UNIVERSITY OF TAMPERE
Master of Science Thesis

Anita Honka

Total Similarity Based Fuzzy Reasoning: Theory and Application

The Department of Mathematics and Statistics
Mathematics
May 2009

Tampereen yliopisto

Matematiikan ja tilastotieteen laitos

HONKA, ANITA: Kokonaissimilaarisuus sumeassa päättelyssä: teoriaa ja soveltamista

Pro gradu -tutkielma, v + 53 sivua.

Matematiikka

Toukokuu 2009

Tiivistelmä

Sumean päättelyn sovelluksissa lähtökohtana on mallintaa reaalimaailman ilmiöitä asiantuntijan ajattelun tavoin. Sumeaan logiikkaan perustuvat päättelysysteemit tarjoavatkin tähän hyvin hyödyllisiä keinoja, sillä ne pystyvät tekemään yksiselitteisiä päätöksiä epätarkan ja puutteellisen tiedon pohjalta, mikä on ominaista juuri ihmisen ajattelukyvylle. Tämän lisäksi sumeat päättelysysteemit ovat rakenteellisesti intuitiivisia, joten niiden ymmärtäminen ei vaadi erityistaitoja. Täten systeemin päättelyketjua on helppo tarkastella. Perinteisten sumeiden päättelysysteemien heikkoutena on kuitenkin, että ne eivät ole matemaattisesti hyvin määriteltyjä.

Tässä pro gradu -tutkielmassa tavoitteena on esitellä perinteisille päättelysysteemeille vaihtoehtoinen, matemaattisesti hyvin määritelty, sumean kokonaissimilaarisuuden käsitteeseen perustuva päättelysysteemi ja sen taustalla olevaa matemaattista teoriaa. Teoriaosuudessa tutustutaan kokonaissimilaarisuuden käsitteen lisäksi algebrallisiin MV- ja Lukasiewicz struktuureihin. Tutkielmassa esitellään myös algoritmi tämän päättelysysteemin rakentamiseksi ja tutkitaan siinä tapahtuvaa päättelyprosessia. Soveltavassa osuudessa päättelyprosessia sovelletaan systeemiin, jonka tarkoituksena on mallintaa liikunnan vaikutusta verenpaineeseen. Tämän mallin tuloksia verrataan vielä vastaavan mallin tuloksiin, johon on sovellettu perinteisempää Sugeno-tyyppistä päättelyprosessia.

Johtopäätelmät: Sumeaan kokonaissimilaarisuuteen perustuva päättelysysteemi on matemaattisesti hyvin perusteltu tapauksissa, joissa ei vaadita sumean disjunktion käsittelyä. Kokonaissimilaarisuus näyttää olevan ainoa sopiva keino sumean konjunktion käsittelyyn. Tutkielmassa esitellyt vertailut kokonaissimilaarisuuteen perustuvan ja Sugeno-tyyppisen mallin tulosten välillä osoittavat, että mallit toimivat jokseenkin samalla tavoin. Tätä ei voitu kuitenkaan arvioida luotettavasti käytettävissä olevilla menetelmillä.

University of Tampere

Department of Mathematics and Statistics

HONKA, ANITA: Total Similarity Based Fuzzy Reasoning: Theory and Application

Master of Science Thesis, v + 53 pages.

Mathematics

May 2009

Abstract

Fuzzy reasoning is an efficient way to model human knowledge. It aims to mimic the decision-making of humans. The main advantage of fuzzy inference systems is that they have the ability to provide a precise conclusion based upon vague, ambiguous or imprecise information. In addition, the structure of fuzzy systems is intuitive in the sense that it does not require any special expertise to be understood. Thus, the inference process of the system is easy to examine. However, the traditional fuzzy inference systems do not have well-defined mathematical foundations.

The main intention of the work described in this Master of Science Thesis is to study an alternative method to perform fuzzy inference that is based on the well-defined concept of total fuzzy similarity relation. First, the theory behind this relation is presented by introducing the important concepts of MV-algebras and Lukasiewicz algebras. Second, the total fuzzy similarity relation is defined. Third, the algorithm for constructing the inference system and the inference process itself are presented and discussed. In the final part of this thesis, the total fuzzy similarity based inference is applied to a system that models the effect of exercise on blood pressure. Basic level comparisons are performed between the outcomes of this model and the outcomes of a corresponding Sugeno type model.

Conclusions: The total fuzzy similarity based inference is mathematically well-defined for systems that do not include inference rules involving fuzzy disjunction. In addition, it seems that total fuzzy similarity is the only appropriate choice for representing fuzzy conjunction in the rules. Finally, the constructed total fuzzy similarity based system seems to perform quite similarly to the corresponding Sugeno type system, though this could not be reliably validated.

Preface

I became acquainted with fuzzy inference systems at VTT, the Technical Research Center of Finland, during the summer 2008. I found the ability of fuzzy inference to model complex real life phenomena in a simplified way resembling human reasoning as intriguing, which gave me the incentive to write this Master of Science Thesis on the subject. However, it was only after my thesis advisor, Docent Esko Turunen, PhD, of the Department of Mathematics at the Tampere University of Technology (TUT), who introduced me to the concept of total fuzzy similarity, that the subject of this thesis got its final shape. For this and all the guidance he provided me, I am deeply grateful.

The fuzzy inference system presented in the application part of this thesis was created at VTT during the fall 2008 as a part of an EU-funded research project, HeartCycle. I am very appreciative of the valuable and constructive feedback Docent Mark van Gils, PhD, and Senior Research Scientist Juha Pärkkä, MSc (Tech.), VTT, gave on my outputs related to the development of the model and the thesis itself. Thank you also for all the encouraging support you provided during this process.

Furthermore, besides Esko Turunen, my sincere thanks go to Professor Lauri Hella, PhD, of the Department of Mathematics and Statistics at the University of Tampere (UTA), for the examination of this thesis.

Tampere, May 2009

Anita Honka

Contents

Introduction

The world surrounding us is full of complex phenomena that are characterized by vagueness and uncertainty. Despite that, human beings are capable to manage this kind of imprecise information and make decisions based on it. Nowadays, the urge to control automatically the real world's phenomena is ever more increasing. However, the complexity of these phenomena is hard or impossible to manage with exact mathematical models, since they do not have the ability to process incomplete knowledge. In 1965, Lofti Zadeh defined the concepts of fuzzy sets and fuzzy logic, which unlike the two-valued logic accepts the concept of partial truth, that is, truth values between the classical values *true* and *false*. This provided the framework for an inference structure that enables the human reasoning capabilities to be applied to artificial modelling systems. These systems are generally referred to as fuzzy inference systems. They are especially developed to deal with uncertainty and imprecision and thus, have the ability to model complex problems with sufficient efficiency, unlike the exact mathematical models [23, 25]. Furthermore, these systems have shown to be successful in modelling several automatic control tasks [19].

Traditionally, fuzzy systems refer either to Mamdani or Sugeno type systems, which differ in the method the final output of the inference is determined. In Mamdani, this process is called defuzzification which involves methods such as Center of Gravity or the least of maximum calculations. Sugeno type models do not require a separate defuzzification method, since the final output is determined already during the aggregation process as the weighted average of the crisp outputs associated with the fired rules [17]. However, these processes of aggregation and defuzzification are artificial in the sense that they are based neither on expert knowledge nor on well-defined mathematical concepts [19].

In this work we present an alternative method to perform fuzzy inference that is based on the well-defined concept of total fuzzy similarity relation first introduced by Turunen et al. in 2002 [16]. The inference of this total similarity based system relies only on expert's knowledge and on well-defined logical concepts. This system has been shown to obtain relatively similar results to Sugeno systems in different control applications such as traffic signal control [19] and real-time water reservoir operation [6].

To begin with we introduce the theory behind total fuzzy similarity relation in Chapter 1 by defining the algebraic structure the relation is valid for. This particular structure is a certain type of complete residuated lattice called an injective Wajsberg or MV-algebra. Fuzzy subsets obey this structure, which allows us to apply the total fuzzy similarity relation on them later on in Chapters 3 and 4 in the context of actual inference systems. Through the way we get familiar with Lukasiewicz algebra, which is a canonical instance of an injective MV-algebra. The Lukasiewicz algebra provides a concrete insight to the concepts of fuzzy subsets and total fuzzy similarity, since it is a structure defined on the real unit interval. We will concentrate only on the essential properties of these structures and present the main results, which are important from the total similarity point of view. Further reading can be found, for instance, from the book of Turunen [25], Mathematics Behind Fuzzy Logic, on which Chapter 1 is mostly based. Another mathematically important reference source is the article [16] of Kukkurainen and Turunen.

In Chapter 2 we present the definitions for the concepts related to fuzzy similarities and to the actual total fuzzy similarity relation. In Chapter 3 we introduce the algorithm for constructing a total fuzzy similarity based inference system and discuss the inference process of the system. In Chapter 4 we apply the introduced inference process to a problem of modelling the effect of exercise on blood pressure. In addition, we perform some very basic level comparisons between the outcomes of this model and the outcomes of a corresponding Sugeno type model. Furthermore, we present the main results of a literature review that was performed to gather the expert knowledge required for the model construction, since after all, expert knowledge is a vital part of fuzzy inference systems. The knowledge was gathered from several clinical exercise trials and meta-analyses [1, 2, 3, 4, 5, 10, 11, 12, 13, 14, 15, 18, 21, 22, 24, 26, 27]. The literature review and large part of the models were developed at VTT by the author in the of context an EU-funded research project, HeartCycle.

We assume the reader is somewhat familiar with the traditional Mamdani and Sugeno fuzzy systems and has the basic knowledge on the areas of set theory and algebra.

Chapter 1

Mathematical Structure of Fuzzy Sets

We begin our study by introducing the concept of fuzzy sets in Section 1.1 since, after all, fuzzy sets are the key part of any fuzzy inference system. Then, we concentrate on the concepts that are essential from the total fuzzy similarity point of view starting from residuated lattices in Section 1.2 and continuing to Wajsberg or MV-algebras in Section 1.3, which are certain kinds of complete residuated lattices. Particularly, we are interested in the structure of an injective Wajsberg algebra, since this structure is required for the total fuzzy similarity relation to be valid. Moreover, a structure called Lukasiewicz algebra is introduced, since it is a canonical example of an injective Wajsberg algebra. Thus, it can provide us with concrete understanding of the presented concepts. We conclude our study of the theory behind total fuzzy similarity with the result stating that a collection of fuzzy subsets have the structure of an injective Wajsberg (or MV-algebra). This enables the total fuzzy similarity relation to be applied on fuzzy sets, which is required by the total fuzzy similarity based inference system.

1.1 Fuzzy Subsets

In the classical set theory a subset A of a non-void set X is referred to as a *crisp set*, which is defined by the *characteristic function* $\mathcal{X}_A : X \mapsto \{0, 1\}$ such that, for all $x \in X$,

$$\mathcal{X}_A(x) = \begin{cases} 1, & x \in A \\ 0, & x \notin A. \end{cases}$$

The set $\{0,1\}$ is the truth value set in classical logic. However, in many-valued logic, which can be seen as a field of fuzzy logic, the truth value set can be more general resulting in the concept of partial truth. Thus, the fuzzy set theory can be considered as a generalization of the classical set theory.

Definition 1.1. Let X and L be crisp non-void sets. An L-valued *fuzzy subset* A of the set X is an ordered couple (X, μ_A), where $\mu_A : X \mapsto L$. The mapping μ_A is called a *membership function*. Cf. [25, p. 118].

The membership function μ_A defines the degree to which an element $x \in X$ belongs to the fuzzy subset A. In practice, the set of membership degrees or truth values L associated with a fuzzy subset refers to the real unit interval $[0,1]$. However, fuzzy set theory is not restricted to these truth values only. On the contrary, the truth value set L has a more general structure, namely the structure of a complete lattice, which we define in the next section. As we proceed, we shall study the structure of L in a general manner and consider the real unit interval $[0,1]$ as an important special case of truth value sets.

Example 1.1. Amongst human beings there are no exact definitions for linguistic terms such as *tall*, *thin*, *fast* speed, *loud* noise or *hot* weather. Some people may feel that the temperature of 26°C is hot, whereas many people might consider it warm only, and the temperature of 30°C as hot. These more or less subjective definitions for hot weather might depend on the climate considered as typical in the home countries of different people. This imprecise definition for hot weather is illustrated below by an example fuzzy subset *hot* associated with the membership function $\mu_{hot}(x)$, where x refers to a certain temperature.

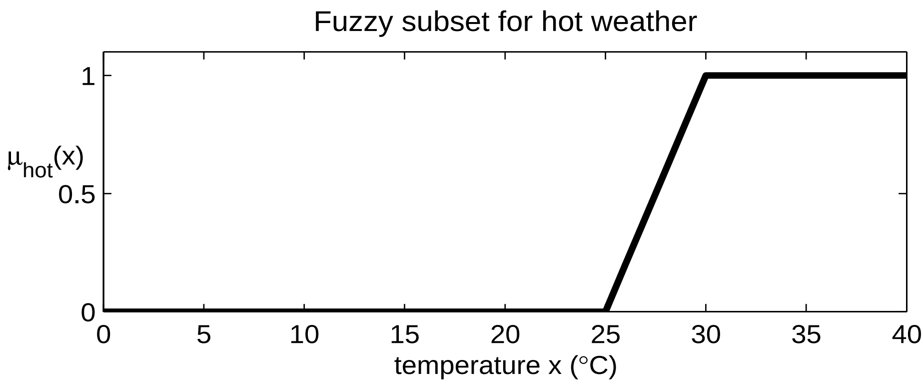

Figure 1.1: Illustration for the imprecise linguistic term *hot weather* by the fuzzy subset *hot*.

From Figure 1.1 we can see that $\mu_{hot}(20) = 0$, $\mu_{hot}(27.5) = 0.5$ and $\mu_{hot}(35) = 1$. This information can be expressed in natural language in the following way: "20°C is certainly not hot", "27.5°C is more or less hot" and "35°C is certainly hot".

1.2 Residuated Lattices

Lattices can be studied from two different aspects: either from the order theory point of view or as an algebraic structure [7]. In definition 1.2 we introduce the concept of a lattice from the aspect of order theory and by theorem 1.1 we conclude that the defined lattice can be presented as a certain algebraic structure. Furthermore, we introduce the important concept of a residuated lattice and get familiar with a Lukasiewicz structure.

Definition 1.2. A *lattice* is a partially ordered set $(L, \leq)$ such that for any $x, y \in L$, $\sup\{x, y\}$ and $\inf\{x, y\}$ exist in L [7, p. 10].

Theorem 1.1. *Let $\vee$ and $\wedge$ be such binary operations on a non-void set L that for any $x, y \in L$, the relation $\leq$ is defined as*

$$x \leq y \Leftrightarrow x \vee y = y \Leftrightarrow x \wedge y = x \tag{1.1}$$

and the following holds for any $x, y, z \in L$

$x \vee x = x,$	$x \wedge x = x$	*idempotency,*	(1.2)
$x \vee y = y \vee x,$	$x \wedge y = y \wedge x$	*commutativity,*	(1.3)
$x \vee (y \vee z) = (x \vee y) \vee z,$	$x \wedge (y \wedge z) = (x \wedge y) \wedge z$	*associativity,*	(1.4)
$x \vee (x \wedge y) = x,$	$x \wedge (x \vee y) = x$	*absorption.*	(1.5)

Then, the algebraic structure $(L, \vee, \wedge)$ is a lattice. Cf. [7, p. 14].

Proof (cf. [7, p. 15]). The operations $\vee$ and $\wedge$ are reflexive according to condition 1.2. The antisymmetry of the operations follows from conditions 1.1 and 1.3 and transitivity from condition 1.4. Thus, the operations $\vee$ and $\wedge$ define a partially ordered set $(L, \leq)$.

We need to show that for any $x, y \in L$, $\sup\{x, y\}$ and $\inf\{x, y\}$ exist in L. Since $\vee$ and $\wedge$ are binary operations on the set L, for any $x, y \in L$, $x \vee y$ and $x \wedge y$ exist in L. We prove that $\sup\{x, y\} = x \vee y$ and $\inf\{x, y\} = x \wedge y$. The following holds based on the properties of associativity, commutativity and idempotency

$$(x \vee y) \vee x = x \vee (y \vee x) = x \vee (x \vee y) = (x \vee x) \vee y = x \vee y.$$

Therefore, by condition 1.1, $x \leq x \vee y$. Similarly, we conclude that $y \leq x \vee y$. Assume $x, y \leq z$. Then $(x \vee y) \vee z = x \vee (y \vee z) = x \vee z = z$. Therefore, $x \vee y \leq z$. Thus, $\sup\{x, y\} = x \vee y$. Furthermore, $x \wedge y \leq x, y$, since by conditions 1.5 and 1.3 $x = x \vee (x \wedge y)$ and $y = y \vee (x \wedge y)$. Assume $z \leq x, y$. Then $x \wedge z = (z \vee x) \wedge z = z$ (by 1.5 and 1.3) and similarly $y \vee z = z$. Now

$$(x \wedge y) \vee z = (x \wedge y) \vee (x \wedge z) = (x \wedge y) \vee (x \wedge (y \wedge z)) = (x \wedge y) \vee ((x \wedge y) \wedge z) = x \wedge y,$$

where the last equility follows from condition 1.5. Therefore, by condition 1.1 $z \leq x \wedge y$. Thus, $\inf\{x, y\} = x \wedge y$. □

Hence, the expressions $(L, \leq)$ and $(L, \vee, \wedge)$ can be both used to describe a lattice. For our purpose it is convenient to study lattices as algebraic structures, so from here onwards the latter expression is used to represent a lattice.

Remark 1.1. The proof of theorem 1.1 shows that in a lattice $(L, \vee, \wedge)$, $\sup\{x, y\}$ and $\sup\{x, y\}$ can be denoted by $x \vee y$ and $x \wedge y$, respectively.

Definition 1.3. A lattice $(L, \vee, \wedge)$ is a *complete lattice*, if for any subset $S \subseteq L$, $\sup S$ and $\inf S$ exist in L [25, p. 5].

Remark 1.2. A complete lattice contains always the least element $\mathbf{0}$ and the greatest element $\mathbf{1}$. This can be seen by setting $S = L$. [25, p. 5]

Example 1.2. The unit interval I is a complete lattice under the usual order, for which $x \vee y = \max\{x, y\}$, $x \wedge y = \min\{x, y\}$, $\mathbf{0} = 0$ and $\mathbf{1} = 1$ (cf. [25, p. 5]). Thus, the truth values set $L = [0, 1]$ often associated with fuzzy subsets, is a complete lattice.

Definition 1.4. Assume a binary operation g defined on a non-void set A. Let the elements $x, y, z \in A$ be arbitrary and $x \leq y$. The operation g is *isotone in the first variable*, if $g(x, z) \leq g(y, z)$ and *isotone in the second variable*, if $g(z, x) \leq g(z, y)$. The operation g is *antitone in the first variable*, if $g(y, z) \leq g(x, z)$ and *antitone in the second variable*, if $g(z, y) \leq g(z, x)$. If g is isotone in both variables, then g is called simply *isotone*. Cf. [25, p. 9].

Definition 1.5. Let $(L, \vee, \wedge)$ be a lattice, where binary operations $\odot$ and $\rightarrow$ are defined such that $\odot$ is associative, commutative and isotone and the condition

$$\forall x, y, z \in L : x \odot y \leq z \Leftrightarrow x \leq y \rightarrow z \tag{1.6}$$

holds. Furthermore, let the elements $\mathbf{0}$ and $\mathbf{1}$ exist in L and

$$\forall x \in L : x \odot \mathbf{1} = x. \tag{1.7}$$

Then, L is a *residuated lattice* which is denoted by $(L, \vee, \wedge, \odot, \rightarrow, \mathbf{0}, \mathbf{1})$. Cf. [25, pp. 9 and 11].

The condition 1.6 is called *residuation* and the operations $\odot$ and $\rightarrow$ are called *product* and *residuum*, respectively [25]. In the real unit interval $[0,1]$ the operation $\odot$ is often referred to as *continuous* t-*norm* with the operation $\rightarrow$ being its *residua* [19].

Theorem 1.2. *In a residuated lattice* $(L, \vee, \wedge, \odot, \rightarrow, \mathbf{0}, \mathbf{1})$*, the residuum* $\rightarrow$ *is antitone in the first variable and isotone in the second variable.*

Proof (cf. [25, p. 10]). Assume $x, y, z \in L$ and $x \leq y$. Since $\odot$ is isotone, we have $(y \rightarrow z) \odot x \leq (y \rightarrow z) \odot y$, where by condition 1.6, $(y \rightarrow z) \odot y \leq z$. Thus, $(y \rightarrow z) \odot x \leq z$, and by condition 1.6, $(y \rightarrow z) \leq (x \rightarrow z)$. Hence, $\rightarrow$ is antitone in the first variable. Furthermore, by condition 1.6, $(z \rightarrow x) \odot z \leq x$, and by the assumption, $(z \rightarrow x) \odot z \leq y$. Thus, by 1.6, $(z \rightarrow x) \leq (z \rightarrow y)$. Hence, $\rightarrow$ is isotone in the second variable. $\square$

Definition 1.6. On the unit interval I, let the binary operations $\odot$ and $\rightarrow$ be defined for any $x, y \in I$ as

$$x \odot y = \max\{0, x + y - 1\}, \tag{1.8}$$

$$x \rightarrow y = \min\{1, 1 - x + y\}. \tag{1.9}$$

Then, we have a *Lukasiewicz structure*. Cf. [25, p. 11].

Theorem 1.3. *Lukasiewicz structure is a residuated lattice* $(I, \min, \max, \odot, \rightarrow, 0, 1)$.

Proof (cf. [25, p. 11]). In the unit interval I it is easy to see that for any $x, y \in I$, $x \vee y = \max\{x, y\}$ and $x \wedge y = \min\{x, y\}$, $\mathbf{0} = 0$, $\mathbf{1} = 1$ and that the operation $\odot$, defined by 1.8 is isotone and commutative. The operation $\odot$ is associative, as $(x \odot y) \odot z = 0 = x \odot (y \odot z)$ if $x + y + z \leq 2$, since then $x + y, y + z \leq 1$. Otherwise, if $x + y + z > 2$, $(x \odot y) \odot z = x + y + z - 2 = x \odot (y \odot z)$. Also, $x \odot 1 = x$ holds for any $x \in I$, since $\max\{0, x + 1 - 1\} = \max\{0, x\} = x$. Furthermore, the condition 1.6 holds, since for any $x, y, z \in I$, $x \odot y = \max\{0, x + y - 1\} \leq z$ iff $x + y - 1 \leq z$ iff $x \leq 1 - y + z$ iff $x \leq \min\{1, 1 - y + z\}$ iff $x \leq y \rightarrow z$. $\square$

Example 1.3. The following introduces couple of other residuated lattices defined on the unit interval I [19].
Gödel structure:

$$x \odot y = \min\{x, y\}, \quad x \rightarrow y = \begin{cases} 1 & \text{if} \quad x \leq y, \\ y & \text{otherwise.} \end{cases}$$

Gaines structure:

$$x \odot y = xy, \quad x \to y = \begin{cases} 1 & \text{if} \quad x \leq y, \\ \frac{y}{x} & \text{otherwise.} \end{cases}$$

In a residuated lattice L, there are two important derived operations

$$x^* = x \to \mathbf{0}, \tag{1.10}$$

$$x \leftrightarrow y = (x \to y) \wedge (y \to x). \tag{1.11}$$

The abbreviation x^* is called the *complement* of an element $x \in L$ and $x \leftrightarrow y$ is called *bi-residuum*. [25]

The operations $\odot$ and $\to$ are the algebraic counterparts of the connectives *conjunction* and *implication* in classical logic, respectively. The complement is the counterpart of *negation* and the bi-residuum is the counterpart of *equivalence* [25]. An operation, which can be regarded as *disjunction*, is introduced in Section 1.3 in the context of Wajsberg and MV-algebras.

In Lukasiewicz algebra the bi-residuum takes the form of (cf. [16])

$$x \leftrightarrow y = 1 - |x - y|. \tag{1.12}$$

Indeed,

$$\begin{aligned} x \leftrightarrow y &\overset{1.11}{=} (x \to y) \wedge (y \to x) \\ &\overset{1.9}{=} \min\{\min\{1, 1 - x + y\}, \min\{1, 1 - y + x\}\} \\ &= \begin{cases} \min\{1, 1 - y + x\}, & \text{if } x \leq y \\ \min\{1 - x + y, 1\}, & \text{otherwise} \end{cases} \\ &= \begin{cases} 1 - y + x, & \text{if } x \leq y \\ 1 - x + y, & \text{otherwise} \end{cases} \\ &= 1 - \max\{x, y\} + \min\{x, y\} \\ &= 1 - |x - y|. \end{aligned}$$

Moreover, it is trivial to see that by the condition 1.10 the complement in Lukasiewicz algebra takes the form of

$$x^* = 1 - x. \tag{1.13}$$

Theorem 1.4. *The bi-residuum has the following properties*

$$x \leftrightarrow x = \mathbf{1} \qquad \textit{reflexivity,} \tag{1.14}$$

$$x \leftrightarrow y = y \leftrightarrow x \qquad \textit{symmetry,} \tag{1.15}$$

$$(x \leftrightarrow y) \odot (y \leftrightarrow z) \leq (x \leftrightarrow z) \qquad \textit{weak transitivity,} \tag{1.16}$$

$$x \leftrightarrow \mathbf{1} = x. \tag{1.17}$$

Proof. See [25, pp. 14 and 156]. □

Bi-residuum can be seen as the equivalence relation of many-valued logic, since it has similar properties to the classical equivalence relation. Furthermore, in Section 2.1 we see that bi-residuum is actually a fuzzy similarity relation.

1.3 Wajsberg Algebras and MV-Algebras

In this section we introduce the set of axioms that are necessary and sufficient for constructing total fuzzy similarity based inference systems. We begin by introducing Wajsberg algebras and see few of the many important consequences of this structure. Particularly, a Wajsberg algebra turns out to be a residuated lattice, which generates a MV-algebra. In fact, there is a one-to-one correspondence between Wajsberg and MV-algebras. Furthermore, we introduce the important concept of an injective Wajsberg algebra and note that a collection of fuzzy sets together with Lukasiewicz algebra are isomorphic to this structure. Thus, by the end of this section we have the structural basis ready for the total fuzzy similarity relation. The proofs for the presented results are out of the scope of this work, since they would require an in-depth analysis of MV-algebras and understanding of the several complicated concepts behind these structures. The proofs can be found e.g. in [25].

Definition 1.7. Let L be a non-void set containing the greatest element $\mathbf{1}$, a binary operation $\rightarrow$ and a unary operation * such that for any $x, y, z \in L$,

$$\mathbf{1} \rightarrow x = x, \tag{1.18}$$

$$(x \rightarrow y) \rightarrow [(y \rightarrow z) \rightarrow (x \rightarrow z)] = \mathbf{1}, \tag{1.19}$$

$$(x \rightarrow y) \rightarrow y = (y \rightarrow x) \rightarrow x, \tag{1.20}$$

$$(x^* \rightarrow y^*) \rightarrow (y \rightarrow x) = \mathbf{1}. \tag{1.21}$$

Then, the structure $(L, \rightarrow, ^*, \mathbf{1})$ is a *Wajsberg algebra.* [25, p. 36]

It can be shown (see [25, pp. 36-39]) that a Wajsberg algebra $(L, \rightarrow, ^*, \mathbf{1})$ generates a residuated lattice $(L, \vee, \wedge, \odot, \rightarrow, \mathbf{0}, \mathbf{1})$, where $\mathbf{0} = \mathbf{1}^*$, and the lattice operations $\leq, \vee, \wedge, \odot$ are defined by

$$\begin{aligned} x \leq y &\Leftrightarrow x \rightarrow y = \mathbf{1}, \\ x \vee y &= (x \rightarrow y) \rightarrow y, \\ x \wedge y &= (x^* \vee y^*)^*, \\ x \odot y &= (x \rightarrow y^*)^*. \end{aligned}$$

Furthermore, on a Wajsberg algebra L we introduce a binary operation $\oplus$, which is defined by [25]

$$\forall x, y \in L : x \oplus y = x^* \rightarrow y.$$

These above definitions for the different operations result in a structure $(L, \oplus, \odot, ^*, \mathbf{0}, \mathbf{1})$ called *MV-algebra*. There is a one-to-one correspondence between Wajsberg and MV-algebras. Therefore, MV-algebras have all the same properties Wajsberg algebras have, and thus, MV-algebras are residuated lattices. [25]

The operation $\oplus$ is important, since it has a key role in the concept of divisible algebra defined below, which in turn is one of the core axioms required for total fuzzy similarity. Moreover, the operation $\oplus$ can be seen as the algebraic counterpart of *disjunction* in classical logic [25]. It is easy to see that in the Lukasiewicz structure the operation $\oplus$ takes the form of

$$x \oplus y = \min\{1, x + y\}. \tag{1.22}$$

Definition 1.8. On a Wajsberg algebra (or MV-algebra) L, let $a, b \in L$ be such elements that the following equalities hold

$$\begin{aligned} (a^* \oplus (n-1)b)^* &= b, \\ nb &= a, \quad \text{where } nb = \underbrace{b \oplus \cdots \oplus b}_{n \text{ times}}, \quad n \in \mathbb{Z}_+. \end{aligned}$$

Then, the element b is called an *n-divisor* of the element a. Cf. [16].

Definition 1.9. Let L be a Wajsberg algebra. If all $a \in L$ have an n-divisor for each $n \in \mathbb{Z}_+$, then L is called *divisible*. [16]

Definition 1.10. A Wajsberg algebra L is *injective*, if [16]

$$L \text{ is complete,} \tag{1.23}$$
$$L \text{ is divisible.} \tag{1.24}$$

The axioms 1.18 - 1.21, 1.23 and 1.24 of an injective Wajsberg algebra are necessary and sufficient for performing total fuzzy similarity based inference on fuzzy sets [16]. In addition, as Boolean algebras are the natural representations of classical set theory, MV-algebras (or Wajsberg algebras) are the natural representations of fuzzy set theory. Thus, MV-algebras are fundamental for fuzzy logic. Also, the concept of injectivity has an essential role in the completeness theorem of fuzzy logic [25] which, however, is out of the scope of our study.

Definition 1.11. A non-void subset S of an algebra L is a *subalgebra* of L, if S is closed with respect to the algebraic operations defined on L [25, p. 53].

Definition 1.12. An algebra $(L, \vee, \wedge)$ is called *linearly ordered* if and only if, for all $x, y \in L$, either $x \leq y$ or $y \leq x$ [20].

Definition 1.13. Let $L_i, i \in \mathcal{I}$ be subalgebras of an algebra L. A *direct product* of the subalgebras is defined as the set

$$\prod_{i \in \mathcal{I}} L_i := \{(x_i)_{i \in \mathcal{I}} | x_i \in L_i, \forall i \in \mathcal{I}\},$$

which is closed with respect to the algebraic operations defined on L.

Example 1.4. The direct product of two sets $X, Y \in \mathbb{R}$ is the cartesian product

$$X \times Y = \{(x, y) | x \in X, y \in Y\}.$$

We are now ready to present the following important theorem.

Theorem 1.5. *Every MV-algebra is isomorphic to a subalgebra of a direct product of linearly ordered MV-algebras.*

Proof. See [20]. □

The theorem 1.5 is especially important, since it is a generalization of the famous *Stone representation theorem* of Boolean algebras, which states that [25]

> Every Boolean algebra is isomorphic to a subalgebra of a direct product of two-element Boolean algebras.

The following theorems 1.6 and 1.7 justify the choice of Lukasiewicz structure as the basis of total fuzzy similarity relation.

Theorem 1.6. *An injective Wajsberg algebra L is structurally isomorphic to a collection $\mathcal{F}$ of $[0, 1]$-valued functions, in other words, to fuzzy sets.*

Proof. See [25]. □

Theorem 1.7. *Lukasiewicz algebra is an injective Wajsberg algebra.*

Proof. See [25, p. 68]. □

Furthermore, it can be shown that the Lukasiewicz algebra is isomorphic to a collection $\mathcal{F}$ of fuzzy sets [25]. Thus, Lukasiewicz structure is an appropriate choice for representing fuzzy sets. Moreover, from theorem 1.7 it follows that the total fuzzy similarity relation is valid for this structure.

Chapter 2

Fuzzy Similarity

In this chapter we study our main interest, the concept of total fuzzy similarity. First, in Section 2.1, we introduce the general definition for a fuzzy relation and discuss a special case of this relation, namely fuzzy similarity, which is fundamental in many-valued reasoning. Then, in Section 2.2, we present the concept of total fuzzy similarity as a certain composition of several fuzzy similarity relations.

2.1 Fuzzy Similarity Relations

Definition 2.1. Let X, Y be two non-void sets. A binary *fuzzy relation* R is a function $R : X \times Y \mapsto L$, where the set L is a complete residuated lattice. Cf. [25, p. 113].

In other words, the fuzzy relation R is a L-valued fuzzy set on $X \times Y$, where the function R corresponds to the membership function $\mu_{X\times Y}$.

Definition 2.2. Let X, Y, W be non-void sets and let L be a complete residuated lattice. In addition, let $R : X \times Y \mapsto L$ and $S : Y \times W \mapsto L$ be binary fuzzy relations. The *composition* of R and S defined by a residuated lattice operation γ is a fuzzy relation $R\gamma S : X \times W \mapsto L$, where (cf. [25, p. 114])

$$\forall x \in X, w \in W : R\gamma S(x, w) = \bigvee_{y\in Y} \gamma(R(x, y), S(y, w)).$$

Definition 2.3. Let X be a non-void set. A fuzzy *similarity* relation S on X is such a binary fuzzy relation that for each $x, y, z \in X$ (cf. [25, p. 116]),

$$S(x, x) = \mathbf{1}, \tag{2.1}$$
$$S(x, y) = S(y, x), \tag{2.2}$$
$$S(x, y) \odot S(y, z) \le S(x, z). \tag{2.3}$$

Remark 2.1. The fuzzy similarity relation is a many-valued equivalence relation, since by the conditions 2.1, 2.2 and 2.3, it is reflexive, symmetric and weakly transitive, respectively.

Theorem 2.1. *Any fuzzy subset $A = (X, \mu_A)$ of a set X generates a fuzzy similarity S_A on X, defined by*

$$S_A(x, y) = \mu_A(x) \leftrightarrow \mu_A(y), \quad \textit{where } x, y \in X.$$

Proof (cf. [19]). By theorem 1.4 the bi-residuum $\leftrightarrow$ satisfies the properties of similarity. □

Remark 2.2. The fuzzy subset A generates also a so-called fuzzy *non-similarity*, which we denote by n-S_A and define as

$$\text{n-}S_A(x, y) = [\mu_A(x) \leftrightarrow \mu_A(y)]^*. \tag{2.4}$$

It is trivial to see that in a Lukasiewicz algebra L, by conditions 1.12 and 1.13, the following holds

$$\forall x, y \in L : (x \leftrightarrow y)^* = |x - y|.$$

Thus, for this structure the representation of the negation of equivalence is distance [16] and for Lukasiewicz-valued fuzzy subsets the distance corresponds to the concept of fuzzy non-similarity.

When conducting fuzzy inference we are interested to know the extent to which an input value $x \in X$ satisfies a certain property A. This is described as a fuzzy subset A of the set X. Let us assume that an element $y \in X$ has the greatest membership degree of $\mathbf{1}$ in the fuzzy subset A and that it is regarded as the ideal element of the subset. Usually, we are interested to know the degree to which the input element x is similar to the ideal element y with respect to the fuzzy subset A. Hence, one should notice that (cf. [16])

$$\text{if } \mu_A(y) = \mathbf{1}, \text{ then } S_A(x, y) = \mu_A(x) \leftrightarrow \mathbf{1} = \mu_A(x).$$

This follows from the property 1.17 of bi-residuum.

2.2 Total Fuzzy Similarity

The notion of total fuzzy similarity gives us a simple way to determine the similarity between any two objects $a, b \in X \times Y \times Z$, which can be characterized by several fuzzy subsets, in other words by different properties. Let us denote here the fuzzy subsets e.g. as A, B, C, which are subsets of the non-void sets X, Y, Z, respectively. Thus, the similarity between two objects a, b (total similarity) is a composition of the partial similarities $S_A(a,b), S_B(a,b), S_C(a,b)$. We shall concretize the concept of objects more in the next chapter.

Since the Lukasiewicz structure L is an injective Wajsberg algebra (by theorem 1.7), it is divisible and thus, by definition 1.9, all the elements $a \in L$ have an n-divisor for each $n \in \mathbb{Z}_+$. In [16] it was proved that the n-divisors of an injective Wajsberg algebra are unique. Hence, the unique n-divisor of an element a may be denoted by a/n.

Moreover, in Lukasiewicz algebra L for elements $a_1, \ldots, a_n \in L$, $n \in \mathbb{Z}_+$, we have

$$\frac{a_1}{n} \oplus \cdots \oplus \frac{a_n}{n} \overset{1.22}{=} \min\{1, \frac{a_1}{n} + \cdots + \frac{a_n}{n}\} = \sum_{i=1}^{n} \frac{a_i}{n} = \frac{1}{n} \sum_{i=1}^{n} a_i,$$

since $a_i \leq 1$ and thus $a_1/n + \cdots + a_n/n \leq 1$ (cf. [16]). We are now ready for the following theorem.

Theorem 2.2. *Let S_i be a Lukasiewicz-valued fuzzy similarity relation on a non-void set $X_i, i = 1, \ldots, n$. Then,*

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

is a Lukasiewicz-valued fuzzy similarity on $X = X_1 \times \cdots \times X_n$. Cf. [25, p. 136].

Proof. The proof follows directly from the assumption that each S_i, $i = 1, \ldots, n$, is a similarity relation. Thus, the relation S is reflexive and symmetric. In addition, since each S_i is transitive, $S(x,y) \odot S(y,z) \leq S(x,z)$ holds for all $x, y, z \in X$. See details from [25, p. 136]. □

Proposition 2.1. *Let S_i be a Lukasiewicz-valued fuzzy similarity relation on a non-void set $X_i, i = 1, \ldots, n$. Then, the weighted mean*

$$S(x,y) = \frac{1}{M} \sum_{i=1}^{n} m_i S_i(x,y),$$

where $M = \sum_{i=1}^{n} m_i$, $m_i \in \mathbb{N}$, is a Lukasiewicz-valued fuzzy similarity on $X = X_1 \times \cdots \times X_n$. Cf. [19].

Proof. The proof is an easy generalization of the proof of theorem 2.2.

The reflexivity and symmetry of S follows from the reflexivity and symmetry of the relations $S_i, i = 1, \ldots, n$. The weak transitivy of S can be shown in the following way. Let A $= S(x,y) \odot S(y,z)$. If $S(x,y)+S(y,z)-1 \leq 0$, then $A = 0$ by condition 1.8. Thus, certainly $S(x,z) \geq A$. Otherwise,

$$
\begin{aligned}
A &\overset{1.8}{=} S(x,y)+S(y,z)-1 \\
&= \frac{1}{M}\sum_{i=1}^{n} m_i S_i(x,y) + \frac{1}{M}\sum_{i=1}^{n} m_i S_i(y,z) - 1 \\
&= \frac{1}{M}\Big(\sum_{i=1}^{n} m_i S_i(x,y) + \sum_{i=1}^{n} m_i S_i(y,z) - M\Big) \\
&= \frac{1}{M}\Big[m_1\Big(S_1(x,y)+S_1(y,z)-1\Big) + \cdots + m_n\Big(S_n(x,y)+S_n(y,z)-1\Big)\Big] \\
&\overset{1.8}{=} \frac{1}{M}\Big[m_1\Big(S_1(x,y)\odot S_1(y,z)\Big) + \cdots + m_n\Big(S_n(x,y)\odot S_n(y,z)\Big)\Big] \\
&\overset{2.3}{\leq} \frac{1}{M}\Big(m_1 S_1(x,z) + \cdots + m_n S_n(x,z)\Big) \\
&= S(x,z).
\end{aligned}
$$

□

The theorem 2.2 presents how the total fuzzy similarity is formed in a Lukasiewicz algebra. Generally, the equation

$$S(x,y) = \frac{S_1(x,y)}{n} \oplus \cdots \oplus \frac{S_n(x,y)}{n} \tag{2.5}$$

holds for any injective Wajsberg algebra [16]. Thus, it defines the concept of total similarity in general. Similarly, the generalization of total fuzzy similarity presented in proposition 2.1 for Lukasiewicz algebra holds for any injective Wajsberg algebra [16].

Example 2.1. The concept of total fuzzy similarity is not applicable in Gödel and Gaines structures, since they are not injective Wajsberg algebras. We demonstrate this for Gödel algebra with a counter example. First, we shall note that in Gödel algebra the binary operations $\leftrightarrow$ and $\oplus$ are interpreted as [9]

$$
\begin{aligned}
x \oplus y &= \max\{x,y\}, \\
x \leftrightarrow y &= \begin{cases} 1 & \text{if } x = y, \\ \min\{x,y\} & \text{otherwise.} \end{cases}
\end{aligned}
$$

Then, let us consider two fuzzy subsets A and B on a set $X \in \mathbb{R}$ such that, for the elements $a, b, c \in X$, the following holds

$$
\begin{array}{ll}
\mu_A(a) = \frac{2}{3}, & \mu_B(a) = 0, \\
\mu_A(b) = \frac{1}{3}, & \mu_B(b) = \frac{2}{3}, \\
\mu_A(c) = 0, & \mu_B(c) = 1.
\end{array}
$$

Figure 2.1 shows an example of how these fuzzy subsets could be illustrated graphically.

Furthermore, let S_A and S_B be Gödel-valued fuzzy similarities generated by the fuzzy subsets A and B, respectively. Thus, for the elements a, b, c, the following holds

$$\begin{array}{ll} S_A(a,b) = \frac{1}{3}, & S_B(a,b) = 0, \\ S_A(a,c) = 0, & S_B(a,c) = 0, \\ S_A(b,c) = 0, & S_B(b,c) = \frac{2}{3}. \end{array}$$

Now, as we combine the fuzzy relations S_A and S_B according to the general equation 2.5 of total fuzzy similarity, we see that the resulting combination is not a fuzzy similarity in Gödel algebra. Particularly, the weak transitivity does not hold. Indeed,

$$\begin{aligned} S(a,b) \odot S(b,c) &= \left[\frac{S_A(a,b)}{2} \oplus \frac{S_B(a,b)}{2}\right] \odot \left[\frac{S_A(b,c)}{2} \oplus \frac{S_B(b,c)}{2}\right] \\ &= \min\left\{\frac{1}{2}\cdot\frac{1}{3}, \frac{1}{2}\cdot\frac{2}{3}\right\} \\ &= \frac{1}{6} \nleq 0 = \frac{S_A(a,c)}{2} \oplus \frac{S_B(a,c)}{2} = S(a,c). \end{aligned}$$

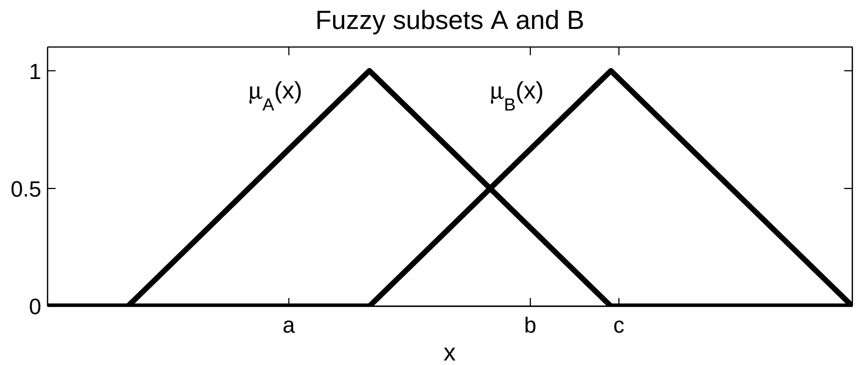

Figure 2.1: An example of the fuzzy subsets A and B.

Chapter 3

Fuzzy Similarity Based Approximate Reasoning

In this chapter we show how the concept of fuzzy similarity can be applied in fuzzy inference systems. It will turn out that such a system is based only on expert knowledge and on the well-defined mathematical concepts that we have been discussing so far. Unlike the traditional fuzzy inference systems, the total fuzzy similarity based system does not require any aggregation or defuzzification methods, which usually do not have a proper mathematical or expert knowledge basis [19]. It is assumed, that the reader is somewhat familiar with the Mamdani and Sugeno fuzzy systems and thus has an understanding about fuzzy inference systems, in general. First, in Section 3.1, we introduce the fuzzy similarity based algorithm to construct the fuzzy inference rules and then, in Section 3.2, we proceed to discuss how the inference is done on that system.

3.1 Algorithm to Construct Fuzzy Inference Systems

Fuzzy inference systems in general are constructed from a finite collection of IF-THEN rules, which are predefined according to the available expert knowledge about the phenomenon to be modelled. Let us consider as an example an inference system S with three input variables and one output variable. Depending on the complexity of the phenomenon in question, the system can have any finite number of input variables [16]. Typically, the rule base of the system S consists of rules of the form

$$\begin{array}{llll}
\text{Rule 1} & \text{IF } x \text{ is } A_1 \text{ and } y \text{ is } B_1 \text{ and } z \text{ is } C_1 & \text{THEN} & w \text{ is } D_1, \\
\text{Rule 2} & \text{IF } x \text{ is } A_2 \text{ and } y \text{ is } B_2 \text{ and } z \text{ is } C_2 & \text{THEN} & w \text{ is } D_2, \\
\vdots & & \vdots & \\
\text{Rule k} & \text{IF } x \text{ is } A_k \text{ and } y \text{ is } B_k \text{ and } z \text{ is } C_k & \text{THEN} & w \text{ is } D_k,
\end{array}$$

where $A_i = (X, \mu_{A_i})$, $B_i = (Y, \mu_{B_i})$, $C_i = (Z, \mu_{C_i})$ and $D_i = (W, \mu_{D_i})$, $i = 1, \dots, k$, are the fuzzy subsets of non-void sets X, Y, Z and W, respectively (cf. [16]). Each rule does not have to involve a fuzzy subset from each input set X, Y, Z. Usually, the output subsets $D_1, \dots, D_k \in L$ are fuzzy, but they can be crisp subsets, too [16]. The product set $\boldsymbol{X} = X \times Y \times Z$ can be seen as the input universe of discourse, which is comprised of objects $\boldsymbol{a} = (x, y, z)$, where $\{x, y, z\}$ is the set of input values. Thus, these input values define the properties of an object $\boldsymbol{a} \in \boldsymbol{X}$.

If one was to model several different aspects of a phenomenon, one might consider having several output variables in his system. However, in such a case the consequent (THEN-part) of each rule should be associated with only one output fuzzy subset, since the output variables should be dependent only on the input variables, not on each other. Thus, the rule base would include separate rule sets for each output variable. In some cases, it might be more convenient to have separate systems altogether for each output variable, especially, if some of the outputs are independent from some of the input variables.

The rule base of a fuzzy inference system is not restricted to rules involving only the fuzzy conjunction, but rules involving also disjunction and negation are possible. In the next section we shall discuss these types of rules separately from the similarity point of view.

The concept of total similarity and its usefulness in reasoning appear to be very intuitive, when this inference method is considered in its application area, in the context of modelling phenomena of real life. Assume one is buying a car. The car is chosen according to certain preferences of the buyer. For instance, the car should be cheap, low in fuel consumption and fast in acceleration. Each of these three criteria can be considered as fuzzy subsets, which define the ideal type of a car. Before making the buying decision, the buyer has compared the properties of several different cars to his preferences on price, fuel consumption and acceleration speed. Finally, based on the similarities between the properties and these preferences, the car most similar to the ideal one is bought. In an inference system implemented to model this specific decision action, the comparison between the properties of a candidate car and an ideal car would correspond to the action of evaluating the partial similarities generated by the above mentioned three fuzzy subsets. The higher the partial similarities are, the higher the degree of total similarity is. In this case, the output of the system would correspond to the car that is most similar to the ideal one. This example system is simple in the sense that it would require only one rule, which involves the fuzzy subsets describing the properties of an ideal car. Typically, a fuzzy inference system would include several rules, where the rule that contributes to the highest degree of total fuzzy similarity defines the output fuzzy subset [16].

Now that we have presented a general form of a rule base and discussed the intuitive nature of total fuzzy similarity, we are ready to introduce the steps required to construct the total fuzzy similarity based inference system S (cf. [16, 19]).

Step 1 Define the IF-THEN rules $1 - k$ of a rule base. Define the input fuzzy subsets A_i, B_i, C_i and the output fuzzy subsets D_i, $i = 1, \ldots, k$, that is, the shapes of the fuzzy subsets.

Step 2 Define weights $m_X, m_Y, m_Z \in \mathbb{N}$ for the input sets X, Y, Z to emphasize the mutual importance of the corresponding input variables x, y, z.

Step 3 Give a criterion on how to distinguish the rules i, j, $i \neq j$ from each other, in case the inputs x_0, y_0, z_0 produce equal degree of total fuzzy similarity in different IF-parts i, j.

Step 4 For each THEN-part i, give a criterion on how to distinguish the crisp outputs $w_0, w_1 \in W$, $w_0 \neq w_1$ that have equal memberships in the output fuzzy subset D_i, that is $\mu_{D_i}(w_0) = \mu_{D_i}(w_1)$.

All the Steps $1 - 4$ should be performed based on expert knowledge. However, the Steps 1 and 2 are flexible in the sense that they do not require complete knowledge to be available: the rule base to be defined in Step 1 can be incomplete, that is, rules can be missing for some combinations of the input fuzzy subsets, since fuzzy inference systems are able to utilize interpolation [16]; if knowledge is lacking of the mutual importance of the input variables, the weights to be defined in Step 2 can be set to one.

If one was to compare the above algorithm to the procedure used to construct a Mamdani or a Sugeno system, one would notice that Step 1 is the only step common to all these three type of systems (similarity, Mamdani and Sugeno). Step 2 enables the incorporation of expert knowledge about some input variables having a greater effect on the output variable than the others to the inference system, if this kind of knowledge is only available. This possibility is missing from Mamdani and Sugeno systems. The knowledge on the mutual importance of input variables can be very useful, resulting in a more accurate model. For instance, if the low price of a car would have the greatest impact on the buying decision amongst other preferences, the car to be chosen might be different from the car chosen based on equally important preferences. The Steps 3 and 4, on the other hand, can be considered to replace the aggregation process included in Mamdani and Sugeno systems and the defuzzification process of Mamdani systems, which are not part of the total similarity based inference as we shall see in the next section. The important difference between the Steps 3 and 4, and the aggregation and defuzzification processes is that the steps are performed based on expert knowledge, whereas the aggregation and defuzzification processes might not be as appropriately justified.

3.2 Inference Process

The inference process presented here is valid for any injective Wajsberg algebra (or MV-algebra), but as Lukasiewicz algebra is a canonical example of such an algebraic structure, we study this process especially from the viewpoint of Lukasiewicz algebra to make it more concrete.

Assume we have an actual input object $\boldsymbol{a}_0 = (x_0, y_0, z_0) \in \boldsymbol{X}$ and in the system S the fuzzy subsets A_i, B_i, C_i, D_i, $i = 1, \dots, k$, each include at least one element that obtains the membership degree of $\mathbf{1} \in L$, in other words, the height of each subset is $\mathbf{1}$($cf.$ [16]). Recall that in the Lukasiewicz structure $\mathbf{1} = 1$. Consider each IF-part i of the rule base as a representation of such an ideal object $\boldsymbol{a}_i = (x_i, y_i, z_i) \in \boldsymbol{X}$ that $\mu_{A_i}(x_i) = \mu_{B_i}(y_i) = \mu_{C_i}(x_i) = 1$. The corresponding output value $w_0 \in W$ is deduced via the following three steps [16, 19], which are to be performed on the general framework of the total similarity based inference system presented in the previous section.

Step A Count the total fuzzy similarities $S_i(\boldsymbol{a}_0, \boldsymbol{a}_i)$ on $\boldsymbol{X}$ between the input values and each IF-part i of the rule base based on the partial similarities $S_{A_i}(x_0, x_i), S_{B_i}(y_0, y_i), S_{C_i}(z_0, z_i)$, which are generated by the input fuzzy subsets A_i, B_i, C_i.

Step B Fire the rule $m \in \{1, \ldots, k\}$ which together with the input object $\boldsymbol{a}_0$ produces the maximal total similarity $S_{max} = \max_{i=1}^{k}\{S_i(\boldsymbol{a}_0, \boldsymbol{a}_i)\}$. If such a rule is not unique, use the criterion given in Step 3 to choose the rule to be fired.

Step C Fire an output value w_0 from the output fuzzy subset D_m corresponding to the THEN-part of the fired rule m such that

$$\mu_{D_m}(w_0) = S_m(\boldsymbol{a}_0, \boldsymbol{a}_m) = S_{max}.$$

If such an output value is not unique, use the criterion given in Step 4 to choose the appropriate output value.

By theorem 2.1, proposition 2.1 and the property 1.17 of bi-residuum, the total similarities for each rule i are counted in Step A as the following weighted means (cf. [19])

$$\begin{aligned}
\frac{m_X \mu_{A_1}(x_0)}{m_X + m_Y + m_Z} + \frac{m_Y \mu_{B_1}(y_0)}{m_X + m_Y + m_Z} + \frac{m_Z \mu_{C_1}(z_0)}{m_X + m_Y + m_Z} &= S_1(\boldsymbol{a}_0, \boldsymbol{a}_1) \\
\frac{m_X \mu_{A_2}(x_0)}{m_X + m_Y + m_Z} + \frac{m_Y \mu_{B_2}(y_0)}{m_X + m_Y + m_Z} + \frac{m_Z \mu_{C_2}(z_0)}{m_X + m_Y + m_Z} &= S_2(\boldsymbol{a}_0, \boldsymbol{a}_2) \\
\vdots \qquad\qquad\qquad & \qquad \vdots \\
\frac{m_X \mu_{A_k}(x_0)}{m_X + m_Y + m_Z} + \frac{m_Y \mu_{B_k}(y_0)}{m_X + m_Y + m_Z} + \frac{m_Z \mu_{C_k}(z_0)}{m_X + m_Y + m_Z} &= S_k(\boldsymbol{a}_0, \boldsymbol{a}_k),
\end{aligned}$$

where the weights $m_X, m_Y, m_Z \in \mathbb{N}$ are given in Step 2 (see the previous section).

The above method for computing the total fuzzy similarities is appropriate only, if our assumption about each fuzzy subset having the height 1 holds, which is, however, a very typical property of fuzzy subsets in any types of fuzzy inference systems. If the heights of fuzzy subsets vary, the total similarity based reasoning presented so far needs to be modified slightly to a more general form. In such a case, the ideal object $\boldsymbol{a}_i$ representing a rule i can consist of elements x_i, y_i, z_i which are not necessarily mapped to the truth value 1 in their corresponding fuzzy subsets A_i, B_i, C_i. However, each of these elements should achieve the highest degree of truth possible in the fuzzy subset it belongs to. Accordingly, by equation 1.12, the formula of total fuzzy similarity for Lukasiewicz algebra takes the following general form for a rule i

$$S_i(\boldsymbol{a}_0, \boldsymbol{a}_i) = \frac{m_X(1 - |\mu_{A_i}(x_0) - \mu_{A_i}(x_i)|)}{m_X + m_Y + m_Z} + \frac{m_Y(1 - |\mu_{B_i}(y_0) - \mu_{B_i}(y_i)|)}{m_X + m_Y + m_Z} + \frac{m_Z(1 - |\mu_{C_i}(z_0) - \mu_{C_i}(z_i)|)}{m_X + m_Y + m_Z}.$$

Moreover, the corresponding formula of total fuzzy similarity for any injective Wajsberg algebra takes the form of

$$S_i(\boldsymbol{a}_0, \boldsymbol{a}_i) = \frac{m_X[\mu_{A_i}(x_0) \leftrightarrow m_X\mu_{A_i}(x_i)]}{m_X + m_Y + m_Z} \oplus \frac{m_Y[\mu_{B_i}(y_0) \leftrightarrow m_X\mu_{B_i}(y_i)]}{m_X + m_Y + m_Z} \oplus \frac{m_Z[\mu_{C_i}(z_0) \leftrightarrow m_X\mu_{C_i}(z_i)]}{m_X + m_Y + m_Z}.$$

Step C has a many-valued logic based theoretical justification, since it can be regarded as an instance of the Generalized Modus Ponens scheme of fuzzy rule of inference. Briefly, a fuzzy rule of inference consists of two components, where the first component operates on formulae and the second component operates on the truth values of the formulae. The following represents the scheme for Generalized Modus Ponens [16, 25]

$$R_{GMP} : \frac{\alpha, (\alpha \text{ imp } \beta)}{\beta}, \frac{a, b}{a \odot b}.$$

In the context of total fuzzy similarity based inference, α and β represent formulae, which correspond to the IF-part and THEN-part of the rule that contributes to the maximal total similarity of the rule base, respectively. The truth value a is the degree of the maximal total similarity and $b = 1$. Let us denote the truth value function as $v : \mathcal{F} \mapsto [0, 1]$, where $\mathcal{F}$ is a set of formulae. Indeed, $v(\alpha \text{ imp } \beta) = b = 1$, since $v(\alpha) = v(\beta)$. The soundness of the conclusion can be verified by $v(\alpha) \odot v(\alpha \text{ imp } \beta) = a \odot b = a \odot 1 = a = v(\beta)$.

Notice that, the (weighted) average of partial similarities, that is, the total similarity, represents fuzzy conjunction in the inference system. One might wonder, why this particular representation involving the operation $\oplus$ is selected for this purpose, though it is mathematically well-defined and intuitive from the similarity point of view. After all, the operation $\oplus$ is the algebraic counterpart of logical disjunction, not conjunction. For instance, other candidates for representing fuzzy conjunction could be the product operation $\odot$, which is the counterpart of logical conjunction, the *min* operation, which is used generally in the traditional systems, and the natural product $\cdot$ defined for real numbers. Let us study this question through the following examples.

Example 3.1. Let S_i be a Lukasiewicz-valued fuzzy similarity relation on a non-void set $X_i, i = 1, \ldots, n$. In addition, let R_A, R_B, R_C be binary fuzzy relations on the product set $X = X_1 \times \cdots \times X_n$ defined as

$$\begin{aligned}
R_A(x, y) &= \frac{1}{n}[S_1(x, y) \odot S_2(x, y) \odot \cdots \odot S_n(x, y)], \\
R_B(x, y) &= \frac{1}{n} \min_{i=1}^{n}\{S_i(x, y)\}, \\
R_C(x, y) &= \frac{1}{n} \prod_{i=1}^{n} S_i(x, y).
\end{aligned}$$

The fuzzy relations R_A, R_B, R_C are not similarity relations, since the condition 2.1 of reflexivity does not hold. Indeed, it is easy to see that, for all $x \in X$,

$$R_A(x, x) = R_B(x, x) = R_C(x, x) = \frac{1}{n} \neq 1.$$

Since R_A, R_B, R_C are not similarity relations, they are inadequate for representing fuzzy conjunction in the total similarity based inference system. The condition of reflexivity fails, since the relations involve the fraction $1/n$. However, it is important that the relation selected to represent fuzzy conjunction is a similarity relation, even when it involves this fraction. Otherwise, the mathematical basis for allowing weighted input variables in the system would become ill-defined. Notice, that the opportunity to define the mutual importance of input variables is based on the divisibility axiom 1.24 of an injective MV-algebra that is fulfilled by the Lukasiewicz-valued inference system. However, if this opportunity is not regarded as important, the product operation $\odot$ could be candidate for representing fuzzy conjunction.

Example 3.2. Let S_i be a Lukasiewicz-valued fuzzy similarity relation on a non-void set $X_i, i = 1, \ldots, n$. Then,

$$S(x, y) = S_1(x, y) \odot S_2(x, y) \odot \cdots \odot S_n(x, y)$$

is a Lukasiewicz-valued fuzzy similarity on $X = X_1 \times \cdots \times X_n$.

To prove this argument, we need to show that the relation S is reflexive, symmetric and weakly transitive. The reflexivity and symmetry of S follows from the reflexivity and symmetry of the relations $S_i, i = 1, \ldots, n$. The weak transitivy of S can be shown in the following way. Let A $= S(x, y) \odot S(y, z)$. If $S(x, y) + S(y, z) - 1 \leq 0$, then $A = 0$ by condition 1.8. Thus, certainly $S(x, z) \geq A$. Otherwise,

$$\begin{aligned}
A &\overset{1.8}{=} S(x, y) + S(y, z) - 1 \\
&= \Big(S_1(x, y) \odot \cdots \odot S_n(x, y)\Big) + \Big(S_1(y, z) \odot \cdots \odot S_n(y, z)\Big) - 1 \\
&= \Big(S_1(x, y) + \cdots + S_n(x, y) - (n - 1)\Big) \\
&+ \Big(S_1(y, z) + \cdots + S_n(y, z) - (n - 1)\Big) - 1 \\
&= \Big(S_1(x, y) + S_1(y, z) - 1\Big) + \cdots + \Big(S_n(x, y) + S_n(y, z) - 1\Big) - n + 1 \\
&\overset{1.8}{=} \Big(S_1(x, y) \odot S_1(y, z)\Big) + \cdots + \Big(S_n(x, y) \odot S_n(y, z)\Big) - n + 1 \\
&\overset{2.3}{\leq} S_1(x, z) + \cdots + S_n(x, z) - n + 1 \\
&= S_1(x, z) \odot \cdots \odot S_n(x, z) \\
&= S(x, z).
\end{aligned}$$

Thus, we have shown that S is a similarity relation.

Since S, as defined in the above example, is a similarity relation, we might consider the product operation $\odot$ as a suitable representation of fuzzy conjunction in the similarity based inference system. However, this would be appropriate only in the situations, when we do not need to define the mutual importance of the input variables, or do not have enough information to do this. Furthermore, in this case, the framework provided with the axioms of an injective MV-algebra would not be fully utilized. In the upcoming example 3.4, we see that the *max* operation is not a similarity relation, even when we do not involve the fraction $1/n$ with it.

The inference method introduced so far for the system to perform Step A, is applicable for rules that involve the fuzzy conjunction only, though rules in general can be constructed from fuzzy disjunction and negation, too. A rule involving fuzzy negation has an intuitive role in the concept of total fuzzy similarity. For instance, consider a rule r of the form

IF x is *not* A_r and y is *not* B_r and z is *not* C_r THEN w is D_r.

The concept of total similarity regarding such a rule can be interpreted as the degree of total non-similarity n-$S_r(\boldsymbol{a_0}, \boldsymbol{a_r})$ between the input object $\boldsymbol{a_0}$ and the ideal object $\boldsymbol{a_r}$, which is computed by combining the degrees of the partial non-similarities n-$S_{A_r}(x_0, x_r)$, n-$S_{B_r}(y_0, y_r)$ and n-$S_{C_r}(z_0, z_r)$ together.

This computation is done similarly to the computation of total similarity, only the partial similarities in the formula need to be substituted with the corresponding partial non-similarities defined by equation 2.4.

The representation of fuzzy disjunction from the total fuzzy similarity point of view is not as intuitive as the representations of conjunction and negation might be. Let us get back to our example of the decision making action related to buying a car. Imagine the buyer having two alternative preferences; he would like to buy a car which is low in fuel consumption *or* fast in acceleration. Preferences like these alternatives to each other might appear inappropriate from the total similarity point of view. It seems that once one of the buyer's preferences is fulfilled, the remaining property ceases to exist as a preference, since he does not care about it anymore. Thus, in this case, we are not interested in the concept of total similarity, which suggests we should find a car which fulfills both the preferences as well as possible. Instead, we are interested only in the partial similarities: the car is perfect, if it fulfills completely the preference on fuel consumption, but fails to fulfill the preference on acceleration speed, and vice versa.

Furthermore, for the Lukasiewicz-valued similarity relations the well-known laws of de Morgan do not hold, when the average of the partial similarities $S_i, i = 1, \ldots, n$ is interpreted as the counterpart of logical conjunction $\wedge$ and $1 - S_i$ is considered as the counterpart of negation $\neg$. This can be shown with the following simple counter example.

Example 3.3. Let us consider two Lukasiewicz-valued similarity relations S_A and S_B on a set X. Then we have,

$$\begin{aligned}\neg(\neg S_A \wedge \neg S_B) &= 1 - \frac{(1 - S_A) + (1 - S_B)}{2} \\ &= 1 - \frac{2 - S_A - S_B}{2} = \frac{S_A + S_B}{2} = S_A \wedge S_B,\end{aligned}$$

which results in a contradiction with the law $\neg(p \wedge q) \Leftrightarrow \neg p \vee \neg q$ of de Morgan.

The example suggests that $S_A \vee S_B = S_A \wedge S_B$, that is, for the Lukasiewicz-valued similarity relations the representations of fuzzy conjunction and disjunction are the same, which is in a disagreement with classical logic. However, it is not clear whether the average of the similarity relations should really be interpreted as the counterpart of logical conjunction, even though it represents fuzzy conjunction. Moreover, this results in the interesting question of how the rules involving fuzzy disjunction should then be managed in the inference process. One possibility would be to sidestep this problem by splitting the rule into several parts. For instance, a rule r of the form

$$\text{IF } x \text{ is } A_r \ \mathit{or} \ y \text{ is } B_r \ \mathit{or} \ z \text{ is } C_r \quad \text{THEN} \quad w \text{ is } D_r.$$

can be split into the three following rules

$$\begin{array}{lll} \text{IF } x \text{ is } A_r & \text{THEN} & w \text{ is } D_r, \\ \text{IF } x \text{ is } B_r & \text{THEN} & w \text{ is } D_r, \\ \text{IF } x \text{ is } C_r & \text{THEN} & w \text{ is } D_r. \end{array}$$

Now, the total similarities to be counted for each rule in Step A refers to counting partial similarities, and by Step B the rule which produces the maximal partial similarity is fired. Step C is performed in the usual way. Thus, in the case of alternative preferences, the inference process would be performed exactly as desired, that is, only the preference that is fulfilled with the greatest degree would count.

The proposed method of performing inference on the rule r is actually equal to computing the maximum of the partial similarities, that is,

$$R_r(\boldsymbol{a}_0, \boldsymbol{a}_r) = \max\{S_{A_r}(x_0, x_r), S_{B_r}(y_0, y_r), S_{C_r}(z_0, z_r)\},$$

where R_r is a binary fuzzy relation on the product set $\boldsymbol{X}$. This suggests maximum as a candidate for representing fuzzy disjunction in the rules. However, the resulting relation R_r is not a similarity relation in Lukasiewicz algebra, which is shown in the following example.

Example 3.4. Let us consider two fuzzy subsets A and B on a set $X \in \mathbb{R}$ such that, for the elements $a, b, c \in X$, the following holds

$$\begin{array}{ll} \mu_A(a) = 1, & \mu_B(a) = 0, \\ \mu_A(b) = \frac{2}{3}, & \mu_B(b) = \frac{1}{4}, \\ \mu_A(c) = 0, & \mu_B(c) = 1. \end{array}$$

Furthermore, let S_A and S_B be Lukasiewicz-valued fuzzy similarity relations generated by the fuzzy subsets A and B, respectively, and let the binary fuzzy relation R on X^2 be defined as $R(x, y) = \max\{S_A(x, y), S_B(x, y)\}$, for all $x, y \in X^2$.

The fuzzy relation R is not a similarity relation in Lukasiewicz algebra, since the weak transitivity does not hold. Indeed,

$$\begin{aligned}
R(a,b) \odot R(b,c) &= \max\Big\{ \max\{S_A(a,b), S_B(a,b)\} \\
&\quad + \max\{S_A(b,c), S_B(b,c)\} - 1, 0\Big\} \\
&= \max\left\{ \max\left\{\frac{2}{3}, \frac{3}{4}\right\} + \max\left\{\frac{1}{3}, \frac{1}{4}\right\} - 1, 0\right\} \\
&= \max\left\{\frac{3}{4} + \frac{1}{3} - 1, 0\right\} \\
&= \frac{1}{12} \not\leq \max\{S_A(a,c), S_B(a,c)\} = 0 = R(a,c).
\end{aligned}$$

In addition, the condition of reflexivity does not hold, when the elements $a, b, c \in X$ are associated with weights, as the results of example 3.1 suggest for the *min* operation.

It seems that applying the total fuzzy similarity approach to the rules involving fuzzy disjunction is not feasible. Firstly, this approach is not intuitive, secondly, it disagrees with classical logic and thirdly, the suitable method proposed for processing these rules does not fulfill the definition of similarity relation. However, this argument is not valid before a firm mathematical basis can be provided for it.

We have now discussed the total fuzzy similarity based inference system thoroughly. As a last point, we consider the aggregation and defuzzification processes specific especially to the Mamdani fuzzy systems from the total fuzzy similarity point of view. As the introduced inference process suggests, the aim of the total similarity based inference is to fire only one rule, that is, the rule that corresponds to the degree of maximal total similarity within the rule base. Thus, an actual aggregation process is not required despite the possibility of the cases occurring where more than one rule contributes to the same degree of maximal total similarity, since this problem is solved through the expert knowledge. Furthermore, since only one rule is fired, only one output fuzzy subset needs to be studied, which diminishes the need of the defuzzification process. However, if more than one crisp value from the output fuzzy subset appears as a candidate for the output value, again expert knowledge is used to choose the desired one. Notice also, that if all the fuzzy subsets of a system are of height 1, the process called fuzzification of inputs in Mamdani and Sugeno systems is actually equivalent to the calculation of partial similarities in the total fuzzy similarity based system.

Finally, we conclude that a fuzzy inference system based on total similarity is well-defined, when it does not include rules involving fuzzy disjunction.

Indeed, Step A is entirely based on the mathematical concept of fuzzy similarity relation. Step B utilizes only expert knowledge in unclear situations and Step C is justified with the Generalized Modus Ponens rule of fuzzy inference and also, uses expert knowledge only to solve unclear situations.

Chapter 4

Application of Total Fuzzy Similarity Based Inference

One of the main objectives of the HeartCycle EU-project is to develop innovative technical support for cardiac patients in order to improve their compliance to the prescribed treatments. For this patient group the treatment of high blood pressure is especially important and regular exercise is considered as a vital part of the treatment. In this chapter we present the total fuzzy similarity based inference system created to model the effect of regular aerobic exercise on resting blood pressure. Aerobic exercise refers to physical activity that activates the cardiovascular system such as brisk walking, jogging or bicycling. In addition to the similarity based system, also a Sugeno type system was developed, which was actually the initial approach to model the phenomenon in question. However, in this chapter we focus mainly on the similarity based system and discuss the Sugeno type system only when we compare the performances of these two systems in Subsection 4.2.2.

In Section 4.1 we introduce the expert knowledge that was gathered for the modelling purposes. This knowledge was acquired by conducting a literature review on the effects of different lifestyle intervention on blood pressure regulation. A summary of the findings concerning the effects of regular aerobic exercise is presented here. These findings are based on reviews and meta-analyses of altogether 201 exercise intervention trials and on several individual clinical studies.

In Section 4.2 we first describe the structure of the similarity based inference system, namely the input and output variables, fuzzy subsets associated with these variables and the rule base of the system. The described structural components apply also to the Sugeno system. In the second part of the section we evaluate the performances of both types of models at a very basic level, since actual data for proper model validation were not available at the time. In this context we also compare the outcomes of these two models.

4.1 Description of the Expert Knowledge

First, we introduce the general properties of the study populations the clinical trials were conducted on, since the models are constructed based on these properties. Thus, the performance of the models should be most reliable for subjects representing these populations. Then, we present the overall blood pressure (BP) changes observed in different studies to provide some idea of the effectiveness of exercise in reducing BP. For the modelling purposes, it is especially interesting to know what are the variables that influence BP outcome and what kind of influence do they have. These variables turn out to be related to certain individual properties and the performed exercise dose. Thus, the focus of this section is to present the results regarding the relationships between these variables and BP change. In the end of this section we summarize these relationships shortly.

4.1.1 Study Population

The overall population studied is diverse including men and women of different ethnicities, ages, body compositions and hypertension statuses (both normotensives and hypertensives). However, most of the subjects are middle-aged Asians or Caucasians having mild or moderate level of hypertension, which refers to systolic blood pressure (SBP) between 140 and 179 mmHg and diastolic blood pressure (DBP) between 90 and 109 mmHg. Only one trial [2] includes subjects with severe hypertension (SBP $>$ 180 mmHg or DBP $>$ 110 mmHg) and some African-Americans are included in studies [27, 10]. Moreover, the exercise trials are mainly conducted on sedentary subjects who have not performed regular exercise for 6-12 months prior to entering the exercise trials. Most of the exercise trials lasted for three or four months.

4.1.2 Overall Blood Pressure Changes

The mean blood pressure change reported in the meta-analyses varies from -20 to +9 mmHg for SBP and from -12 to +11 mmHg for DBP [5, 27, 8]. None of the individual studies included in the conducted literature review reported an increase in blood pressure, but otherwise their results fit into these ranges, with one exception that reported even larger reductions [2]. Certainly, individual factors are responsible for this inconsistency, but also the varying study methods and exercise conditions affect the measured BP outcomes. However, it is clear that there are non-responders to exercise-induced blood pressure regulation. Despite this, the meta-analyses indicate that in general regular exercise does reduce blood pressure in all subjects independent of individual characteristics. Cornelissen and Fagard [5] reported in their meta-analysis on 72 trials that in the overall population the resting SBP/DBP was reduced on the average by 3.0/2.4 mmHg. Fagard [8] reported similar reductions over 44 trials. Whelton et al. [27] reported a slightly greater mean decrease in BP of 3.8/2.6 mmHg over 52 trials.

4.1.3 Individual Properties

The available studies provide information on how some individual properties influence the exercise-induced BP response. This information regards blood pressure level, gender, age, ethnicity and body composition. Next, we discuss the relationships of these properties to blood pressure change.

Regular exercise reduces blood pressure substantially more in hypertensive than in pre-hypertensive (SBP $120-139$ mmHg and DBP $80-89$ mmHg) or normotensive (SBP < 120 mmHg and DBP < 80) subjects [5, 27, 1, 22]. Hagberg et al. [10] performed a review on 62 trials, which included only hypertensive populations (mean baseline BP 153/97 mmHg). They reported that the average BP reduction observed in responders was 10.6/8.2 mmHg. Approximately 20% of this overall population was non-responders, since they did not manage to decrease their blood pressure with exercise. Cornelissen and Fagard had divided their overall population into hypertensive (28 trials), pre-hypertensive (33 trials) and normotensive (15 trials) subgroups. They reported similar results for the hypertensive subgroup as Hagberg et al. For both the pre-hypertensive and normotensive subgroups the authors observed an approximate average reduction in blood pressure of 2/2 mmHg, which is substantially less than reported for the hypertensive population. However, it should be noted that unlike Hagberg et al., Cornelissen and Fagard included also the non-responders into their results, which obviously decreases the average results.

The results of Cornelissen and Fagard indicate that pre-hypertensive and normotensive subjects reduce their BP with exercise to similar extent. However, the normotensive subgroup was statistically significantly younger than the pre-hypertensive subgroup, which might weaken the reliability of this conclusion. Whelton et al. report an average reduction of 4.0/2.3 mmHg for a subgroup including both normotensive and pre-hypertensive subjects (27 trials). The subgroup properties are not specified in this meta-analysis.

Though, it seems that pre-hypertensive and normotensive subjects have a very similar BP response, among hypertensive subjects the effect of different BP levels is clear. It seems that the higher the hypertension status is, the greater the BP reductions are, which we conclude based on the results of trial [2]. However, the reliability of this study is not conclusive.

The effect of gender on the blood pressure response to exercise was studied by Hagberg et al. The results indicate that hypertensive women reduce their blood pressure somewhat more and more consistently than men. All the women in the population were SBP responders showing an average SBP reduction of 14.7 mmHg, whereas the men who were responders (72%) reduced their SBP by 8.7 mmHg. The same trend was reported for DBP with women showing an average reduction of 10.5 mmHg (89% responders) and men 7.8 mmHg (82% responders).

In the hypertensive population middle-aged ($41-60$ years) subjects tend to have somewhat greater and more consistent falls in SBP due to exercise-training than younger ($21-40$ years) and older (60+ years) subjects [10, 12]. For DBP, Hagberg et al. suggest that age does not influence the exercise-induced reductions. However, an individual clinical trial [12] reported approximately twice as great reductions in DBP in the younger ($30-49$ year) than older ($50-69$ years) subjects. This result can be considered reliable, since the trial was conducted on a sufficiently large study sample (n=109) and it was especially designed to study the effect of age on BP response.

Whelton et al. studied the effect of ethnicity on blood pressure response to exercise in an overall population that included both normotensive and hypertensive subjects. According to their results, African-Americans tend to benefit the most from regular exercise, Asians the second most and Caucasians the least in terms of SBP. Furthermore, the SBP response of African-Americans was reported to be approximately three-fold and Asians two-fold greater compared to the response of Caucasians. However, these substantial differences should be interpreted cautiously, since the baseline BP values for these subgroups are not specified and the number of Asian and African-American subjects is relatively small. Hagberg et al. reported the a similar trend for hypertensive Asians and Caucasians, but with smaller differences: Asian subjects reduced their SBP 40% (∼1.5-fold) more than Caucasians.

For DBP, Whelton et al. concluded that Asian subjects show two times greater reductions than Caucasian and African-American subjects, whereas African-Americans and Caucasians have similar DBP response. However, Hagberg et al. did not report any significant differences in the DBP response between hypertensive Asian and Caucasian subjects, but the reduction seems to be more consistent with Asians (85% responders) than Caucasians (75% responders). The results of Hagberg et al. regarding the differences between Asians and Caucasians might be more reliable than the results of Whelton et al., since the number of Asians is sufficient in the sample of the former.

None of the clinical studies reported significant differences in the exercise-induced blood pressure changes between normal and overweight populations. Though, this was not fully explored, since the number of obese individuals was small in the study groups. Thus, the effect of baseline body composition (weight, body mass index, abdominal fat) on exercise-induced BP response remains still to be studied.

4.1.4 Exercise Dose

Presumably, the blood pressure benefits of exercise are dependent on the exercise dose, which is a combination of intensity and duration of the training sessions, and exercise frequency and type. In this subsection we see, how these exercise components are related to blood pressure.

Light to moderate intensity exercise of about 50% of the maximal oxygen consumption (VO_2max) seems to be as effective in reducing blood pressure as training at a vigorous intensity of about 75% of VO_2max, if not even more effective [10, 1, 14, 13]. Exercise intensity describes the effort or energy cost associated with physical activity. Moderate intensity exercise is equivalent to a brisk walk that noticeably accelerates the heart rate [11]. Hagberg et al. noticed in their review that the studies using training intensities less than 70% of VO_2max achieved substantially greater reductions in SBP than studies using higher intensities. For DBP the effect of different exercise intensities was not as obvious. Aroll and Beaglehole [1] reported in their review slightly greater BP reductions with lower exercise intensities especially for the elderly. Also, the American College of Sports Medicine (ACSM) and the American Heart Association (AHA) recommend moderate intensity exercise of 40-60% of VO_2max for hypertensive individuals [18].

There is some, but weak evidence from the available clinical studies that daily training sessions produce greater blood pressure benefits than an exercise frequency of three times per week. Aroll and Beaglehole reported the results of a small trial (n=13), which compared the effectiveness of 3 and 7 times per week training frequencies in reducing the blood pressure of hypertensive subjects. The fall in SBP was 45% and in DBP 22% greater for the group having daily training sessions than for the other group having three weekly training sessions. In addition, the individual trial [2] reported huge reductions in both SBP and DBP with daily exercise, but as mentioned in Subsection 4.1.3, the reliability of this study might be questionable. However, daily exercise is also recommended by ACSM and AHA for hypertensive individuals [18]. Thus, it seems that daily exercise is more beneficial in reducing blood pressure in hypertensive subjects than less frequent exercise, but more evidence is needed to validate this conclusion.

All types of aerobic exercise (walking, jogging, running, cycling, swimming and recreational sports) were reported to be successful in reducing blood pressure. In addition, Aroll and Beaglehole concluded that also resistance training can reduce blood pressure.

The available individual clinical trials were not able to show whether the exercise duration has an effect on blood pressure outcome, since none of them studied this relation. However, the meta-analyses [5, 27, 8] reported that BP response is not related with exercise duration, but this is most likely due to the homogeneous exercise duration performed in the different trials, since most of the trials had exercise sessions of similar length (30-45 min).

4.1.5 Summary of the Expert Knowledge

As a conclusion to this section we summarize the relationships between the exercise-induced BP reduction and the properties (variables) that affect the BP response. The rule bases of the models discussed in the next section are constructed mainly according to these relationships.

BP status

- The higher the hypertension status is, the greater the BP reduction is.
- Hypertensive subjects reduce BP substantially more than pre-hypertensive or normotensive subjects.
- Pre-hypertensive and normotensive subjects reduce BP similarly.

Gender

- Women reduce BP more than men.

Age

- Young subjects reduce SBP slightly more than old subjects, whereas middle-aged subjects reduce SBP the most.
- Younger subjects reduce DBP more than older subjects.

Ethnicity

- Asian subjects reduce SBP more than Caucasian subjects.
- Hypertensive Asians and Caucasians reduce DBP similarly.
- Pre-hypertensive and normotensive Asians reduce DBP more than Caucasians.

Intensity

- Low to moderate intensity exercise reduces BP more than high intensity exercise.

Frequency

- Daily exercise reduces BP more than exercise performed 3 times/wk.

4.2 Description of the Inference System

The knowledge presented in the previous section is rather vague providing only imprecise relationships between the BP change and the properties that influence blood pressure. However, since fuzzy reasoning has the ability to manage this kind of imprecision characteristic of human knowledge, it is an appropriate method for us to apply to the knowledge we have.

In Subsection 4.2.1 we describe the structure of the total fuzzy similarity based inference system constructed according to the vague expert knowledge we have. The structure, as well as the inference process of this system, are built according to the steps defined in Chapter 3. One should keep in mind that the input and output variables, membership functions and rules to be presented shortly are the same for the Sugeno type system, too. Only the inference processes differ from each other in the two types of systems. Both systems were implemented with MATLAB (the MathWorks, Inc; Natick, Massachusetts). The Fuzzy Toolbox of MATLAB provided direct tools for creating the structure and inference process of the Sugeno type system. However, the inference process of the similarity based system had to be built from scratch, but the structure initially built for Sugeno could be reused for the similarity system.

In Subsection 4.2.2 we perform a very basic assessment of the performance abilities of these two types of models by comparing their outcomes to the expert knowledge. In the same connection we compare the outcomes of these models to each other. Since proper data were not available for the validation of the models, we can only consider the possible differences in their performances, but cannot make any conclusions on which of the systems provide the most correct outcomes.

4.2.1 Structure of the Model

The individual properties that were recognized to influence significantly blood pressure are BP status, gender, age and ethnicity, whereas the important properties related to the performed exercise dose are exercise intensity (% VO_2max) and the weekly exercise frequency. Thus, these properties are considered as the input variables of the model, whereas the BP change is considered as the output of the model. The BP change predicted by the model should be achieved within 3 to 4 months after initiating the regular exercise plan, which comprises aerobic training sessions of 30-45 minutes of duration (see Subsections 4.1.1 and 4.1.4) with the values of intensity and frequency given as input. The predicted BP change corresponds to the average BP reduction that should be observed in the population of people who have the individual properties given as inputs. Moreover, since most of the study populations were sedentary at baseline, the model is applicable only to sedentary people, who are becoming physically active.

The model is divided into several submodels for simplicity reasons: SBP response and DBP response are predicted in different models and separate submodels are constructed also for Caucasian women, Caucasian men, Asian women and Asian men. Submodels for African-American subjects were not created, since the overall study population of the clinical trials did not represent them sufficiently. Instead of going through all of these submodels separately, we have chosen to present two submodels as examples, namely the SBP and DBP models created for Caucasian women. The SBP and DBP submodels created for the other ethnicity and gender combinations are very similar to these example submodels. Only the membership functions of the output variable differ in these models.

The first step in constructing a fuzzy inference system is to define the membership functions and the rule base of the system. The membership functions for the input variables were fairly simple to define from the available knowledge: The fuzzy subsets associated with different input variables were easily identified from the linguistic terms used in the clinical trials. In addition, it was clear which values of a fuzzy subset should obtain the highest membership degree 1. Defining the shapes of the membership functions was only a matter of choice and both Gaussian and triangular shapes were used. Indeed, preliminary tests of the submodels indicated that the selected shape made hardly any difference in their outcome.

Figures 4.1, 4.2 and 4.3 represent the fuzzy subsets of each input variable of the SBP and DBP submodels. The input variables exercise intensity and frequency are same for these submodels, but the variable containing the information of blood pressure status is different for them. SBP status and DBP status are the inputs for the SBP and DBP submodels, respectively. Moreover, the number of fuzzy subsets for the variables BP status and age differ in the two submodels, since the preliminary tests of the submodels indicated that the missing fuzzy subsets had hardly any effect on the outcomes of the submodels. Thus, the subsets were regarded as unnecessary at the time. The submodels do not require ethnicity and gender as inputs, since this information is already incorporated into them.

For SBP and DBP submodels the output variable BP change refers to the predicted SBP and DBP change, respectively. Rather precise information on BP change could be extracted from the clinical trials, since it was expressed as the average BP reduction observed in different populations. This precise information was maintained by creating constant membership functions for the output variable that correspond to the average BP reductions. The range

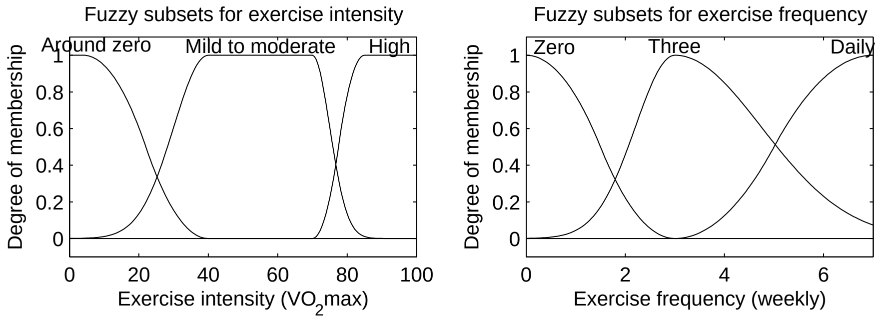

Figure 4.1: Fuzzy subsets for input variables exercise intensity and frequency. These variables are common for both SBP and DBP submodels.

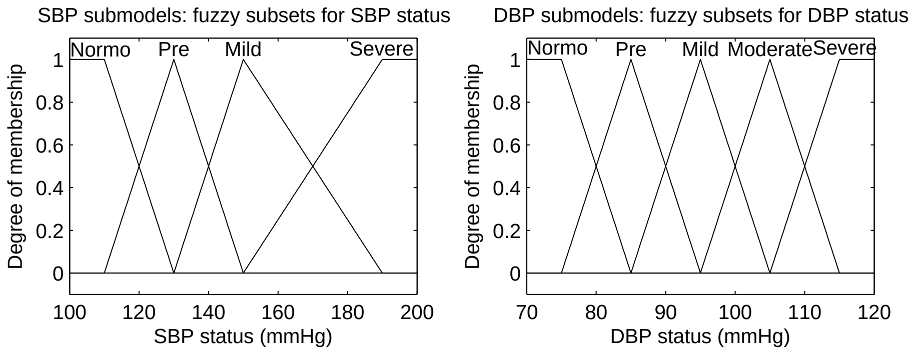

Figure 4.2: Fuzzy subsets for the input variable BP status. The variable is different for SBP and DBP submodels.

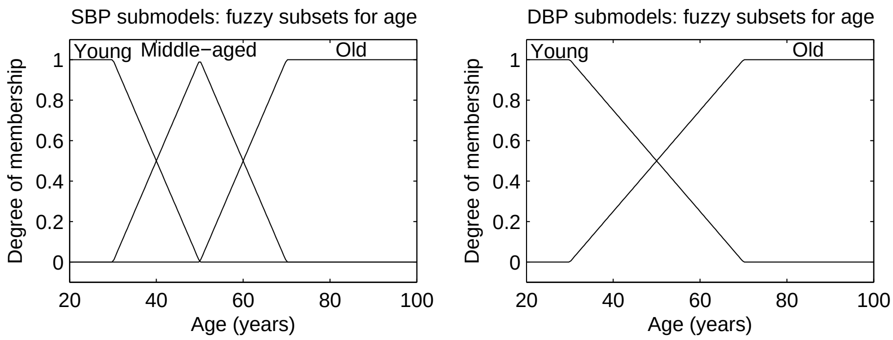

Figure 4.3: Fuzzy subsets for the input variable age. The membership functions and the number of fuzzy subsets are differenct for SBP and DBP submodels.

of these constant values varies for each submodel according to the ethnicity and gender properties. For instance, in the SBP submodel for Caucasian women the SBP reduction varies between -50 and 0 mmHg, whereas for Caucasian men the reduction varies from -30 to 0 mmHg. This is consistent with the expert knowledge which indicates that women show greater exercise-induced BP reductions than men. The SBP submodels have 37 constant values for the output variable SBP change, and the DBP submodels have 31 constant values for the output variable DBP change. For our example submodels, the output value range of the SBP submodel is -50 $-$ 0 mmHg as already mentioned and for the DBP submodel it is -30 $-$ 0 mmHg.

In a complete rulebase the number of rules would be high: $108 = (3 \times 3 \times 4 \times 3)$ for SBP submodels and $90 = (3 \times 3 \times 5 \times 2)$ for DBP submodels. This number of rules would increase to 135 for both submodels, if the currently missing fuzzy subsets would be regarded as essential and thus included in the models. However, a truly complete rulebase is not required for our submodels, since the exercise related inputs allow us to simplify the rule base a bit as will be explained in the following.

Certainly, a person that does not exercise at all cannot show a BP response to exercise. This assumption can be expressed by the terms used in the models in the following way: *if the intensity is around zero or the frequency is zero, the BP decrease is 0/0 mmHg*. Furthermore, the available knowledge suggests that daily exercise results in greater BP reductions than exercise performed less frequently, but high intensity exercise contributes to smaller BP reductions than low or mild intensity exercise. From this we reason that if high intensity exercise is too strenuous to decrease BP effectively, it would be even more strenuous when performed daily, and thus should not result in any lower BP. On the contrary, BP could even increase. However, since this conclusion cannot be verified by the available expert knowledge, the rules involving the fuzzy subset of high intensity were formulated to indicate that the exercise frequency does not have an effect on BP. This might be a bit weaker assumption to make than the original one.

Based on these two assumptions, the number of rules in the SBP models was reduced to 38 rules and in the DBP submodels to 32 rules. Though, the current number of rules in the submodels is notably less than it would be with the complete rule bases, it is still quite high. Hence, we will not present all the individual rules, but rather give a general idea on how the rules look like. As an example we present the rules included in the SBP and DBP submodels of Caucasian women that involve the mild BP status. The rules are similar for the other hypertension statuses also. However, the presented rules have the most solid expert knowledge basis, since most of the study populations included mild hypertensives.

The one rule common for both submodels is

IF intensity is *around zero* OR frequency is *zero*	THEN SBP reduction is 0/0,

which is divided into the following two separate rules for the reasons explained in Section 3.2

IF intensity is *around zero*	THEN BP reduction is 0/0,
IF frequency is *zero*	THEN BP reduction is 0/0.

The rules for the SBP submodel for mild hypertensive Caucasian women are

IF intensity is *low/moderate* AND frequency is *three* AND SBP status is *mild* AND age is *young*	THEN SBP reduction is 7.5,
IF intensity is *low/moderate* AND frequency is *three* AND SBP status is *mild* AND age is *middle*	THEN SBP reduction is 10.2,
IF intensity is *low/moderate* AND frequency is *three* AND SBP status is *mild* AND age is *old*	THEN SBP reduction is 6.6,
IF intensity is *low/moderate* AND frequency is *daily* AND SBP status is *mild* AND age is *young*	THEN SBP reduction is 10.9,
IF intensity is *low/moderate* AND frequency is *daily* AND SBP status is *mild* AND age is *middle*	THEN SBP reduction is 22.5,
IF intensity is *low/moderate* AND frequency is *daily* AND SBP status is *mild* AND age is *old*	THEN SBP reduction is 9.6,
IF intensity is *high* AND frequency is NOT *zero* AND SBP status is *mild* AND age is *young*	THEN SBP reduction is 5.2,
IF intensity is *high* AND frequency is NOT *zero* AND SBP status is *mild* AND age is *middle*	THEN SBP reduction is 7.1,
IF intensity is *high* AND frequency is NOT *zero* AND SBP status is *mild* AND age is *old*	THEN SBP reduction is 4.6.

The rules for the DBP submodel for mild hypertensive Caucasian women are

IF intensity is *low/moderate* AND frequency is *three* AND DBP status is *mild* AND age is *young*	THEN DBP reduction is 22.8,
IF intensity is *low/moderate* AND frequency is *three* AND DBP status is *mild* AND age is *old*	THEN DBP reduction is 5.6,
IF intensity is *low/moderate* AND frequency is *daily* AND DBP status is *mild* AND age is *young*	THEN DBP reduction is 12.7,
IF intensity is *low/moderate* AND frequency is *daily* AND DBP status is *mild* AND age is *old*	THEN DBP reduction is 6.8,
IF intensity is *high* AND frequency is NOT *zero* AND DBP status is *mild* AND age is *young*	THEN DBP reduction is 8.2,
IF intensity is *high* AND frequency is NOT *zero* AND DBP status is *mild* AND age is *old*	THEN DBP reduction is 4.4.

Constructing the rules from the vague expert knowledge was multi-phased and tedious. We will not explain this process in detail, but rather discuss briefly the outline of it. The imprecise relationships listed in Subsection 4.1.5 were processed a bit further in order to get slightly more detailed relationships. This was done by comparing the average BP reductions reported for different subgroups to the BP response of subgroups that differed at most by one property. For instance, the relationship between age and SBP reduction could be expressed in the following more precise form: *The SBP reduction for young subjects is 0.74-fold and for old subjects 0.65-fold compared to the SBP reduction observed in middle-aged subjects.*

The resulting precise relationships gave further insight about the shapes of most of the input-output relations: BP reduction seems to be in such a non-linear relation with BP status that the higher the hypertension status is, the greater the BP reduction is; SBP reduction seems to be in a U-shaped relation with age and DBP decrease seems to be linearly related to age; Also, the relation between BP decrease and exercise frequency seems to be linear. Furthermore, the average BP reductions were derived for the consequent part of each rule by utilizing the available information on the average BP reductions observed for different subgroups, the reported percentage proportion of the different properties appearing in the subgroups and the input-output relationships.

It is important to point out that part of the rules might not be fully reliable, since the available knowledge is not extensive enough to cover all the property combinations of the input variables. For instance, very little information was available for severely hypertensive subgroups. Furthermore, there is no guarantee that the subgroups compared in the process of developing the precise relationships differed by one property only, since the subgroup characteristics were not fully provided in all the clinical trials. Especially, the information on the proportions of different properties appearing in the study populations was incomplete. Thus, part of the rules had to be based on assumptions and generalizations.

We have now covered the Step 1 of the algorithm described in Section 3.1. The other steps are covered as follows: Step 2 is settled by setting the weight for each input variable to 1, since the available knowledge does not provide any information on the mutual importance of the input variables. In case the input values produce equal degree of total fuzzy similarity for several rules, the average of the fired output values is selected as the final output. This covers Step 3 of the algorithm. Step 4 does not require any criteria to be defined, since the output subsets are always crisp values.

The inference process of the similarity based system is performed according to the Steps $A-C$ described in Section 3.2 by comparing the similarity of the properties of a person, that is the input values, to the ideal values defined by each rule. In the submodels the concept of an input object refers to a person having the properties of exercise intensity and frequency, SBP/DBP status and age. The rule that involves the properties that are most similar to the properties of that person is fired and the crisp value of SBP/DBP reduction associated with that rule is given as output. If several rules are fired, the criterion given in Step 3 is used.

4.2.2 Model Performance

Validation data were not available for the proper evaluation of the submodels' performances. Thus, the evaluations were conducted at a very basic level. The assessment process consisted of two parts. First, it was tested whether the submodels predict according to the implemented rules. This was helpful in identifying the possible implementation errors, but also in considering whether the current rules were sufficient for the submodel in question. In the second part of the assessment, the outcomes of the submodels were compared to the knowledge available on the input-output variable relationships. These two types of evaluations were performed by comparing visually the output curves constructed by the submodels to the expected curves. In the following we present these evaluated curves and discuss the resulting conclusions for the both similarity based and Sugeno type submodels.

In order to test that the submodels predict as implemented, an input data set was created including the ideal input values corresponding to the antecedent of each available rule. Recall, that an ideal value or element of a fuzzy subset, is the element that obtains the highest membership degree in the subset. In our case, the ideal value of a fuzzy subset obtains the membership degree 1. The input data set includes also the ideal input values associated with the rules that were left out from the submodels. For SBP submodels, these rules involve the fuzzy subset for moderate SBP status and for DBP submodels the rules involve the middle-aged fuzzy subset as can be concluded from Figures 4.2 and 4.3, which present the fuzzy subsets for each input variable. The outcome of a submodel generated with the input data set was then compared to the expected outcome. Here, the expected outcome refers to the output data set constructed directly from the crisp output values associated with the consequent of each rule. Unlike the input data set, the output data set depends on the submodel in question.

Figure 4.4 shows the expected and predicted outcomes of the similarity based SBP and DBP submodels of Caucasian women. Figure 4.5 presents the corresponding information for the Sugeno type inference. The predicted output values are generated with the ideal input data set and the expected outputs are data points plotted from the output data set. It seems that the outcomes of the SBP and DBP submodels are very close to the expected for the both types of models. Indeed, the figures show that the similarity and Sugeno type submodels perform very similarly with the same ideal input data set. The differences visible between the expected and predicted output values for both types of models result from the rules that were left out from the submodels. However, these differences are very small, which confirms that the missing rules (and subsets) are not essential for the performances of the submodels from the aspect of the present evaluation.

Next, we study how the submodels generate the relationships between the output variable BP reduction and each of the input variables. Figures 4.6- 4.9 show the curves of these input-output relations generated by the SBP and DBP submodels of the both model types with randomly selected reference input values. The main shapes of the curves are similar for other reference input values, too. Again, the submodels for Caucasian women are used as examples.

In general, the input-output relations generated by the submodels seem to be as expected. We know from Subsection 4.1.5 that low to moderate intensity exercise reduces blood pressure more effectively than high intensity exercise. Also, we made the logical assumption in Subsection 4.2.1 that zero

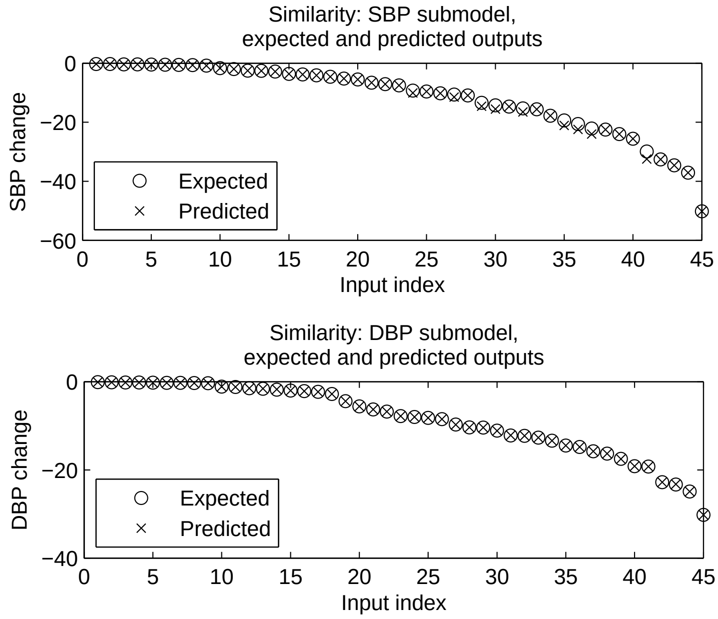

Figure 4.4: The expected and predicted outcomes of the similarity based SBP and DBP submodels of Caucasian women corresponding to the ideal input data set.

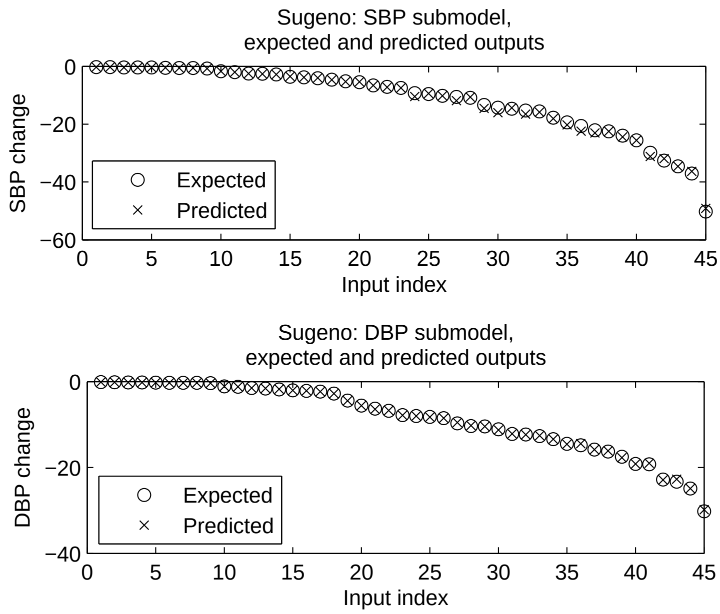

Figure 4.5: The expected and predicted outcomes of the Sugeno type SBP and DBP submodels of Caucasian women corresponding to the ideal input data set.

exercise intensity does not improve blood pressure at all. The both similarity based and Sugeno type submodels seem to follow this vague knowledge well as can be seen from Figure 4.6.

For the BP change-exercise frequency relation we expect to see a decreasing curve, since daily exercise reduces blood pressure more effectively than exercise performed 3 times/week and exercising zero times/week should not improve BP at all. The similarity based and Sugeno type submodels both fulfill this expectation as can be seen from Figure 4.7. However, as mentioned in Subsection 4.2.1, the information on the more precise relationships acquired by processing the available expert knowledge a bit further suggests that the relation should be linear. The curves created by the Sugeno submodels seem to be a bit closer to linear than the corresponding curves of the similarity based submodels. Though, the curves of both the models are still far from linearity. In any case, we cannot draw any reliable conclusions on the performance abilities of these models based on this aspect, since the assumption of linearility is not a validated fact.

The curves for BP change-BP status relations should slant down steeper as the hypertension status grows, since the higher the hypertension status is, the larger the BP decrease is. From Figure 4.8 we see that the both types of submodels generate curves that resemble this relation, at least slightly. Even the stepwise curves created by the similarity based submodels seem to get steeper as hypertension status increases, at least for SBP. The curve of the similarity based SBP submodel shows one step less than the curve of the DBP submodel, because the fuzzy subset for moderate SBP status is missing from the SBP submodel.

According to the further processed expert knowledge, SBP decrease seems to be in a U-shaped relation with age, and the relation between DBP decrease and age seems to be linear. The curves of the both type of SBP submodels in Figure 4.9 seem to represent the SBP related knowledge rather well. Furthermore, the curve of the Sugeno type DBP submodel is fairly close to linear between 30 and 70 years of age. Even though the stepwise curve generated by the similarity based DBP submodel is far from linear, it is acceptable, since it follows the unprocessed and thus the most reliable expert knowledge, which indicates that younger subjects reduce DBP more than older subjects (see Subsection 4.1.5).

The input-output relationship curves are noticeably smoother for the Sugeno type submodels than for the similarity based submodels. The reason for this follows directly from the differences in the inference process of these two types of systems. In the total fuzzy similarity based inference, usually only one rule is fired. However, for the Sugeno type systems, it is typical that several rules are fired at the same time and the final output is then

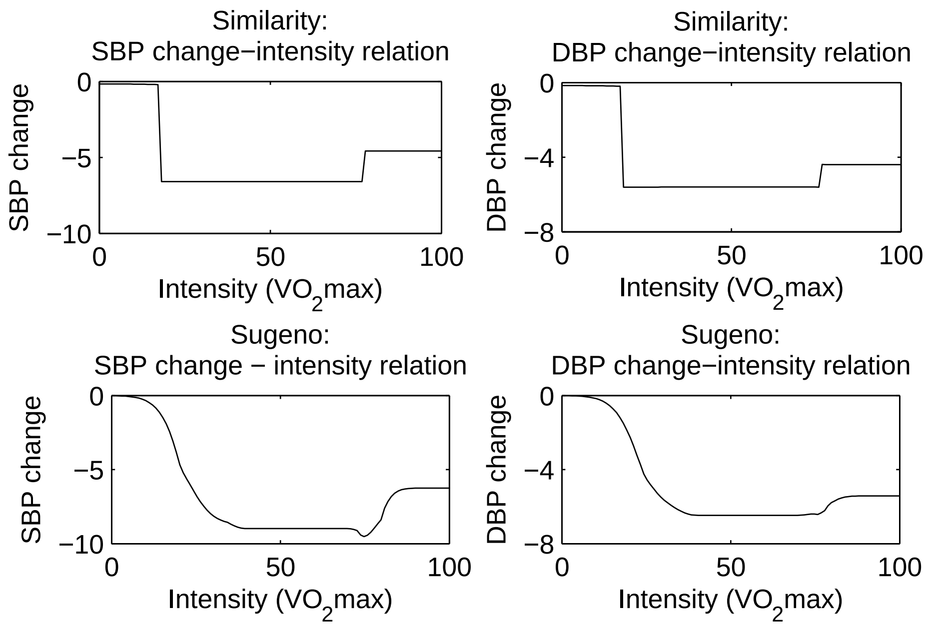

Figure 4.6: The BP change-intensity relation generated by the similarity based and Sugeno type SBP and DBP submodels of Caucasian women with reference input values frequency = 3, SBP/DBP status = 156/98, age = 75.

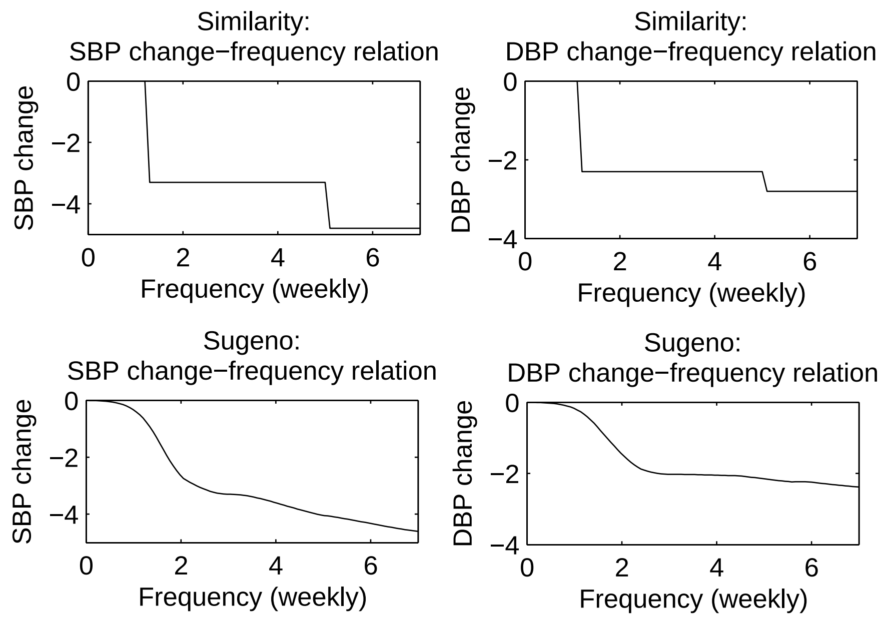

Figure 4.7: The BP change-frequency relation generated by the similarity based and Sugeno type SBP and DBP submodels of Caucasian women with reference input values intensity = 45, SBP/DBP status = 130/85, age = 40.

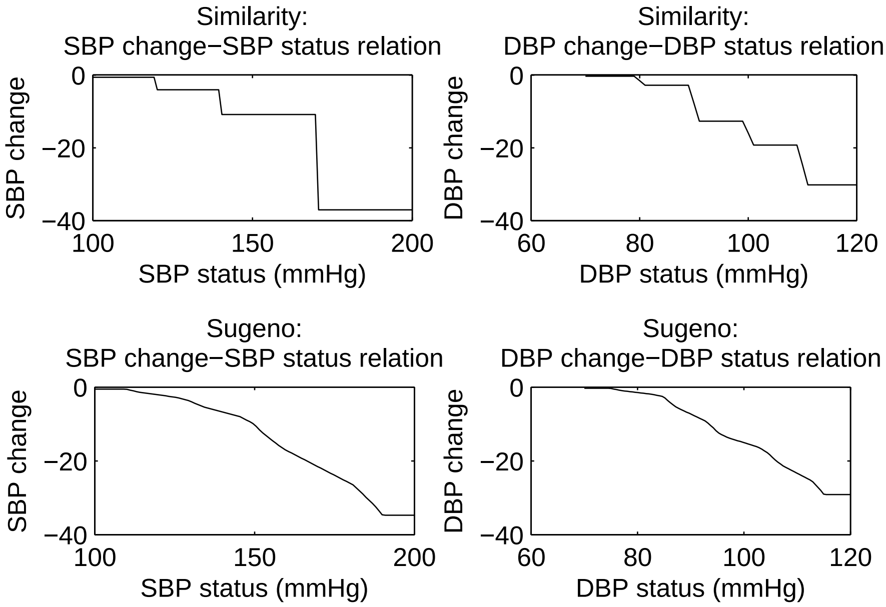

Figure 4.8: The BP change-BP status relation generated by the similarity based and Sugeno type SBP and DBP submodels of Caucasian women with reference input values intensity = 40, frequency = 6, age = 30.

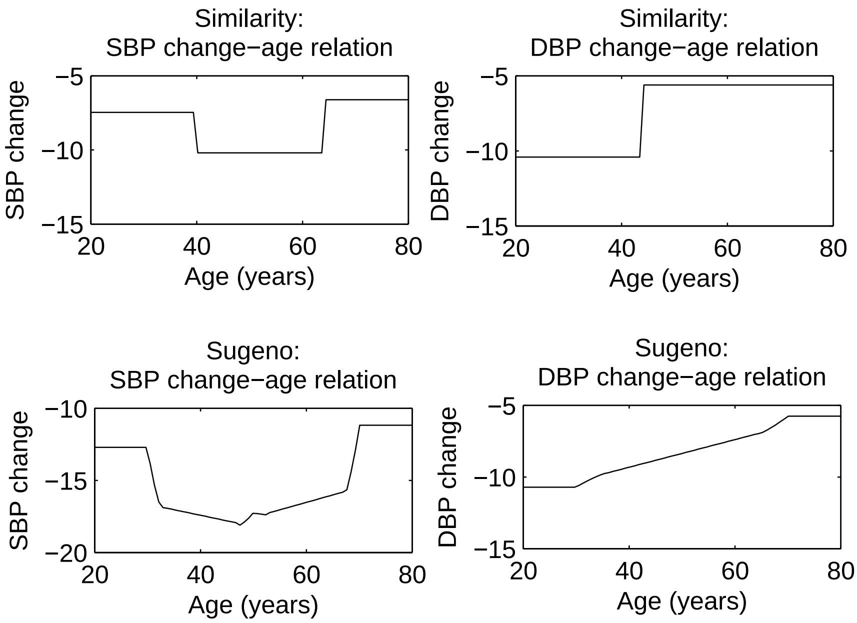

Figure 4.9: The BP change-age relation generated by the similarity based and Sugeno type SBP and DBP submodels of Caucasian women with reference input values intensity = 65, frequency = 4, SBP/DBP status = 155/95.

calculated as the weighted average of the outputs associated with the fired rules [17]. In our case, as we generated the input-output relations by the similarity based submodels, several consecutive input value sets ended up to fire the same rule. This produces the stepwise shaped curves generated by these submodels. For our Sugeno type submodels the situation is a bit different. Recall, that the structure of the two systems is identical. Thus, the both similarity based and Sugeno type submodels do end up firing the same rule for the same set of input values. The difference is that for the similarity based submodels, this rule is the only one to be fired, whereas for the Sugeno type submodels, it is one rule among other fired rules. Now, in the Sugeno type submodels the weights of the fired rules vary a bit for each consecutive set of input values. Thus, the resulting output values associated with the consecutive input value sets are slightly different from each other, which explains the smoother curves generated by the Sugeno type submodels. As we know, similarity based systems do not associate weights with the output values, though the degree of maximal total similarity might vary for different input value sets.

The same mechanism is responsible for the any other differences evident in the outputs of these two systems. For instance, the most obvious differences in the outcomes can be noticed from Figures 4.6 and 4.9, which show the SBP change-intensity and SBP change-age relationships generated by both systems, respectively. For these relationships, the Sugeno type submodels seem to predict significantly greater SBP reductions than the similarity based submodels. This is because the additional rules fired in the Sugeno SBP submodels are associated with greater SBP reductions than the rule fired commonly for the both systems, which results in the greater SBP reductions predicted by the Sugeno submodels than by the similarity based submodels. However, since we do not have actual validation data for the submodels, we cannot make any conclusions on which of the two types of systems models the phenomenon the best.

Conclusions

We have studied the theory behind total fuzzy similarity relation, presented the algorithm for constructing a total fuzzy similarity based system, discussed the inference process of the system and applied it to a real life problem. In addition, we presented the expert knowledge the described fuzzy system is based on and compared the performance of the total similarity based system to the performance of the corresponding Sugeno type system.

We attempted to find a suitable alternative representation of fuzzy conjunction in the inference rules, since the current representation involving the operation $\oplus$ did not seem intuitive in the sense that it is actually the counterpart of logical disjunction. We studied the operations $\odot$, $\cdot$ and *min* associated with the fraction $1/n$ as candidate representations of fuzzy conjunction. However, none of these candidates managed to fulfill the properties of similarity relation, but alone the operation $\odot$ did manage to fulfill these properties. Therefore, the product operation $\odot$ could be an alternative representation of fuzzy conjunction. However, this representation does not provide the opportunity to define the mutual importance of the input variables in the system. Hence, the framework provided with the structure of an injective MV-algebra would not be fully utilized. Thus, we conclude that the average of partial similarities is the most appropriate representation of fuzzy conjunction in the total similarity based system.

We discussed also the rules involving fuzzy negation and disjunction. The representation of fuzzy negation seems to be straightforward. However, for fuzzy disjunction this is not the case, since applying the total similarity approach on rules involving fuzzy disjunction seems to be inappropriate for the following reasons: the approach is not intuitive for fuzzy disjunction and it disagrees with the de Morgan's laws of classical logic, when the average of partial similarities is regarded as the counterpart of logical conjunction. We proposed *max* operation as an appropriate representation of fuzzy disjunction, since it is intuitive and performs the inference as desired. Though, it does not fulfill the properties of similarity relation. Thus, it remains to be studied whether a representation that fulfills these properties exists for fuzzy disjunction at all and if not, what this would mean for the inference process. Based on the present knowledge, we conclude that the total similarity based inference system is logically well-defined for systems that do not include rules involving fuzzy disjunction.

Finally, the comparisons we performed between the outcomes of the constructed similarity based and Sugeno type systems suggest that the two systems perform quite similarly. However, since we did not have actual validation data to test the models, the evaluations were performed at a very basic level. Thus, any reliable conclusions on the performances of these models cannot be drawn at the moment. However, the models can be studied further in the future, in the case validation data become available.

References

[1] Arroll B., Beaglehole R. *Does physical activity lower blood pressure: A critical review of the clinical trials*, J. Clin. Epidemiol. 45(5) (1992) 439-447.

[2] Cade R., Mars D., Wagemaker H. *Effect of aerobic exercise training on patients with systemic arterial hypertension*, Am. J. Med. 77(5) (1984) 785-790.

[3] Church T. S., Earnest C. P., Skinner J. S., Blair S. N. *Effects of different doses of physical activity oncardiorespiratory fitness among sedentary, overweight or obese postmenopausal women with elevated blood pressure: A randomized controlled trial*, J. Am. Med. Assoc. 297(19) (2007) 2081-2091.

[4] Cleroux J., Kouame N., Nadeau A., Coulombe D., Lacourciere Y. *Aftereffects of exercise on regional and systemic hemodynamics in hypertension*, Hypertension 19(2) (1992) 183-191.

[5] Cornelissen V. A., Fagard R. H. *Effects of endurance training on blood pressure, blood pressure-regulating mechanisms, and cardiovascular risk factors*, Hypertension 46(4) (2005) 667-675.

[6] Dubrovin T., Jolma A., Turunen E. *Fuzzy model for real-time reservoir operation*, J. Water Resour. Plng. and Mgmt. 128(1) (2002) 66-73.

[7] Eronen M. *Hilateoriaa*, 1st ed., Lecture notes of a course on Lattice Theory at the University of Tampere (in finnish), 1999.

[8] Fagard R. H. *Effects of exercise, diet and their combination on blood pressure*, J. Hum. Hypertens. 19(suppl. 3) (2005).

[9] Gottwald S. *Mathematical fuzzy logic as a tool for the treatment of vague information*, Inf. Sci. 172(1-2) (2005) 41-71.

[10] Hagberg J. M., Park J., Brown M. D. *The role of exercise training in the treatment of hypertension: An update*, Sports Med. 30(3) (2000) 193-206.

[11] Haskell W. L., Lee I., Pate R. R., Powell K. E., Blair S. N., Franklin B. A., MacEra C. A., Heath G. W., Thompson P. D., Bauman A. *Physical activity and public health: Updated recommendation for adults from the American College of Sports Medicine and the American Heart Association*, Med. Sci. Sports. Exerc. 39(8) (2007) 1423-1434.

[12] Ishikawa K., Ohta T., Zhang J., Hashimoto S., Tanaka H. *Influence of age and gender on exercise training-induced blood pressure reduction in systemic hypertension*, Am. J. Cardiol. 84(2) (1999) 192-6.

[13] Kesäniemi Y. A., Danforth E. J., Jensen M. D., Kopelman P. G., Lefebvre P., Reeder B. A. *Dose-response issues concerning physical activity and health: An evidence-based symposium*, Med. Sci. Sports. Exerc. 33(suppl. 6) (2001).

[14] Kiyonaga A., Arakawa K., Tanaka H., Shindo M. *Blood pressure and hormonal responses to aerobic exercise*, Hypertension 7(1) (1985) 125-131.

[15] Kokkinos P. F., Narayan P., Colleran J. A., Pittaras A., Notargiacomo A., Reda D., Papademetriou V. *Effects of regular exercise on blood pressure and left ventricular hypertrophy in African-American men with severe hypertension*, New Engl. J. Med. 333(22) (1995) 1462-1467.

[16] Kukkurainen P., Turunen E. *Many-valued similarity reasoning: An axiomatic approach*, Multiple-Valued Logic 8 (2002) 751-760.

[17] The MathWorks, Inc. *MATLAB Fuzzy Logic Toolbox 2, User's Guide* [WWW], The MathWorks homepage www.mathworks.com. Revised for version 2.2.9, 2009, [Referred to 1.5.2009] Available from http://www.mathworks.com/access/helpdesk/help/toolbox/fuzzy/.

[18] Nelson M. E., Rejeski W. J., Blair S. N., Duncan P. W., Judge J. O., King A. C., Macera C. A., Castaneda-Sceppa C. *Physical activity and public health in older adults: Recommendation from the American College of Sports Medicine and the American Heart Association*, Circulation 116(9) (2007) 1094-1105.

[19] Niittymäki J., Turunen E. *Traffic signal control on similarity logic reasoning*, Fuzzy Sets and Systems 133(1) (2003) 109-131.

[20] Rothenberg R. *Distribution theorems in MV-algebra*, Technical report, 2007.

[21] Seals D. R., Reiling M. J. *Effect of regular exercise on 24-hour arterial pressure in older hypertensive humans*, Hypertension 18(5) (1991) 583-592.

[22] Seals D. R., Silverman H. G., Reiling M. J., Davy K. P. *Effect of regular aerobic exercise on elevated blood pressure in postmenopausal women*, Am. J. Cardiol. 80(1) (1997) 49-55.

[23] Sivanandam S. N., Sumathi S., Deepa S. N. *Introduction to Fuzzy Logic using MATLAB*, Springer-Verlag, Heidelberg, 2007.

[24] Stewart K. J., Bacher A. C., Turner K. L., Fleg J. L., Hees P. S., Shapiro E. P., Tayback M., Ouyang P. *Effect of exercise on blood pressure in older persons: A randomized controlled trial*, Arch. Intern. Med. 165(7) (2005) 756-762.

[25] Turunen E. *Mathematics Behind Fuzzy Logic, Advances in Soft Computing*. Physica-Verlag, Heidelberg, 1999.

[26] Urata H., Tanabe Y., Kiyonaga A. *Antihypertensive and volume-depleting effects of mild exercise on essential hypertension*, Hypertension 9(3) (1987) 245-252.

[27] Whelton S. P., Chin A., Xin X., He J. *Effect of aerobic exercise on blood pressure: A meta-analysis of randomized, controlled trials*, Ann. Intern. Med. 136(7) (2002) 493-503.