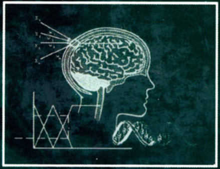
The CRC Press International Series on Computational Intelligence

D. Dumitrescu
B. Lazzerini
L.C. Jain



FUZZY SETS AND THEIR APPLICATION TO CLUSTERING AND TRAINING

CRC PRESS

FUZZY SETS AND THEIR APPLICATION TO CLUSTERING AND TRAINING

The CRC Press

International Series on Computational Intelligence

Series Editor
L.C. Jain, Ph.D., M.E., B.E. (Hons), Fellow I.E. (Australia)

L.C. Jain, R.P. Johnson, Y. Takefuji, and L.A. Zadeh
Knowledge-Based Intelligent Techniques in Industry

L.C. Jain and C.W. de Silva
Intelligent Adaptive Control: Industrial Applications in the Applied Computational Intelligence Set

L.C. Jain and N.M. Martin
Fusion of Neural Networks, Fuzzy Systems, and Genetic Algorithms: Industrial Applications

H.-N. Teodorescu, A. Kandel, and L.C. Jain
Fuzzy and Neuro-Fuzzy Systems in Medicine

C.L. Karr and L.M. Freeman
Industrial Applications of Genetic Algorithms

L.C. Jain and B. Lazzerini
Knowledge-Based Intelligent Techniques in Character Recognition

L.C. Jain and V. Vemuri
Industrial Applications of Neural Networks

H.-N. Teodorescu, A. Kandel, and L.C. Jain
Soft Computing in Human-Related Sciences

B. Lazzerini, D. Dumitrescu, L.C. Jain, and A. Dumitrescu
Evolutionary Computing and Applications

B. Lazzerini, D. Dumitrescu, and L.C. Jain
Fuzzy Sets and Their Application to Clustering and Training

L.C. Jain, U. Halici, I. Hayashi, S.B. Lee, and S. Tsutsui
Intelligent Biometric Techniques in Fingerprint and Face Recognition

Z. Chen
Computational Intelligence for Decision Support

L.C. Jain
Evolution of Engineering and Information Systems and Their Applications

H.-N. Teodorescu and A. Kandel
Dynamic Fuzzy Systems and Chaos Applications

L. Medsker and L.C. Jain
Recurrent Neural Networks: Design and Applications

L.C. Jain and A.M. Fanelli
Recent Advances in Artifical Neural Networks: Design and Applications

M. Russo and L.C. Jain
Fuzzy Learning and Applications

J. Liu and J. Wu
Multiagent Robotic Systems

M. Kennedy, R. Rovatti, and G. Setti
Chaotic Electronics in Telecommunications

H.-N. Teodorescu and L.C. Jain
Intelligent Systems and Techniques in Rehabilitation Engineering

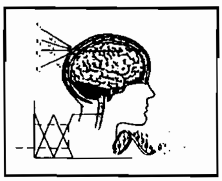
I. Baturone, A. Barriga, C. Jimenez-Fernandez, D. Lopez, and S. Sanchez-Solano
Microelectronics Design of Fuzzy Logic-Based Systems

T. Nishida
Dynamic Knowledge Interaction

C.L. Karr
Practical Applications of Computational Intelligence for Adaptive Control

Ricardo Salem Zebulum, Marco Aurélio C. Pacheco, and Marley Maria B.R. Vellasco
Evolutionary Electronics: Automatic Design of Electronic Circuit and Systems by Genetic Algorithms

FUZZY SETS AND THEIR APPLICATION TO CLUSTERING AND TRAINING



D. Dumitrescu
University of Cluj-Napoca

B. Lazzerini
University of Pisa

L.C. Jain
University of South Australia

CRC Press
Boca Raton London New York Washington, D.C.

Library of Congress Cataloging-in-Publication Data

Dumitrescu, D.(Dimitru), 1949—
Fuzzy sets and their application to clustering and training / D. Dumitrescu, B. Lazzerini, L. C. Jain.
p. cm.
Includes bibliographical references and index.
ISBN 0-8493-0589 (alk. paper)
1. Fuzzy sets. 2. Cluster analysis. I. Lazzerini, Beatrice, 1953— II. Jain, L. C. (Laxmi Chandra), 1926— III. Title.
QA248.5 .D86 2000
511.3′22—dc21 99-088763
CIP

Direct all inquiries to CRC Press LLC, 2000 N.W. Corporate Blvd., Boca Raton, Florida 33431.

International Standard Book Number 0-8493-0589-6
Library of Congress Card Number 99-088763

To my mother and in memory of my father

Dan-Dumitru Dumitrescu

To our families and all of our students

Beatrice Lazzerini
Lakhmi Jain

Acknowledgments

Our thanks are due to our families and all persons who helped us to accomplish the work concerning this book.

We are particularly grateful to all the active researchers in the field of Fuzzy Set Theory and its applications, who sent us copies of their papers.

We also thank Geta Bonda for her kind help in preparing the manuscript.

Preface

Fuzzy Set Theory (FST) and its underlying fuzzy logic represent one of the most significant scientific and cultural paradigms emerging in the second half of this century.

This paradigm has a definite 'postmodern' flavour as it represents an alternative to the positivist view of the world.

FST paradigm imposed itself on the scientific community who became aware that common sense concepts and approximate reasoning (mainly based on simple, intuitive rules) have a great theoretical and technological potential and an important explanatory power.

It became evident that many parts of natural and artificial reality may adequately be described in terms that tolerate the ambiguity and imprecision specific to the reality itself.

These approximate descriptions may be sufficient for many practical purposes. A 'complete', rigorous description of complex systems is usually impossible or too costly.

It also became manifest that traditional descriptions of complex systems are sometimes possible only at the cost of over-simplification.

On the other hand, the human mind and natural languages can perfectly cope with ambiguity and imprecision.

FST significantly enlarged the frames of the mathematical approach to incorporate imprecise concept descriptions and imprecise (or approximate) reasoning, and to treat them in a definitely rigorous manner.

Reality usually has an intrinsic, non-probabilistic ambiguity. If we want to remove this ambiguity we have to resort to over-simplified descriptions.

On one hand, the grid of traditional mathematics and scientific concepts may be too coarse or too restrictive. On the other hand, a complete and exact description of a system could have prohibitive cost, be difficult to manage, or even be useless due to its complexity. Further, predictive or control results based on such description might not be obtained in real time.

It is important to observe that ambiguity, vagueness and imprecision represent only one side of FST. Many important theoretical models and useful applications have been obtained by taking only this aspect into account.

The dual aspect of ambiguity-tolerating reasoning is the nuanced reasoning. With FST we became aware that in order to describe reality plausibly we have to bypass the strong, brittle Yes-No dichotomy.

FST represents a useful set theoretical model of multi-valued logics. Multi-valued logics and FST cross-fertilize each other. Moreover, some multi-valued logics are now considered as special cases of fuzzy logic (in the wide sense).

Clustering is the very first application of FST. The reason for this is that fuzzy clustering does not require very sophisticated conceptual or mathematical tools.

On the other hand, fuzzy clusters represent natural models of fuzzy concepts. We may speak, for instance, about the class of useful books in a library, or the class of clever students in a school, and so on.

Fuzzy classes (or clusters) are the simplest and most natural examples of fuzzy set.

Moreover, it is evident that most real-world classes are fuzzy rather than crisp.

Goals of the book

The main goals of the book are:

(*i*) to offer a general, comprehensive introduction to Fuzzy Set Theory with a special emphasis on the notions and results needed for training and clustering purposes;

(*ii*) to give an adequate and unitary mathematical framework for fuzzy classification and clustering;

(*iii*) to provide a general methodology to develop fuzzy training and classification methods;

(*iv*) to provide a general method to obtain a large variety of fuzzy clustering algorithms;

(*v*) to offer a comprehensive introduction to the fields of fuzzy learning classifiers and fuzzy clustering;

(*vi*) to present some basic fuzzy clustering algorithms treated in a unified manner;

(*vii*) to present a hierarchical fuzzy clustering method able to detetct hierarchically organized cluster structures without any *a priori* knowledge of the optimal number of clusters in the data set.

Structure of the book

The book is structured in four parts.

Part I (Chapters 1-6) describes fundamental aspects of Fuzzy Set Theory. This part is intended to offer an introduction, both simple and self-contained, to mathematical notions and results concerning fuzzy sets, fuzzy partitions and related concepts.

Part II (Chapters 7-8) is dedicated to fuzzy learning machines (fuzzy classifiers) able to learn from fuzzy data. Fuzzy learning classifiers may use the outputs of a fuzzy clustering algorithm as training sets.

Some basic training algorithms are described and their convergence properties are investigated.

Part III (Chapters 9-16) deals with fuzzy partitional prototype-based clustering. The main topics within this part concern fuzzy clustering with point and linear prototypes, adaptive clustering, validity functionals and convergence properties of clustering procedures.

Part IV (Chapters 17-19) is dedicated to fuzzy discriminant analysis and fuzzy hierarchical clustering.

The content of each chapter is now briefly described.

Chapter 1 contains the basic notions about fuzzy sets. The framework is that of triangular norms (t-norms) and conorms (t-conorms), and their generators. Several families of t-norms and t-conorms are considered. Various definitions of union, intersection and complement operations are considered.

The *ordinal sum*-based method to obtain new t-norms and t-conorms is addressed.

An axiomatic definition of complementation operator is given and some of its instances are taken into account.

Chapter 2 explores the properties of set operations induced by various operators, particularly T_o, S_o and T_∞, S_∞.

A particular attention is paid to the equivalence between the concepts of binary fuzzy partitions and fuzzy partition of unity (Ruspini's early definition of fuzzy partition).

It is proved that the two concepts are completely equivalent if and only if the set operations are induced by T_∞ and S_∞.

The notion of fuzzy n-partition, for $n > 2$, is investigated. Other topics of this chapter are:

(i) refinement relation for fuzzy partitions;

(ii) algebraic join of two fuzzy partitions.

The algebraic structure of the family $L(X)$ of the fuzzy sets on a fixed universe X is studied in *Chapter 3*. The key concepts are those of partially ordered set, lattice, residuated lattice and multi-valued algebra (MV-algebra). It is emphasized that T_∞, S_∞ seem to be suitable to define set operations for fuzzy sets, whereas T_o, S_o are suitable to describe the order relation on $L(X)$.

The framework of MV-algebras makes clear the non-competitive relationship of the pairs (T_o, S_o) and (T_∞, S_∞), and their different meanings and specific roles.

Residuated lattices also represent a powerful algebraic tool for characterizing fuzzy sets and their underlying logic.

The notion of a basic triple is considered. Roughly speaking, (T, S, C) is a basic triple if C is a complement operation, T and S are C-dual and S and C have the same generator. The use of basic triples in defining set operations ensures a deep coherence of the resulting fuzzy set theory.

The connection of the notions of basic triple and residuated implication is investigated. In this respect, a matching operator is defined. A basic triple and a matching operator can generate a residuated lattice via a residuated implication.

Chapter 4 begins with a presentation of the metric concepts for fuzzy sets. There are several definitions of distance between fuzzy sets. The definition used in this chapter is not the standard one. We preferred it both for its being a natural extension of the classical notion and for its fitness to clustering and training purposes.

The other topics of this chapter are:

(*i*) distance between fuzzy points;

(*ii*) diameter of a fuzzy set;

(*iii*) fuzzy ball;

(*iv*) bounded fuzzy set;

(*v*) distance in a fuzzy class.

Chapter 5 deals with the notions of entropy and informational energy of fuzzy partitions. These notions are based on a suitable concept of fuzzy measures.

Let us suppose the atoms of a fuzzy partition P describe the outcomes of an experiment. The entropy of the fuzzy partition P measures the information obtained (or the uncertainty removed) by performing the experiment associated with this fuzzy partition.

Chapter 6 is devoted to the characterization of fuzzy sets using fuzziness and nonfuzziness measures. These measures give a global characterization of the uncertainty/certainty associated with a fuzzy description of a situation, experiment, etc. Some particular fuzziness and nonfuzziness measures for fuzzy sets defined on finite or infinite universes are considered.

Several correlation coefficients of fuzzy sets are also taken into account.

Fuzzy learning classifiers, considered in *Chapter 7*, may deal with data that are either erroneous or containing atypical points. This kind of classifiers are robust and may cope with the non-separability of the training sets.

Chapter 7 proposes a new paradigm for fuzzy neural network training. Within this paradigm, classifiers able to learn fuzzy training classes may be considered. A fuzzy perceptron is considered as an example of this paradigm. The convergence of the fuzzy perceptron training procedure is studied.

A robust variant of the fuzzy perceptron is considered.

Several generalizations of the fuzzy perceptron model are also taken into account. The fuzzy pocket (FP) algorithm is such a generalization. FP algorithm is able to obtain an approximate separation hyperplane in the case of arbitrary non-separable training classes.

Chapter 8 is dedicated to fuzzy training procedures based on squared error criterion functions.

A fuzzy relaxation algorithm is derived. Some variants of this algorithm are also considered.

Other training procedures considered in this chapter are fuzzy relatives of some well-known classical learning algorithms. These procedures are: Fuzzy MSE method, Fuzzy Widrow-Hoff algorithm and several variants of Fuzzy Ho-Kashyap algorithm.

The convergence properties of the fuzzy training algorithms presented in this chapter are studied. Most of the models and algorithms in Part II are original.

The main idea in *Chapter 9* is that the sub-cluster structure of a fuzzy class may also contain some important useful clustering information. To detect this structure a well-known alternating optimization method is applied to a squared error objective function. The objective function is not guessed, but it is derived using a general method based on the local distance with respect to a fuzzy set.

A Generalized Fuzzy n-Means (GFNM) algorithm for detecting the sub-cluster structure of a fuzzy class is derived.

In order to detect unequal size clusters correctly, the use of several adaptive distances is proposed. A modified GFNM algorithm is considered. Other topics in this chapter are:

- (*i*) data normalization using mean and variance of fuzzy classes;
- (*ii*) use of local distances for clustering purposes within the GFNM algorithm.

In *Chapter 10*, the infinite family of (G)FNM algorithms is considered. The limit properties of this family are given.

Other topics in this chapter are:

- (*i*) reformulated version of the (G)FNM algorithm;
- (*ii*) clustering with L_p metric;
- (*iii*) clustering with set prototypes.

Chapter 11 mainly concerns the detection of linear (sub)clusters of a fuzzy class. Two clustering methods are considered. They are alternating optimization and principal component analysis of a fuzzy class.

In *Chapter 12* various families of adaptive fuzzy clustering algorithms are considered. These standard families are generalized to detect the cluster substructure of a fuzzy class.

Prototype-based partitional clustering algorithms are used. Some algorithms adopt a variable metric inducing matrix.

The main algorithms in this chapter are:

(*i*) adaptive FNM;

(*ii*) shell algorithms;

(*iii*) adaptive fuzzy n-shells algorithms.

In *Chapter 13*, other algorithms to detect spherical, elliptical or planar shaped clusters are considered. Several variants of these algorithms are presented. These algorithms use modified distance functions like:

(*i*) distance generated by an unconstrained distance-inducing matrix (AFNSU family of algorithms);

(*ii*) algebraic distance;

(*iii*) exponential distance.

The main classes of algorithms are:

(*i*) AFNSU family;

(*ii*) ellipsoidal shell-clustering;

(*iii*) fuzzy maximum likelihood;

(*iv*) Gath-Geva algorithm;

(*v*) robust fuzzy clustering algorithms.

The cluster validity problem is addressed in *Chapter 14*. To detect the optimal cluster number, validity functionals are used. Validity functionals give a numerical expression of the quality of a fuzzy partition. The intuitive idea is that a good fuzzy partition is not a very fuzzy one.

In this chapter, some well-known validity functionals, like partition coefficient and classification entropy, are considered.

Using mean and variance, standardized and normalized versions of these functionals are defined. Other validity functionals studied in this chapter are coupling coefficient and proportion exponent.

In *Chapter 15*, a wide range of validity functionals are considered. These functionals are based on various principles. Some of them represent uniform data validity functionals. Most of them are geometric validity functionals that intend to reflect the actual structure of the data set.

Geometric functionals are also related to the cluster shapes. Examples of geometric validity functionals are:

(*i*) fuzzy partition density;

(*ii*) fuzzy partition volume;

(*iii*) class inertia;

(*iv*) separation index.

Other geometric functionals include class prototypes and data set points.

Convergence of the FNM fuzzy clustering algorithms using point prototypes is studied in *Chapter 16*.

Both local and global convergence properties are studied. Some well-known convergence results for FNM are considered. These results still remain valid for the GFNM algorithm.

The convergence theorems considered in this chapter may represent the framework for a general convergence theory for prototype-based partitional fuzzy clustering algorithms.

Part IV contains an original approach to fuzzy discriminant analysis and hierarchical clustering.

Chapter 17 introduces fuzzy scatter matrices and Fisher discriminant vector for two fuzzy classes. Discriminant axes for $n > 2$ fuzzy classes are analyzed. Some fuzzy scattering criteria are considered for clustering purposes.

The classical optimization methods seem not to be suitable to optimize the obtained objective functions. Evolutionary algorithms (particularly genetic algorithms) are suggested as an ideal tool to optimize the scatter objective functions.

Chapter 18 addresses the problem of fuzzy hierarchical clustering. Hierarchical clustering methods may detect the built-in hierarchical structure in a data set and also detect the optimal cluster number in a given data set.

A divisive method to detect a fuzzy cluster hierarchy is proposed. At each decomposition level, only 'real' clusters are retained. The obtained hierarchy is binary and the method is sufficiently flexible and robust. It may deal with arbitrary cluster structures. Each node in the decomposition tree corresponds to a fuzzy class. For each such node, a fuzzy one-level algorithm to detect the fuzzy partition of the corresponding fuzzy class is used.

The quality of each binary fuzzy partition is measured by its *polarization degree.* Using the entropy of a fuzzy partition a stability degree of a fuzzy hierarchy is proposed.

A hierarchical clustering procedure using structural, entropy-based information is also considered.

Chapter 19 addresses the problem of *simultaneous clustering* (SC). The aim of SC is to find simultaneously a fuzzy partition of the data set and a fuzzy partition of the characteristics (features) describing the data samples. The two fuzzy partitions have to be mutually relevant.

The problem of simultaneous clustering has a wide range of practical applications. It is also important for data mining, data reduction techniques and for classifier design for large data sets.

An algorithm for simultaneously detecting one-level fuzzy cluster structures of data and of the corresponding features is proposed. This algorithm is then used to obtain a simultaneous hierarchical clustering structure. It is noteworthy that both algorithms are intrinsically fuzzy, i.e., they do not allow hard versions.

The book contains an extensive unified treatment of fuzzy sets, fuzzy clustering and fuzzy training models.

The book is intended for those who are interested in intelligent computation models and, in particular, in pattern recognition, data mining, clustering and classifier design. It may also be used as an introduction to basic concepts of fuzzy sets, fuzzy clustering and supervised fuzzy training. The book may be useful for scientists from various fields (such as chemistry, physics, biology, economics, and engineering) interested in data analysis.

The book may also be used as a textbook in a one-semester postgraduate course in pattern recognition and fuzzy technology.

Contents

Part I

Basic aspects of fuzzy set theory

Chapter 1

Fuzzy Sets

1.1 Introduction

Fuzzy sets represent a suitable mathematical tool for the modeling of imprecision and vagueness. In general, vagueness is associated with the difficulty of making precise affirmations about a certain domain. On the other hand, in fuzzy set theory, the strong alternative *yes* - *no* is indefinitely nuanced. From this point of view, fuzzy set theory is not only a theory dealing with ambiguity and vagueness. It is also a theory of the nuance reasoning. Pascal's *l'esprit de finesse* and *l'esprit géométrique* are jointed in this theory.

In this chapter, some basic notions and results in fuzzy set theory are presented.

The set operations are introduced using the general formalism of triangular norms (t-norms) and triangular conorms (t-conorms) as is developed in the theory of probabilistic metric spaces. Some families of t-norms and t-conorms and their relationships are considered. The additive generators of t-norms and t-conorms are defined.

The method of ordinal sums to construct new t-norms (t-conorms) from a given family of t-norms (t-conorms) is presented.

The use of t-norms and t-conorms connects fuzzy set theory with the algebraic theory of semigroups.

The solutions of an important functional equation arising from information theory may be characterized using the concept of ordinal sum (Frank's theorem). Frank's theorem will be used in chapter 2 to prove the uniqueness of the pair (T_∞, S_∞).

Other topics of this chapter are:

- conditions for the uniqueness of the standard set connectives T_o, S_o;
- axiomatic definition of fuzzy complement;
- generator of a fuzzy complement;
- C-dual connectives;
- equilibrium and dual point of a complement.

1.2 Fuzzy sets and fuzzy points

1.2.1 Basic definitions

Definition

Let X be a non-empty set considered to be the *universe of discourse.* A *fuzzy set* is a pair (X, A), where $A : X \to I$ and $I = [0, 1]$. A is called the *membership function.*

Remark

In what follows we will consider a fixed universe X. Therefore we may identify a fuzzy set with its membership function.

The family of all fuzzy sets on the universe X will be denoted by $L(X)$. Thus

$$L(X) = \{A \mid A : X \to I\}.$$

The notion of fuzzy set has been introduced by L.A. Zadeh (see [43]).

$A(x)$ is the *membership degree* of x to A. It may also be interpreted as the plausibility degree of the affirmation 'x belongs to A'. If $A(x) = 0$, x is 'definitely not in A' and if $A(x) = 1$, x is 'definitely in A'. The intermediate cases are 'fuzzy'.

Definition

The fuzzy set A is called *non-ambiguous* (or *crisp*) if $A(x) \in \{0, 1\}$.

Remark

The characteristic function of every classical set is thus a non-ambiguous fuzzy set.

Definition

The *empty set* $\emptyset$ is defined as

$$\emptyset(x) = 0, \quad \forall x \in X.$$

Definition

If A is from $L(X)$, the *complement* of A is the fuzzy set $\bar{A}$ defined as

$$\bar{A}(x) = 1 - A(x), \quad \forall x \in X.$$

Remark

The complement of $\emptyset$ is the fuzzy set 1_X. In what follows we will also denote 1_X by X.

Let A and B be fuzzy sets on X.

Definition

The *equality* between A and B on X is determined by the usual equality of mappings, i.e.,

$$A = B \iff A(x) = B(x), \quad \forall x \in X.$$

Definition

The *inclusion* relation between fuzzy sets is defined pointwise, i.e., we have

$$A \subseteq B \iff A(x) \leq B(x), \quad \forall x \in X.$$

Definition

The *product* of the fuzzy sets A, B is the fuzzy set AB defined by

$$(AB)(x) = A(x) \cdot B(x), \quad \forall x \in X.$$

Definition

The *difference* of A and B is the fuzzy set $A - B$ defined by

$$(A - B)(x) = \max\left(A(x) - B(x), 0\right), \quad \forall x \in X.$$

1.2.2 Fuzzy points. Level sets of a fuzzy set

Definition

Let $t \in [0,1]$ and A be a fuzzy set on X. We define the *t-level* (or *t-cut*) of A as the ordinary set

$$A^t = \{x \in X \mid A(x) \geq t\}.$$

Definition

The *strong t-level* (or *strong t-cut*) of A is defined by

$$A^{t^*} = \{x \in X \mid A(x) > t\}.$$

Definition

The *support* of a fuzzy set A on X, denoted supp A, is the ordinary subset of X given by

$$\text{supp } A = \{x \in X \mid A(x) > 0\}.$$

Definition

A fuzzy set on X is called a *fuzzy point*, or a *fuzzy singleton*, if and only if it takes the value 0 for all points in X except one.

Remarks

(*i*) A fuzzy set is a fuzzy point if and only if its support reduces to a point in X.

(*ii*) A fuzzy point is completely determined by its support, say $\{y\}$, and its value b at y. We will use the notation f_y^b for this fuzzy point.

Thus we have

$$f_y^b(x) = \begin{cases} b, & \text{if } x = y \\ 0, & \text{if } x \neq y \end{cases}$$

for every x in X.

Definition

A fuzzy singleton f_y^b *belongs* to a fuzzy set A if and only if $f_y^b(x) \leq A(x)$, for each x in X. We denote this membership relation between a fuzzy singleton and a fuzzy set by $f_y^b \in A$.

Remark

$f_y^b \in A$ if and only if $b \leq A(y)$.

Definition

A fuzzy point f_x^a is called *crisp* if and only if $a = 1$.

1.2.3 Fuzzy points and the inclusion relation

The relation between fuzzy points and the inclusion of fuzzy sets is given by the following proposition.

Proposition

Let A and B be two fuzzy sets on X. Then we have

(i) $A \subseteq B$ if and only if the implication

$$f_x^a \in A \Rightarrow f_x^a \in B$$

holds for each fuzzy point f_x^a that belongs to A.

(ii) $A = B$ if and only if the equivalence

$$f_x^a \in A \iff f_x^a \in B$$

holds for each fuzzy point f_x^a on X.

Proof.

(i) Let us consider $A \subseteq B$. For each fuzzy point f_x^a that belongs to A we have $a \leq A(x) \leq B(x)$. Thus f_x^a belongs to B. Conversely, let x be an arbitrary point in X and let us assume that

$$f_x^a \in A \Rightarrow f_x^a \in B.$$

It follows that

$$a \leq A(x) \Rightarrow a \leq B(x),$$

i.e.,

$$A(x) \leq B(x), \quad \forall x \in X$$

and thus $A \subseteq B$.

(ii) It is obvious. □

1.3 Axioms for operations on fuzzy sets

1.3.1 Basic requirements for set operators

The intersection and the union of two fuzzy sets may be defined by using two functions $F, G : I \times I \to I$. The set operations may be defined pointwise as follows:

$$(A \cap B)(x) = F(A(x), B(x)), \quad \forall x \in X,$$

$$(A \cup B)(x) = G(A(x), B(x)), \quad \forall x \in X.$$

Fuzzy set operations must satisfy some natural requirements. These requirements are listed below.

(*i*) The operations induced by F and G must reduce to crisp (usual) set operations when the sets are non-ambiguous.

(*ii*) The fuzzy set operations must satisfy the boundary conditions:

$$\begin{aligned} A \cap X &= A, \\ A \cap \emptyset &= \emptyset, \end{aligned}$$

$$\begin{aligned} A \cup X &= X, \\ A \cup \emptyset &= A. \end{aligned}$$

(*iii*) The set operations must be associative and commutative. The associativity allows us to extend the operations of fuzzy sets to more than two sets.

(*iv*) The functions F and G must be monotone with respect to both variables.

(*v*) The fuzzy set operations must satisfy standard De Morgan laws:

$$\overline{A \cap B} = \bar{A} \cup \bar{B},$$

$$\overline{A \cup B} = \bar{A} \cap \bar{B}.$$

Remark

From condition (*i*) we obtain:

$$\begin{aligned}
F(1,1) &= 1, \\
\\
F(0,0) &= F(0,1) \\
&= F(1,0) \\
&= 0, \\
\\
G(0,0) &= 0, \\
\\
G(0,1) &= G(1,0) \\
&= G(1,1) \\
&= 1.
\end{aligned}$$

From condition (ii) we have

$$\begin{aligned}
F(a,1) &= a, \\
F(a,0) &= 0, \quad \forall a \in I, \\
\\
G(a,1) &= 1, \\
G(a,0) &= a, \quad \forall a \in I.
\end{aligned}$$

Conditions (v) may be written

$$1 - F(A(x), B(x)) = G(1 - A(x), 1 - B(x)),$$

and

$$1 - G(A(x), B(x)) = F(1 - A(x), 1 - B(x)),$$

for every x from X.

It follows that the functions F and G satisfy the requirements:

$$F(a,b) = 1 - G(1-a, 1-b),$$

$$G(a,b) = 1 - F(1-a, 1-b),$$

for every a, b in [0,1].

It is easy to see that the last two conditions are equivalent.

1.3.2 Axioms for the set operators

Now we are able to give the axioms for the set operators of the fuzzy sets. The minimal requirements are given by the following list of axioms.

Axiom 1

$$\begin{aligned} F(1,1) &= 1, \\ F(0,0) &= F(0,1) \\ &= F(1,0) \\ &= 0, \end{aligned}$$

$$\begin{aligned} G(0,0) &= 0, \\ G(0,1) &= G(1,0) \\ &= G(1,1) \\ &= 1. \end{aligned}$$

Axiom 2

$$\begin{aligned} F(a,1) &= a, \\ F(a,0) &= 0, \quad \forall a \in I, \\ \\ G(a,1) &= 1, \\ G(a,0) &= a, \quad \forall a \in I. \end{aligned}$$

Axiom 3 Commutativity:

$$\begin{aligned} F(a,b) &= F(b,a), \quad \forall a,b \in I, \\ G(a,b) &= G(b,a), \quad \forall a,b \in I. \end{aligned}$$

Axiom 4 Associativity:

$$\begin{aligned} F\left(F(a,b),c\right) &= F\left(a,F(b,c)\right), \quad \forall a,b,c \in I, \\ G\left(G(a,b),c\right) &= G\left(a,G(b,c)\right), \quad \forall a,b,c \in I. \end{aligned}$$

Axiom 5 Monotony:

$$a \le a',\ b \le b' \ \Rightarrow\ F(a,b) \le F(a',b'),$$

$$a \le a',\ b \le b' \ \Rightarrow\ G(a,b) \le G(a',b'),$$

where a, b, a', b' are from I.

Axiom 6 De Morgan law:

$$F(a,b) = 1 - G(1-a, 1-b), \quad \forall a, b \in I.$$

Remark

These axioms suggest us to consider for F a triangular norm (t-norm) and for G a triangular conorm (t-conorm).

We recall that t-norms have been introduced in the context of probabilistic metric spaces.

1.4 Triangular norms and conorms

In this section, we will recall some basic properties of *t-norms* and *t-conorms*. These properties will be very useful to develop a theory of fuzzy sets.

Triangular norms have been studied extensively by Schweizer and Sklar [34], Ling [27], Kimberling [24], Frank [16] and others.

1.4.1 Definition of t-norms. Archimedean t-norms

Definition

A t-norm is a two-argument function

$$T : I \times I \to I$$

fulfilling the axioms:

(i) $T(a,1) = a, \quad \forall a \in I$ (boundary condition), (*ii*) $T(a,b) \leq T(u,v)$ if $a \leq u,\ b \leq v$ (monotony),

(*iii*) $T(a,b) = T(b,a)$ (commutativity),

(*iv*) $T\left(T(a,b),c\right) = T\left(a,T(b,c)\right)$ (associativity).

Remarks

(*1*) From axioms (*i*) and (*ii*) we have

$$0 \leq T(0,a) \leq T(0,1) = 0,$$

and thus

$$T(0,a) = T(a,0) = 0, \quad \forall a \in I.$$

From these two axioms also one gets

$$T(a,a) \leq T(a,1) = a, \quad \forall a \in I.$$

(*2*) It is easy to see that the pair (I,T) is an Abelian semigroup with unity.

(*3*) If " $\leq$ " is the natural order relation on I, then the triple $(I,T,\leq)$ is an Abelian ordered semigroup with unity.

Definition

A t-norm T is said to be *Archimedean* if it fulfills the condition

$$T(a,a) < a, \quad \forall a \in (0,1).$$

Remark

The t-norm T is Archimedean if and only if T has no interior idempotents. This means that there exists no $a \in (0, 1)$ for which

$$T(a, a) = a.$$

1.4.2 Definition of t-conorms. Archimedean t-conorms

Definition

Let T be a t-norm. The two-place function

$$S : I \times I \to I$$

defined by

$$S(a, b) = 1 - T(1 - a, 1 - b), \qquad \forall a, b \in I,$$

is called a *t-conorm* (or the dual of T).

Remarks

(α) If S is a t-conorm then S is monotone, commutative, associative and

$$S(a, 0) = a.$$

(β) We also have

$$S(a, 1) = 1$$

and for every $a \in I$

$$S(a, a) \geq a.$$

Definition

A t-conorm S is called *Archimedean* if and only if it fulfills the condition

$$S(a, a) > a, \quad \forall a \in (0, 1).$$

Remark

If T is an Archimedean t-norm then its dual t-conorm is also Archimedean.

Concerning Archimedean t-norms and t-conorms, in the next section we will give two representation theorems.

1.4.3 Pseudo-inverse and additive generators

Definition

Let f be a continuous and strictly decreasing function $f : [u, v] \to [0, \infty]$.

The *pseudo-inverse* of f is the function $f^{(-1)} : [0, \infty] \to [u, v]$, defined by

$$f^{(-1)}(x) = \begin{cases} v, & \text{if } x \in [0, f(v)] \\ f^{-1}(x), & \text{if } x \in (f(v), f(u)) \\ u, & \text{if } x \in [f(u), \infty], \end{cases}$$

where f^{-1} is the ordinary inverse of f.

Theorem (Ling, 1965) [27]

A function $T : I \times I \to I$ is an Archimedean t-norm if and only if there exists a continuous and strictly decreasing function $f : [0,1] \to [0,\infty]$, with $f(1) = 0$, such that T may be represented as

$$T(a,b) = f^{(-1)}\left(f(a) + f(b)\right) \quad \forall a, b \in [0,1].$$

Moreover, T is *strict*, i.e., is strictly decreasing in (0,1), if and only if $f(0) = +\infty$.

For Archimedean t-conorms we have an analogous result given by the following theorem.

Theorem [27]

A function $S : I \times I \to I$ is an Archimedean t-conorm if and only if there exists a continuous and strictly increasing function $g : [0,1] \to [0,\infty]$, with $g(0) = 0$, such that S may be represented as

$$S(a,b) = g^{(-1)}\left(g(a) + g(b)\right), \quad \forall a, b \in [0,1].$$

Moreover, S is *strict*, i.e., is strictly increasing in (0,1), if and only if $g(1) = +\infty$.

Definition

A strictly decreasing continuous function f that satisfies the condition

$$T(a,b) = f^{(-1)}\left(f(a) + f(b)\right)$$

is called an *additive generator* of T.

Remark

An additive generator of T is unique except for a positive factor, i.e., if f is an additive generator of T then af, $a > 0$, is also an additive generator.

We may now reformulate Ling's theorem concerning Archimedean t-norms as follows.

Theorem

A function $T : I \times I \to I$ is an Archimedean t-norm if and only if T admits an additive generator.

Definition

A strictly increasing continuous function g, with $g(0) = 0$, is an additive generator of the t-conorm S if and only if

$$S(a,b) = g^{(-1)}\left(g(a) + g(b)\right), \quad \forall a, b \in [0,1].$$

Example

The generator of t-conorm S_∞ is the identity function on [0,1].

1.4.4 Frank's fundamental family of t-norms and t-conorms

The most interesting t-norms and t-conorms are listed below:

$$T_o(x,y) = \min(x,y),$$

$$S_o(x,y) = \max(x,y),$$

$$T_1(x,y) = xy,$$

$$S_1(x,y) = x + y - xy,$$

$$T_\infty(x,y) = \max(x+y-1,0),$$

$$S_\infty(x,y) = \min(x+y,1),$$

$$T_s(x,y) = \log_s\left(1 + \frac{(s^x-1)(s^y-1)}{s-1}\right), \; s>0, \; s \neq 1,$$

$$S_s(x,y) = 1 - \log_s\left(1 + \frac{(s^{1-x}-1)(s^{1-y}-1)}{s-1}\right), \; s>0, \; s \neq 1.$$

Remark

The families T_s and S_s, $0 \leq s \leq \infty$, are considered in [16].

Proposition

The previous t-norms form a single family in the sense that

$$T_o = \lim_{s \to 0, s>0} T_s,$$

and

$$T_i = \lim_{s \to i} T_s, \quad i = 1, \infty.$$

For the proof see [16].

Remarks

(*i*) A similar result holds for the t-conorms of the family $\{S_s \mid s \geq 0\}$.

(*ii*) It is easy to see that T_o is not Archimedean and T_∞ as well as T_s, $0 < s < \infty$, are Archimedean.

1.4.5 Other families of t-norms and t-conorms

Some other examples of families of t-norms and related dual t-conorms are reported below.

Yager [41] :

$$T_p^1(x,y) = 1 - \min\left[1, ((1-x)^p + (1-y)^p)^{1/p}\right],$$

$$S_p^1(x,y) = \min\left(1, (x^p + y^p)^{1/p}\right), \quad p > 0.$$

Hamacher [19], [20]:

$$T_a^2(x,y) = \frac{xy}{a + (1-a)(x+y-xy)},$$

$$S_a^2(x,y) = \frac{(a-2)xy + x + y}{1 + (a-1)xy}, \quad a > 0.$$

Schweizer and Sklar [36]:

$$T_r^3(x,y) = \left[\max\left(0, x^r + y^r - 1\right)\right]^{1/r},$$

$$S_r^3(x,y) = 1 - \left[\max\left(0, (1-x)^r + (1-y)^r - 1\right)\right]^{1/r}, \quad r \neq 0.$$

Sugeno [39]:

$$T_b^4(x,y) = \max\left[0, (1+b)(x+y-1) - bxy\right],$$

$$S_b^4(x,y) = \min(1, x+y+bxy), \quad b > -1.$$

Dubois and Prade [10]:

$$T_c^5(x,y) = \frac{xy}{\max(x,y,c)},$$

$$S_c^5(x,y) = \frac{x+y-xy-\min(x,y,1-c)}{1-\min(x,y,1-c)}, \quad c \in (0,1).$$

1.4.6 Relationship between Frank's family and other families

Let us consider the t-norm T_w defined as: