Lappeenrannan teknillinen yliopisto
Lappeenranta University of Technology

Kalle Saastamoinen

MANY VALUED ALGEBRAIC STRUCTURES AS MEASURES OF COMPARISON

Thesis for the degree of Doctor of Philosophy to be presented with due permission for public examination and criticism in the Auditorium of the Student Union House at Lappeenranta University of Technology, Lappeenranta, Finland, on the 7^{th} of November 2008, at noon.

Acta Universitatis
Lappeenrantaensis
322

Supervisor Professor Jorma K. Mattila
Lappeenranta University of Technology
Department of Mathematics and Physics
Lappeenranta, Finland

Reviewer Professor József Dombi
József Attila University
Department of Applied Informatics
Szeged, Hungary

Reviewer Professor Bernard De Baets
Ghent University
Department of Applied Mathematics, Biometrics and Process Control
Coupure links 653, B-9000 Gent, Belgium

Opponent Dr. Péter Majlender
Department of Management and Organization
Hanken School of Economics
Casa Academica, Perhonkatu 6B
FI-00100 Helsinki, Finland

ISBN 978-952-214-645-8 (paperback)
ISBN 978-952-214-646-5 (PDF)
ISSN 1456-4491

Lappeenrannan teknillinen yliopisto
Digipaino 2008

Acknowledgements

I would like to express my gratitude to professor Jorma Mattila, my supervisor, for his support during this research and to the reviewers of this thesis professors József Dombi and Bernard De Baets for their valuable comments, which helped to improve the thesis.

I also wish to acknowledge the following people for their personal support and co-operation: researchers Jaakko Ketola, Jouni Sampo, Esko Turunen and Pasi Luukka, my wife Lydia Saastamoinen, my children Lauri, Mikko, Maija and their mother Satu Johansson, my father Pekka Saastamoinen, my mother Tuulikki Saastamoinen, my brothers Sakari Saastamoinen and Jaakko Saastamoinen.

Special thanks to the Laboratory of Applied Mathematics in Lappeenranta. I also feel gratitude to Elizabeth Nyman for her excellent work in correcting language and the following instances for financial support: the ECSE graduate school, The Foundation of Lahja and Lauri Hotinen, The Foundation of Lappeenranta University of Technology, Lappeenranta University of Technology, the East Finland Virtual University, and the Väisälä foundation.

Thank you all. I hope that in the future we can continue our fruitful co-operation.

Lappeenranta, Finland
October 18, 2008

Kalle Saastamoinen

Abstract

Kalle Saastamoinen

MANY VALUED ALGEBRAIC STRUCTURES AS MEASURES OF COMPARISON

Lappeenranta, 2008
86 p.

Acta Universitatis Lappeenrantaensis 322
Diss. Lappeenranta University of Technology

ISBN 978-952-214-645-8 (paperback), 978-952-214-646-5 (PDF)
ISSN 1456-4491

This thesis studies the properties and usability of operators called t-norms, t-conorms, uninorms, as well as many valued implications and equivalences. Into these operators, weights and a generalized mean are embedded for aggregation, and they are used for comparison tasks and for this reason they are referred to as comparison measures. The thesis illustrates how these operators can be weighted with a differential evolution and aggregated with a generalized mean, and the kinds of measures of comparison that can be achieved from this procedure. New operators suitable for comparison measures are suggested. These operators are combination measures based on the use of t-norms and t-conorms, the generalized 3Π-uninorm and pseudo equivalence measures based on S-type implications.

The empirical part of this thesis demonstrates how these new comparison measures work in the field of classification, for example, in the classification of medical data. The second application area is from the field of sports medicine and it represents an expert system for defining an athlete's aerobic and anaerobic thresholds.

The core of this thesis offers definitions for comparison measures and illustrates that there is no actual difference in the results achieved in comparison tasks, by the use of comparison measures based on distance, versus comparison measures based on many valued logical structures. The approach has been highly practical in this thesis and all usage of the measures has been validated mainly by practical testing. In general, many different types of operators suitable for comparison tasks have been presented in fuzzy logic literature and there has been little or no experimental work with these operators.

Keywords: Fuzzy logic, Operators, Comparison, Expert systems, Classification, Medical data.

UDC 510.644 : 519.234.8 : 004.891

Contents

1 Introduction

William James has written the following about the sense of sameness [40]: "This sense of sameness is the very keel and backbone of our thinking." The fields of problem solving, categorization, data mining, classification, memory retrieval, inductive reasoning, and cognitive processes in general require that the matter of how to assess sameness is understood. Sometimes, in practice, comparison measures used to measure sameness in the field of soft computing are based on a very intuitive understanding of the theoretical backgrounds of mathematics or a naive idea of coupling the measure of sameness to the Euclidean distance. Taking measures intuitively can be a crucial mistake and can lead to so-called black-box systems, which can work but no one is able to say how and why they work. An example of these kinds of systems are neural networks [38]. In particular, neural networks are called black-box systems since for the most part the user cannot explain how the learning from input data was done or how performance can be consistently ensured. One of the first things to consider in the systems, which do any kind of comparison, should be the comparison measures used. In this thesis some many-valued structures are presented, which are then combined with weights, and the values achieved are then aggregated with a generalized mean. These structures are called comparison measures, which can be used in various different applications and have many good properties for further analysis.

The standard approach is to create mathematical models with a kind of logic where every axiom, sentence, connective etc. is based on the classical, Aristotelian, bivalent logic, which is concerned with the interpretations of two values, 0 and 1. The real world, however, is not so black and white, and therefore models offered by bivalent logic are sometimes inaccurate from the beginning, and this inaccuracy originates from the basic nature of this logic. Fuzzy logic is considered in this thesis as the type of many valued logic where every interpretation is in the closed interval of 0 and 1 and used connectives is derived from the many valued logics. It is this nature of offering an infinite number of possible interpretations between 0 and 1 which makes fuzzy logic, in general, more practically suitable than a bivalent approach.

1.1 Fuzzy Set Theory

Fuzzy set theory is an active research area, highly mathematical in its nature. It can provide a robust and consistent foundation for information processing, including pattern-formatted information processing. It plays at least two roles in the pattern recognition. In one role, it serves as an interface between the linguistic variables, seemingly preferred by humans and the quantitative characterizations appropriate for machines. In this role, it might also serve as a bridge between symbolic processing of artificial intelligence and the parallel distributed processing approaches favoured by adaptive pattern recognition. In another role, it emphasizes the possibility-distribution interpretation of the concept of fuzziness. The value of this role is that it legitimizes and provides a meaningful interpretation for some distributions that are believed to be useful, but that might be difficult to justify on the basis of the objective probabilities. The two roles are not distinct, but the differences are interesting and worth noting [67].

Interesting application of the fuzzy set theory in classification and clustering is relevance feedback in information retrieval systems. In these methods, the user's subjective judgments are represented using numbers or distributions in a measurement space and fuzzy sets provide an appropriate framework for the mathematical modeling of this process. The fuzzy classifier presented in this thesis can be used as the base for the fuzzy relevance feedback method with small modifications. Although fuzzy information retrieval methods are shown to be very efficient in many applications, they are not very well-known among information retrieval researchers outside the fuzzy research community. Good sources of information retrieval and, in particular, of fuzzy information retrieval are e.g. [64] and [2].

The expressions *fuzzy set, t-norm, t-conorm, uninorm, similarity, S-type equivalence, aggregation, averaging operator and generalized mean* appear many times in this thesis. Here at the beginning of the thesis, I give a short definition of each of them. Firstly, the difference is defined between the classical crisp set vs. fuzzy set [48].

The classical set theory deals with the crisp sets, where each member x of the universal set X is a member or non-member of the subset in question, therefore when $A \subset X$, then for all $x \in X$ it remains true that either $x \in A$ or $x \notin A$. The crisp set is a sort of degenerated fuzzy set, which can be defined by a function called *characteristic function* that declares which elements of X are members of the set and which are not.

Definition 1 *Set A is defined by its characteristic function χ_A, as follows:*

$$\chi_A : X \to \{0, 1\} .$$

The fuzzy set theory deals with fuzzy sets, where each member x of the universal set X has a certain membership grade in the set in question. The term fuzzy set was first used by Lotfi Zadeh [115]. The fuzzy set theory deals with situations where the ordinary set theory is not accurate enough for the problems at hand.

The membership function assigns to each member of the universal set X a certain value which falls within a specific range. The most commonly used range is the unit interval $[0, 1]$. The value indicates the membership grade of the member x of the universal set X in the set in question.

Definition 2 *Fuzzy set A by its membership function μ_A is as follows:*

$$\mu_A : X \to [0, 1] .$$

In a fuzzy set the membership values $\mu_A : X \to [0, 1]$ may be interpreted in terms of truth values of any propositions that can be considered to belong to some set, and fuzzy set operators in terms of logical connectives in many-valued logic. This provides a formulation of fuzzy set theory based on many-valued logic [47]. In this thesis, I will study this kind of set theoretic based fuzzy logic.

There are also other approaches to fuzzy logic. Klir studied in [47] modal logic [11] as a basis for the development of the fuzzy set theory. This offers a research direction, in which fuzzy sets with non-truth-functional set-theoretic operators are studied. Rough sets [69] can also be treated as a special class of fuzzy sets, in which membership functions are interpreted in terms of conditional probabilities [109].

For more details, see for example [48], [115], [119], [35], [3].

1.2 Fuzzy Negations, Fuzzy Intersections and Fuzzy Unions

When different objects are compared, perhaps the most important and frequently used logic operators are the negation, conjunction and disjunction. These operators can, in a set theoretic view, be seen as complement, intersection, and union. Standard fuzzy negation $\neg$ is most often used and it is of the form $1 - a$ for $\neg a$. Functions, which qualify as fuzzy intersections and unions of any given pair of fuzzy sets are t-norms and t-conorms. Originally, t-norms and t-conorms were used in the field of statistical metric spaces as the tool for generalizing the classical triangular inequality [62], [93].

Definition 3 *A t-norm is a binary operation $T : [0, 1]^2 \to [0, 1]$ that satisfies at least the following for all $x, y, z \in [0, 1]$:*

a1) $T(x, 1) = x$ *(boundary condition),*

b) $T(x, y) = T(y, x)$ *(commutativity),*

c) $y \leq z$ *implies* $T(x, y) \leq T(x, z)$ *(monotonicity),*

d) $T(x, T(y, z)) = T(T(x, y), z)$ *(associativity).*

The only difference between t-norms and t-conorms is the choice of the neutral element in the boundary condition. For t-conorms, this is defined by the following:

a2) $S(x,0) = x$ (boundary condition).

The simple properties above ensure that fuzzy sets aggregated by t-norms denoted by T, and t-conorms denoted by S are intuitively acceptable as meaningful fuzzy intersections and unions of any given pair of fuzzy sets.

The following definition [94] for t-conorms is equivalent to the axiomatic definition given above. This same definition shows the duality that combines t-norms and t-conorms together:

Definition 4 *A function $S : [0,1]^2 \rightarrow [0,1]$ is a dual t-conorm of t-norm such that for all $(x,y) \in [0,1]^2$ both of the following equivalent equalities holds*

$$S(x,y) = 1 - T(1-x, 1-y), \tag{1}$$

$$T(x,y) = 1 - S(1-x, 1-y). \tag{2}$$

From the definition (4) it can be seen that t-norms and t-conorms also always have dual forms. The t-conorm given by (1) is called the dual t-conorm of T, and correspondingly, the t-norm given by (2) is called the dual t-norm of S [46]. One can see that the definition (4) is the generalization of the normal De Morgan's theorem [15] into many valued logics. Duality expressed in (1) allows us to change properties of t-norms into the corresponding properties of t-conorms. Duality also changes the order, that is, if two t-norms are ordered as $T_1 \leq T_2$, then corresponding t-conorms are ordered as $S_1 \geq S_2$.

Some important additional requirements that restrict the class of t-norms and t-conorms and that are needed in this thesis are continuity and Archimedean property. These can be expressed as follows:

Definition 5 *T and S are continuous functions (continuity). A t-norm $T : [0,1]^2 \rightarrow [0,1]$ is continuous if for all convergent sequences $(x_n)_{n\in\mathbb{N}} (y_n)_{n\in\mathbb{N}} \in [0,1]^{\mathbb{N}}$ we have*

$$T\left(\lim_{n\to\infty} x_n, \lim_{n\to\infty} y_n\right) = \lim_{n\to\infty} T(x_n, y_n)$$

Definition 6 *A continuous t-norm (correspondingly for t-conorm) that satisfies $T(x,x) < x$ (for t-conorm $S(x,x) > x$), when $x \in]0,1[$ is an Archimedean t-norm (correspondingly Archimedean t-conorm).*

e) A continuous t-norm (correspondingly for t-conorm) that is continuous and increasing is called strict.

Continuity prevents a situation in which a very small change in the membership grade of either set A or set B would produce a large (discontinuous) change in the membership grade of the intersection or union. From this consideration, it can be seen that in practice it is preferable to use t-norms and t-conorms that are continuous. Archimedean property clearly guarantees that t-norms and t-conorms are intuitionally acceptable to be intersections and unions, when membership degrees are in the interval $]0,1[$. Non-Archimedean t-norms are also not guaranteed to be measurable [8].

Theorem 7 *Representation theorem for t-norms: Let T be a binary operation on the unit interval. Then, T is an Archimedean t-norm iff there exists a rational function f such that*

$$T(x,y) = f^{-1}(f(x) + f(y)) \tag{3}$$

for all $x, y \in [0,1]$.

Proof. See Schweizer and Sklar, 1963 [95] and Ling, 1965 [51]. ■

The following are examples of three parameterized classes of rational functions (decreasing generators) and the corresponding classes of t-norms. In each case, the parameter is used as a subscript of f and T to distinguish different generators and t-norms in each class.

Example 8 *Schweizer and Sklar, 1963 [95]: The class of decreasing generators distinguished by parameter p is defined by*

$$f_p(x) = 1 - x^p \qquad (p \neq 0). \tag{4}$$

$$f_p^{-1}(z) = \begin{cases} 1, \text{ when } z \in \,]-\infty, 0[\\ (1-z)^{\frac{1}{p}}, \text{ when } z \in [0,1] \\ 0, \text{ when } z \in \,]1, \infty[\end{cases} \tag{5}$$

and it is obtained corresponding class of t-norms by applying representation theorem of t-norms

$$\begin{aligned} T_p(x,y) &= f_p^{-1}(f_p(x) + f_p(y)) \\ &= f_p^{-1}(2 - x^p - y^p) \\ &= \begin{cases} (x^p + y^p - 1)^{\frac{1}{p}}, \text{ when } 2 - x^p - y^p \in [0,1] \\ 0, \text{ otherwise} \end{cases} \end{aligned}$$

so

$$T_p(x,y) = (\max(0, x^p + y^p - 1))^{\frac{1}{p}}. \tag{6}$$

Example 9 *Yager, 1980 [110]: Given a class of decreasing generators with parameter w*

$$f_w(x) = (1-x)^w \qquad (w>0), \tag{7}$$

one obtains

$$f_w^{-1}(z) = \begin{cases} 1 - z^{\frac{1}{w}}, \text{ when } z \in [0,1] \\ 0, \text{ when } z \in]1,\infty[\end{cases} \tag{8}$$

and it is reached corresponding class of t-norms by applying characterization theorem of t-norms

$$\begin{aligned} T_w(x,y) &= f_w^{-1}(f_w(a) + f_w(y)) \\ &= f_w^{-1}((1-x)^w + (1-y)^w) \\ &= \begin{cases} 1 - ((1-x)^w + (1-y)^w)^{\frac{1}{w}}, \text{ when } (1-x)^w + (1-y)^w \in [0,1] \\ 0, \text{ otherwise} \end{cases} \end{aligned}$$

so

$$T_w(x,y) = 1 - \min\left(1, [(1-x)^w + (1-y)^w]^{\frac{1}{w}}\right). \tag{9}$$

Example 10 *Frank, 1979 [29]: This class of t-norms is based on the class of decreasing generators*

$$f_s(x) = -\ln\frac{s^x - 1}{s - 1} \qquad (s>0, s\neq 1), \tag{10}$$

whose pseudo-inverses are given by

$$f_s^{-1}(z) = \log_s\left(1 + (s-1)e^{-z}\right). \tag{11}$$

and corresponding class of t-norms is obtained by applying characterization theorem of t-norms

$$\begin{aligned} T_s(x,y) &= f_s^{-1}(f_s(x) + f_s(y)) \\ &= f_s^{-1}\left(-\ln\tfrac{(s^x-1)(s^y-1)}{(s-1)^2}\right) \\ &= \log_s\left[1 + (s-1)\tfrac{(s^x-1)(s^y-1)}{(s-1)^2}\right] \end{aligned}$$

so

$$T_w(x,y) = \log_s\left[1 + \frac{(s^x-1)(s^y-1)}{s-1}\right]. \tag{12}$$

Remark 11 *Yager class of fuzzy intersections [110] among some other intersections like Dombi (1982) [16]covers the whole range of fuzzy t-norms.*

By choosing different pairs of t-norms T and associated t-conorms $S(x,y) = 1 - T(1-x, 1-y)$, for all $x, y \in [0,1]$ one can derive distinct fuzzy set systems.

Example 12 *The following are examples of some t-norms that are frequently used as fuzzy conjunctions (each defined for all $x, y \in [0,1]$).*

Table 1: Examples of t-norms

Name	**t-norm**
Minimum (largest)	$T_{min}(x,y) = \min(x,y)$
Algebraic product	$T_{Aprod}(x,y) = xy$
Bounded product	$T_{Ł}(x,y) = \max(0, x+y-1)$
Drastic product (smallest)	$T_{Dprod}(x,y) = \begin{cases} x, \text{ when } y=1 \\ y, \text{ when } x=1 \\ 0, \text{ otherwise} \end{cases}$
Hamacher $p \geq 0$	$T_H(x,y) = \frac{xy}{p+(1-p)(x+y-xy)}$
Dombi $p > 0$ **(1982) [16]**	$T_D(x,y) = \frac{1}{1+\left(\left(\frac{1}{x}-1\right)^p+\left(\frac{1}{y}-1\right)^p\right)^{\frac{1}{p}}}$
Frank $p > 0$**,** $p \neq 1$ **(1979) [29]**	$T_F(x,y) = \log_p\left(1+\frac{(p^x-1)(p^y-1)}{p-1}\right)$
Yager $p \geq 1$ **(1980) [110]**	$T_Y(x,y) = 1-\left(1 \wedge ((1-x)^p+(1-y)^p)^{\frac{1}{p}}\right)$
Schweizer & Sklar's 1^{st} $p > 0$ **(1963) [95]**	$T_{SS}^1(x,y) = (0 \vee (x^p+y^p-1))^{\frac{1}{p}}$
Yu $p > -1$ **(1985) [114]**	$T_{Yu}(x,y) = \max(0, (1+p)(x+y-1)-pxy)$

Example 13 *The following are examples of associated t-conorms that are frequently used as fuzzy disjunctions (each defined for all* $x, y \in [0,1]$*).*

Below is the table of properties of some t-norms and t-conorms that are used in this thesis.

Norm	Archimedean	continuous	strict
T_{min}, S_{max}	no	yes	no
$T_{Ł}$, $S_{Ł}$	yes	yes	no
T_{Aprod}, S_{Asum}	yes	yes	yes
T_D, S_D	yes	yes	yes
T_Y, S_Y	yes	yes	no
T_F, S_F	yes	yes	yes
T_{SS}^1, S_{SS}^1	yes	yes	no
T_{Yu}, S_{Yu}	yes	yes	no

Below are limits of some t-norms and t-conorms that are used in this thesis. These show limits for the valuations that comparison measures created from these t-norms and t-conorms can get.

1. Dombi:

$\lim\limits_{p \to 0} T_D = T_{Dprod}$ and $\lim\limits_{p \to 0} S_D = S_{Dsum}$
$\lim\limits_{p \to 1} T_D = \lim\limits_{p \to 0} T_H$ and $\lim\limits_{p \to 1} S_D = \lim\limits_{p \to 0} S_H$
$\lim\limits_{p \to \infty} T_D = T_{min}$ and $\lim\limits_{p \to \infty} S_D = S_{max}$

Table 2: Examples of t-conorms

Name	**t-conorm**
Maximum (smallest)	$S_{max}(x,y) = \max(x,y)$
Algebraic sum	$S_{Asum}(x,y) = x + y - xy$
Bounded sum	$S_{\text{L}}(x,y) = \min(1, x+y)$
Drastic sum (largest)	$S_{Dsum}(x,y) = \begin{cases} x, \text{ when } y=0 \\ y, \text{ when } x=0 \\ 1, \text{ otherwise} \end{cases}$
Hamacher $p \geq 0$	$S_H(x,y) = \frac{x+y-xy-(1-p)xy}{1-(1-p)xy}$
Dombi $p > 0$ **(1982) [16]**	$S_D(x,y) = \frac{1}{1+\left(\left(\frac{1}{x}-1\right)^{-p}+\left(\frac{1}{y}-1\right)^{-p}\right)^{-\frac{1}{p}}}$
Frank $p > 0$**,** $p \neq 1$ **(1979) [29]**	$S_F(x,y) = 1 - \log_p\left(1 + \frac{\left(p^{1-x}-1\right)\left(p^{1-y}-1\right)}{p-1}\right)$
Yager $p \geq 1$ **(1980) [110]**	$S_Y(x,y) = 1 \wedge (x^p + y^p)^{\frac{1}{p}}$
Schweizer & Sklar's 1^{st} $p > 0$ **(1963) [95]**	$S_{SS}^1(x,y) = 1 - (0 \vee ((1-x)^p + (1-y)^p - 1))^{\frac{1}{p}}$
Yu $p > -1$ **(1985) [114]**	$S_{Yu}(x,y) = \min(1, x+y+pxy)$

2. Yager:

$\lim_{p\to 0} T_Y = T_{Dprod}$ and $\lim_{p\to 0} S_Y = S_{Dsum}$
$\lim_{p\to 1} T_Y = T_{\text{L}}$ and $\lim_{p\to 1} S_Y = S_{\text{L}}$
$\lim_{p\to\infty} T_Y = T_{min}$ and $\lim_{p\to\infty} S_Y = S_{max}$

3. Frank:

$\lim_{p\to 0} T_F = T_{min}$ and $\lim_{p\to 0} S_F = S_{max}$
$\lim_{p\to 1} T_F = T_{Aprod}$ and $\lim_{p\to 1} S_F = S_{Asum}$
$\lim_{p\to\infty} T_F = T_{\text{L}}$ and $\lim_{p\to\infty} S_F = S_{\text{L}}$

4. SS1:

$\lim_{p\to 0} T_{SS}^1 = T_{Aprod}$ and $\lim_{p\to 0} S_{SS}^1 = S_{Asum}$
$\lim_{p\to 1} T_{SS}^1 = T_{\text{L}}$ and $\lim_{p\to 1} S_{SS}^1 = S_{\text{L}}$
$\lim_{p\to\infty} T_{SS}^1 = T_{Dprod}$ and $\lim_{p\to\infty} S_{SS}^1 = S_{Dsum}$

5. Yu:

$\lim_{p\to -1} T_{Yu} = T_{Aprod}$ and $\lim_{p\to -1} S_{Yu} = S_{Asum}$
$T_0 = T_{\text{L}}$ and $S_0 = S_{\text{L}}$
$\lim_{p\to\infty} T_{Yu} = T_{Dprod}$ and $\lim_{p\to\infty} S_{Yu} = S_{Dsum}$

Almost every disjunction or conjunction, many-valued or crisp, compensates the values and gives a compensated valuation of these values. If many compensated valuations are combined by a generalized mean and then added into every valuation of its own weight it reaches very flexible and adaptive combined measures, which can be used for comparison as will be seen later in this thesis.

1.3 Aggregation

Aggregation operators are used to combine various degrees of membership into one numerical value, or, more generally, several fuzzy sets are combined in a desirable way to produce a single fuzzy set. An important class of aggregation operators are averaging operators.

Definition 14 *A mapping $f : [0,1]^n \to [0,1]$ is called an aggregation function f if it fulfills the following conditions*

a) $f(0,0,...,0) = 0$ *and* $f(1,1,...,1) = 1$ *(boundary conditions).*

b) For any pair of n-tuples $\langle a_1, a_2, \ldots, a_n \rangle$ and $\langle b_1, b_2, \ldots, b_n \rangle$ such that $a_i, b_i \in [0,1]$ for all $i \in \mathbb{N}$, if $a_i \le b_i$ for all $i \in \mathbb{N}$, then

$$f(a_1, a_2, \ldots, a_n) \le f(b_1, b_2, \ldots, b_n);$$

that is, f is monotonic increasing in all its arguments.

c) f is a continuous function.

Two important additional requirements are the following ones:

d) f is a symmetric function in all its arguments; that is,

$$f(a_1, a_2, \ldots, a_n) = f\left(a_{p(1)}, a_{p(2)}, \ldots, a_{p(n)}\right)$$

for any permutation p on N_n.

e) f is an idempotent function; that is,

$$f(a, a, \ldots, a) = a$$

for all $a \in [0,1]$.

Aggregation operations that are idempotent are called *averaging operations.* For more details see [16], [18] and [26]

The following definition is for aggregation operators in general [16]:

Definition 15 *For the aggregation of a number n of arguments, it holds that*

$$A\left(a_1,\ldots,a_n\right)=f^{-1}\left(\sum_{i=1}^{n}f\left(a_i\right)\right), \tag{13}$$

where a_i denotes arguments and f is a generator function as defined in [16].

Thus, if a generator function is chosen to be $f\left(a_i\right)=a_i^m$, where $m\in\mathbb{R}$ and $m\neq 0$, which is invertible, the following form for aggregation of number n of arguments a_i is reached:

$$A\left(a_1,\ldots,a_n\right)=\left(\sum_{i=1}^{n}a_i^m\right)^{\frac{1}{m}}. \tag{14}$$

Obviously, this is the same as taking the L_p-norm from the arguments. This can be transformed by generalization into the weighted form of the generalized mean [37].

Definition 16 *The weighted form of the generalized mean operator of dimension n, $m\neq 0$ and $\sum_{i=1}^{n}w_i=1$ is in the form*

$$A\left(a_1,\ldots,a_n\right)=\left(\sum_{i=1}^{n}w_i a_i^m\right)^{\frac{1}{m}}, \tag{15}$$

where a_i denotes arguments, w_i denotes weights and m denotes the parameter used to define the mean.

The formula above corresponds to the formulation found in [37, 21].

Lemma 17 *By setting $w_i=\frac{1}{n}$, $\forall i\in\mathbb{N}$ it is reached normal formulation of the generalized mean.*

From the [37, 21, 113] one can find that the weighted form of the generalized mean defined in (15) is an averaging operator. It also possesses several other good properties listed in [21].

A highly compelling feature of the fuzzy set theory is to provide categories for the sets of measures called aggregation connectives [18, 26]. In the papers [82, 83, 84, 85] I used a generalized mean as the aggregative operator in order to compensate and combine values. These articles introduced several new generalized measures starting from t-norms, t-conorms, 3Π-uninorm and pseudo equivalences based on the use of

S-implications. These measures are aggregated with a generalized mean and called Generalized Weighted T-norm measure in (44) and (47), also a new parameterized 3Π-operator is given in (61). All along several new comparison measures based on S-type pseudo equivalences have been created and presented in (36), (39), (42) and (43).

The generalized mean [37, 21] or quasilinear mean [18] is an aggregation operator that belongs to the class called averaging or mean operators. Other well-known averaging operators include OWA operators [111]. The name averaging operator originates from the fact that it combines arguments by giving them some kind of compensation value. Averaging operators can also be considered as extending the space of the universal quantifier $\forall$ (for all) and $\exists$ (at least one) from the pair $\forall, \exists$ to the interval $[\forall, \exists]$. More recently, the generalized mean has been implemented in OWA operators in [113] to obtain a generalized version of these operators. This GOWA operator seems to be an important special case of the use of the generalized mean.

Definition 18 *Grade of compensation: Since the generalized mean increases monotonically with respect to m, the grade of compensation is achieved by any strictly monotone increasing transformation γ of the compensation parameter m from $[-\infty, \infty]$ onto $[0, 1]$.*

Example 19 *The grade of compensation can be achieved for instance by*

$$\gamma = 0.5\left(1 + \frac{m}{1+|m|}\right)$$

or

$$\gamma = 0.5\left(1 + \frac{2}{\pi}\arctan(m)\right).$$

Both yield:

$$\begin{array}{lll} \gamma = 0.00 & \textit{for minimum} & m \to -\infty \\ \gamma = 0.25 & \textit{for harmonic mean} & m = -1 \\ \gamma = 0.50 & \textit{for geometric mean} & m = 0 \\ \gamma = 0.75 & \textit{for arithmetic mean} & m = 1 \\ \gamma = 1.00 & \textit{for maximum} & m \to \infty, \end{array}$$

where $\gamma = 0$ and $\gamma = 1$ characterize min and max type of operators [115], which are only t-norms and t-conorms that are idempotent and distributive [7]. From this it is seen that t-norms and t-conorms provide a lower and upper bound for averaging operators.

Compensative property clearly means that the generalized mean is a valuable tool for the combining values. It also provides more freedom than the use of well-known arithmetic, geometric or harmonic means. The coming chapters illustrate how the generalized mean has been used for combining operators such as t-norms, t-conorms, uninorms, S-type implications and many valued equivalences.

1.4 Uninorms

Uninorms are an important generalization of triangular norms and conorms, having a neutral element lying anywhere in the unit interval. József Dombi introduced aggregative operators in 1982 [16], which form a class of representable uninorms. Ronald Yager first used the term uninorms in 1996 [112]. Uninorms are very closely related to the t-norms and t-conorms because of their neutral element. The main difference between fuzzy intersections and unions and uninorms is that the latter have a neutral element e lying in the interval between 0 and 1. The case $e = 1$ leads back to t-norms and the case $e = 0$ leads back to t-conorms. For more details, see [25].

Definition 20 *A mapping $U : [0,1]^2 \to [0,1]$ is called a uninorm if it is a commutative, monotonic and associative operator that satisfies*

$$(\exists e \in [0,1])\,(\forall x \in [0,1])\,(U(e,x) = x) \tag{16}$$

element e is unique and is called the neutral element of U.

1.5 Fuzzy Implications

Implications $x \to y$ are naturally suitable for decision-making.

Remark 21 *In general, a fuzzy implication I is a function of the form*

$$I : [0,1] \times [0,1] \to [0,1]$$

Which defines the valuation of the conditional proposition "if x then y" for all x and y.

In classical logic $x, y \in \{0,1\}$, implications can be defined in several distinct forms. While these forms are equivalent in classical logic, their extension into fuzzy logic are not equivalent and result in distinct classes of fuzzy implications. This fact makes the concept of fuzzy implication somewhat complicated and several definitions for fuzzy implications have been given, see for example, [23], [48]. There is however a consensus that every fuzzy implication should be an extension of the classical implication. This means that they should at least fulfill the following $I : [0,1] \times [0,1] \to [0,1]$ is a fuzzy implication if and only if $I(1,1) = I(0,1) = I(0,0) = 1$ and $I(1,0) = 0$. One can see that fuzzy implications are obtained by generalizing the implication operator of classical logic. That is, they collapse to the classical implication when truth values are restricted to 0 and 1.

Identifying various properties of the classical implication and generalizing them appropriately leads to the following acceptable properties (compare for example [23], [48]):

Properties 22 *a) If $x \leq z$ then $I(x,y) \geq I(z,y)$ (monotonicity in first argument).*

b) If $y \leq z$ then $I(x,y) \leq I(x,z)$ (monotonicity in second argument).

c) $I(0,y) = 1$ (dominance of falsity).

d) $I(x,1) = 1$ (neutrality of truth).

e) $I(x,x) = 1$ (identity).

f) $I(x, I(y,z)) = I(y, I(x,z))$ (exchange principle).

g) If $x \leq y$ then $I(x,y) = 1$ (boundary condition).

h) $I(x,y) \geq y$

i) I is a continuous function (continuity).

j) $I(x,y) = I(\neg(y), \neg(x))$. This is valid when $\neg$ is a standard fuzzy negation.

Definition 23 *Fuzzy implications that results from the use of the formula $x \rightarrow y \equiv S(\neg x, y)$, define an implication class called S-implications. Here $\neg$ denotes a standard fuzzy negation that is of the form $1 - x$, and S denotes used t-conorm.*

Example 24 *Kleene-Dienes, Reichenbach, Łukasiewicz, the largest S-implications are basic S-implications. The order of these implications are Kleene-Dienes $\leq$ Reichenbach $\leq$ Łukasiewicz $\leq$ Largest S-implication. Later in this thesis, the use of these type of implications as the measures for comparison is demonstrated.*

Definition 25 *Fuzzy implications that results from the use formula $x \rightarrow y \equiv \sup\{z \in [0,1] \,|T(x,z) \leq y\}$ are normally referred to in the literature as a class called R-implications.*

There are papers published in 19^{th} century that are in some degree relevant to the many-valued logic like H. McColl [59] and C.S. Peirce [70]. However it is the opinion of many authors such as Petr Hájek [35] that Jan Łukasiewicz [53] was the first one who investigated systematically many-valued logics in the 1920's. In 1935, Morchaj Wajsberg showed that infinite valued sentential logic was complete with respect to the axioms conjectured by Łukasiewicz. Wajsberg's proof has never been published. However, in 1958 A. Rose and J.B. Rosser [72] were the first ones to prove that Łukasiewicz propositional calculus was complete and in the same year C.C. Chang introduced MV-algebras [10] which allows for another completeness proof of Łukasiewicz's logic. In 1979 Jan Pavelka [68] published a paper in which he generalized Łukasiewicz's logic.

1.6 Fuzzy Equivalence Relations and their Connection with Pseudo-Metrics

Definition 26 ***Pseudo-metric****, on a set X, is a mapping $d : X \times X \to [0, \infty[$ such that $\forall x, y, z \in X$ the following conditions holds true:*

1. $d(x, x) = 0$
2. $d(x, y) = d(y, x)$
3. $d(x, z) \leq d(x, y) + d(y, z)$

Equivalence is naturally suitable for the comparison of different objects. In soft computing, there is a variety of measures that have been applied to comparison. Some of these measures are set-theoretical, distance based, logical or heuristical. Similarity relations are used many times as synonyms for measures, which are used for comparison. The well-known definition for similarity relations in the field of fuzzy logic is the one presented by Lotfi Zadeh [116]. There is a close link between this notion of similarity and that of distance (see for example [27] and [104]). This definition is a fuzzy logical generalization of the equivalence relation, and it is also an inverse pseudo-metric. The definition given by Zadeh [116] coincides with the following one:

Definition 27 *A mapping $E : X^2 \to [0, 1]$ is called fuzzy equivalence relation with respect to t-norm T or just T-equivalence if it fulfills the following conditions, where $x, y, z \in [0, 1]$, $\mathbf{0} = minE\langle x, y\rangle$ and $\mathbf{1} = maxE\langle x, y\rangle$:*

1. $\forall x \in X : E\langle x, x\rangle = \mathbf{1}$ *or* $\forall x \in X : E\langle x, \neg x\rangle = \mathbf{0}$
2. $\forall x, y \in X : E\langle x, y\rangle = E\langle y, x\rangle$
3. $\forall x, y, z \in X : T\langle E\langle x, y\rangle, E\langle y, z\rangle\rangle \leq E\langle x, z\rangle$

E is also called a *similarity relation*, a *fuzzy equivalence* [45], *an indistinguishability operator* [100], *fuzzy equality* [39] or a *proximity relation* [19]. In the case of the definition (27), it is assumed that the similarity relation used is reflexive, symmetric and t-transitive. By replacing the set $[0,1]$ with the two-element set $\{0,1\}$, the similarity relation (27) coincides with the usual equivalence relation.

It is known that pseudo-metrics bounded by one and fuzzy equivalence relations with respect to the bounded product $\max\{a+b-1,0\}$ are dual concepts [45], which means that E is a fuzzy equivalence relation on X with respect to a bounded product if and only if $1-E$ is a pseudo-metric in X. So fuzzy equivalence $S\langle x,y\rangle$ now acquires the form $E\langle x,y\rangle = 1-d(x,y)$. In a case where the pseudo-metric d is not bounded by one, this property is enforced by considering the pseudo-metric $\hat{d} = \min\{d(x,y),1\}$, which coincides with d for 'small' distances. Thus any pseudo-metric induces an equivalence relation of X with respect to the bounded product by

$$E\langle x,y\rangle = 1-\hat{d} = 1-\min\{d(x,y),1\}. \tag{17}$$

The following theorem that was presented by Siegfried Gottwald in [32] shows that t-norms used has to be bigger or equal than the bounded product in order to maintain pseudo-metricity in fuzzy equivalence. In his results, Gottwald used fuzzy equivalences $T((a\to b),(b\to a))$ where implications are of the type R look definition (25).

Theorem 28 *$d\langle x,y\rangle = 1-E\langle x,y\rangle$ is a pseudo-metric if and only if used t-norm T in $T((a\to b),(b\to a))$ is larger or equal to the bounded product T_L.*

Proof. The proof is presented in [32]. ■

From the equation $E\langle x,y\rangle = T((x\to y),(y\to x))$ one can see that created logical equivalences depend on the used implication as well as used t-norms. It is also known that some R- and S-implications (25), (23) overlap. This is the case for example with the so called Łukasiewicz implication which is the same in both of its R- and S-forms. Use of certain left continuous t-norms called nilpotent minimum which satisfy conditions given in [24] also leads to the implications where S- and R-forms always overlap.

It is well-known that sometimes the demands of reflexivity, symmetricity or t-transitivity can be unsuitable for the problems or data sets at hand [101, 92, 14, 41]. It is over-simplyfing to consider the similarity relation as defined in the definition (27) as the only measure for sameness or comparison. The meaning of sameness is wider than the definition of this pseudo-metric based similarity relation. For example, in this thesis logical equivalences are created which are not reflexive. Symmetricity and triangular inequality have also been questioned in for example [102], and sometimes non-symmetric measures are called directed similarities [28]. Tversky has especially shown that measures of similarity that conform to human perception do not satisfy the usual properties of a metric [101, 102].

In general, Mathematical machinery used in engineering leans heavily on the use of metrics. An example of this would be the embedding of signals of interest in metric spaces, after which it is possible to measure the similarity or dissimilarity of signals, fidelity and dissortion, as distance. In practice it is always an application that decides the choice of an axiom-set and the flavor of the mathematics used. There are, however, many metric spaces that can be used to measure distance for example:

Manhattan:

$$d(x,y) = \sum_{i=1}^{N} |x_i - y_i| \tag{18}$$

Euclidean:

$$d(x,y) = \sqrt{\sum_{i=1}^{N} (x_i - y_i)^2} \tag{19}$$

Cumulative Euclidean:

$$d(x,y) = \sqrt{\sum_{i=1}^{N} \left(\sum_{u=1}^{i} x_u - \sum_{u=1}^{i} y_u \right)^2} \tag{20}$$

Minkowsky:

$$d(x,y) = \left(\sum_{i=1}^{N} |x_i - y_i|^p \right)^{\frac{1}{p}} \tag{21}$$

Cumulative Minkowsky:

$$d(x,y) = \left(\sum_{i=1}^{N} \left| \sum_{u=1}^{i} x_u - \sum_{u=1}^{i} y_u \right|^p \right)^{\frac{1}{p}} \tag{22}$$

Landmover:

$$d(x,y) = \sum_{i=1}^{N} \left| \sum_{u=1}^{i} x_u - \sum_{u=1}^{i} y_u \right| \tag{23}$$

Mahalanobis:

$$d(x,y) = (x-y)^T C^{-1} (x-y), \tag{24}$$

where C is the covariance matrix.

Modified G:

$$d(x,y) = 2\left\{\left[\sum_{f=x,y}\sum_{i=1}^{N} f_i \log f_i\right] - \left[\sum_{f=x,y}\left(\sum_{i=1}^{N} f_i\right)\log\left(\sum_{i=1}^{N} f_i\right)\right] - \left[\sum_{i=1}^{N}\left(\sum_{f=x,y} f_i\right)\log\left(\sum_{f=x,y} f_i\right)\right] + \left[\left(\sum_{f=x,y}\sum_{i=1}^{N} f_i\right)\log\sum_{f=x,y}\sum_{i=1}^{N} f_i\right]\right\} \tag{25}$$

Of the commonly used distance functions the most common ones are Euclidean corresponding to the norm L_2 and Manhattan corresponding to the norm L_1, these are both special cases of the Minkowsky, the norm of which is L_p. As the degree of the norm (p) increases, the weight of a large difference between values of a single attributes increases. Thus, the Euclidean distance weights have large differences in comparison with the Manhattan distance. Both the Euclidean and Manhattan distances are calculated separately for each dimension, and thus, they are not very good measures of similarity between objects where attributes are correlated and ordered. In this kind of situation The Cumulative Euclidean and Landmover distances can be used, because they measure the difference of cumulative values of attribute vectors. Thus the Cumulative Euclidean and Landmover distances can be considered to measure the spatial concentration of the values in the attribute vector and the order of the attributes affects the value of the distance. Therefore, with ordered data these measures are likely to provide better results than the standard Euclidean and Manhattan distances.

To prevent the effects of unbalanced ranges of different dimensions, the attributes are often normalized. The normalization can be performed by dividing the attribute values by the range (maximum-minimum) of the corresponding attribute. Another possibility is to perform the normalization using the standard deviation instead of the range.

Most of these distance functions, performed previously, consider the attributes to be non-correlated. For cross-correlated attributes, statistical properties of a data set can be used to reduce the effect of correlations. One such distance function with a statistical factor, the correlation matrix is the Mahalanobis distance. Eventually, calculation of the correlation matrix in the Mahalanobis distance needs considerable data, the exact amount depending on the length of the attribute vectors and the variation between them. If there is not enough variation in the data, numerical calculation of the inverse of covariance matrix may be impossible because of the singularity of the matrix (ill-posed problem). A comprehensive study of the Mahalanobis distance with a limited sample size can be found in [99]. Furthermore, the removal of the correlation effect should be computed among the samples of the same class rather than over samples in all classes. Such a calculation procedure is again problematic if the classes are unknown.

Another statistical method, the log likelihood ratio G, measures the degree to which observed data corresponds to an expected distribution. As a distance function the basic G metric has the drawback of depending on the order of variables (non-symmetric metric), i.e., which variable is considered to be the sample and which the expected distribution. Sokal and Rohlf [98] introduced a modified G test, which is a two-way test of interaction or heterogeneity.

With statistical metrics such as Mahalanobis distance and Modified G, there are some limitations and they cannot always be used. With the Mahalanobis distance, the covariance matrix is calculated and this cannot always be done because of the singularity of the matrix. With Modified G there is a logarithm involved and this causes problems when values are close to zero. These kinds of metrics have been tested in, for example, [76], [56], and [80], but the results have never been promising.

1.7 Lattices, Algebras and Similarities

To commence with, these are the underlying definitions.

Definition 29 *(Cartesian Product Space). Let* (x, y) *be an* ordered pair*, where* $x \in X$ *and* $y \in Y$*, the* Cartesian product *is defined as the set:*

$$X \times Y = \{(x, y) : x \in X,\ \ y \in Y\} \tag{26}$$

Definition 30 *(Binary relation). Any subset* $R \subseteq X \times Y$ *defines a binary relation between the elements of* X *and* Y *:*

$$R = \{(x, y) \in X \times Y : R\,(x, y)\, holds\} \tag{27}$$

A relation is a multi-valued correspondence:

$$\begin{aligned} R : X \times Y &\rightarrow \{0, 1\} \\ (x, y) &\rightarrow R\,(x, y\} \end{aligned} \tag{28}$$

Definition 31 *A* binary relation *is a* quasi-order *if it is* reflexive *and* transitive. *If it is also* anti-symmetric *then* binary relation *is* partial order. *If* quasi-order *is* symmetric *it is an* equivalence relation.

Definition 32 *A partially ordered set or poset is a set L on which an order relation $\leq$ has been defined. Of course, on a set L various order relations can be defined. If in a poset L either $x \leq y$ or $y \leq x$ for each $x, y \in L$, then L is linear and is called a chain. In such a case, the order $\leq$ is a total order and L is referred to as linearly ordered.*

Definition 33 *A **lattice** is a poset L such that for any $x, y \in L$, $x \cap y$ and $x \cup y$ exits in L. $x \cap y$ is called meet, $x \cup y$ is called join of x and y. A lattice L is a (countable) complete lattice if $\bigcup \{x \mid x \in X\}$ and $\bigcap \{x \mid x \in X\}$ exist in L for any (countable) subset $X \subseteq L$. A lattice is often denoted by $\langle L, \leq, \cap, \cup \rangle$.*

Example 34 *The unit interval I, for example, is a complete lattice under the usual order of $x \cup y = \max\{x, y\}$ and $x \cap y = \min\{x, y\}$.*

Definition 35 *A lattice is called **residuated** if it contains the greatest element $\mathbf{1}$, and binary operations $\odot$ (**continuous t-norm**) and $\rightarrow$ (called **residuum**) such that the following conditions hold true*

1. $\odot$ is associative, commutative and isotone
2. $a \odot \mathbf{1} = a$ for all elements $a \in L$ and
3. for all elements $a, b, c \in L$, $a \odot c \leq b$ if and only if $c \leq (a \rightarrow b)$.

Definition 36 *Basic Logic (BL) algebra is a residuated lattice $\langle L, \leq, \wedge, \vee, \odot, \rightarrow, 0, 1 \rangle$, where $\forall x, y \in L$ holds:*

1. $x \wedge y = x \odot (x \rightarrow y)$
2. $x \vee y = [(x \rightarrow y) \rightarrow y] \wedge [(y \rightarrow x) \rightarrow x]$
3. $(x \rightarrow y) \vee (y \rightarrow x) = \mathbf{1}$

Remark 37 *Each continuous t-norm determines a Basic Logic algebra in the unit interval* $[0,1]$ *with its standard linear ordering as Hájek has shown in [35].*

The following demonstrates how to use similarity in order to find similar pairs. Here a chosen situation is examined where features of different objects can be expressed in values between [0,1]. Let X be the set of m objects. If the similarity value of the features are known $f_1, ..., f_n$ between objects, the object can be chosen that has the highest total similarity value. The problem is to find for object x_i a similar object x_j, where $1 \leq i, j \leq m$ and $i \neq j$. By choosing for example Łukasiewicz-structure for features of the objects n similarities are achieved for comparing the two objects (x_1, x_2)

$$S_{f_i}\langle x_1, x_2\rangle = E(x_1(f_i), x_2(f_i)), \tag{29}$$

where $x_1, x_2 \in X$ and $i \in \{1, ..., n\}$. Because Łukasiewicz-structure is chosen for the membership of objects, the similarity can be defined as follows

$$S\langle x_1, x_2\rangle = \frac{1}{n}\sum_{i=1}^{n} E(x_1(f_i), x_2(f_i)). \tag{30}$$

Different non-zero weights $(W_1, .., W_n)$ can also be given to the different features in order to obtain the following formula, which again meets the definition of the similarity.

$$S\langle x_1, x_2\rangle = \frac{\sum_{i=1}^{n} W_i E(x_1(f_i), x_2(f_i))}{\sum_{i=1}^{n} W_i}. \tag{31}$$

In the ordinary Łukasiewicz-structure equivalence relation $E(x, y)$ as well as similarity $S(x, y)$ is defined as

$$E(x, y) = 1 - |x - y| = S(x, y). \tag{32}$$

In the case of the so called generalized Łukasiewicz-structure [48], the equivalence relation (or similarity in case of Łukasiewicz) is more complicated, i.e.

$$E(x, y) = (1 - |x^p - y^p|)^{\frac{1}{p}} = S(x, y). \tag{33}$$

This similarity has a clear connection with Minkowsky-metrics.

Lemma 38 *Consider* n *Lukasiewicz valued fuzzy similarities* S_i, $i = 1, ..., n$ *on a set* X. *Then* $\cdots$

$$S\langle x, y\rangle = \frac{1}{n}\sum_{i=1}^{n} S_i\langle x, y\rangle$$

is a Lukasiewicz valued similarity on X.

Proof. As all S_i, $i = 1, ..., n$ are reflexive and symmetric consequently S is also. The transitivity of S can be seen in the following way. Let $A = S\langle x, y\rangle \odot S\langle y, z\rangle$.

Then

$$A = \left(\frac{1}{n}\sum_{i=1}^{n} S_i \langle x, y\rangle\right) \odot \left(\frac{1}{n}\sum_{i=1}^{n} S_i \langle y, z\rangle\right)$$
$$= \frac{1}{n}\sum_{i=1}^{n} S_i \langle x, y\rangle + \frac{1}{n}\sum_{i=1}^{n} S_i \langle y, z\rangle - 1$$
$$= \frac{1}{n}\left(\sum_{i=1}^{n} S_i \langle x, y\rangle + \sum_{i=1}^{n} S_i \langle y, z\rangle - n\right)$$
$$= \frac{1}{n}\left[(S_1 \langle x, y\rangle + S_1 \langle y, z\rangle - 1) + \cdots + (S_n \langle x, y\rangle + S_n \langle y, z\rangle - 1)\right]$$
$$= \frac{1}{n}\left[(S_1 \langle x, y\rangle \odot S_1 \langle y, z\rangle) + \cdots + (S_n \langle x, y\rangle \odot S_n \langle y, z\rangle)\right]$$
$$\leq (S_1 \langle x, z\rangle + \cdots + S_n \langle x, z\rangle)$$
$$= S \langle x, z\rangle \quad \blacksquare$$

Since generalized mean is a monotonically increasing aggregation operator the following result can be concluded::

Corollary 39 *Consider n Łukasiewicz valued fuzzy similarities S_i, $i = 1, ..., n$ on a set X. Then*

$$S \langle x, y\rangle = \left[\frac{1}{n}\sum_{i=1}^{n} (S_i \langle x, y\rangle)^p\right]^{\frac{1}{p}}$$

is a Łukasiewicz valued similarity on X.

1.8 Classification

Since classification is used to test all comparison measures in this thesis then a description of this classification procedure and data sets are given first.

Much of the fuzzy set theory's original inspiration and further developments originate from the problems of pattern classification and cluster analysis. Essentially, this is the reason why classification is chosen to be the test bench for many valued logic based comparison measures in this thesis. In classification, the question is not whether a given object is or is not a member of a class, but the degree to which the object belongs to the class. This means that most classes in real situations are fuzzy in nature [117]. This fuzzy nature of real world classification problems may shed some light on the general problem of decision making [66].

Many times a given set of data is already grouped into classes, and the problem is then to predict to which class the new data items belong. This is normally referred to as a classification problem. The first set of data is referred to as a training set, while this new set of data is referred to as a test set [36]. Classification is seen here as a comparison between the training set and test set. A classification procedure demonstrated here could be categorized as a supervised learning method [20] or instance based learning.

Classification procedure could also be categorized as locally weighted learning, that uses weighted learning to average, interpolate between, extrapolate from, or otherwise combine training data [105]. This can be seen from the following definition [1]. Lazy learning method means (defers) processing of training data until a query needs to be answered. When training data is stored into the memory and relevant data in the database is found to answer a particular query, lazy learning is sometimes referred to as memory-based learning. Relevance is often measured using some distance function, with nearby points having high relevance. Since the query to be answered is known during processing of training data, training query-specific local models is possible in lazy learning. Weighted average local models average the outputs of nearby points, inversely weighted by their distance to the query point and locally weighted regression fits with a surface on nearby points using a distance weighted regression. Automatic tuning of the learning algorithm's parameters to specific tasks or data sets is part of the previous methods. Therefore, this classification model could be called something like similitude based weighted averaging, based on minimizing difference between test data and learning data. It is important to note the following two facts [1]:

1. Local learning is critically dependant on the distance function. Distance function in locally weighted learning does not need to satisfy the formal mathematical requirements for a distance metric. The relative importance of the input dimensions in generating the distance measurement depends on how the inputs are scaled (normalized).
2. Distance functions can be asymmetric and nonlinear, so that a distance along a particular dimension can depend on whether the query point's value for the dimension is larger or smaller than the stored point's value for that dimension. The distance along the dimension can also depend on the values being compared.

Measures have been tested in this thesis with six different data sets which are available from [103]. The data sets chosen for the test were: Ionosphere, Iris, Pima, Post, Thyroid gland and Wine. These are derived from the fields of medicine, biology, geology and engineering. These learning sets differ greatly in the magnitude of instances and number of predictive attribute values. By classification several new generalized measures will be tested, such as The Generalized Weighted T-norm measure (44) and (47), a new parameterized 3Π-operator (61) and several new measures based on S-type pseudo equivalences were also tested (36), (39), (42) and (43).

The stability of the measures has been tested by classification for different weight values. Each parameter value p and m has been tested and differential evolution is used to find the right weights. Classification has also been tested with weight values which were randomly chosen 200 times to find out how stable these measures really were. At each point the maximum, minimum, mean values and variation has been calculated . Tables of results are appended to this thesis in (3), (4), (5), (6), (7), (8), (9), (10), (11), (12), (13), (14).

1.8.1 Description of the Comparison Measure Based Classifier

Objects, each characterized by one feature vector in $[0,1]^n$, is classified into different classes. The assumption that the vectors belong to $[0,1]^n$ is not restrictive since the appropriate shift and normalization can be done for any space $[a,b]^n$. The comparison measures can be used to compare objects to classes. Below is the used classifier in the algorithmic form:

Require: *data*
  scale *data* between $[0,1]$
Require: *test*,*learn*$[1...n]$,*weights*,*dim*
  for $i = 1$ to n **do**
    $idealvec[i] = IDEAL[learn[i]]$
    $maxcomp[i] = \left(\frac{1}{\dim}\right)^{1/m} \left(\sum_{j=1}^{\dim} weights\,[j]\,(CM\,(idealvec\,[i,j]\,,test\,[j]))^m\right)^{1/m}$
  end for
  $class = \arg\max_i maxcomp[i]$

In the algorithm, the comparison measure CM with a generalized mean is used. $IDEAL[i]$ is the vector that best characterizes the class i and here the generalized mean vector of the class as an $IDEAL$-operator has been used.

Evolutionary algorithm is used because of its diversity and robustness to find weights in classification process, information about evolutionary algorithms in general can be found for example from [31], [60], [63] and [33]. Obviously, other optimizers can be used as well. Evolutionary algorithm is based on differential evolution [71]. Differential Evolution (DE) is a method of mathematical optimization of multidimensional functions and belongs to the class of evolution strategy optimizers.

The classification task can also be described as in the flowchart (1).

One can see that with respect to L, the number of classes, the classification time is $\mathcal{O}(L)$ after the parameters have been fixed and ideal vectors calculated. However,

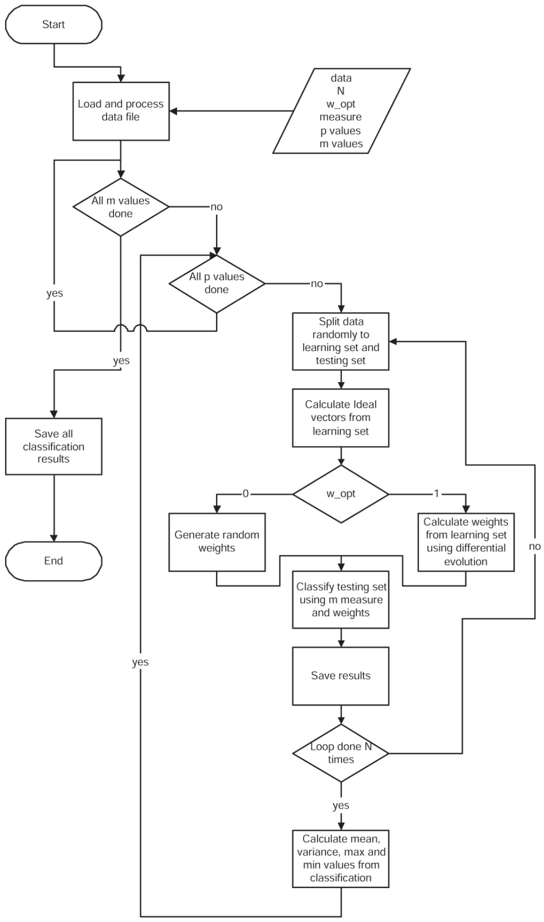

Figure 1: Simplified Flow Chart of the Classification Procedure

finding good parameters can be difficult, but it seems that all measures presented in this thesis are relatively stable with respect to the parameters, which is demonstrated later. The advantage of using this classifier here in this thesis is that its results mainly depend on which comparison measure is chosen to be used.

1.8.2 Data Sets

In the classification tasks presented in my articles [82, 83, 84, 85], the data is in all cases divided into learning sets and test sets. Weights are estimated with differential evolution. The only main difference between classification tasks is the use of different types of measures of comparison.

- **Iris**: Perhaps the best-known database to be found in the pattern recognition literature. The number of attributes is 4 and the class. The data set contains 3 classes of 50 instances each, where each class refers to a type of iris plant.

- **Ionosphere**: This is radar data where the targets were free electrons in the ionosphere. Here are two classes: "Good" and "Bad". "Good" radar returns are those showing evidence of some type of structure in the ionosphere. "Bad" returns are those that do not; their signals pass through the ionosphere. The number of instances is 351. The number of attributes is 34 plus the class attribute.

- **Pima**: The diagnostic, binary-valued variable investigated is whether the patient shows signs of diabetes. All the instances here are females of Pima Indian heritage who are at least 21 years old. The number of Instances is 768. The number of attributes is 8 plus the class. Class 1 (negative for diabetes) 500, Class 2 (positive for diabetes) 268. Reported 76 % classification result.

- **Post Operative**: Task of this database is to determine where patients in a postoperative recovery area should be sent to next. The attributes correspond roughly to body temperature measurements. The number of Instances is 90. The number of attributes is 9 including the decision (class attribute). Attribute 8 has 3 missing values.

- **Thyroid**: Five laboratory tests are used to try to predict whether a patient's thyroid belongs to the class euthyroidism, hypothyroidism or hyperthyroidism. The diagnosis (the class label) was based on a complete medical record, including anamnesis, scan etc. The number of instances is 215. The number of attributes is 5 plus the class. Class 1 (normal) 150, Class 2 (hyper) 35 and Class 3 (hypo) 30.

- **Wine**: The data is the result of a chemical analysis of wines grown in the same region in Italy but derived from three different cultivars. The analysis determined the quantities of 13 constituents found in each of the three types of wines. The number and deviation of instances: class 1 59, class 2 71, class 3 48.

In classification tasks, measures based on t-norms, t-conorms, uninorms, fuzzy implications, fuzzy equivalence relations and fuzzy pseudo equivalence relations, which have been combined with weights and aggregated by a generalized mean, are tested. The results achieved can be seen from the tables (3), (4), (5), (6), (7), (8), (9), (10), (11), (12), (13), (14). Classifications, in all tasks are tested with weights that were randomly selected 200 times (RND) and with weights which were optimized 10 times (DE) for each p- and m-value.

1.9 Short Notations

It is known that Łukasiewicz logic is theoretically the best choice from multi-valued logic used with fuzzy logic in the sense that it possesses the completeness theorem [68] and all the theorems that are true in first-order logic are also true in Łukasiewicz logic [65]. Furthermore, all multi-valued logics, which possess completeness theorem are isomorphic to the Łukasiewics logic [68]. In this sense Pavelka showed that the only good fuzzy implication rule operator in $[0, 1]$, as far as R-implications are concerned, is the Łukasiewics's implication, up to the isomorphism. However there have been reports of bad undecidability of Łukasiewics logic [34].

In this thesis, the approach generally has been highly practical and all usage of measures presented has been validated by practical testing. That is, all comparison measures created have been validated with experimental data. Therefore this thesis can be a valuable tool especially for practitioners. It will concentrate on presenting new comparison measures and a number of ways of creating and using structures, such as the combination of t-norms and t-conorms, the generalized 3Π-uninorm and S-type equivalences created from S-type implications. In all cases, the measures created are implemented with the generalized mean. At the end of this thesis, how these comparison measures have been applied to the classification, and to sports medicine in defining an athlete's aerobic and anaerobic thresholds will be presented. For fuzzification in classification tasks only fuzzy comparison measures created and presented in this thesis are to be used. The only case where values are 'really' fuzzified is when the athlete's anaerobic and aerobic thresholds are estimated.

During the process of writing this thesis the following 25 articles has been published by me [54, 73, 55, 74, 56, 75, 76, 80, 57, 58, 79, 81, 77, 78, 86, 82, 84, 83, 85, 43, 87, 88, 89, 90, 91].

2 Logical Comparison Measures

It is a common belief that measures for comparison should hold true for some properties of metric spaces. This belief originates from the blinkered view that the comparison of objects should always have something to do with distance. This has been questioned in many papers [101, 92, 28, 14, 102]. In practice, it seems that properties of distance have little or no affect at all on the results that can be achieved from the use of different comparison measures. This becomes empirically clear when one looks at the test results presented in this thesis.

Definition 40 *A set function g defined on X, where X is a fuzzy set and has the following properties is called a fuzzy measure:*

1. $g(\emptyset) = 0,\ g(X) = 1$
2. *If* $A, B \in X$ *and* $A \subseteq B$ *then* $g(A) \leq g(B)$
3. *If* $A_n \in B,\ A_1 \subseteq A_2 \subseteq \ldots \subseteq A_{n-1} \subseteq A_n$ *then* $\lim_{n\to\infty} g(A_n) = g\left(\lim_{n\to\infty} A_n\right)$

A general definition is given below of what in this thesis is meant by a comparison measure.

It is suggested here that the comparison measures used in fuzzy sets, where comparison is done feature by feature and then these comparisons are aggregated, could actually be any measures which fulfil the following properties:

Properties 41

1. *The comparison measure used has a clear logical structure e.g. it is an Archimedean t-norm or t-conorm (like Frank (49), (54)) or S-equivalence ((35), (42), (39)).*
2. *The comparison measure is monotone. This condition ensures that a decrease (or increase) in any values that are to be compared cannot produce an increase (or decrease) in the comparison result.*
3. *The comparison measure is associative. This guarantees that the final comparison results are independent of the grouping of the arguments and that one can expand these comparison to more than two arguments.*
4. *The comparison measure is continuous. This guarantees that one can safely compute with the values that are to be compare.*

In this chapter comparison measures are derived starting from the logical structures. Some of the resulting structures can be categorized as metric based comparison measures. Comparison measures based on metric axioms formally called similarities are measures which are reflexive, symmetric and transitive. Basically, they are a type of one minus distance. However most of the measures achieved in this chapter do not fulfil the demands of similarity, and they can be categorized as logic based measures. The usefulness of these measures is validated by practical testing.

The idea behind using logical structures instead of, for example, simple distances lies in the fact that logical structures always have some kind of linguistic content inside them. For example t-norms and t-conorms can be seen as corresponding to the words "and" and "or", equivalence as corresponding to the expression "if and only if". One can see that just by using these logical measures it is possible to give some linguistic meaning to the comparison procedure.

Some criteria for comparison measures are suggested here. The following criteria are almost the same as Lowen gives for aggregation operators [52] and originally they are presented by Bellman-Giertz [7]. It has also been suggested that not all of these criteria are necessary [42]. One can, however, see that the criteria by Bellman R. and Giertz M. also applies well to the comparison measures presented in this thesis.

Criterion 42

1. ***Axiomatic strength***. *It is suggested here that the operator is better if the axioms the operator satisfies are less limiting, this is equivalent to Lowen [52]. It is seen that depending on the choice of the logical structure used this will fit well with the definition given in (41).*

2. ***Flexibility***. *Through the flexibility three things are met that are of an empirical fit, adaptability and compensation. Adaptability comes from the fact that all comparison measures created in this thesis are parameterized. Compensation property follows from the use of a generalized mean to combine the different values. Empirical fit follows then from the three things and these are the use of logical structures, adaptability and compensation. Empirical fit can naturally only finally be proven by empirical testing, as is done in this thesis.*

3. ***Numerical efficiency***. *Some operators such as min and max are numerically more efficient than, for example, Frank's t-norm and t-conorm. In large problems this will always be problematic to some degree. However, it is gradually becoming less of problem as computers computing power is constantly increasing.*

4. ***Range of compensation***. *In general, the larger the range of compensation the better the compensatory operator. In some comparison measures presented in this thesis the range of compensation has been increased by combining t-norms and t-conorms and in all comparison measures a generalized mean has been used.*

5. ***Aggregating behavior of the comparison measure****. Aggregating behavior can in the comparison measures presented here, be adjusted by the use of proper mean value in the generalized mean. For example, if a parameter value of 0 is used with a generalized mean a geometric mean will be obtained, which is to say that one attains the product of the values and subsequently each value "added" normally decreases the resulting aggregate degrees of membership.*

6. ***Required scale level of membership functions****. Comparison measures presented in this thesis have very little restrictions concerning scale levels.*

2.1 Equivalences and Implications as the Measures for Comparison

The logical operation of a many-valued equivalence is commonly used when the comparison of two fuzzy propositions $a, b \in [0, 1]$ are required. Equivalence can then be interpreted to define the valuation of the two-way conditional proposition "a if and only if b". For this reason, it is naturally suitable for the comparison of different objects. The implication can logically be interpreted to define the conditional proposition "if a then b". Implications are naturally suitable for decision-making and therefore they are generally widely used in approximate reasoning. Rule based classifiers are quite popular in classification processes [50] and they are normally used as counterparts for fuzzy control systems.

One way of extending the implication is to first use the classical logic formula $x \to y \equiv \neg x \vee y$ for all $x, y \in \{0, 1\}$. This is done by interpreting the disjunction as a t-conorm and negation by the use of a standard fuzzy complement ($\neg x \equiv 1 - x$). This results in defining the implication with the formula $a \to b \equiv S(\neg a, b)$ for all $a, b \in [0, 1]$, which gives rise to the family of many valued implications called S-implications. Equivalences used in this thesis are of the form $a \leftrightarrow b \equiv T(a \to b, b \to a)$. On the other hand, fuzzy implications that result from the use of the formula $a \to b \equiv \sup\{x \in [0, 1] \,|T(a, x) \leq b\}$ are normally referred to in the literature as a class called R-implications.

Jan Łukasiewicz used only implication and negation, when he studied many-valued logic. However O. Frink Jr. [30] used the term for the first time. "Łukasiewicz arithmetical conjunction" from the $T(a, b) = \max\{0, a + b - 1\}$ is also known as bounded difference. Nowadays, it is common to call bounded difference as Łukasiewicz conjunction or Łukasiewicz t-norm and bounded sum $S(a, b) = \min\{1, a + b\}$ as Łukasiewicz disjunction or Łukasiewicz t-conorm see, for example, Hájek [35], Klir [48], Kundu [49] etc.

Łukasiewicz used implication given by $a \to b = min\{1, 1 - a + b\}$, where $a, b \in [0, 1]$

and negation $\neg a = 1 - a$ as primitive connectives when he created his many valued structure [53]. Jorma K. Mattila has shown that if one uses definitions as Łukasiewicz used one should use only *min* as conjunctions and *max* as disjunctions [61].

Here S-type implications and T-norms are used to create equivalences. The S-type implication is a combination of T-conorm and the negation approach here is reversed to the approach used by Łukasiewicz [53]. It is known that operations, t-norms, t-conorms and negation should be selected very carefully for, example, in the case of Łukasiewics correct 'and' operation for the creation of S-type equivalence is $max\{0, a+b-1\}$ and correct 'or' operation for the creation of S-type is $min\{1, a+b\}$ [35], Klir [48], Kundu [49] etc.

2.1.1 Comparison Measures Based on S-Type Implications

Comparison measures, which arise from the fuzzy set theoretic class of implications called S-implications are now presented. In addition, comparison measures based on the functional form of implications and first presented by Smetz and Magrez, 1987 in [97] are also presented.

Articles [76, 77] studies the use of the Łukasiewicz type of equivalence, with means and weights. In the article [84] this study was taken further by the use of a generalized mean and weights. In the equations presented in [84] implications have been parameterized by replacing the variables in the formulas by their exponential forms. When these kinds of formulations are used one can see that they hold true for all the properties of the corresponding implications, since the exponent is a monotonic operator. In this thesis it will be demonstrated that these new many-valued equivalences and implications, which have been combined with weights and then aggregated with a generalized mean in order to make them comparison measures, are able to give competitive results when they are tested in classification tasks.

The following procedure given is for defining logical equivalences from implications .

1. Firstly take an implication of the type $x \to y \equiv S(\neg x, y)$. Ordinary Łukasiewicz implication $I_{\text{Ł}}(x, y) = \min(1, 1 - x + y)$ is obtained by the use of a bounded sum as a t-conorm. Another way of defining this implication is to employ (close to Galois correspondence) the formula $a \to b \equiv \max\{x \in \{0,1\} \,|a \wedge x \leq b\}$ for all $a, b \in \{0,1\}$ and interpreting the conjunction as a *min*-operator, which gives the original definition of the Łukasiewicz implication. In cases where the conjunction is interpreted as t-norm, this results in defining implications in fuzzy logic by the formula $a \to b \equiv \sup\{x \in [0,1] \,|T(a,x) \leq b\}$ for all $a, b \in [0,1]$ and the bounded difference $\max(0, a+b-1)$ as a t-norm again achieves the Łukasiewicz implication. From this it can be seen that S-type and R-type

Łukasiewicz implications coincide.

2. Secondly a parameterized form is created from the implications chosen by setting variables into exponential forms. For example by setting $x := x^p$ and $y := y^p$ to the ordinary Łukasiewicz implication, leads to the following formula $I_{\text{L}}(x^p, y^p) = \min(1, 1 - x^p + y^p)$.

3. Thirdly equation $x \leftrightarrow y \equiv T((x \rightarrow y), (y \rightarrow x))$ will be used, where T refers to the any t-norm and $\rightarrow$ to any implication. For example, in the case of Łukasiewicz one may choose to use a bounded product as the conjunction which leads to the following equivalence equation $E_{\text{L}}(x, y) = 1 - |x^p - y^p|$.

4. In the fourth step of this operator the weights w_i, $i \in N$ will be added, and these weighted equivalences are combined with the generalized mean m. In the example case here, this leads to the equation $E_{\text{L}}(x_i, y_i) = \left(\sum_{i=1}^{n} w_i \left(1 - |x_i^p - y_i^p|\right)^m\right)^{\frac{1}{m}}$.

It can easily be seen that, in general, this kind of equivalence does not hold true for the definition of similarity, i.e. that they are reflexive, symmetric and transitive. In this approach, the 'right' measure is selected using the metamathematical structure of logic instead of the syntactical structure of mathematics. The second approach which has been used here has been a sort of combination of the syntactic approach originating from mathematics and the semantic approach originating from the many valued logic.

1. First the equation $I(x, y) = f^{-1}(f(1) - f(x) + f(y))$ is used, that maps $I : [0,1]^2 \rightarrow [0,1]$, where $f : [0,1] \rightarrow [0, \infty[$ is a strictly increasing continuous function such that $f(0) = 0$, for all $x, y \in [0,1]$ [97]. The implications that follow from the usage of the previous equation have many important properties, and therefore some of these are mentioned here. The first property is the monotonicity in the first and second argument, which means that the truth value of many valued implications increase if the truth value of the antecedent decreases or the truth value of the consequent increases. The second property is that it is continuous, which is a property that ensures that small changes in the truth values of the antecedent or the consequent do not produce large (discontinuous) changes in the truth values of many valued implications. The third is that it is bounded, which means that many valued implications are true if and only if the consequent is at least as true as the antecedent. For example, one can select $f(x) = x^p$, which leads to the implication of the form $I(x^p, y^p) = \min\left\{1, (1 - x^p + y^p)^{\frac{1}{p}}\right\}$, also known as the pseudo-Łukasiewicz type 2 implication [48], and if one allows $p \in [-\infty, \infty]$ this can also be seen as a Schweizer and Sklar type 1 implication [107, 48].

2. In the next phase two many valued implications are combined by the use of a proper many valued conjunction into the equivalence of the form $x \leftrightarrow y \equiv T((x \to y), (y \to x))$. In the example case here fuzzy conjunction of the form $(\max\{x^p + y^p - 1, 0\})^{\frac{1}{p}}$, where $p \in [-\infty, \infty]$ is chosen or one can also select min to be the fuzzy conjunction. These both lead to the equivalence equation of the form $E(x, y) = (1 - |x^p - y^p|)^{\frac{1}{p}}$, where $p \in [-\infty, \infty]$.

3. In this third step, the previous step is further generalized by using weights w_i, $i \in N$ and a generalized mean m. In the example case this leads to the equation $E(x_i, y_i) = \left(\sum_{i=1}^{n} w_i \left(1 - |x_i^p - y_i^p|\right)^{\frac{m}{p}}\right)^{\frac{1}{m}}$.

This thesis demonstrates that there is no significant difference between the results achieved by comparison measures done by logical manipulations of equations versus the comparison measures achieved, in the first place, by using functional properties.

2.1.2 Created Equivalences, Pseudo Equivalences and Implications as Comparison Measures

It was noted that since S-type implications were used as basic algebraic structure, the implications were defined by the formula $a \to b \equiv S(\neg a, b)$ for all $a, b \in [0, 1]$, where S is a t-conorm, and a dual t-norm of this t-conorm was used for the creation of pseudo equivalences as comparison measures since this seems to be the obvious choice.

1. **Kleene-Dienes Based Comparison Measure**

Kleene-Dienes implication [48] is obtained by using standard fuzzy disjunction as a t-conorm:

$$I_{K-D}(x, y) = \max(1 - x, y)$$

One can create a parameterized form of Kleene-Dienes by setting $x := x^p$ and $y := y^p$, which leads to the parameterized Kleene-Dienes implication:

$$I_{K-D}(x, y) = \max(1 - x^p, y^p) \tag{34}$$

A standard fuzzy conjunction was then used to combine two implications $I_{K-D}(x, y) = \max(1 - x^p, y^p)$ and $I_{K-D}(y, x) = \max(1 - y^p, x^p)$ from this the following was achieved

Corollary 43 *The logical equivalence based on parameterized Kleene-Dienes implication:*

$$E_{K-D}(x, y) = \min(\max(1 - x^p, y^p), \max(1 - y^p, x^p)) \tag{35}$$

Since for example $E_{K-D}(x,x) \neq 1$ in general (35) can not fulfill reflexivity, one can see that (35) is not a similarity relation as defined in [116]. For (35) weights can be applied and combine values achieved with a generalized mean. From this procedure more values are obtained for evaluation. This approach has been proven to be practically effective in many previous studies [77, 83, 85, 113].

Definition 44 *Comparison measure based on the Kleene-Dienes implication:*

$$E_{K-D}(f_1(i), f_2(i)) = \left(\sum_{i=1}^{n} w_i \left(E_{K-D}(f_1(i), f_2(i)) \right)^m \right)^{\frac{1}{m}} \tag{36}$$

2. **Reichenbach Based Comparison Measure**

Reichenbach implication [48] is obtained by using the algebraic sum as a t-conorm:

$$I_R(x,y) = 1 - x + xy$$

A parameterized form of Reichenbach implication can be set by setting $x := x^p$ and $y := y^p$, which leads to the equation

$$I_R(x,y) = 1 - x^p + x^p y^p \tag{37}$$

The algebraic product is then used as a conjunction to combine two implications $I_R(x,y) = 1 - x^p + x^p y^p$ and $I_R(y,x) = 1 - y^p + x^p y^p$ from this one will reach

Corollary 45 *The logical equivalence based on parameterized Reichenbach implication:*

$$E_R(x,y) = (1 - x^p + x^p y^p)(1 - y^p + x^p y^p) \tag{38}$$

The formula above cannot fulfil reflexivity, so it cannot be considered as a similarity in the manner defined by Zadeh [116]. To this formula weights can also be applied and combined values with a generalized mean in order to obtain an extra parameter, which then can be used to obtain more values for evaluation.

Definition 46 *Comparison measure based on the parameterized Reichenbach implication:*

$$E_R(f_1(i), f_2(i)) = \left(\sum_{i=1}^{n} w_i \left(E_R(f_1(i), f_2(i)) \right)^m \right)^{\frac{1}{m}} \tag{39}$$

3. **Łukasiewicz Based Comparison Measure**

Łukasiewicz S-type implication [48] is obtained by using a bounded sum as a t-conorm:

$$I_L(x, y) = \min(1, 1 - x + y)$$

A parameterized form of the Łukasiewicz implication is obtained by setting $x := x^p$ and $y := y^p$, which leads to the equation

$$I_L(x, y) = \min(1, 1 - x^p + y^p) \tag{40}$$

The algebraic bounded product is then used as a conjunction to combine two implications $I_L(x, y) = \min(1, 1 - x^p + y^p)$ and $I_L(y, x) = \min(1, 1 - y^p + x^p)$ from this procedure one attains

Corollary 47 *The logical equivalence based on parameterized Łukasiewicz implication:*

$$E_L(x, y) = 1 - |x^p - y^p| \tag{41}$$

The equation above is reflexive, symmetric and transitive as are normal pseudo type Łukasiewicz structures. When weights are added and then values aggregated with a generalized mean the following form of equation is obtained:

Definition 48 *Comparison measure based on the parameterized Łukasiewicz implication:*

$$E_L(f_1(i), f_2(i)) = \left(\sum_{i=1}^{n} w_i \left(E_L(f_1(i), f_2(i))\right)^m\right)^{\frac{1}{m}} \tag{42}$$

4. **Combined Lukasiewicz and Shweizer & Sklar Based Comparison Measure**

Here the comparison measure is created which arises from the functional definition for the implications given in [97]. It is noted that pseudo Łukasiewicz type 2 [48] and Shweizer and Sklar type 1 [107, 48] implications form almost the same equivalence relation when these equivalences are formed by using a fuzzy conjunction such as *min* to combine corresponding implications.

This Łukasiewicz equivalence is included in the corresponding Shweizer & Sklar equivalence by taking the parameter values which go from negative to positive, thus $p \in]-\infty, \infty[$.

Definition 49 *Comparison measure based on Shweizer & Sklar - Łukasiewicz:*

$$E_{SSL}(f_1(i), f_2(i)) = \left(\sum_{i=1}^{n} w_i \left(1 - |f_1^p(i) - f_2^p(i)|\right)^{\frac{m}{p}}\right)^{\frac{1}{m}} \tag{43}$$

2.1.3 Some Statistical Results

Below, the tables are displayed of the variances, maximums and averages from the classification results obtained from the classifications, and the results are briefly discuss.

Table 3: Classification Variances With Optimized (DE) and Randomized (RND) Weights 1 for Equivalences and Implications

	Ionosphere	Iris	Pima
Kl-Die eqv$_{DE}$	$[0, 0.0057163]$	$[0.00011062, 0.010827]$	$[0, 0.0089028]$
Kl-Die eqv$_{RND}$	$[0, 0.066186]$	$[0.00057889, 0.0086307]$	$[0, 0.0206]$
Kl-Die imp$_{DE}$	$[0.0001277, 0.046692]$	$[0, 0.025025]$	$[0, 0.0035672]$
Kl-Die imp$_{RND}$	$[0, 0.040445]$	$[0, 0.027154]$	$[0, 0.022037]$
Łuka ekv$_{DE}$	$[0, 0.014592]$	$[0, 0.0038894]$	$[0, 0.0029255]$
Łuka ekv$_{RND}$	$[0, 0.067938]$	$[0.00038825, 0.0036141]$	$[0.00030067, 0.020311]$
Łuka imp$_{DE}$	$[0.00010366, 0.0053217]$	$[0.00016593, 0.01266]$	$[0, 0.022284]$
Łuka imp$_{RND}$	$[0, 0.0025634]$	$[0.0011248, 0.01024]$	$[0, 0.019975]$
Rbach ekv$_{DE}$	$[0.00011514, 0.033759]$	$[0, 0.02963]$	$[0, 0.0066011]$
Rbach ekv$_{RND}$	$[0, 0.065684]$	$[0, 0.02215]$	$[0, 0.01895]$
Rbach imp$_{DE}$	$[0.00013236, 0.045485]$	$[0, 0.027702]$	$[0, 0.014974]$
Rbach imp$_{RND}$	$[0, 0.038049]$	$[0, 0.024116]$	$[0, 0.021102]$
SS ekv$_{DE}$	$[0, 0.00624]$	$[0, 0.0039131]$	$[0, 0.0036583]$
SS ekv$_{RND}$	$[0, 0.006687]$	$[0.00041094, 0.0034994]$	$[0, 0.0254]$

Table 4: Classification Variances With Optimized (DE) and Randomized (RND) Weights 2 for Equivalences and Implications

	Post	Thyroid	Wine
Kl-Die eqv$_{DE}$	$[0.00044444, 0.038195]$	$[0, 0.10974]$	$[0.00011523, 0.017827]$
Kl-Die imp$_{DE}$	$[0, 0.040829]$	$[0, 0.11805]$	$[0, 0.018705]$
Kl-Die eqv$_{RND}$	$[0.00012785, 0.036564]$	$[0.00077909, 0.079777]$	$[0.0010296, 0.012229]$
Kl-Die imp$_{RND}$	$[0.0017656, 0.10286]$	$[0, 0.023163]$	$[0, 0.01122]$
Łuka ekv$_{DE}$	$[0.00041701, 0.033498]$	$[0, 0.11064]$	$[0.00011111, 0.004561]$
Łuka imp$_{DE}$	$[0.00035665, 0.033997]$	$[0, 0.10951]$	$[0.00022085, 0.010557]$
Łuka ekv$_{RND}$	$[0.00034662, 0.015962]$	$[0.00049578, 0.096275]$	$[0.00091196, 0.0032172]$
Łuka imp$_{RND}$	$[0.0090717, 0.034177]$	$[0.0022482, 0.08285]$	$[0.00099499, 0.0087482]$
Rbach ekv$_{DE}$	$[0, 0.04085]$	$[0, 0.080061]$	$[0, 0.023259]$
Rbach imp$_{DE}$	$[0, 0.040077]$	$[0, 0.086187]$	$[0, 0.023808]$
Rbach ekv$_{RND}$	$[0.0001409, 0.098257]$	$[0.0002145, 0.079288]$	$[0, 0.0241]$
Rbach imp$_{RND}$	$[0.0088538, 0.097995]$	$[0, 0.052141]$	$[0, 0.010503]$
SS ekv$_{DE}$	$[0.0012949, 0.03341]$	$[0, 0.11955]$	$0, 0.0043676]$
SS ekv$_{RND}$	$[0.0039882, 0.016374]$	$[0.00045619, 0.097096]$	$[0.00095514, 0.0031575]$

Table 5: Maximal Classification Results with Optimized (DE) and Randomized (RND) Weights for Equivalences and Implications

	Iono	Iris	Pima	Post	Thyroid	Wine
Kl-Die eqv$_{DE}$	0.92045	1	0.79688	0.77778	1	1
Kl-Die imp$_{DE}$	0.875	1	0.78385	0.77778	0.96296	0.97778
Kl-Die eqv$_{RND}$	0.91477	1	0.80469	0.75556	1	1
Kl-Die imp$_{RND}$	0.84659	0.66667	0.78906	0.75556	0.81481	0.91111
Łuka ekv$_{DE}$	0.93182	1	0.79427	0.77778	1	1
Łuka imp$_{DE}$	0.85227	0.96	0.77083	0.8	1	0.94444
Łuka ekv$_{RND}$	0.9375	1	0.8099	0.73333	1	1
Łuka imp$_{RND}$	0.84659	0.89333	0.74219	0.77778	1	0.95556
Rbach ekv$_{DE}$	0.93182	1	0.79688	0.8	1	1
Rbach imp$_{DE}$	0.88636	1	0.77604	0.8	0.99074	0.97778
Rbach ekv$_{RND}$	0.92614	1	0.80729	0.77778	1	1
Rbach imp$_{RND}$	0.85227	0.66667	0.6849	0.77778	0.98148	0.9
SS ekv$_{DE}$	0.89773	1	0.80208	0.77778	1	1
SS ekv$_{RND}$	0.90909	1	0.8151	0.73333	1	1

Table 6: Mean Classification Results with Optimized (DE) and Randomized (RND) Weights for Equivalences and Implications

	Iono	Iris	Pima	Post	Thyroid	Wine
Kl-Die eqv$_{DE}$	0.79007	0.86727	0.73059	0.61517	0.88318	0.90853
Kl-Die imp$_{DE}$	0.78037	0.62923	0.68996	0.68962	0.67952	0.62989
Kl-Die eqv$_{RND}$	0.79479	0.86293	0.68817	0.53682	0.8487	0.90526
Kl-Die imp$_{RND}$	0.75613	0.58612	0.66238	0.64956	0.49774	0.46769
Łuka ekv$_{DE}$	0.78627	0.97073	0.72927	0.58266	0.92598	0.93849
Łuka imp$_{DE}$	0.77261	0.8481	0.70728	0.67291	0.84262	0.83737
Łuka ekv$_{RND}$	0.79199	0.98617	0.74453	0.52977	0.92166	0.95437
Łuka imp$_{RND}$	0.75426	0.81669	0.58129	0.65336	0.85509	0.79738
Rbach ekv$_{DE}$	0.79891	0.84002	0.7458	0.63128	0.84133	0.88149
Rbach imp$_{DE}$	0.78547	0.74251	0.70939	0.69027	0.63493	0.79292
Rbach ekv$_{RND}$	0.80155	0.80921	0.71335	0.56499	0.777	0.85355
Rbach imp$_{RND}$	0.75602	0.62764	0.5041	0.67733	0.36739	0.61571
SS ekv$_{DE}$	0.79102	0.97048	0.73304	0.58396	0.93213	0.94992
SS ekv$_{RND}$	0.79755	0.98837	0.7477	0.58824	0.92596	0.96354

2.2 T-norms and T-conorms as the Measures for Comparison

In the paper [82] measures have been defined based on the use of the generalized mean, weights, t-norms and t-conorms. Below these results are added to the definition of the combination measure of the t-norm and t-conorm.

Connectives play an important role when trying to model reality by equations. For example, when linguistic interpretations such as "AND" or "OR" are used for connectives in conjunction and disjunction, quite often this does not require or mean crisp connectives, but that these connectives are only needed to some degree. In such cases connectives called t-norms or t-conorms may be used. The t-norm gives minimum compensation, while the t-conorm gives maximum compensation. This means that t-norms tend to give more value for the low values, while t-conorms give more value for the high values in the interval in which they are used. In practice, neither of these connectives fit the collected data appropriately. There is still a lot of information that is left in between of these two connectives. An important issue when dealing with t-norms and t-conorms is the question of how to combine them in a meaningful way, since neither of these connectives alone give a general compensation for the values where they are adapted. For this reason one should use a measure that somewhat compensates for this gap in values between these two norms. [21] shows how the generalized mean works as the compensative connective between minimum and maximum connectives. The scope of aggregation operators is demonstrated in Figure (2).

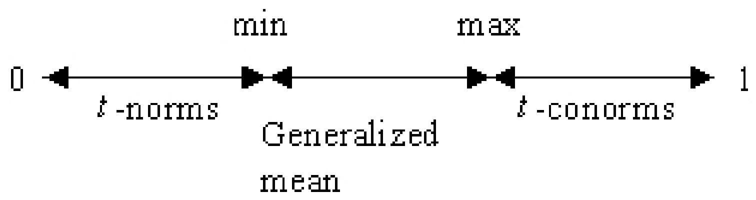

Figure 2: Compensation of t-norms and t-conorms

The first researchers to try the compensation of t-norms and t-conorms were Zimmermann and Zysno in [118]. They used the weighted geometric mean in order to compensate the gap between fuzzy intersections and unions. When one uses the geometric mean equal compensation is allocated to the all values, and problems might occur if some of the values combined are relatively very low or high.

2.2.1 Combined Measure for Comparison Based on T-norms and T-conorms

In (44) a general definition is given for the comparison measure based on t-norms and t-conorms, where fuzzy unions and intersections are combined by the generalized mean and weights.

It is clear that the comparison measure (44) should be able to give at least equal results to any comparison measure made from fuzzy intersections (45) or unions (46) alone, as long as the weights are chosen properly. Differential evolution is used for the weight selection problem. The t-norms and t-conorms used should be continuous for all vectors $x_1, x_2 \in [0,1]$. Archimedean t-norms and t-conorms are a good choice since they are continuous and monotonic [8].

The motivation for the equations (44), (45) and (46) is that when a generalized mean and weights for aggregation are used, it is possible to go through all the possible variations of the magnitude of the combined values of fuzzy intersections and unions.

Definition 50 *A combined comparison measure based on a fuzzy intersection and a fuzzy union with weights aggregated by a generalized mean:*

$$F_p\left(x_1,x_2\right)=\left(\sum_{i=1}^{n}\left(w_i(\omega_{Ti}T_i\left(x_1,x_2\right))+\left(1-w_i\right)\left(\omega_{Si}S_i\left(x_1,x_2\right)\right)\right)^m\right)^{\frac{1}{m}}, \tag{44}$$

where $i=1,\ldots,n$, p is a parameter combined with the corresponding class of weighted fuzzy intersections T_i and unions S_i, ω_{Ti}, ω_{Si} and w_i are weights.

It can be seen that if one sets $w_i = 1$, $\forall i \in \mathbb{Z}$ the following equation is reached:

Definition 51 *A comparison measure based on fuzzy intersection with weights and a generalized mean:*

$$C_p\left(x_1,x_2\right)=\left(\sum_{i=1}^{n}\left(\omega_{Ti}T_i\left(x_1,x_2\right)\right)^m\right)^{\frac{1}{m}}, \tag{45}$$

where p is a parameter combined with the corresponding class of weighted fuzzy intersections and $i=1,\ldots,n$ and ω_{Ti} is the weight belonging to the fuzzy intersection.

If one sets $w_i = 0$, $\forall i \in \mathbb{Z}$ the following equation is reached:

Definition 52 *Comparison measure based on a fuzzy union with weights and a generalized mean:*

$$D_p\left(x_1,x_2\right)=\left(\sum_{i=1}^{n}\left(\omega_{Si}S_i\left(x_1,x_2\right)\right)^m\right)^{\frac{1}{m}}, \tag{46}$$

where p is a parameter combined with the corresponding class of weighted fuzzy unions and $i = 1, \ldots, n$ and ω_{Si} is the weight belonging to the fuzzy union.

From the formulation of (44), (45) and (46) it can be seen that the combined measure of fuzzy intersections and unions includes all the valuations that weighted t-norms and t-conorms can give and in this way will always give results at least as good as any combined t-norms or t-conorms alone, as long as weights are chosen in a sensible way. In the experimental part of this thesis results will be presented on how these comparison measures worked in classification tasks with some well-known t-norms and t-conorms.

2.2.2 Created Comparison Measures From T-norms and T-conorms

The following is a brief representation of the algebraic equations that can be created by combining weights into some important t-norms and t-conorms and then the combining values are given that were achieved by aggregating them with a generalized mean.

Definition 53 *Combined comparison measure based on the t-norm and t-conorm with a generalized mean and weights:*

$$F_p \langle f_1(i), f_2(i) \rangle = \left(\sum_{i=1}^{n} \left(w_i T_i^p \langle f_1(i), f_2(i) \rangle + (1 - w_i) \left(S_i^p \langle f_1(i), f_2(i) \rangle \right) \right)^m \right)^{\frac{1}{m}}, \tag{47}$$

where $i = 1, \ldots, n$, p is a parameter combined to the corresponding class of fuzzy intersections T_i and unions S_i and w_i are weights and $i = 1, \ldots, n$.

The comparison measure (47) has been tested by combining it with the following comparison measures (48), (49), (50), (51), (52), (53), (54), (55), (56) and (57). The measure (47) has been tested without weights ω_{ci} and ω_{di}, since the weighting process was too time consuming with differential evolution. All the comparison measures mentioned in this sub-chapter have been tested in classification tasks. T-norms and t-conorms are tested with weights, where a generalized mean has been used to aggregate and compensate the values.

Parameterized families of t-norms and t-conorms are used here. These families are the Dombi family, Frank family, Schweizer-Sklar family, Yager family and Yu family. The Frank and Schweizer-Sklar families of t-norms are also copula families [96] so they have some good statistical properties see Fisher 1997 [22]. The Yager, Dombi and Yu families are among the most popular families for modelling the intersection.

Yager also has several applications for the fuzzy set theory in the context of fuzzy numbers.

From these t-norms and t-conorm families have been created for the following comparison measures.

Definition 54 *Measure based on Dombi (1982), [16] class of t-norm with generalized mean and weights:*

$$C_D \left\langle f_1(i), f_2(i) \right\rangle = \left(\sum_{i=1}^{n} \omega_{ci} \left(1 + \left[\left(\frac{1}{f_1(i)} - 1 \right)^p + \left(\frac{1}{f_2(i)} - 1 \right)^p \right]^{\frac{1}{p}} \right)^{-m} \right)^{\frac{1}{m}}, \tag{48}$$

where $p > 0$ and $i = 1, \ldots, n$.

Definition 55 *Measure based on Frank (1979) [29] class of t-norm with generalized mean and weights:*

$$C_F \left\langle f_1(i), f_2(i) \right\rangle = \left(\sum_{i=1}^{n} \omega_{ci} \left(\log_p \left[1 + \frac{\left(p^{f_1(i)} - 1\right)\left(p^{f_2(i)} - 1\right)}{p-1} \right] \right)^m \right)^{\frac{1}{m}}, \tag{49}$$

where $p > 0$, $p \neq 1$ and $i = 1, \ldots, n$.

Definition 56 *Measure based on Schweizer & Sklar 1 (1963) [95] class of t-norm with generalized mean and weights:*

$$C_{SS} \left\langle f_1(i), f_2(i) \right\rangle = \left(\sum_{i=1}^{n} \omega_{ci} \left(\max \left\{ 0, (f_1(i))^p + (f_2(i))^p - 1 \right\} \right)^{\frac{m}{p}} \right)^{\frac{1}{m}}, \tag{50}$$

where $p \neq 0$ and $i = 1, \ldots, n$.

Definition 57 *Measure based on Yager (1980) [110] class of t-norm with generalized mean and weights:*

$$C_Y \left\langle f_1(i), f_2(i) \right\rangle = \left(\sum_{i=1}^{n} \omega_{ci} \left(1 - \min \left\{ 1, \left[(1 - f_1(i))^p + (1 - f_2(i))^p \right]^{\frac{1}{p}} \right\} \right)^m \right)^{\frac{1}{m}}, \tag{51}$$

where $p > 0$ and $i = 1, \ldots, n$.

Definition 58 *Measure based on Yu (1985) [114] class of t-norm with generalized mean and weights:*

$$C_{Yu} \left\langle f_1(i), f_2(i) \right\rangle = \left(\sum_{i=1}^{n} \omega_{ci} \left(\max \left\{ 0, (1+p)\left(f_1(i) + f_2(i) - 1\right) - p \cdot f_1(i) f_2(i) \right\} \right)^m \right)^{\frac{1}{m}}, \tag{52}$$

where $p > -1$ and $i = 1, \ldots, n$.

Definition 59 *Measure based on Dombi (1982) [16] class of t-conorm with generalized mean and weights:*

$$D_D \langle f_1(i), f_2(i) \rangle = \left(\sum_{i=1}^{n} \omega_{di} \left(1 + \left[\left(\frac{1}{f_1(i)} - 1 \right)^{-p} + \left(\frac{1}{f_2(i)} - 1 \right)^{-p} \right]^{-\frac{1}{p}} \right)^{-m} \right)^{\frac{1}{m}}, \tag{53}$$

where $p > 0$ *and* $i = 1, \ldots, n$.

Definition 60 *Measure based on Frank (1979) [29] class of t-conorm with generalized mean and weights:*

$$D_F \langle f_1(i), f_2(i) \rangle = \left(\sum_{i=1}^{n} \omega_{di} \left(1 - \log_p \left[1 + \frac{\left(p^{1-f_1(i)} - 1\right)\left(p^{1-f_2(i)} - 1\right)}{p - 1} \right] \right)^{m} \right)^{\frac{1}{m}}, \tag{54}$$

where $p > 0$, $p \neq 1$ *and* $i = 1, \ldots, n$.

Definition 61 *Measure based on Schweizer & Sklar 1 (1963) [95] class of t-conorm with generalized mean and weights:*

$$D_{SS} \langle f_1(i), f_2(i) \rangle = \left(\sum_{i=1}^{n} \omega_{di} \left(1 - \left(\max\left\{0, (f_1(i))^p + (f_2(i))^p - 1\right\} \right)^{\frac{1}{p}} \right)^{m} \right)^{\frac{1}{m}}, \tag{55}$$

where $p \neq 0$ *and* $i = 1, \ldots, n$.

Definition 62 *Measure based on Yager (1980) [110] class of t-conorm with generalized mean and weights:*

$$D_Y \langle f_1(i), f_2(i) \rangle = \left(\sum_{i=1}^{n} \omega_{di} \left(\min\left\{ 1, \left[(f_1(i))^p + (f_2(i))^p \right]^{\frac{1}{p}} \right\} \right)^{m} \right)^{\frac{1}{m}}, \tag{56}$$

where $p > 0$ *and* $i = 1, \ldots, n$.

Definition 63 *Measure based on Yu (1985) [114] class of t-conorm with generalized mean and weights:*

$$D_{Yu} \langle f_1(i), f_2(i) \rangle = \left(\sum_{i=1}^{n} \omega_{di} \left(\min\left\{ 1, f_1(i) + f_2(i) + p \cdot f_1(i) f_2(i) \right\} \right)^{m} \right)^{\frac{1}{m}}, \tag{57}$$

where $p > -1$ *and* $i = 1, \ldots, n$.

2.2.3 Some Statistical Results

Table 7: Classification Variances With Optimized (DE) and Randomized (RND) Weights 1 for T-norms and T-conorms

	Ionosphere	Iris	Pima
Dombi t_{DE}	$[0, 0.022195]$	$[0, 0.031565]$	$[0, 0.0071898]$
Dombi tco_{DE}	$[0, 0.042921]$	$[0, 0.067319]$	$[0, 0.031025]$
Dombi $combo_{DE}$	$[0, 0.062152]$	$[0, 0.10634]$	$[0, 0.024517]$
Dombi t_{RND}	$[0, 0.015997]$	$[0, 0.022814]$	$[0, 0.022922]$
Dombi tco_{RND}	$[0, 0.022625]$	$[0, 0.0068198]$	$[0, 0.013764]$
Dombi $combo_{RND}$	$[0, 0.030018]$	$[0, 0.006287]$	$[0, 0.013362]$
Frank t_{DE}	$[0, 0.0035511]$	$[0, 0.034497]$	$[0, 0.0070928]$
Frank tco_{DE}	$[0.00022419, 0.045307]$	$[0, 0.039982]$	$[0, 0.017554]$
Frank $combo_{DE}$	$[0, 0.0075833]$	$[0, 0.022331]$	$[0, 0.016104]$
Frank t_{RND}	$[0, 0.0018317]$	$[0, 0.024154]$	$[0, 0.018401]$
Frank tco_{RND}	$[0, 0.036171]$	$[0, 0.021666]$	$[0, 0.023652]$
Frank $combo_{RND}$	$[0, 0.034384]$	$[0.0027093, 0.047313]$	$[0, 0.026613]$
SS1 t_{DE}	$[0, 0.0039888]$	$[0, 0.02259]$	$[0, 0.022478]$
SS1 tco_{DE}	$[0, 0.046902]$	$[0, 0.037894]$	$[0, 0.0029705]$
SS1 $combo_{DE}$	$[0, 0.0066302]$	$[0, 0.037651]$	$[0, 0.0036146]$
SS1 t_{RND}	$[0, 0.0017585]$	$[0, 0.024946]$	$[0, 0.020029]$
SS1 tco_{RND}	$[0, 0.019612]$	$[0, 0.025581]$	$[0, 0.0097393]$
SS1 $combo_{RND}$	$[0.0013732, 0.035834]$	$[0.0010476, 0.04591]$	$[0, 0.022846]$
Yager t_{DE}	$[0, 0.0044626]$	$[0, 0.020585]$	$[0, 0.014673]$
Yager tco_{DE}	$[0, 0.081635]$	$[0, 0.063378]$	$[0, 0.022963]$
Yager $combo_{DE}$	$[0, 0.0588]$	$[0, 0.10281]$	$[0, 0.03236]$
Yager t_{RND}	$[0, 0.001915]$	$[0, 0.024094]$	$[0, 0.018067]$
Yager tco_{RND}	$[0, 0.071424]$	$[0, 0.0035212]$	$[0, 0.014494]$
Yager $combo_{RND}$	$[0, 0.019078]$	$[0, 0.012974]$	$[0, 0.014832]$
Yu t_{DE}	$[0, 0.0040806]$	$[0, 0.032128]$	$[0, 0.019328]$
Yu tco_{DE}	$[0, 0.070861]$	$[0, 0.030864]$	$[0, 0.0092253]$
Yu $combo_{DE}$	$[0, 0.062489]$	$[0, 0.043615]$	$[0, 0.03739]$
Yu t_{RND}	$[0, 0.001886]$	$[0, 0.026196]$	$[0, 0.0163]$
Yu tco_{RND}	$[0, 0.060143]$	$[0, 0.0028854]$	$[0, 0.0066133]$
Yu $combo_{RND}$	$[0, 0.0043759]$	$[0, 0.014497]$	$[0, 0.0061772]$

Table 8: Classification Variances With Optimized (DE) and Randomized (RND) Weights 2 for T-norms and T-conorms

	Post	Thyroid	Wine
Dombi t_{DE}	$[0, 0.047276]$	$[0, 0.11967]$	$[0.00023182, 0.024347]$
Dombi tco_{DE}	$[0, 0.034546]$	$[0, 0.11996]$	$[0, 0.045152]$
Dombi $combo_{DE}$	$[0, 0.057004]$	$[0, 0.10742]$	$[0, 0.053108]$
Dombi t_{RND}	$[0, 0.076472]$	$[0, 0.074297]$	$[0, 0.011611]$
Dombi tco_{RND}	$[0, 0.031498]$	$[0, 0.0026813]$	$[0, 0.014835]$
Dombi $combo_{RND}$	$[0.0040719, 0.069764]$	$[0, 0.019775]$	$[0, 0.016622]$
Frank t_{DE}	$[0, 0.043764]$	$[0, 0.075766]$	$[0, 0.030317]$
Frank tco_{DE}	$[0.00046091, 0.050283]$	$[0.0001105, 0.072303]$	$[0.00013169, 0.028911]$
Frank $combo_{DE}$	$[0, 0.027852]$	$[0, 0.074909]$	$[0.00013855, 0.022344]$
Frank t_{RND}	$[0.013187, 0.07996]$	$[0, 0.013771]$	$[0, 0.011413]$
Frank tco_{RND}	$[0.00064341, 0.026005]$	$[0.00039394, 0.090032]$	$[0.0036436, 0.022923]$
Frank $combo_{RND}$	$[0.021903, 0.080807]$	$[0.00053232, 0.098421]$	$[0.0027008, 0.045572]$
SS1 t_{DE}	$[0, 0.030837]$	$[0, 0.066443]$	$[0, 0.025582]$
SS1 tco_{DE}	$[0, 0.069882]$	$[0, 0.063047]$	$[0, 0.031364]$
SS1 $combo_{DE}$	$[0, 0.028143]$	$[0, 0.067086]$	$[0.00016461, 0.024951]$
SS1 t_{RND}	$[0, 0.073282]$	$[0, 0.024676]$	$[0, 0.011099]$
SS1 tco_{RND}	$[0, 0.043788]$	$[0, 0.060266]$	$[0, 0.02044]$
SS1 $combo_{RND}$	$[0.02529, 0.081395]$	$[0.00020141, 0.048737]$	$[0.0014516, 0.02555]$
Yager t_{DE}	$[0, 0.030228]$	$[0, 0.072021]$	$[0, 0.028956]$
Yager tco_{DE}	$[0, 0.033257]$	$[0, 0.10218]$	$[0, 0.044012]$
Yager $combo_{DE}$	$[0, 0.041706]$	$[0, 0.11024]$	$[0, 0.046147]$
Yager t_{RND}	$[0, 0.070296]$	$[0, 0.041101]$	$[0, 0.0134]$
Yager tco_{RND}	$[0, 0.026539]$	$[0, 0.067553]$	$[0, 0.017146]$
Yager $combo_{RND}$	$[0.0059061, 0.071157]$	$[0.00011029, 0.074523]$	$[0, 0.017791]$
Yu t_{DE}	$[0, 0.030091]$	$[0, 0.077271]$	$[0, 0.019638]$
Yu tco_{DE}	$[0, 0]$	$[0, 0.12428]$	$[0, 0.035365]$
Yu $combo_{DE}$	$[0, 0.030486]$	$[80, 0.095002]$	$[0.00035117, 0.025158]$
Yu t_{RND}	$[0, 0.073576]$	$[0, 0.0035122]$	$[0, 0.010529]$
Yu tco_{RND}	$[0, 0]$	$[0.00019166, 0.0055831]$	$[0, 0.011344]$
Yu $combo_{RND}$	$[0.019738, 0.073204]$	$[0.00026761, 0.036305]$	$[0, 0.014517]$

Table 9: Maximal Classification Results with Optimized (DE) and Randomized (RND) Weights for T-norms and T-conorms

	Iono	Iris	Pima	Post	Thyroid	Wine
Dombi t_{DE}	0.89773	0.93333	0.75521	0.8	1	0.97778
Dombi tco_{DE}	0.88068	0.94667	0.78125	0.82222	0.98148	0.93333
Dombi $combo_{DE}$	0.90909	1	0.77604	0.8	1	0.98889
Dombi t_{RND}	0.85227	0.72	0.7474	0.77778	1	0.94444
Dombi tco_{RND}	0.67614	0.33333	0.76042	0.82222	0.69444	0.35556
Dombi $combo_{RND}$	0.88068	0.66667	0.77604	0.82222	0.92593	0.76667
Frank t_{DE}	0.90341	0.98667	0.74479	0.8	0.99074	0.97778
Frank tco_{DE}	0.93182	0.90667	0.79948	0.8	0.99074	0.96667
Frank $combo_{DE}$	0.94886	1	0.80729	0.77778	1	1
Frank t_{RND}	0.875	0.66667	0.72135	0.84444	0.87037	0.9
Frank tco_{RND}	0.89773	0.66667	0.79167	0.77778	0.98148	0.91111
Frank $combo_{RND}$	0.94318	1	0.8151	0.82222	1	1
SS1 t_{DE}	0.88636	0.97333	0.77865	0.77778	0.92593	0.94444
SS1 tco_{DE}	0.88068	0.97333	0.79427	0.77778	0.90741	0.93333
SS1 $combo_{DE}$	0.93182	1	0.79167	0.8	0.91667	0.96667
SS1 t_{RND}	0.88068	0.66667	0.78385	0.8	0.82407	0.86667
SS1 tco_{RND}	0.86364	0.66667	0.78906	0.82222	0.72222	0.84444
SS1 $combo_{RND}$	0.89773	0.97333	0.78906	0.8	0.90741	0.93333
Yager t_{DE}	0.90341	0.97333	0.78385	0.8	0.99074	0.98889
Yager tco_{DE}	0.89773	0.97333	0.77865	0.8	0.99074	0.97778
Yager $combo_{DE}$	0.94886	1	0.78125	0.8	0.99074	0.98889
Yager t_{RND}	0.875	0.66667	0.77083	0.77778	0.98148	0.95556
Yager tco_{RND}	0.86364	0.41333	0.71354	0.8	0.69444	0.66667
Yager $combo_{RND}$	0.86364	0.66667	0.75521	0.82222	0.75926	0.88889
Yu t_{DE}	0.89773	0.98667	0.78125	0.8	0.84259	0.95556
Yu tco_{DE}	0.90341	0.66667	0.73698	0.022222	0.96296	1
Yu $combo_{DE}$	0.92614	1	0.79167	0.82222	0.9537	0.97778
Yu t_{RND}	0.86932	0.66667	0.79167	0.77778	0.7963	0.87778
Yu tco_{RND}	0.85227	0.33333	0.72656	0.022222	0.64815	0.47778
Yu $combo_{RND}$	0.86932	0.66667	0.73958	0.8	0.94444	0.8

Table 10: Mean Classification Results with Optimized (DE) and Randomized (RND) Weights for T-norms and T-conorms

	Iono	Iris	Pima	Post	Thyroid	Wine
Dombi t_{DE}	0.81003	0.70245	0.63523	0.71632	0.75537	0.7972
Dombi tco_{DE}	0.59766	0.69219	0.70636	0.71072	0.58936	0.71955
Dombi $combo_{DE}$	0.56841	0.82143	0.6429	0.71131	0.74274	0.82988
Dombi t_{RND}	0.68855	0.50687	0.51795	0.72117	0.57869	0.64626
Dombi tco_{RND}	0.4228	0.33285	0.49723	0.67868	0.13854	0.30177
Dombi $combo_{RND}$	0.50663	0.37502	0.52403	0.73466	0.29497	0.40796
Frank t_{DE}	0.58678	0.73116	0.6521	0.72316	0.36156	0.68854
Frank tco_{DE}	0.82523	0.69186	0.71644	0.72687	0.81572	0.8369
Frank $combo_{DE}$	0.8368	0.91677	0.74105	0.72103	0.83698	0.90467
Frank t_{RND}	0.54345	0.46915	0.51067	0.72808	0.24405	0.49268
Frank tco_{RND}	0.5183	0.41616	0.64671	0.5054	0.6633	0.64959
Frank $combo_{RND}$	0.79859	0.7161	0.72448	0.7252	0.75595	0.87217
SS1 t_{DE}	0.57586	0.56898	0.6616	0.43899	0.57813	0.44556
SS1 tco_{DE}	0.56161	0.57267	0.69433	0.69862	0.59627	0.53976
SS1 $combo_{DE}$	0.80951	0.69026	0.69809	0.72104	0.6099	0.70736
SS1 t_{RND}	0.54718	0.51817	0.64657	0.43461	0.58367	0.41418
SS1 tco_{RND}	0.38858	0.33992	0.66522	0.68328	0.50116	0.39618
SS1 $combo_{RND}$	0.75039	0.62171	0.69545	0.73047	0.55252	0.54878
Yager t_{DE}	0.58224	0.66718	0.64665	0.68692	0.67863	0.7041
Yager tco_{DE}	0.5832	0.65698	0.69679	0.60789	0.59708	0.70006
Yager $combo_{DE}$	0.78778	0.78848	0.61902	0.72001	0.64222	0.76967
Yager t_{RND}	0.54963	0.48627	0.52905	0.693	0.64195	0.55067
Yager tco_{RND}	0.48709	0.30432	0.50614	0.599	0.15576	0.33845
Yager $combo_{RND}$	0.53824	0.39725	0.5192	0.72894	0.32447	0.42781
Yu t_{DE}	0.58429	0.76449	0.66028	0.46831	0.47884	0.53567
Yu tco_{DE}	0.58484	0.47904	0.67926	0.022222	0.61134	0.76706
Yu $combo_{DE}$	0.75966	0.80896	0.64811	0.72136	0.63387	0.74479
Yu t_{RND}	0.54521	0.66425	0.66475	0.49218	0.47256	0.48777
Yu tco_{RND}	0.53801	0.25542	0.53917	0.022222	0.19901	0.34205
Yu $combo_{RND}$	0.53366	0.46593	0.54644	0.72011	0.47866	0.55863

2.3 Generalized 3Π-uninorm as The Measure for Comparison

3Π-uninorm was presented for the first time in [16] by József Dombi. In the papers [83] and [90] a generalized form of the 3Π-uninorm is given. In this subsection these results are recalled. The uninorm was defined in (20) and the representation theorem for uninorms can be defined as follows:

Theorem 64 *Suppose U is an almost continuous uninorm with a neutral element $e \in \,]0,1[$. Then there exists a strictly increasing continuous function $h : [0,1] \to \mathbb{R}$ with $h(e) = 0$ such that representation*

$$U(x,y) = h^{-1}(h(x) + h(y)) \tag{58}$$

holds true if and only if the following two conditions are satisfied:

a) U is strictly increasing on the open unit square;
b) U is self-dual with respect to a strong negation N with fixed point e.

Proof. Look [25] for proof. ■

The only class of uninorms that can be represented in the form (58) are aggregative operators that József Dombi introduced in 1982 [16], [25], [26].

Definition 65 *A multiplicative form of the representation theorem for uninorms can be derived for uninorms by setting h as an additive generator of the uninorm U and $f(x) = \exp h(x)$, which is strictly increasing continuous function from $[0,1]$ to $[0,\infty]$ such that*

$$\begin{array}{l} U(x_1, x_2) = h^{-1}(h(x_1) + h(x_2)) = \\ f^{-1}(f(x_1) f(x_2)) \end{array} \tag{59}$$

holds true.

Corollary 66 *If $f(x^p) = \frac{x^p}{1-x^p}$ is set then $f^{-1}(x^p) = \frac{x^p}{1+x^p}$ is obtained and from this it follows*

$$\begin{array}{l} U(x_1^p, x_2^p) = \\ \begin{cases} 0 \text{ if } x_1 = 0 \text{ and } x_2 = 1 \text{ or } x_1 = 1 \text{ and } x_2 = 0 \\ \frac{x_1^p x_2^p}{x_1^p x_2^p + (1-x_1^p)(1-x_2^p)} \text{ otherwise} \end{cases} \end{array}. \tag{60}$$

This is the parameterized form of the 3π-uninorm, $p > 0$ and neutral element $e = \sqrt[p]{\frac{1}{2}}$.

Lemma 67 *A parameterized form of the 3π-uninorm can be defined by the following formula*

$$U\left(x_1^p, x_2^p\right) = \begin{cases} 0 \text{ if } x_1 = 0 \text{ and } x_2 = 1 \text{ or } x_1 = 1 \text{ and } x_2 = 0 \\ \frac{x_1^p x_2^p}{x_1^p x_2^p + \left(1 - x_1^p\right)\left(1 - x_2^p\right)} \text{ otherwise} \end{cases} \tag{61}$$

$p > 0$ and neutral element $e = \sqrt[p]{\frac{1}{2}}$.

Proof. The commutativity is obvious. The mapping is also still increasing since the power value p holds true for the monotonicity. Associativity follows on from the associativity of $U(x_1, x_2)$. It is also an easy calculation to show that the neutral element $e = \sqrt[p]{\frac{1}{2}}$. ■

In the measure (62) two vectors $x_1, x_2 \in [0,1]^n$ are given:the elements of which present the properties of two objects. These objects are to be compared by the 3π-uninorm with a generalized mean. The parameters for measures are the mean value $m \in \mathbb{R} \setminus \{0\}$, weights $w = (w_1, \ldots, w_n)$ and parameter value p.

Theorem 68 *The weighted, generalized mean compensated form of the parameterized 3π-uninorm, is a function $U_{3\pi} : ([0,1]^n)^2 \to [0,1]$ defined as:*

$$\begin{aligned} &U_{3\pi}\left(x_1^p, x_2^p; m, w\right) = \\ &\begin{cases} 0 \text{ if } x_1 = 0 \text{ and } x_2 = 1 \text{ or } x_1 = 1 \text{ and } x_2 = 0 \\ \left(\sum\limits_{i=1}^{n} w_i \left(\frac{x_1^p(i) x_2^p(i)}{x_1^p(i) x_2^p(i) + \left(1 - x_1^p(i)\right)\left(1 - x_2^p(i)\right)}\right)^m\right)^{\frac{1}{m}} \end{cases}, \end{aligned} \tag{62}$$

where $p > 0$, $0 \leq w_i \leq 1$ and $\frac{1}{n}\sum\limits_{i=1}^{n} w_i = 1$. In fact, the way the weights are normalized is not important in classification since it affects only the range of the measure, not the order of values.

Proof. Since the generalized mean is a continuous and monotonic operator, when $m \in \mathbb{R} \setminus \{0\}$ all uninorm properties stay intact. ■

2.3.1 Some Statistical Results

Table 11: Classification Variances With Optimized (DE) and Randomized (RND) Weights 1 for 3Π-uninorm

	Ionosphere	Iris	Pima
3Π-uni$_{DE}$	$[0, 0.035158]$	$[0, 0.026825]$	$[0, 0.011224]$
3Π-uni$_{RND}$	$[0, 0.015214]$	$[0, 0.022218]$	$[0, 0.020157]$

Table 12: Classification Variances With Optimized (DE) and Randomized (RND) Weights 2 for 3Π-uninorm

	Post	Thyroid	Wine
3Π-uni$_{DE}$	$[0.00013717, 0.035276]$	$[0, 0.079961]$	$[0, 0.032368]$
3Π-uni$_{RND}$	$[0.0097227, 0.030767]$	$[0, 0.006699]$	$[0, 0.015109]$

Table 13: Maximal Classification Results with Optimized (DE) and Randomized (RND) Weights for 3Π-uninorm

	Iono	Iris	Pima	Post	Thyroid	Wine
3Π-uni$_{DE}$	0.90909	0.97333	0.78385	0.82222	1	0.97778
3Π-uni$_{RND}$	0.82955	0.66667	0.6901	0.82222	0.83333	0.9

Table 14: Mean Classification Results with Optimized (DE) and Randomized (RND) Weights for 3Π-uninorm

	Iono	Iris	Pima	Post	Thyroid	Wine
3Π-uni$_{DE}$	0.59147	0.61591	0.60158	0.6983	0.29837	0.56514
3Π-uni$_{RND}$	0.56686	0.46184	0.51414	0.67234	0.20889	0.43728

2.4 Results From Classification Tasks

From the tables (3), (4), (5), (6), (7), (8), (9), (10), (11), (12), (13), (14) it can be seen that variances, all the way through, were relatively small; normally maximum variance is just a few per cent. This can also be seen from the mean classification results.

The results achieved by the different equivalences based comparison measures give no great differences. Accordingly, it seems that it does not matter much, in practice, whether one uses logically or functionally formulated equivalences in classification. In the Post-operative data set, implication based comparison measures worked better than equivalence based ones. In all other cases, equivalence based comparison measures mostly gave significantly better classification results than implication based ones.

It can also be seen that in most cases combined comparison measure (47) from t-norms and t-conorms give better results than plain t-norms or t-conorms. This is of course dependent on the data sets because t-norms give more value for small values and t-conorms gives more value for large values. However, in 37 of 60 mean classification cases and in 41 of 60 maximum classification cases the best classification results were achieved using a combined measure. In many cases classification results also improve a great deal by the use of combination of t-norms and t-conorms. For example, this is the case where the t-norms and t-conorms of Frank (49) and (54) or Schweizer & Sklar 1 (50) and (55) are used. It is also worth noting that in all of the classification tasks presented here Frank based comparison measures gave the best classification results, which are the same or better than those coming from the pseudo equivalences. All but one of these best results achieved by Frank come from the use of a combined measure (47) of t-norm (49) and t-conorm (54).

3 Applications

The benefits that one can obtain by applying an algebraic form of many-valued logical operators with weights and by combining these with a generalized mean in comparison tasks becomes apparent when these comparison measures are used to model the systems which are somewhat graded in their nature. Examples of these kinds of systems are expert systems and decision analysis, where the management of uncertainty often plays an important part.

This section will present how the comparison measures that have been briefly introduced in previous chapters, work in application areas concerning the defining of an athlete's aerobic and anaerobic thresholds and medical data classification.

3.1 Defining an Athlete's Aerobic and Anaerobic Thresholds

Described here is an expert system for defining an athlete's aerobic and anaerobic thresholds that successfully mimics the decision-making done by sport medicine professionals [85]. The functionality of this system is based on the fuzzy comparison measure, generalized mean, membership functions and differential evolution. Differential evolution is used to tune the parameters in the comparison measure. The comparison measure is based on the use of fuzzy equivalences and a modification factor that tunes the shape of the membership function in hand. The measure presented is especially suitable for expert systems. The system is tested in order to show that the results do not show any statistically significant differences from the values estimated by experts.

The increase in the blood lactate, above a certain exercise intensity, has been recognized since the early 20th century [17]. The term anaerobic threshold was first used by Wasserman and McIlroy [106] when the blood lactate is about $2.0\frac{mmol}{l}$. Later, Kinderman et al [44] used a blood lactate of $4.0\frac{mmol}{l}$ to describe the aerobic-anaerobic threshold. The terminology used has been the source of continued controversy and debate in sports medicine. A somewhat stable practice nowadays is to call Wasserman's threshold aerobic and Kinderman's threshold anaerobic. These thresholds can be defined in a laboratory with commonly accepted criteria. However, in this case a human expert is always needed to define these thresholds, and moreover, most of the laboratories have done some small modifications to their criteria so that the results are not comparable with each other even from the same data. People's subjective opinions and the weighting of different variables and criteria also create considerable differences in the thresholds that are related to the client. Here, the goal has been to establish a simulation model which computes the threshold values mentioned, by mimicking expert judgments for thresholds values.

It is of vital importance for top athletes to know their individual aerobic and anaerobic thresholds. This is due to the fact that their workouts will be more efficient and can be more focused on different parts of endurance if the thresholds are known. The basic aerobic endurance is improved with workouts when the pulse does not exceed the aerobic threshold. Aerobic speed endurance, on the other hand, is improved when the pulse stays between these thresholds. Finally, the maximal aerobic endurance is improved when the pulse is over the anaerobic threshold. The thresholds are given in heartbeats per minute.

It has been shown that there are no statistically significant difference between the results given by this system versus the results given by sport's medicine experts. The fuzzy decision-making system presented in this thesis is based on the use of membership functions of fuzzy criteria set by experts, fuzzy equivalence, the generalized mean, weights and modifying factors. The generalized mean, weights and modifying factors are used as parameters in the comparison measure. Differential evolution is used to find such values for these parameters that the thresholds defined by this system, compared to the values defined by experts are as close to each other as possible. The correct thresholds are, by assumption, to be found at the point where the combination of comparison measure gives the maximal value for the pulse.

3.1.1 System Definition

The measurement data sets were provided by KIHU - the Research Institute for Olympic Sports. Each file in the set contains an athlete's measurements during an incremental workout. In total, 154 data files were used, all of which included 11 variables. The variables used for defining the aerobic and anaerobic thresholds are the following: 1) the content of lactic acid in the capillary blood, 2) ventilation, 3) the consumption of oxygen, 4) the production of carbon dioxide, 5) the relative amount of oxygen in the respiration air, 6) the relative amount of carbon dioxide in the respiration air, 7) the ventilation equivalent for oxygen, 8) the ventilation equivalent for carbon dioxide and 9) the respiration quotient. The last two were the pulse in beats per minute and times per minute.

The criteria deduced from the experts' instructions for defining the aerobic threshold were the following:

1. The pulse is about 40 beats per minute below the maximal pulse.
2. The content of lactic acid in the capillary blood begins to rise.
3. The content of lactic acid in the capillary blood is about $1.0 - 2.5$ mmol per liter.

4. The ventilation begins to rise from the initial level.
5. The relative amount of oxygen in the respiration air reaches its maximum.
6. The ventilation equivalent for oxygen is the lowest.
7. The lactic acid divided by the consumption of oxygen is the lowest.

When the load is raised over the aerobic threshold, the muscles start to work at the aerobic-anaerobic level. If the load is raised enough, the anaerobic production of energy will increase over the point where the ability of the muscles to remove lactic acid and to control the acidity is insufficient. This point is the anaerobic threshold.

The corresponding criteria for defining the anaerobic threshold were the following:

1. The pulse is about 15 beats per minute below the maximal pulse.
2. The content of lactic acid in the capillary blood is about $2.5 - 4.0$ mmol per liter.
3. The content of lactic acid in the capillary blood begins to raise radically.
4. The ventilation equivalent for carbon dioxide changes radically.
5. The ventilation equivalent for oxygen begins to rise radically.
6. The relative amount of oxygen in the respiration air begins to drop.

This information was the starting point from which the system for the thresholds' definition was created. Because most of the criteria were vague, fuzzy logic offered suitable tools for the modeling of this kind of system.

3.1.2 Fuzzy Decision Making Model

Linear interpolation is used to interpolate the measurement data along the pulse. Various different interpolation methods were also tested but they gave significantly poorer results.

Setting the criteria for the aerobic and anaerobic threshold was taken from the criteria presented. Using the membership functions, which were Gaussian, triangular and trapezoidal, fuzzified these criteria. These are chosen to imitate an expert's judgment as a grade of certainty with the corresponding data.

In the equations below, weights w_i and membership values x_i, y_i and μ have been normalized to the values between $[0, 1]$, $\forall i$ and $m \neq 0$.

One result concerning the choice of comparison measure, is that if one supposes that the values set by experts are ideal, that is valuation $x_i = 1$, for all i then the following formula will always be reached:

$$E(x_i, y_i) = \left(\sum_{i=1}^{n} w_i y_i^m \right)^{\frac{1}{m}}, \tag{63}$$

where $E(x_i, y_i)$ is the total sameness between all fuzzified criteria for the corresponding data. This follows from the following:

1. implications have neutrality of the truth $I(1, b) = b$ and boundary condition ii) $I(a, b) = 1$ iff $a \leq b$
2. every t-norm must satisfy the boundary condition, that is $T(1, y_i) = y_i$.

When the membership value $y_i = \mu_{Aet_i}^{p_i}$ for aerobic and $y_i = \mu_{Ant_i}^{p_i}$ for anaerobic are set, then the modifying factor p_i tunes the shape of the membership function i. This corresponds to the real situation where for, example, experts' subjective opinions affect the final decision. Index i shows the number of corresponding membership functions, which mimic the criteria. This gives the following two equations:

$$E_{Aet}\left(x_i, \mu_{Aet_i}^{p_i}\right) = \left[\sum_{i=1}^{n} w_i \mu_{Aet_i}^{p_i \cdot m} \right]^{\frac{1}{m}} \tag{64}$$

for the aerobic threshold and

$$E_{Ant}\left(x_i, \mu_{Ant_i}^{p_i}\right) = \left[\sum_{i=1}^{n} w_i \mu_{Ant_i}^{p_i \cdot m} \right]^{\frac{1}{m}} \tag{65}$$

for the anaerobic threshold.

3.1.3 Use of Differential Evolution for Finding the Right Parameters

Differential evolution (DE) is used for finding the correct parameter values for the system. DE is a simple population based stochastic function minimizer. The objective of DE is to iterate each member of the population and compare its value to the trial member value, and the superior member stays for the next iteration. The evolution strategy defines the way in which a trial member is generated. DE tries to seek parameters that will give the maximal similarity compared to the values set by experts. This is done so that DE tries to minimize the value of the objective function with trial member values. The objective function is the total difference between the thresholds defined by experts and the thresholds defined by similarity used here for all learning data sets. Finally, DE gives the optimal parameter values. The basic action of used differential evolution is demonstrated in figure (3). Figures (4) and (5) show that differential evolution reached the optimal solution rather quickly. The solution, in this case, is the mean error of classification with the current training data set.

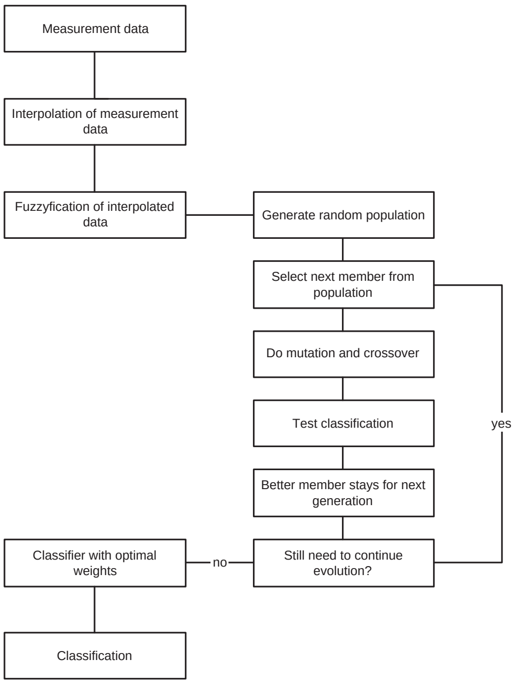

Figure 3: Simplified computational model for DE

After the optimization, the parameters shown in Table (15) were reached for the aerobic threshold and the parameters shown in Table (16) for the anaerobic threshold.

Table 15: Optimized parameters, aerobic

i	$\mathbf{w_i}$	$\mathbf{p_i}$	**m**
1	3.1394	9.4039	0.2307
2	45.0996	3.0604	
3	25.7102	4.7183	
4	18.4844	0.6633	
5	0.2418	0.6976	
6	7.2322	6.0769	
7	0.0925	5.6312	

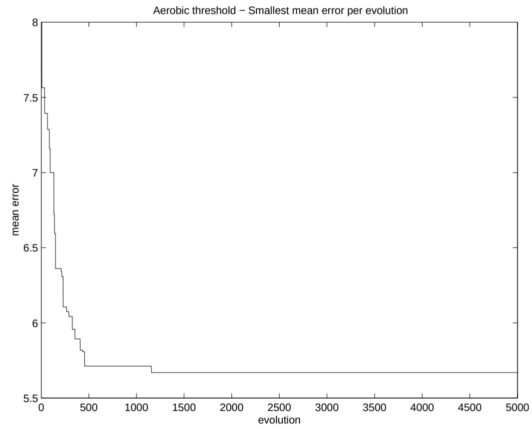

Figure 4: Mean error development for aerobic

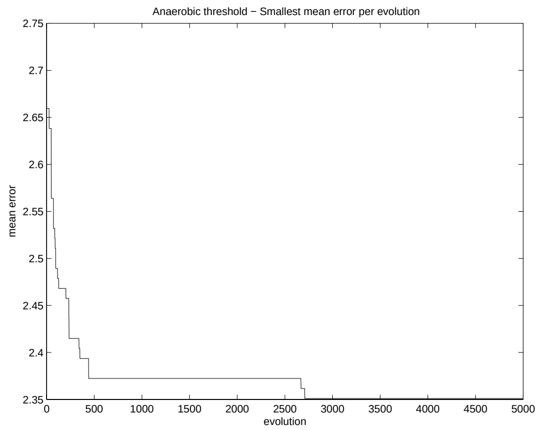

Figure 5: Mean error development for anaerobic

3.1.4 Final Model

The making of this computational model consisted of the following steps:

1. The interpolation of the measurement data, which interpolates data on the pulse and the ventilation of oxygen.

2. Setting the criteria for aerobic and anaerobic thresholds in co-operation with sports medicine experts.

3. The selection of the membership functions which correspond as well as possible to the fuzzy criteria set by experts.

4. The fuzzification of the interpolated data which fuzzifies data with the membership functions.

Table 16: Optimized parameters, anaerobic

i	$\mathbf{w_i}$	$\mathbf{p_i}$	**m**
1	0.0967	9.3537	0.4517
2	33.4295	9.7802	
3	23.6398	3.0888	
4	6.9080	0.8098	
5	34.5696	2.8369	
6	1.3564	0.1598	

5. Counting the partial similarities and combining them to the total similarity.
6. Estimating the parameters for partial similarities with differential evolution and combining the best parameters into the model.
7. The model is ready for use.

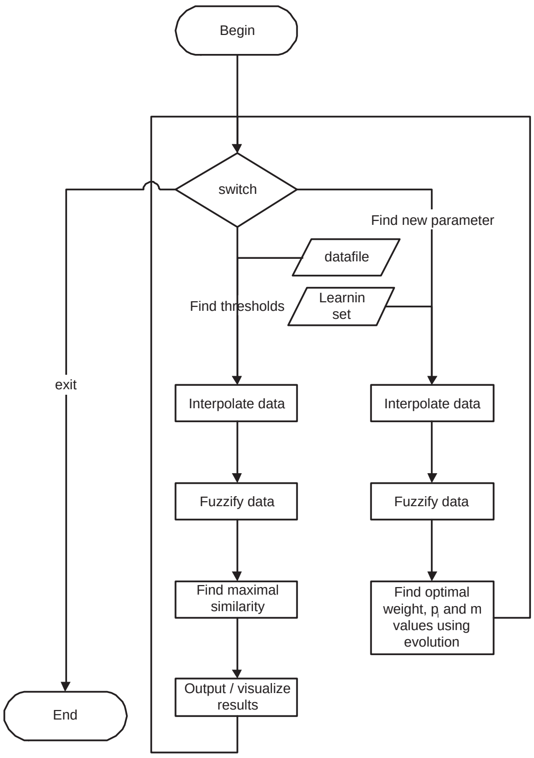

Figure 6: Computational model for the system

From the flow chart of the final model in Figure (6) it can be seen that with some measurement data the computations needed can be done in a quite straightforward manner because the interpolation used is linear, and finding the right evaluation result is a simple task of finding the maximal combination of membership values. This means that once the correct parameter values are estimated, this system works rapidly.

3.1.5 Results and Discussion

The following section considers the statistical effects of the parameters in equivalence for the thresholds estimated by this system versus the thresholds given by experts. Tables (17) and (18) show the comparison statistics from the optimal values set by experts and calculations with three different mean measures, arithmetic, geometric and harmonic with and without weights vs. newest model. In this latest model, in addition to weights an individual parameter p_i for different membership values and the generalized mean were also used. The following notations have been used in tables (17) and (18). The word **expert** means a thresholds heartbeat value, which is defined by the expert. Letters **A**, **G** and **H** illustrate the corresponding arithmetic, geometric and harmonic averages that are used in the measure to achieve estimates for the heartbeat, **noW** means that no optimized weights have been used, so that all weights equal 1 and **W** means the weight. The word **final** means the heartbeat thresholds that have been obtained from the most recent model. Test called the multiple range test were used, which in this case was Fisher's least significant difference (LSD) method, an approach suggested by the statistician R. A. Fisher. This procedure carries out pair-wise comparisons using the variances for the groups chosen for the test.

Table 17: Fisher's least significant differences (LSD), aerobic (* denotes a statistically significant difference.)

expert	-	final	1.6755
AnoW	-	final	*-12.2185
AW	-	final	-0.450331
GnoW	-	final	*-20.6755
GW	-	final	*-20.2517
HnoW	-	final	*-6.1457
HW	-	final	*-3.72848

It can be noticed that the model with individual parameter values and a generalized mean give the best results. One can also notice that in most of the cases the results that obtained by using ordinary means such as arithmetic, geometric or harmonic usually differ significantly from the results obtained with optimized mean value. In all cases, the weighted version was better than the non-weighted. For the aerobic threshold, the results did not show any statistically significant difference when

Table 18: Fisher's least significant differences (LSD), anaerobic (* denotes a statistically significant difference.)

expert	-	final	0.298013
AnoW	-	final	-0.966887
AW	-	final	-0.543046
GnoW	-	final	*-59.2119
GW	-	final	*-59.2119
HnoW	-	final	-0.211921
HW	-	final	0.152318

weights were applied for the similarity measure at a 95.0 % confidence level. For the anaerobic threshold the results were better when weights were used, but already the un-weighted version of this system worked so well that there was no statistically significant difference to the expert values at a 95.0 % confidence level.

All the way through, the results were very good and there were no statistically significant differences between the results estimated by this system and the results estimated by experts. In conclusion, this model can be regarded as a fully developed working system to determine an athlete's aerobic and anaerobic thresholds.

Obviously the use of three parameters - weights, modification factors and a generalized mean - improved the accuracy of the forecasted thresholds.

3.2 Medical Data Classification Using Comparison Measures

In this chapter, the kind of results that can be established using comparison measures in classification of some well-known medical data sets, are studied. Results are compared in comparison with some known results and one can see from the results presented that the comparison measures used are able to produce the results better; pointing out the importance of choosing the right comparison measures.

3.2.1 Data sets and used comparison measures

Medical data sets are to be tested with equivalence based comparison measures that have been presented in this thesis so they are of the form $E(x, y) = T(S(\neg x, y), S(\neg y, x))$, where T and S present the corresponding pairs of t-norms and t-conorms. The comparison measures tested were the combined Łukasiewicz and Shweizer & Sklar based comparison measure with a generalized mean presented in (43), the Kleene-Dienes based comparison measure with a generalized mean presented in (36), the Reichenbach based comparison measure with a generalized mean presented in (39), the Łukasiewicz based comparison measure with a generalized mean presented in (42). The data sets chosen for the test were: Pima Indians diabetes, Thyroid gland, BUPA liver disorders and Wisconsin Diagnostic Breast Cancer (WDBC). They were all derived from medical sources.

- **Pima, row 1**: The diagnostic, binary-valued variable investigated is whether the patient shows signs of diabetes. All instances here are females of Pima Indian heritage of at least 21 years of age. The number of instances is 768. The number of attributes is 8 plus the class. Class 1 (negative for diabetes) 500, Class 2 (positive for diabetes) 268.

  Past usage Zhou Z.H. and Jiang Y. [120] etc.

- **Thyroid, row 2**: Five laboratory tests are used to try to predict whether a patient's thyroid belongs to the class euthyroidism, hypothyroidism or hyperthyroidism. The diagnosis (the class label) was based on a complete medical record, including anamnesis, scan etc. The number of instances is 215. The number of attributes is 5 plus the class. Class 1 (normal) 150, Class 2 (hyper) 35 and Class 3 (hypo) 30.

  Past usage Coomans, D. et. al. [12], [13] etc.

- **BUPA, row 3**: The first 5 variables are all blood tests which are thought to be sensitive to liver disorders that might arise from excessive alcohol consumption. Each line in the bupa.data file constitutes the record of a single male individual. The number of instances is 345. The number of attributes is 6 plus the class. Class 1 145 and Class 2 200.

  Past usage Bologna G. [9] etc.

- **WDBC, row 4**: Wisconsin Diagnostic Breast Cancer. The features are computed from a digitized image of a fine needle aspirate (FNA) of a breast mass. They describe characteristics of the cell nuclei present in the image. The number of instances is 569. The number of attributes is 30, including the mean, standard error and maximum value from 10 different measures, plus the class. Class 1 (benign) 357 and 2 (malignant -dangerous) 212.

  Past usage, Wolberg W. H. and Mangasarian O. L. [108].

3.2.2 Classification

Classification in this study was carried out by using a comparison measure based classifier described in the introduction part of this thesis in Figure (1). For each parameter value p and m, corresponding weight values were randomly chosen 200

Table 19: Maximum classification accuracy

	SS1	KD	RB	Ł
Pima	0.8021	0.7969	0.7969	0.7943
Thyroid	1.0000	1.0000	1.0000	1.0000
BUPA	0.9942	0.9942	0.9942	0.9942
WDBC	0.9860	0.9825	0.9825	0.9895

times. Results achieved are compared with the results that were achieved in the articles [120], [12], [13], [9] and [108].

The following are very short descriptions of classifiers that are used for comparison of classification results here, although better descriptions can be found from Bologna G. [9].

- C4.5 is a popular decision tree classifier.
- LDA is described as a linear discriminant analysis classifier.
- MLP is a multi-layer perceptron based classifier.
- DIMLP is a discrete interpretable multi-layer perceptron based classifier.

3.2.3 Results and discussion

Below is a table displaying the maximum and average classification results and a brief discussion of the results.

In the tables (19), (20) and (21) SS1 mean that the combined Łukasiewicz and Schweizer & Sklar based comparison measure presented is used in (43), KD means that the Kleene-Dienes based comparison measure presented in (36) is used, RB means that the Reichenbach based comparison measure presented in (39) is used and Ł means that the Łukasiewicz based comparison measure presented in (42) is used. From Table (19) the maximal classification result can be seen using the optimal parameters of p and m. From Table (20) can be seen the average classification results, that are the means of classification results over all p- and m-values. From Table (21) can be seen the variances of classification results, which show that all equivalences are also able to give quite stable results.

From the Table (22) it can be seen that equivalence based comparison measures are able to give better results than other methods used before for classification in Bologna G. [9], where reported results were better than results reported in [120], [12], [13] or

Table 20: Average classification accuracy

	SS1	KD	RB	L
Pima	0.7659	0.7607	0.7682	0.7643
Thyroid	0.9741	0.9694	0.9667	0.9769
BUPA	0.9249	0.6861	0.7110	0.6884
WDBC	0.9611	0.9493	0.9369	0.9490

Table 21: Maximum variance in classification

	SS1	KD	RB	L
Pima	0.0037	0.0089	0.0066	0.0029
Thyroid	0.1196	0.1097	0.0801	0.1106
BUPA	0.0750	0.0194	0.0255	0.0203
WDBC	0.0051	0.0212	0.0338	0.0295

[108]. Except when Coomans D. uses in [13] a different kernel density methods, some of which achieve 100 % correct classification for Thyroid data.

It can seen from the (22) that the best result for the Pima data set is 3.42 % better than best result achieved in the articles used for comparison. For the thyroid data set the best result is 3.76 % better than best result achieved in the articles used for comparison. For the BUPA data set the best result is 28.71 % better than best result achieved in the articles used for comparison. For the WNBC data set the best result is 1.52 % better than best result achieved in the articles used for comparison.

Figures (7) and (8) show typical classification results from which one can see that comparison measures give large areas where results remain very stable. Classification results and the best p- and m-values are dependent on the data set used.

Table 22: Classifier comparison

	LDA	C4.5	MLP	DIMLP	CM
Pima	0.7640	0.7380	0.7638	0.7679	**0.8021**
Thyroid	0.8134	0.9326	0.9624	0.9486	**1.000**
BUPA	0.6912	0.6657	0.7023	0.7071	**0.9942**
WDBC	0.9719	0.9406	0.9743	0.9692	**0.9895**

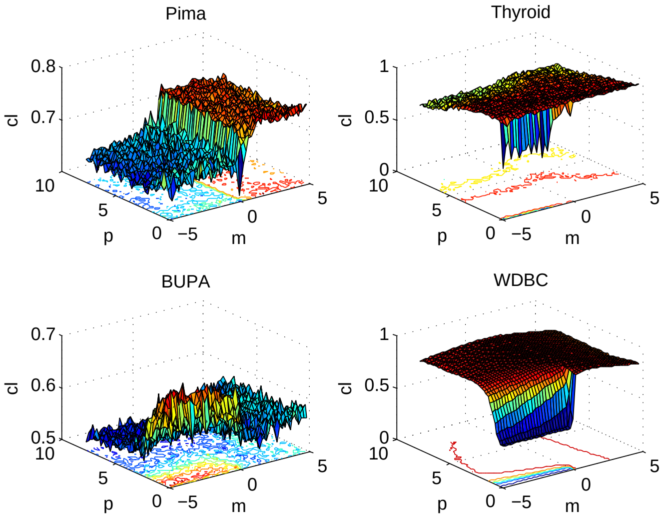

Figure 7: Average classification results using Łukasiewics comparison measure

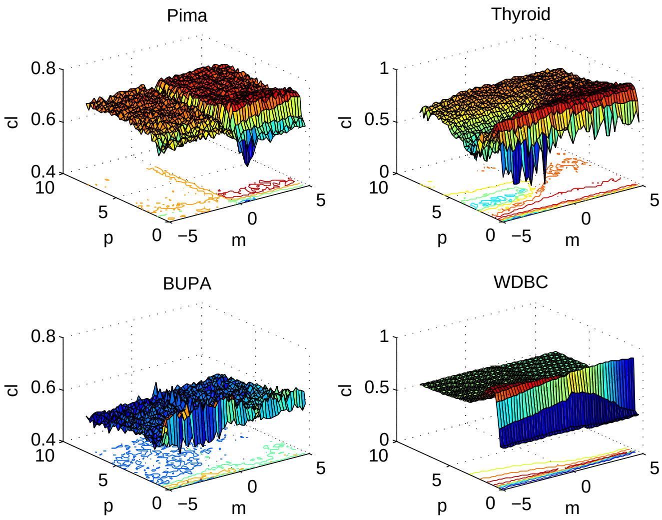

Figure 8: Average classification results using Reichenbach comparison measure

In this chapter illustrates that the results achieved by using comparison measures presented in this thesis, see Table (22), were either better or the same [13] in classification tasks chosen as comparisons, than the classification results presented in articles [120], [12], [9] and [108]. One can also see from Figure (7) and Table (21) that normally these classification results remain quite stable with any parameter values.

4 Conclusions

In basically every area of human life some kind of measures for comparison are needed. In soft computing, these areas include classification, pattern recognition, clustering, expert systems, medical diagnosis systems, decision support systems, and fuzzy control. Recently, new areas have been generated, for example, in web-search engines, where information retrieval is of high importance. There are several different logical operators presented in the literature of fuzzy logic for example in [48], [119], [35], [3] and [18]. Some have also suggested ways to create operators directly from the data at hand [4], [5] and [6]. Many authors have presented operators that can be used for comparison purposes and even claim that these work well in practice, but normally they show no data to support the claim that they actually work in practice. In this thesis, all comparison measures have been tested with some practical data, and it has been shown that they in fact give reasonable results when applied.

In classification and the development of expert systems, the problem of choosing the right functions for comparison is often faced. When data has different dependencies, different operators should be used. Usually the simplest operators are selected, which are not normally the optimal choice. As a solution to this problem this thesis has offered different approaches for creating comparison measures, on a logically sound basis. It has been shown that these comparison measures give reasonable results. Even if these comparison measures always only concentrated on a few pieces of data at a time, and not, for example, on membership functions or deviations, they gave good results, which supports the claim that perhaps people handle the concept of sameness in a similar way [121].

4.1 Comparison Measures

It is also suggested that the comparison measures used in a fuzzy sets, where comparison is done feature by feature and then these comparisons are aggregated, could actually be any measures which fulfills the properties (41), listed below:

1. The comparison measure used has a clear logical structure e.g. it is an Archimedean t-norm or t-conorm (like Frank (49), (54)) or S-equivalence ((35), (42), (39)).

2. The comparison measure is monotone. This condition ensures that a decrease (or increase) in any values that are to be compared cannot produce an increase (or decrease) in the comparison result.

3. The comparison measure is associative. This guarantees that the final comparison results are independent of the grouping of the arguments and that one can expand these comparisons to more than two arguments.

4. The comparison measure is continuous. This guarantees that one can safely compute with the values that are to be compared.

Three methods to create comparison measures from t-norms and t-conorms, implications and equivalences and from 3Π-uninorm are presented in this thesis.

Furthermore, several new generalized measures it is introduced starting from t-norms, t-conorms, 3Π-uninorm and pseudo equivalences based on the use of S-implications. The Generalized Weighted T-norm measure (44) and (47) were created. A new parameterized 3Π-operator was given (61). Several new measures based on S-type pseudo equivalences were also given (36), (39), (42) and (43).

In order to give freedom and adaptability for the comparison tasks weights and generalized mean are used with comparison measures. From the generalized mean one can find *min* and *max* operators by giving the generalized mean the lowest and correspondingly highest compensation values. This is a theoretically interesting property since these are also the upper and lower bounds of the t-norm and t-conorm ranges, respectively, and these operators are also the only ones satisfying distributivity and idempotency [7]. One can also see that comparison measures quite satisfactorily fulfil the criteria presented in (42). Properties of the created comparison measures are mainly dependent of the properties of the used t-norms and t-conorms. For example, continuity follows directly from continuity of the used t-norms and t-conorms that is if they are continuous then the following comparison measures will also be. One has to of course be careful with m-value 0, when using a generalized mean.

Due the algorithm that is used, optimal mean values wrt. data and comparison measures will automatically be selected.

4.2 Practical Results

One can see that the classification presented in (6) and in (1) is a quite straightforward instance based classification. Classification is seen as a cyclic comparison between the training set and test set. The parameter values for the operators and the generalized mean are predefined for a selected interval with steps. For weight optimization differential evolution was used. Differential evolution gives a weight vector where every attribute or property of the data has its own weight. Actually, this weight vector can be used for reading how important different properties measured are for the conclusion; this ought to be valuable data for practitioners. It is also noted that this procedure gives comparison measures high adaptability for the data in hand.

It has been shown that the comparison measures introduced in this thesis consistently give good and stable results in classification, which can be seen from the following tables which show variances, means and maximal results (3), (4), (5), (6), (7), (8), (9), (10), (11), (12), (13), (14). For example, when one uses a combined measure (47), that is the Generalized Weighted T-norm operator in the classification, the results improved a great deal more than by using t-norms or t-conorms alone. For example, combined comparison measures (47) based on Frank type of t-norm (49) and t-conorm (54) gave the best classification results, which are the same or better than those attained from the pseudo equivalences. All but one of these best results achieved by Frank measures came from the use of a combined measure. Pseudo equivalences also gave good and stable results in classification. One can also see that the improvements in classification results due to changing to the right comparison measures were quite significant.

From the tested combined comparison measures (47) use of a combination of Frank type t-norm and t-conorm is recommended. When using logical equivalences Łukasiewicz type (42) is to be recommended since it is computationally more effective than the comparison measure based on Shweizer & Sklar - Łukasiewicz (43).

In general, logical equivalences give better results than combined comparison measures therefore logical equivalences are recommended over combined comparison measures.

The comparison measures created were tested with data coming from several different disciplines, but one particular type of data namely the medical data was given special attention. It was shown that the results achieved by using these simple comparison measures, see table (22), were mainly better in classification tasks chosen as an example of comparison than most of the classification results found in [9]. One can also see from Figure (7) and Table (21) that normally these classification results remained quite stable with any parameter values.

Comparison measures were also applied to the expert system and a working application for defining an athlete's aerobic and anaerobic thresholds was created. Differential evolution presented in figure (3) was now used to tune free parameters of used comparison measures. Figures (4) and (5) show that differential evolution reached the optimal parameter values rather rapidly.

One result concerning the choice of comparison measure was that if it is supposed that the values set by experts are ideal, that is valuation $x_i = 1$, for all i the resulting formula is always:

$$E\left(x_i, y_i\right) = \left(\sum_{i=1}^{n} w_i y_i^m\right)^{\frac{1}{m}}, \tag{66}$$

where $E\left(x_i, y_i\right)$ is the total sameness between all fuzzified criteria for the corresponding data. This follows from the following:

1. implications have neutrality of the truth $I(1, b) = b$ and boundary condition $I(a, b) = 1$ iff $a \leq b$

2. every t-norm must satisfy the boundary condition, that is $T(1, y_i) = y_i$.

The results were consistently very good and there were no statistically significant differences between the results estimated by this system and the results estimated by experts. In conclusion, it can be stated that a working system to determine an athlete's aerobic and anaerobic thresholds was developed.

This thesis presented some guidelines concerning which measures used for comparison ought to have, as well as creating several comparison measures and presenting two application areas. The results achieved were consistently relatively good.

I hope that this thesis will be used for the guidance of practitioners in the creation and use of fuzzy logical comparison measures and an inspiration for the theoretical study of comparison.

References

[1] Atkeson C.G., Moore A.W. and Schaal S., Locally Weighted Learning, *Artificial Intelligence Review*, Springer Netherlands, vol. 11, pp. 1-5, 1997.

[2] Baeza-Yates R., Ribeiro-Neto B., *Modern Information Retrieval*, Addison Wesley/ACM Press, New York, 1999.

[3] Bandemer H. and Näther W., *Fuzzy data analysis*, Kluwer Acad. Publishers, Dordrecht, 1992.

[4] Beliakov G., Definition of general aggregation operators through similarity relations, *Fuzzy Sets and Systems*, 114 (3), pp. 437 - 453, 2000.

[5] Beliakov G., How to build aggregation operators from data, *Int. J. Intell. Syst.*, 18(8), pp. 903-923, 2003.

[6] Beliakov G., Mesiar R. and Valaskova V., Fitting Generated Aggregation Operators To Empirical Data, *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 12(2), pp. 219-236, 2004.

[7] Bellman R. and Giertz M., On the Analytic Formalism of the Theory of Fuzzy Sets, *Information Sciences*, 5, pp. 149-165, 1973.

[8] Bilgiç T. and Türkşen I.B., Measurement-theoretic Justification of Connectives in Fuzzy Set Theory, *Fuzzy Sets and Systems*, 76 (3), pp. 289-308, 1995.

[9] Bologna G., A Study on Rule Extraction from Neural Networks Applied to Medical Databases, *Proceedings of The 4^{th} European Principles and Practice of Knowledge Discovery in Databases*, Lyon, France, 2000.

[10] Chang C.C., Algebraic analysis of many-valued logics, *Trans. Amer. Math. Soc.*, 88, pp. 467-490, 1958.

[11] Chellas B.F., *Modal Logic: An Introduction*, Cambridge University Press, Cambridge, 1980.

[12] Coomans D., Broeckaert M., Jonckheer M. and Massart D.L., Comparison of Multivariate Discriminant Techniques for Clinical Data - Application to the Thyroid Functional State, *Meth. Inform. Med.* 22, pp. 93-101, 1983.

[13] Coomans D. and Broeckaert I, *Potential pattern recognition in chemical and medical decision making*, John Wiley & Sons, Inc. New York, NY, USA, 1986.

[14] De Cock M. and Kerre E. E., Why Fuzzy T-Equivalence Relations do Not Resolve the Poincaré Paradox, and Related Issues, *Fuzzy Sets and Systems*, 133(2), pp. 181-192, 2003.

[15] De Morgan A., *Formal logic or The calculus of inference: necessary and probable*, Taylor and Walton, London, 1847, Elibron Classics series, King's College London, 2005.

[16] Dombi J., Basic Concepts for a theory of evaluation: The aggregative operator, *European Journal of Operational Research*, 10, pp. 282-293, 1982.

[17] Douglas C.G., Respiration and circulation with variations in bodily activity, *Lancet*, 1, pp. 213-218, pp. 265-269, 1927.

[18] Dubois D. and Prade H., A review of fuzzy set aggregation connectives, *Information Sciences*, 36, pp. 85-121, 1985.

[19] Dubois D. and Prade H., *Similarity-Based Approximate Reasoning*, in: Computational Intelligence Imitating Life, Zurada, J.M., Marks II, R.J., Robinson, C.J., eds., IEEE Press, New York, pp. 69-80, 1994.

[20] Duda R. and Hart P., *Pattern Classification and Scene Analysis*, John Wiley & Sons, 1973.

[21] Dyckhoff H. and Pedrycz W., Generalized Means as Model of Compensative Connectives, *Fuzzy Sets and Systems*, 14, pp. 143-154, 1984.

[22] Fisher N.I., Copulas, In Kotz S., Read C.B., and Banks D.L. eds., Update Vol. 1, pp. 159-163, *Encyclopedia of Statistical Sciences*, John Wiley & Sons, New York, 1997.

[23] Fodor J., On fuzzy implication operators, *Fuzzy Sets and Systems*, vol. 42 (3), pp. 293-300, 1991.

[24] Fodor J., Nilpotent minimum and related connectives for fuzzy logic, *Proceedings of 1995 IEEE International Conference on Fuzzy Systems and The Second International Fuzzy Engineering Symposium*, vol. 4, pp. 2077-2082, 1995.

[25] Fodor J., Yager R.R. and Rybalov A., Structure of uninorms, *International Journal of Uncertainty, Fuzziness and Knowledge Based Systems*, 5, pp. 411-427, 1997.

[26] Fodor J. and Yager R. R., Fuzzy set-theoretic operators and quantifiers, in: *Fundamentals of Fuzzy Sets*, edited by Dubois, D. and Prade, H., Kluwer Academic Publishers: Norwell, Ma, pp. 125-193, 2000.

[27] Formato F., Gerla G.and Scarpati L., Fuzzy Subgroups and Similarities, *Soft Computing*, 3, 1999.

[28] France R. K., Weights and measures: An axiomatic model for similarity computations, *Technical report, Virginia Polytechnic Institute and State University*, 1993.

[29] Frank M.J., On the simultaneous associativity of $F(x,y)$ and $x+y-F(x,y)$, *Aeqationes Math.*, (19), pp. 194-226, 1979.

[30] Frink O. Jr., New Algebras of Logic, *The American Mathematical Monthly*, 45, pp. 210-219, 1938.

[31] Goldberg D.E., Real-coded genetic algorithms, virtual alphabets, and blocking, *Technical Report 9001*, University of Illinois at Urbana- Champain, 1990.

[32] Gottwald S., Approximate solutions of fuzzy relational equations and a characterization of t-norms that define metrics for fuzzy sets, *Fuzzy Sets and Systems*, 75(2), pp. 189 - 201, 1995.

[33] Grefenstette J.J., Optimization of control parameters for genetic algorithms, *IEEE Transactions on Systems, Man and Cybernetics*, 16(1), pp. 122-128, 1986.

[34] Hájek P., Fuzzy logic and arithmetical hierarchy, *Fuzzy Sets and Systems*, 73, pp. 359-363, 1995.

[35] Hájek P., *Metamathematics of fuzzy logic*, Kluwer Acad. Publishers, Dordrecht, 1998.

[36] Hastie T. and Tibshirani R., *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, Springer Series in Statistics, Springer, New York, 2001.

[37] Hardy G., Littlewood J. and Polya G., *Inequalities*, Cambridge University Press, (1934), second edition, 1952.

[38] Haykin S. *Neural Networks: A Comprehensive Foundation, 2^{nd} Edition*, Prentice Hall, 1998.

[39] Höhle U. and Stout L.N., Foundations of Fuzzy Sets, *Fuzzy Sets and Systems*, 40(2), pp. 257-296, 1991.

[40] James W., *The Principles of Psychology*, Dover: New York, Vol. 1, Chapter 7, 1890/1950.

[41] Jain R., Murthy S. N., Tran L. and Chatterjee S., Similarity Measures for Image Databases, *SPIE Proceedings, Storage and Retrieval for Image and Video Databases*, 58-65 1995.

[42] Kang T. and Chen G., Modifications of Bellman-Giertz's theorem, *Fuzzy Sets and Systems*, Vol. 94, Issue 3 (16), pp. 349-353, 1998.

[43] Ketola J., Saastamoinen K. and Turunen E., Tuning the Parameters for the Decision Making System in Order to Define Athlete's Aerobic and Anaerobic Thresholds, *Proceedings of the International Conference on Control, Automation and Systems, ICCAS 2004*.

[44] Kinderman W., Simon G. and Keul J., The significance of the aerobic-anaerobic transition for the determination of work load intensities during endurance training, *Eur. J. Appl. Physiol.*, 42, pp. 25-34, 1979.

[45] Klawonn F. and Castro J.L., Similarity in Fuzzy Reasoning, *Mathware and Soft Computing*, 3(2), pp. 197-228, 1995.

[46] Klement E. P., Mesiar R. and Endre P., Triangular norms: Basic notions and properties, in: Klement E. P. and Mesiar R.(eds.), *Logical, algebraic, analytic and probabilistic aspects of triangular norms*, Elsevier, pp. 17-60, 2005.

[47] Klir G.J., Multivalued logics versus modal logics: alternative frameworks for uncertainty modelling, in: P.P. Wang (Ed.), *Advances in Fuzzy Theory and Technology*, Department of Electrical Engineering, Duke University, Durham, North Carolina, pp. 3-47, 1994.

[48] Klir G.J., Yuan B., *Fuzzy Sets and Fuzzy Logic: Theory and Applications*, Prentice Hall, PTR, pp. 304-321, 1995.

[49] Kundu S. and Chen J., Fuzzy logic or Lukasiewicz logic: A clarification, *Fuzzy Sets and Systems*, 95, pp. 369-379, 1998.

[50] Kuncheva L.I., How good are fuzzy if-then classifiers?, *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, 30 (4), pp. 501-509, 2000.

[51] Ling C.H., Representation of associative functions, *Publ. Math. Debrecen*, 12, pp. 189-212, 1965.

[52] Lowen R.R., *Fuzzy Set Theory: Basic Concepts, Techniques, and Bibliography*, Kluwer Acad. Publishers, Dordrecht, 1996.

[53] Łukasiewicz J., *Selected Works Studies in Logic and The Foundations of Mathematics*, North-Holland, Amsterdam, 1970.

[54] Luukka P., Saastamoinen K. and Könönen V., A Classifier Based on the Maximal fuzzy Similarity in the Generalized Łukasiewicz Structure, *Proceedings of The 10th IEEE International Conference on Fuzzy Systems*, Melbourne, Australia, ISBN: 0-7803-7294-X, 2001.

[55] Luukka P. and Saastamoinen K., Fuzzy Similarity Based Mutation in Genetic Algorithm, *Joint 1st International Conference on Soft Computing and Intelligent Systems and 3rd International Symposium on Advanced Intelligent Systems*, Tsukuba Japan, Oct 21-25, 2002.

[56] Luukka P. and Saastamoinen K., Fuzzy Similarity Based Classification in the Normal Łukasiewicz-Structure, *Proceedings of KES 2002, Knowledge-based Intelligent Information Engineering Systems and Allied Technologies, Frontiers in Artificial Intelligence and Applications*, Vol. 82, ISBN: 1 58603 280 1, pp. 974-981, 2002.

[57] Luukka P., Saastamoinen K. and Louro S., Classification of High Dimensional Data with Arithmetic, Geometric and Harmonic Fuzzy Similarity, *Part I - Proceedings of the 10^{th} IFSA World Congress*, ISBN: 975-518-208-X, pp. 257-260, 2003.

[58] Luukka P. and Saastamoinen K., Fuzzy similarity based classification in normal Łukasiewicz algebra with geometric and harmonic mean, *Proceedings of the 3^{rd} Eusflat Conference*, ISBN: 3-9808089-4-7, pp. 242-245, 2003.

[59] MacColl H., The calculus of equivalent statements, *Proceedings of the London Mathematical Society*, pp. 156-183, 1896-1877.

[60] Mantel B., Periaux J. and Sefrioui M., Gradient and genetic optimizers for aerodynamic desing, *ICIAM 95 Conference*, Hamburgh, 1995.

[61] Mattila J., On Łukasiewicz Modifier Logic, *Journal of Advanced Computational Intelligence and Intelligent Information*, Vol. 9, 5, 2005.

[62] Menger K., Statistical Metrics, *CProc. Nat. Acad. Sci.*, U.S.A., 37, pp. 535-537, 1942.

[63] Michalewics Z., *Genetic algorithms + data structures = evolution programs Artificial Intelligence*, Springer-Verlag, New York, 1992.

[64] Miyamoto S., *Fuzzy Sets in Information Retrieval and Cluster Analysis*, Kluwer Academic Publishers, Netherlands, 1990.

[65] Novak V., On the syntactico-semantical completeness of first-order fuzzy logic, I, II, *Kybernetika*, 26, pp. 47-66, pp. 134-152, 1990.

[66] Pal S.K. and D.K. Dutta-Majumder, *Fuzzy Mathematical Approach to Pattern Recognition* John Wiley & Sons (Halsted), N. Y. 1986.

[67] Pao Y.-H., *The Pattern Recognition and Neural Networks*, Addison-Wesley Publishing Company, Inc, 1989.

[68] Pavelka J., On fuzzy logic l, II, III, *Z. Math. Loqik*, 25, pp. 45-52, pp. 119-134, pp. 447-464, pp. 1979.

[69] Pawlak Z., *Rough Sets: Theoretical Aspects of Reasoning About Data*, Kluwer, 1991.

[70] Peirce C.S., On the Algebra of Logic: A Contribution to the Philosophy of Notation, *American Journal of Mathematics*, Vol. 7, pp. 180-202, 1885.

[71] Price K.V., Storn R.M., Lampinen J.A., *Differential Evolution - A Practical Approach to Global Optimization*, Springer, Natural Computing Series, 2005.

[72] Rose A., Rosser J.B., Fragments of Many-Valued Statement Calculi, *Transactions of the American Mathematical Society*, 87, pp. 1-53, 1958.

[73] Saastamoinen K. and Luukka P., Classification in the Łukasiewicz-Structure with Different Means, *Joint 1st International Conference on Soft Computing and Intelligent Systems and 3rd International Symposium on Advanced Intelligent Systems*, Tsukuba Japan, Oct 21-25, 2002.

[74] Saastamoinen K. and Luukka P., Comparison of the Fuzzy Similarity Based Classification in the Normal and the Generalized Łukasiewicz-Structure, *Proceedings of KES 2002, Knowledge-based Intelligent Information Engineering Systems and Allied Technologies, Frontiers in Artificial Intelligence and Applications*, Vol. 82, ISBN: 1 58603 280 1, pp. 974-981, 2002.

[75] Saastamoinen K. and Luukka P. and Könönen V., Weighted Similarity Classifier in Generalized Łukasiewicz Structure, *Advanced in Intelligent Systems, Fuzzy Systems, Evolutionary Computation, - Interlaken Switzerland, A Series of Reference Books and Textbooks*, ISBN: 960 8052 49 1, pp. 128-135, 2002.

[76] Saastamoinen K., Könönen V. and Luukka P., A Classifier Based on the Fuzzy Similarity in the Łukasiewicz-Structure with Different Metrics, *Proceedings of the FUZZ-IEEE 2002 Conference*, Hawaii USA, 2002.

[77] Saastamoinen K. and Luukka P., Testing Continuous t-norm called Łukasiewicz Algebra with Different Means in Classification, *Proceedings of the FUZZ-IEEE 2003 Conference*, St Louis, USA.

[78] Saastamoinen K. and Sampo J., Use of Hybrid Method in Wavelets Bases Selection for Signal Compression, *Proceedings of the SPIE Annual Meeting 2003*, San Diego, USA.

[79] Saastamoinen K. and Luukka P., Comparison of the Use of Different Similarities Based on t-norms in the Classification Tasks, *Proceedings of the 3rd Eusflat Conference*, ISBN: 3-9808089-4-7, pp. 246-247, 2003.

[80] Saastamoinen K. and Luukka P., Cumulative Similarity Measure Based Classifier in the Łukasiewicz-Structure, *Part I - Proceedings of the 10th IFSA World Congress*, ISBN: 975-518-208-X, pp. 249-252, 2003.

[81] Saastamoinen K. and Sampo J., Use of Different Similarity Measures with Hybrid Method in Wavelets Bases Selection for Signal Compression, *Proceedings of the Fourth International Conference on Intelligent Technologies (InTech'03)*. ISBN: 974-658-151-1, pp. 227-232, 2003.

[82] Saastamoinen K., On the Use of Generalized Mean with T-norms and T-conorms, *Proceedings of the IEEE 2004 Conference on Cybernetic and Intelligent Systems*, Singapore, 2004.

[83] Saastamoinen K. and Sampo J., On General Class of Parameterized $3\pi-$ Uninorm Based Comparison, *WSEAS TRANSACTIONS on MATHEMATICS*, 3(3), pp. 482-486, 2004.

[84] Saastamoinen K., Semantic Study of the Use of Parameterized S-Implications and Equivalences in Comparison, *Proceedings of the IEEE 2004 Conference on Cybernetic and Intelligent Systems*, Singapore.

[85] Saastamoinen K. and Ketola J., Defining Athlete's Anaerobic and Aerobic Thresholds by Using Similarity Measures and Differential Evolution, *Proceedings of the IEEE SMC 2004 conference*, Hague, Netherlands.

[86] Saastamoinen K., Survey of the Use of Uninorms and Some Related Structures as Similarities, *Fuzziness in Finland 2004*.

[87] Saastamoinen K. and Ketola J., Using Generalized Combination Measure from Dombi and Yager type of T-norms and T-conorms in Classification, *Proceedings of the ECTI-CON 2005 conference*.

[88] Saastamoinen K. and Ketola J., Fuzzy Logic and Differential Evolution Based Expert System for Defining Top Athlete's Aerobic and Anaerobic Threshold, *Journal of Advanced Computational Intelligence and Intelligent Informatics*, Vol 9(5), 2005.

[89] Saastamoinen K. and Ketola J., Medical Data Classification using Logical Similarity Based Measures, *Proceedings of the IEEE 2006 Conference on Cybernetic and Intelligent Systems*.

[90] Saastamoinen K., On New General Class of $3\pi-$ Uninorms, *Proceedings of the IEEE-ICCC 2006*.

[91] Saastamoinen K., *Many Valued Algebraic Structures as the Measures for Comparison*, Frontiers in Artificial Intelligence and Applications, Vol. 149, pp. 23-42, 2006.

[92] Santini S. and Jain R., Similarity Measures, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21(9), pp. 871-883, 1999.

[93] Schweizer B. and Sklar A., Statistical Metric Spaces, *Pacific J. Math.*, 10, pp. 313-334, 1960.

[94] Schweizer B. and Sklar A., Associative functions and statistical triangle inequalities, *Publ. Math Debrecen*, 8, pp. 169-186, 1961.

[95] Schweizer B. and Sklar A., Associative functions and abstract semigroups, *Publ. Math. Debrecen*, 10, pp. 69-81, 1963.

[96] Sklar A., Fonctions de répartition *á* n dimensions et leurs marges, *Publ. Inst. Statist. Univ. Paris*, 8, pp. 229-231, 1959.

[97] Smets P. and Magrez P., Implication in fuzzy logic, *Int. J. Approx. Reasoning*, 1(4), pp. 327-347, 1987.

[98] Sokal R.R. and Rohlf F.J., *Biometry*, W.H. Freeman, New York, 1969.

[99] Takeshita T., Nozawa S. and Kimura F, On the bias of Mahalanobis distance due to limited sample size effect, *Proceedings of the Second International Conference on Document and Recognition*, pp. 171-174, 1993.

[100] Trillas E. and Valverde L., *An Inquiry into Indistinguishability Operators*, in Aspects of Vaguenes, Skala, H.J., Termini, S., Trillas, E., eds., Reidel, Dordrecht, 231–256, 1984.

[101] Tversky A. and Krantz D.H., The Dimensional Representation and the Metric Structure of Similarity Data, *Journal of Mathematical Psychology*, pp. 572-596, 1970.

[102] Tversky A., Features of similarity, *Psychological Review*, 84(4), pp. 327-352, 1977.

[103] UCI ML Rep., various authors. UCI Repository of Machine Learning Databases network document. *http://www.ics.uci.edu/˜mlearn/MLRepository.html*, Accessed March 14, 2004.

[104] Valverde L., On the Structure of F-Indistinguishability Operators, *Fuzzy Sets and Systems*, 17, 1985.

[105] Vapnik V. and Bottou L., Local Algorithms for Pattern Recognition and Dependencies Estimation, *Neural Computation*, 5 (6), pp. 893-909, 1993.

[106] Wasserman K. and McIlroy M.B., Detecting the threshold of anaerobic metabolism in cardiac patients during exercise, *Am. J. Cardiol.*, 14, pp. 844-852, 1964.

[107] Whalen T., Parameterized R-implications, *Fuzzy Sets and Systems*, 134(2), pp. 231-281, 2003.

[108] Wolberg W. H. and Mangasarian O. L., Multisurface method of pattern separation for medical diagnosis applied to breast cytology, *Proceedings of the National Academy of Sciences of the United States of America*, , Vol. 87, No. 23, pp. 9193-9196, 1990.

[109] Wong S.K.M. and Ziarko W., Comparison of the probabilistic approximate classification and the fuzzy set model, *Fuzzy Sets and Systems*, 21, pp. 357-362, 1987.

[110] Yager R.R., On a General Class of Fuzzy Connectives, Fuzzy Sets and Systems 4, pp. 235-242, 1980.

[111] Yager R.R., On ordered weighted averaging aggregation operators in multicriteria decision making, *IEEE Transactions on Systems, Man and Cybernetics*, 18, pp. 183-190, 1988.

[112] Yager, R.R. and Rybalov, A., Uninorm aggregation operators, *Fuzzy Sets and Systems*, 80, pp. 111-120, 1996.

[113] Yager R.R., Generalized OWA aggregation operators, *Fuzzy Optimization and Decision Making*, 3, pp. 93-107, 2004.

[114] Yu Y., Triangular norms and TNF-sigma algebras, *Fuzzy Sets and Systems*, 16, pp. 251-264 1985.

[115] Zadeh L.A., Fuzzy Sets, *Information and Control*, 8, pp. 338-353, 1965.

[116] Zadeh L.A., Similarity relations and fuzzy orderings, *Inform. Sci.*, 3, pp. 177-200, 1971.

[117] Zadeh L.A., *Fuzzy Sets and Their Application to Pattern Classification And Clustering Analysis*, in J. Van Ryzin (Ed.): Classification and Clustering, Academic Press, pp. 251-299, 1977.

[118] Zimmermann H.-J. and Zysno P., Latent connectives in human decision making, *Fuzzy Sets and Systems*, 4, pp. 37-51, 1980.

[119] Zimmermann H.-J., *Fuzzy Set Theory - and its Applications*, Kluwer, Boston, 3^{rd} edition, 1996.

[120] Zhou Z.H. and Jiang Y., NeC4.5: neural ensemble based C4. 5, *IEEE Transactions on Knowledge and Data Engineering*, Vol. 16, 6, pp. 770-773, 2004.

[121] Zwick R., Carlstein E. and Budescu D.V., Measures of similarity among fuzzy concepts: A comparative analysis, *International Journal of Approximate Reasoning*, 1, pp. 221-242, 1987.

278. NEDEOGLO, NATALIA. Investigation of interaction between native and impurity defects in ZnSe. 2007. Diss.

279. KÄRKKÄINEN, ANTTI. Dynamic simulations of rotors during drop on retainer bearings. 2007. Diss.

280. KARPOVA, TATJANA. Aqueous photocatalytic oxidation of steroid estrogens. 2007. Diss.

281. SHIPILOVA, OLGA. Particle transport method for convection-diffusion-reaction problems. 2007. Diss.

282. ILONEN, JARMO. Supervised local image feature detection. 2007. Diss.

283. BOTAR-JID, CLAUDIU CRISTIAN. Selective catalytic reduction of nitrogen oxides with ammonia in forced unsteady state reactors. Case based and mathematical model simulation reasoning. 2007. Diss.

284. KINNUNEN, JANNE. Direct-on-line axial flux permanent magnet synchronous generator static and dynamic performance. 2007. Diss.

285. VALTONEN, MIKKO. Performance characteristics of an axial-flux solid-rotor-core induction motor. 2007. Diss.

286. PUNNONEN, PEKKA. Impingement jet cooling of end windings in a high-speed electric machine. 2007. Diss.

287. KÄRRI, TIMO. Timing of capacity change: Models for capital intensive industry. 2007. Diss.

288. TUPPURA, ANNI. Market entry order and competitive advantage of the firm. 2007. Diss.

289. TARKIAINEN, ANSSI. Field sales management control: Towards a multi-level theory. 2007. Diss.

290. HUANG, JUN. Analysis of industrial granular flow applications by using advanced collision models. 2007. Diss.

291. SJÖMAN, ELINA. Purification and fractionation by nanofiltration in dairy and sugar and sweetener industry applications. 2007. Diss.

292. AHO, TUOMO. Electromagnetic design of a solid steel rotor motor for demanding operation environments. 2007. Diss.

293. PURHONEN, HEIKKI. Experimental thermal hydraulic studies on the enhancement of safety of LWRs. 2007. Diss.

294. KENGPOL, ATHAKORN. An evaluation of ICTs investment using decision support systems: Case applications from distributor´s and end user´s perspective group decision. 2007. Diss.

295. LASHKUL, ALEXANDER. Quantum transport phenomena and shallow impurity states in CdSb. 2007. Diss.

296. JASTRZĘBSKI, RAFAŁ PIOTR. Design and implementation of FPGA-based LQ control of active magnetic bearings. 2007. Diss.

297. GRÖNLUND, TANJA. Development of advanced silicon radiation detectors for harsh radiation environment. 2007. Diss.

298. RUOKONEN, MIKA. Market orientation in rapidly internationalizing small companies – evidence from the software industry. 2008. Diss.

299. OIKARINEN, TUIJA. Organisatorinen oppiminen – tapaustutkimus oppimisprosessien jännitteistä teollisuusyrityksessä. 2008. Diss.

300. KARHULA, JUKKA. Cardan gear mechanism versus slider-crank mechanism in pumps and engines. 2008. Diss.

301. RAJAMÄKI, PEKKA. Fusion weld metal solidification: Continuum from weld interface to centerline. 2008. Diss.

302. KACHINA, ANNA. Gas-phase photocatalytic oxidation of volatile organic compounds. 2008. Diss.

303. VIRTANEN, PERTTU. Evolution, practice and theory of European database IP law. 2008.

304. TANNINEN, KATI. Diffusion of administrative innovation: TQM implementation and effectiveness in a global organization. 2008. Diss.

305. PUISTO, ANTTI. The initial oxidation of transition metal surfaces. 2008. Diss.

306. FELLMAN, ANNA. The effects of some variables on CO_2 laser-MAG hybrid welding. 2008. Diss.

307. KALLIOINEN, MARI. Regenerated cellulose ultrafiltration membranes in the treatment of pulp and paper mill process waters. 2008. Diss.

308. PELTOLA, SATU. Capability matrix – identifying and evaluating the key capabilities of purchasing and supply management. 2008. Diss.

309. HONKAPURO, SAMULI. Performance benchmarking and incentive regulation – considerations of directing signals for electricity distribution companies. 2008. Diss.

310. KORHONEN, KIRSI. Facilitating coordination improvement efforts in cross-functional process development programs. 2008. Diss.

311. RITVANEN, VIRPI. Purchasing and supply management capabilities in Finnish medium-sized enterprises. 2008. Diss.

312. PYNNÖNEN, MIKKO. Customer driven business model – connecting customer value to firm resources in ICT value networks. 2008. Diss.

313. AL NAZER, RAMI. Flexible multibody simulation approach in the dynamic analysis of bone strains during physical activity. 2008. Diss.

314. The Proceedings of the 7th MiNEMA Workshop. Middleware for Network Eccentric and Mobile Applications. Ed. by Pekka Jäppinen, Jouni Ikonen and Jari Porras. 2008.

315. VÄÄTÄNEN, JUHA. Russian enterprise restructuring – the effect of privatisation and market liberalisation on the performance of large enterprises. 2008. Diss.

316. DABAGHMESHIN, MAHSA. Modeling the transport phenomena within the arterial wall: porous media approach. 2008. Diss.

317. HAIMALA, JUHA. Supplier´s position in project marketing networks. 2008. Diss.

318. UOTILA, TUOMO. The use of future-oriented knowledge in regional innovation processes: research on knowledge generation, transfer and conversion. 2008. Diss.

319. LAPPALAINEN, TOMMI. Validation of plant dynamic model by online and laboratory measurements – a tool to predict online COD loads out of production of mechanical printing papers. 2008. Diss.

320. KOSONEN, ANTTI. Power line communication in motor cables of variable-speed electric drives – analysis and implementation. 2008. Diss.