functions have the usual trapezoidal shapes are defined on the interval $[0, 100]$, the range of the base variable. Each of them expresses a fuzzy restriction on this range.

**Membership Functions**

**Definition 1.2.4:** Membership function of fuzzy set $A$ on the universe of discourse $X$ is defined as $\mu_A : X \to [0, 1]$, where every member of $X$ is mapped to the value between 0 and 1. This value is called degree of membership or membership value, quantifies the grade of membership of the element in $X$ to the fuzzy set $A$.

The membership functions can be constructed in the following ways:

(a). Employing proficient knowledge,

(b). Explicitly, based on observations collected in advance and processed appropriately (e. g., by statistical methods),

(c). Systematically, by properly preferred functions (e. g., probabilistic distribution).

The types of membership function are as follows:

**Triangular Membership Function**

Triangular membership function is stipulated by three parameters, defined by

$$
\mu_A(x) = \begin{cases} 
0, & x \leq a \ & \& x \geq b, \\
\frac{x-a}{m-a}, & a < x \leq m, \\
\frac{b-x}{b-m}, & m < x < b,
\end{cases}
$$

where $a$ and $b$ are lower and upper limit, respectively and number $m$ is defined as $a < m < b$. 

16
Figure 1.6: Common membership functions: (a) Triangular, (b) Trapezoidal, (c) Gaussian

Trapezoidal Membership Function

Trapezoidal membership function is stipulated by four parameters, defined by

\[
\mu_A(x) = \begin{cases} 
0, & x < a \text{ & } x > d, \\
\frac{x-a}{m-a}, & a \leq x \leq b, \\
1, & b \leq x \leq c \\
\frac{b-x}{b-m}, & c \leq x \leq d, 
\end{cases}
\]

where \( a \) and \( d \) are lower and upper limit, respectively and lower and upper support limit \( b \) and \( c \) are defined such that \( a < b < c < d \).

Gaussian Membership Function

Gaussian membership is determined by two parameters \((m, \delta)\) defined as

\[
\mu_A(x) = \exp\left( -\frac{(x-m)^2}{\delta^2} \right),
\]

where \( m \) denotes the centre value and \( \delta \) shows the standard deviation (width \( \delta > 0 \)) of the function, respectively.
Fuzzification

Fuzzification is an important concept in the fuzzy logic theory. Fuzzification is the process where the crisp quantities are converted to fuzzy (crisp to fuzzy). By identifying some of the uncertainties present in the crisp values, we form the fuzzy values. The conversion of fuzzy values is represented by the membership functions.

In any practical applications, in industries, etc., measurement of voltage, current, temperature, etc., there might be a negligible error. This causes imprecision in the data. This imprecision can be represented by the membership functions. Hence, fuzzification is performed.

Thus, fuzzification process may involve assigning membership values for the given crisp quantities. There are various methods to assign the membership values or the membership functions to fuzzy variables. The assignment can be just done by intuition or by using some algorithms or logical procedures. The methods for assigning the membership values are listed as follows:

(a). Intuition, (b). Inference,
(c). Rank ordering, (d). Angular fuzzy sets,
(e). Neural networks, (f). Genetic algorithms, and
(g). Inductive seasoning.

Defuzzification

Defuzzification means the fuzzy to crisp conversions. The fuzzy results generated cannot be used as such to the applications, hence, it is necessary to convert the fuzzy quantities into crisp quantities for further processing. This can be achieved by using defuzzification process. The defuzzification has the capability to reduce a fuzzy to a crisp single-valued quantity or as a set or converting to the form in
which fuzzy quantity is present. Defuzzification can also be called as “rounding off” method. Defuzzification reduces the collection of membership function values in to a single sealer quantity.

There are other various defuzzification methods employed to convert the fuzzy quantities into crisp quantities. The output of an entire fuzzy process can be union of two or more fuzzy membership functions. They are:

(a). $\alpha-$cut procedure, (b). Max-membership principle,
(c). Centroids method, (d). Weighted average method,
(e). Mean–max membership, (f). First of maxima or last of maxima,
(g). Centre of sums, and (h). Centre of largest area.

\[ \square \]

1.3 Intuitionistic Fuzzy Set

Let $X$ be the set of all countries with elective governments. Assume that we know for every country $x \in X$ the percentage of the electorate who have voted for the corresponding government. Let it be denoted by $P(x)$ and let $\mu(x) = \frac{P(x)}{100}$. Let $\mu(x) = 1 - \nu(x)$. This number corresponds to that part of electorate who have not voted for the government. By means of the fuzzy set theory we cannot consider this value in more detail. However, if we define $\nu(x)$ as the number of votes given to parties or persons outside the government, then we can show the part of electorate who have not voted at all and the corresponding number will be $1 - \mu(x) - \nu(x)$.

Definition 1.3.1 [Atanassov (1986)]. An intuitionistic fuzzy set (IFS) $A$ in a finite universe of discourse $X = \{x_1, x_2, ..., x_n\}$ is given by

$$ A = \{ (x, \mu_A(x), \nu_A(x)) : x \in X \} , $$

where $\mu_A : X \to [0, 1]$ is the degree of membership and $\nu_A : X \to [0, 1]$ is the
degree of non-membership of $x \in X$ to $A$ such that

$$0 \leq \mu_A(x) + \nu_A(x) \leq 1, \forall x \in X.$$  

For an IFS $A$ in $X$ we call the intuitionistic index (or hesitancy degree) of an element $x \in X$ in $A$ the following expression:

$$\pi_A(x) = 1 - \mu_A(x) - \nu_A(x).$$

It is evident that $0 \leq \pi_A(x) \leq 1, \forall x \in X$ [Atanassov (1986, 1999)].

Obviously, an FS $A$ defined on $X$ can be represented as the following intuitionistic fuzzy set:

$$A = \{ (x, \mu_A(x), 1 - \mu_A(x)) : x \in X \}, \text{ with } \pi_A(x) = 0, \forall x \in X.$$  

The complement set of $A$ is denoted by $A^c$ and is defined as

$$A^c = \{ (x, \nu_A(x), \mu_A(x)) : x \in X \}.$$  

A geometric interpretation of intuitionistic fuzzy sets and fuzzy sets is presented in Figure 1.7 which summarizes considerations presented in Szmidt and Kacprzyk (2000). Basically, it should be meant as follows. An intuitionistic fuzzy set $X$ is mapped into the triangle $ABD$ in such a manner that each element of $X$ corresponds to an element of $ABD$-in Figure 1.7, as an example, a point $x' \in ABD$ corresponding to $x \in X$ is marked (the values of $\mu_i(x)$, $\nu_i(x)$, $\pi_i(x)$ fulfill (A)). When $\pi_i$ is equal 0, then $\mu_i + \nu_i = 1$. In Figure 1.7, this condition is fulfilled only on the segment $AB$. Segment $AB$ may be therefore viewed to represent a fuzzy set.

The orthogonal projection of the triangle $ABD$ gives the representation of an intuitionistic fuzzy set on the plane. (The orthogonal projection transfers $x' \in ABD$ into $x'' \in ABC$.) The interior of the triangle $ABC = ABD'$ is the area where $\pi > 0$. Segment $AB$ represents a fuzzy set described by two parameters: $\mu$ and $\nu$.  

20
Figure 1.7: A geometrical interpretation of an intuitionistic fuzzy set [Szmidt and Kacprzyk (2000)]

The orthogonal projection of the segment AB on the axis $\mu$ (the segment [0, 1] is only considered) gives the fuzzy set represented by one parameter $\mu$ only. (The orthogonal projection transfers $x'' \in ABC$ into $x''' \in CA$.)

Operations on Intuitionistic Fuzzy Sets

Atanassov (1983, 1986, 1994) defined operations over intuitionistic fuzzy sets. Here, we discuss some of their basic properties:

For every two IFSs A and B, the following operations are defined as follows: $A \subset B \text{ iff } (\forall x \in X) \left( \mu_A(x) \leq \mu_B(x) \& \nu_A(x) \geq \nu_B(x) \right)$,
\[ A = B \text{ iff } (\forall x \in X)(\mu_A(x) = \mu_B(x) \land \nu_A(x) = \nu_B(x)), \]
\[ \overline{A} = \{ (x, \nu_A(x), \mu_A(x)) : x \in X \}, \]
\[ A \cup B = \{ (x, \max(\mu_A(x), \mu_B(x)), \min(\nu_A(x), \nu_B(x)) : x \in X \}, \]
\[ A \cap B = \{ (x, \min(\mu_A(x), \mu_B(x)), \max(\nu_A(x), \nu_B(x)) : x \in X \}, \]
\[ A + B = \{ (x, \mu_A(x) + \mu_B(x) - \mu_A(x) \cdot \mu_B(x), \nu_A(x) \cdot \nu_B(x)) : x \in X \}, \]
\[ A.B = \{ (x, \mu_A(x) \cdot \mu_B(x), \nu_A(x) + \nu_B(x) - \nu_A(x) \cdot \nu_B(x)) : x \in X \}, \]
\[ A@B = \left\{ \left( x, \frac{\mu_A(x) + \mu_B(x)}{2}, \frac{\nu_A(x) + \nu_B(x)}{2} \right) : x \in X \right\}, \]
\[ A\#B = \left\{ \left( x, \sqrt{\mu_A(x) \cdot \mu_B(x)}, \sqrt{\nu_A(x) \cdot \nu_B(x)} \right) : x \in X \right\}, \]
\[ A \ast B = \left\{ \left( x, \frac{\mu_A(x) + \mu_B(x) + \mu_A(x) \cdot \mu_B(x)}{\mu_A(x) \cdot \mu_B(x) + 1}, \frac{\nu_A(x) + \nu_B(x) + \nu_A(x) \cdot \nu_B(x)}{\nu_A(x) \cdot \nu_B(x) + 1} \right) : x \in X \right\}, \]
\[ \Box A = \{ (x, \mu_A(x), 1 - \mu_A(x)) : x \in X \}, \]
\[ \Diamond A = \{ (x, 1 - \nu_A(x), \nu_A(x)) : x \in X \}, \]
\[ C(A) = \{ (x, K, L) : x \in X \}, \text{ where } K = \max_{x \in X} \mu_A(x), L = \min_{x \in X} \nu_A(x), \]
\[ I(A) = \{ (x, k, l) : x \in X \}, \text{ where } k = \min_{x \in X} \mu_A(x), l = \max_{x \in X} \nu_A(x). \]

**Example 1.3:** Let \( X = \{a, b, c, d, e\} \) be the universe of discourse and let \( A \) and \( B \) be two intuitionistic fuzzy sets defined over \( X \). Then
\[ A = \{ \langle a, 0.5, 0.3 \rangle, \langle b, 0.1, 0.7 \rangle, \langle c, 1.0, 0.0 \rangle, \langle d, 0.0, 0.0 \rangle, \langle e, 0.0, 1.0 \rangle \}, \]
\[ B = \{ \langle a, 0.7, 0.1 \rangle, \langle b, 0.3, 0.2 \rangle, \langle c, 0.5, 0.5 \rangle, \langle d, 0.2, 0.2 \rangle, \langle e, 1.0, 0.0 \rangle \}, \]
we obtain
\[ \overline{A} = \{ \langle a, 0.3, 0.5 \rangle, \langle b, 0.7, 0.1 \rangle, \langle c, 0.0, 1.0 \rangle, \langle d, 0.0, 0.0 \rangle, \langle e, 1.0, 0.0 \rangle \}, \]
\[ A \cap B = \{ \langle a, 0.5, 0.3 \rangle, \langle b, 0.1, 0.7 \rangle, \langle c, 0.5, 0.5 \rangle, \langle d, 0.0, 0.2 \rangle, \langle e, 0.0, 1.0 \rangle \}, \]
\[ A \cup B = \{ \langle a, 0.7, 0.1 \rangle, \langle b, 0.3, 0.2 \rangle, \langle c, 1.0, 0.0 \rangle, \langle d, 0.2, 0.0 \rangle, \langle e, 1.0, 0.0 \rangle \}, \]
\[ A + B = \{ \langle a, 0.85, 0.03 \rangle, \langle b, 0.37, 0.14 \rangle, \langle c, 1.0, 0.0 \rangle, \langle d, 0.2, 0.0 \rangle, \langle e, 1.0, 0.0 \rangle \}, \]
\[ A.B = \{ \langle a, 0.35, 0.37 \rangle, \langle b, 0.03, 0.76 \rangle, \langle c, 0.5, 0.5 \rangle, \langle d, 0.0, 0.2 \rangle, \langle e, 0.0, 1.0 \rangle \}, \]
\[ A@B = \{ \langle a, 0.6, 0.2 \rangle, \langle b, 0.2, 0.45 \rangle, \langle c, 0.75, 0.25 \rangle, \langle d, 0.1, 0.1 \rangle, \langle e, 0.5, 0.5 \rangle \}, \]
\[ A\#B = \{ \langle a, 0.591, 0.173 \rangle, \langle b, 0.173, 0.374 \rangle, \langle c, 0.07, 0.0 \rangle, \langle d, 0.0, 0.0 \rangle, \langle e, 0.0, 0.0 \rangle \}, \]
\[ A \ast B = \{ \langle a, 0.444, 0.194 \rangle, \langle b, 0.194, 0.394 \rangle, \langle c, 0.5, 0.5 \rangle, \langle d, 0.1, 0.1 \rangle, \langle e, 0.5, 0.5 \rangle \}, \]
\[ A\#B = \{ \langle a, 0.58, 0.15 \rangle, \langle b, 0.15, 0.31 \rangle, \langle c, 0.66, 0.0 \rangle, \langle d, 0.0, 0.0 \rangle, \langle e, 0.0, 0.0 \rangle \}. \]
1.4 Fuzzy Uncertainty and Probability

In a thrown of a die, obtaining the outcome in the top face is uncertainty called randomness. On the other hand the outcome is described using the probability distribution on the six faces die. Uncertainty that arises due to chance is called Probabilistic Uncertainty (PU).

A digital image of the top face is analyze by an artificial vision system, which is more complex situation due to insufficient information provided by system that suggest the top face is either a 5 or 6, although cannot be further precise. Mentioned type of uncertainty is arises due to restrictions of the substantiation assembly system. The discussed situation reflects ambiguity in specifying the exact solution, and is called non-specificity by Yager; Pal and Bezdek (1994) preferred to use the alternate term Resolutional Uncertainty (RU). If we are certain that the top face is either a 5 or 6, this case involves only non-specificity. Additional, the information from vision system might also supply a certainty factor. For instance, the system might suggest that the top face is either a 5 or a 6 with belief of 0.8. In this case the chance uncertainty due to the top face can take any value is also discussed, consequently the system contain both PU and RU.

Now, we are supposed to interpret the top face of the die is high (or low). Intentioned us about other third kind of uncertainty, which due to linguistic imprecision or vagueness also termed as Fuzzy Uncertainty (FU). This type of uncertainty is different from PU and RU due to blunt (not sharply defined) boundaries situation, instead the belongingness of elements or events to crisp sets. Sometimes the fuzziness uncertainty is customary related with probabilistic uncertainty. For instance, in a throw of die the occurrence of a 6 supports the fuzzy event HIGH more than a 3 does, but there is still an element of chance about the outcome of a throw, so the system contains PU and FU. Furthermore,
it is obvious that RU can also come into sight in this third case, so the most complex systems may reveal all three types.

Further, fuzzy sets provide a mathematical way to represent vagueness and fuzziness in humanistic like approaches. For example, Suppose two glasses of liquid lies on a table. The liquid in the first glass is described to you as having a 95% chance of being healthful and good. The liquid in the second glass is described as having a 0.95 membership in the class of “healthful and good” liquids. Which glass would you select, keeping in mind that the first glass has a 5% chance of being filled with non healthful liquids, including poisons [Bezdek (1993)]?

What philosophical distinction can be made regarding the forms of information?

Suppose we are allowed to measure the test the liquids in the glasses. The prior probability of 0.95 in each case becomes a posterior probability of 1.0 or 0; that is, the liquid is either being or not. However, the membership value of 0.95, which measures the extent to which the drinkability of the liquid is “healthful and good,” remains 0.95 after measuring or testing. The example illustrates very clearly the difference in the information content between chance and fuzziness.

This brings us to the clearest distinction between fuzziness and chance. Fuzziness describes the lack of distinction of an event, whereas chance describes the uncertainty in the occurrence of the event. The event will occur or not occur; but is the description of the event clear enough to measure its occurrence or non occurrence?

Fuzzy and probability are different ways of expressing uncertainty. While both fuzzy and probability theory can be used to represent subjective belief, fuzzy set theory uses the concept of fuzzy set membership (i.e., how much a variable is in a set), probability theory uses the concept of subjective probability (i.e., how probable a variable is in a set). While this distinction is mostly philosophical, the
fuzzy-logic-derived possibility measure is inherently different from the probability measure; hence, they are not directly equivalent.

Fuzzy uncertainty differs from probabilistic uncertainty because it deals with the situations, where the boundaries are not sharply defined. Probabilistic uncertainties are not due to ambiguity about set-boundaries, but rather about the belongingness of elements or events to crisp sets.

However, many statisticians are of the opinion that only one kind of mathematical uncertainty is needed and thus, fuzzy logic is unnecessary. Lotfi A. Zadeh (1965) argued that fuzzy logic is different in character from probability and is not a replacement for it. He fuzzified probability to fuzzy probability and also generalized it to what is called possibility theory. Note, however, that fuzzy logic is not controversial to probability but rather complementary.

\[ \text{1.5 Fuzzy Information and Discrimination Measures} \]

Let \( X = \{x_1, x_2, ..., x_n\} \) be a discrete random variable with probability distribution \( P = \{(p_1, p_2, ..., p_n) : p_i \geq 0; \sum_{i=1}^{n} p_i = 1\} \) in an experiment. Shannon’s (1948) gave a mathematical formulation to measure uncertainty of the randomness in a probability distribution and information contained in an experiment as

\[
H_S(P) = - \sum_{i=1}^{n} p_i \log p_i, \tag{1.5.1}
\]

which is called Shannon’s entropy.

The expression (1.5.1) indicates the uncertainty due to probabilistic nature of the phenomenon concerned. On the advice of the famous mathematician-physicists John Von Nuemann, Shannon’s called the expression (1.5.1) as a mea-
sure of entropy because it resembled the expression of entropy in thermodynamics. Meanwhile, there was no connection between the two entropies—thermodynamics entropy and information theoretic entropy, but later some links were discovered and information theoretic entropy was found useful in the study of thermodynamics.

Zadeh (1965) developed the concept of fuzzy set and defined the entropy of a fuzzy set, which is different from the classical Shannon’s entropy as no probabilistic concept is needed in order to define it. It may be noted that fuzzy entropy deals with vagueness and ambiguous uncertainties, while Shannon (1948) entropy deals with randomness (probabilistic) of uncertainties. Fuzzy entropy is a measure of fuzziness of a set which arises from the intrinsic ambiguity or vagueness carried by the fuzzy set.

Let \( X = \{x_1, x_2, \ldots, x_n\} \) be the members of the universe of discourse, then all membership values \( \mu_A(x_i); \ i = 1 (1) n \) lie between 0 and 1, but these are not probabilities because their sum is not unity. However,

\[
\psi_A(x_i) = \frac{\mu_A(x_i)}{\sum_{i=1}^{n} \mu_A(x_i)}, \ i = 1 (1) n, \tag{1.5.2}
\]

is probability distribution. On bearing in mind (1.5.2), Kauffman (1975) proposed entropy of a fuzzy set \( A \) having \( n \) support points \( x_i; \ i = 1 (1) n \) by

\[
H_K(A) = -\frac{1}{\log n} \sum_{i=1}^{n} \psi_A(x_i) \log \psi_A(x_i). \tag{1.5.3}
\]

Fuzziness is found in our decision, in our language and in the way we process information. The main use of information is to remove uncertainty and fuzziness. In fact, we measure information supplied by the amount of probabilistic uncertainty removed in an experiment. The measure of uncertainty removed is called as a measure of information while measure of fuzziness is the measure of vagueness and ambiguity of uncertainties.

In fuzzy set theory, the entropy is a measure of fuzziness which means the amount
of average ambiguity or difficulty in making a decision whether an element belongs to a set or not.

Definition 1.5.1 [De Luca and Termini (1972)]: A measure of fuzziness $H(A)$ in a fuzzy set should satisfy the following properties:

(F1). **(Crispness):** $H(A)$ is minimum iff $A$ is crisp set, i.e., $\mu_A(x_i) = 0$ or $1$ for all $x_i \in X$;

(F2). **(Maximality):** $H(A)$ is maximum if and only if $A$ is the fuzziest set, i.e., $\mu_A(x_i) = 0.5$ for all $x_i \in X$;

(F3). **(Resolution):** $H(A) \geq H(A^*)$, where $A^*$ is a sharpened version of $A$;

(F4). **(Duality):** $H(A) = H(A^c)$, where $A^c$ is a complement of $A$.

Corresponding to Shannon’s (1948) entropy of probability distribution De Luca and Termini (1972) defined the following measure of fuzzy entropy:

$$H_{DT}(A) = -\sum_{i=1}^{n} [\mu_A(x_i) \ln \mu_A(x_i) + (1 - \mu_A(x_i)) \ln(1 - \mu_A(x_i))] .$$

The fuzzy information given by (1.5.4) measures uncertainty due to vagueness and ambiguity, while probabilistic entropy measures uncertainty due to the information being available in terms of probability distribution. It may be noted that fuzzy equivalent sets have the same fuzzy information and also two sets may have the same fuzzy information without being fuzzy equivalent.

On the other hand, Loo (1977) developed a general mathematical form for measuring fuzziness as

$$H_L(A) = F \left[ \sum_{i=1}^{n} c_i f_i(\mu_A(x_i)) \right] ,$$

where $c_i \in \mathbb{R}^+$, $f_i$ is a real valued function such that $f_i(0) = f_i(1) = 0$ and $f_i(u) = f_i(1 - u)$ for $u \in [0, 1]$. Here, $f_i(u)$ is a strictly increasing function on $[0, 0.5]$. 

27
In addition, Yager (1979) and Kosko (1986) introduced information measure of a fuzzy set on the distance; from the set to its complement set, the set to its nearest non fuzzy set and from the set to its farthest non fuzzy set.

For intuitionistic fuzzy sets, different definition and interpretation of information measure has been proposed by Burillo and Bustince (1996). Later on, Szmidt and Kacprzyk (2001) extended the axioms of De Luca and Termini (1972) for information measure of IFSs.

**Definition 1.5.2 [Szmidt and Kacprzyk (2001)]:** A real valued function 
\( h: IFS(X) \rightarrow [0, 1] \) is called an information measure for IFSs if it satisfies the following axioms:

(P1). \( h(A) = (\text{minimum}) \) iff \( A \) is a crisp set;

(P2). \( h(A) = 1 \) (maximum) iff \( \mu_A(x_i) = \nu_A(x_i) \) for all \( x_i \in X \);

(P3). \( h(A) \leq h(B) \) and if \( A \) is less fuzzy than \( B \), i.e.,
\[
\mu_A(x_i) \leq \mu_B(x_i) \text{ and } \nu_A(x_i) \geq \nu_B(x_i) \text{ for } \mu_B(x_i) \leq \nu_B(x_i) \text{ or }
\mu_A(x_i) \geq \mu_B(x_i) \text{ and } \nu_A(x_i) \leq \nu_B(x_i) \text{ for } \mu_B(x_i) \geq \nu_B(x_i) \text{ for any } x_i \in X;
\]

(P4). \( h(A) = h(A^c) \).

Hung and Yang (2006) developed axiomatic definitions for IFSs by exploiting the concept of probability. Vlachos and Sergiadis (2007) introduced the following information measure \( E_{VS} \) corresponding to cross entropy measure:

\[
E_{VS}(A) = -\frac{1}{n \ln 2} \sum_{i=1}^{n} \left[ \mu_A(x_i) \ln \mu_A(x_i) + \nu_A(x_i) \ln \nu_A(x_i) \right] -(1 - \pi_A(x_i)) \ln (1 - \pi_A(x_i)) - \pi_A(x_i) \ln 2. \tag{1.5.6}
\]

The intuitionistic fuzzy information measure based on trigonometric function function was developed by Ye (2010); Xia and Xu (2012) derived a cross information measure for IFSs while Verma and Sharma (2013) proposed information measure based on exponential function.
Fuzzy Discrimination Measures

Let

\[ P = \left\{ (p_1, p_2, \ldots, p_n) : p_i \geq 0; \sum_{i=1}^{n} p_i = 1 \right\} \]

and

\[ Q = \left\{ (q_1, q_2, \ldots, q_n) : q_i \geq 0; \sum_{i=1}^{n} q_i = 1 \right\} \]

be two probability distribution of discrete random variable. The discrimination measure of \( P \) from \( Q \) is denoted by \( D(P : Q) \) and satisfies the following axioms:

(a). \( D(P : Q) \geq 0; \)

(b). \( D(P : Q) = 0, \text{ iff } P = Q; \)

(c). \( D(P : X) = H(X) - H(P), \) where \( X \) is the uniform probability distribution and \( H(P) \) is the measure of probabilistic theory.

Kullback and Leibler (1951) proposed the cross entropy measure of \( p \) from the \( q \) as

\[ D_{KL}(P : Q) = \sum_{i=1}^{n} p(x) \ln \frac{p(x)}{q(x)}, \]

which measures the amount of discrimination of \( p \) from the \( q \) [Kullback (1951)].

Kullback (1959) also suggested the measure of symmetric divergence as

\[ J_K(P : Q) = \sum_{i=1}^{n} (p(x) - q(x)) \ln \frac{p(x)}{q(x)}. \]

The discrimination measure of fuzzy set \( A \) about fuzzy set \( B \) is defined as function \( I(A, B) \) satisfying the following axioms:

(a). \( I(A, B) \geq 0; \)

(b). \( I(A, B) = 0, \text{ iff } A = B; \)
(c). $I(A, A_F) = H(A_F) - H(A)$, where $A_F$ is the fuzziest set, i.e., all membership values are 0.5 and $H(A)$ is the fuzzy information measure of $FS(X)$.

Let $X$ be the set of supports $x_i$; $\forall i = 1 \cdots n$ and let $A$ and $B$ be two fuzzy sets with same supporting points. Let $\mu_A(x_1), ..., \mu_A(x_n)$ and $\mu_B(x_1), ..., \mu_B(x_n)$ be two fuzzy vectors of fuzzy sets $A$ and $B$, respectively, where $0 \leq \mu_A(x_i) \leq 1$ and $0 \leq \mu_B(x_i) \leq 1$. The simplest measures of fuzzy directed divergence and symmetric divergence as suggested by Bhandari and Pal (1993) are

$$I_{BP}(A, B) = \sum_{i=1}^{n} \left[ \mu_A(x_i) \ln \frac{\mu_A(x_i)}{\mu_B(x_i)} + (1 - \mu_A(x_i)) \ln \frac{(1 - \mu_A(x_i))}{(1 - \mu_B(x_i))} \right],$$

and

$$D_{BP}(A, B) = I(A, B) + I(B, A) = \sum_{i=1}^{n} [\mu_A(x_i) - \mu_B(x_i)] \ln \left[ \frac{\mu_A(x_i)(1 - \mu_B(x_i))}{\mu_B(x_i)(1 - \mu_A(x_i))} \right].$$

It may be noted that $D(A, B)$ is symmetric with respect to $\mu_A(x_i)$ and $\mu_B(x_i)$. It also holds

(C1). $D(A, B) \geq 0$;

(C2). $D(A, B) = 0$, iff $A = B$;

(C3). $D(A, B) = D(B, A)$;

and $D(A, B)$ does not hold the triangle inequality of a metric. Therefore, $D(A, B)$ is called pseudo metric. Consequently, if we assume $B = A_F$ (the fuzziest set), i.e., $\mu_B(x_i) = 0.5 \forall i$, then we obtain

$$I(A, A_F) = n \ln 2 - \left[ - \sum_{i=1}^{n} \{\mu_A(x_i) \ln \mu_A(x_i) + (1 - \mu_A(x_i)) \ln(1 - \mu_A(x_i))\} \right].$$

It implies


This informative distance between $A$ and $A_F$ represents a measure of non fuzziness in the set $A$. \(\Box\)
1.6 Distance and Similarity Measures for FSs and IFSs

A distance measure of two fuzzy sets is a measure which describes the difference between fuzzy sets. Liu (1992) was the first to develop the definitions of distance measure.

**Definition 1.6.1 [Liu (1992)]:** A real valued function $d : FS(X) \times FS(X) \rightarrow \mathbb{R}^+ \cup \{0\}$ is called distance measure on $FS(X)$ if it satisfies the following postulate:

- **(D1):** $d(A, B) = d(B, A)$, $\forall A, B \in FSs(X)$;
- **(D2):** $d(A, A) = 0$, $\forall A \in FS(X)$;
- **(D3):** $d(G, G^c) = \max_{A,B \in FS(X)} d(A, B)$, $\forall G \in P(X)$;
- **(D4):** $d(A, B) \leq d(A, C)$ and $d(B, C) \leq d(A, C)$, $\forall A, B, C \in FSs(X)$ with the condition $A \subseteq B \subseteq C$.

**Example 1.4:** Let $X = \{x_1, x_2, ..., x_n\}$ and

$$D_p(A, B) = \left( \sum_{i=1}^{n} |\mu_A(x_i) - \mu_B(x_i)|^p \right)^{1/p}, \ \forall A, B \in FSs(X), \ p \in \{1, 2, ...\}.$$  

Obviously, $D_p(A, B)$ is a distance measure on $FS(X)$, since it satisfies all the postulate (D1)-(D4) of distance measure.

**Definition 1.6.2:** If a distance measure $d$ on $FS(X)$ satisfies $\max_{A,B \in FS(X)} d(A, B) = 1$, then $d$ is called the normal distance measure on $FS(X)$.

**Example 1.5:** If $d$ is a distance measure on $FS(X)$, then

$$\overline{d}(A, B) = \frac{d(A, B)}{\max_{C,D \in FS(X)} d(C, D)}, \ \forall A, B \in FSs(X),$$

is a normal distance measure on $FS(X)$.

**Definition 1.6.3:** Let $d$ be a distance measure on $FS(X)$, then $d$ is called a
\(\sigma\)-distance measure if for any \(A, B \in FSs(X)\) and \(G \in P(X)\) holds

\[
d(A, B) = d(A \cap G, B \cap G) + d(A \cap G^c, B \cap G^c).
\]

In addition, based on the axiom definition of distance measure, Fan et al. (1999, 2001) developed some formulae of fuzzy entropy induced by distance measure and also some developed properties of distance measure.

The concept of similarity and distance measures for fuzzy sets introduced and provided a computation formula by Wang (1983), and other researcher developed the measure that indicates the similarity between fuzzy sets. Zwick et al. (1987) reviewed geometric distance and Hausdorff metrics representing similarity measures among fuzzy sets. Turksen and Zhong (1988) applied the similarity measures between fuzzy sets for approximate analogical reasoning. Kaufman and Rousseeuw (1990) proposed some examples to illustrate traditional similarity measure applied in hierarchical cluster analysis. Wang et al. (1995) conducted a comparative study of similarity measure. The similarity of fuzzy sets is now extensively applied in many fields such as fuzzy clustering, image processing, fuzzy reasoning and fuzzy neural network.

A similarity measure is an important tool for determining the degree of similarity between two objects. Pappis and Karacapilidis (1993) proposed three similarity measures based on union and intersection operations, the maximum difference method, and the difference and sum of membership grades. Buckley and Hayashi (1993) used a similarity measure between fuzzy sets to determine the rule to be fired for matching in fuzzy control and neural networks. Wang (1997) presented two similarity measures between fuzzy sets and elements. Liu (1992) and Fan and Xie (1999) provided the axiomatic definition and properties of similarity measures between fuzzy sets.

**Definition 1.6.4 [Liu (1992)]:** A real function \(s : FS(X) \times FS(X) \rightarrow \mathbb{R}^+ \cup \{0\}\) is called a similarity measure on \(FS(X)\) if it satisfies the following postulate:
(S1). \( s(A, B) = s(B, A), \forall A, B \in FSSs(X); \)

(S2). \( s(G, G^c) = 0, \forall G \in P(X); \)

(S3). \( s(C, C) = \max_{A,B \in FSs(X)} s(A, B), \forall C \in FS(X); \)

(S4). \( s(A, B) \geq s(A, C) \) and \( s(B, C) \geq s(A, C), \forall A, B, C \in FSSs(X) \)

with the condition \( A \subseteq B \subseteq C. \)

There exists a one-to-one correlation between all distance measures and all similarity measures. A distance measure \( d \) and its corresponding similarity measure \( s \) satisfy \( d + s = 1 \). It is easy to see that the similarity measure and distance measure are two dual concepts.

For intuitionistic fuzzy sets, Li and Cheng (2002) introduced the concept of similarity measure and applied it to pattern recognition; Liang et al. (2003) and Mitchell (2003) improved the similarity measure of IFS introduced by Li and Cheng (2002); Hung and Yang (2004, 2007, 2008) investigated the similarity measure of IFS based on Hausdorff metric and \( L_p \) distance, respectively; and Szmidt and Kacprzyk (2005) applied the similarity measure of IFS in group decision making.

**Definition 1.6.5 [Li & Cheng (2002), Mitchell (2003) and Xu & Chen (2008)]:** A real function \( S : IFS(X) \times IFS(X) \to [0, 1] \) is called the similarity measure on \( IFS(X) \), if it satisfies the following properties:

(M1). \( S(A, A^c) = 0 \) if \( A \) is a crisp set;

(M2). \( S(A, B) = 1 \Leftrightarrow A = B; \)

(M3). \( S(A, B) = S(B, A); \)

(M4). For all \( A, B, C \in IFSs(X) \), if \( A \subseteq B \subseteq C, \) then \( S(A, C) \leq S(A, B) \) and \( S(A, C) \leq S(B, C). \)
Many studies focus on the similarity measure and the entropy for IFS. However, few studies focus on the relationship among the similarity measure and the entropy for IFS, especially on the systematic transformation of the entropy into the similarity measure for IFS and vice versa. This is important because more formulae could be used to calculate the similarity measure and the entropy of IFS. It is important to unite these two numerical indexes.

1.7 Fuzzy Mean Code Word Length

Let \( S = \{s_1, s_2, ..., s_n\} \) be a source alphabet and let \( A = \{a_1, a_2, ..., a_m\} \) be a code alphabet, which is also the input alphabet of some channel. Then to transmit text written in the source alphabet in the form of code alphabetic character, we have to associate a code alphabet word to each source alphabet word.

An encoding function is a function or a mapping \( \varphi : S \rightarrow A \) determines a code. The code determined by \( \varphi \) is said to be unambiguous if and only if \( \varphi \) is one-to-one (injective), Otherwise, the code is ambiguous.

Let \( s_i \rightarrow w_i \in A, i = 1 (1) n \) be an encoding scheme for a source alphabet \( S = \{s_1, s_2, ..., s_n\} \). Let \( P = \{p_1, p_2, ..., p_n\} \) be the set of probabilities associated with each source letters respectively such that \( \sum_{i=1}^{n} p_i = 1 \), then average code word length is given as

\[
\sum_{i=1}^{n} p_i n_i,
\]

where \( n_i \) is the number of code alphabets in \( w_i \).

Actually “average code word length” is average length of a code word. \( L \) is, in fact, the average value of the random variable associated with the experiment of randomly selecting a source letter from the source text. If we have \( N \) source letter, then expected number of code required to encode a source text consisting of \( N \) source letters is \( L \times N \).
Hölder’s Inequality: If $x_i, y_i > 0; i = 1 \ldots n$ and $\frac{1}{p} + \frac{1}{q} = 1, p > 1$, then the following inequality holds:

$$\sum_{i=1}^{n} x_i y_i \leq \left[ \sum_{i=1}^{n} x_i^p \right]^{\frac{1}{p}} \left[ \sum_{i=1}^{n} y_i^q \right]^{\frac{1}{q}}.$$

1.8 Summary of the Thesis

The thesis is assembled in seven chapters, starting with the general introduction and historical background in the first chapter. The detailed descriptions of the remaining chapters are as follows:

In Chapter 2, new logarithmic fuzzy information measures are introduced and studied their validity. Logarithmic fuzzy discrimination measures are defined and their validity has been studied. Properties of logarithmic fuzzy discrimination measures are listed. Further, the applications to pattern recognition and medical diagnosis of the proposed measures are discussed.

In Chapter 3, two new exponential fuzzy information measures are introduced, studied their validity and established the relations among the various existing and propose fuzzy information measures. To show the effectiveness of the proposed measure, it is compared with the existing measures. Two exponential fuzzy discrimination measures are defined and characterized. Properties of these exponential fuzzy discrimination measures are studied. The applications of these exponential fuzzy symmetric discrimination measures to the pattern recognition and diagnosis of crop disease are examined.

In Chapter 4, some basic definitions related to the fuzzy sets, fuzzy information, distance and similarity measures are studied. New trigonometric and tangent inverse trigonometric information measures for fuzzy sets are proposed.
and studied their validity. A new formula of trigonometric fuzzy distance measure is proposed. Fuzzy information measure induced by fuzzy distance measure is defined. Further, the characterization of $\sigma$- distance measure defined on fuzzy sets as a metric is studied. Two new weighted trigonometric and hyperbolic fuzzy information measures are developed and checked their validity. Finally, the applications of trigonometric and hyperbolic fuzzy information measures to maximum weighted fuzzy information principle are illustrated.

In Chapter 5, fuzzy mean code word lengths are introduced. The bounds of the generalized fuzzy mean code word length of degree $\beta$ in terms of fuzzy information measure and the bounds of the generalized fuzzy mean code word length of type $(\alpha, \beta)$ in terms of fuzzy information measure are studied. The monotonic behaviors of generalized fuzzy mean code word lengths are discussed.

In Chapter 6, new exponential intuitionistic fuzzy information measure is introduced and checked their validity. Numerical example is listed and comparisons are tabulated between new and existing intuitionistic fuzzy information measures, followed by the application of the proposed IFIMWAO method in assessment of service quality of vehicle insurance companies.

In Chapter 7, new trigonometric similarity and intuitionistic fuzzy information measures are developed. Empirical illustration is listed and comparisons are tabulated between new and existing intuitionistic fuzzy information measures. The proposed algorithm IFMWAO (based on intuitionistic fuzzy information and similarity measures) is applied to the physical education teaching assessment.