FLEXIBLE NEURO-FUZZY SYSTEMS

STRUCTURES, LEARNING AND PERFORMANCE EVALUATION

Leszek Rutkowski

Kluwer Academic Publishers

FLEXIBLE NEURO-FUZZY SYSTEMS

Structures, Learning and Performance Evaluation

CRACKED TRADING SOFTWARE

70+ DVD's FOR SALE & EXCHANGE

www.traders-software.com

www.forex-warez.com

www.trading-software-collection.com

www.tradestation-download-free.com

Contacts

andreybbrv@gmail.com
andreybbrv@yandex.ru
Skype: andreybbrv

THE KLUWER INTERNATIONAL SERIES IN ENGINEERING AND COMPUTER SCIENCE

FLEXIBLE NEURO-FUZZY SYSTEMS

Structures, Learning and Performance Evaluation

by

Leszek Rutkowski
Technical University of Czestochowa
Poland

KLUWER ACADEMIC PUBLISHERS
NEW YORK, BOSTON, DORDRECHT, LONDON, MOSCOW

eBook ISBN: 1-4020-8043-3
Print ISBN: 1-4020-8042-5

Created in the United States of America

Visit Kluwer Online at: http://kluweronline.com
and Kluwer's eBookstore at: http://ebooks.kluweronline.com

This book is dedicated to
Professor Lotfi Zadeh

This page intentionally left blank

Contents

This page intentionally left blank

Foreword

To write a foreword to Professor Rutkowski's opus "*Flexible Neuro-Fuzzy Systems,*" or FNFS for short, was a challenging task. Today, there exists an extensive literature on neuro-fuzzy systems, but Professor Rutkowski's work goes far beyond what is in print. FNFS ventures into new territory and opens the door to new directions in research and new application areas.

First, a bit of history. The concept of a neuro-fuzzy system is rooted in the pioneering work of H. Takagi and I. Hayashi, who in 1988 obtained a basic patent in Japan, assigned to Matsushita, describing a system in which techniques drawn from fuzzy logic and neural networks were used in combination to obtain superior performance. The basic idea underlying their patent was to exploit the learning capability of neural networks for enhancing the performance of fuzzy rule-based systems. Today, neuro-fuzzy systems are employed in most of the consumer products manufactured in Japan.

In the years which followed, the concept of a neuro-fuzzy system was broadened in various ways. In particular, a basic idea pioneered by Arabshahi et al was to start with a neuro-based algorithm such as the backpropagation algorithm, and improve its performance by employing fuzzy if-then rules for adaptive adjustment of parameters. What should be noted is that the basic idea underlying this approach is applicable to any type of algorithm in which human expertise plays an essential role in choosing parameter-values and controlling their variation as a function of performance. In such applications, fuzzy if-then rules are employed as a language for describing human expertise.

Another important direction which emerged in the early nineties was rooted in the realization that a fuzzy rule-based system could be viewed as a multilayer network in which the nodes are (a) the antecedents and consequents of fuzzy if-then rules; and (b) the conjunctive and disjunctive connectives. The membership functions of antecedents and consequents are assumed to be triangular or trapezoidal. The problem is to optimize the values of parameters of such membership function through minimization of mean-squared error, as in the backpropagation algorithm. The problem is solved through the use of gradient techniques which are very similar to those associated with backpropagation. It is this similarity that underlies the use of the label "*neuro-fuzzy,*" in describing systems of this type. A prominent example is the ANFIS system developed by Roger Jaing, a student of mine who conceived ANFIS as a part of his doctoral dissertation at UC Berkeley.

Neuro-fuzzy systems, which are the focus of attention in Professor Rutkowski's work, are, basically, in the ANFIS spirit. There is, however, an important difference. In Professor Rutkowski's systems, the connectives and everything else are flexible in the sense that they have a variable structure, which is adjusted in the course of training. The flexibility of Professor Rutkowski's systems, call then FNFS's, has the potential for a major improvement in performance compared to that of neuro-fuzzy systems with a fixed structure.

In another important departure from convention, Professor Rutkowski employs weighted t-norms and t-conorms instead of the simple *"and"* and *"or"* connectives used in existing neuro-fuzzy systems. Flexible use of such connectives has an important bearing on performance. Throughout the book, Professor Rutkowski's analysis is conducted at a high level of mathematical sophistication and in great detail. Extensive computer simulation is employed to verify results of analysis.

An issue that receives a great deal of attention relates to the use of what is commonly referred to as Mamdani-type reasoning vs. logical reasoning. In what follows, I should like to comment on this issue since it is a source of a great deal of misunderstanding and confusion.

The crux of the issue relates to interpretation of the proposition *"if X is A the Y is B,"* where X and Y are linguistic variables, and A and B are the linguistic values of X and Y, respectively. The source of confusion is that *"if X is A then Y is B,"* can be interpreted in two different ways. The first, and simpler way, is to interpret *"if X is A then Y is B,"* as *"X is A and Y is B"* or, equivalently, as (X,Y) is A×B, where A×B is the Cartesian product of A and B. Thus, in this interpretation, *"if X is A then Y is B"* is a joint constraint on X

and Y. A source of confusion is that Mamdani and Assilian used this interpretation in their seminal 1974 paper, but referred to it as implication, which it is not, rather than as a joint constraint.

An alternative way is to interpret *"if X is A then Y is B,"* as a conditional constraint or, equivalently, as an implication, with the understanding that there are many ways in which implication may be defined. What should be noted is that, generally, we are concerned with interpretation of a collection of fuzzy if-then rules, that is, a rule set, rather than an isolated rule. When *"if X is A then Y is B,"* is interpreted as a joint constraint, the concomitant interpretation of the rule set is the disjunction of interpretations of its constituent rules, leading to the concept of a fuzzy graph, described in my 1974 paper *"On the Analysis of Large Scale Systems,"* Systems Approaches and Environment Problems, H. Gottinger (ed.), 23-37, Gottingen: Vandenhoeck and Ruprecht. Alternatively, when the conditional constraint interpretation is used, interpretations of constituent rules are combined conjunctively.

When response to a given input is sought, the joint constraint interpretation is distributive, while the conditional constraint interpretation, is not. Simplicity resulting from distributivity is the principal reason why Mamdani's approach, which is based on the joint constraint interpretation, is in preponderant use in applications. A more detailed discussion may be found in my paper, *"Fuzzy logic and the calculi of fuzzy rules and fuzzy graphs"* Multiple-Valued Logic 1, 1-38, 1996. An important concept within Professor Rutkowski's theory is that of flexible compromise neuro-fuzzy systems. In such systems, simultaneous appearance of Mamdani-type and logical-type reasoning is allowed.

To say that Professor Rutkowski's work is a major contribution to the theory and application of neuro-fuzzy systems is an understatement. The wealth of new ideas, the thoroughness of analysis, the attention to detail, the use of computer simulation, the problems at the end of each chapter, and high expository skill, combine to make Professor Rutkowski's work a must reading for anyone interested in the conception, design and utilization of intelligent systems. Professor Rutkowski and the publisher, Kluwer, deserve a loud applause.

Lotfi A. Zadeh
December 22, 2003
UC Berkeley

This page intentionally left blank

Chapter 1

INTRODUCTION

Over the last decade fuzzy sets and fuzzy logic introduced in 1965 by Lotfi Zadeh [113] have been used in a wide range of problem domains including process control, image processing, pattern recognition and classification, management, economics and decision making. Specific applications include washing-machine automation, camcorder focusing, TV colour tuning, automobile transmissions and subway operations [29]. We have also been witnessing a rapid development in the area of neural networks (see e.g. [93, 94, 135]). Both fuzzy systems and neural networks, along with probabilistic methods [1, 20, 67], evolutionary algorithms [23, 59], rough sets [69, 70] and uncertain variables [6, 7, 8], constitute a consortium of soft computing techniques [1, 39, 42]. These techniques are often used in combination. For example, fuzzy inference systems are frequently converted into connectionist structures called neuro-fuzzy systems which exhibit advantages of neural networks and fuzzy systems. In literature various neuro fuzzy systems have been developed (see e.g. [11, 12, 13, 15, 18, 24, 25, 26, 34, 35, 37, 40, 49, 50, 52, 53, 54, 55, 60, 61, 65, 72, 75, 76, 80-85, 100]). They combine the natural language description of fuzzy systems and the learning properties of neural networks. Some of them are known in literature under short names such as ANFIS [33], ANNBFIS [15], DENFIS [41], FALCON [51], GARIC [3], NEFCLASS [62], NEFPROX [62, 64], SANFIS [99] and others. The most popular designs of neuro-fuzzy structures fall into one of the following categories, depending on the connective between the antecedent and the consequent in fuzzy rules:

(i) Takagi-Sugeno method – consequents are functions of inputs,

(ii) Mamdani-type reasoning method – consequents and antecedents are related by the min operator or generally by a t-norm,
(iii) Logical-type reasoning method - consequents and antecedents are related by fuzzy implications, e.g. binary, Łukasiewicz, Zadeh and others (see Chapter 2).

These models are illustrated in Fig. 1.1.

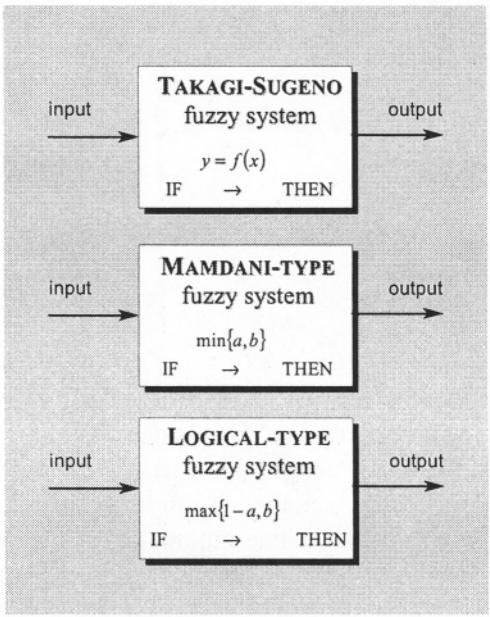

Fig. 1.1. Examples of typical fuzzy systems

It should be noted that most applications are dominated by the Mamdani-type fuzzy reasoning. Moreover, in a specific singleton case, the Takagi-Sugeno method is reduced to the Mamdani method. On the other hand, there is a widely held belief about the inferiority of the logical method comparing with the Mamdani method. However, it was emphasized by Yager [106, 107] that "*no formal reason exists for the preponderant use of the Mamdani method in fuzzy logic control as opposed to the logical method other than inertia.*" Moreover, Yager said [108] that "*as a matter of fact the Mamdani approach has some disadvantages: its inability to distinguish more specific information in the face of vague information and the requirement of having the antecedents of the rules span the whole input space.*" This statement was an inspiration for the author to determine the type of fuzzy inference (Mamdani or logical) in the process of learning. In Fig. 1.2 we illustrate the process of learning of fuzzy inference.

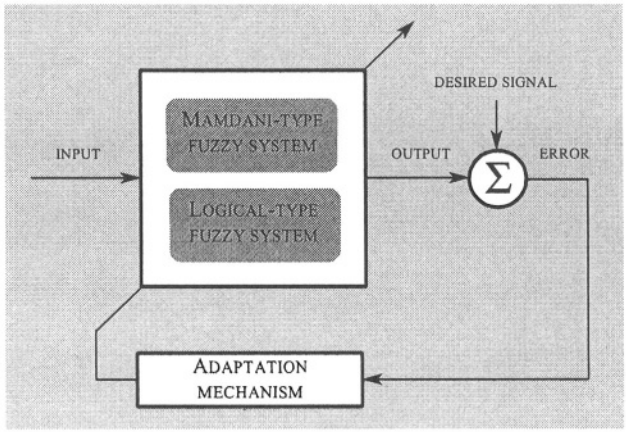

Fig. 1.2. Adaptation of fuzzy reasoning to data

To our best knowledge such a concept has not yet been proposed in literature by other authors. Neuro-fuzzy systems have been treated separately (Mamdani-type or logical-type) in a number of books and papers, see e.g. [15, 28, 76]. A type of the system has never been determined in the process of learning. We solve that problem in two different ways:

a) We propose (see Chapter 5) a new class of neuro-fuzzy systems characterized by automatic determination of a fuzzy inference (Mamdani / logical) in the process of learning. Consequently, the structure of the system is determined in the process of learning. This class is based on the definition of an H-function which becomes a t-norm or t-conorm depending on a certain parameter ν which can be found in the process of learning. We refer to this class as to OR-type fuzzy systems.

b) We develop (see Chapter 6) AND-type neuro-fuzzy inference systems by making use of the concept of flexible structures studied by Yager and Filev [108]. The AND-type fuzzy inference systems exhibit simultaneously Mamdani and logical type inferences.

It is well know that introducing additional parameters to be tuned in neuro-fuzzy systems improves their performance and they are able to better represent the patterns encoded in the data. Therefore, in this book, in addition to automatic determination of a system type we introduce several flexibility concepts in the design of neuro-fuzzy systems, in particular we introduce:

- softness to fuzzy implication operators, to the aggregation of rules and to the connectives of antecedents,
- certainty weights to the aggregation of rules and to the connectives of antecedents,

- parameterized families of t-norms and t-conorms to fuzzy implication operators, to the aggregation of rules and to the connectives of antecedents

in both AND-type and OR-type neuro-fuzzy inference systems.

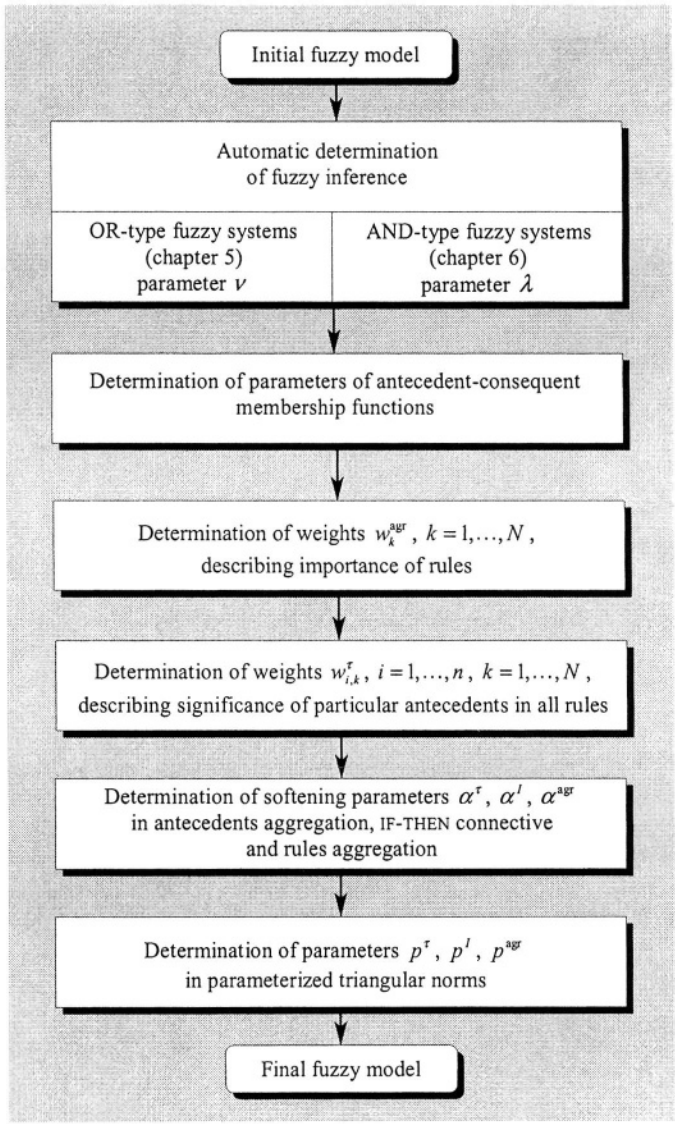

Fig. 1.3. Design of flexible fuzzy system

In Fig. 1.3 we show a design process of flexible neuro-fuzzy systems developed in this book. In the book, we study a wide class of fuzzy systems trained by the backpropagation method. Following other authors we call

them neuro fuzzy inference systems (NFIS). To emphasize their main feature flexibility, we also use the name FLEXNFIS.

The book consists of nine chapters. Chapter 2 provides an introduction to the theory of fuzzy sets. In Chapter 3 we describe fuzzy systems with the Mamdani-type inference and logical-type inference. Moreover, we introduce the concept of the generalized neuro-fuzzy system. Chapter 4 concentrates on incorporating various flexibility parameters into the design of neuro-fuzzy systems. In Chapter 5 we develop the concept of adjustable quasi-triangular norms and quasi-implications. Based on these concepts, flexible OR-type neuro-fuzzy systems are presented. Chapter 6 deals with flexible compromise neuro-fuzzy systems. They are characterized by the simultaneous appearance of Mamdani-type and logical-type reasoning. In Chapters 7 and 8 we fix the system type (Mamdani and logical, respectively) and present explicitly connectionist structures of the Mamdani-type and logical-type systems. In Chapter 9 we compare all flexible neuro-fuzzy structures studied in the book and give conclusions and directions for future research. The material presented in the book is illustrated by many computer simulations. Through these simulations we show that Mamdani-type systems are more suitable for approximation problems, whereas logical-type systems may be preferred for classification problems. Moreover, we observe that the most influential parameters in FLEXNFIS are certainty weights (introduced in this paper in a novel way) in connectives of antecedents and in aggregations of rules. They significantly improve the performance of NFIS in the process of learning.

It should be noted that the book provides a framework for the unification, construction and development of neuro-fuzzy systems. It presents complete algorithms in a systematic and structured fashion, easing the understanding and implementation. Besides, the book contains numerous examples and exercises following each chapter. Another strength of the book is that it provides tools for possible applications in business and economics, medicine and bioengineering, automatic control, robotics, decision theory and expert systems.

The author thanks Professor Lotfi Zadeh for his Foreword. The author gratefully acknowledges the material quoted from his previous works published by IEEE and Springer-Verlag. For a complete listing of quoted articles the reader is referred to the "References". The author also gratefully acknowledges Krzysztof Cpałka and Robert Nowicki, his former Ph. D. students, for the material being a part of our joint research and included in the book. I also want to express appreciations to my editor Susan Lagerstrom-Fife, assistant editor Sharon Palleschi and other staff members at Kluwer Academic Publishers.

This page intentionally left blank

Chapter 2

ELEMENTS OF THE THEORY OF FUZZY SETS

2.1. INTRODUCTION

Traditional knowledge representation is based on bivalent logic. However, human thinking and behaviour is strictly connected with imprecision and uncertainty. The traditional Boolean algebra, characterized by categorical values of truth and falsehood, is not able to cope with such problems. Fuzzy logic is an extension of the traditional logic to intermediate and approximate values. The concept of fuzzy logic was proposed and developed by Zadeh [113-134]. In this chapter we recall the basic definitions and properties of fuzzy sets and fuzzy reasoning which will be useful in the next chapters.

2.2. BASIC DEFINITIONS

Definition 2.1.
A fuzzy set A defined in space $\mathbf{X}$ is a set of pairs:

$$A = \{(\mathbf{x}, \mu_A(\mathbf{x})), \mathbf{x} \in \mathbf{X}\}, \ \forall \mathbf{x} \in \mathbf{X} \tag{2.1}$$

where $\mu_A : \mathbf{X} \to [0,1]$ is a membership function of the fuzzy set A, which for every element $\mathbf{x} \in \mathbf{X}$ assigns its membership degree $\mu_A(\mathbf{x}) \in [0,1]$ to the fuzzy set A. The set $\mathbf{X}$ is called a universe of discourse and we write $A \subseteq \mathbf{X}$.

The fuzzy set A is completely determined by the set of pairs (2.1). When the universe of discourse is discrete and finite with cardinality n, that is $\mathbf{X} = \{x_1,...,x_n\}$, the fuzzy set A can be represented as

$$A = \sum_{i=1}^{n} \mu_A(x_i)/x_i = \mu_A(x_1)/x_1 + \cdots + \mu_A(x_n)/x_n \tag{2.2}$$

or equivalently

$$A = \sum_{i=1}^{n} \frac{\mu_A(x_i)}{x_i} = \frac{\mu_A(x_1)}{x_1} + \cdots + \frac{\mu_A(x_n)}{x_n} \tag{2.3}$$

When the universe of discourse $\mathbf{X}$ is an interval of real numbers, the fuzzy set A can be expressed as

$$A = \int_{\mathbf{X}} \mu_A(x)/x \tag{2.4}$$

or

$$A = \int_{\mathbf{X}} \frac{\mu_A(x)}{x} \tag{2.5}$$

respectively. Symbols $\sum, +, \int$ in formulas (2.2)-(2.5) refer to the set union, not to the arithmetic summation. Similarly, there is no arithmetic division in these formulas. This symbolic notation is used in order to connect an element and its membership value. The same notation can be applied in the multidimensional case.

Example 2.1.
Let $\mathbf{X}$ = {4,5,6,7,8,9,10,11}. We will define a set of natural numbers "close to 8" by making use of the following fuzzy set

$$A = \frac{0.3}{4} + \frac{0.5}{5} + \frac{0.7}{6} + \frac{0.85}{7} + \frac{1}{8} + \frac{0.85}{9} + \frac{0.7}{10} + \frac{0.5}{11} \tag{2.6}$$

We will now present frequently used membership functions defined in $\mathbf{X} \subset \mathbf{R}$.

a) Singleton Membership Function

Singleton is a membership function taking value 1 only in one point $\overline{\mathbf{x}}$ of the universe of discourse $\mathbf{X}$ and 0 otherwise

$$\mu_{\text{Singl}}(\mathbf{x}) = \begin{cases} 1 & \text{for} \quad \mathbf{x} = \overline{\mathbf{x}} \\ 0 & \text{for} \quad \mathbf{x} = \overline{\mathbf{x}} \end{cases}. \tag{2.7}$$

b) Gaussian Membership Function

A Gaussian membership function has two parameters: $\bar{x}$ -responsible for its center and σ -responsible for its width

$$\mu_{\text{Gauss}}(x) = \exp\left(-\left(\frac{x-\bar{x}}{\sigma}\right)^2\right). \tag{2.8}$$

c) Generalized Bell Membership Function

A generalized bell membership function has three parameters: a -responsible for its width, c -responsible for its center and b -responsible for its slopes

$$\mu_{\text{Bell}}(x) = \frac{1}{1+\left|\frac{x-c}{a}\right|^{2b}}. \tag{2.9}$$

d) Sigmoidal Membership Function

A sigmoidal membership function has two parameters: a -responsible for its slope at the crossover point $x = c$

$$\mu_{\text{Sigm}}(x) = \frac{1}{1+\exp[-a(x-c)]}. \tag{2.10}$$

e) Triangular Membership Function

A triangular membership function is fully described by three parameters $\{a,b,c\}$, $a < b < c$

$$\mu_{\text{Triangle}}(x) = \begin{cases} 0, & x \le a \\ \frac{x-a}{b-a}, & a \le x \le b \\ \frac{c-x}{c-b}, & b \le x \le c \\ 0, & c \le x \end{cases}, \tag{2.11}$$

or by an alternative formula to Equation (2.11)

$$\mu_{\text{Triangle}} = (x) = \max\left(\min\left(\frac{x-a}{b-a}, \frac{c-a}{c-b}\right), 0\right) \tag{2.12}$$

f) Trapezoidal Membership Function

A trapezoidal membership function is fully described by three parameters $\{a,b,c,d\}$, $a<b<c<d$

$$\mu_{\text{Trap}}(x)=\begin{cases}0, & x\le a\\ \dfrac{x-a}{b-a}, & a\le x\le b\\ 1, & b\le x\le c\\ \dfrac{d-x}{d-c}, & c\le x\le d\\ 0, & d\le x\end{cases}, \tag{2.13}$$

or by an alternative formula to Equation (2.13)

$$\mu_{\text{Trap}}(x)=\max\left(\min\left(\frac{x-a}{b-a},1,\frac{d-x}{d-c}\right),0\right). \tag{2.14}$$

Typical membership functions are depicted in Fig. 2.1

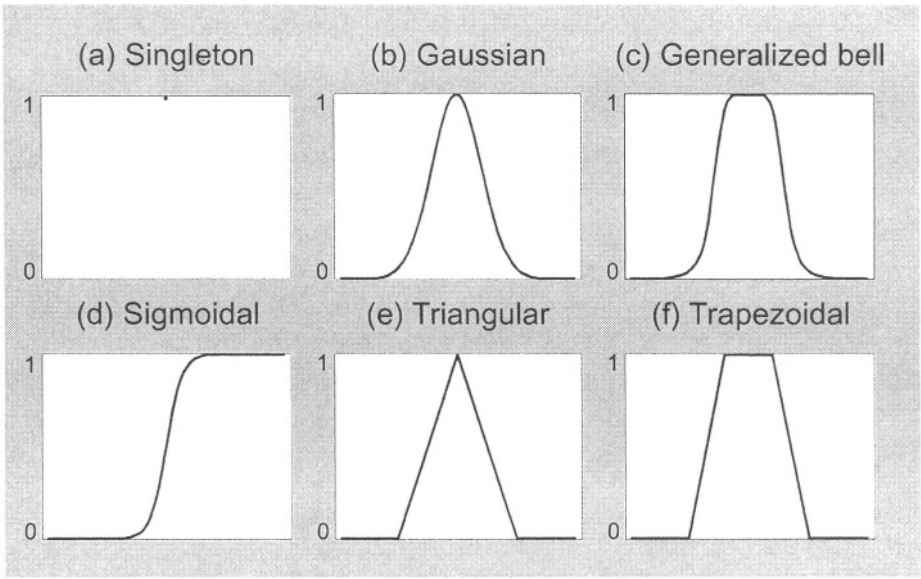

Fig. 2.1. Examples of membership functions

Definition 2.2.
The height of a fuzzy set A is a maximal value of its membership function

$$\mathrm{hgt}(A) = \sup_{\mathbf{x} \in \mathbf{X}} \mu_A(\mathbf{x}). \tag{2.15}$$

If $\mathrm{hgt}(A) = 1$, then A is called a normal fuzzy set, otherwise it is subnormal.

Example 2.2.
If $\mathbf{X} = \{2,3,4,5\}$ and

$$A = \frac{0.4}{2} + \frac{0.9}{4} + \frac{0.1}{5} \tag{2.16}$$

then $\mathrm{hgt}(A) = 0.9$.

Definition 2.3.
The support of a fuzzy set A is a set of all points $\mathbf{x}$ in $\mathbf{X}$ such that $\mu_A(\mathbf{x}) > 0$, i.e.

$$\mathrm{supp}(A) = \{\mathbf{x} \in \mathbf{X} \mid \mu_A(\mathbf{x}) > 0\} \tag{2.17}$$

Example 2.3.
If $\mathbf{X} = \{1,2,3,4,5\}$ and

$$A = \frac{0.3}{1} + \frac{0.4}{3} + \frac{0.6}{4} \tag{2.18}$$

then $\mathrm{supp}(A) = \{1,3,4\}$.

Definition 2.4.
The core of a fuzzy set A is a set of all points $\mathbf{x}$ in $\mathbf{X}$ such that $\mu_A(\mathbf{x}) = 1$

$$\mathrm{core}(A) = \{\mathbf{x} \in \mathbf{X} \mid \mu_A(\mathbf{x}) = 1\} \tag{2.19}$$

In other words, we can say that a fuzzy set is normal if its core is nonempty.

Example 2.4.
If $\mathbf{X} = \{1,2,3\}$ and

$$A = \frac{0.2}{1} + \frac{1}{2} + \frac{0.3}{3} \tag{2.20}$$

then $\mathrm{core}(A) = 2$.

Definition 2.5.
A fuzzy set A is convex if and only if for any $\mathbf{x}_1, \mathbf{x}_2 \in \mathbf{X}^n$ and any $\lambda \in [0,1]$

$$\mu_A(\lambda \mathbf{x}_1 + (1-\lambda)\mathbf{x}_2) \geq \min\{\mu_A(\mathbf{x}_1), \mu_A(\mathbf{x}_2)\} \tag{2.21}$$

Definition 2.6.
A fuzzy set A is a fuzzy number when $\mathbf{X} \subset \mathbf{R}^n$ and A is normal and convex.

Definition 2.7.
A fuzzy set A is symmetric if its membership function is symmetric around a certain point c

$$\mu_A(c+\Delta\mathbf{x}) = \mu_A(c-\Delta\mathbf{x}) \tag{2.22}$$

with $c+\Delta\mathbf{x} \in \mathbf{X}$ and $c-\Delta\mathbf{x} \in \mathbf{X}$.

Definition 2.8.
The cardinality of a fuzzy set A is defined as

$$|A| = \int_{\mathbf{x}} \mu_A(\mathbf{x})\, d\mathbf{x} \tag{2.23}$$

and in case of a discrete universe the integration is replaced by summation.

Definition 2.9.
The concentration of a fuzzy set A in $\mathbf{X}$, denoted by CON(A), is defined by

$$\mu_{\mathrm{CON}(A)}(\mathbf{x}) = (\mu_A(\mathbf{x}))^2 \tag{2.24}$$

for all $\mathbf{x} \in \mathbf{X}$.

Example 2.5.
If $\mathbf{X} = \{1,2,3,4\}$ and

$$A = \frac{0.3}{1} + \frac{0.6}{2} + \frac{1}{4} \tag{2.25}$$

then $\mathrm{CON}(A) = \dfrac{0.09}{1} + \dfrac{0.36}{2} + \dfrac{1}{4}$.

Definition 2.10.
The dilation of a fuzzy set A in $\mathbf{X}$, denoted by DIL(A), is defined by

$$\mu_{\mathrm{DIL}(A)}(\mathbf{x}) = (\mu_A(\mathbf{x}))^{0.5} \tag{2.26}$$

for all $x \in \mathbf{X}$.

Example 2.6.
If $\mathbf{X} = \{1,2,3\}$ and

$$A = \frac{0.16}{1} + \frac{1}{2} + \frac{0.64}{3} \tag{2.27}$$

then $\mu_{\mathrm{DIL}(A)}(\mathbf{x}) = \frac{0.4}{1} + \frac{1}{2} + \frac{0.8}{3}$.

Definition 2.11.
If A is a fuzzy set in $\mathbf{X}$, then its cylindrical extension in $\mathbf{X}\times\mathbf{Y}$ is a fuzzy set $ce(A, \mathbf{X}\times\mathbf{Y})$ defined by

$$ce(A, \mathbf{X}\times\mathbf{Y}) = \int_{\mathbf{X}\times\mathbf{Y}} \frac{\mu_A(\mathbf{x})}{(\mathbf{x}, \mathbf{y})} \tag{2.28}$$

2.3. TRIANGULAR NORMS AND NEGATIONS

Definition 2.12.
A t-norm is a function T of two variables

$$T{:}[0,1]\times[0,1] \to [0,1], \tag{2.29}$$

satisfying the following conditions:
1) T is monotonic

$$T\{a,c\} \le T\{b,d\} \text{ whenever } a \le b \text{ and } c \le d, \tag{2.30}$$

2) T is commutative

$$T\{a,b\} = T\{b,a\}, \tag{2.31}$$

3) T is associative

$$T\{T\{a,b\},c\} = T\{a,T\{b,c\}\}, \tag{2.32}$$

4) T satisfies boundary conditions

$$T\{a,0\} = 0,\ T\{a,1\} = a, \tag{2.33}$$

where $a,b,c,d \in [0,1]$.

Further, a t-norm on arguments a and b will be denoted by

$$T\{a,b\} = a \overset{T}{*} b. \tag{2.34}$$

The associativity condition allows to extend Definition 2.12 to $n > 2$ arguments, i.e.

$$\overset{n}{\underset{i=1}{T}}\{a_i\} = T\left\{\overset{n-1}{\underset{i=1}{T}}\{a_i\}, a_n\right\} = T\{a_1, a_2, \ldots, a_n\} = T\{\mathbf{a}\} = a_1 \overset{T}{*} a_2 \overset{T}{*} \ldots \overset{T}{*} a_n \tag{2.35}$$

Definition 2.13.
A t-conorm is a function S of two variables

$$S{:}[0,1]\times[0,1] \to [0,1], \tag{2.36}$$

that satisfies the following conditions:

1) S is monotonic

$$S\{a,c\} \le S\{b,d\} \text{ whenever } a \le b \text{ and } c \le d, \tag{2.37}$$

2) S is commutative

$$S\{a,b\} = S\{b,a\}, \tag{2.38}$$

3) S is associative

$$S\{S\{a,b\},c\} = S\{a,S\{b,c\}\}, \tag{2.39}$$

4) S satisfies boundary conditions

$$S\{a,0\} = a, \; S\{a,1\} = 1, \tag{2.40}$$

where $a,b,c,d \in [0,1]$.

Further, a t-conorm on arguments a and b will be denoted by

$$S\{a,b\} = a \overset{S}{*} b. \tag{2.41}$$

The associativity condition allows to extend Definition 2.13 to $n > 2$ arguments, i.e.

$$\overset{n}{\underset{i=1}{S}}\{a_i\} = S\left\{\overset{n-1}{\underset{i=1}{S}}\{a_i\}, a_n\right\} = S\{a_1,a_2,\ldots,a_n\} = S\{\mathbf{a}\} = a_1 \overset{S}{*} a_2 \overset{S}{*} \ldots \overset{S}{*} a_n \tag{2.42}$$

Example 2.7.

The min/max triangular norms are defined as follows:

$$T_M\{a_1,a_2\} = \min\{a_1,a_2\} \tag{2.43}$$

$$S_M\{a_1,a_2\} = \max\{a_1,a_2\} \tag{2.44}$$

$$T_M\{a_1,a_2,\ldots,a_n\} = \min_{i=1,\ldots,n}\{a_i\} \tag{2.45}$$

$$S_M\{a_1,a_2,\ldots,a_n\} = \max_{i=1,\ldots,n}\{a_i\} \tag{2.46}$$

We will also refer to T_M and S_M as to Zadeh's triangular norms.

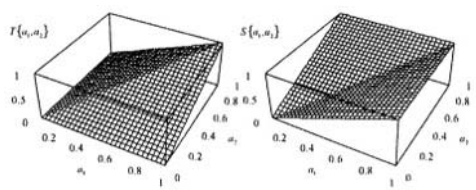

Fig. 2.2. Min/max triangular norms

The min/max triangular norms are depicted in Fig. 2.2.

Example 2.8.
The algebraic triangular norms are defined as follows:

$$T_P\{a_1,a_2\} = a_1 a_2 \tag{2.47}$$

$$S_P\{a_1,a_2\} = a_1 + a_2 - a_1 a_2 \tag{2.48}$$

$$T_P\{a_1,a_2,\ldots,a_n\} = \prod_{i=1}^{n} a_i \tag{2.49}$$

$$S_P\{a_1,a_2,\ldots,a_n\} = 1 - \prod_{i=1}^{n} (1 - a_i) \tag{2.50}$$

The algebraic triangular norms are depicted in Fig. 2.3

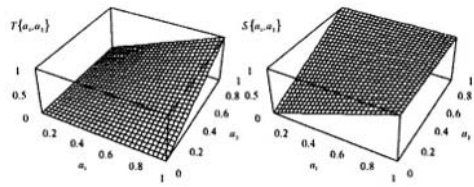

Fig. 2.3. Algebraic triangular norms

The t-norm (2.47) is also known under the name of a product t-norm (or algebraic product), whereas the t-conorm (2.48) is known under the name of a probabilistic sum (or algebraic sum).

Example 2.9.
The bounded triangular norms are defined as follows:

$$T_L\{a_1,a_2\} = \max\{a_1 + a_2 - 1, 0\} \tag{2.51}$$

$$S_L\{a_1,a_2\} = \min\{a_1 + a_2, 1\} \tag{2.52}$$

$$T_L\{a_1,a_2,\ldots,a_n\} = \max\left\{\sum_{i=1}^{n} a_i - (n-1), 0\right\} \tag{2.53}$$

$$S_L\{a_1,a_2,\ldots,a_n\} = \min\left\{\sum_{i=1}^{n} a_i, 1\right\} \tag{2.54}$$

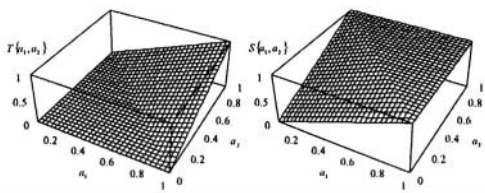

Fig. 2.4. Bounded triangular norms

The t-norm (2.51) and (2.52) t-conorm are also known under the name of the Łukasiewicz triangular norms.

Example 2.10.
The drastic triangular norms are defined as follows:

$$T_D\{a_1,a_2\}=\begin{cases}0 & \text{if} \quad S_M\{a_1,a_2\}<1\\ T_M\{a_1,a_2\} & \text{if} \quad S_M\{a_1,a_2\}=1\end{cases} \tag{2.55}$$

$$S_D\{a_1,a_2\}=\begin{cases}1 & \text{if} \quad T_M\{a_1,a_2\}>0\\ S_M\{a_1,a_2\} & \text{if} \quad T_M\{a_1,a_2\}=0\end{cases} \tag{2.56}$$

$$T_D\{a_1,a_2,\ldots,a_n\}=\begin{cases}0 & \text{if} \quad S_M\{a_1,a_2,\ldots,a_n\}<1\\ T_M\{a_1,a_2,\ldots,a_n\} & \text{if} \quad S_M\{a_1,a_2,\ldots,a_n\}=1\end{cases} \tag{2.57}$$

$$S_D\{a_1,a_2,\ldots,a_n\}=\begin{cases}1 & \text{if} \quad T_M\{a_1,a_2,\ldots,a_n\}>0\\ S_M\{a_1,a_2,\ldots,a_n\} & \text{if} \quad T_M\{a_1,a_2,\ldots,a_n\}=0\end{cases} \tag{2.58}$$

The drastic triangular norms are depicted in Fig. 2.5. Observe that formulas (2.55) and (2.56) can be alternatively expressed by:

$$T_D\{a_1,a_2\}=\begin{cases}a_1 & \text{if} \quad a_2=1\\ a_2 & \text{if} \quad a_1=1\\ 0 & \text{if} \quad a_1,a_2\neq 1\end{cases} \tag{2.59}$$

$$S_D\{a_1,a_2\}=\begin{cases}a_1 & \text{if} \quad a_2=0\\ a_2 & \text{if} \quad a_1=0\\ 0 & \text{if} \quad a_1,a_2\neq 0\end{cases} \tag{2.60}$$

respectively.

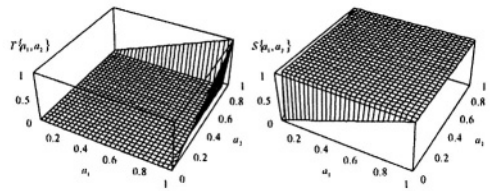

Fig. 2.5. Drastic triangular norms

All t-norms and t-conorms satisfy the following conditions:

$$S_M\{a_1,a_2,\ldots,a_n\}\le S\{a_1,a_2,\ldots,a_n\}\le S_D\{a_1,a_2,\ldots,a_n\} \tag{2.61}$$

$$T_D\{a_1,a_2,\ldots,a_n\}\le T\{a_1,a_2,\ldots,a_n\}\le T_M\{a_1,a_2,\ldots,a_n\} \tag{2.62}$$

Definition 2.14.
A t-norm and t-conorm are dual if they satisfy conditions:

$$\overset{n}{\underset{i=1}{S}}\{a_i\} = 1 - \overset{n}{\underset{i=1}{T}}\{1 - a_i\} \tag{2.63}$$

$$\overset{n}{\underset{i=1}{T}}\{a_i\} = 1 - \overset{n}{\underset{i=1}{S}}\{1 - a_i\} \tag{2.64}$$

It is easily seen that condition (2.63) and (2.64) are De Morgan's laws, well known in the classical set theory.

Definition 2.15.
(i) A non-increasing function $N:[0,1]\to[0,1]$ is called a negation if $N(0) = 1$ and $N(1) = 0$.
(ii) A negation $N:[0,1]\to[0,1]$ is called a strict negation if N is continuous and strictly decreasing.
(iii) A strict negation $N:[0,1]\to[0,1]$ is called a strong negation if it is an involution, i.e., if $N(N(a)) = a$.

Example 2.11.
Zadeh's negation is defined by

$$N(a) = 1 - a \tag{2.65}$$

and depicted in Fig. 2.6.

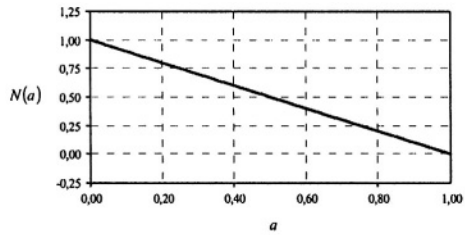

Fig. 2.6. Zadeh's negation

Example 2.12.
Yager's negation is given by

$$N(a) = \left(1 - a^p\right)^{\frac{1}{p}}, \quad p > 0 \tag{2.66}$$

and illustrated in Fig. 2.7.

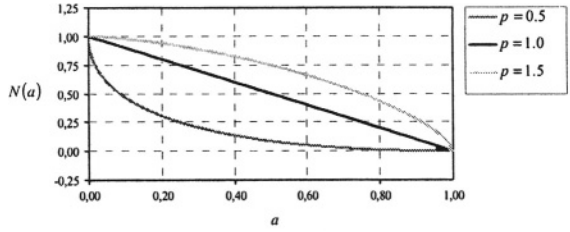

Fig. 2.7. Yager's negation

Example 2.13.
Sugeno's negation is given by

$$N(a)=\frac{1\text{-}a}{1+pa}\ ,\ p>-1 \tag{2.67}$$

and depicted in Fig. 2.8.

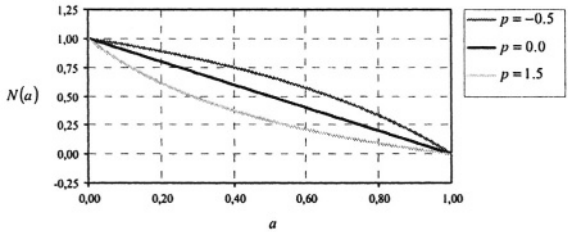

Fig. 2.8. Sugeno's negation

2.4. OPERATIONS ON FUZZY SETS

Definition 2.16.
The intersection of two fuzzy sets $A, B \subseteq \mathbf{X}$ is a fuzzy set denoted by $A \cap B$ whose membership function is given by

$$\mu_{A\cap B}(\mathbf{x})=\mu_A(\mathbf{x}) \overset{T}{*} \mu_A(\mathbf{x}) \tag{2.68}$$

where $\overset{T}{*}$ is any t-norm.

The intersection of n fuzzy sets $A_1, A_2, ..., A_n \subseteq \mathbf{X}$ is a fuzzy set denoted by $A = A_1 \cap A_2 \cap ... A_n = \bigcap_{i=1}^{n} A_i$ with a membership function defined by

$$\mu_A(\mathbf{x}) = \mathop{T}_{i=1}^{n} \mu_{A_i}(\mathbf{x}) \tag{2.69}$$

Definition 2.17.
The union of two fuzzy sets $A, B \subseteq \mathbf{X}$ is a fuzzy set denoted by $A \cup B$ whose membership function is given by

$$\mu_{A \cup B}(\mathbf{x}) = \mu_A(\mathbf{x}) \overset{S}{*} \mu_A(\mathbf{x}) \tag{2.70}$$

where $\overset{S}{*}$ is any t-conorm.

The union of n fuzzy sets $A_1, A_2, ..., A_n \subseteq \mathbf{X}$ is a fuzzy set denoted by $A = A_1 \cup A_2 \cup ... A_n = \bigcup_{i=1}^{n} A_i$ with a membership function defined by

$$\mu_A(\mathbf{x}) = \mathop{S}_{i=1}^{n} \mu_{A_i}(\mathbf{x}) \tag{2.71}$$

In Figures 2.9 and 2.10 we illustrate Definitions 2.16 and 2.17 for min/max and algebraic triangular norms, respectively.

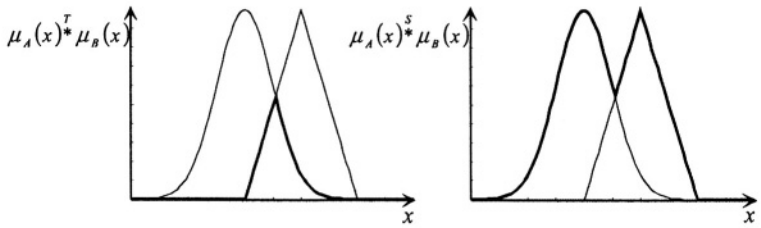

Fig. 2.9. Intersection and union operations based on min/max triangular norms

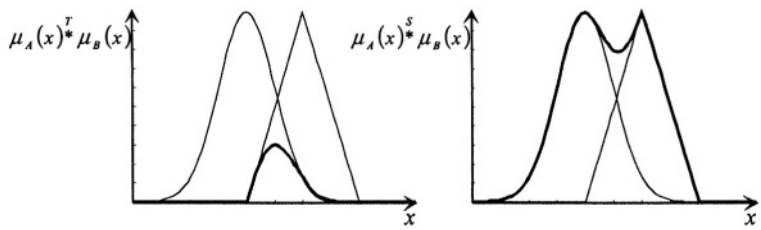

Fig. 2.10. Intersection and union operations based on algebraic triangular norms

Definition 2.18.

The complement of a fuzzy set A, denoted by $\tilde{A}$, is defined by

$$\mu_{\tilde{A}}(\mathbf{x}) = 1 - \mu_A(\mathbf{x}) \tag{2.72}$$

An example is shown in Fig. 2.11.

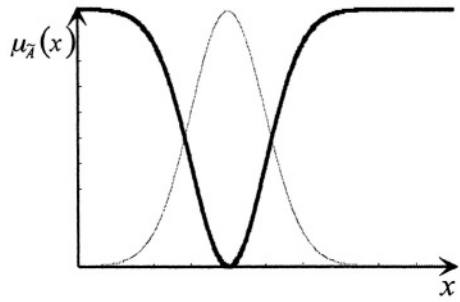

Fig. 2.11. Complement operation

Definition 2.19.

The Cartesian product of two fuzzy sets $A \subseteq \mathbf{X}$ and $B \subseteq \mathbf{Y}$, is denoted by $A \times B$, and defined as

$$\mu_{A\times B}(\mathbf{x},\mathbf{y}) = \min\{\mu_A(\mathbf{x}), \mu_B(\mathbf{y})\} \tag{2.73}$$

or

$$\mu_{A\times B}(\mathbf{x},\mathbf{y}) = \mu_A(\mathbf{x}) \cdot \mu_B(\mathbf{y}) \tag{2.74}$$

where $\mathbf{x} \in \mathbf{X}$ and $\mathbf{y} \in \mathbf{Y}$.

The Cartesian product of n fuzzy sets $A_1 \subseteq \mathbf{X}_1$, $A_2 \subseteq \mathbf{X}_2, \ldots, A_n \subseteq \mathbf{X}_n$, is denoted by $A_1 \times A_2 \times \ldots \times A_n$, and defined as

$$\mu_{A_1\times A_2\times\ldots\times A_n}(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n) = \min_{i=1\ldots n}\{\mu_{A_i}(\mathbf{x}_i)\} \tag{2.75}$$

or

$$\mu_{A_1\times A_2\times\ldots\times A_n}(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n) = \prod_{i=1}^{n} \mu_{A_i}(\mathbf{x}_i) \tag{2.76}$$

In Figures 2.12 and 2.13 we depict the Cartesian product defined by the minimum operator and the product operator, respectively. In a general case the Cartesian product is defined by a t-norm.

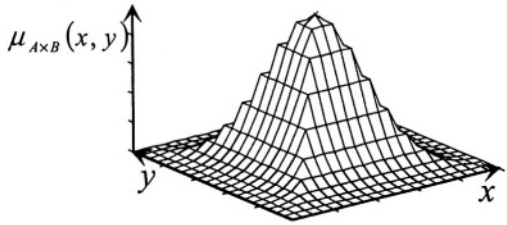

Fig. 2.12. Cartesian product defined by minimum operator

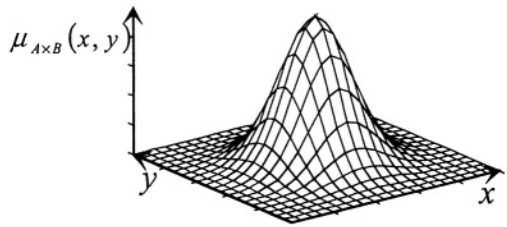

Fig. 2.13. Cartesian product defined by product operator

2.5. FUZZY RELATIONS

The most important definitions in fuzzy reasoning are concerned with fuzzy relations.

Definition 2.20.
Let $\mathbf{X}$ and $\mathbf{Y}$ be two universes of discourse. Binary fuzzy relations, denoted by R, are fuzzy sets which map each element in $\mathbf{X}\times\mathbf{Y}$ to a membership grade $\mu_R(\mathbf{x},\mathbf{y})$, where $\mathbf{x}\in\mathbf{X}$ and $\mathbf{y}\in\mathbf{Y}$. We write

$$R=\sum_{\mathbf{X}\times\mathbf{Y}}\frac{\mu_R(\mathbf{x},\mathbf{y})}{(\mathbf{x},\mathbf{y})} \tag{2.77}$$

or

$$R=\int_{\mathbf{X}\times\mathbf{Y}}\frac{\mu_R(\mathbf{x},\mathbf{y})}{(\mathbf{x},\mathbf{y})} \tag{2.78}$$

in discrete and continuous cases, respectively.

Example 2.14.
We apply Definition 2.20 to describe the imprecise statement that "y is about equal x". Let $\mathbf{X} = \{3,4\}$ and $\mathbf{Y} = \{4,5\}$. A fuzzy relation R can be defined as follows

$$R = \frac{1}{(4,4)} + \frac{0.8}{(4,3)} + \frac{0.8}{(5,4)} + \frac{0.6}{(5,3)} \tag{2.79}$$

Fuzzy relation (2.79) can be expressed in the matrix-form

$$R = \begin{bmatrix} 0.8 & 0.6 \\ 1 & 0.8 \end{bmatrix}$$

The next definition shows how to reduce the dimension of the product space by taking the supremum of the membership function over the domain of the variable corresponding to the dimension to be eliminated.

Definition 2.21.
Let R be a fuzzy relation defined on the Cartesian product $\mathbf{X}\times\mathbf{Y}$. The projection of R onto $\mathbf{Y}$ is defined as

$$\text{proj}(R;\mathbf{Y}) = \int_{\mathbf{y}\in\mathbf{Y}} \frac{\sup_{\mathbf{x}\in\mathbf{X}} \mu_R(\mathbf{x},\mathbf{y})}{\mathbf{y}}$$

where $\mu_R(\mathbf{x},\mathbf{y})$ is the membership grade of the pair $(\mathbf{x},\mathbf{y})$ to the fuzzy relation R.

Definition 2.22.
Let R and S be fuzzy relations in $\mathbf{X}\times\mathbf{Y}$ and $\mathbf{Y}\times\mathbf{Z}$, respectively. The sup-T composition of R and S is a fuzzy relation denoted by $R \circ S \subseteq \mathbf{X}\times\mathbf{Z}$ and defined by

$$\mu_{R\circ S}(\mathbf{x},\mathbf{z}) = \sup_{\mathbf{y}\in\mathbf{Y}} \left\{ \mu_R(\mathbf{x},\mathbf{y}) \overset{T}{*} \mu_S(\mathbf{y},\mathbf{z}) \right\} \tag{2.80}$$

Definition 2.23.
The sup-T composition of a fuzzy set $A \subseteq \mathbf{X}$ and fuzzy relation $R \subseteq \mathbf{X}\times\mathbf{Y}$ is a fuzzy set

$$B = A \circ R \subseteq \mathbf{Y} \tag{2.81}$$

defined by

$$\mu_B(\mathbf{y}) = \mu_{A\circ R}(\mathbf{y}) = \sup_{\mathbf{x}\in\mathbf{X}} \left\{ \mu_A(\mathbf{x}) \overset{T}{*} \mu_R(\mathbf{x},\mathbf{y}) \right\} \tag{2.82}$$

It should be noted that using Definitions 2.11 and 2.21, composition (2.81) can be alternatively expressed as follows

$$B = A \circ R = \text{proj}\{(ce(A; \mathbf{X}\times\mathbf{Y}) \cap R); \mathbf{Y}\} \tag{2.83}$$

2.6. FUZZY REASONING

The basic rule of inference in classical logic is modus ponens. The compositional rule of inference describes a composition of a fuzzy set and a fuzzy relation. Fuzzy rule

$$\text{IF} \quad x \quad \text{is} \quad A \quad \text{THEN} \quad y \quad \text{is} \quad B \tag{2.84}$$

is represented by a fuzzy relation R. Having given input linguistic value A', we can infer an output fuzzy set B' by the composition of the fuzzy set A' and the relation R. The generalized modus ponens is the extension of the conventional modus ponens tautology, to allow partial fulfilment of the premises:

Premise	x is A'
Implication	IF x is A THEN y is B
Conclusion	y is B'

where A, B, A', B' are fuzzy sets and x, y are linguistic variables. Following formulas (2.81) and (2.82), a fuzzy set B' is defined by

$$B' = A' \circ R = A' \circ (A \rightarrow B) \tag{2.85}$$

and

$$\mu_{B'}(y) = \mu_{A' \circ R}(y) = \sup_{x \in \mathbf{X}} \left\{ \mu_{A'}(x) \overset{T}{*} \mu_R(x, y) \right\} \tag{2.86}$$

The problem is to determine the membership function of the fuzzy relation described by

$$\mu_R(x, y) = \mu_{A \rightarrow B}(x, y) \tag{2.87}$$

based on the knowledge of $\mu_A(x)$ and $\mu_B(y)$. We denote

$$\mu_{A \rightarrow B}(x, y) = I(\mu_A(x), \mu_B(y)) \tag{2.88}$$

where $I(\cdot)$ is a fuzzy implication given in Definition 2.24 (see Fodor [21]).

Definition 2.24.

A fuzzy implication is a function $I : [0,1]^2 \rightarrow [0,1]$ satisfying the following conditions:

(I1) if $a_1 \leq a_3$, then $I(a_1, a_2) \geq I(a_3, a_2)$, for all $a_1, a_2, a_3 \in [0,1]$,

(I2) if $a_2 \le a_3$, then $I(a_1,a_2) \le I(a_1,a_3)$, for all $a_1,a_2,a_3 \in [0,1]$,

(I3) $I(0,a_2)=1$, for all $a_2 \in [0,1]$ (falsity implies anything),

(I4) $I(a_1,1)=1$, for all $a_1 \in [0,1]$ (anything implies tautology),

(I5) $I(1,0) = 0$ (Booleanity).

Selected fuzzy implications satisfying all or some of the above conditions are listed in Table 2.1.

Table 2.1 Fuzzy implications

No	Name	Implication $I(a,b)$
1	Kleene-Dienes (binary)	$\max\{1-a,b\}$
2	Łukasiewicz	$\min\{1,1-a+b\}$
3	Reichenbach	$1-a+a\cdot b$
4	Fodor	$\begin{cases} 1 & \text{if } a \le b \\ \max\{1-a,b\} & \text{if } a > b \end{cases}$
5	Rescher	$\begin{cases} 1 & \text{if } a \le b \\ 0 & \text{if } a > b \end{cases}$
6	Goguen	$\begin{cases} 1 & \text{if } a = 0 \\ \min\left\{1,\frac{b}{a}\right\} & \text{if } a > 0 \end{cases}$
7	Gödel	$\begin{cases} 1 & \text{if } a \le b \\ b & \text{if } a > b \end{cases}$
8	Yager	$\begin{cases} 1 & \text{if } a = 0 \\ b^a & \text{if } a > 0 \end{cases}$
9	Zadeh	$\max\{\min\{a,b\},1-a\}$
10	Willmott	$\min\begin{cases} \max\{1-a,b\}, \\ \max\{a,1-b,\min\{1-a,b\}\} \end{cases}$
11	Dubois-Prade	$\begin{cases} 1-a & \text{if } b = 0 \\ b & \text{if } a = 1 \\ 1 & \text{if } \textit{otherwise} \end{cases}$

In this table, implications 1-4 are examples of an S-implication associated with a t-conorm

$$I(a,b) = S\{1-a,b\} \tag{2.89}$$

Neuro-fuzzy systems based on fuzzy implications given in Table 2.1 are called logical systems. Implications 6 and 7 belong to a group of R-implications associated with the t-norm T and given by

$$I(a,b) = \sup_z \{z \mid T\{a,z\} \le b\}, \quad a,b \in [0,1] \tag{2.90}$$

The Zadeh implication belongs to a group of Q-implications given by

$$I(a,b) = S\{N(a),T\{a,b\}\}, \quad a,b \in [0,1] \tag{2.91}$$

It is easy to verify that S-implications and R-implications satisfy all the conditions of Definition 2.24. However, the Zadeh implication violates conditions I1 and I4, whereas the Willmott implication violates conditions I1, I3 and I4. In practice we frequently use Mamdani-type operators (see Section 3.1) which do not satisfy conditions of Definition 2.24.

2.7. PROBLEMS

Problem 2.1. Let $\mathbf{X} = \{1, 2, 3, 4, 5\}$ and

$$A = \frac{0.8}{2} + \frac{1}{3} + \frac{0.7}{4}, \tag{2.92}$$

$$B = \frac{0.6}{3} + \frac{1}{4} + \frac{0.4}{5} \tag{2.93}$$

Determine the intersection and union of fuzzy sets A and B using the min/max triangular norms.

Problem 2.2. Let $\mathbf{X} = \{1, 2, 3, 4, 5\}$ and

$$A = \frac{0.5}{2} + \frac{1}{3} + \frac{0.8}{5} \tag{2.94}$$

Find the complement $\tilde{A}$ of A and determine

$$A \cap \tilde{A} \text{ and } A \cup \tilde{A} \tag{2.95}$$

using the min/max triangular norms. Discuss the results and compare with traditional logic.

Problem 2.3. Repeat Problem 2.2 using the algebraic triangular norms.

Problem 2.4. Let $\mathbf{X} = \{2,4\}$, $\mathbf{Y} = \{2,4,5\}$ and

$$A = \frac{0.6}{2} + \frac{0.8}{4}, \tag{2.96}$$

$$B = \frac{0.4}{2} + \frac{0.9}{4} + \frac{0.2}{5} \tag{2.97}$$

Find the Cartesian product of A and B.

Problem 2.5. Let $\mathbf{X} = \{x_1, x_2\}$ and $\mathbf{Y} = \{y_1, y_2\}$. Let

$$R = \begin{bmatrix} 0.3 & 0.6 \\ 0.7 & 1 \end{bmatrix} \tag{2.98}$$

be the fuzzy relation defined on the Cartesian product $\mathbf{X}\times\mathbf{Y}$. Find the projection of R onto $\mathbf{Y}$.

Problem 2.6. $\mathbf{X}=\{x_1,x_2\}$, $\mathbf{Y}=\{y_1,y_2\}$ and $\mathbf{Z}=\{z_1,z_2,z_3\}$. Define the fuzzy relations

$$R=\begin{bmatrix}0.3 & 0.6\\ 0.7 & 1\end{bmatrix} \tag{2.99}$$

and

$$S=\begin{bmatrix}0.4 & 0.7 & 0.9\\ 0.8 & 1 & 0.5\end{bmatrix} \tag{2.100}$$

in $\mathbf{X}\times\mathbf{Y}$ and $\mathbf{Y}\times\mathbf{Z}$, respectively. Determine the sup-T composition of R and S using the min t-norm.

Problem 2.7. Let $\mathbf{X}=\{x_1,x_2,x_3\}$ and $\mathbf{Y}=\{y_1,y_2\}$. Let

$$A=\frac{0.5}{x_1}+\frac{1}{x_2}+\frac{0.7}{x_1} \tag{2.101}$$

and

$$R=\begin{bmatrix}0.6 & 0.8\\ 0.3 & 1\\ 0.9 & 0.4\end{bmatrix} \tag{2.102}$$

where $A\subseteq\mathbf{X}$ and $R\subseteq\mathbf{X}\times\mathbf{Y}$ Determine the sup-T composition of fuzzy set A and fuzzy relation R.

Problem 2.8. Show that for the singleton fuzzy set A' formula (2.86) takes the form

$$\mu_{B'}(y)=I(\mu_A(\bar{x}),\mu_B(y)) \tag{2.103}$$

Problem 2.9. Verify if the min/max triangular norms are dual.

Problem 2.10. Prove that all t-norms and t-conorms satisfy conditions (2.61) and (2.62).

Problem 2.11. Verify if the Rescher and Yager fuzzy implications satisfy conditions of Definition 2.24.

Chapter 3

FUZZY INFERENCE SYSTEMS

3.1. INTRODUCTION

In up-to-date literature two approaches have been proposed to design fuzzy systems having linguistic descriptions of inputs and outputs. The fundamental differences between them is explained in the Foreword to this book written by Professor Lotfi Zadeh.

(i) The first approach, called the Mamdani method, uses a conjunction for inference and a disjunction to aggregate individual rules. In the Mamdani approach, the most widely used operators measuring the truth of the relation between the input and output are the following

$$I(a,b) = \min\{a,b\} \tag{3.1}$$

and

$$I(a,b) = a \cdot b \tag{3.2}$$

or more generally

$$I(a,b) = T\{a,b\} \tag{3.3}$$

It should be emphasized that formulas (3.1) and (3.2) do not satisfy the conditions of a fuzzy implication formulated by Fodor [21]. Following Mendel [57, 58], we refer to (3.1) and (3.2) as to "engineering implications" contrary to the fuzzy implications satisfying the axiomatic definition (see Definition 2.24).

The aggregation is performed by the application of a t-conorm

$$S\{a_1,a_2,\ldots,a_n\} = S\{\mathbf{a}\} = a_1 \overset{S}{*} a_2 \overset{S}{*} \ldots \overset{S}{*} a_n = \overset{n}{\underset{i=1}{S}}\{a_i\} \tag{3.4}$$

e.g.,

$$S\{\mathbf{a}\} = \max_{i=1,\ldots,n}\{a_i\} \tag{3.5}$$

It should also be noted that in most cases the aggregation of rules is performed as part of the defuzzification (see, e.g. [58, 100]).

(ii) The second paradigm applies fuzzy implications to the inference and the conjunction to the aggregation. Instead of using the "engineering implication" (3.3) we design fuzzy systems based on the fuzzy implications defined in Section 2.6.

For fuzzy systems with a logical implication, the aggregation is realized by a t-norm

$$T\{a_1, a_2, \ldots, a_n\} = T\{\mathbf{a}\} = a_1 \overset{T}{*} a_2 \overset{T}{*} \ldots \overset{T}{*} a_n = \overset{n}{\underset{i=1}{T}}\{a_i\} \tag{3.6}$$

e.g.,

$$T\{\mathbf{a}\} = \min_{i=1,\ldots,n}\{a_i\} \tag{3.7}$$

It should be noted that the aggregation of antecedents in each rule is performed by the same formula (3.6) for both Mamdani and logical-type systems.

In the next sections we present a formal description of neuro-fuzzy systems and illustrate Mamdani-type and logical-type reasoning on various "engineering" and fuzzy implications. A generalized neuro-fuzzy inference system will be introduced, which reflects a structure of all neuro-fuzzy systems studied in this book. We will also provide data sets on which all systems, developed in the next chapters, are tested and compared.

3.2. DESCRIPTION OF FUZZY INFERENCE SYSTEMS

In this book, we consider multi-input-single-output fuzzy NFIS mapping $\mathbf{X} \to \mathbf{Y}$, where $\mathbf{X} \subset \mathbf{R}^n$ and $\mathbf{Y} \subset \mathbf{R}$. The system (see Fig. 3.1) is composed of a fuzzifier, a fuzzy rule base, a fuzzy inference engine and a defuzzifier.

The fuzzifier performs a mapping from the observed crisp input space $\mathbf{X} \subset \mathbf{R}^n$ to a fuzzy set defined in $\mathbf{X}$. The most commonly used fuzzifier is the singleton fuzzifier which maps $\bar{\mathbf{x}} = [\bar{x}_1, \ldots, \bar{x}_n] \in \mathbf{X}$ into fuzzy set $A' \subseteq \mathbf{X}$ characterized by the membership function

$$\mu_{A'}(\mathbf{x}) = \begin{cases} 1 & \text{if} \quad \mathbf{x} = \bar{\mathbf{x}} \\ 0 & \text{if} \quad \mathbf{x} \neq \bar{\mathbf{x}} \end{cases} \tag{3.8}$$

The fuzzy rule base consists of a collection of N fuzzy IF-THEN rules, aggregated by the disjunction or the conjunction, in the form

$$R^{(k)}:\begin{cases} \text{IF} & x_1 \text{ is } A_1^k & \text{AND} \\ & x_2 \text{ is } A_2^k & \text{AND} \quad \ldots \\ & x_n \text{ is } A_n^k & \\ \text{THEN} & y \text{ is } B^k & \end{cases} \tag{3.9}$$

or

$$R^{(k)}: \text{IF } \mathbf{x} \text{ is } A^k \text{ THEN } y \text{ is } B^k \tag{3.10}$$

where $\mathbf{x}=[x_1,\ldots,x_n]\in \mathbf{X}$, $y\in \mathbf{Y}$, $A^k = A_1^k \times A_2^k \times \ldots \times A_n^k$, $A_1^k, A_2^k, \ldots, A_n^k$ are fuzzy sets characterized by membership functions $\mu_{A_i^k}(x_i)$, $i=1,\ldots,n$, $k=1,\ldots,N$, whereas B^k are fuzzy sets characterized by membership functions $\mu_{B^k}(y)$, $k=1,\ldots,N$. The firing strength of the k-th rule, $k=1,...,N$, is defined by

$$\tau_k(\bar{\mathbf{x}}) = \mathop{T}_{i=1}^{n}\left\{\mu_{A_i^k}(\bar{x}_i)\right\} = \mu_{A^k}(\bar{\mathbf{x}}) \tag{3.11}$$

In the book notations τ_k and $\mu_{A^k}(\bar{\mathbf{x}})$ will be used interchangeably.

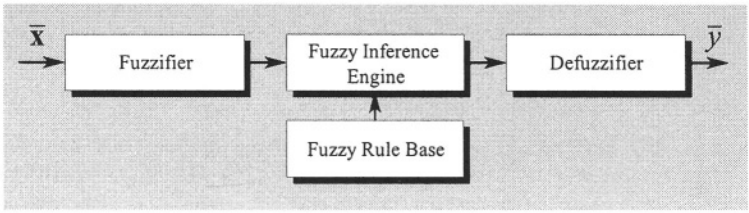

Fig. 3.1. Fuzzy inference system

The fuzzy inference engine determines the mapping from the fuzzy sets in the input space $\mathbf{X}$ to the fuzzy sets in the output space $\mathbf{Y}$. Each of N rules (3.9) determines a fuzzy set $\bar{B}^k \subseteq \mathbf{Y}$ given by the compositional rule of inference

$$\bar{B}^k = A' \circ \left(A^k \rightarrow B^k\right) \tag{3.12}$$

where $A^k = A_1^k \times A_2^k \times \ldots \times A_n^k$.

Fuzzy sets $\bar{B}^k$, according to formula (3.12), are characterized by membership functions expressed by the *sup-star* composition

$$\mu_{\bar{B}^k}(y) = \sup_{\mathbf{x}\in\mathbf{X}}\left\{\mu_{A'}(\mathbf{x}) * \mu_{A_1^k \times \ldots \times A_n^k \rightarrow B^k}(\mathbf{x}, y)\right\} \tag{3.13}$$

where $*^T$ can be any operator in the class of t-norms. It is easily seen that for a crisp input $\mathbf{x} = [x_1, \ldots, x_n] \in \mathbf{X}$, i.e., the singleton fuzzifier (3.8), formula (3.13) becomes

$$\begin{aligned} \mu_{\overline{B}^k}(y) &= \mu_{A_1^k \times \ldots \times A_n^k \to B^k}(\overline{\mathbf{x}}, y) \\ &= \mu_{A^k \to B^k}(\overline{\mathbf{x}}, y) \\ &= I\left(\mu_{A^k}(\overline{\mathbf{x}}), \mu_{B^k}(y)\right) \end{aligned} \tag{3.14}$$

where $I(\cdot)$ is an "engineering implication" or fuzzy implication. More precisely,

$$I\left(\mu_{A^k}(\overline{\mathbf{x}}), \mu_{B^k}(y)\right) = \begin{cases} I_{eng}\left(\mu_{A^k}(\overline{\mathbf{x}}), \mu_{B^k}(y)\right) \\ \text{for } \textit{the Mamdani approach} \\ I_{fuzzy}\left(\mu_{A^k}(\overline{\mathbf{x}}), \mu_{B^k}(y)\right) \\ \text{for } \textit{the logical approach} \end{cases} \tag{3.15}$$

As we mentioned in Section 3.1, in the Mamdani approach

$$I_{eng}\left(\mu_{A^k}(\overline{\mathbf{x}}), \mu_{B^k}(y)\right) = T\left\{\mu_{A^k}(\overline{\mathbf{x}}), \mu_{B^k}(y)\right\} \tag{3.16}$$

In the logical approach we apply fuzzy implications listed in Table 2.1. Other definitions of fuzzy implications can be found in [14]. The Kleene-Dienes, Łukasiewicz, Reichenbach and Fodor implications are examples of the S-implication given by

$$I_{fuzzy}\left(\mu_{A^k}(\overline{\mathbf{x}}), \mu_{B^k}(y)\right) = S\left\{1 - \mu_{A^k}(\overline{\mathbf{x}}), \mu_{B^k}(y)\right\} \tag{3.17}$$

Obviously, an S-implication can be generated by various t-conorms given in Section 2.3 and Table 4.1 (see also Section 4.4 for a definition of the S-implication generated by the Dombi t-conorm).

The aggregation operator, applied in order to obtain the fuzzy set B' based on fuzzy sets $\overline{B}^k$, is the t-norm or t-conorm operator, depending on the type of fuzzy implication. In Table 3.1 we describe connectives in the Mamdani approach and logical approach. In case of the Mamdani approach, the aggregation is carried out by

$$B' = \bigcup_{k=1}^{N} \overline{B}^k \tag{3.18}$$

The membership function of B' is computed by the use of a t-conorm, that is

$$\mu_{B'}(y) = \underset{k=1}{\overset{N}{S}} \mu_{\overline{B}^k}(y) \tag{3.19}$$

When we use the logical model, the aggregation is carried out by

$$B' = \bigcap_{k=1}^{N} \overline{B}^k \tag{3.20}$$

The membership function of B' is determined by the use of a t-norm, i.e.

$$\mu_{B'}(y) = \underset{k=1}{\overset{N}{T}} \left\{ \mu_{\overline{B}^k}(y) \right\} \tag{3.21}$$

As a result of the fuzzy reasoning we obtain the fuzzy set B'.

Table 3.1 Operations in fuzzy inference

System type	Aggregation of antecedents	Implication	Aggregation of rules
Mamdani	t-norm	„engineering" – t-norm	t-conorm
Logical	t-norm	logical	t-norm

The defuzzifier performs a mapping from the fuzzy set B' to a crisp point $\overline{y}$ in $\mathbf{Y} \subset \mathbf{R}$. The COA (center of area) method is defined by the following formula

$$\overline{y} = \frac{\int_{\mathbf{Y}} y \cdot \mu_{B'}(y) dy}{\int_{\mathbf{Y}} \mu_{B'}(y) dy} \tag{3.22}$$

or by

$$\overline{y} = \frac{\sum_{r=1}^{N} \overline{y}^r \cdot \mu_{B'}(\overline{y}^r)}{\sum_{r=1}^{N} \mu_{B'}(\overline{y}^r)} \tag{3.23}$$

in the discrete form, where $\overline{y}^r$ are centers of the membership functions $\mu_{B^r}(y)$, i.e., for $r = 1,\ldots,N$

$$\mu_{B^r}(\overline{y}^r) = \max_{y \in \mathbf{Y}} \left\{ \mu_{B^r}(y) \right\} \tag{3.24}$$

For other definitions of the defuzzifier the reader is referred to [58].

***Remark** 3.1:* Formula (3.14) has been derived assuming a singleton fuzzifier given by (3.8). The fuzzifier characterized by the membership function that equals 1 for $\mathbf{x} = \bar{\mathbf{x}}$ and decreases from 1 as $\mathbf{x}$ moves away from $\bar{\mathbf{x}}$ is called a non-singleton fuzzifier. An example of the non-singleton fuzzifier is the fuzzy set A' with the Gaussian membership function

$$\mu_{A'}(\mathbf{x}) = \exp\left(-\frac{(\mathbf{x}-\bar{\mathbf{x}})^T(\mathbf{x}-\bar{\mathbf{x}})}{\sigma^2}\right) \tag{3.25}$$

It should by noted that formula (3.23) is also applicable to systems with a non-singleton fuzzifier. However, in such a case equation (3.14) is not valid and the *sup-star* composition (3.13) should be determined by the use of formula (2.83). According to that definition we should follow the following steps:

a) Find the cylindrical extension of A' on $\mathbf{X}\times\mathbf{Y}$, given by

$$\mathrm{ce}(A';\mathbf{X}\times\mathbf{Y}) = \int_{\mathbf{X}\times\mathbf{Y}} \frac{\mu_{A'}(\mathbf{x})}{(\mathbf{x},y)} \tag{3.26}$$

b) Find the intersection, denoted by c^k, of $\mathrm{ce}(A';\mathbf{X}\times\mathbf{Y})$ and fuzzy relation $A^k \to B^k$, i.e.

$$c^k = \mathrm{ce}(A';\mathbf{X}\times\mathbf{Y}) \cap (A^k \to B^k), \quad k = 1,\ldots,N \tag{3.27}$$

c) Find the projection of c^k on $\mathbf{X}\times\mathbf{Y}$, i.e.

$$\begin{aligned} \bar{B}^k &= A' \circ (A^k \to B^k) \\ &= \mathrm{proj}(c^k, \mathbf{Y}) \\ &= \int_{\mathbf{Y}} \frac{\sup_{\mathbf{x}\in\mathbf{X}} \mu_{c^k}(\mathbf{x},y)}{y} \end{aligned} \tag{3.28}$$

We will illustrate these steps on a number of examples in Sections 3.3 and 3.4.

3.3. MAMDANI-TYPE INFERENCE

The first rule in this approach is the minimum rule, called also the Mamdani rule. It is defined by equation

$$\mu_{A^k\to B^k}(\mathbf{x},y) = \min\{\mu_{A^k}(\mathbf{x}), \mu_{B^k}(y)\} \tag{3.29}$$

Figures 3.2-a, b and 3.3-a, b depict an exemplary reasoning process for this fuzzy relation. They describe two cases. The first case shown in Fig. 3.2-a, b concerns singleton fuzzification (3.8), i.e. the reasoning expressed by (3.14). In Fig. 3.2-a we assume one crisp input, n = 1, and three rules, N = 3. The results of fuzzy reasoning for each rule are represented by fuzzy sets $\bar{B}^1$, $\bar{B}^2$

and $\overline{B}^3$. These sets are aggregated in Fig. 3.2b by the use of three methods: the maximum t-conorm, the algebraic sum t-conorm and the bounded sum t-conorm. Figures 3.3-a and 3.3-b concern a non-singleton case and show the complexity of such a reasoning process. In the non-singleton case the reasoning is described in Remark 3.1 and the aggregation is performed as in the crisp case.

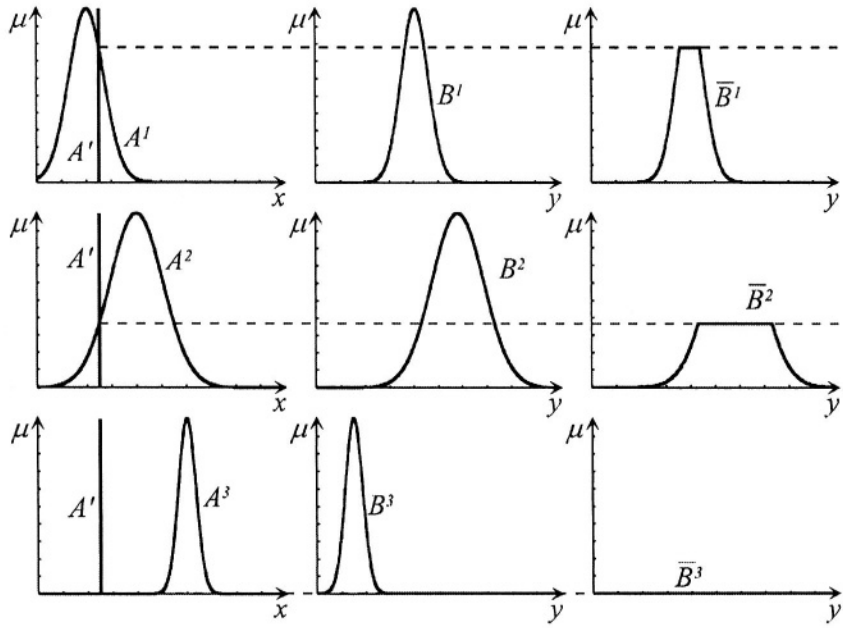

Fig. 3.2-a. Illustration of fuzzy inference based on the Mamdani rule for a crisp input

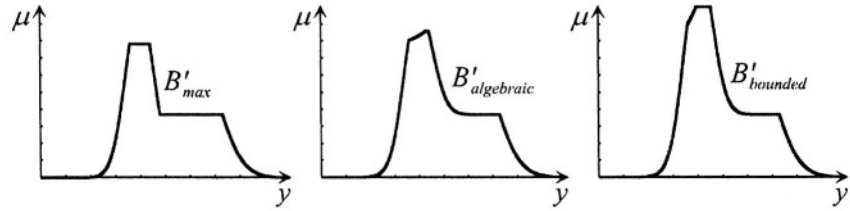

Fig. 3.2b. Illustration of fuzzy inference based on the Mamdani rule for a crisp input-aggregation

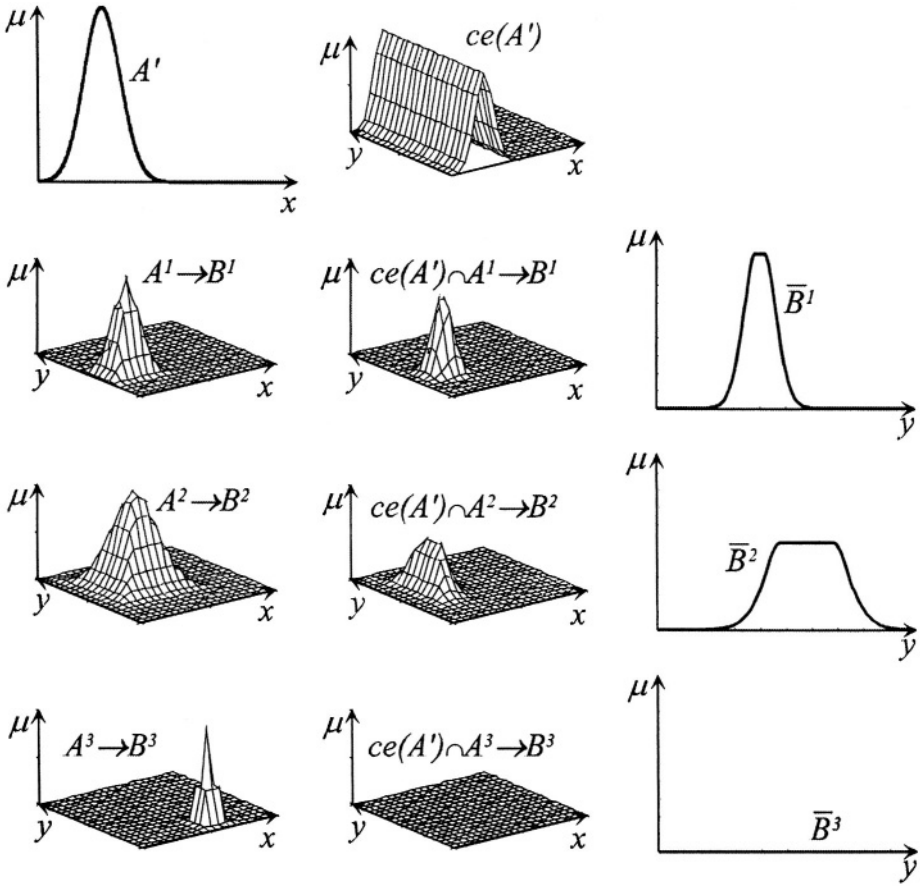

Fig. 3.3-a. Illustration of fuzzy inference based on the Mamdani rule-general case

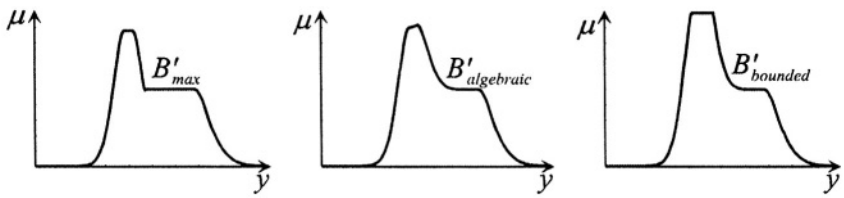

Fig. 3.3b. Illustration of fuzzy inference based on the Mamdani rule for a general case – aggregation

The second fuzzy reasoning rule belonging to the discussed group is algebraic product known under the name of the Larsen rule given by

$$\mu_{A^k \to B^k}(\mathbf{x}, y) = \mu_{A^k}(\mathbf{x}) \cdot \mu_{B^k}(y) \tag{3.30}$$

Similarly to the Mamdani case, the Larsen reasoning process is shown in Figures 3.4-a, b and 3.5-a, b. Figures 3.4-a, b concern the singleton fuzzy set A', whereas Figures 3.5-a, b depict a more general case with the non-singleton set A'.

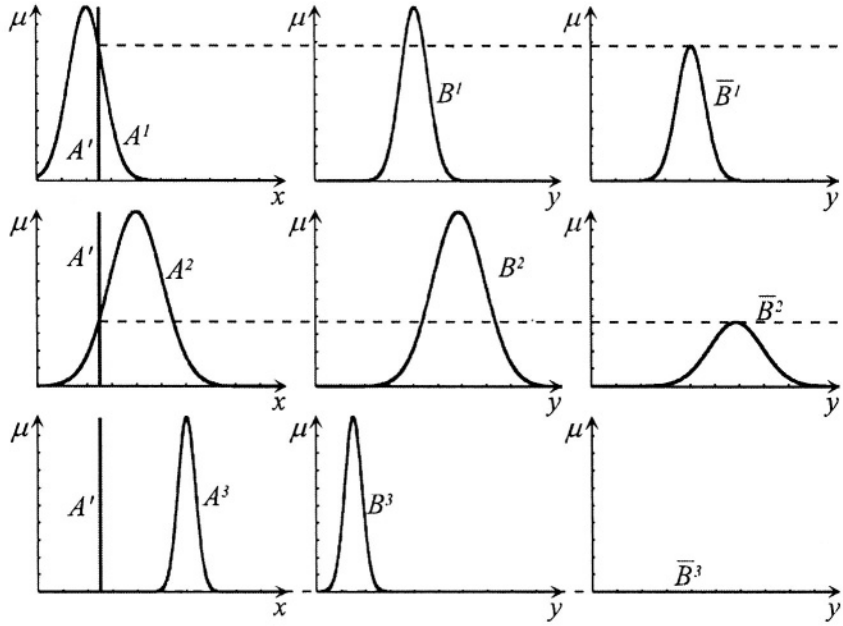

Fig. 3.4-a. Illustration of fuzzy inference based on the Larsen rule for a crisp input

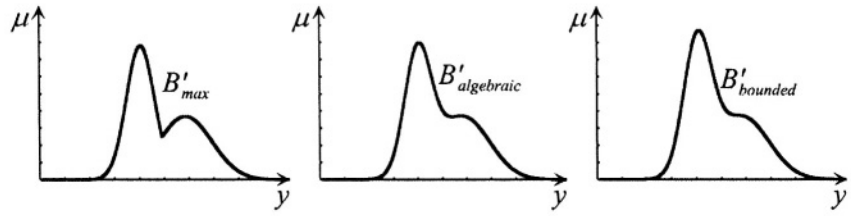

Fig. 3.4-b. Illustration of fuzzy inference based on the Larsen rule for a crisp input-aggregation

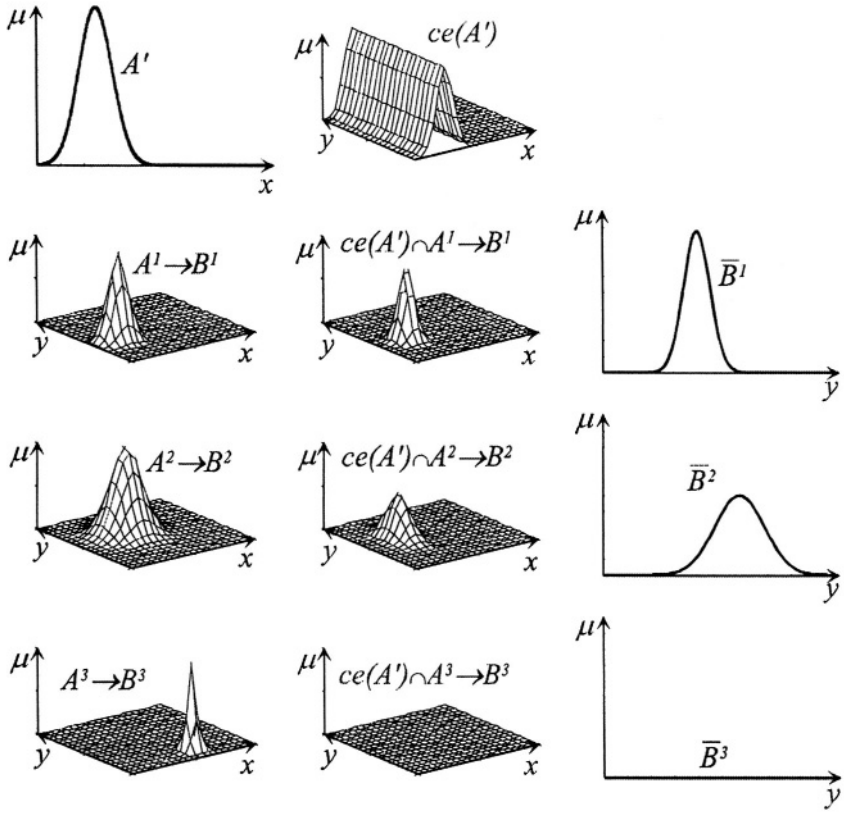

Fig. 3.5-a. Illustration of fuzzy inference based on the Larsen rule-general case

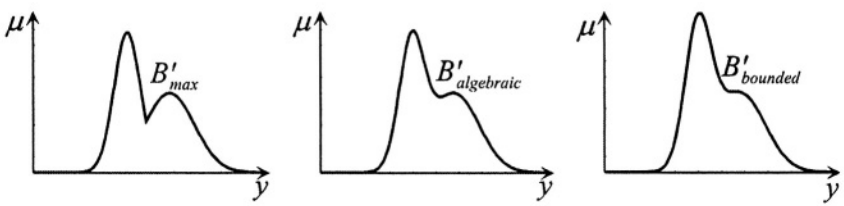

Fig. 3.5-b. Illustration of fuzzy inference based on the Larsen rule for a general case – aggregation

Obviously, many other t-norms represent "engineering implications". The reader can easily analyse a fuzzy reasoning based on various t-norms analogously to the Mamdani rule (3.29) and to the Larsen rule (3.30).

3.4. LOGICAL-TYPE INFERENCE

Membership functions of the relation $A^k \to B^k$ used in the logical-type inference are a generalization of implications in the classic (non-fuzzy) propositional logic, therefore we call them fuzzy implications (see Definition 2.24). One of the simplest fuzzy implications is the binary implication (Kleene-Dienes implication), defined by

$$\mu_{A^k \to B^k}(\mathbf{x}, y) = \max\left\{1 - \mu_{A^k}(\mathbf{x}), \mu_{B^k}(y)\right\} \tag{3.31}$$

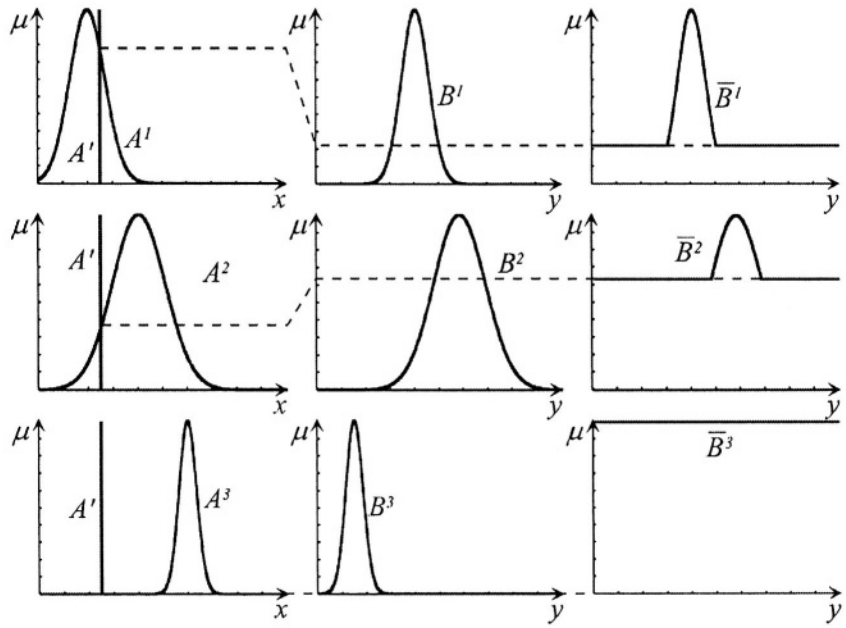

Fig. 3.6-a. Illustration of fuzzy inference based on the binary (Kleene-Dienes) inference for a crisp input

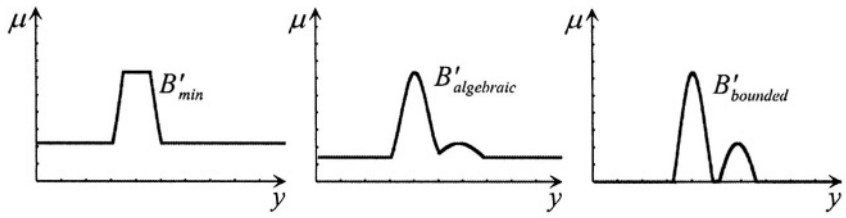

Fig. 3.6-b. Illustration of fuzzy inference based on the binary (Kleene-Dienes) inference for a crisp input-aggregation

Figures 3.6-a, b and 3.7-a, b show an exemplary reasoning process. Figures 3.6-a, b depict fuzzy reasoning in the case of the singleton fuzzification, whereas Figures 3.7-a, b concern the non-singleton fuzzification using a Gaussian membership function.

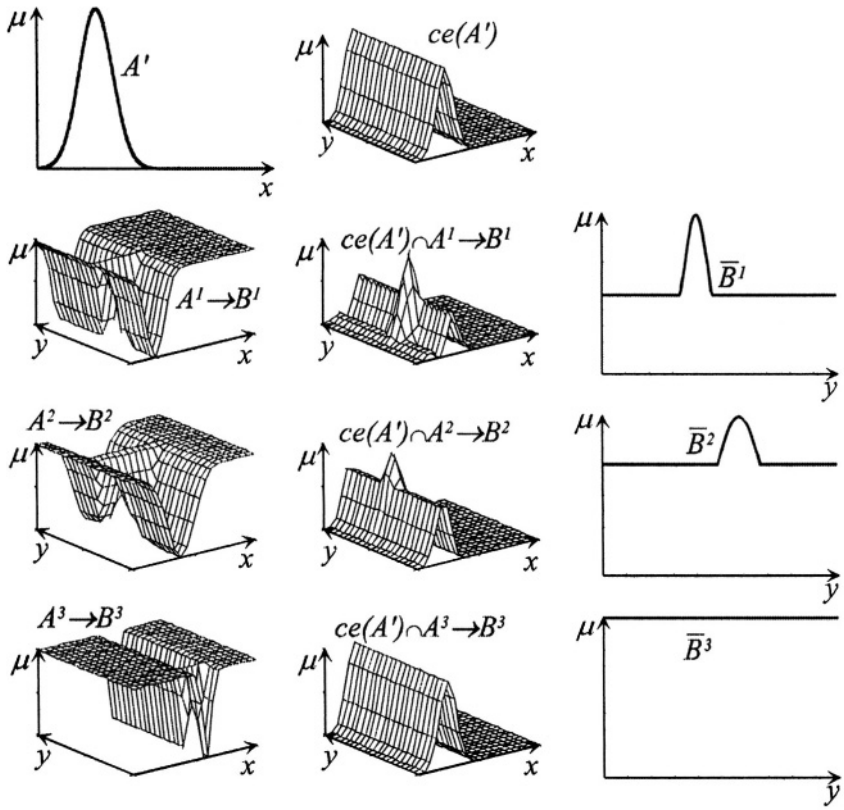

Fig. 3.7-a. Illustration of fuzzy inference based on the binary (Kleene-Dienes) inference-general case

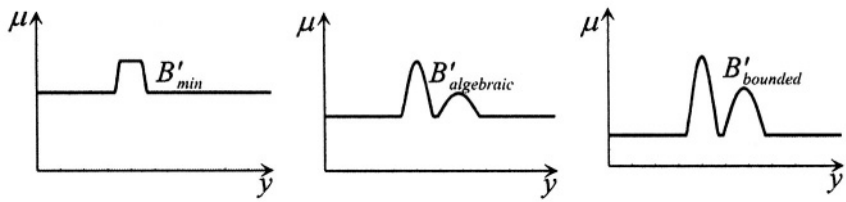

Fig. 3.7-b. Illustration of fuzzy inference based on the binary (Kleene-Dienes) inference for a general case – aggregation

Another well known fuzzy implication is the **Łukasiewicz** implication, given by

$$\mu_{A^k \to B^k}(\mathbf{x}, y) = \min\{1, 1 - \mu_{A^k}(\mathbf{x}) + \mu_{B^k}(y)\} \tag{3.32}$$

The reasoning process by this implication is depicted in Fig. 3.8-a, b for the singleton set A', and in Fig. 3.9-a, b for the Gaussian set A'.

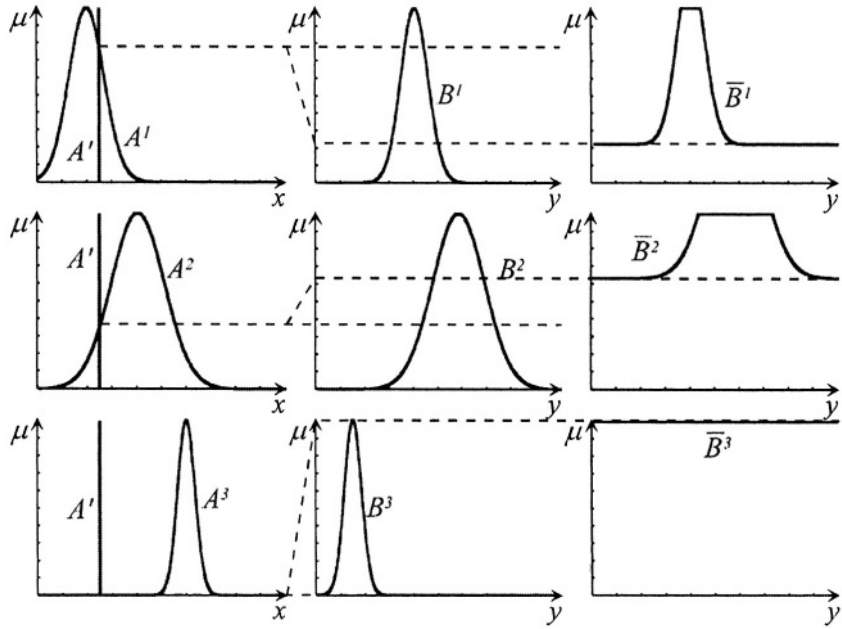

Fig. 3.8-a. Illustration of fuzzy inference based on the Łukasiewicz inference for a crisp input

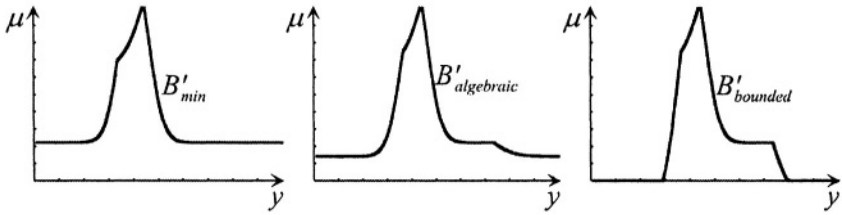

Fig. 3.8b. Illustration of fuzzy inference based on the Łukasiewicz inference for a crisp input-aggregation

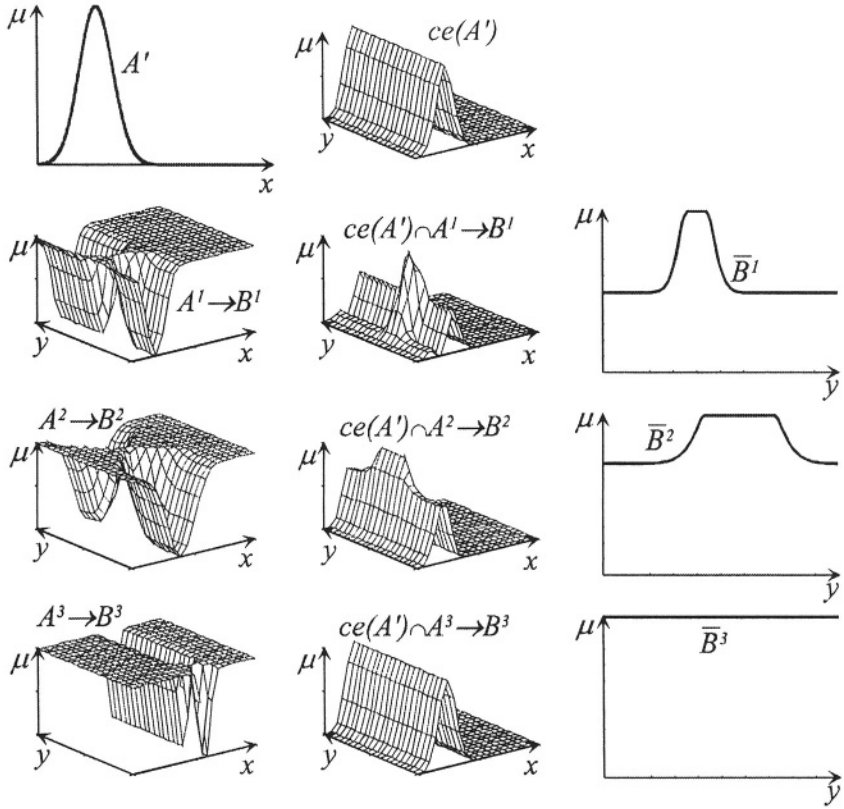

Fig. 3.9-a. Illustration of fuzzy inference based on the Łukasiewicz inference-general case

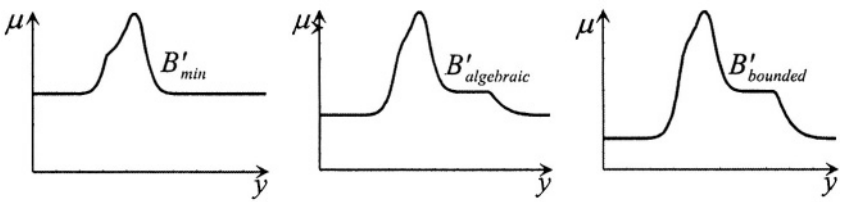

Fig. 3.9b. Illustration of fuzzy inference based on the Łukasiewicz inference for a general case – aggregation

In a similar way we may illustrate the fuzzy inference based on other fuzzy implications given in Table 2.1.

3.5. GENERALIZED NEURO-FUZZY SYSTEM

In this section we generalize the Mamdani-type and the logical-type approach to a neuro-fuzzy system design.

a) Mamdani-type neuro-fuzzy systems

In this approach, function $I(\cdot)$ given by (3.14) is a t-norm (e.g., minimum, algebraic or Dombi), i.e.

$$I\left(\mu_{A^k}(\bar{\mathbf{x}}),\mu_{B^k}(\bar{y}^r)\right)=T\left\{\mu_{A^k}(\bar{\mathbf{x}}),\mu_{B^k}(\bar{y}^r)\right\} \tag{3.33}$$

The aggregated output fuzzy set $B'\subseteq\mathbf{Y}$ is given by

$$\mu_{B'}(\bar{y}^r)=\underset{k=1}{\overset{N}{S}}\left\{\mu_{\bar{B}^k}(\bar{y}^r)\right\}=\underset{k=1}{\overset{N}{S}}\left\{T\left\{\mu_{A^k}(\bar{\mathbf{x}}),\mu_{B^k}(\bar{y}^r)\right\}\right\} \tag{3.34}$$

Consequently, formula (3.23) takes the form

$$\bar{y}=\frac{\sum_{r=1}^{N}\bar{y}^r\cdot\underset{k=1}{\overset{N}{S}}\left\{T\left\{\underset{i=1}{\overset{n}{T}}\left\{\mu_{A_i^k}(\bar{x}_i)\right\},\mu_{B^k}(\bar{y}^r)\right\}\right\}}{\sum_{r=1}^{N}\underset{k=1}{\overset{N}{S}}\left\{T\left\{\underset{i=1}{\overset{n}{T}}\left\{\mu_{A_i^k}(\bar{x}_i)\right\},\mu_{B^k}(\bar{y}^r)\right\}\right\}} \tag{3.35}$$

Obviously, the t-norms used to connect the antecedents in the rules and in the "engineering implication" do not have to be the same. Besides, they can be chosen as differentiable functions as e.g. Dombi families (see Section 4.4).

b) Logical-type neuro-fuzzy systems

In this approach, function $I(\cdot)$ given by (3.14) is a fuzzy implication (see Table. 2.1), i.e.

$$I\left(\mu_{A^k}(\bar{\mathbf{x}}),\mu_{B^k}(\bar{y}^r)\right)=I_{fuzzy}\left(\mu_{A^k}(\bar{\mathbf{x}}),\mu_{B^k}(\bar{y}^r)\right) \tag{3.36}$$

The aggregated output fuzzy set $B'\subseteq\mathbf{Y}$ is given by

$$\mu_{B'}(\bar{y}^r)=\underset{k=1}{\overset{N}{T}}\left\{\mu_{\bar{B}^k}(\bar{y}^r)\right\}=\underset{k=1}{\overset{N}{T}}\left\{I_{fuzzy}\left(\mu_{A^k}(\bar{\mathbf{x}}),\mu_{B^k}(\bar{y}^r)\right)\right\} \tag{3.37}$$

and formula (3.23) becomes

$$\bar{y}=\frac{\sum_{r=1}^{N}\bar{y}^r\cdot\underset{k=1}{\overset{N}{T}}\left\{I_{fuzzy}\left(\underset{i=1}{\overset{n}{T}}\left\{\mu_{A_i^k}(\bar{x}_i)\right\},\mu_{B^k}(\bar{y}^r)\right)\right\}}{\sum_{r=1}^{N}\underset{k=1}{\overset{N}{T}}\left\{I_{fuzzy}\left(\underset{i=1}{\overset{n}{T}}\left\{\mu_{A_i^k}(\bar{x}_i)\right\},\mu_{B^k}(\bar{y}^r)\right)\right\}} \tag{3.38}$$

Now, we generalize both approaches described in points a) and b) and propose a general architecture of NFIS. It is easily seen that systems (3.35) and (3.38) can be presented in the form

$$\bar{y} = f(\bar{\mathbf{x}}) = \frac{\sum_{r=1}^{N} \bar{y}^r \cdot \mathrm{agr}_r(\bar{\mathbf{x}}, \bar{y}^r)}{\sum_{r=1}^{N} \mathrm{agr}_r(\bar{\mathbf{x}}, \bar{y}^r)} \tag{3.39}$$

where

$$\mathrm{agr}_r(\bar{\mathbf{x}}, \bar{y}^r) = \begin{cases} \overset{N}{\underset{k=1}{S}} \{I_{k,r}(\bar{\mathbf{x}}, \bar{y}^r)\}, \\ \text{for } \textit{the Mamdani approach} \\ \overset{N}{\underset{k=1}{T}} \{I_{k,r}(\bar{\mathbf{x}}, \bar{y}^r)\}, \\ \text{for } \textit{the logical approach} \end{cases} \tag{3.40}$$

$$I_{k,r}(\bar{\mathbf{x}}, \bar{y}^r) = \begin{cases} T\{\tau_k(\bar{\mathbf{x}}), \mu_{B^k}(\bar{y}^r)\}, \\ \text{for } \textit{the Mamdani approach} \\ I_{fuzzy}(\tau_k(\bar{\mathbf{x}}), \mu_{B^k}(\bar{y}^r)), \\ \text{for } \textit{the logical approach} \end{cases} \tag{3.41}$$

Moreover, the firing strength of rules has already been defined by

$$\tau_k(\bar{\mathbf{x}}) = \overset{n}{\underset{i=1}{T}} \{\mu_{A_i^k}(\bar{x}_i)\}$$

The general architecture of system (3.39) is depicted in Fig. 3.10.

Remark 3.2: It should be emphasized that formula (3.39) and the scheme depicted in Fig. 3.10 are applicable to all the systems, flexible and nonflexible, studied in this book with different definitions of $\mathrm{agr}_r(\bar{\mathbf{x}}, \bar{y}^r)$ and $I_{k,r}(\bar{\mathbf{x}}, \bar{y}^r)$. The nonflexible systems are described by (3.39), (3.40), (3.41) and (3.11), whereas the flexible systems by (3.39) and $\mathrm{agr}_r(\bar{\mathbf{x}}, \bar{y}^r)$, $I_{k,r}(\bar{\mathbf{x}}, \bar{y}^r)$, $\tau_k(\bar{\mathbf{x}})$ defined in Chapters 5 and 6. How we define the aggregation operator $\mathrm{agr}_r(\bar{\mathbf{x}}, \bar{y}^r)$ and the implication operator $I_{k,r}(\bar{\mathbf{x}}, \bar{y}^r)$, depends on the particular class of the system.

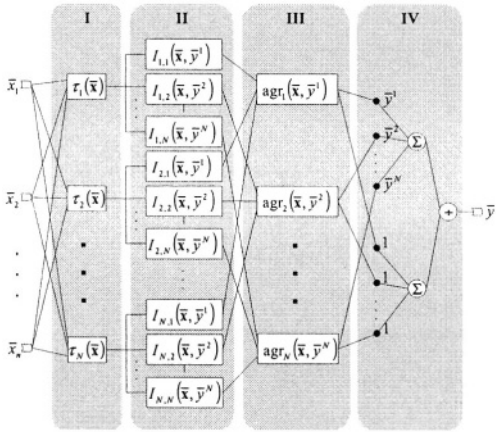

Fig. 3.10. General architecture of NFIS studied in the book (flexible and nonflexible)

Remark 3.3: If an S-implication is used, i.e.

$$I_{k,r}\left(\overline{\mathbf{x}},\overline{y}^{r}\right)=S\left\{N\left(\mu_{A^{k}}\left(\overline{\mathbf{x}}\right)\right),\mu_{B^{k}}\left(\overline{y}^{r}\right)\right\} \tag{3.42}$$

then the aggregated output fuzzy set $B'\subseteq\mathbf{Y}$ is given by

$$\mu_{B'}\left(\overline{y}^{r}\right)=\underset{k=1}{\overset{N}{T}}\left\{I_{k,r}\left(\overline{\mathbf{x}},\overline{y}^{r}\right)\right\}=\underset{k=1}{\overset{N}{T}}\left\{S\left\{N\left(\mu_{A^{k}}\left(\overline{\mathbf{x}}\right)\right),\mu_{B^{k}}\left(\overline{y}^{r}\right)\right\}\right\} \tag{3.43}$$

Consequently, formula (3.38) becomes

$$\overline{y}=\frac{\sum_{r=1}^{N}\overline{y}^{r}\cdot\underset{k=1}{\overset{N}{T}}\left\{S\left\{N\left(\underset{i=1}{\overset{n}{T}}\left\{\mu_{A_{i}^{k}}\left(\overline{x}_{i}\right)\right\}\right),\mu_{B^{k}}\left(\overline{y}^{r}\right)\right\}\right\}}{\sum_{r=1}^{N}\underset{k=1}{\overset{N}{T}}\left\{S\left\{N\left(\underset{i=1}{\overset{n}{T}}\left\{\mu_{A_{i}^{k}}\left(\overline{x}_{i}\right)\right\}\right),\mu_{B^{k}}\left(\overline{y}^{r}\right)\right\}\right\}} \tag{3.44}$$

Remark 3.4: We will explain how to modify formula (3.39) to solve multi-classification problems. Let $[x_1,...,x_n]$ be the vector of features of an object v. Let $\Omega=\{\omega_1,...,\omega_M\}$ be a set of classes. The knowledge is represented by a set of N rules in the form

$$R^{(k)}:\begin{cases}\text{IF} & x_1 \text{ is } A_1^k \text{ AND}\\ & x_2 \text{ is } A_2^k \text{ AND } \ldots\\ & x_n \text{ is } A_n^k\\ \text{THEN} & v\in\omega_1\left(z_1^k\right),\\ & v\in\omega_2\left(z_2^k\right), \quad \ldots,\\ & v\in\omega_M\left(z_M^k\right),\end{cases} \tag{3.45}$$

where z_j^k, $j = 1,..., M$, $k = 1,\ldots,N$, are interpreted as "support" for class ω_j given by rule $R^{(k)}$. We will now redefine description (3.45). Let us introduce vector $\mathbf{z} = [z_1,...,z_M]$, where z_j, $j = 1,\ldots,M$, is the "support" for class ω_j given by all M rules. We can scale the support values to the interval [0,1], so that z_j is the membership degree of an object ν to class ω_j according to all M rules. The rules are represented by

$$R^{(k)}: \begin{cases} \text{IF} & x_1 & \text{is} & A_1^k & \text{AND} & \\ & x_2 & \text{is} & A_2^k & \text{AND} & \ldots \\ & x_n & \text{is} & A_n^k & & \\ \text{THEN} & z_1 & \text{is} & B_1^k & \text{AND} & \\ & z_2 & \text{is} & B_2^k & \text{AND} & \ldots \\ & z_M & \text{is} & B_M^k & & \end{cases} \tag{3.46}$$

and formula (3.39) adopted for classification takes the form

$$\bar{z}_j = \frac{\sum_{r=1}^{N} \bar{z}_j^r \operatorname{agr}_r\left(\bar{\mathbf{x}}, \bar{z}_j^r\right)}{\sum_{r=1}^{N} \operatorname{agr}_r\left(\bar{\mathbf{x}}, \bar{z}_j^r\right)} \tag{3.47}$$

where $\bar{z}_j^r$ are centers of fuzzy sets $B_j^r, j = 1,\ldots,M, r = 1,..,N$.

Remark 3.5: The general description (3.39) is not applicable to the Takagi-Sugeno model [95] given by the following fuzzy rules

$$R^{(k)}: \begin{cases} \text{IF} & x_1 & \text{is} & A_1^k & \text{AND} & \\ & x_2 & \text{is} & A_2^k & \text{AND} & \ldots \\ & x_n & \text{is} & A_n^k & & \\ \text{THEN} & y^k = f^{(k)}(x_1, x_2, \ldots, x_n) & & & & \end{cases} \tag{3.48}$$

In this model, the consequences, contrary to model (3.9), are not fuzzy sets; they are linear or non-linear functions of inputs. The aggregation in the Takagi-Sugeno model is described by formula

$$\bar{y} = f(\bar{\mathbf{x}}) = \frac{\sum_{k=1}^{N} f^{(k)}(\bar{\mathbf{x}}) \cdot \mu_{A^k}(\bar{\mathbf{x}})}{\sum_{k=1}^{N} \mu_{A^k}(\bar{\mathbf{x}})} \tag{3.49}$$

3.6. DATA SETS USED IN THE BOOK

The data sets used to evaluate the performance of the neuro-fuzzy systems developed in this book are listed in Table 3.2. It shows the name of the data, the problem to be solved (classification or approximation), the number of inputs, the length of the training sequence and the length of the testing sequence.

Table 3.2 Benchmark data specification

Data	Problem	Inputs	Training data	Testing data
Box and Jenkins Gas Furnace	approximation	6	296	-
Chemical Plant	approximation	3	70	-
Glass Identification	classification	9	150	64
Ionosphere	classification	33	246	105
Iris	classification	4	105	45
Modeling of Static Nonlinear Function (HANG)	approximation	2	50	-
Monk 1	classification	6	124	432
Monk 2	classification	6	169	432
Monk 3	classification	6	124	432
Nonlinear Dynamic Plant	approximation	2	200	200
Pima Indians Diabetes	classification	8	576	192
Rice Taste	approximation	5	75	30
Wine Recognition	classification	13	125	53
Wisconsin Breast Cancer	classification	9	478	205

We will now describe in detail the problems which appear in the first column of Table 3.2.

Box and Jenkins Gas Furnace problem [5]

The Box and Jenkins Gas Furnace data consists of 296 measurements of the gas furnace system: the input measurement $u(k)$ is the gas flow rate into the furnace and the output measurement $y(k)$ is the CO_2 concentration in the outlet gas. The sampling interval is 9 s. We wish to determine a fuzzy model of the gas furnace system. In our simulations we assume that $y(t) = f(y(t-1), y(t-2), y(t-3), y(t-4), u(t-1), u(t-2))$.

Chemical Plant problem [92]

We deal with a model of an operator's control of a chemical plant. The plant produces polymers by polymerisating some monomers. Since the start-up of the plant is very complicated, men have to perform the manual operations at the plant. Three continuous inputs are chosen for controlling the system: monomer concentration, change of monomer concentration and monomer flow rate. The output is the set point for the monomer flow rate.

Glass Identification problem [98]

The Glass Identification problem contains 214 instances and each instance is described by nine attributes (RI: refractive index, Na: sodium, Mg: magnesium, Al: aluminium, Si: silicon, K: potassium, Ca: calcium, Ba: barium, Fe: iron). All attributes are continuous. There are two classes: the window glass and the non-window glass. In our experiments, all sets are divided into a learning sequence (150 sets) and a testing sequence (64 sets). The study of the classification of the types of glass was motivated by criminological investigation. At the scene of the crime, the glass left can be used as evidence if it is correctly identified.

Ionosphere problem [98]

This radar data was collected by a system in Goose Bay, Labrador. This system consists of a phased array of 16 high-frequency antennas with total transmitted power in the order of 6.4 kW. The targets were free electrons in the ionosphere. The database is composed of 33 continuous attributes plus the class variable, using 351 examples. In our experiments, all sets are divided into a learning sequence (246 sets) and a testing sequence (105 sets).

Iris problem [98]

The Iris data is a common benchmark in classification and pattern classification studies. It contains 50 measurements of four features (sepal length in cm, sepal width in cm, petal length in cm, petal width in cm) from each of the following three species: iris setosa, iris versicolor, and iris virginica. In our experiments, all sets are divided into a learning sequence (105 sets) and a testing sequence (45 sets).

Modeling of Static Nonlinear Function (HANG) problem [92]

The problem is to approximate a nonlinear function given by

$$y(x_1, x_2) = \left(1 + x_1^{-2} + x_2^{-1.5}\right)^2 \tag{3.50}$$

We obtained 50 input-output data by sampling the input range $x_1, x_2 \in [1,5]$.

The Three Monks' problems [98]

The three monks' problems are artificial, small problems designed to test machine learning algorithms. Each of the three monks' problems requires determining whether an object described by six features (head shape, body shape, is smiling, holding, jacket colour, has tie) is a monk or not.

There are 432 combinations of the six symbolic attributes. In the first problem (Monk1), 124 cases were randomly selected for the training set, in the second problem (Monk2) 169 cases, and in the third problem (Monk3) 122 cases, of which 5% were misclassifications introducing some noise in the data.

Nonlinear Dynamic Plant problem [103]

We consider the second-order nonlinear plant described by

$$y(k) = g(y(k-1), y(k-2)) + u(k) \tag{3.51}$$

with

$$g(y(k-1), y(k-2)) = \frac{y(k-1)y(k-2)(y(k-1)-0.5)}{1+y^2(k-1)+y^2(k-2)} \tag{3.52}$$

The goal is to approximate the nonlinear component $g(y(k-1),y(k-2))$ of the plant with a fuzzy model. In [103], 400 simulated data were generated from the plant model (3.52). Starting from the equilibrium state (0,0), 200 samples of the identification data were obtained with a random input signal $u(k)$ uniformly distributed in [–1.5,1.5] , followed by 200 samples of evaluation data obtained using a sinusoidal input signal $u(k) = \sin(2\pi k/25)$.

Pima Indians Diabetes problem [98]

The Pima Indians Diabetes data contains two classes, eight attributes (number of times pregnant, plasma glucose concentration in an oral glucose tolerance test, diastolic blood pressure (mm Hg), triceps skin fold thickness (mm), 2-hour serum insulin (mu U/ml), body mass index (weight in $\text{kg/(height in m)}^2$), diabetes pedigree function, age (years)). We consider 768 instances, 500 (65.1%) healthy and 268 (34.9%) diabetes cases. All patients were females at least 21 years old of Pima Indian heritage, living near Phoenix, Arizona. It should be noted that about 33% of this population suffers from diabetes. In our experiments, all sets are divided into a learning sequence (576 sets) and a testing sequence (192 sets).

Rice Taste problem [30, 66]

The Rice Taste data contains 105 instances and each instance is described by five attributes (inputs): flavour, appearance, taste, stickiness, toughness, and overall evaluation. The output is the overall evaluation of the rice taste. The problem is to find a non-linear mapping between inputs and

the output. In simulations the inputs-output pairs of the rice taste data were normalized in the interval [0,1]

Wine Recognition problem [98]

The Wine data contains a chemical analysis of 178 wines grown in the same region of Italy but derived from three different vineyards. The 13 continuous attributes available for classification are: alcohol, malic acid, ash, alkalinity of ash, magnesium, total phenols, flavanoids, nonflavanoid phenols, proanthocyanins, colour intensity, hue, OD280/OD315 of diluted wines and proline. In our experiments all sets are divided into a learning sequence (125 sets) and a testing sequence (53 sets). The problem is to classify wine samples based on the learning sequence.

Wisconsin Breast Cancer problem [98]

The Wisconsin Breast Cancer data contains 699 instances (of which 16 instances have a single missing attribute) and each instance is described by nine attributes (clump thickness, uniformity of cell size, uniformity of cell shape, marginal adhesion, single epithelial cell size, bare nuclei, bland chromatin, normal nucleoli, mitoses). We removed those 16 instances and used the remaining 683 instances. Out of 683 data samples, 444 cases represent benign breast cancer and 239 cases describe malignant breast cancer. The problem is to classify whether a new case is a benign (class 1) or malignant (class 2) type of cancer. In our experiments, all sets are divided into a learning sequence (478 sets) and a testing sequence (205 sets).

Remark 3.6: Because of the space limitation in this book, the results in Chapters 5-8 will be illustrated on four benchmarks given in Table 3.2:

- Box and Jenkins Gas Furnace,
- Glass Identification,
- Modeling of Static Nonlinear Function (HANG),
- Wisconsin Breast Cancer.

However, in Chapter 9 all the benchmarks will be used for the comparison of flexible neuro-fuzzy systems.

3.7. SUMMARY AND DISCUSSION

Several authors (e.g., Jager [32], Mendel [58]) reported problems with the application of logical implications to NFIS. A major problem is caused by the indeterminant part of the membership function. We illustrated such a situation in Section 3.4 (see e.g. Fig. 3.7-a, b). The indicated problem can

be easily resolved by the application of a modified center of gravity defuzzifier

$$\bar{y} = \frac{\int\limits_{Y_\alpha} y \cdot (\mu_{B'}(y) - \alpha) dy}{\int\limits_{Y_\alpha} (\mu_{B'}(y) - \alpha) dy} \tag{3.53}$$

where

$$\mathbf{Y}_\alpha = \{y \in \mathbf{Y} | \mu_{B'}(y) \geq \alpha\} \tag{3.54}$$

value $\alpha \in [0,1]$ describes the indeterminacy that accompanies the corresponding part of the information. It is easily seen that in order to eliminate the indeterminant part of the membership function $\mu_{B'}(y)$, the informative part has to be parallely shifted downward by value a. Neuro-fuzzy inference systems of a logical-type with defuzzifier (3.53) have been studied by Czogała and Łęski [15].

3.8. PROBLEMS

Problem 3.1. Consider fuzzy sets A', A and B depicted in Fig. 3.11.

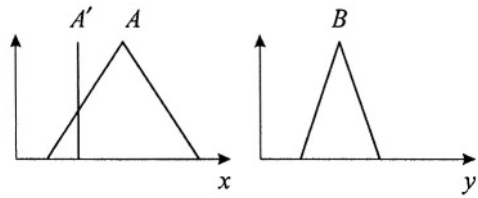

Fig. 3.11. Illustration to Problem 3.1

Find fuzzy set B' as a result of fuzzy reasoning according to the Mamdani and Larsen rules.

Problem 3.2. Repeat Problem 3.1 using the binary and Łukasiewicz fuzzy implications.

Problem 3.3. Consider a fuzzy system with a single input and two rules depicted in Fig. 3.12. Perform a fuzzy inference based on the Mamdani rule. Choose an appropriate triangular norm for the aggregation.

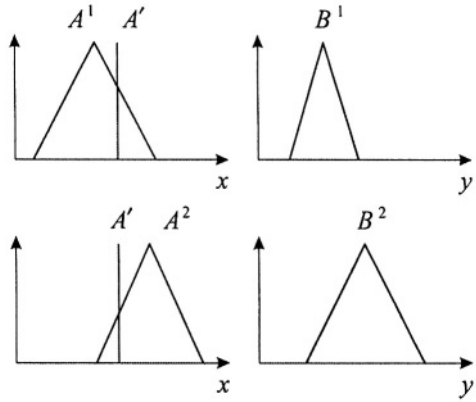

Fig. 3.12. Illustration to Problem 3.3

Problem 3.4. Repeat Problem 3.3 using the binary fuzzy implication.

Problem 3.5. Consider a Takagi-Sugeno fuzzy system with the following 3 rules:

$$\text{IF } x \text{ is } A^1 \text{ THEN } y = 30 + x$$

$$\text{IF } x \text{ is } A^2 \text{ THEN } y = 2x$$

$$\text{IF } x \text{ is } A^3 \text{ THEN } y = 70 - 3x$$

The fuzzy sets A^1, A^2 and A^3 are depicted in Fig. 3.13 Find the mapping $\overline{y} = f(\overline{x})$ and determine $\overline{y}$ for $\overline{x} = 10$.

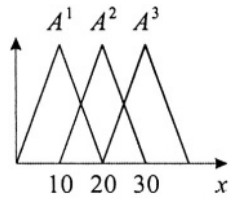

Fig. 3.13. Illustration to Problem 3.5

Chapter 4

FLEXIBILITY IN FUZZY SYSTEMS

4.1. INTRODUCTION

In the previous works on neuro-fuzzy systems it was assumed that fuzzy inference (Mamdani or logical) was fixed in advance and during the design process only the parameters of the membership functions were optimized to meet the design requirements. On the other hand it is well known that introducing additional parameters to be tuned in the system usually improves its performance. The system is able to better represent the patterns encoded in the data. In this chapter we present various concepts leading to the designing of flexible neuro-fuzzy systems, characterized by many parameters determined in the process of learning. Based on our propositions various flexible structures will be developed in Chapters 5-8.

4.2. WEIGHTED TRIANGULAR NORMS

Any construction of fuzzy systems relies on triangular norms. They are used to connect the antecedents in the individual rules, aggregate the rules and express the relation between the antecedents and the consequent. In this section we incorporate the weights into the construction of triangular norms. The idea of weighted triangular norms will be explained on the two-dimensional case and next extended to the multidimensional case. The weighted t-norm in the two-dimensional case is defined as follows

$$T^*\{a_1, a_2; w_1, w_2\} = T\{1 - w_1(1 - a_1), 1 - w_2(1 - a_2)\} \tag{4.1}$$

Parameters a_1 and a_2 can be interpreted as antecedents of the rule. The weights w_1 and w_2 are corresponding certainties (credibilities) of both antecedents in (4.1).

Observe that:

(i) If $w_1 = w_2 = 1$ then the weighted t-norm is reduced to the standard t-norm. In the context of linguistic values we assign the truth to both antecedents a_1 and a_2 of the rule.

(ii) If $w_1 = 0$ then

$$\begin{aligned} T^*\{a_1, a_2; 0, w_2\} &= T\{1, 1 - w_2(1 - a_2)\} \\ &= 1 - w_2(1 - a_2) \end{aligned} \tag{4.2}$$

Therefore, antecedent a_1 is discarded since its certainty is equal to 0. Similarly if $w_2 = 0$ then antecedent a_2 vanishes, i.e.

$$\begin{aligned} T^*\{a_1, a_2; w_1, 0\} &= T\{1 - w_1(1 - a_1), 1\} \\ &= 1 - w_1(1 - a_1) \end{aligned} \tag{4.3}$$

(iii) If $0 < w_1 < 1$ and $0 < w_2 < 1$ then we assume a partial certainty of antecedents a_1 and a_2.

The t-conorm corresponding to the t-norm (4.1) is defined as follows

$$S^*\{a_1, a_2; w_1, w_2\} = S\{w_1 a_1, w_2 a_2\} \tag{4.4}$$

In the same spirit we propose the weighted triangular norms

$$S^*\{a_1, a_2; w_1^{\text{agr}}, w_2^{\text{agr}}\} = S\{w_1^{\text{agr}} a_1, w_2^{\text{agr}} a_2\} \tag{4.5}$$

and

$$T^*\{a_1, a_2; w_1^{\text{agr}}, w_2^{\text{agr}}\} = T\{1 - w_1^{\text{agr}}(1 - a_1), 1 - w_2^{\text{agr}}(1 - a_2)\} \tag{4.6}$$

to aggregate individual rules (for $N = 2$) in Mamdani-type and logical-type systems, respectively. The weights w_1 and w_2 in (4.1), as well as w_1^{agr} and w_2^{agr} in (4.5) or (4.6), can be found in the process of learning subject to constraints w_1, w_2, w_1^{agr}, $w_2^{\text{agr}} \in [0,1]$.

In the next chapters we will apply the weighted t-norm

$$T^*\{a_1, \ldots, a_n; w_1^r, \ldots, w_n^r\} = \mathop{T}_{i=1}^{n}\{1 - w_i^r(1 - a_i)\} \tag{4.7}$$

to connect the antecedents in each rule, $k = 1,...,N$, and the weighted t-norm and t-conorm:

$$T^*\left\{a_1,\ldots,a_N;w_1^{\text{agr}},\ldots,w_N^{\text{agr}}\right\}=\underset{k=1}{\overset{N}{T}}\left\{1-w_k^{\text{agr}}\left(1-a_k\right)\right\} \tag{4.8}$$

$$S^*\left\{a_1,\ldots,a_N;w_1^{\text{agr}},\ldots,w_N^{\text{agr}}\right\}=\underset{k=1}{\overset{N}{S}}\left\{w_k^{\text{agr}}a_k\right\} \tag{4.9}$$

to aggregate the individual rules in the logical and Mamdani models, respectively. It is easily seen that formula (4.7) can be applied to the evaluation of an importance of input linguistic values, and the weighted t-norm (4.8) or t-conorm (4.9) to a selection of important rules. The results will be depicted in the form of diagrams. In Fig. 4.1 we show an example of a diagram for a fuzzy system having four rules ($N = 4$) and two inputs ($n = 2$) described by:

$$R^1:\left[\text{IF } x_1 \text{ is } A_1^1\left(w_{1,1}^{\tau}\right)\text{AND } x_2 \text{ is } A_2^1\left(w_{2,1}^{\tau}\right)\text{THEN } y \text{ is } B^1\right]w_1^{\text{agr}}$$

$$R^2:\left[\text{IF } x_1 \text{ is } A_1^2\left(w_{1,2}^{\tau}\right)\text{AND } x_2 \text{ is } A_2^2\left(w_{2,2}^{\tau}\right)\text{THEN } y \text{ is } B^2\right]w_2^{\text{agr}}$$

$$R^3:\left[\text{IF } x_1 \text{ is } A_1^3\left(w_{1;3}^{\tau}\right)\text{AND } x_2 \text{ is } A_2^3\left(w_{2,3}^{\tau}\right)\text{THEN } y \text{ is } B^3\right]w_3^{\text{agr}}$$

$$R^4:\left[\text{IF } x_1 \text{ is } A_1^4\left(w_{1,4}^{\tau}\right)\text{AND } x_2 \text{ is } A_2^4\left(w_{2,4}^{\tau}\right)\text{THEN } y \text{ is } B^4\right]w_4^{\text{agr}}$$

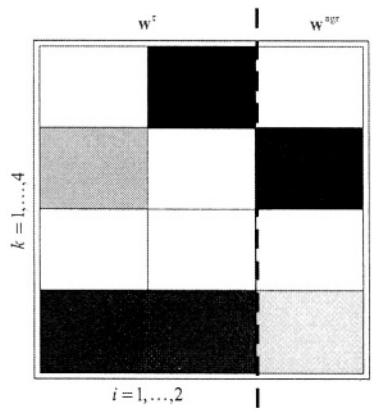

Fig. 4.1. Exemplary weights representation in a fuzzy system with four rules and two inputs (dark areas correspond to low values of weights and vice versa)

Observe that the third rule is "weaker" than the others and the linguistic value A_2^4 corresponds to a low value of $w_{2,4}^{\tau}$. The designing of neuro-fuzzy systems should be a compromise between the accuracy of the model and its transparency. The measure of accuracy is usually the RMSE-criterion (approximation problems) and the percentage of correct or wrong decisions (classification problems). The measure of transparency is the number and

form of fuzzy rules obtained. It was mentioned by several authors (see e.g. [2, 28]) that the lack of transparency is a major drawback of many neuro-fuzzy systems. Most designers focus their effort on approximation accuracy, while the issue of transparency has received less attention. In this context our method of weighted triangular norms seems to be a promising tool for extracting both transparent and accurate rule-based knowledge from empirical data. More specifically, diagrams (weights representation like that shown in Fig. 4.1) can be used for analysis and pruning of the fuzzy-rule bases in all the simulation examples. Note that our application of weights in NFIS is different from those studied in [19, 31, 63, 97, 111].

Example 4.1. (An example of Zadeh triangular norms with weighted arguments)
The Zadeh triangular norms with weighted arguments are based on classical Zadeh triangular norms given by (2.43) and (2.44). The Zadeh t-norm with weighted arguments is described as follows

$$T^*\{a_1, a_2; w_1, w_2\} = \min\{1 - w_1(1 - a_1), 1 - w_2(1 - a_2)\} \tag{4.10}$$

The 3D plots of function (4.10) are depicted in Fig. 4.2.

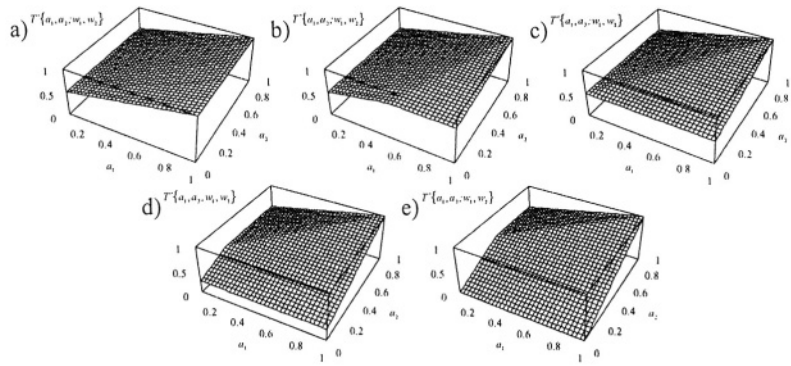

Fig. 4.2. 3D plots of function (4.10)
for $w_1 = 0.50$ and a) $w_2 = 0.00$, b) $w_2 = 0.25$, c) $w_2 = 0.50$, d) $w_2 = 0.75$, e) $w_2 = 1.00$

The Zadeh t-conorm with weighted arguments is given by

$$S^*\{a_1, a_2; w_1, w_2\} = \max\{w_1 a_1, w_2 a_2\} \tag{4.11}$$

The 3D plots of function (4.11) are presented in Fig. 4.3.

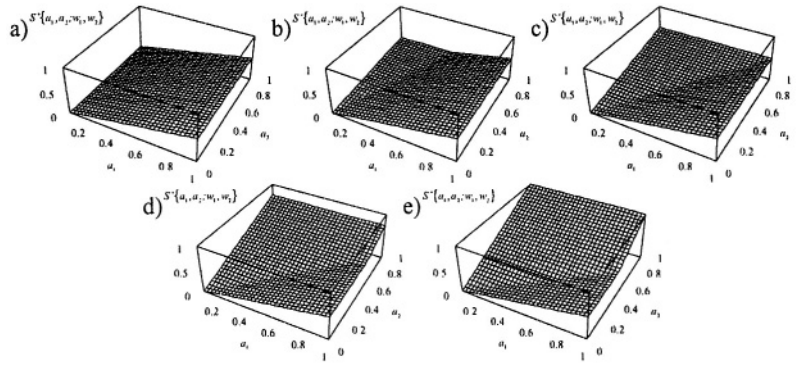

Fig. 4.3. 3D plots of function (4.11) for $w_1 = 0.50$ and a) $w_2 = 0.00$, b) $w_2 = 0.25$, c) $w_2 = 0.50$, d) $w_2 = 0.75$, e) $w_2 = 1.00$

Example 4.2. (An example of algebraic triangular norms with weighted arguments)

The algebraic triangular norms with weighted arguments are based on classical algebraic triangular norms given by (2.47) and (2.48). The algebraic t-norm with weighted arguments is described as follows

$$T^*\{a_1, a_2; w_1, w_2\} = (1 - w_1(1 - a_1))(1 - w_2(1 - a_2)) \tag{4.12}$$

The 3D plots of function (4.12) are depicted in Fig. 4.4.

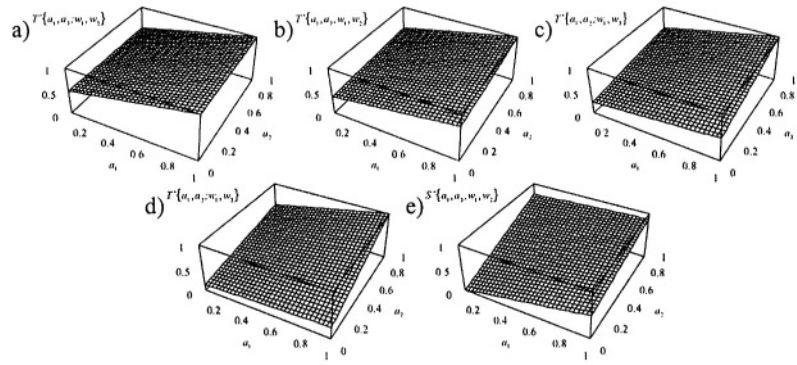

Fig. 4.4. 3D plots of function (4.12) for $w_1 = 0.50$ and a) $w_2 = 0.00$, b) $w_2 = 0.25$, c) $w_2 = 0.50$, d) $w_2 = 0.75$, e) $w_2 = 1.00$

The algebraic t-conorm with weighted arguments is given by

$$S^*\{a_1, a_2; w_1, w_2\} = w_1 a_1 + w_2 a_2 - w_1 a_1 w_2 a_2 \tag{4.13}$$

The 3D plots of function (4.13) are presented in Fig. 4.5.

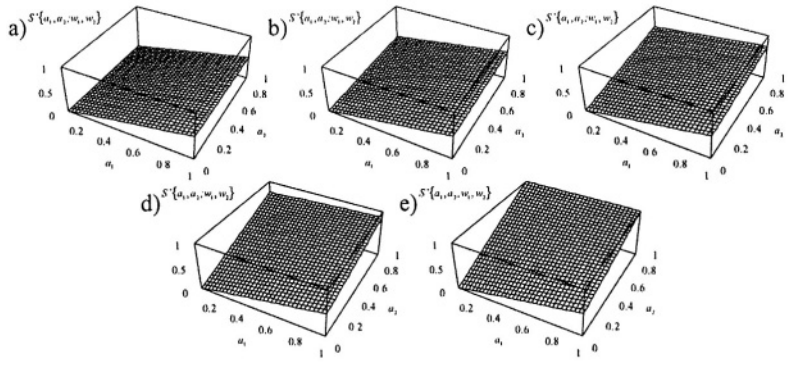

Fig. 4.5. 3D plots of function (4.13)
for $w_1 = 0.50$ and a) $w_2 = 0.00$, b) $w_2 = 0.25$, c) $w_2 = 0.50$, d) $w_2 = 0.75$, e) $w_2 = 1.00$

Example 4.3. (An example of bounded triangular norms with weighted arguments)
The bounded triangular norms with weighted arguments are based on classical bounded triangular norms given by (2.51) and (2.52). The bounded t-norm with weighted arguments is described as follows

$$T^*\{a_1, a_2; w_1, w_2\} = \max\{1 - w_1(1 - a_1) + 1 - w_2(1 - a_2) - 1, 0\} \qquad (4.14)$$

The 3D plots of function (4.14) are depicted in Fig. 4.6.

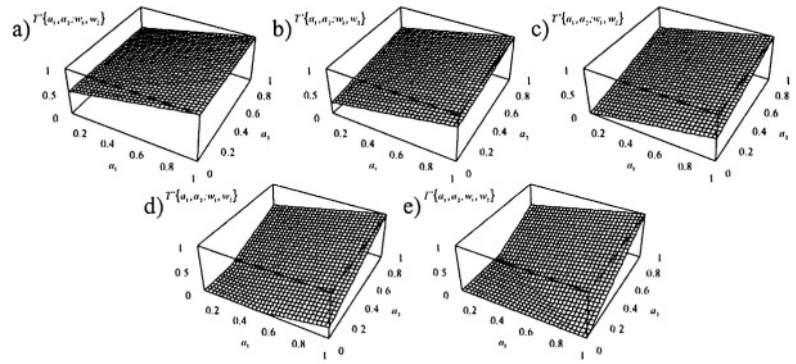

Fig. 4.6. 3D plots of function (4.14)
for $w_1 = 0.50$ and a) $w_2 = 0.00$, b) $w_2 = 0.25$, c) $w_2 = 0.50$, d) $w_2 = 0.75$, e) $w_2 = 1.00$

The bounded t-conorm with weighted arguments is given by

$$S^*\{a_1, a_2; w_1, w_2\} = \min\{w_1 a_1 + w_2 a_2, 1\} \qquad (4.15)$$

The 3D plots of function (4.15) are presented in Fig. 4.7.

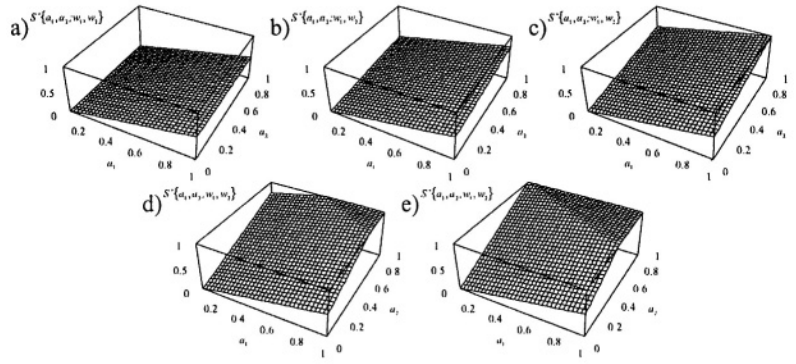

Fig. 4.7. 3D plots of function (4.15)
for $w_1 = 0.50$ and a) $w_2 = 0.00$, b) $w_2 = 0.25$, c) $w_2 = 0.50$, d) $w_2 = 0.75$, e) $w_2 = 1.00$

Example 4.4. (An example of drastic triangular norms with weighted arguments)
The drastic triangular norms with weighted arguments are based on classical drastic triangular norms given by (2.55) and (2.56). The drastic t-norm with weighted arguments is described as follows

$$T^*\{a_1, a_2; w_1, w_2\} = \begin{cases} 1 - w_1(1 - a_1) & \text{if } 1 - w_2(1 - a_2) = 1 \\ 1 - w_2(1 - a_2) & \text{if } 1 - w_1(1 - a_1) = 1 \\ 0 & \textit{otherwise} \end{cases} \tag{4.16}$$

The 3D plots of function (4.16) are depicted in Fig. 4.8.

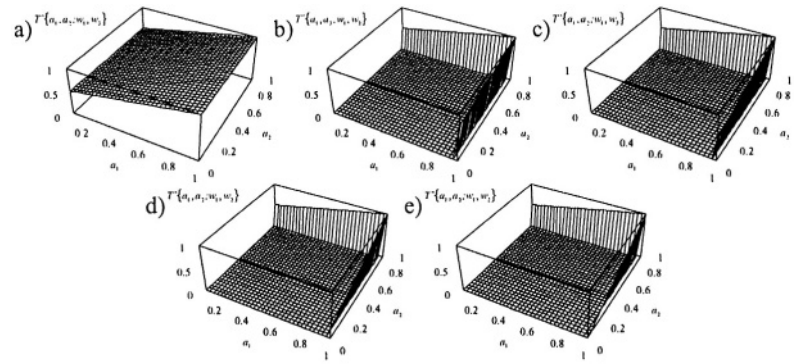

Fig. 4.8. 3D plots of function (4.16)
for $w_1 = 0.50$ and a) $w_2 = 0.00$, b) $w_2 = 0.25$, c) $w_2 = 0.50$, d) $w_2 = 0.75$, e) $w_2 = 1.00$

The drastic t-conorm with weighted arguments is given by

$$S^*\{a_1,a_2;w_1,w_2\}=\begin{cases} w_1a_1 & \text{if} \quad w_2a_2=0 \\ w_2a_2 & \text{if} \quad w_1a_1=0 \\ 1 & \text{if} \quad otherwise \end{cases} \tag{4.17}$$

The 3D plots of function (4.17) are presented in Fig. 4.9.

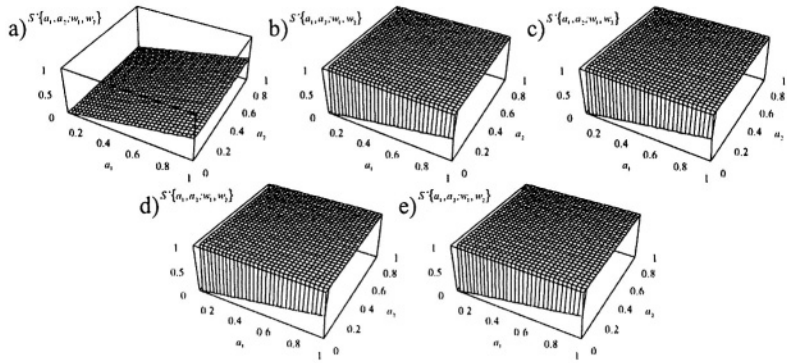

Fig. 4.9. 3D plots of function (4.17)
for $w_1 = 0.50$ and a) $w_2 = 0.00$, b) $w_2 = 0.25$, c) $w_2 = 0.50$, d) $w_2 = 0.75$, e) $w_2 = 1.00$

4.3. SOFT FUZZY NORMS

In this sections we recall a concept of soft fuzzy norms proposed by Yager and Filev [108]. Let $a_1,\ldots,a_n$ be numbers in the unit interval that are to be aggregated. The soft version of triangular norms suggested by Yager and Filev is defined by

$$\tilde{T}\{\mathbf{a};\alpha\}=(1-\alpha)\frac{1}{n}\sum_{i=1}^{n}a_i+\alpha\overset{n}{\underset{i=1}{T}}\{a_i\} \tag{4.18}$$

and

$$\tilde{S}\{\mathbf{a};\alpha\}=(1-\alpha)\frac{1}{n}\sum_{i=1}^{n}a_i+\alpha\overset{n}{\underset{i=1}{S}}\{a_i\} \tag{4.19}$$

where $\alpha\in[0,1]$. They allow to balance between the arithmetic average aggregator and the triangular norm aggregator depending on parameter α.

Example 4.5. (An example of soft Zadeh triangular norms)
The soft Zadeh triangular norms based on classical Zadeh triangular norms given by (2.43) and (2.44) and formulas (4.18) and (4.19). The soft Zadeh t-norm is described as follows

$$\widetilde{T}\{a_1, a_2; \alpha\} = (1-\alpha)\frac{1}{2}(a_1 + a_2) + \alpha \min\{a_1, a_2\} \tag{4.20}$$

The 3D plots of function (4.20) are depicted in Fig. 4.10.

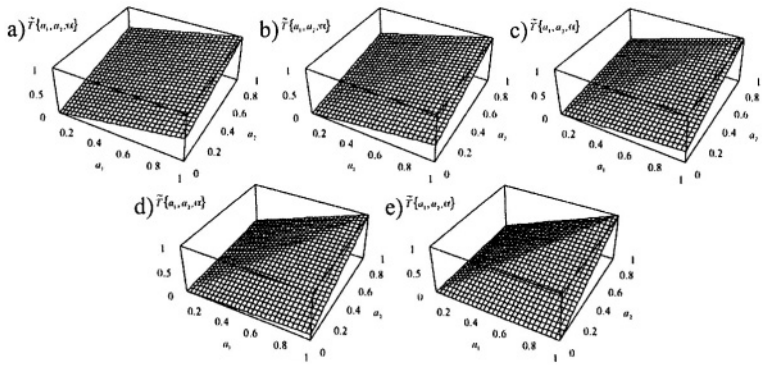

Fig. 4.10. 3D plots of function (4.20)
for a) $\alpha = 0.00$, b) $\alpha = 0.25$, c) $\alpha = 0.50$, d) $\alpha = 0.75$, e) $\alpha = 1.00$

The soft Zadeh t-conorm is given by

$$\widetilde{S}\{a_1, a_2; \alpha\} = (1-\alpha)\frac{1}{2}(a_1 + a_2) + \alpha \max\{a_1, a_2\} \tag{4.21}$$

The 3D plots of function (4.21) are presented in Fig. 4.11.

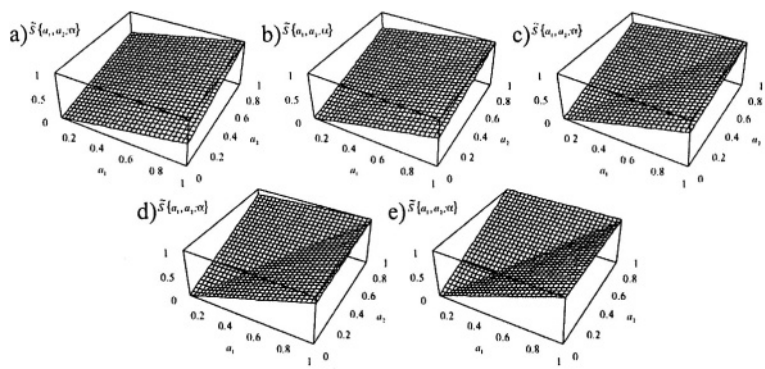

Fig. 4.11. 3D plots of function (4.21)
for a) $\alpha = 0.00$, b) $\alpha = 0.25$, c) $\alpha = 0.50$, d) $\alpha = 0.75$, e) $\alpha = 1.00$

Example 4.6. (An example of soft algebraic triangular norms)
The soft algebraic triangular norms are based on classical algebraic triangular norms given by (2.47) and (2.48) and formulas (4.18) and (4.19). The soft algebraic t-norm is described as follows

$$\tilde{T}\{a_1, a_2; \alpha\} = (1-\alpha)\frac{1}{2}(a_1 + a_2) + \alpha a_1 a_2 \tag{4.22}$$

The 3D plots of function (4.22) are depicted in Fig. 4.12.

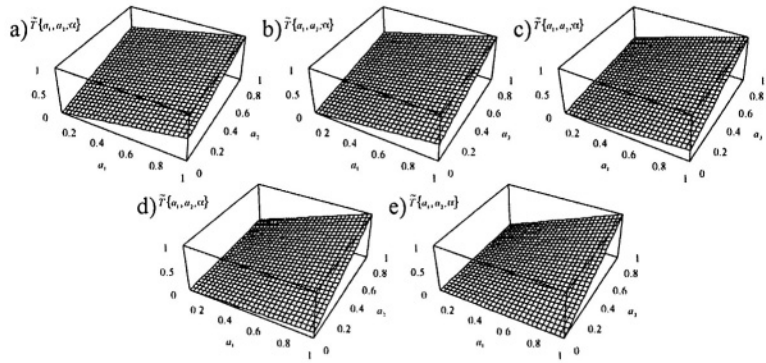

Fig. 4.12. 3D plots of function (4.22)
for a) $\alpha = 0.00$, b) $\alpha = 0.25$, c) $\alpha = 0.50$, d) $\alpha = 0.75$, e) $\alpha = 1.00$

The soft algebraic t-conorm is given by

$$\tilde{S}\{a_1, a_2; \alpha\} = (1-\alpha)\frac{1}{2}(a_1 + a_2) + \alpha(a_1 + a_2 - a_1 a_2) \tag{4.23}$$

The 3D plots of function (4.23) are presented in Fig. 4.13.

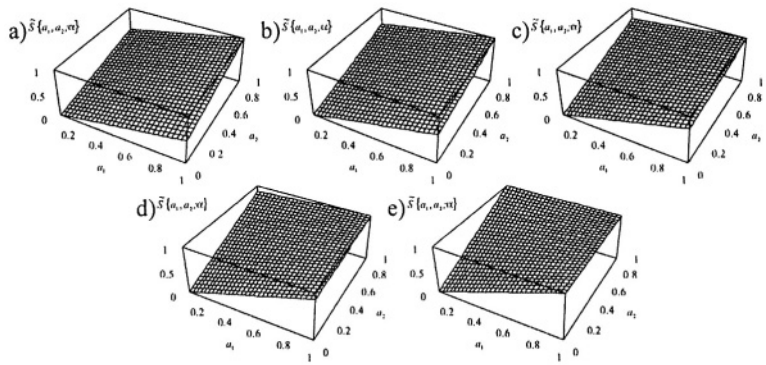

Fig. 4.13. 3D plots of function (4.23)
for a) $\alpha = 0.00$, b) $\alpha = 0.25$, c) $\alpha = 0.50$, d) $\alpha = 0.75$, e) $\alpha = 1.00$

Example 4.7. (An example of soft bounded triangular norms)
The soft bounded triangular norms are based on classical bounded triangular norms given by (2.51) and (2.52) and formulas (4.18) and (4.19). The soft bounded t-norm is described as follows

$$\widetilde{T}\{a_1, a_2; \alpha\} = (1-\alpha)\frac{1}{2}(a_1 + a_2) + \alpha \max\{a_1 + a_2 - 1, 0\} \tag{4.24}$$

The 3D plots of function (4.24) are depicted in Fig. 4.14.

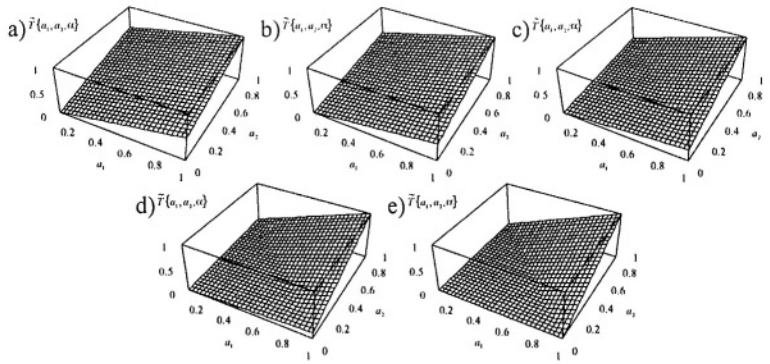

Fig. 4.14. 3D plots of function (4.24)
for a) $\alpha = 0.00$, b) $\alpha = 0.25$, c) $\alpha = 0.50$, d) $\alpha = 0.75$, e) $\alpha = 1.00$

The soft bounded t-conorm is given by

$$\widetilde{S}\{a_1, a_2; \alpha\} = (1-\alpha)\frac{1}{2}(a_1 + a_2) + \alpha \min\{a_1 + a_2, 1\} \tag{4.25}$$

The 3D plots of function (4.25) are presented in Fig. 4.15.

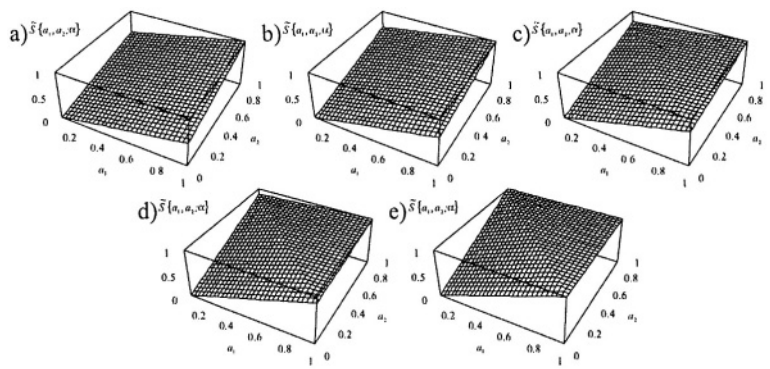

Fig. 4.15. 3D plots of function (4.25)
for a) $\alpha = 0.00$, b) $\alpha = 0.25$, c) $\alpha = 0.50$, d) $\alpha = 0.75$, e) $\alpha = 1.00$

Example 4.8. (An example of soft drastic triangular norms)
The soft drastic triangular norms are based on classical drastic triangular norms are given by (2.55) and (2.56) and formulas (4.18) and (4.19). The soft drastic t-norm is described as follows

$$\widetilde{T}\{a_1, a_2; \alpha\} = (1-\alpha)\frac{1}{2}(a_1 + a_2) + \alpha T_D\{a_1, a_2\} \tag{4.26}$$

The 3D plots of function (4.26) are depicted in Fig. 4.16.

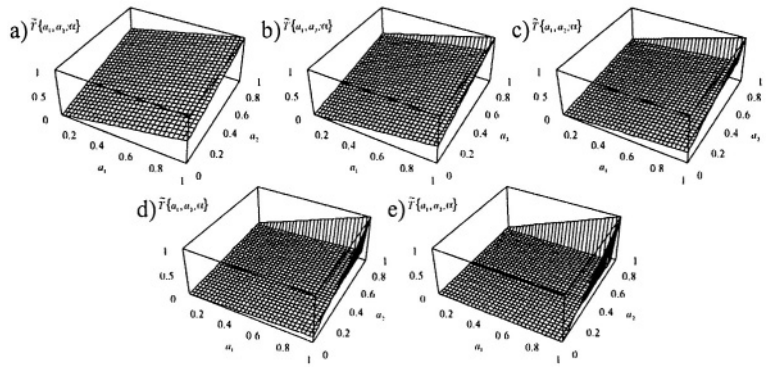

Fig. 4.16. 3D plots of function (4.26)
for a) $\alpha = 0.00$, b) $\alpha = 0.25$, c) $\alpha = 0.50$, d) $\alpha = 0.75$, e) $\alpha = 1.00$

The soft drastic t-conorm is given by

$$\widetilde{S}\{a_1, a_2; \alpha\} = (1-\alpha)\frac{1}{2}(a_1 + a_2) + \alpha S_D\{a_1, a_2\} \tag{4.27}$$

The 3D plots of function (4.27) are presented in Fig. 4.17.

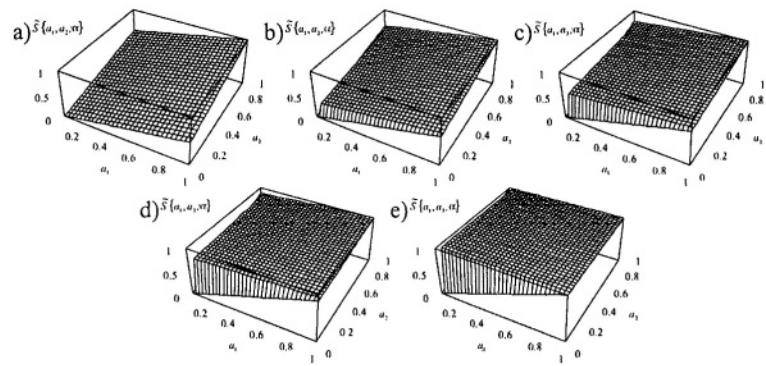

Fig. 4.17. 3D plots of function (4.27)
for a) $\alpha = 0.00$, b) $\alpha = 0.25$, c) $\alpha = 0.50$, d) $\alpha = 0.75$, e) $\alpha = 1.00$

In the same spirit we define the softening of the "engineering implication" by

$$\tilde{I}(a,b;\beta)=(1-\beta)\frac{1}{2}(a+b)+\beta T\{a,b\} \tag{4.28}$$

and softening of the S-fuzzy implication by

$$\tilde{I}(a,b;\beta)=(1-\beta)\frac{1}{2}(1-a+b)+\beta S\{1-a,b\} \tag{4.29}$$

where $\beta\in[0,1]$.

Example 4.9. (An example of soft Kleene-Dienes S-implication)
The soft Kleene-Dienes S-implication is based on the Zadeh t-conorm given by (2.44), the Zadeh negation given by (2.65) and formula (4.29). The soft Kleene-Dienes S-implication is described as follows

$$\tilde{I}(a,b;\beta)=(1-\beta)\frac{1}{2}(1-a+b)+\beta\max\{1-a,b\} \tag{4.30}$$

The 3D plots of function (4.30) are depicted in Fig. 4.18.

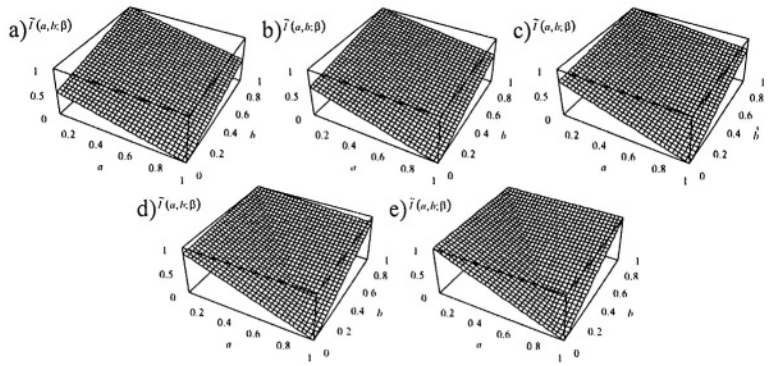

Fig. 4.18. 3D plots of function (4.30)
for a) $\beta=0.00$, b) $\beta=0.25$, c) $\beta=0.50$, d) $\beta=0.75$, e) $\beta=1.00$

Example 4.10. (An example of soft Reichenbach S-implication)
The soft Reichenbach S-implication is based on the algebraic t-conorm given by (2.48), the Zadeh negation given by (2.65) and formula (4.29). The soft Reichenbach S-implication is described as follows

$$\tilde{I}(a,b;\beta)=(1-\beta)\frac{1}{2}(1-a+b)+\beta(1\text{-}a+ab) \tag{4.31}$$

The 3D plots of function (4.31) are depicted in Fig. 4.19.

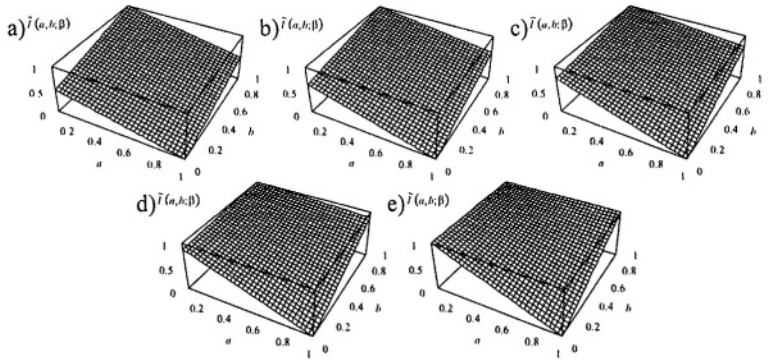

Fig. 4.19. 3D plots of function (4.31)
for a) $\beta = 0.00$, b) $\beta = 0.25$, c) $\beta = 0.50$, d) $\beta = 0.75$, e) $\beta = 1.00$

Example 4.11. (An example of soft Łukasiewicz S-implication)
The soft Łukasiewicz S-implication is based on the bounded t-conorm given by (2.52), the Zadeh negation given by (2.65) and formula (4.29). The soft Łukasiewicz S-implication is described as follows

$$\tilde{I}(a,b;\beta) = (1-\beta)\frac{1}{2}(1-a+b) + \beta \min\{1,1-a+b\} \tag{4.32}$$

The 3D plots of function (4.32) are depicted in Fig. 4.20.

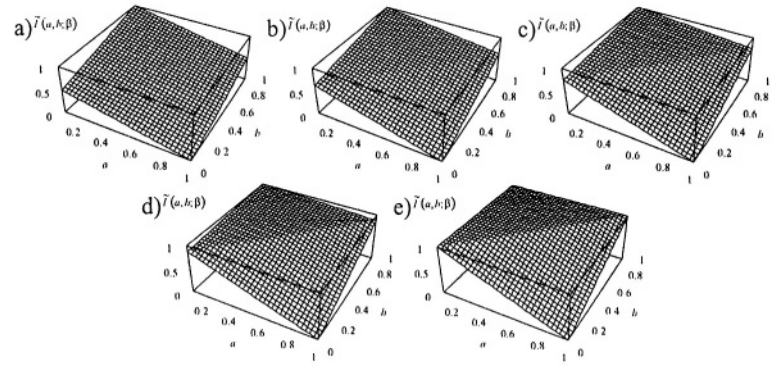

Fig. 4.20. 3D plots of function (4.32)
for a) $\beta = 0.00$, b) $\beta = 0.25$, c) $\beta = 0.50$, d) $\beta = 0.75$, e) $\beta = 1.00$

In the design of Mamdani-type systems we use the following soft triangular norms:

(i) $\tilde{T}_1\{\mathbf{a};\alpha^\tau\} = (1-\alpha^\tau)\frac{1}{n}\sum_{i=1}^{n} a_i + \alpha^\tau \overset{n}{\underset{i=1}{T}}\{a_i\}$ to connect the antecedents,

(ii) $\tilde{T}_2\{b_1, b_2; \alpha^I\} = (1-\alpha^I)\frac{1}{2}(b_1 + b_2) + \alpha^I T\{b_1, b_2\}$ to connect the antecedent and the consequent,

(iii) $\tilde{S}\{\mathbf{c}; \alpha^{agr}\} = (1-\alpha^{agr})\frac{1}{N}\sum_{k=1}^{N} c_k + \alpha^{agr} \overset{N}{\underset{k=1}{S}}\{c_k\}$ to aggregate individual rules

where n is the number of inputs and N is the number of rules.

In the design of logical-type systems based on S-implications we use the following soft triangular norms:

(i) $\tilde{T}_1\{\mathbf{a}; \alpha^\tau\} = (1-\alpha^\tau)\frac{1}{n}\sum_{i=1}^{n} a_i + \alpha^\tau \overset{n}{\underset{i=1}{T}}\{a_i\}$ to connect the antecedents,

(ii) $\tilde{S}\{b_1, b_2; \alpha^I\} = (1-\alpha^I)\frac{1}{2}(1-b_1 + b_2) + \alpha^I S\{1-b_1, b_2\}$ to connect the antecedent and the consequent,

(iii) $\tilde{T}_2\{\mathbf{c}; \alpha^{agr}\} = (1-\alpha^{agr})\frac{1}{N}\sum_{k=1}^{N} c_k + \alpha^{agr} \overset{N}{\underset{k=1}{T}}\{c_k\}$ to aggregate individual rules.

Parameters α^τ, α^I, and α^{agr} can be found in the process of learning.

4.4. PARAMETERIZED TRIANGULAR NORMS

As we mentioned in Section 4.2 any construction of fuzzy systems relies on triangular norms. Most fuzzy inference structures studied in literature employ the triangular norms described in Section 2.3. There is only a little knowledge within the engineering community about the so-called parameterized families of t-norm and t-conorms. They include the Dombi, Hamacher, Yager, Frank, Weber I, Weber II, Dubois and Prade, Schweizer I, Schweizer II, Schweizer III, Mizumoto IV-X families [45, 56]. These parameterized triangular norms are listed in Table 4.1. We use notation $\ddot{T}\{a_1, a_2, \ldots, a_n; p\}$ and $\ddot{S}\{a_1, a_2, \ldots, a_n; p\}$ for parameterized triangular norms. The hyperplanes corresponding to them can be adjusted in the process of learning of parameter p.

Table 4.1 Parameterised triangular norms

Name	t-norm s-norm	p
Dombi	$\vec{T}\{a_1,a_2;p\}=\dfrac{1}{1+\left(\left(\dfrac{1}{a_1}-1\right)^p+\left(\dfrac{1}{a_2}-1\right)^p\right)^{\frac{1}{p}}}$ $\vec{S}\{a_1,a_2;p\}=\dfrac{1}{1+\left(\left(\dfrac{1}{a_1}-1\right)^{-p}+\left(\dfrac{1}{a_2}-1\right)^{-p}\right)^{-\frac{1}{p}}}$	$p>0$
Hamacher	$\vec{T}\{a_1,a_2;p\}=\dfrac{a_1a_2}{p+(1-p)(a_1+a_2-a_1a_2)}$ $\vec{S}\{a_1,a_2;p\}=\dfrac{a_1+a_2-a_1a_2-(1-p)a_1a_2}{1-(1-p)a_1a_2}$	$p\geq 0$
Yager	$\vec{T}\{a_1,a_2;p\}=1-\min\left\{1,\sqrt[p]{(1-a_1)^p+(1-a_2)^p}\right\}$ $\vec{S}\{a_1,a_2;p\}=\min\left\{1,\sqrt[p]{a_1{}^p+a_2{}^p}\right\}$	$p>0$
Frank	$\vec{T}\{a_1,a_2;p\}=\log_p\left(1+\dfrac{(p^{a_1}-1)(p^{a_2}-1)}{p-1}\right)$ $\vec{S}\{a_1,a_2;p\}=1-\log_p\left(1+\dfrac{(p^{1-a_1}-1)(p^{1-a_2}-1)}{p-1}\right)$	$p>0$, $p\neq 1$
Weber I	$\vec{T}\{a_1,a_2;p\}=\max\left\{0,\dfrac{a_1+a_2-1+pa_1a_2}{1+p}\right\}$ $\vec{S}\{a_1,a_2;p\}=\min\left\{1,\dfrac{(1+p)(a_1+a_2)-pa_1a_2}{1+p}\right\}$	$p>-1$
Weber II	$\vec{T}\{a_1,a_2;p\}=\max\{0,(1+p)a_1+(1+p)a_2-pa_1a_2-(1+p)\}$ $\vec{S}\{a_1,a_2;p\}=\min\{1,a_1+a_2+pa_1a_2\}$	$p>-1$
Dubois and Prade	$\vec{T}\{a_1,a_2;p\}=\dfrac{a_1a_2}{\max\{a_1,a_2,p\}}$ $\vec{S}\{a_1,a_2;p\}=1-\dfrac{(1-a_1)(1-a_2)}{\max\{(1-a_1),(1-a_2),p\}}$	$0\leq p\leq 1$
Schweizer I	$\vec{T}\{a_1,a_2;p\}=\sqrt[p]{\max\{0,a_1{}^p+a_2{}^p-1\}}$ $\vec{S}\{a_1,a_2;p\}=1-\sqrt[p]{\max\{0,(1-a_1)^p+(1-a_2)^p-1\}}$	$p>0$
Schweizer II	$\vec{T}\{a_1,a_2;p\}=\dfrac{1}{\sqrt[p]{\dfrac{1}{a_1{}^p}+\dfrac{1}{a_1{}^p}-1}}$ $\vec{S}\{a_1,a_2;p\}=1-\dfrac{1}{\sqrt[p]{\dfrac{1}{(1-a_1)^p}+\dfrac{1}{(1-a_1)^p}-1}}$	$p>0$

Name	t-norm s-norm	p
Schweizer III	$\vec{T}\{a_1,a_2;p\}=1-\sqrt[p]{(1-a_1)^p+(1-a_2)^p-(1-a_1)^p(1-a_2)^p}$ $\vec{S}\{a_1,a_2;p\}=\sqrt[p]{a_1{}^p+a_2{}^p-a_1{}^p a_2{}^p}$	$p>0$
Mizumoto IV	$\vec{T}\{a_1,a_2;p\}=\begin{cases}\frac{1}{p}\left(\frac{1}{\max\left(1,\frac{1}{1+pa_1}+\frac{1}{1+pa_2}-\frac{1}{1+p}\right)}-1\right) & \text{if } -1<p<0\\ \frac{1}{p}\left(\frac{1}{\min\left(1,\frac{1}{1+pa_1}+\frac{1}{1+pa_2}-\frac{1}{1+p}\right)}-1\right) & \text{if } 0<p\end{cases}$ $\vec{S}\{a_1,a_2;p\}=\begin{cases}1-\frac{1}{p}\left(\frac{1}{\max\left(1,\frac{1}{1+p(1-a_1)}+\frac{1}{1+p(1-a_2)}-\frac{1}{1+p}\right)}-1\right) & \text{if } -1<p<0\\ 1-\frac{1}{p}\left(\frac{1}{\min\left(1,\frac{1}{1+p(1-a_1)}+\frac{1}{1+p(1-a_2)}-\frac{1}{1+p}\right)}-1\right) & \text{if } 0<p\end{cases}$	$p>-1$
Mizumoto V	$\vec{T}\{a_1,a_2;p\}=1-\log_p\left(\min\left\{p,p^{1-a_1}+p^{1-a_2}-1\right\}\right)$ $\vec{S}\{a_1,a_2;p\}=\log_p\left(\min\left\{p,p^{a_1}+p^{a_2}-1\right\}\right)$	$p>1$
Mizumoto VI	$\vec{T}\{a_1,a_2;p\}=1-\frac{1}{p}\ln\left(\min\left\{e^p,e^{p(1-a_1)}+e^{p(1-a_2)}-1\right\}\right)$ $\vec{S}\{a_1,a_2;p\}=\frac{1}{p}\ln\left(\min\left\{e^p,e^{pa_1}+e^{pa_2}-1\right\}\right)$	$p>0$
Mizumoto VII	$\vec{T}\{a_1,a_2;p\}=\sqrt[p]{1-\ln\left(\min\left\{e,e^{1-a_1{}^p}+e^{1-a_2{}^p}-1\right\}\right)}$ $\vec{S}\{a_1,a_2;p\}=1-\sqrt[p]{1-\ln\left(\min\left\{e,e^{1-(1-a_1)^p}+e^{1-(1-a_2)^p}-1\right\}\right)}$	$p>0$
Mizumoto VII	$\vec{T}\{a_1,a_2;p\}=\frac{1}{\log_p\left(p^{\frac{1}{a_1}}+p^{\frac{1}{a_2}}-p\right)}$ $\vec{S}\{a_1,a_2;p\}=1-\frac{1}{\log_p\left(p^{\frac{1}{1-a_1}}+p^{\frac{1}{1-a_2}}-p\right)}$	$p>1$
Mizumoto IX	$\vec{T}\{a_1,a_2;p\}=\frac{1}{\frac{1}{p}\ln\left(e^{\frac{p}{a_1}}+e^{\frac{p}{a_2}}-e^p\right)}$ $\vec{S}\{a_1,a_2;p\}=1-\frac{1}{\frac{1}{p}\ln\left(e^{\frac{p}{1-a_1}}+e^{\frac{p}{1-a_2}}-e^p\right)}$	$p>0$

Name	t-norm s-norm	p
Mizumoto X	$\ddot{T}\{a_1,a_2;p\}=\dfrac{1}{\sqrt[p]{\ln\left(e^{\frac{1}{a_1^p}}+e^{\frac{1}{a_2^p}}-e\right)}}$ $\ddot{S}\{a_1,a_2;p\}=1-\dfrac{1}{\sqrt[p]{\ln\left(e^{\frac{1}{(1-a_1)^p}}+e^{\frac{1}{(1-a_2)^p}}-e\right)}}$	$p>0$

As an example we present the Dombi family of parameterized triangular norms. The t-norm and t-conorm are given as follows:

(i) the Dombi t-norm

$$\ddot{T}\{\mathbf{a};p\}=\begin{cases} \textit{drastic t-norm} & \text{if} \quad p=0 \\ \left(1+\left(\sum_{i=1}^{n}\left(\dfrac{1-a_i}{a_i}\right)^p\right)^{\frac{1}{p}}\right)^{-1} & \text{if} \quad p\in(0,\infty) \\ \textit{Zadeh t-norm} & \text{if} \quad p=\infty \end{cases} \tag{4.33}$$

where $\ddot{T}$ stands for a t-norm of the Dombi family parameterized by p.

(ii) the Dombi t-conorm

$$\ddot{S}\{\mathbf{a};p\}=\begin{cases} \textit{drastic t-conorm} & \text{if} \quad p=0 \\ 1-\left(1+\left(\sum_{i=1}^{n}\left(\dfrac{a_i}{1-a_i}\right)^p\right)^{\frac{1}{p}}\right)^{-1} & \text{if} \quad p\in(0,\infty) \\ \textit{Zadeh t-conorm} & \text{if} \quad p=\infty \end{cases} \tag{4.34}$$

where $\ddot{S}$ stands for a t-conorm of the Dombi family parameterized by p.

Obviously formula (4.33) defines the "engineering implication" for $n = 2$. Combining the S-implication and (4.34) we get the fuzzy S-implication generated by the Dombi family

$$\ddot{I}(a,b;p)=1-\left(1+\left(\left(\frac{1-a}{a}\right)^p+\left(\frac{b}{1-b}\right)^p\right)^{\frac{1}{p}}\right)^{-1} \tag{4.35}$$

In the design of Mamdani-type systems we use the following parameterised triangular norms:

(i) $\vec{T}_1\{a_1, a_2, \ldots, a_n; p^\tau\}$ to connect the antecedents,

(ii) $\vec{T}_2\{b_1, b_2; p^I\}$ to connect the antecedent and the consequent,

(iii) $\vec{S}\{c_1, c_2, \ldots, c_N; p^{\text{agr}}\}$ to aggregate individual rules

where n is the number of inputs and N is the number of rules.

In the design of logical-type systems based on S-implications we use the following parameterised triangular norms:

(i) $\vec{T}_1\{a_1, a_2, \ldots, a_n; p^\tau\}$ to connect the antecedents,

(ii) $\vec{S}\{1 - b_1, b_2; p^I\}$ to connect the antecedent and the consequent,

(iii) $\vec{T}_2\{c_1, c_2, \ldots, c_N; p^{\text{agr}}\}$ to aggregate individual rules.

Parameters p^τ, p^I, and p^{agr} can be found in the process of learning. The NFIS realized by parameterized families of t-norms and t-conorms are studied in Chapters 5 and 6.

4.5. OR-TYPE SYSTEMS

In this section we introduce the concept of an OR-type system which smoothly switches between the Mamdani-type and the logical-type systems. Details and mathematical background will be provided in Chapter 5. The OR-type NFIS have been proposed by Rutkowski and Cpałka [81]. Depending on a certain parameter ν this class of systems exhibits "more Mamdani" $(0 < \nu < 0.5)$ or "more logical" $(0.5 < \nu < 1)$ behaviour. At the boundaries the system becomes of the Mamdani-type $(\nu = 0)$ or of the logical-type $(\nu = 1)$. The definition of OR-type systems heavily relies on the concept of the H-function (see Section 5.3). The H-function exhibits the behaviour of fuzzy norms. More precisely, it is a t-norm for $\nu = 0$ and a t-conorm for $\nu = 1$. For $0 < \nu < 0.5$ the H-function resembles a t-norm and for $0.5 < \nu < 1$ the H-function resembles a t-conorm. In a similar spirit we construct OR-type implications. The parameter ν can be found in the process of learning subject to the constraint $0 \le \nu \le 1$. In Chapter 5 based on the input-output data we learn a system type starting from $\nu = 0.5$ as an initial value. The behaviour of the OR-type systems, depending on parameter ν, is shown in Table 4.2.

Table 4.2 OR-type system

v	System
0	Mamdani type
1	logical type
0.5	undefined
(0,0.5)	"more Mamdani"
(0.5,1)	"more logical"

The OR-type NFIS are studied in detail in Chapter 5.

4.6. COMPROMISE SYSTEMS

The Mamdani and the logical systems lead to different results and, in literature, there are no formal proofs as to which of them is superior. Therefore, Yager and Filev [108] proposed to combine both methods. The AND-type compromise neuro-fuzzy systems are characterized by the simultaneous appearance of Mamdani-type and logical-type systems. In this book we study the following combination of "engineering" and fuzzy implications

$$I(a,b)=(1-\lambda)I_{eng}(a,b)+\lambda I_{fuzzy}(a,b) \tag{4.36}$$

For an S-implication as the fuzzy-implication and a t-norm as the "engineering implication", formula (4.36) takes the form

$$I(a,b)=(1-\lambda)T\{a,b\}+\lambda S\{1-a,b\} \tag{4.37}$$

In Chapter 6 we develop compromise NFIS based on formula (4.37). It should be emphasized that parameter λ can be found in the process of learning subject to the constraint $0\le\lambda\le 1$. In Chapter 6, based on the input-output data, we learn a system type starting from $\lambda=0.5$ as the initial value. The behaviour of the AND-type compromise systems, depending on parameter λ, is depicted in Table 4.3.

Table 4.3 AND-type system

λ	System
$\lambda = 0$	Mamdani type
$\lambda = 1$	logical type
$\lambda \in [0,1]$	compromise (Mamdani AND logical)

Example 4.12. (An example of AND-type compromise implication generated by the Zadeh t-norm)
We consider the AND-type compromise implication based on the Zadeh t-norm (2.43) and the Kleene-Dienes S-implication (3.31). It is given by

$$I(a,b) = (1-\lambda)\min\{a,b\} + \lambda \max\{1-a,b\} \tag{4.38}$$

The 3D plots of function (4.38) are depicted in Fig. 4.21.

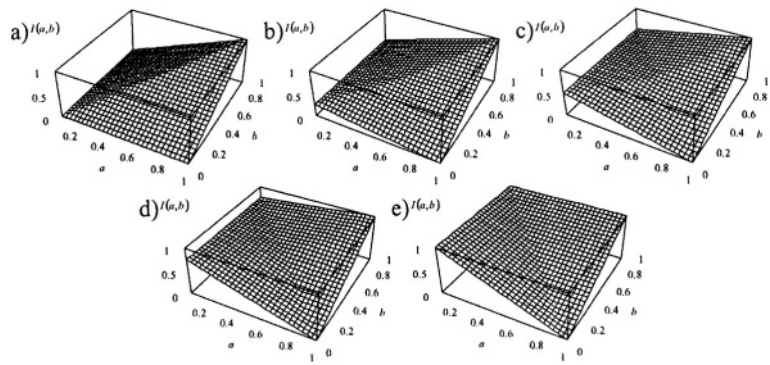

Fig. 4.21. 3D plots of function (4.38)
for a) $\lambda = 0.00$, b) $\lambda = 0.25$, c) $\lambda = 0.50$, d) $\lambda = 0.75$, e) $\lambda = 1.00$

Example 4.13. (An example of AND-type compromise implication generated by the algebraic t-norm)
We consider the AND-type compromise implication based on the algebraic t-norm (2.47) and the Reichenbach S-implication given in Table 2.1. It is defined by

$$I(a,b) = (1-\lambda)ab + \lambda(1-a+ab) \tag{4.39}$$

The 3D plots of function (4.39) are depicted in Fig. 4.22.

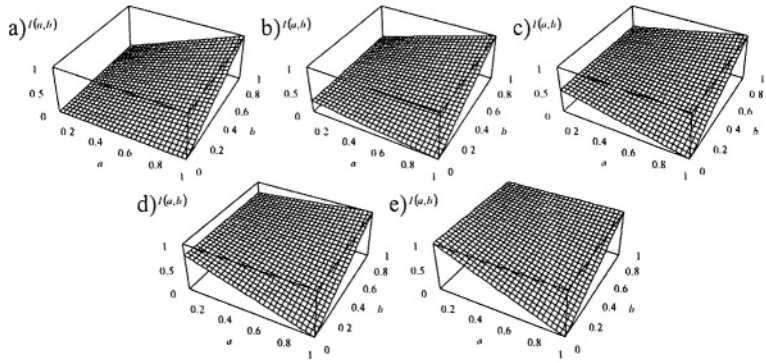

Fig. 4.22. 3D plots of function (4.39)
for a) $\lambda = 0.00$, b) $\lambda = 0.25$, c) $\lambda = 0.50$, d) $\lambda = 0.75$, e) $\lambda = 1.00$

Example 4.14. (An example of AND-type compromise implication generated by the bounded t-norm)
We consider the AND-type compromise implication based on the bounded t-norm (2.51) and the Łukasiewicz S-implication (3.32). It is given by

$$I(a,b) = (1-\lambda)\max\{a+b-1,0\} + \lambda \min\{1,1-a+b\} \tag{4.40}$$

The 3D plots of function (4.40) are depicted in Fig. 4.23.

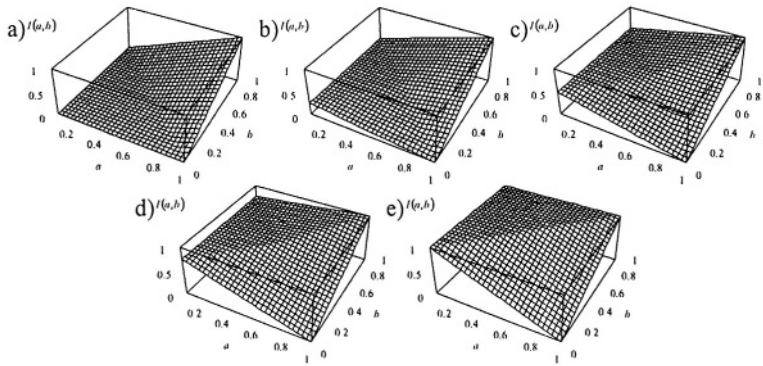

Fig. 4.23. 3D plots of function (4.40)
for a) $\lambda = 0.00$, b) $\lambda = 0.25$, c) $\lambda = 0.50$, d) $\lambda = 0.75$, e) $\lambda = 1.00$

Example 4.15. (An example of AND-type compromise implication generated by the drastic t-norm)
We consider the AND-type compromise implication based on the drastic t-norm (2.55) and the Dubois-Prade S-implication given in Table 2.1. It is defined as follows

$$I(a,b)=(1-\lambda)I_{eng}(a,b)+\lambda I_{fuzzy}(a,b) \tag{4.41}$$

where

$$I_{eng}(a,b)=\begin{cases} a & \text{if } b=1 \\ b & \text{if } a=1 \\ 0 & otherwise \end{cases} \tag{4.42}$$

and

$$I_{fuzzy}(a,b)=\begin{cases} 1-a & \text{if } b=0 \\ b & \text{if } a=1 \\ 1 & otherwise \end{cases} \tag{4.43}$$

The 3D plots of function (4.41) are depicted in Fig. 4.24.

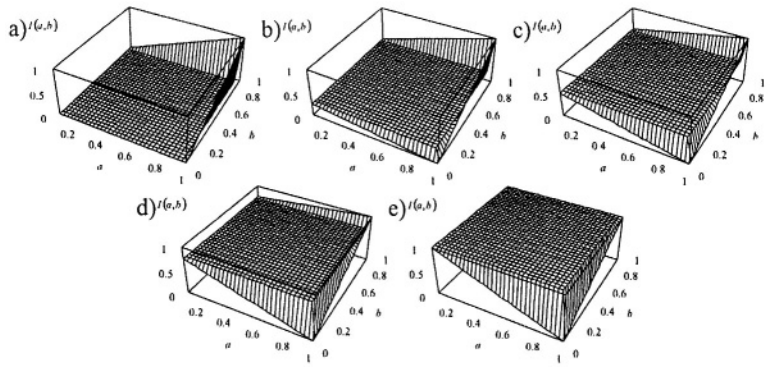

Fig. 4.24. 3D plots of function (4.41)
for a) $\lambda = 0.00$, b) $\lambda = 0.25$, c) $\lambda = 0.50$, d) $\lambda = 0.75$, e) $\lambda = 1.00$

4.7. SUMMARY AND DISCUSSION

In this chapter we explained how to incorporate flexibility into the designing of neuro-fuzzy systems. The flexibility is represented by various parameters allowing to better represent the patterns encoded in the data.

We note that other concepts can be applied to design flexible neuro-fuzzy structures. Among them it is worth to mention various

aggregation operators, e.g. the geometric mean or the harmonic mean, the so-called exponential compensatory operator [45] or the weighted minimum and maximum operators [19].

4.8. PROBLEMS

Problem 4.1. Combine soft triangular norms with the weighted t-norm, and define the weighted soft triangular norms.

Problem 4.2. Combine soft triangular t-conorms with the weighted t-conorm, and define the weighted soft triangular t-conorms.

Problem 4.3. Define the Dombi t-norm with weighted arguments.

Problem 4.4. Define the soft Dombi t-norm and t-conorm.

Problem 4.5. Define the S-implication generated by the Yager t-conorm.

Problem 4.6. Define the soft S-implication generated by the Dombi t-conorm.

Problem 4.7. Prove that

$$T\{a_1,a_2,\ldots,a_n\} \le \tilde{T}\{a_1,a_2,\ldots,a_n;\alpha\} \le \text{avg}(a_1,a_2,\ldots,a_n)$$

Problem 4.8. Prove that

$$\text{avg}(a_1,a_2,\ldots,a_n) \le \tilde{S}\{a_1,a_2,\ldots,a_n;\alpha\} \le S\{a_1,a_2,\ldots,a_n\}$$

Problem 4.9. Prove that

$$T_D\{a_1,a_2,\ldots,a_n\} \le \ddot{T}\{a_1,a_2,\ldots,a_n;p\} \le T_M\{a_1,a_2,\ldots,a_n\}$$

Problem 4.10. Prove that

$$S_M\{a_1,a_2,\ldots,a_n\} \le \ddot{S}\{a_1,a_2,\ldots,a_n;p\} \le S_D\{a_1,a_2,\ldots,a_n\}$$

Problem 4.11. Prove that

$$T_D\{a_1,a_2\} \le \ddot{I}_{eng}\{a_1,a_2;p\} \le T_M\{a_1,a_2\}$$

Problem 4.12. Prove that

$$S_M\{N(a_1),a_2\} \le \ddot{I}_{fuzzy}\{a_1,a_2;p\} \le S_D\{N(a_1),a_2\}$$

Chapter 5

FLEXIBLE OR-TYPE NEURO-FUZZY SYSTEMS

5.1. INTRODUCTION

Triangular norms (t-norms and t-conorms) have been used to model the intersection and union of fuzzy sets, the logical conjunction and disjunction and the fuzzy preference. For excellent surveys and overviews of various aggregation operators the reader is referred to [4, 9, 14, 45]. Apart from various traditional triangular norms several modifications and extensions have been proposed [10, 22, 36], e.g. the ordinal sum, the uniform, the nullnorms and the t-operator. These definitions are potentially useful to construct various new neuro-fuzzy systems, however in this chapter we introduce another class of functions called quasi-triangular norms. They are denoted by H and parameterized by parameter ν: $H(a_1, a_2, \ldots, a_n; \nu)$. From the construction of function H it follows that it becomes a t-norm for $\nu = 0$ and a dual t-conorm (S-norm) for $\nu = 1$. For $0 < \nu < 0.5$ function H resembles a t-norm and for $0.5 < \nu < 1$ it resembles a t-conorm. In this chapter we also propose adjustable quasi-implications. Most neuro-fuzzy systems proposed in the past decade employ "engineering implications" (this terminology is suggested in [57] and [58]) defined by a t-norm (e.g. minimum or product). In our proposition a quasi-implication $I(a,b;\nu)$ varies between an "engineering implication" $T\{a,b\}$ and an S-implication as ν goes from 0 to 1. Moreover, it satisfies

$$I(a,b;\nu) = \begin{cases} I_{eng}(a,b) = T\{a,b\} & \text{for} \quad \nu = 0 \\ I_{fuzzy}(a,b) = S\{1-a,b\} & \text{for} \quad \nu = 1 \end{cases} \tag{5.1}$$

assuming that T and S are dual. It should be emphasized that parameter $\nu \in [0,1]$ can be learned from the data. Consequently, we do not know the type of the system in advance. In this chapter we propose a new class of neuro-fuzzy systems. The fuzzy inference (Mamdani-type or logical-type) of such systems is determined in the process of learning. More precisely after learning one of the following two possible structures is established depending on parameter ν

a) The neuro-fuzzy system with an "engineering implication" operator (t-norm) to describe the relation between the antecedents and the consequent, and with a t-conorm for the aggregation of rules.
b) The neuro-fuzzy system with an S-implication operator to describe the relation between the antecedents and the consequent, and with a t-norm for the aggregation of rules.

As we have already mentioned in the Introduction, such a concept has not yet been proposed in literature by other authors. This chapter is organized as follows. The problem description is given in Section 5.2. In Section 5.3 we introduce the concept of adjustable quasi-triangular norms. In Section 5.4 we present adjustable quasi-implications. In Sections 5.5-5.7 a new class of neuro-fuzzy systems is proposed. In Section 5.8 learning procedures are derived to learn parameter ν (type of the system), parameters of membership functions, flexibility parameters and weights described in Chapter 4. In Section 5.9 simulation examples are given. We will use the following notations (see the block diagrams in Chapter 5 and 6)

$$\begin{aligned} G(a_1,a_2;\alpha) &= (1-\alpha)a_1 + \alpha a_2 \\ G(b_1,b_2;\alpha) &= (1-\alpha)b_1 + \alpha b_2 \end{aligned} \tag{5.2}$$

where $\alpha \in [0,1]$.

5.2. PROBLEM DESCRIPTION

Let $\overline{\mathbf{x}}(t) \in \mathbf{X} \subset \mathbf{R}^n$, $y(t) \in \mathbf{R} \subset \mathbf{Y}$ and $d(t) \in \mathbf{R} \subset \mathbf{Y}$, $t = 1,2,\ldots$, be a sequence of inputs, outputs and desirable outputs, respectively. In this chapter we address the following design and learning problems

a) Out first problem is to design a neuro-fuzzy system realizing the mapping $f: \mathbf{X} \to \mathbf{Y}$ such that a fuzzy inference, described by a quasi-implication $I(a,b;\nu)$, varies between an "engineering implication" $T\{a,b\}$ and an S-implication depending on parameter ν. The construction of quasi-implication $I(a,b;\nu)$ is described in Section 5.4. Moreover, we will incorporate: (i) softness to implication operators, to the aggregation of rules and to the connectives of

antecedents; (ii) certainty weights to the aggregation of rules and to the connectives of antecedents; and (iii) parameterized families of t-norms and t-conorms to implication operators, to the aggregation of rules and to the connectives of antecedents. All the neuro-fuzzy systems described in this chapter are in the form (3.39), i.e.

$$\bar{y} = f(\bar{\mathbf{x}}) = \frac{\sum_{r=1}^{N} \bar{y}^r \cdot \mathrm{agr}_r(\bar{\mathbf{x}}, \bar{y}^r)}{\sum_{r=1}^{N} \mathrm{agr}_r(\bar{\mathbf{x}}, \bar{y}^r)}$$

b) Based on the learning sequence $(\bar{\mathbf{x}}(1), d(1)), (\bar{\mathbf{x}}(2), d(2)), \ldots$ we wish to determine all parameters (including the system type represented by value of ν) and weights of NFIS such that

$$e(t) = \frac{1}{2}[f(\bar{\mathbf{x}}(t)) - d(t)]^2 \tag{5.3}$$

is minimized. The steepest descent optimization algorithm with constraints can be applied to solve this problem.

5.3. ADJUSTABLE QUASI-TRIANGULAR NORMS

We start with a definition which is a generalization of a strong negation (see Definition 2.15).

Definition 5.1. (Compromise operator)
Function

$$\tilde{N}_\nu : [0,1] \to [0,1] \tag{5.4}$$

given by

$$\begin{aligned} \tilde{N}_\nu(a) &= (1-\nu)N(a) + \nu N(N(a)) \\ &= (1-\nu)N(a) + \nu a \end{aligned} \tag{5.5}$$

is called a compromise operator where $\nu \in [0,1]$ and $N(a) = \tilde{N}_0(a) = 1 - a$.

Observe that

$$\tilde{N}_\nu(a) = \begin{cases} N(a) & \text{for} \quad \nu = 0 \\ \dfrac{1}{2} & \text{for} \quad \nu = \dfrac{1}{2} \\ a & \text{for} \quad \nu = 1 \end{cases} \tag{5.6}$$

Obviously function $\tilde{N}_\nu$ is a strong negation for $\nu = 0$. The 3D plot of function (5.5) is depicted in Fig. 5.1.

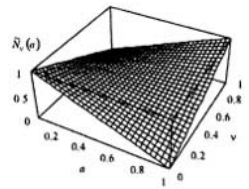

Fig. 5.1. 3D plot of function (5.5)

Definition 5.2. (H-function)
Function

$$H:[0,1]^n \to [0,1] \tag{5.7}$$

given by

$$H(\mathbf{a};\nu)=\tilde{N}_\nu\left(\overset{n}{\underset{i=1}{S}}\{\tilde{N}_\nu(a_i)\}\right)=\tilde{N}_{1-\nu}\left(\overset{n}{\underset{i=1}{T}}\{\tilde{N}_{1-\nu}(a_i)\}\right) \tag{5.8}$$

is called an H-function where $\nu \in [0,1]$.

Theorem 5.1: Let T and S be dual triangular norms. Function H defined by (5.8) varies between a t-norm and a t-conorm as ν goes from 0 to 1.

Proof: From this assumption it follows that

$$T\{\mathbf{a}\}=N(S\{N(a_1),N(a_2),\ldots,N(a_n)\}) \tag{5.9}$$

For $\nu=0$ formula (5.9) can be rewritten with the notation of the compromise operator (5.5)

$$T\{\mathbf{a}\}=\tilde{N}_0(S\{\tilde{N}_0(a_1),\tilde{N}_0(a_2),\ldots,\tilde{N}_0(a_n)\}) \tag{5.10}$$

Apparently

$$S\{\mathbf{a}\}=\tilde{N}_1(S\{\tilde{N}_1(a_1),\tilde{N}_1(a_2),\ldots,\tilde{N}_1(a_n)\}) \tag{5.11}$$

for $\nu=1$. The right-hand sides of (5.10) and (5.11) can be written as follows

$$H(\mathbf{a};\nu)=\tilde{N}_\nu\left(\overset{n}{\underset{i=1}{S}}\{\tilde{N}_\nu(a_i)\}\right) \tag{5.12}$$

for $\nu=0$ and $\nu=1$, respectively. If parameter ν changes from 0 to 1, then the result is established.

Remark 5.1: Observe that

$$H(\mathbf{a};\nu)=\begin{cases} T\{\mathbf{a}\} & \text{for} \quad \nu=0 \\ \dfrac{1}{2} & \text{for} \quad \nu=\dfrac{1}{2} \\ S\{\mathbf{a}\} & \text{for} \quad \nu=1 \end{cases} \tag{5.13}$$

It is easily seen, that for $0 < \nu < 0.5$ the H-function resembles a t-norm and for $0.5 < \nu < 1$ the H-function resembles a t-conorm.

Example 5.1. (An example of the H-function generated by the Zadeh triangular norms)
We will show how to switch smoothly from a t-norm to a t-conorm by making use of Definition 5.2. Let $n = 2$ and the standard min-norm and max-conorm be chosen:

$$H(a_1, a_2; 0) = T\{a_1, a_2\} = \min\{a_1, a_2\} \tag{5.14}$$

$$H(a_1, a_2; 1) = S\{a_1, a_2\} = \max\{a_1, a_2\} \tag{5.15}$$

The H-function generated by formulas (5.14) or (5.15) takes the form

$$\begin{aligned} H(a_1, a_2; \nu) &= \tilde{N}_{1-\nu}\left(\min\left\{\tilde{N}_{1-\nu}(a_1), \tilde{N}_{1-\nu}(a_2)\right\}\right) \\ &= \tilde{N}_{\nu}\left(\max\left\{\tilde{N}_{\nu}(a_1), \tilde{N}_{\nu}(a_2)\right\}\right) \end{aligned} \tag{5.16}$$

and varies from (5.14) to (5.15) as ν goes from zero to one. In Fig. 5.2, we illustrate function (5.16) for $\nu = 0.00$, $\nu = 0.15$, $\nu = 0.50$, $\nu = 0.85$, $\nu = 1.00$.

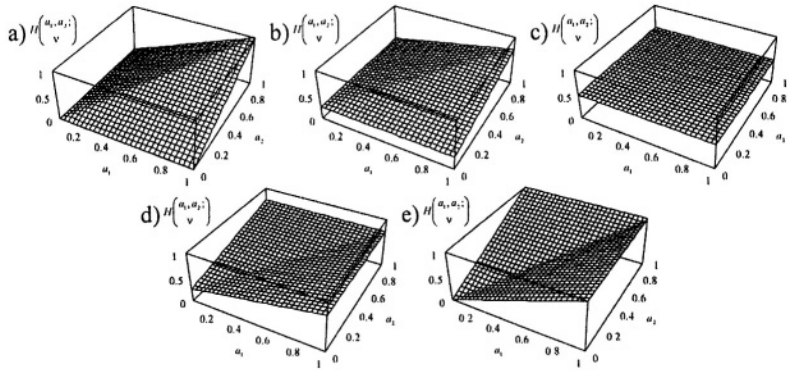

Fig. 5.2. 3D plots of function (5.16)
for a) $\nu = 0.00$, b) $\nu = 0.15$, c) $\nu = 0.50$, d) $\nu = 0.85$, e) $\nu = 1.00$

Example 5.2. (An example of the H-function generated by the algebraic triangular norms)
We will apply Theorem 5.1 to illustrate (for $n = 2$) how to switch between the algebraic t-norm

$$T\{a_1, a_2\} = H(a_1, a_2; 0) = a_1 a_2 \tag{5.17}$$

and the algebraic t-conorm

$$S\{a_1, a_2\} = H(a_1, a_2; 1) = a_1 + a_2 - a_1 a_2 \tag{5.18}$$

The H-function generated by formulas (5.17) or (5.18) takes the form

$$\begin{aligned} H(a_1,a_2;\nu) &= \tilde{N}_{1-\nu}\left(\tilde{N}_{1-\nu}(a_1)\tilde{N}_{1-\nu}(a_1)\right) \\ &= \tilde{N}_{\nu}\left(1-\left(1-\tilde{N}_{\nu}(a_1)\right)\left(1-\tilde{N}_{\nu}(a_1)\right)\right) \end{aligned} \tag{5.19}$$

and varies from (5.17) to (5.18) as ν goes from zero to one. In Fig. 5.3, we illustrate function (5.19) for $\nu = 0.00$, $\nu = 0.15$, $\nu = 0.50$, $\nu = 0.85$, $\nu = 1.00$.

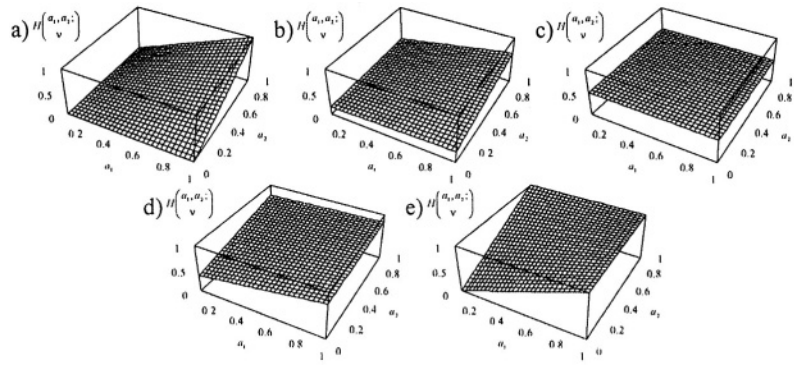

Fig. 5.3. 3D plots of function (5.19)
for a) $\nu = 0.00$, b) $\nu = 0.15$, c) $\nu = 0.50$, d) $\nu = 0.85$, e) $\nu = 1.00$

Example 5.3. (An example of the H-function generated by the bounded triangular norms)
For the bounded t-norm

$$T\{a_1,a_2\} = H(a_1,a_2;0) = \max\{a_1+a_2-1,0\} \tag{5.20}$$

and the corresponding t-conorm

$$S\{a_1,a_2\} = H(a_1,a_2;1) = \min\{a_1+a_2,1\} \tag{5.21}$$

the H-function generated by formula (5.20) or (5.21) takes the form

$$\begin{aligned} H(a_1,a_2;\nu) &= \tilde{N}_{1-\nu}\left(\max\left\{\tilde{N}_{1-\nu}(a_1)+\tilde{N}_{1-\nu}(a_1)-1,0\right\}\right) \\ &= \tilde{N}_{\nu}\left(\min\left\{\tilde{N}_{\nu}(a_1)+\tilde{N}_{\nu}(a_1),1\right\}\right) \end{aligned} \tag{5.22}$$

and varies from (5.20) to (5.21) and ν goes from 0 to 1. In Fig. 5.4 we depict function (5.22) for a) $\nu = 0.00$, b) $\nu = 0.15$, c) $\nu = 0.50$, d) $\nu = 0.85$, e) $\nu = 1.00$.

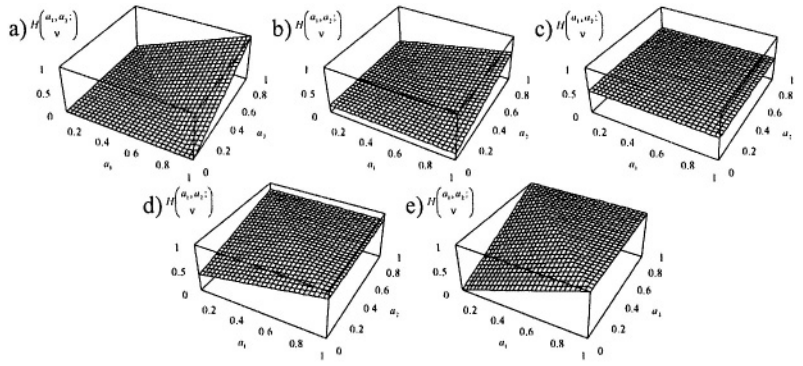

Fig. 5.4. 3D plots of function (5.22)
for a) $\nu = 0.00$, b) $\nu = 0.15$, c) $\nu = 0.50$, d) $\nu = 0.85$, e) $\nu = 1.00$

Example 5.4. (An example of the H-function generated by the drastic triangular norms)
The drastic H-function based on drastic triangular norms given by (2.55) and (2.56) and formula (5.8). The drastic H-function is described as follows

$$\begin{aligned} H(a_1,a_2;\nu) &= \widetilde{N}_{1-\nu}\left(\begin{cases} \widetilde{N}_{1-\nu}(a_1) & \text{if } \widetilde{N}_{1-\nu}(a_2)=1 \\ \widetilde{N}_{1-\nu}(a_2) & \text{if } \widetilde{N}_{1-\nu}(a_1)=1 \\ 0 & \textit{otherwise} \end{cases}\right) \\ &= \widetilde{N}_{\nu}\left(\begin{cases} \widetilde{N}_{\nu}(a_1) & \text{if } \widetilde{N}_{\nu}(a_2)=0 \\ \widetilde{N}_{\nu}(a_2) & \text{if } \widetilde{N}_{\nu}(a_1)=0 \\ 1 & \textit{otherwise} \end{cases}\right) \end{aligned} \tag{5.23}$$

The 3D plots of function (5.23) are depicted in Fig. 5.5.

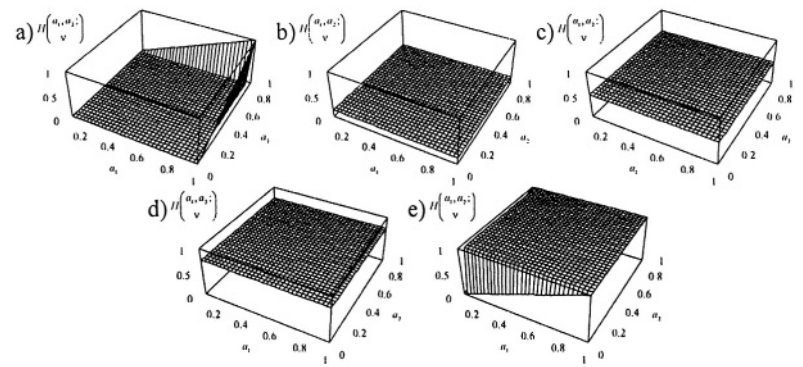

Fig. 5.5. 3D plots of function (5.23)
for a) $\nu = 0.00$, b) $\nu = 0.15$, c) $\nu = 0.50$, d) $\nu = 0.85$, e) $\nu = 1.00$

5.4. ADJUSTABLE QUASI-IMPLICATIONS

The next theorem shows how to switch between an "engineering implication" (defined by a t-norm) and an S-implication.

Theorem 5.2: Let T and S be dual triangular norms. Then

$$I(a,b;\nu) = H\left(\widetilde{N}_{1-\nu}(a), b;\nu\right) \tag{5.24}$$

switches between an "engineering implication"

$$I_{eng}(a,b) = I(a,b;0) = T\{a,b\} \tag{5.25}$$

and an S-implication

$$I_{fuzzy}(a,b) = I(a,b;1) = S\{1-a,b\} \tag{5.26}$$

Proof: Theorem 5.2 is a straightforward consequence of Theorem 5.1.

Example 5.5. (An example of the H-implication generated by the Zadeh triangular norms)
We will define the H-implication generated by the Zadeh triangular norms and based on formula (5.24). Let

$$\begin{aligned} I_{eng}(a,b) &= H(a,b;0) \\ &= T\{a,b\} \\ &= \min\{a,b\} \end{aligned} \tag{5.27}$$

and

$$\begin{aligned} I_{fuzzy}(a,b) &= H\left(\widetilde{N}_0(a), b;1\right) \\ &= S\{N(a),b\} \\ &= \max\{N(a),b\} \end{aligned} \tag{5.28}$$

Then

$$I(a,b;\nu) = H\left(\widetilde{N}_{1-\nu}(a), b;\nu\right) \tag{5.29}$$

goes from (5.27) to (5.28) as ν varies from 0 to 1.

The 3D plots of function (5.29) are depicted in Fig. 5.6.

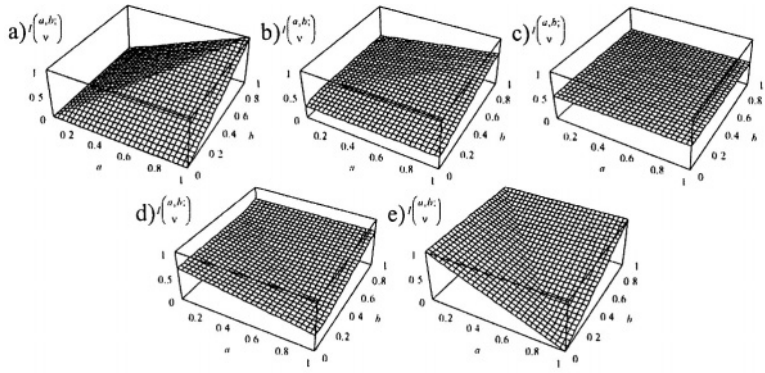

Fig. 5.6. 3D plots of function (5.29)
for a) $\nu = 0.00$, b) $\nu = 0.15$, c) $\nu = 0.50$, d) $\nu = 0.85$, e) $\nu = 1.00$

Example 5.6. (An example of the H-implication generated by the algebraic triangular norms)
We will define the H-implication generated by the algebraic triangular norms and based on formula (5.24). Let

$$\begin{aligned} I_{eng}(a,b) &= H(a,b;0) \\ &= T\{a,b\} \\ &= ab \end{aligned} \tag{5.30}$$

and

$$\begin{aligned} I_{fuzzy}(a,b) &= H\left(\tilde{N}_0(a),b;1\right) \\ &= S\{N(a),b\} \\ &= 1-a+ab \end{aligned} \tag{5.31}$$

Then

$$I(a,b;\nu) = H\left(\tilde{N}_{1-\nu}(a),b;\nu\right) \tag{5.32}$$

goes from (5.30) to (5.31) as ν varies from 0 to 1.

The 3D plots of function (5.32) are depicted in Fig. 5.7.

Example 5.7. (An example of the H-implication generated by the bounded triangular norms)
We will define the H-implication generated by the bounded triangular norms and based on formula (5.24). Let

$$\begin{aligned} I_{eng}(a,b) &= H(a,b;0) \\ &= T\{a,b\} \\ &= \max\{a+b-1,0\} \end{aligned} \tag{5.33}$$

and

$$\begin{aligned} I_{fuzzy}(a,b) &= H\left(\widetilde{N}_0(a),b;1\right) \\ &= S\{N(a),b\} \\ &= \min\{1-a+b,1\} \end{aligned} \tag{5.34}$$

Then

$$I(a,b;\nu) = H\left(\widetilde{N}_{1-\nu}(a),b;\nu\right) \tag{5.35}$$

goes from (5.33) to (5.34) as ν varies from 0 to 1.

The 3D plots of function (5.35) are depicted in Fig. 5.8.

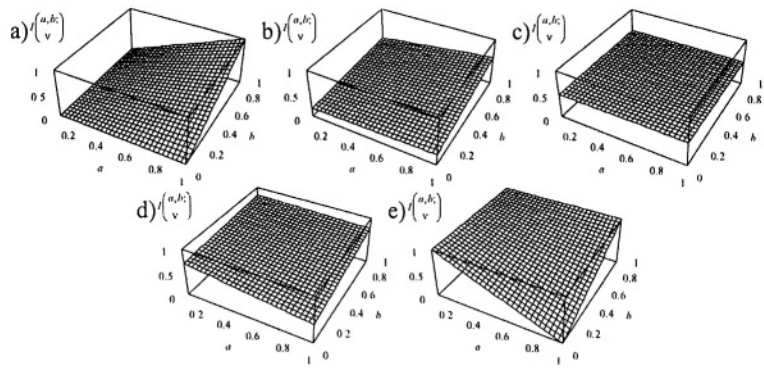

Fig. 5.7. 3D plots of function (5.32)
for a) $\nu = 0.00$, b) $\nu = 0.15$, c) $\nu = 0.50$, d) $\nu = 0.85$, e) $\nu = 1.00$

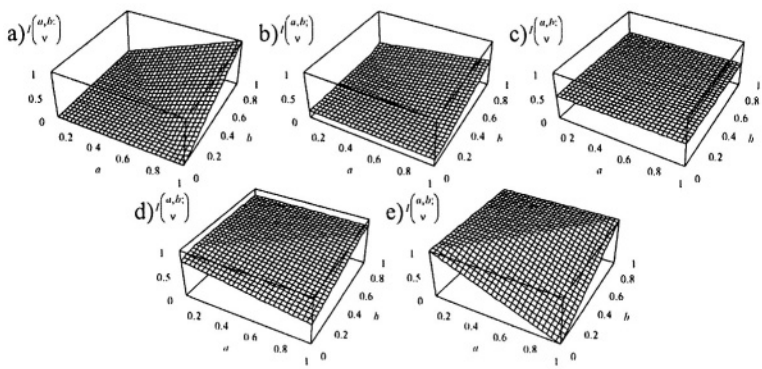

Fig. 5.8. 3D plots of function (5.35)
for a) $\nu = 0.00$, b) $\nu = 0.15$, c) $\nu = 0.50$, d) $\nu = 0.85$, e) $\nu = 1.00$

Example 5.8. (An example of the H-implication generated by the drastic triangular norms)
We will define the H-implication generated by the drastic triangular norms and based on formula (5.24). Let

$$\begin{aligned} I_{eng}(a,b) &= H(a,b;0) \\ &= T\{a,b\} \\ &= \begin{cases} a & \text{if} \quad b=1 \\ b & \text{if} \quad a=1 \\ 0 & \textit{otherwise} \end{cases} \end{aligned} \tag{5.36}$$

and

$$\begin{aligned} I_{fuzzy}(a,b) &= H\left(\tilde{N}_0(a),b;1\right) \\ &= S\{N(a),b\} \\ &= \begin{cases} 1-a & \text{if} \quad b=0 \\ b & \text{if} \quad a=1 \\ 1 & \textit{otherwise} \end{cases} \end{aligned} \tag{5.37}$$

Then

$$I(a,b;\nu) = H\left(\tilde{N}_{1-\nu}(a),b;\nu\right) \tag{5.38}$$

goes from (5.36) to (5.37) as ν varies from 0 to 1.

The 3D plots of function (5.38) are depicted in Fig. 5.9.

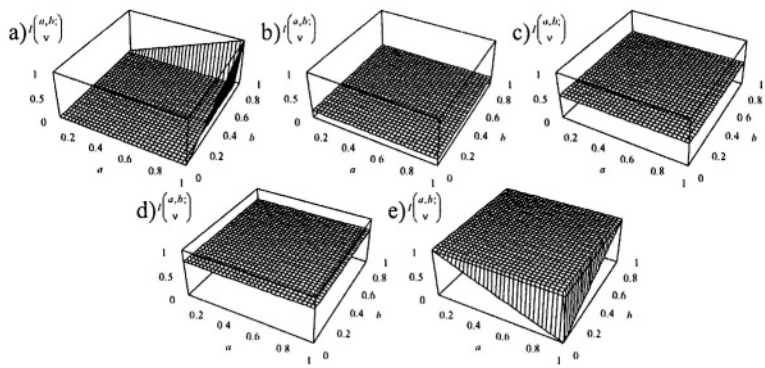

Fig. 5.9. 3D plots of function (5.38)
for a) $\nu = 0.00$, b) $\nu = 0.15$, c) $\nu = 0.50$, d) $\nu = 0.85$, e) $\nu = 1.00$

5.5. BASIC FLEXIBLE SYSTEMS

In this section we will apply the concept of an H-function for constructing the basic flexible neuro-fuzzy inference systems. By making use of that concept we will establish a new fuzzy inference allowing to switch smoothly between the Mamdani-type and the logical-type inference and vice-versa. Therefore, our basic flexible neuro-fuzzy system generalizes commonly used fuzzy inference systems of the Mamdani and the logical-type. Moreover, our approach allows to exclude insufficient performance of the NFIS resulted from incorrectly assumed fuzzy reasoning. The flexible OR-type systems exhibit "more Mamdani" or "more logical" behaviour depending on parameter ν appearing in the construction of function *H*. The system becomes of Mamdani-type for $\nu = 0$ and of logical-type for $\nu = 1$. The fuzzy inference (fuzzy quasi-implication) is realized based on function *H* with parameter ν, whereas the aggregation is based on function *H* with parameter $1-\nu$. The aggregation of the antecedents is based on a t-norm which is equivalent to an H-function with parameter $\nu = 0$. Finally, the basic neuro-fuzzy system of an OR-type is given as follows:

OR I

$$\tau_k(\bar{\mathbf{x}}) = H\left(\begin{matrix}\mu_{A_1^k}(\bar{x}_1),\ldots,\mu_{A_n^k}(\bar{x}_n)\ ;\\ 0\end{matrix}\right) \tag{5.39}$$

$$I_{k,r}(\bar{\mathbf{x}},\bar{y}^r) = H\left(\begin{matrix}\tilde{N}_{1-\nu}(\tau_k(\bar{\mathbf{x}})),\mu_{B^k}(\bar{y}^r)\ ;\\ \nu\end{matrix}\right) \tag{5.40}$$

$$\mathrm{agr}_r(\bar{\mathbf{x}},\bar{y}^r) = H\left(\begin{matrix}I_{1,r}(\bar{\mathbf{x}},\bar{y}^r),\ldots,I_{N,r}(\bar{\mathbf{x}},\bar{y}^r)\ ;\\ 1-\nu\end{matrix}\right) \tag{5.41}$$

Observe that system (5.39)-(5.41) is of the Mamdani-type for $\nu = 0$, like the Mamdani-type for $\nu \in (0, 0.5)$, undetermined for $\nu = 0.5$, like the logical-type for $\nu \in (0.5, 1)$ and the logical-type for $\nu = 1$. It is worth noticing that parameter ν can be learned and consequently the type of the system can be determined in the process of learning. In Table 5.1 we depict quasi-implication and aggregation operators used in system (5.39)-(5.41).

Table 5.1 Implication and aggregation operators as ν varies from 0 to 1

Parameter ν	Implication	Aggregation
$\nu = 0$	$T\{a,b\}$	S-norm
$\nu = 1$	$S\{1-a,b\}$	T-norm
$0 < \nu < 1$	$H(\tilde{N}_{1-\nu}(a),b;\nu)$	$H(a,b;1-\nu)$
$\nu = 0.5$	$H(0.5,b;0.5) = 0.5$	$H(a,b;0.5) = 0.5$

In Fig. 5.10, 5.11 and 5.12 we depict block diagrams realizing the aggregation of antecedents $\tau_k(\bar{\mathbf{x}})$ given by (5.39), the quasi-implication $I_{k,r}(\bar{\mathbf{x}}, \bar{y}^r)$ given by (5.40) and the aggregation $\mathbf{agr}_r(\bar{\mathbf{x}}, \bar{y}^r)$ given by (5.41), respectively.

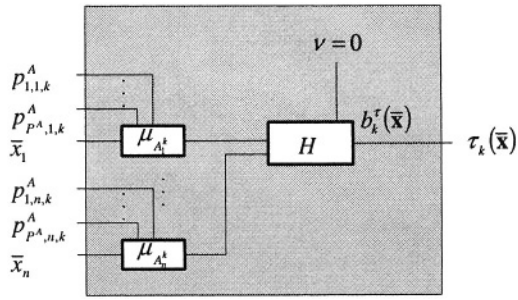

Fig. 5.10. Connection of antecedents in basic OR-type NFIS (OR I)

The firing strength of rules $\tau_k(\bar{\mathbf{x}})$, $k = 1,....,N$, following Fig. 5.10, is determined by an aggregation, based on an H-function parameterized by $\nu = 0$, of antecedent membership function values $\mu_{A_i^k}(\bar{x}_i)$ calculated for all inputs $\bar{x}_i$, $i = 1,...,n$. It is assumed that the membership functions are characterized by P^A parameters $p^A_{u,i,k}$, $u = 1,...,P^A$, $i = 1,...,n$, $k = 1,...,N$.

The quasi-implication $I_{k,r}(\bar{\mathbf{x}}, \bar{y}^r)$, $k = 1,...,N$, $r = 1,...,N$, following Fig. 5.11, is determined by an aggregation based on an H-function parameterized by $\nu \in [0,1]$. The first argument of the aggregation operation is the firing strength of rules $\tau_k(\bar{\mathbf{x}})$, $k = 1,...,N$. The second argument of the aggregation operation is the value of consequence membership functions $\mu_{B^k}(\bar{y}^r)$, $k = 1,...,N$, $r = 1,...,N$, determined in point $\bar{y}^r$, $r = 1,...,N$, at

which consequent membership functions μ_{B^k}, $k = 1,\ldots,N$, attain their maximum equal to 1. It is assumed that each output membership function μ_{B^k} is characterized by P^B parameters $p_{u,k}^B$, $u = 1,\ldots,P^B$, $k = 1,\ldots,N$.

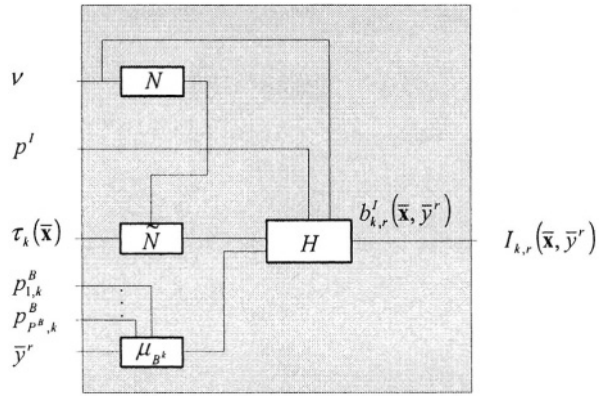

Fig. 5.11. Implication operator in basic OR-type NFIS (OR I)

The quasi-triangular aggregation $\mathrm{agr}_r(\bar{\mathbf{x}},\bar{y}^r)$, $r = 1,\ldots,N$, following Fig. 5.12, is determined by an aggregation, based on an H-function parameterized by $1-\nu$, $\nu \in [0,1]$, of quasi-implications $I_{k,r}(\bar{\mathbf{x}},\bar{y}^r)$, $k = 1,\ldots,N$, $r = 1,\ldots,N$, determined in point $\bar{y}^r$, $r = 1,\ldots,N$, at which consequent membership functions μ_{B^k}, $k = 1,\ldots,N$, attain their maximum equal to 1.

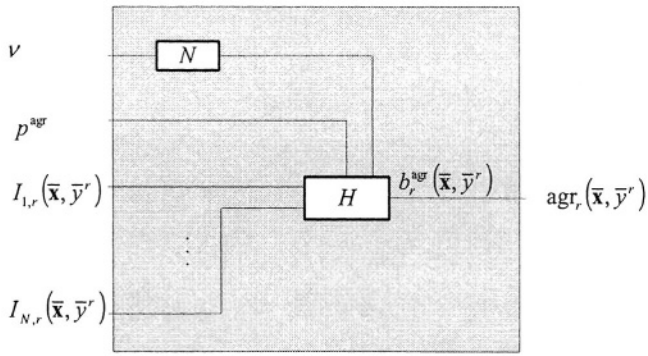

Fig. 5.12. Aggregation of rules in basic OR-type NFIS (OR I)

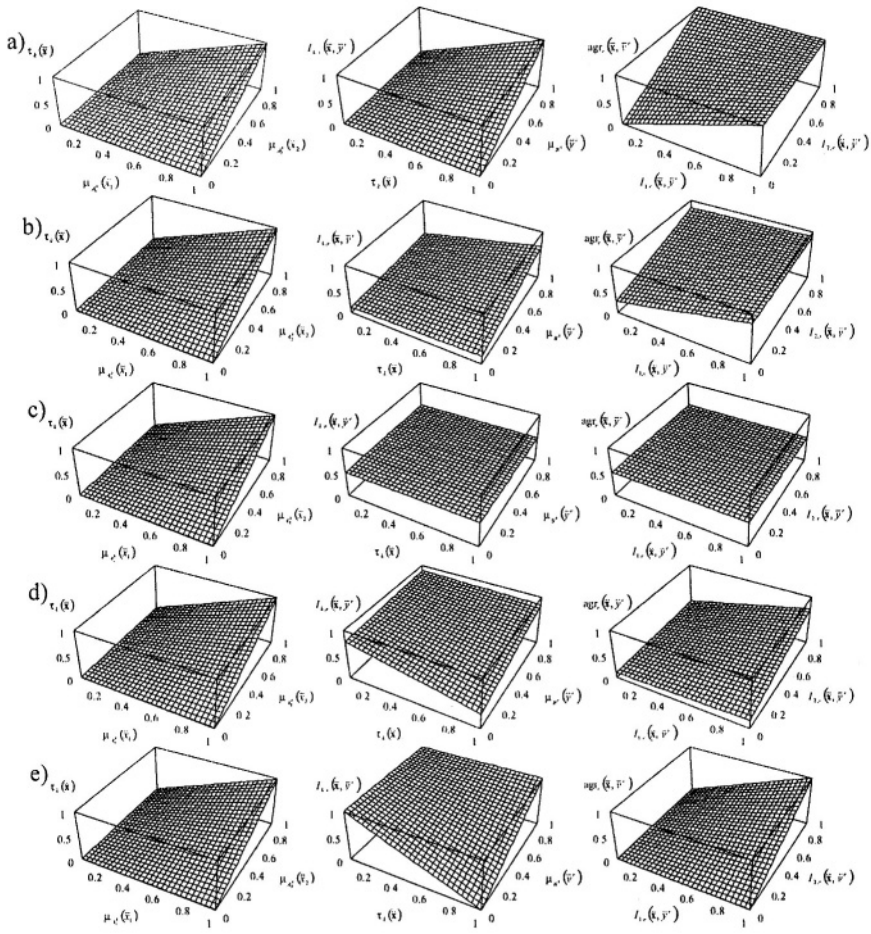

Fig. 5.13. 3D plots of basic OR-type NFIS (OR I)
for $n = 2$, $N = 2$ and a) $\nu = 0.00$, b) $\nu = 0.15$, c) $\nu = 0.50$, d) $\nu = 0.85$, e) $\nu = 1.00$

In Fig. 5.13 we depict the shape of the hyperplanes described by formulas (5.39)-(5.41) for $n = 2$, $N = 2$ and the varying parameter ν of function H generated by the algebraic triangular norms.

5.6. SOFT FLEXIBLE SYSTEMS

In this chapter we propose soft NFIS based on soft fuzzy norms (4.18) and (4.19). These systems are characterized by:

(i) soft strength of firing controlled by parameter α^{τ},

(ii) soft implication controlled by parameter α^{I},

(iii) soft aggregation of rules controlled by parameter α^{agr}.

Moreover, we assume that fuzzy norms (and H-functions) in the connection of antecedents, implication and aggregation of rules are parameterised by parameters p^{τ}, p^{I}, p^{agr}, respectively. We use notation $\ddot{H}(\cdot)$ to indicate parameterised families analogously to (4.33) and (4.34).

The soft compromise NFIS of the OR-type are defined as follows:

OR II

$$\tau_k(\bar{\mathbf{x}}) = \begin{pmatrix} (1-\alpha^{\tau})\text{avg}\left(\mu_{A_1^k}(\bar{x}_1),\ldots,\mu_{A_n^k}(\bar{x}_n)\right)+ \\ +\alpha^{\tau}\ddot{H}\begin{pmatrix}\mu_{A_1^k}(\bar{x}_1),\ldots,\mu_{A_n^k}(\bar{x}_n)\ ; \\ p^{\tau},0\end{pmatrix}\end{pmatrix} \tag{5.42}$$

$$I_{k,r}(\bar{\mathbf{x}},\bar{y}^r) = \begin{pmatrix} (1-\alpha^{I})\text{avg}\left(\tilde{N}_{1-\nu}(\tau_k(\bar{\mathbf{x}})),\mu_{B^k}(\bar{y}^r)\right)+ \\ +\alpha^{I}\ddot{H}\begin{pmatrix}\tilde{N}_{1-\nu}(\tau_k(\bar{\mathbf{x}})),\mu_{B^k}(\bar{y}^r)\ ; \\ p^{I},\nu\end{pmatrix}\end{pmatrix} \tag{5.43}$$

$$\text{agr}_r(\bar{\mathbf{x}},\bar{y}^r) = \begin{pmatrix} (1-\alpha^{\text{agr}})\text{avg}\left(I_{1,r}(\bar{\mathbf{x}},\bar{y}^r),\ldots,I_{N,r}(\bar{\mathbf{x}},\bar{y}^r)\right)+ \\ +\alpha^{\text{agr}}\ddot{H}\begin{pmatrix}I_{1,r}(\bar{\mathbf{x}},\bar{y}^r),\ldots,I_{N,r}(\bar{\mathbf{x}},\bar{y}^r)\ ; \\ p^{\text{agr}},1-\nu\end{pmatrix}\end{pmatrix} \tag{5.44}$$

Observe that system (5.42)-(5.44) is:

(i) soft Mamdani-type NFIS for $\nu = 0$,

(ii) soft logical-type NFIS for $\nu = 1$,

(iii) soft like Mamdani-type NFIS for $0 < \nu < 0.5$,

(iv) soft like logical-type NFIS for $0.5 < v < 1$,
(v) undetermined for $v = 0.5$.

It is easily seen that the above system reduces to a basic flexible system described in Section 5.5 if $\alpha^{\tau} = \alpha^{I} = \alpha^{\mathrm{agr}} = 1$ and parameterized triangular norms are replaced by standard triangular norms.

In Fig. 5.14, 5.15 and 5.16 we depict block diagrams realizing the aggregation of antecedents $\tau_k(\bar{\mathbf{x}})$ given by (5.42), the quasi-implication $I_{k,r}(\bar{\mathbf{x}}, \bar{y}^r)$ given by (5.43) and the aggregation $\mathrm{agr}_r(\bar{\mathbf{x}}, \bar{y}^r)$ given by (5.44), respectively.

The firing strength of rules $\tau_k(\bar{\mathbf{x}})$, $k = 1,...,N$, following Fig. 5.14 is determined by a soft composition, controlled by parameter α^{τ}, of two components $a_k^{\tau}(\bar{\mathbf{x}})$ and $b_k^{\tau}(\bar{\mathbf{x}})$. Component $a_k^{\tau}(\bar{\mathbf{x}})$ is the arithmetic average of the antecedents' membership function values $\mu_{A_i^k}(\bar{x}_i)$, $k = 1,...,N$, $i = 1,...,n$, determined for all inputs $\bar{x}_i$, $i = 1,...,n$. Component $b_k^{\tau}(\bar{\mathbf{x}})$ is determined analogously as in the basic flexible NFIS described in Section 5.5. The dominance of one component over another depends on the value of parameter α^{τ} .

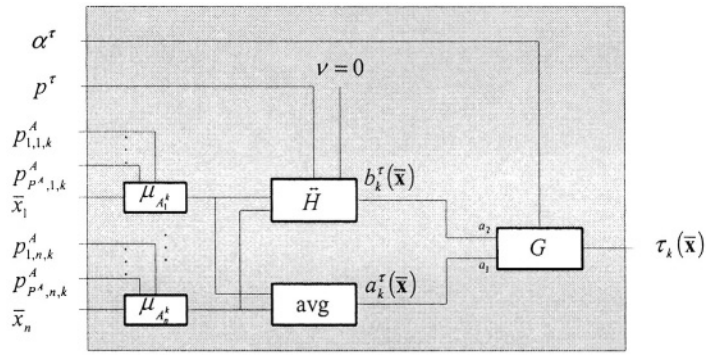

Fig. 5.14. Connection of antecedents in soft OR-type NFIS (OR II)

The quasi-implication $I_{k,r}(\bar{\mathbf{x}}, \bar{y}^r)$, $k = 1,...,N$, $r = 1,...,N$, following Fig. 5.15, is determined by a soft composition, controlled by parameter α^{I}, of two components $a_{k,r}^{I}(\bar{\mathbf{x}}, \bar{y}^r)$ and $b_{k,r}^{I}(\bar{\mathbf{x}}, \bar{y}^r)$. Component $a_{k,r}^{I}(\bar{\mathbf{x}}, \bar{y}^r)$ is the arithmetic average of the compromise operator (5.5), operated on the firing rule strength $\tau_k(\bar{\mathbf{x}})$, $k = 1,...,N$, and the value of consequence membership

functions $\mu_{B^k}(\bar{y}^r)$, $k = 1,\ldots,N$, $r = 1,\ldots,N$, determined in point $\bar{y}^r$, $r = 1,\ldots,N$, at which functions μ_{B^k} attain their maximum equal to 1. Component $b^I_{k,r}(\bar{\mathbf{x}}, \bar{y}^r)$ is determined analogously as in the basic flexible systems described in Section 5.5. The dominance of one component over another depends on the value of parameter α^r.

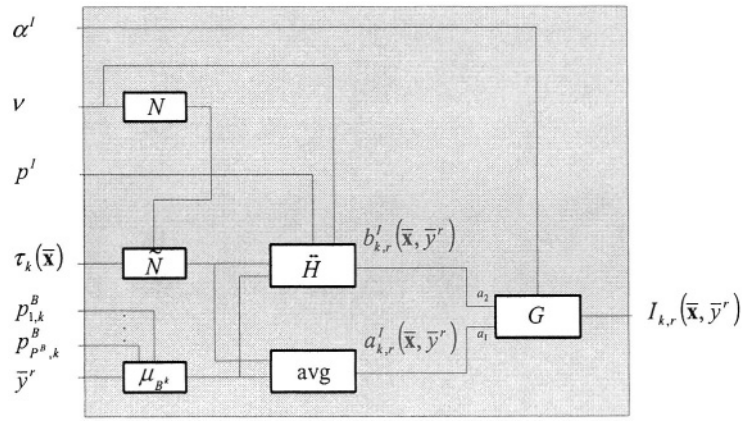

Fig. 5.15. Implication operator in soft OR-type NFIS (OR II)

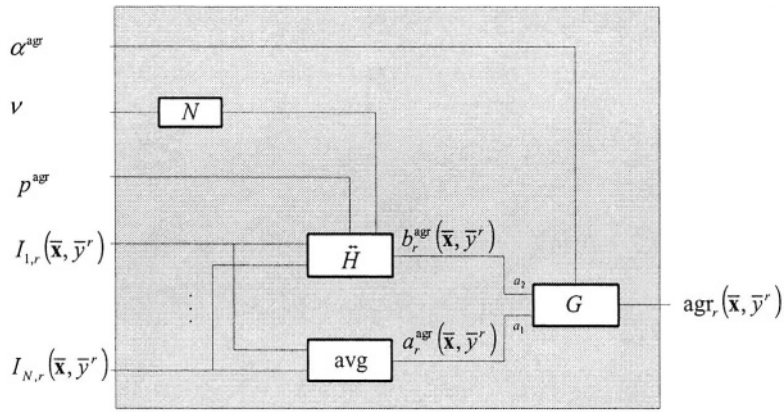

Fig. 5.16. Aggregation of rules in soft OR-type NFIS (OR II)

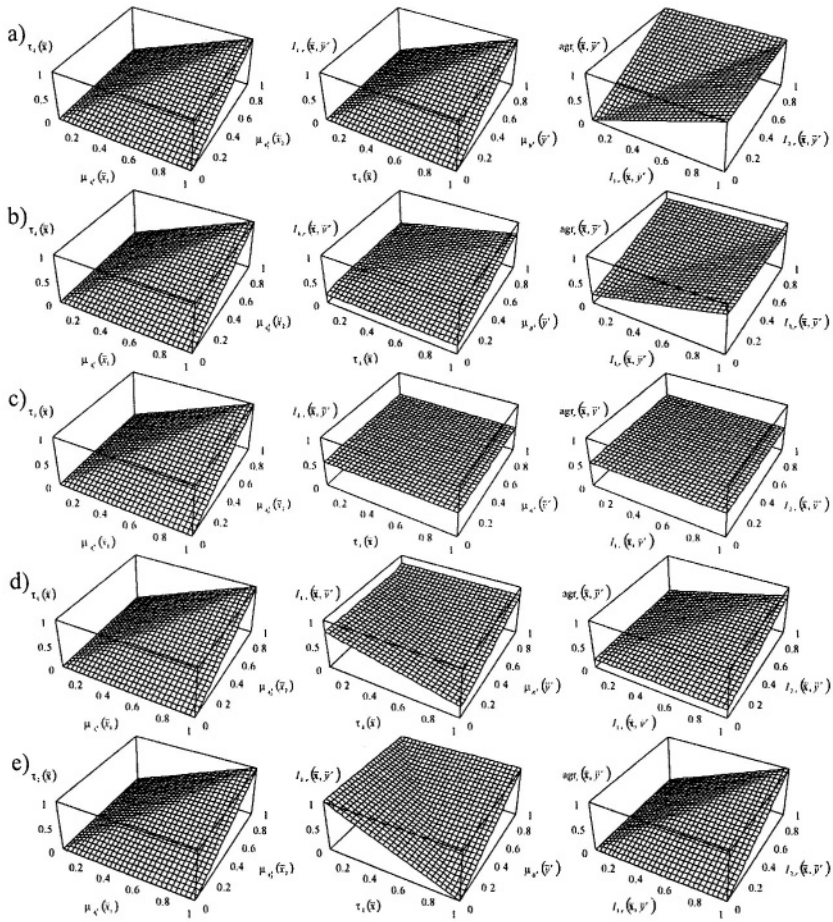

Fig. 5.17. 3D plots of soft OR-type NFIS (OR II)
for $n = 2$, $N = 2$, $\alpha^{r} = \alpha^{l} = \alpha^{agr} = 1.0$, $p^{r} = p^{l} = p^{agr} = 10$
and a) $\nu = 0.00$, b) $\nu = 0.15$, c) $\nu = 0.50$, d) $\nu = 0.85$, e) $\nu = 1.00$

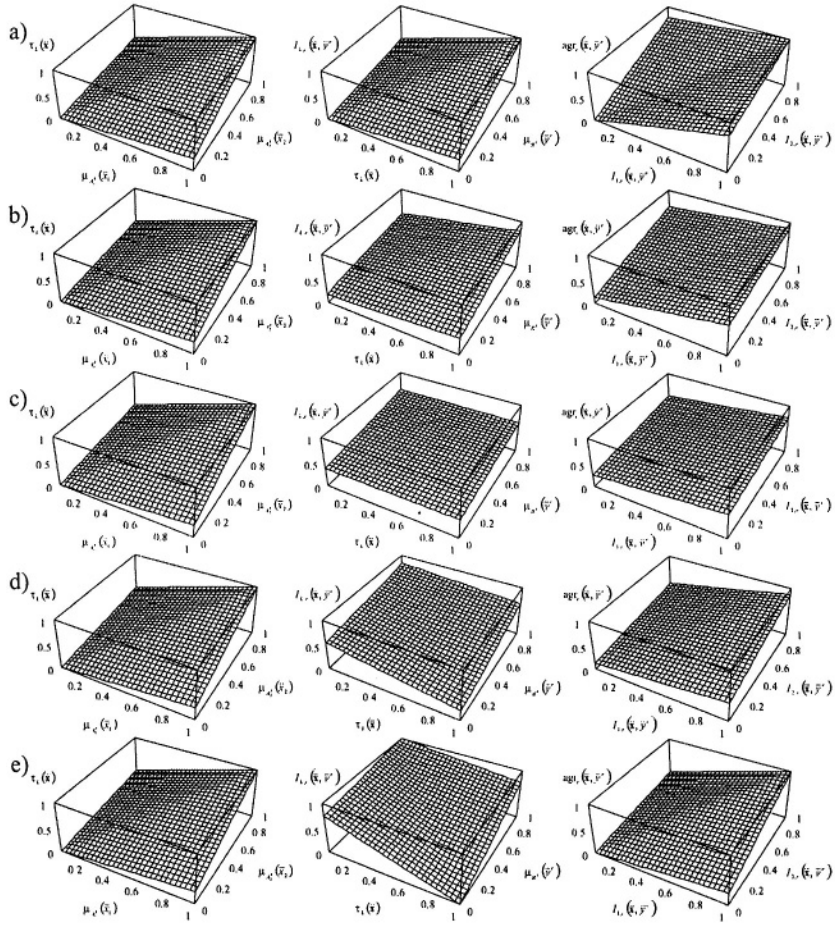

Fig. 5.18. 3D plots of soft OR-type NFIS (OR II)
for $n = 2$, $N = 2$, $\alpha^r = \alpha^l = \alpha^{agr} = 0.5$, $p^r = p^l = p^{agr} = 10$
and a) $\nu = 0.00$, b) $\nu = 0.15$, c) $\nu = 0.50$, d) $\nu = 0.85$, e) $\nu = 1.00$

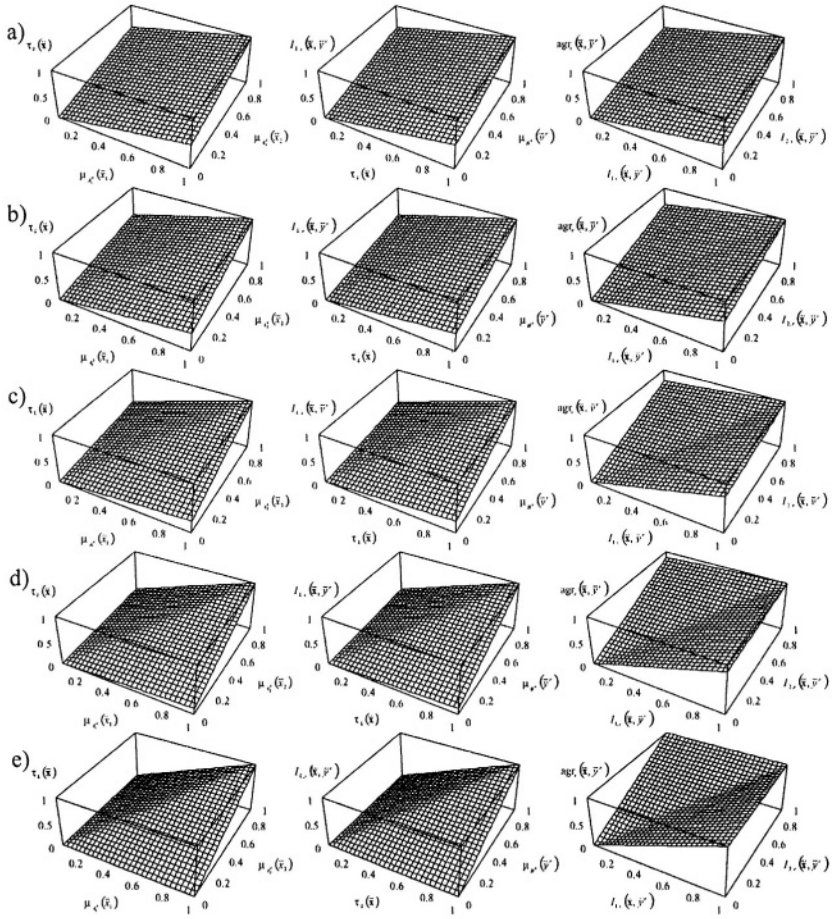

Fig. 5.19. 3D plots of soft OR-type NFIS (OR II)
for $n = 2$, $N = 2$, $p^{\tau} = p^{I} = p^{agr} = 10$, $v = 0$
and a) $\alpha^{\tau} = \alpha^{I} = \alpha^{agr} = 0.00$, b) $\alpha^{\tau} = \alpha^{I} = \alpha^{agr} = 0.25$, c) $\alpha^{\tau} = \alpha^{I} = \alpha^{agr} = 0.50$,
d) $\alpha^{\tau} = \alpha^{I} = \alpha^{agr} = 0.75$, e) $\alpha^{\tau} = \alpha^{I} = \alpha^{agr} = 1.00$

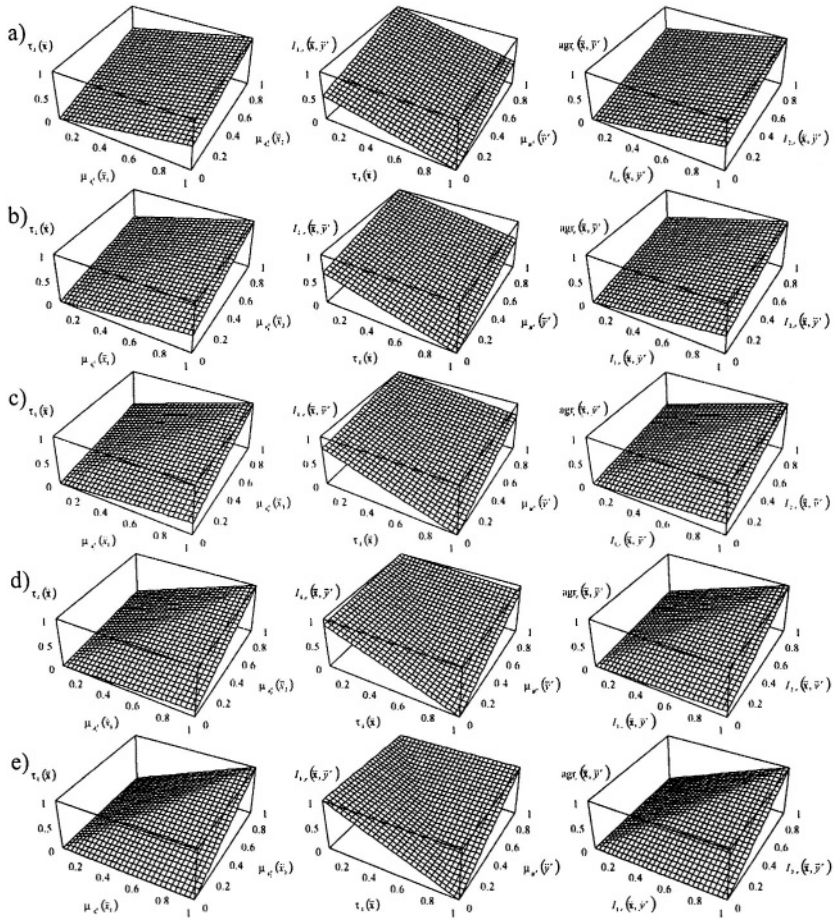

Fig. 5.20. 3D plots of soft OR-type NFIS (OR II)
for $n = 2$, $N = 2$, $p^{\tau} = p^{I} = p^{\text{agr}} = 10$, $\nu = 1$
and a) $\alpha^{\tau} = \alpha^{I} = \alpha^{\text{agr}} = 0.00$, b) $\alpha^{\tau} = \alpha^{I} = \alpha^{\text{agr}} = 0.25$, c) $\alpha^{\tau} = \alpha^{I} = \alpha^{\text{agr}} = 0.50$,
d) $\alpha^{\tau} = \alpha^{I} = \alpha^{\text{agr}} = 0.75$, e) $\alpha^{\tau} = \alpha^{I} = \alpha^{\text{agr}} = 1.00$

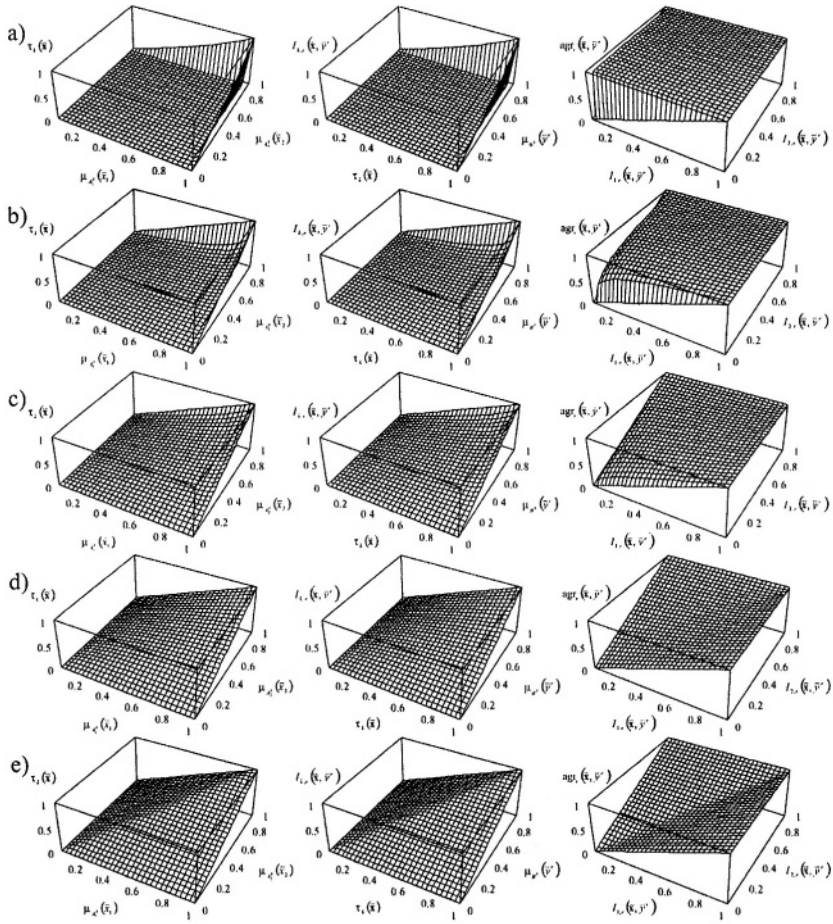

Fig. 5.21. 3D plots of soft OR-type NFIS (OR II)
for $n = 2$, $N = 2$, $\alpha^{r} = \alpha^{l} = \alpha^{agr} = 1.00$, $\nu = 0$
and a) $p^{r} = p^{l} = p^{agr} = 0.10$, b) $p^{r} = p^{l} = p^{agr} = 0.25$, c) $p^{r} = p^{l} = p^{agr} = 0.50$,
d) $p^{r} = p^{l} = p^{agr} = 1.00$, e) $p^{r} = p^{l} = p^{agr} = 10.00$

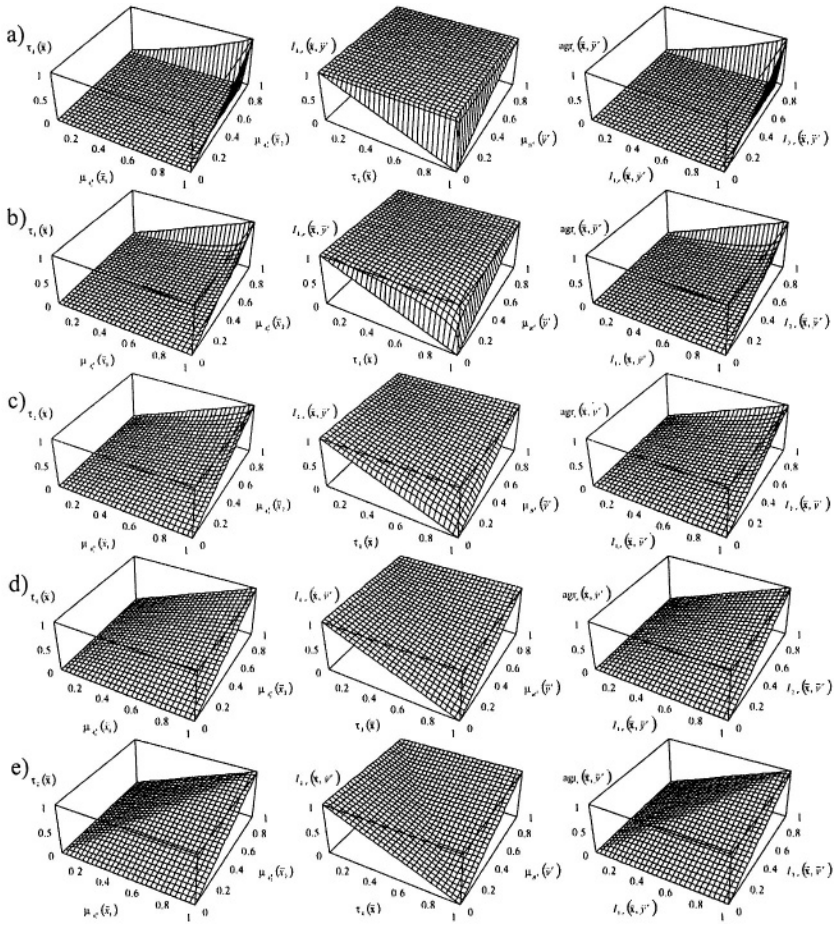

Fig. 5.22. 3D plots of soft OR-type NFIS (OR II)
for $n = 2$, $N = 2$, $\alpha^r = \alpha^l = \alpha^{agr} = 1.00$, $v = 1$
and a) $p^r = p^l = p^{agr} = 0.10$, b) $p^r = p^l = p^{agr} = 0.25$, c) $p^r = p^l = p^{agr} = 0.50$,
d) $p^r = p^l = p^{agr} = 1.00$, e) $p^r = p^l = p^{agr} = 10.00$

The aggregation operator $\mathbf{agr}_r\left(\bar{\mathbf{x}}, \bar{y}^r\right)$, $r = 1,\ldots,N$, following Fig. 5.16 is determined by a soft composition of two components $a_r^{\mathrm{agr}}\left(\bar{\mathbf{x}}, \bar{y}^r\right)$ and $b_r^{\mathrm{agr}}\left(\bar{\mathbf{x}}, \bar{y}^r\right)$. Component $a_r^{\mathrm{agr}}\left(\bar{\mathbf{x}}, \bar{y}^r\right)$ is the average of quasi-implications $I_{k,r}\left(\bar{\mathbf{x}}, \bar{y}^r\right)$, $k=1,\ldots,N$, $r = 1,\ldots,N$, determined in points $\bar{y}^r$, $r = 1,\ldots,N$, at which consequent membership functions μ_{B^k}, $k = 1,\ldots,N$, attain their maximum equal to 1. Component $b_r^{\mathrm{agr}}\left(\bar{\mathbf{x}}, \bar{y}^r\right)$ is determined analogously as in Section 5.5. The dominance of one component over another depends on the value of parameter α^{agr}.

In Fig. 5.17-5.22 we depict the shape of the hyperplanes described by formulas (5.42)-(5.44) for $n = 2$ and $N = 2$. In all the cases the function H is generated by the Dombi triangular norms.

5.7. WEIGHTED FLEXIBLE SYSTEMS

Our concept of weighted triangular norms, presented in Section 4.2, allows to assign different degrees of credibility (importance) to fuzzy rules. Moreover, we incorporate weights to describe the significance of particular antecedents in all rules. More precisely, we insert:

(i) weights to the aggregation operator of the rules $w_k^{\mathrm{agr}} \in [0,1]$, $k = 1,\ldots,N$,

(ii) weights to the antecedents $w_{i,k}^{\tau} \in i[0,1]$, $i= 1,\ldots,n$, $k = 1,\ldots,N$

in system ORII.

Consequently, we get the weighted soft NFIS of the OR-type:

OR III

$$\tau_k\left(\bar{\mathbf{x}}\right) = \begin{pmatrix} \left(1-\alpha^{\tau}\right)\mathrm{avg}\left(\mu_{A_1^k}\left(\bar{x}_1\right),\ldots,\mu_{A_n^k}\left(\bar{x}_n\right)\right)+ \\ +\alpha^{\tau}\ddot{H}^*\begin{pmatrix}\mu_{A_1^k}\left(\bar{x}_1\right),\ldots,\mu_{A_n^k}\left(\bar{x}_n\right);\\ w_{1,k}^{\tau},\ldots,w_{n,k}^{\tau},p^{\tau},0\end{pmatrix}\end{pmatrix} \tag{5.45}$$

$$I_{k,r}\left(\bar{\mathbf{x}}, \bar{y}^r\right) = \begin{pmatrix} \left(1-\alpha^{I}\right)\mathrm{avg}\left(\tilde{N}_{1-\nu}\left(\tau_k\left(\bar{\mathbf{x}}\right)\right),\mu_{B^k}\left(\bar{y}^r\right)\right)+ \\ +\alpha^{I}\ddot{H}\begin{pmatrix}\tilde{N}_{1-\nu}\left(\tau_k\left(\bar{\mathbf{x}}\right)\right),\mu_{B^k}\left(\bar{y}^r\right);\\ p^{I},\nu\end{pmatrix}\end{pmatrix} \tag{5.46}$$

$$\mathrm{agr}_r(\bar{\mathbf{x}},\bar{y}^r)=\begin{pmatrix}(1-\alpha^{\mathrm{agr}})\mathrm{avg}(I_{1,r}(\bar{\mathbf{x}},\bar{y}^r),\ldots,I_{N,r}(\bar{\mathbf{x}},\bar{y}^r))+\\ +\alpha^{\mathrm{agr}}\ddot{H}^*\begin{pmatrix}I_{1,r}(\bar{\mathbf{x}},\bar{y}^r),\ldots,I_{N,r}(\bar{\mathbf{x}},\bar{y}^r)\ ;\\ w_1^{\mathrm{agr}},\ldots,w_N^{\mathrm{agr}},p^{\mathrm{agr}},1-\nu\end{pmatrix}\end{pmatrix} \tag{5.47}$$

In the ORIII system we use parameterised families $\ddot{H}(\cdot)$ and parameterised families with weights $\ddot{H}^*(\cdot)$ analogously to formula (4.1). More specifically, in (5.45) and (5.47) we use the following definition

$$\ddot{H}^*\begin{pmatrix}a_1,\ldots,a_n;\\ w_1,\ldots,w_n,p,\nu\end{pmatrix}=\ddot{H}\begin{pmatrix}\mathrm{arg}_1(a_1,w_1,\nu),\ldots,\mathrm{arg}_n(a_n,w_n,\nu)\ ;\\ p,\nu\end{pmatrix} \tag{5.48}$$

where

$$\mathrm{arg}_i(a_i,w_i,\nu)=(1-\nu)(1-w_i(1-a_i))+\nu w_i a_i \tag{5.49}$$

It is easily seen that the above system reduces to a soft compromise system described in Section 5.6 if $w_{i,k}^{\tau}=1$, $i=1,\ldots,n$, $k=1,\ldots,N$ and $w_k^{\mathrm{agr}}=1$, $k=1,\ldots,N$.

In Fig. 5.23 and 5.24 we depict block diagrams realizing the quasi-implication $I_{k,r}(\bar{\mathbf{x}},\bar{y}^r)$ given by (5.46) and aggregation $\mathrm{agr}_r(\bar{\mathbf{x}},\bar{y}^r)$ given by (5.47), respectively. The block diagram realizing the aggregation of antecedents $\tau_k(\bar{\mathbf{x}})$ is shown in Fig. 5.14.

The firing strength of rules $\tau_k(\bar{\mathbf{x}})$, $k=1,...,N$, following Fig. 5.23, is determined by a soft composition, controlled by parameter α^{τ}, of two components $a_k^{\tau}(\bar{\mathbf{x}})$ and $b_k^{\tau}(\bar{\mathbf{x}})$. Component $a_k^{\tau}(\bar{\mathbf{x}})$ is the arithmetic average of antecedent membership function values $\mu_{A_i^k}(\bar{x}_i)$, $k=1,...,N$, $i=1,...,n$, determined for all inputs $\bar{x}_i i$, $=1,\ldots,n$. Component $b_k^{\tau}(\bar{\mathbf{x}})$ is determined by the weighted aggregation (see Section 4.2) with weights $w_{i,k}^{\tau}$, based on an H-function parameterized by $\nu=0$, of antecedent membership function values $\mu_{A_i^k}(\bar{x}_i)$, $k=1,\ldots,N$, $i=1,\ldots,n$, calculated for all inputs $\bar{x}_i$, $i=1,...,n$. The dominance of one component over another depends on the value of parameter α^{τ}. The weights $w_{i,k}^{\tau}$ can be interpreted as credibilities of antecedents and can be applied to evaluate the importance of input linguistic labels.

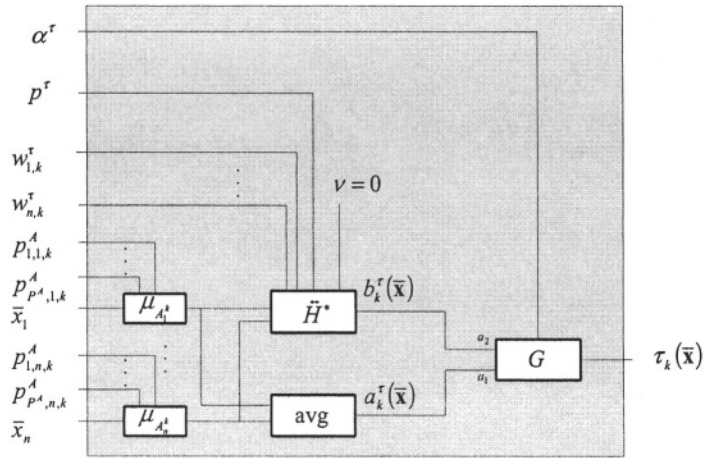

Fig. 5.23. Connection of antecedents in weighted soft OR-type NFIS (OR III)

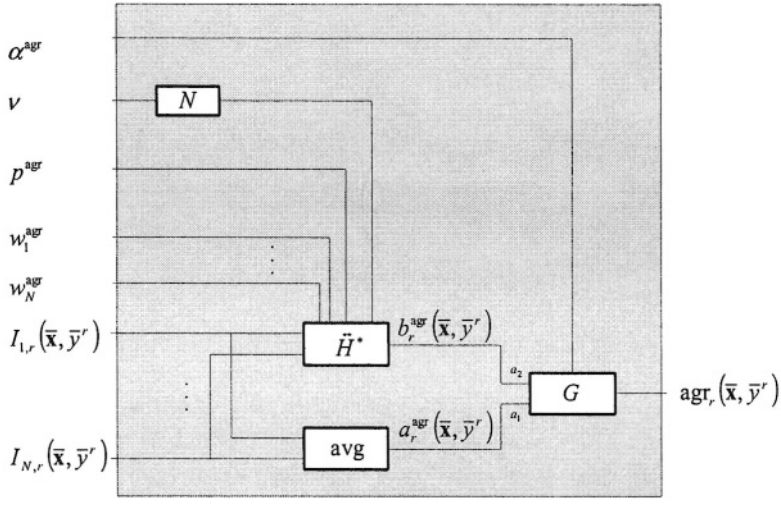

Fig. 5.24. Aggregation of rules in weighted soft OR-type NFIS (OR III)

The aggregation operator $\text{agr}_r(\bar{\mathbf{x}},\bar{y}^r)$, $r = 1,\ldots,N$, following Fig. 5.24, is determined by a soft composition of two components $a_r^{agr}(\bar{\mathbf{x}},\bar{y}^r)$ and $b_r^{agr}(\bar{\mathbf{x}},\bar{y}^r)$. Component $a_r^{agr}(\bar{\mathbf{x}},\bar{y}^r)$ is the average of quasi-implications $I_{k,r}(\bar{\mathbf{x}},\bar{y}^r)$, $k = 1,\ldots,N$, $r = 1,\ldots,N$, determined in points $\bar{y}^r$, $r = 1,\ldots,N$, at which consequent membership functions μ_{B^k}, $k = 1,\ldots,N$, attain their maximum equal to 1. Component $b_r^{agr}(\bar{\mathbf{x}},\bar{y}^r)$ is determined analogously as in

Section 5.5, however certainty weights w_k^{agr}, $k = 1,\ldots, N$, are additionally incorporated into the NFIS structure. The dominance of one component over another depends on the value of parameter α^{agr}. Weights w_k^{agr} can be interpreted as credibilities of rules.

In Fig. 5.25-5.28 we depict the shape of the hyperplanes described by formulas (5.45)-(5.47) for $n = 2$ and $N = 2$. In all the cases the function H is generated by the Dombi triangular norms.

5.8. LEARNING ALGORITHMS

It is well known that the basic concept of the backpropagation algorithm, commonly used to train neural networks, can also be applied to any feedforward network. Let $\bar{\mathbf{x}}(t) \in \mathbf{R}^n$ and $d(t) \in \mathbf{R}$ be a sequence of inputs and desirable output signals, respectively. Based on the learning sequence $(\bar{\mathbf{x}}(1), d(1)), (\bar{\mathbf{x}}(2), d(2)), \ldots$ we wish to determine all parameters (including the system's type ν) and weights of NFIS such that

$$e(t) = \frac{1}{2}\left[f(\bar{\mathbf{x}}(t)) - d(t)\right]^2 \tag{5.50}$$

is minimized, were $f(\cdot)$ is given (3.48). The steepest descent optimization algorithm can be applied to solve this problem. For instance, parameters $\bar{y}^r$, $r = 1,\ldots, N$, are trained by the iterative procedure

$$\bar{y}^r(t+1) = \bar{y}^r(t) - \eta \frac{\partial e(t)}{\partial \bar{y}^r(t)} \tag{5.51}$$

Directly calculating partial derivatives in recursion (5.51) is rather complicated. Therefore, we recall that our NFIS has a layered architecture (see Fig. 5.29) and apply the idea of the backpropagation method to train the system. The exact recursions derived in this section reflect that idea, however, they are not a copy of the standard backpropagation used to train multi-layer neural networks. We will apply the gradient optimization with constraints in order to optimize:

- $\nu \in [0,1]$,
- $\alpha^r \in [0,1]$, $\alpha^l \in [0,1]$, $\alpha^{\text{agr}} \in [0,1]$,
- $p^r \in [0,\infty)$, $p^l \in [0,\infty)$, $p^{\text{agr}} \in [0,\infty)$,
- $w_{i,k}^r \in [0,1]$, $i = 1,\ldots, n$, $k = 1,\ldots, N$,
- $w_k^{\text{agr}} \in [0,1]$, $k = 1,\ldots, N$.

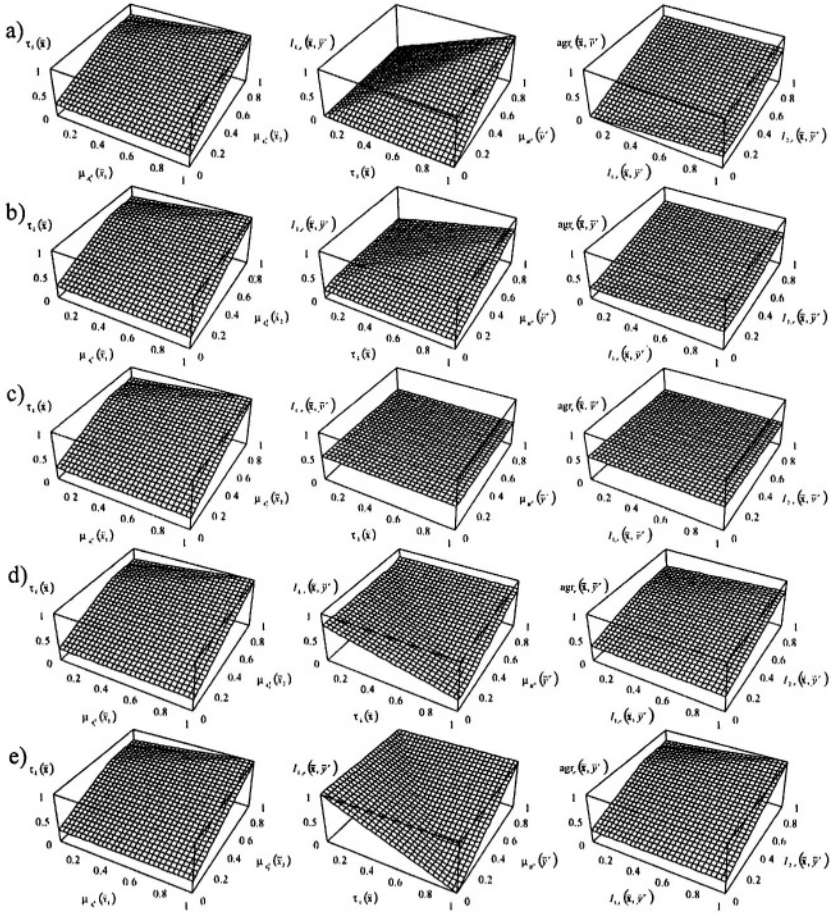

Fig. 5.25. 3D plots of weighted soft OR-type NFIS (OR III) for $n = 2$, $N = 2$, $\alpha^{\tau} = \alpha^{I} = \alpha^{agr} = 1$, $p^{\tau} = p^{I} = p^{agr} = 10$, $w_{1,k}^{\tau} = w_1^{agr} = 0.25$, $w_{2,k}^{\tau} = w_2^{agr} = 0.75$, and a) $\nu = 0.00$, b) $\nu = 0.15$, c) $\nu = 0.50$, d) $\nu = 0.85$, e) $\nu = 1.00$

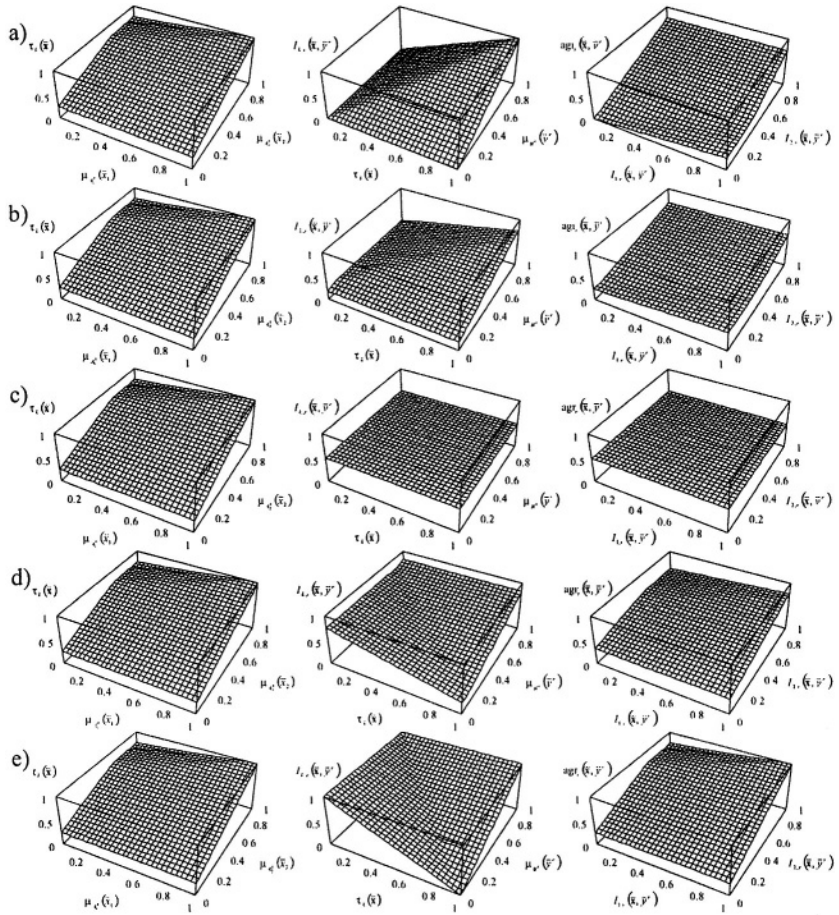

Fig. 5.26. 3D plots of weighted soft OR-type NFIS (OR III) for $n = 2$, $N = 2$, $\alpha^{\tau} = \alpha^{I} = \alpha^{agr} = 0.5$, $p^{\tau} = p^{I} = p^{agr} = 10$, $w_{1,k}^{\tau} = w_{1}^{agr} = 0.25$, $w_{2,k}^{\tau} = w_{2}^{agr} = 0.75$, and a) $\nu = 0.00$, b) $\nu = 0.15$, c) $\nu = 0.50$, d) $\nu = 0.85$, e) $\nu = 1.00$

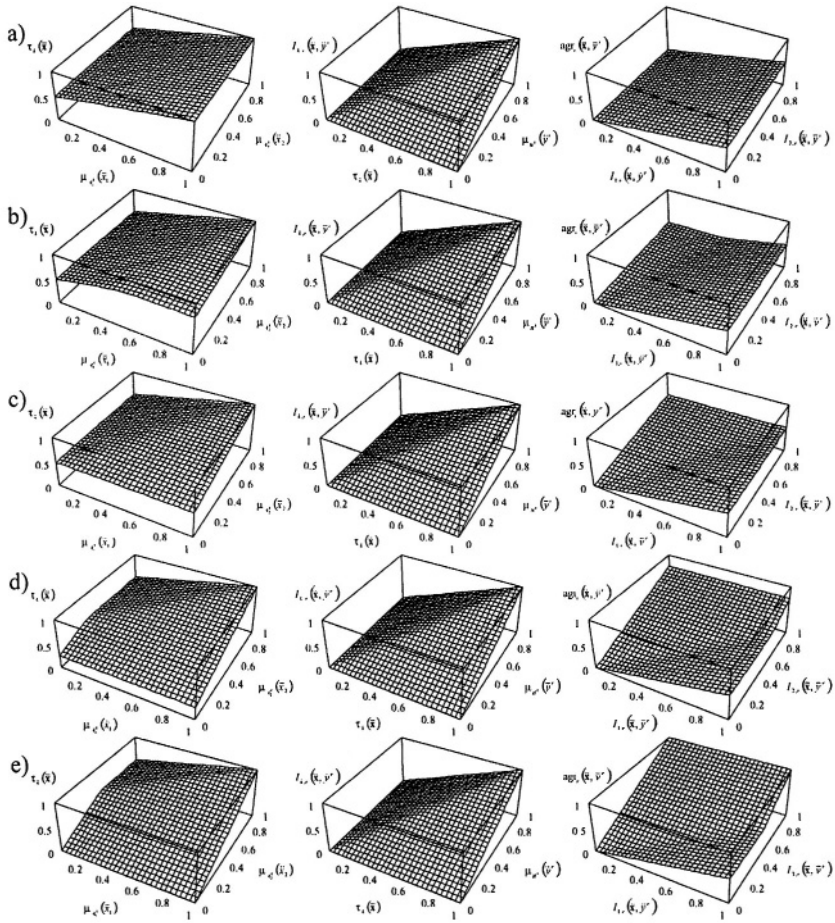

Fig. 5.27. 3D plots of weighted soft OR-type NFIS (OR III) for $n = 2$, $N = 2$, $\alpha^{\tau} = \alpha^{I} = \alpha^{agr} = 1$, $p^{\tau} = p^{I} = p^{agr} = 10$, $v = 0$ $w_{1,k}^{\tau} = w_1^{agr} = 0.50$, and a) $w_{2,k}^{\tau} = w_2^{agr} = 0.00$, b) $w_{2,k}^{\tau} = w_2^{agr} = 0.25$, c) $w_{2,k}^{\tau} = w_2^{agr} = 0.50$, d) $w_{2,k}^{\tau} = w_2^{agr} = 0.75$, e) $w_{2,k}^{\tau} = w_2^{agr} = 1.00$

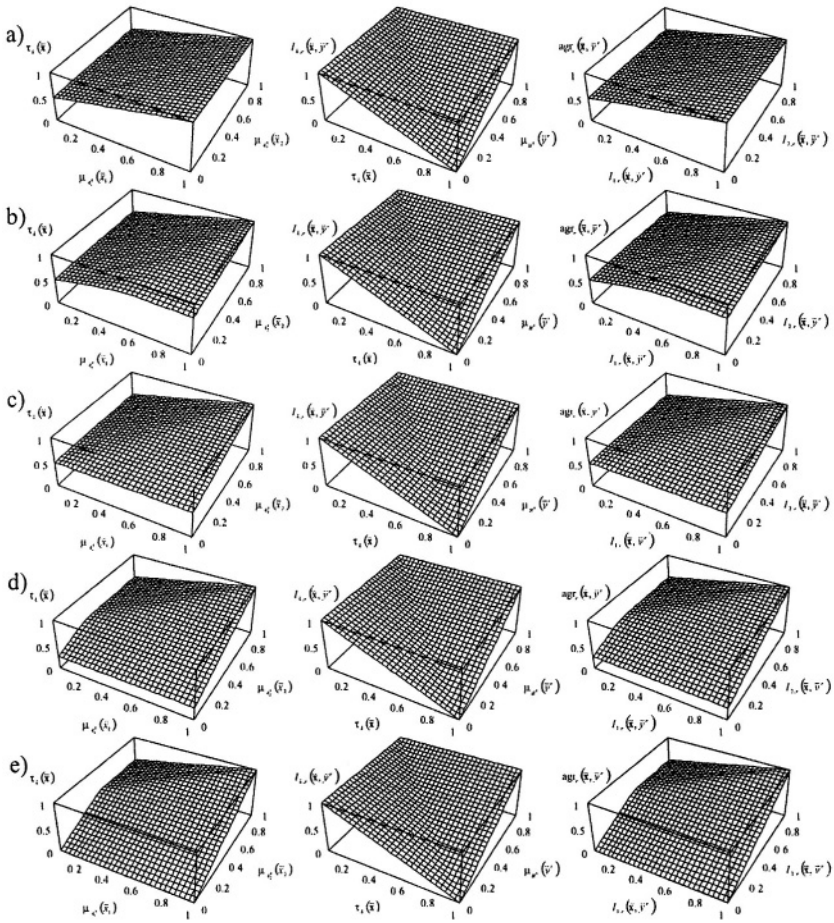

Fig. 5.28. 3D plots of weighted soft OR-type NFIS (OR III) for $n = 2$, $N = 2$, $\alpha^{\tau} = \alpha^{I} = \alpha^{\text{agr}} = 1$, $p^{\tau} = p^{I} = p^{\text{agr}} = 10$, $\nu = 1$ $w_{1,k}^{\tau} = w_1^{\text{agr}} = 0.50$, and a) $w_{2,k}^{\tau} = w_2^{\text{agr}} = 0.00$, b) $w_{2,k}^{\tau} = w_2^{\text{agr}} = 0.25$, c) $w_{2,k}^{\tau} = w_2^{\text{agr}} = 0.50$, d) $w_{2,k}^{\tau} = w_2^{\text{agr}} = 0.75$, e) $w_{2,k}^{\tau} = w_2^{\text{agr}} = 1.00$

The constraint functions, see formulas (33) and (35) in the Appendix, will be used to perform the optimization of the above parameters and weights. Moreover, we will find in the process of learning parameters of the membership functions $\mu_{A_i^k}(x_i)$ and $\mu_{B^k}(y)$, $i = 1,\ldots,n$, $k = 1,\ldots,N$:

- $p_{u,i,k}^{A}$, $u = 1,\ldots,P^{A}$, $i = 1,\ldots,n$, $k = 1,\ldots,N$,
- $p_{u,k}^{B}$, $u = 2,\ldots,P^{B}$, $k = 1,\ldots,N$,
- $p_{1,k}^{B} = \bar{y}^{k}$, $k = 1,\ldots,N$.

Remark 5.2: We will explain the notation used in this book (Chapter 5 and 6) on a simple example of a single neuron given by:

$$y = f(s),\ s = \sum_{i=1}^{n} x_i w_i$$

where f is a sigmoidal function, x_i and w_i, $i = 1,\ldots,n$, are inputs and weights, respectively. Let d be the desired output signal. Then we write

$$\varepsilon^{f} = \varepsilon = y - d$$

and

$$\varepsilon^{s} = \varepsilon^{f}\{s\} = \varepsilon^{f}\frac{\partial f(s)}{\partial s} = (y-d)f'(s)$$

i.e. $\varepsilon^{f}\{s\}$ is the error transferred from the output block f to the summation block s. We will use the analogous notation depicted in Fig. 5.29.

a) General learning procedures

Parameter ν and other parameters of the NFIS are updated by the recursive procedures:

$$\nu(t+1) = \nu(t) - \eta\Delta\nu(t) \tag{5.52}$$

$$\alpha^{\tau}(t+1) = \alpha^{\tau}(t) - \eta\Delta\alpha^{\tau}(t) \tag{5.53}$$

$$\alpha^{l}(t+1) = \alpha^{l}(t) - \eta\Delta\alpha^{l}(t) \tag{5.54}$$

$$\alpha^{\mathrm{agr}}(t+1) = \alpha^{\mathrm{agr}}(t) - \eta\Delta\alpha^{\mathrm{agr}}(t) \tag{5.55}$$

$$p^{\tau}(t+1) = p^{\tau}(t) - \eta\Delta p^{\tau}(t) \tag{5.56}$$

$$p^{l}(t+1) = p^{l}(t) - \eta\Delta p^{l}(t) \tag{5.57}$$

$$p^{\mathrm{agr}}(t+1) = p^{\mathrm{agr}}(t) - \eta\Delta p^{\mathrm{agr}}(t) \tag{5.58}$$

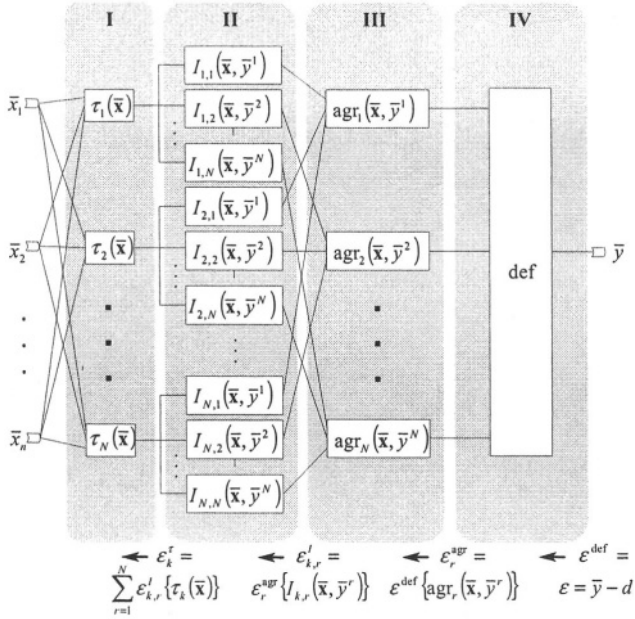

Fig. 5.29. The scheme of NFIS

$$w_{i,k}^{\tau}(t+1) = w_{i,k}^{\tau}(t) - \eta \Delta w_{i,k}^{\tau}(t) \tag{5.59}$$

$$w_r^{\text{agr}}(t+1) = w_r^{\text{agr}}(t) - \eta \Delta w_r^{\text{agr}}(t) \tag{5.60}$$

$$p_{u,i,k}^{A}(t+1) = p_{u,i,k}^{A}(t) - \eta \Delta p_{u,i,k}^{A}(t) \tag{5.61}$$

$$p_{u,k}^{B}(t+1) = p_{u,k}^{B}(t) - \eta \Delta p_{u,k}^{B}(t) \;;\; u = 2,\ldots,P^B \tag{5.62}$$

$$\bar{y}^r(t+1) = p_{1,r}^{B}(t+1) = \bar{y}^r(t) - \eta \Delta \bar{y}^r(t) \tag{5.63}$$

where corresponding corrections Δ are given by:

$$\Delta v = \sum_{k=1}^{N}\sum_{r=1}^{N} \varepsilon_{k,r}^{I}\{v\} + \sum_{r=1}^{N} \varepsilon_r^{\text{agr}}\{v\} \tag{5.64}$$

$$\Delta \alpha^{\tau} = \sum_{k=1}^{N} \varepsilon_k^{\tau}\{\alpha^{\tau}\} \tag{5.65}$$

$$\Delta \alpha^{I} = \sum_{k=1}^{N}\sum_{r=1}^{N} \varepsilon_{k,r}^{I}\{\alpha^{I}\} \tag{5.66}$$

$$\Delta \alpha^{\text{agr}} = \sum_{r=1}^{N} \varepsilon_r^{\text{agr}}\{\alpha^{\text{agr}}\} \tag{5.67}$$

$$\Delta p^{\tau} = \sum_{k=1}^{N} \varepsilon_k^{\tau}\left\{p^{\tau}\right\} \tag{5.68}$$

$$\Delta p^{I} = \sum_{k=1}^{N}\sum_{r=1}^{N} \varepsilon_{k,r}^{I}\left\{p^{I}\right\} \tag{5.69}$$

$$\Delta p^{\text{agr}} = \sum_{r=1}^{N} \varepsilon_r^{\text{agr}}\left\{p^{\text{agr}}\right\} \tag{5.70}$$

$$\Delta w_{i,k}^{\tau} = \varepsilon_k^{\tau}\left\{w_{i,k}^{\tau}\right\} \tag{5.71}$$

$$\Delta w_r^{\text{agr}} = \varepsilon_r^{\text{agr}}\left\{w_r^{\text{agr}}\right\} \tag{5.72}$$

$$\Delta p_{u,i,k}^{A} = \varepsilon_k^{\tau}\left\{p_{u,i,k}^{A}\right\} \tag{5.73}$$

$$\Delta p_{u,k}^{B} = \sum_{r=1}^{N} \varepsilon_{k,r}^{I}\left\{p_{u,k}^{B}\right\} \;;\; u = 2,\ldots,P^{B} \tag{5.74}$$

$$\Delta \bar{y}^{r} = \Delta p_{1,r}^{B} = \varepsilon^{\text{def}}\left\{\bar{y}^{r}\right\} + \sum_{k=1}^{N} \varepsilon_{k,r}^{I}\left\{\bar{y}^{r}\right\} + \sum_{k=1}^{N} \varepsilon_{r,k}^{I}\left\{p_{1,r}^{B}\right\} \tag{5.75}$$

The errors propagated through the net are indicated at the bottom of Fig. 5.29 and satisfy the following relations:

$$\varepsilon_k^{\tau} = \sum_{r=1}^{N} \varepsilon_{k,r}^{I}\left\{\tau_k(\bar{\mathbf{x}})\right\} \tag{5.76}$$

$$\varepsilon_{k,r}^{I} = \varepsilon_r^{\text{agr}}\left\{I_{k,r}(\bar{\mathbf{x}}, \bar{y}^{r})\right\} \tag{5.77}$$

$$\varepsilon_r^{\text{agr}} = \varepsilon^{\text{def}}\left\{\text{agr}_r(\bar{\mathbf{x}}, \bar{y}^{r})\right\} \tag{5.78}$$

$$\varepsilon^{\text{def}} = \varepsilon = \bar{y} - d \tag{5.79}$$

b) Block of rules' activation

The errors propagated through the block of rules' activation (see Fig. 5.30) are given by:

$$\varepsilon_k^{\tau}\left\{\alpha^{\tau}\right\} = \varepsilon_k^{\tau}\frac{\partial \tau_k(\bar{\mathbf{x}})}{\partial \alpha^{\tau}} \tag{5.80}$$

$$\varepsilon_k^{\tau}\left\{p^{\tau}\right\} = \varepsilon_k^{\tau}\frac{\partial \tau_k(\bar{\mathbf{x}})}{\partial b_k^{\tau}(\bar{\mathbf{x}})}\frac{\partial b_k^{\tau}(\bar{\mathbf{x}})}{\partial p^{\tau}} \tag{5.81}$$

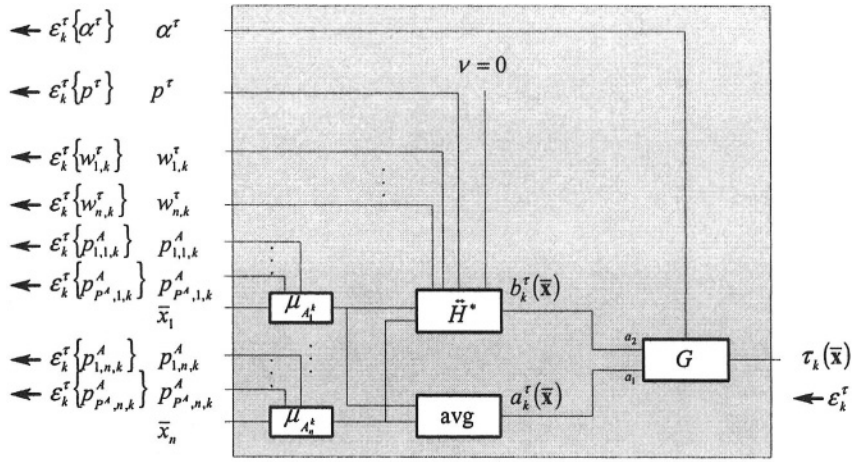

Fig. 5.30. Block of rules' activation

$$\varepsilon_k^\tau\left\{w_{i,k}^\tau\right\} = \varepsilon_k^\tau \frac{\partial \tau_k(\bar{\mathbf{x}})}{\partial b_k^\tau(\bar{\mathbf{x}})} \frac{\partial b_k^\tau(\bar{\mathbf{x}})}{\partial w_{i,k}^\tau} \tag{5.82}$$

$$\varepsilon_k^\tau\left\{p_{u,i,k}^A\right\} = \varepsilon_k^\tau \left(\frac{\partial \tau_k(\bar{\mathbf{x}})}{\partial b_k^\tau(\bar{\mathbf{x}})} \frac{\partial b_k^\tau(\bar{\mathbf{x}})}{\partial p_{u,i,k}^A} + \frac{\partial \tau_k(\bar{\mathbf{x}})}{\partial a_k^\tau(\bar{\mathbf{x}})} \frac{\partial a_k^\tau(\bar{\mathbf{x}})}{\partial p_{u,i,k}^A} \right) \frac{\partial \mu_{A_i^t}(\bar{x}_i)}{\partial p_{u,i,k}^A} \tag{5.83}$$

where

$$\frac{\partial \tau_k(\bar{\mathbf{x}})}{\partial a_k^\tau(\bar{\mathbf{x}})} = \frac{\partial}{\partial a_k^\tau(\bar{\mathbf{x}})} G\left(\begin{matrix} a_k^\tau(\bar{\mathbf{x}}), b_k^\tau(\bar{\mathbf{x}}) \ ; \\ \alpha^\tau \end{matrix} \right) \tag{5.84}$$

$$\frac{\partial \tau_k(\bar{\mathbf{x}})}{\partial b_k^\tau(\bar{\mathbf{x}})} = \frac{\partial}{\partial b_k^\tau(\bar{\mathbf{x}})} G\left(\begin{matrix} a_k^\tau(\bar{\mathbf{x}}), b_k^\tau(\bar{\mathbf{x}}) \ ; \\ \alpha^\tau \end{matrix} \right) \tag{5.85}$$

$$\frac{\partial \tau_k(\bar{\mathbf{x}})}{\partial \alpha^\tau} = \frac{\partial}{\partial \alpha^\tau} G\left(\begin{matrix} a_k^\tau(\bar{\mathbf{x}}), b_k^\tau(\bar{\mathbf{x}}) \ ; \\ \alpha^\tau \end{matrix} \right) \tag{5.86}$$

$$\frac{\partial a_k^\tau(\bar{\mathbf{x}})}{\partial \mu_{A_i^t}(\bar{x}_i)} = \frac{\partial}{\partial \mu_{A_i^t}(\bar{x}_i)} \operatorname{avg}\left(\mu_{A_1^t}(\bar{x}_1), \ldots, \mu_{A_n^t}(\bar{x}_n)\right) \tag{5.87}$$

$$\frac{\partial b_k^{\tau}(\bar{\mathbf{x}})}{\partial p^{\tau}} = \frac{\partial}{\partial p^{\tau}} \ddot{H}^* \begin{pmatrix} \mu_{A_1^k}(\bar{x}_1), \ldots, \mu_{A_n^k}(\bar{x}_n) ; \\ w_{1,k}^{\tau}, \ldots, w_{n,k}^{\tau}, p^{\tau}, 0 \end{pmatrix} \tag{5.88}$$

$$\frac{\partial b_k^{\tau}(\bar{\mathbf{x}})}{\partial w_{i,k}^{\tau}} = \frac{\partial}{\partial w_{i,k}^{\tau}} \ddot{H}^* \begin{pmatrix} \mu_{A_1^k}(\bar{x}_1), \ldots, \mu_{A_n^k}(\bar{x}_n) ; \\ w_{1,k}^{\tau}, \ldots, w_{n,k}^{\tau}, p^{\tau}, 0 \end{pmatrix} \tag{5.89}$$

$$\frac{\partial b_k^{\tau}(\bar{\mathbf{x}})}{\partial \mu_{A_i^k}(\bar{x}_i)} = \frac{\partial}{\partial \mu_{A_i^k}(\bar{x}_i)} \ddot{H}^* \begin{pmatrix} \mu_{A_1^k}(\bar{x}_1), \ldots, \mu_{A_n^k}(\bar{x}_n) ; \\ w_{1,k}^{\tau}, \ldots, w_{n,k}^{\tau}, p^{\tau}, 0 \end{pmatrix} \tag{5.90}$$

The exact form of (5.84)-(5.90) and other derivatives depends on the used H-function and input-output membership functions. We explain details and give some examples in the Appendix.

c) Block of implications

The errors propagated through the block of implications (see Fig. 5.31) are determined as follows:

$$\varepsilon_{k,r}^{I}\{\nu\} = \varepsilon_{k,r}^{I} \left(\begin{array}{l} \dfrac{\partial I_{k,r}(\bar{\mathbf{x}}, \bar{y}^r)}{\partial b_{k,r}^{I}(\bar{\mathbf{x}}, \bar{y}^r)} \dfrac{\partial b_{k,r}^{I}(\bar{\mathbf{x}}, \bar{y}^r)}{\partial \nu} + \\ + \left(\dfrac{\partial I_{k,r}(\bar{\mathbf{x}}, \bar{y}^r)}{\partial b_{k,r}^{I}(\bar{\mathbf{x}}, \bar{y}^r)} \dfrac{\partial b_{k,r}^{I}(\bar{\mathbf{x}}, \bar{y}^r)}{\partial \widetilde{N}_{1-\nu}(\tau_k(\bar{\mathbf{x}}))} + \right. \\ \left. + \dfrac{\partial I_{k,r}(\bar{\mathbf{x}}, \bar{y}^r)}{\partial a_{k,r}^{I}(\bar{\mathbf{x}}, \bar{y}^r)} \dfrac{\partial a_{k,r}^{I}(\bar{\mathbf{x}}, \bar{y}^r)}{\partial \widetilde{N}_{1-\nu}(\tau_k(\bar{\mathbf{x}}))} \right) \cdot \\ \cdot \dfrac{\partial \widetilde{N}_{1-\nu}(\tau_k(\bar{\mathbf{x}}))}{\partial (1-\nu)} \dfrac{\partial N(\nu)}{\partial \nu} \end{array} \right) \tag{5.91}$$

$$\varepsilon_{k,r}^{I}\{\alpha^I\} = \varepsilon_{k,r}^{I} \frac{\partial I_{k,r}(\bar{\mathbf{x}}, \bar{y}^r)}{\partial \alpha^I} \tag{5.92}$$

$$\varepsilon_{k,r}^{I}\{p^I\} = \varepsilon_{k,r}^{I} \frac{\partial I_{k,r}(\bar{\mathbf{x}}, \bar{y}^r)}{\partial b_{k,r}^{I}(\bar{\mathbf{x}}, \bar{y}^r)} \frac{\partial b_{k,r}^{I}(\bar{\mathbf{x}}, \bar{y}^r)}{\partial p^I} \tag{5.93}$$

$$\varepsilon_{k,r}^{I}\{p_{u,k}^{B}\} = \varepsilon_{k,r}^{I} \left(\begin{array}{l} \dfrac{\partial I_{k,r}(\bar{\mathbf{x}}, \bar{y}^r)}{\partial b_{k,r}^{I}(\bar{\mathbf{x}}, \bar{y}^r)} \dfrac{\partial b_{k,r}^{I}(\bar{\mathbf{x}}, \bar{y}^r)}{\partial \mu_{B^k}(\bar{y}^r)} + \\ + \dfrac{\partial I_{k,r}(\bar{\mathbf{x}}, \bar{y}^r)}{\partial a_{k,r}^{I}(\bar{\mathbf{x}}, \bar{y}^r)} \dfrac{\partial a_{k,r}^{I}(\bar{\mathbf{x}}, \bar{y}^r)}{\partial \mu_{B^k}(\bar{y}^r)} \end{array} \right) \frac{\partial \mu_{B^k}(\bar{y}^r)}{\partial p_{u,k}^{B}} \tag{5.94}$$

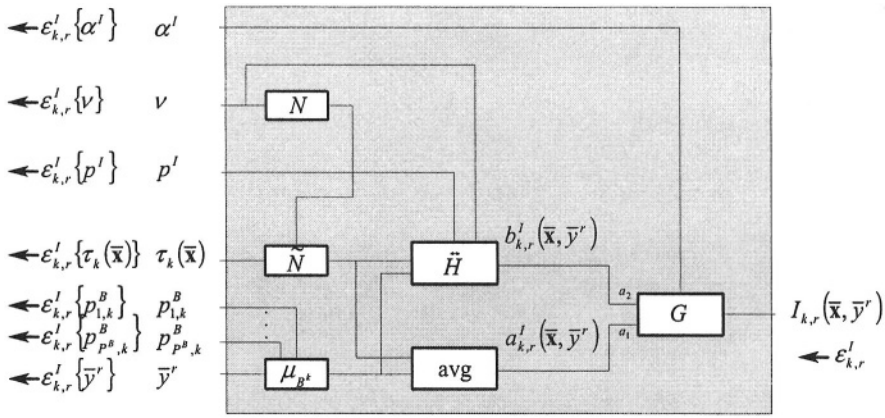

Fig. 5.31. Block of implications

$$\varepsilon^I_{k,r}\{\bar{y}^r\}=\varepsilon^I_{k,r}\left(\begin{array}{l}\dfrac{\partial I_{k,r}(\bar{\mathbf{x}},\bar{y}^r)}{\partial b^I_{k,r}(\bar{\mathbf{x}},\bar{y}^r)}\dfrac{\partial b^I_{k,r}(\bar{\mathbf{x}},\bar{y}^r)}{\partial \mu_{B^k}(\bar{y}^r)}+\\ +\dfrac{\partial I_{k,r}(\bar{\mathbf{x}},\bar{y}^r)}{\partial a^I_{k,r}(\bar{\mathbf{x}},\bar{y}^r)}\dfrac{\partial a^I_{k,r}(\bar{\mathbf{x}},\bar{y}^r)}{\partial \mu_{B^k}(\bar{y}^r)}\end{array}\right)\frac{\partial \mu_{B^k}(\bar{y}^r)}{\partial \bar{y}^r} \tag{5.95}$$

$$\varepsilon^I_{k,r}\{\tau_k(\bar{\mathbf{x}})\}=\varepsilon^I_{k,r}\left(\begin{array}{l}\dfrac{\partial I_{k,r}(\bar{\mathbf{x}},\bar{y}^r)}{\partial b^I_{k,r}(\bar{\mathbf{x}},\bar{y}^r)}\dfrac{\partial b^I_{k,r}(\bar{\mathbf{x}},\bar{y}^r)}{\partial \tilde{N}_{1-\nu}(\tau_k(\bar{\mathbf{x}}))}+\\ +\dfrac{\partial I_{k,r}(\bar{\mathbf{x}},\bar{y}^r)}{\partial a^I_{k,r}(\bar{\mathbf{x}},\bar{y}^r)}\dfrac{\partial a^I_{k,r}(\bar{\mathbf{x}},\bar{y}^r)}{\partial \tilde{N}_{1-\nu}(\tau_k(\bar{\mathbf{x}}))}\end{array}\right)\frac{\partial \tilde{N}_{1-\nu}(\tau_k(\bar{\mathbf{x}}))}{\partial \tau_k(\bar{\mathbf{x}})} \tag{5.96}$$

where

$$\frac{\partial I_{k,r}(\bar{\mathbf{x}},\bar{y}^r)}{\partial a^I_{k,r}(\bar{\mathbf{x}},\bar{y}^r)}=\frac{\partial}{\partial a^I_{k,r}(\bar{\mathbf{x}},\bar{y}^r)}G\left(\begin{array}{c}a^I_{k,r}(\bar{\mathbf{x}},\bar{y}^r),b^I_{k,r}(\bar{\mathbf{x}},\bar{y}^r);\\ \alpha^I\end{array}\right) \tag{5.97}$$

$$\frac{\partial I_{k,r}(\bar{\mathbf{x}},\bar{y}^r)}{\partial b^I_{k,r}(\bar{\mathbf{x}},\bar{y}^r)}=\frac{\partial}{\partial b^I_{k,r}(\bar{\mathbf{x}},\bar{y}^r)}G\left(\begin{array}{c}a^I_{k,r}(\bar{\mathbf{x}},\bar{y}^r),b^I_{k,r}(\bar{\mathbf{x}},\bar{y}^r);\\ \alpha^I\end{array}\right) \tag{5.98}$$

$$\frac{\partial I_{k,r}(\bar{\mathbf{x}},\bar{y}^r)}{\partial \alpha^I}=\frac{\partial}{\partial \alpha^I}G\left(\begin{array}{c}a^I_{k,r}(\bar{\mathbf{x}},\bar{y}^r),b^I_{k,r}(\bar{\mathbf{x}},\bar{y}^r);\\ \alpha^I\end{array}\right) \tag{5.99}$$

$$\frac{\partial a^I_{k,r}(\bar{\mathbf{x}},\bar{y}^r)}{\partial \tilde{N}_{1-\nu}(\tau_k(\bar{\mathbf{x}}))}=\frac{\partial}{\partial \tilde{N}_{1-\nu}(\tau_k(\bar{\mathbf{x}}))}\mathrm{avg}\left(\tilde{N}_{1-\nu}(\tau_k(\bar{\mathbf{x}})),\mu_{B^k}(\bar{y}^r)\right) \tag{5.100}$$

$$\frac{\partial a_{k,r}^{l}\left(\bar{\mathbf{x}},\bar{y}^{r}\right)}{\partial \mu_{B^{k}}\left(\bar{y}^{r}\right)}=\frac{\partial}{\partial \mu_{B^{k}}\left(\bar{y}^{r}\right)}\mathrm{avg}\left(\tilde{N}_{1-\nu}\left(\tau_{k}\left(\bar{\mathbf{x}}\right)\right),\mu_{B^{k}}\left(\bar{y}^{r}\right)\right) \tag{5.101}$$

$$\frac{\partial b_{k,r}^{l}\left(\bar{\mathbf{x}},\bar{y}^{r}\right)}{\partial \nu}=\frac{\partial}{\partial \nu}\ddot{H}\left(\begin{matrix}\tilde{N}_{1-\nu}\left(\tau_{k}\left(\bar{\mathbf{x}}\right)\right),\mu_{B^{k}}\left(\bar{y}^{r}\right);\\ p^{l},\nu\end{matrix}\right) \tag{5.102}$$

$$\frac{\partial b_{k,r}^{l}\left(\bar{\mathbf{x}},\bar{y}^{r}\right)}{\partial p^{l}}=\frac{\partial}{\partial p^{l}}\ddot{H}\left(\begin{matrix}\tilde{N}_{1-\nu}\left(\tau_{k}\left(\bar{\mathbf{x}}\right)\right),\mu_{B^{k}}\left(\bar{y}^{r}\right);\\ p^{l},\nu\end{matrix}\right) \tag{5.103}$$

$$\frac{\partial b_{k,r}^{l}\left(\bar{\mathbf{x}},\bar{y}^{r}\right)}{\partial \tilde{N}_{1-\nu}\left(\tau_{k}\left(\bar{\mathbf{x}}\right)\right)}=\frac{\partial}{\partial \tilde{N}_{1-\nu}\left(\tau_{k}\left(\bar{\mathbf{x}}\right)\right)}\ddot{H}\left(\begin{matrix}\tilde{N}_{1-\nu}\left(\tau_{k}\left(\bar{\mathbf{x}}\right)\right),\mu_{B^{k}}\left(\bar{y}^{r}\right);\\ p^{l},\nu\end{matrix}\right) \tag{5.104}$$

$$\frac{\partial b_{k,r}^{l}\left(\bar{\mathbf{x}},\bar{y}^{r}\right)}{\partial \mu_{B^{k}}\left(\bar{y}^{r}\right)}=\frac{\partial}{\partial \mu_{B^{k}}\left(\bar{y}^{r}\right)}\ddot{H}\left(\begin{matrix}\tilde{N}_{1-\nu}\left(\tau_{k}\left(\bar{\mathbf{x}}\right)\right),\mu_{B^{k}}\left(\bar{y}^{r}\right);\\ p^{l},\nu\end{matrix}\right) \tag{5.105}$$

d) Block of aggregation

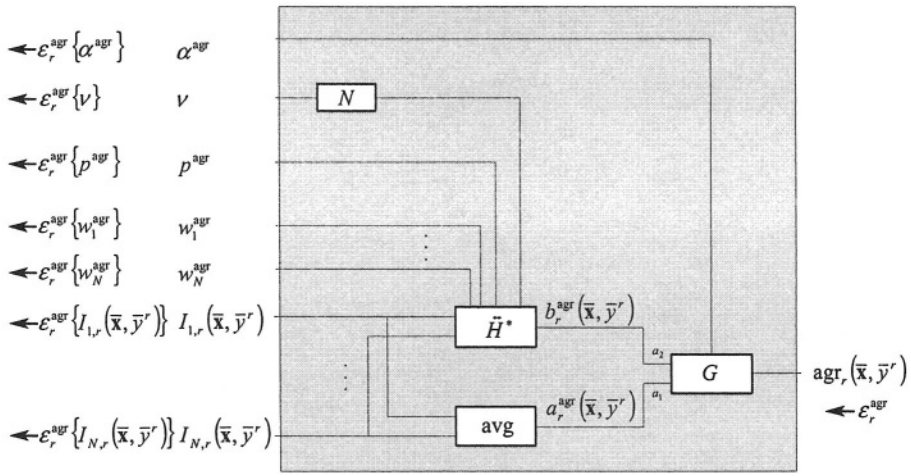

Fig. 5.32. Block of aggregation

The errors propagated through the block of aggregation (see Fig. 5.32) are given by:

$$\varepsilon_{r}^{\mathrm{agr}}\{\nu\}=\varepsilon_{r}^{\mathrm{agr}}\frac{\partial \mathrm{agr}_{r}\left(\bar{\mathbf{x}},\bar{y}^{r}\right)}{\partial b_{r}^{\mathrm{agr}}\left(\bar{\mathbf{x}},\bar{y}^{r}\right)}\frac{\partial b_{r}^{\mathrm{agr}}\left(\bar{\mathbf{x}},\bar{y}^{r}\right)}{\partial\left(1-\nu\right)}\frac{\partial N(\nu)}{\partial \nu} \tag{5.106}$$

$$\varepsilon_r^{\text{agr}}\left\{\alpha^{\text{agr}}\right\}=\varepsilon_r^{\text{agr}}\frac{\partial\,\text{agr}_r\left(\bar{\mathbf{x}},\bar{y}^r\right)}{\partial\alpha^{\text{agr}}} \tag{5.107}$$

$$\varepsilon_r^{\text{agr}}\left\{p^{\text{agr}}\right\}=\varepsilon_r^{\text{agr}}\frac{\partial\,\text{agr}_r\left(\bar{\mathbf{x}},\bar{y}^r\right)}{\partial b_r^{\text{agr}}\left(\bar{\mathbf{x}},\bar{y}^r\right)}\frac{\partial b_r^{\text{agr}}\left(\bar{\mathbf{x}},\bar{y}^r\right)}{\partial p^{\text{agr}}} \tag{5.108}$$

$$\varepsilon_r^{\text{agr}}\left\{w_r^{\text{agr}}\right\}=\varepsilon_r^{\text{agr}}\frac{\partial\,\text{agr}_r\left(\bar{\mathbf{x}},\bar{y}^r\right)}{\partial b_r^{\text{agr}}\left(\bar{\mathbf{x}},\bar{y}^r\right)}\frac{\partial b_r^{\text{agr}}\left(\bar{\mathbf{x}},\bar{y}^r\right)}{\partial w_r^{\text{agr}}} \tag{5.109}$$

$$\varepsilon_r^{\text{agr}}\left\{I_{k,r}\left(\bar{\mathbf{x}},\bar{y}^r\right)\right\}=\varepsilon_r^{\text{agr}}\left(\begin{array}{l}\dfrac{\partial\,\text{agr}_r\left(\bar{\mathbf{x}},\bar{y}^r\right)}{\partial b_r^{\text{agr}}\left(\bar{\mathbf{x}},\bar{y}^r\right)}\dfrac{\partial b_r^{\text{agr}}\left(\bar{\mathbf{x}},\bar{y}^r\right)}{\partial I_{k,r}\left(\bar{\mathbf{x}},\bar{y}^r\right)}+\\ +\dfrac{\partial\,\text{agr}_r\left(\bar{\mathbf{x}},\bar{y}^r\right)}{\partial a_r^{\text{agr}}\left(\bar{\mathbf{x}},\bar{y}^r\right)}\dfrac{\partial a_r^{\text{agr}}\left(\bar{\mathbf{x}},\bar{y}^r\right)}{\partial I_{k,r}\left(\bar{\mathbf{x}},\bar{y}^r\right)}\end{array}\right) \tag{5.110}$$

where

$$\frac{\partial\,\text{agr}_r\left(\bar{\mathbf{x}},\bar{y}^r\right)}{\partial a_r^{\text{agr}}\left(\bar{\mathbf{x}},\bar{y}^r\right)}=\frac{\partial}{\partial a_r^{\text{agr}}\left(\bar{\mathbf{x}},\bar{y}^r\right)}G\left(\begin{array}{c}a_r^{\text{agr}}\left(\bar{\mathbf{x}},\bar{y}^r\right),b_r^{\text{agr}}\left(\bar{\mathbf{x}},\bar{y}^r\right);\\ \alpha^{\text{agr}}\end{array}\right) \tag{5.111}$$

$$\frac{\partial\,\text{agr}_r\left(\bar{\mathbf{x}},\bar{y}^r\right)}{\partial b_r^{\text{agr}}\left(\bar{\mathbf{x}},\bar{y}^r\right)}=\frac{\partial}{\partial b_r^{\text{agr}}\left(\bar{\mathbf{x}},\bar{y}^r\right)}G\left(\begin{array}{c}a_r^{\text{agr}}\left(\bar{\mathbf{x}},\bar{y}^r\right),b_r^{\text{agr}}\left(\bar{\mathbf{x}},\bar{y}^r\right);\\ \alpha^{\text{agr}}\end{array}\right) \tag{5.112}$$

$$\frac{\partial\,\text{agr}_r\left(\bar{\mathbf{x}},\bar{y}^r\right)}{\partial\alpha^{\text{agr}}}=\frac{\partial}{\partial\alpha^{\text{agr}}}G\left(\begin{array}{c}a_r^{\text{agr}}\left(\bar{\mathbf{x}},\bar{y}^r\right),b_r^{\text{agr}}\left(\bar{\mathbf{x}},\bar{y}^r\right);\\ \alpha^{\text{agr}}\end{array}\right) \tag{5.113}$$

$$\frac{\partial a_r^{\text{agr}}\left(\bar{\mathbf{x}},\bar{y}^r\right)}{\partial I_{k,r}\left(\bar{\mathbf{x}},\bar{y}^r\right)}=\frac{\partial}{\partial I_{k,r}\left(\bar{\mathbf{x}},\bar{y}^r\right)}\text{avg}\left(I_{1,r}\left(\bar{\mathbf{x}},\bar{y}^r\right),\ldots,I_{N,r}\left(\bar{\mathbf{x}},\bar{y}^r\right)\right) \tag{5.114}$$

$$\frac{\partial b_r^{\text{agr}}\left(\bar{\mathbf{x}},\bar{y}^r\right)}{\partial(1-\nu)}=\frac{\partial}{\partial(1-\nu)}\ddot{H}^*\left(\begin{array}{c}I_{1,r}\left(\bar{\mathbf{x}},\bar{y}^r\right),\ldots,I_{N,r}\left(\bar{\mathbf{x}},\bar{y}^r\right);\\ w_1^{\text{agr}},\ldots,w_N^{\text{agr}},p^{\text{agr}},1-\nu\end{array}\right) \tag{5.115}$$

$$\frac{\partial b_r^{\text{agr}}\left(\bar{\mathbf{x}},\bar{y}^r\right)}{\partial p^{\text{agr}}}=\frac{\partial}{\partial p^{\text{agr}}}\ddot{H}^*\left(\begin{array}{c}I_{1,r}\left(\bar{\mathbf{x}},\bar{y}^r\right),\ldots,I_{N,r}\left(\bar{\mathbf{x}},\bar{y}^r\right);\\ w_1^{\text{agr}},\ldots,w_N^{\text{agr}},p^{\text{agr}},1-\nu\end{array}\right) \tag{5.116}$$

$$\frac{\partial b_r^{\text{agr}}\left(\bar{\mathbf{x}},\bar{y}^r\right)}{\partial w_r^{\text{agr}}}=\frac{\partial}{\partial w_r^{\text{agr}}}\ddot{H}^*\left(\begin{array}{c}I_{1,r}\left(\bar{\mathbf{x}},\bar{y}^r\right),\ldots,I_{N,r}\left(\bar{\mathbf{x}},\bar{y}^r\right);\\ w_1^{\text{agr}},\ldots,w_N^{\text{agr}},p^{\text{agr}},1-\nu\end{array}\right) \tag{5.117}$$

$$\frac{\partial b_r^{\text{agr}}\left(\bar{\mathbf{x}},\bar{y}^r\right)}{\partial I_{k,r}\left(\bar{\mathbf{x}},\bar{y}^r\right)}=\frac{\partial}{\partial I_{k,r}\left(\bar{\mathbf{x}},\bar{y}^r\right)}\ddot{H}^*\left(\begin{array}{c} I_{1,r}\left(\bar{\mathbf{x}},\bar{y}^r\right),\ldots,I_{N,r}\left(\bar{\mathbf{x}},\bar{y}^r\right);\\ w_1^{\text{agr}},\ldots,w_N^{\text{agr}},p^{\text{agr}},1-\nu\end{array}\right) \tag{5.118}$$

e) Block of defuzzification

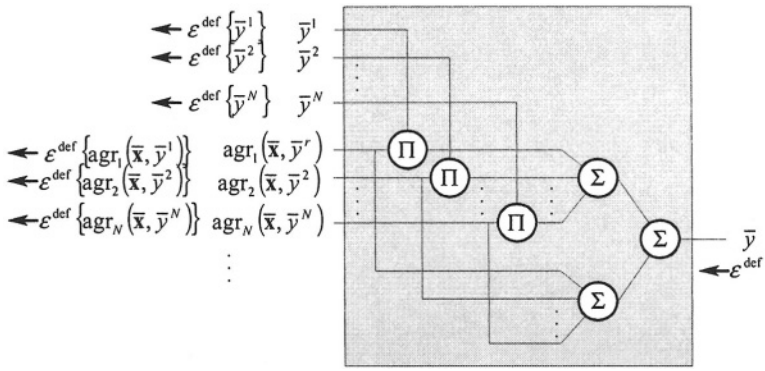

Fig. 5.33. Block of defuzzification

The errors propagated through the block of defuzzification (see Fig. 5.33) are given by:

$$\varepsilon^{\text{def}}\left\{\bar{y}^r\right\}=\varepsilon^{\text{def}}\frac{\partial}{\partial\bar{y}^r}\text{def}\left(\begin{array}{c}\text{agr}_1\left(\bar{\mathbf{x}},\bar{y}^1\right),\ldots,\text{agr}_N\left(\bar{\mathbf{x}},\bar{y}^N\right);\\ \bar{y}^1,\ldots,\bar{y}^N\end{array}\right) \tag{5.119}$$

$$\varepsilon^{\text{def}}\left\{\text{agr}_r\left(\bar{\mathbf{x}},\bar{y}^r\right)\right\}=\varepsilon^{\text{def}}\frac{\partial\,\text{def}\left(\begin{array}{c}\text{agr}_1\left(\bar{\mathbf{x}},\bar{y}^1\right),\ldots,\text{agr}_N\left(\bar{\mathbf{x}},\bar{y}^N\right);\\ \bar{y}^1,\ldots,\bar{y}^N\end{array}\right)}{\partial\,\text{agr}_r\left(\bar{\mathbf{x}},\bar{y}^r\right)} \tag{5.120}$$

5.9. SIMULATION RESULTS

In this section we present four simulations of the OR-type neuro-fuzzy systems. We use benchmarks described in Section 3.6. Each of the four simulations are designed in the same fashion:

(i) In the first experiment, based on the input-output data, we learn the parameters of the membership functions and a system type $\nu\in[0,1]$ of the basic flexible system described in Section 5.5. It will be seen that

the optimal values of ν, determined by a gradient procedure, are either zero or one.

(ii) In the second experiment, we learn the parameters of the membership functions of the basic flexible system described in Section 5.5 choosing value ν as opposite to that obtained in experiment (i). Obviously, we expect a worse performance of the neuro-fuzzy system comparing with experiment (i).

(iii) In the third experiment, we learn the parameters of the membership functions, system type $\nu \in [0,1]$ and soft parameters $\alpha^{\tau} \in [0,1]$, $\alpha^{l} \in [0,1]$, $\alpha^{\text{agr}} \in [0,1]$ of the soft flexible system described in Section 5.6 assuming that classical (not-parameterised) triangular norms are applied.

(iv) In the fourth experiment, we learn the same parameters as in the third experiment and, moreover, the weights $w_{i,k}^{\tau} \in [0,1]$, $= 1,\ldots,n$, $k = 1,\ldots,N$, in the antecedents of rules and weights $w_k^{\text{agr}} \in [0,1]$, $k = 1,\ldots,N$, of the aggregation operator of rules. In all diagrams (weights representation) we separate $w_{i,k}^{\tau} \in [0,1]$, $i = 1,\ldots,n$, $k = 1,\ldots,N$, from $w_k^{\text{agr}} \in [0,1]$, $k = 1,\ldots, N$, by a vertical dashed line.

In each of the above simulations we apply the Zadeh H-implication and the algebraic H-implication described in examples 5.5 and 5.6, respectively. In separate experiments we repeat simulations (i)-(iv) replacing the Zadeh H-implication and the algebraic H-implication by quasi-implications generated by parameterised triangular norms: the Dombi H-implication and the Yager H-implication. In these simulations we additionally incorporate parameters $p^{\tau} \in [0,\infty)$, $p^{l} \in [0,\infty)$, $p^{\text{agr}} \in [0,\infty)$.

Box and Jenkins Gas Furnace problem

The experimental results for the Box and Jenkins Gas Furnace problem are depicted in Tables 5.2 and 5.3 for the not-parameterised (Zadeh and algebraic) and parameterised (Dombi and Yager) H-functions, respectively. For experiment (iv) the final values (after learning) of weights $w_{i,k}^{\tau} \in [0,1]$ and $w_k^{\text{agr}} \in [0,1]$, $i = 1,\ldots,6$, $k = 1,\ldots,4$, are shown in Fig. 5.34 (Zadeh and algebraic H-functions) and Fig. 5.35 (Dombi and Yager H-functions).

Table 5.2 Experimental results

OR-TYPE NFIS WITH NON-PARAMETRISED H-FUNCTIONS (BOX AND JENKINS GAS FURNACE PROBLEM)						
Experiment number	Name of flexibility parameter	Initial values	Final values after learning		RMSE (learning sequence)	
			Zadeh H-function	Algebraic H-function	Zadeh H-function	Algebraic H-function
i	ν	0.5	0.0000	0.0000	0.5001	0.4973
ii	ν	1	-	-	0.6332	0.6217
iii	ν	0.5	0.0000	0.0000	0.4118	0.4019
	α^r	1	0.9734	0.9921		
	α^l	1	0.9958	0.9967		
	α^{agr}	1	0.9789	0.9899		
iv	ν	0.5	0.0000	0.0000	0.2412	0.2407
	α^r	1	0.9902	0.9691		
	α^l	1	0.9565	0.9754		
	α^{agr}	1	0.9999	0.9883		
	$\mathbf{w}^r$	**1**	Fig. 5.34-a	Fig. 5.34-b		
	$\mathbf{w}^{agr}$	**1**	Fig. 5.34-a	Fig. 5.34-b		

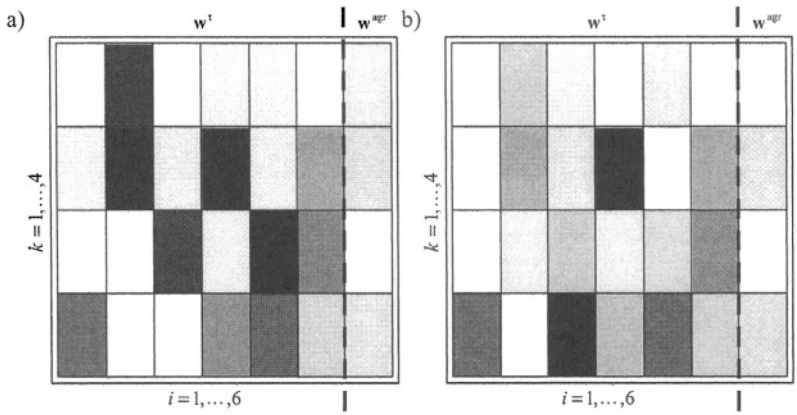

Fig. 5.34. Weights representation in the Modeling of Box and Jenkins Gas Furnace problem for OR-type NFIS and a) Zadeh H-function, b) algebraic H-function

Table 5.3 Experimental results

OR-TYPE NFIS WITH PARAMETRISED H-FUNCTIONS (BOX AND JENKINS GAS FURNACE PROBLEM)						
Experiment number	Name of flexibility parameter	Initial values	Final values after learning		RMSE (learning sequence)	
			Dombi H-function	Yager H-function	Dombi H-function	Yager H-function
i	ν	0.5	0.0000	0.0000	0.5102	0.5079
ii	ν	1	-	-	0.6500	0.6420
iii	ν	0.5	0.0000	0.0000	0.4325	0.4232
	p^r	10	10.2725	23.4276		
	p^l	10	9.9951	22.5107		
	p^{agr}	10	10.4641	2.8213		
	α^r	1	0.9999	0.9999		
	α^l	1	0.9999	0.9999		
	α^{agr}	1	0.9999	0.9999		
iv	ν	0.5	0.0000	0.0000	0.2453	0.2416
	p^r	10	12.1621	22.8896		
	p^l	10	8.7344	22.0345		
	p^{agr}	10	11.8385	2.2820		
	α^r	1	0.9553	0.9715		
	α^l	1	0.9768	0.9582		
	α^{agr}	1	0.9472	0.9663		
	$\mathbf{w}^r$	**1**	Fig. 5.35-a	Fig. 5.35-b		
	$\mathbf{w}^{agr}$	**1**	Fig. 5.35-a	Fig. 5.35-b		

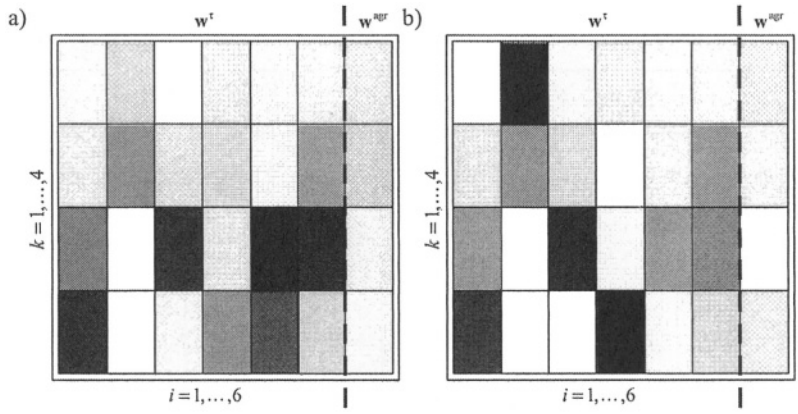

Fig. 5.35. Weights representation in the Modeling of Box and Jenkins Gas Furnace problem for OR-type NFIS and a) Dombi H-function, b) Yager H-function

Glass Identification problem

The experimental results for the Glass Identification problem are depicted in Tables 5.4 and 5.5 for the not-parameterised (Zadeh and algebraic) and parameterised (Dombi and Yager) H-functions, respectively. For experiment (iv) the final values (after learning) of weights $w_{i,k}^r \in [0,1]$ and $w_k^{agr} \in [0,1]$, $i = 1,...,9$, $k = 1,...,2$, are shown in Fig. 5.37 (Dombi and Yager H-functions).

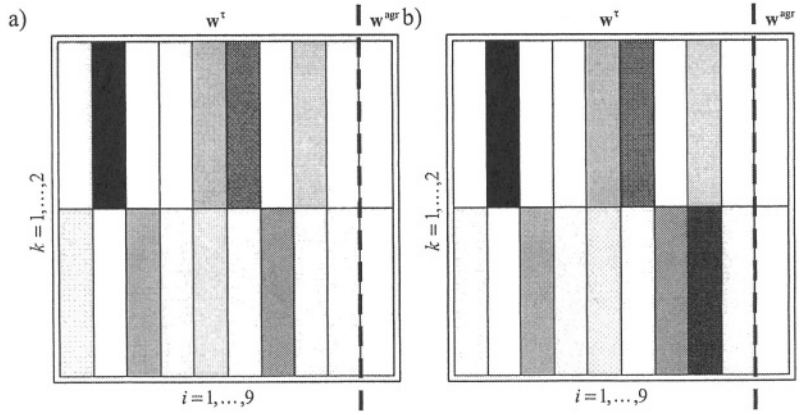

Fig. 5.36. Weights representation in the Glass Identification problem for OR-type NFIS and a) Zadeh H-function, b) algebraic H-function

Table 5.4 Experimental results

OR-TYPE NFIS WITH NON-PARAMETRISED H-FUNCTIONS (GLASS IDENTIFICATION PROBLEM)								
Experiment number	Name of flexibility parameter	Initial values	Final values after learning		Mistakes [%] (learning sequence)		Mistakes [%] (testing sequence)	
			Zadeh H-function	Algebraic H-function	Zadeh H-function	Algebraic H-function	Zadeh H-function	Algebraic H-function
i	ν	0.5	1.0000	1.0000	3.33	3.33	3.13	3.13
ii	ν	0	-	-	4.00	4.67	3.13	3.13
iii	ν	0.5	1.0000	1.0000	2.67	2.67	3.13	3.13
	$\alpha^{\wedge}$	1	0.0163	0.0019				
	$\alpha^{\vee}$	1	0.9939	0.9970				
	α^{agr}	1	0.9554	0.9912				
iv	ν	0.5	1.0000	1.0000	2.67	2.67	1.56	1.56
	$\alpha^{\wedge}$	1	0.0038	0.0137				
	$\alpha^{\vee}$	1	0.9858	0.9937				
	α^{agr}	1	0.8674	0.9693				
	$\mathbf{w}^{\wedge}$	**1**	Fig. 5.36-a	Fig. 5.36-b				
	$\mathbf{w}^{agr}$	**1**	Fig. 5.36-a	Fig. 5.36-b				

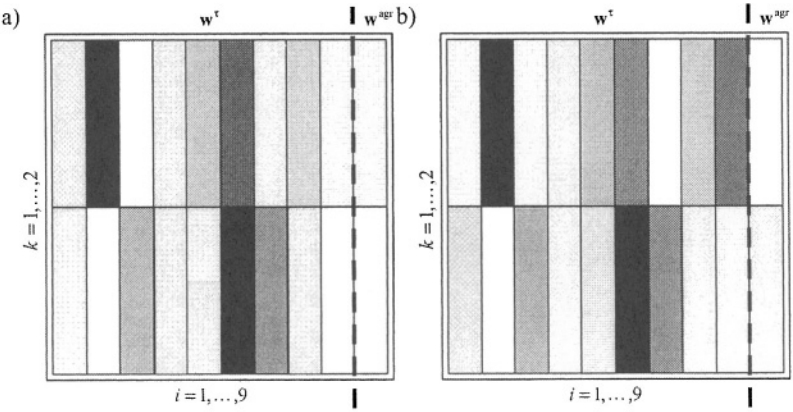

Fig. 5.37. Weights representation in the Glass Identification problem for OR-type NFIS and a) Dombi H-function, b) Yager H-function

Table 5.5 Experimental results

OR-TYPE NFIS WITH PARAMETRISED H-FUNCTIONS (GLASS IDENTIFICATION PROBLEM)								
Experiment number	Name of flexibility parameter	Initial values	Final values after learning		Mistakes [%] (learning sequence)		Mistakes [%] (testing sequence)	
			Dombi H-function	Yager H-function	Dombi H-function	Yager H-function	Dombi H-function	Yager H-function
i	ν	0.5	1.0000	1.0000	3.33	3.33	3.13	3.13
ii	ν	0	-	-	4.00	4.67	3.13	3.13
iii	ν	0.5	1.0000	1.0000	2.67	2.67	1.56	1.56
	p^{τ}	10	9.7496	12.5239				
	p^{l}	10	10.0006	9.9965				
	p^{agr}	10	9.9999	9.9920				
	α^{τ}	1	0.0302	0.1122				
	α^{l}	1	0.9173	0.9413				
	α^{agr}	1	0.9934	0.9973				
iv	ν	0.5	1.0000	1.0000	2.00	2.00	1.56	1.56
	p^{τ}	10	9.1328	12.1261				
	p^{l}	10	10.0601	9.8597				
	p^{agr}	10	10.3097	9.9544				
	α^{τ}	1	0.0948	0.1280				
	α^{l}	1	0.8896	0.9349				
	α^{agr}	1	0.9600	0.9695				
	$\mathbf{w}^{\tau}$	**1**	Fig. 5.37-a	Fig. 5.37-b				
	$\mathbf{w}^{agr}$	**1**	Fig. 5.37-a	Fig. 5.37-b				

Modeling of Static Nonlinear Function (HANG) problem

The experimental results for the Modeling of the Static Nonlinear Function problem are shown in Tables 5.6 and 5.7 for the not-parameterised (Zadeh and algebraic) and parameterised (Dombi and Yager) H-functions, respectively. For experiment (iv) the final values (after learning) of weights $w_{i,k}^{\tau} \in [0,1]$ and $w_{k}^{agr} \in [0,1]$, $i = 1,\ldots,2$, $k = 1,...,5$, are depicted in Fig. 5.38 (Zadeh and algebraic H-functions) and Fig. 5.39 (Dombi and Yager H-functions).

Table 5.6 Experimental results

OR-TYPE NFIS WITH NON-PARAMETRISED H-FUNCTIONS (MODELING OF STATIC NONLINEAR FUNCTION PROBLEM)						
Experiment number	**Name of flexibility parameter**	**Initial values**	**Final values after learning**		**RMSE (learning sequence)**	
			Zadeh H-function	**Algebraic H-function**	**Zadeh H-function**	**Algebraic H-function**
i	ν	0.5	0.0000	0.0000	0.1162	0.1059
ii	ν	1	-	-	0.1315	0.1205
iii	ν	0.5	0.0000	0.0000	0.0992	0.0662
	α^{τ}	1	0.0259	0.9935		
	$\alpha^{\prime}$	1	0.9999	0.9973		
	α^{agr}	1	0.9361	0.9978		
iv	ν	0.5	0.0000	0.0000	0.0816	0.0485
	α^{τ}	1	0.0021	0.9841		
	$\alpha^{\prime}$	1	0.9466	0.9367		
	α^{agr}	1	0.8806	0.9973		
	$\mathbf{w}^{\tau}$	**1**	Fig. 5.38-a	Fig. 5.38-b		
	$\mathbf{w}^{\mathrm{agr}}$	**1**	Fig. 5.38-a	Fig. 5.38-b		

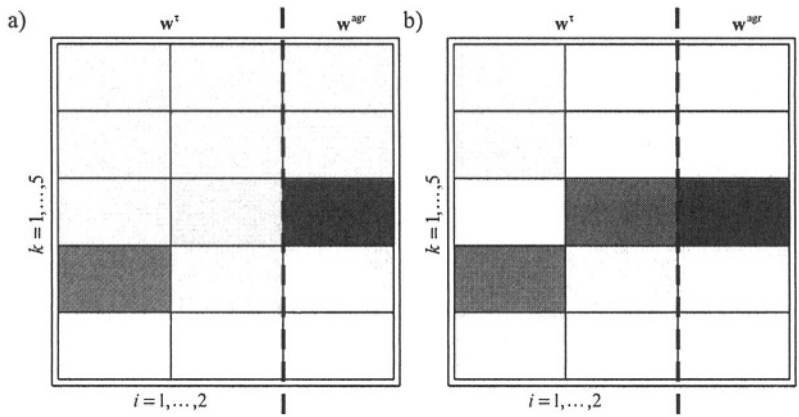

Fig. 5.38. Weights representation in the HANG problem for OR-type NFIS and a) Zadeh H-function, b) algebraic H-function

Table 5.7 Experimental results

OR-TYPE NFIS WITH PARAMETRISED H-FUNCTIONS (MODELING OF STATIC NONLINEAR FUNCTION PROBLEM)						
Experiment number	**Name of flexibility parameter**	**Initial values**	**Final values after learning**		**RMSE (learning sequence)**	
			Dombi H-function	**Yager H-function**	**Dombi H-function**	**Yager H-function**
i	ν	0.5	0.0000	0.0000	0.1307	0.1021
ii	ν	1	-	-	0.1578	0.1265
iii	ν	0.5	0.0000	0.0000	0.1245	0.0857
	p^r	10	17.6026	4.9607		
	p^l	10	10.0986	11.3656		
	p^{agr}	10	10.3006	3.1707		
	α^r	1	0.0153	0.9983		
	α^l	1	0.9999	0.9997		
	α^{agr}	1	0.9971	0.9962		
iv	ν	0.5	0.0000	0.0000	0.1093	0.0739
	p^r	10	17.4917	5.4504		
	p^l	10	10.0221	11.7908		
	p^{agr}	10	10.4474	3.8058		
	α^r	1	0.0576	0.9469		
	α^l	1	0.9836	0.9327		
	α^{agr}	1	0.9672	0.9448		
	$\mathbf{w}^r$	**1**	Fig. 5.39-a	Fig. 5.39-b		
	$\mathbf{w}^{agr}$	**1**	Fig. 5.39-a	Fig. 5.39-b		

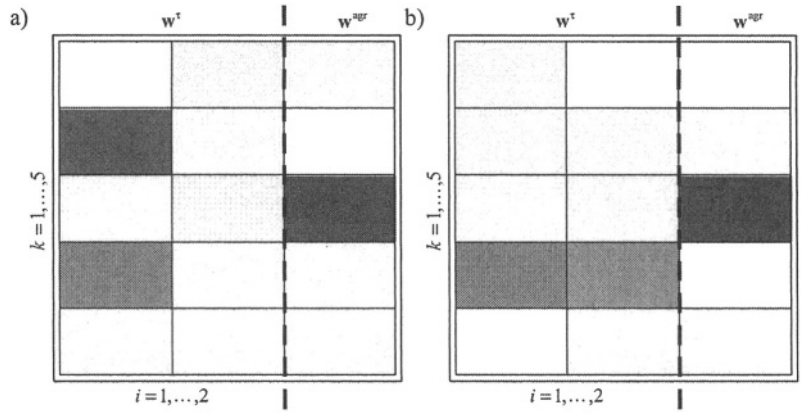

Fig. 5.39. Weights representation in the HANG problem for OR-type NFIS and a) Dombi H-function, b) Yager H-function

Wisconsin Breast Cancer problem

The experimental results for the Wisconsin Breast Cancer problem are shown in Tables 5.8 and 5.9 for the not-parameterised (Zadeh and algebraic) and parameterised (Dombi and Yager) H-functions, respectively. For experiment (iv) the final values (after learning) of weights $w_{i,k}^r \in [0,1]$ and $w_k^{agr} \in [0,1]$, $i=1,\ldots,9$, $k=1,\ldots,2$, are shown in Fig. 5.40 (Zadeh and algebraic H-functions) and Fig. 5.41 (Dombi and Yager H-functions).

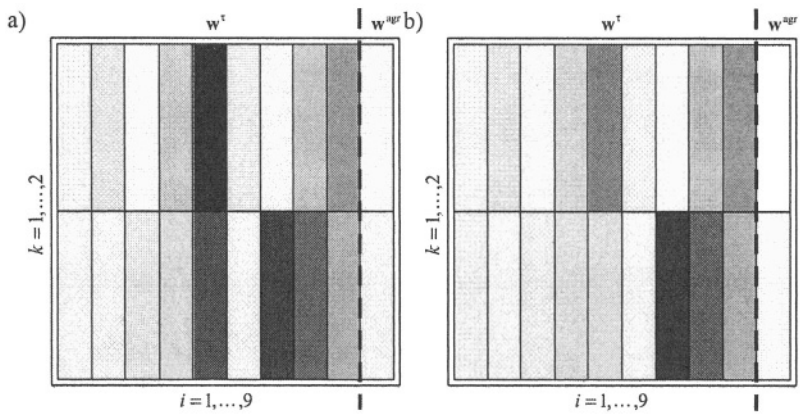

Fig. 5.40. Weights representation in the Wisconsin Breast Cancer problem for OR-type NFIS and a) Zadeh H-function, b) algebraic H-function

Table 5.8 Experimental results

OR-TYPE NFIS WITH NON-PARAMETRISED H-FUNCTIONS (WISCONSIN BREAST CANCER PROBLEM)								
Experiment number	Name of flexibility parameter	Initial values	Final values after learning		Mistakes [%] (learning sequence)		Mistakes [%] (testing sequence)	
			Zadeh H-function	Algebraic H-function	Zadeh H-function	Algebraic H-function	Zadeh H-function	Algebraic H-function
i	ν	0.5	1.0000	1.0000	2.93	2.51	1.95	1.95
ii	ν	0	-	-	3.14	2.51	2.44	2.44
iii	ν	0.5	1.0000	1.0000	2.30	2.09	1.95	1.46
	α^{τ}	1	0.0001	0.0001				
	$\alpha^{\perp}$	1	0.9532	0.9598				
	α^{agr}	1	0.9991	0.9999				
iv	ν	0.5	1.0000	1.0000	1.88	1.88	1.46	1.46
	α^{τ}	1	0.0396	0.0569				
	$\alpha^{\perp}$	1	0.9774	0.9942				
	α^{agr}	1	0.9888	0.9241				
	$\mathbf{w}^{\tau}$	**1**	Fig. 5.40-a	Fig. 5.40-b				
	$\mathbf{w}^{agr}$	**1**	Fig. 5.40-a	Fig. 5.40-b				

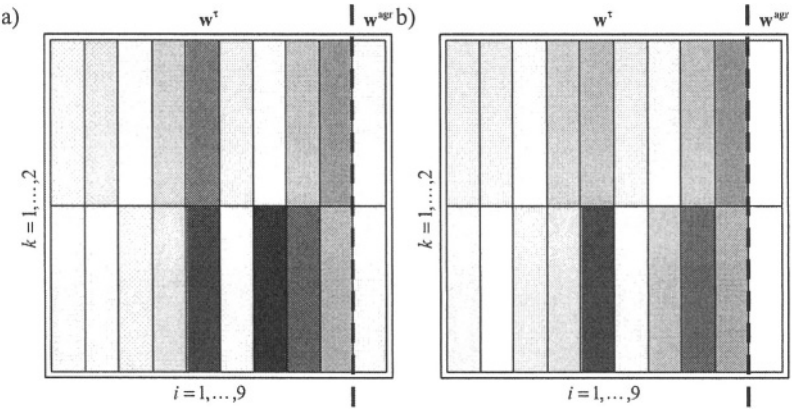

Fig. 5.41. Weights representation in the Wisconsin Breast Cancer problem for OR-type NFIS and a) Dombi H-function, b) Yager H-function

Table 5.9 Experimental results

OR-TYPE NFIS WITH PARAMETRISED H-FUNCTIONS (WISCONSIN BREAST CANCER PROBLEM)								
Experiment number	Name of flexibility parameter	Initial values	Final values after learning		Mistakes [%] (learning sequence)		Mistakes [%] (testing sequence)	
			Dombi H-function	Yager H-function	Dombi H-function	Yager H-function	Dombi H-function	Yager H-function
i	ν	0.5	1.0000	1.0000	2.30	2.09	1.95	1.95
ii	ν	0	-	-	2.51	2.30	2.44	1.95
iii	ν	0.5	1.0000	1.0000	2.09	2.09	1.46	1.46
	p^r	10	9.9996	10.3369				
	p^l	10	10.0002	3.8732				
	p^{agr}	10	9.3716	11.4615				
	α^r	1	0.0001	0.0025				
	α^l	1	0.9554	0.9608				
	α^{agr}	1	0.9993	0.9851				
iv	ν	0.5	1.0000	1.0000	1.67	1.67	1.46	1.46
	p^r	10	6.8385	16.1151				
	p^l	10	10.6122	6.6922				
	p^{agr}	10	14.3388	8.7812				
	α^r	1	0.0451	0.0842				
	α^l	1	0.9677	0.9533				
	α^{agr}	1	0.9431	0.9660				
	$\mathbf{w}^r$	**1**	Fig. 5.41-a	Fig. 5.41-b				
	$\mathbf{w}^{agr}$	**1**	Fig. 5.41-a	Fig. 5.41-b				

5.10. SUMMARY AND DISCUSSION

In this chapter we introduced the concept of adjustable quasi-triangular norms. They switch smoothly between t-norms and t-conorms depending on the value of parameter ν. In a similar spirit we proposed the idea of quasi-implications which vary between "engineering implications" and fuzzy s-implications as ν goes from 0 to 1. Based on these ideas we developed several neuro-fuzzy systems characterized by automatic determination in the process of learning of the fuzzy inference (Mamdani type or logical type).

Moreover, we incorporated various flexibility parameters into the construction of neuro-fuzzy systems. From simulations performed in Section 5.9 we found out that the best results were obtained for weighted flexible neuro-fuzzy systems.

5.11. PROBLEMS

Problem 5.1. Derive an H-function generated by the Dombi t-norm.

Problem 5.2. Derive an H-function generated by the Yager t-conorm.

Problem 5.3. Derive a quasi-implication generated by the Dombi t-norm.

Problem 5.4. Derive a quasi-implication generated by the Yager t-conorm.

Problem 5.5. Derive the basic flexible systems based on the H-function generated by the Zadeh t-norm. Plot the corresponding hyperplanes.

Problem 5.6. Derive the soft flexible system based on the H-function generated by the Zadeh t-norm. Plot the corresponding hyperplanes.

Problem 5.7. Derive the weighted flexible system based on the H-function generated by the Zadeh t-norm. Plot the corresponding hyperplanes.

Problem 5.8. Apply the weighted minimum and maximum operations [19]

$$\min\{a_1,\ldots,a_n;w_1,\ldots,w_n\}=\bigwedge_{i=1}^{n}[(1-w_i)\vee a_i]$$

$$\max\{a_1,\ldots,a_n;w_1,\ldots,w_n\}=\bigvee_{i=1}^{n}[w_i\wedge a_i]$$

to construct a weighted flexible NFIS.

Problem 5.9. Apply the geometric mean

$$G(a_1,\ldots,a_n)=\left(\prod_{i=1}^{n}a_i\right)^{\frac{1}{n}}$$

to construct a soft flexible system.

This page intentionally left blank

Chapter 6

FLEXIBLE COMPROMISE AND-TYPE NEURO-FUZZY SYSTEMS

6.1. INTRODUCTION

In Chapter 5 we constructed a new class of adjustable quasi-triangular norms and corresponding quasi-implications. After learning, a Mamdani-type or a logical-type neuro-fuzzy system has been established depending on parameter ν in function $H(\mathbf{a};\nu)$. In this chapter we assume that fuzzy inference is characterized by the simultaneous appearance of Mamdani-type and logical-type reasoning. We will develop another class of flexible neuro-fuzzy systems.

This chapter is organized as follows. The problem description is given in Section 6.2. In Sections 6.3-6.5 a new class of neuro-fuzzy systems is presented. In Section 6.6 learning procedures are derived to learn parameter λ (type of the system), parameters of membership functions, flexibility parameters and weights described in Chapter 4. In Section 6.7 simulation examples are given.

6.2. PROBLEM DESCRIPTION

Let $\bar{\mathbf{x}}(t) \in \mathbf{X} \subset \mathbf{R}^n$, $y(t) \in \mathbf{R} \subset \mathbf{Y}$ and $d(t) \in \mathbf{R} \subset \mathbf{Y}$, $t = 1,2,\ldots$, be a sequence of inputs, outputs and desirable outputs, respectively. In this chapter we address the following design and learning problems:

a) Our first problem is to design a neuro-fuzzy system realizing a mapping $f : \mathbf{X} \to \mathbf{Y}$ such that a fuzzy inference is described by the following combination of "engineering" and fuzzy implications

$$I(a,b;\lambda) = (1-\lambda) I_{eng}(a,b) + \lambda I_{fuzzy}(a,b) \tag{6.1}$$

We assume that an "engineering implication" $I_{eng}(a,b)$ is represented by a t-norm whereas a fuzzy implication $I_{fuzzy}(a,b)$ is represented by an S-implication, i.e.

$$I(a,b;\lambda) = (1-\lambda) T\{a,b\} + \lambda S\{1-a,b\} \tag{6.2}$$

where T and S are dual triangular norms.

b) Moreover, we will incorporate: (i) softness to implication operators, to the aggregation of rules and to the connectives of antecedents; (ii) certainty weights to the aggregation of rules and to the connectives of antecedents; and (iii) parameterized families of t-norms and t-conorms to implication operators, to the aggregation of rules and to the connectives of antecedents. All the neuro-fuzzy systems described in this chapter are in the form (3.39), i.e.

$$\bar{y} = f(\bar{\mathbf{x}}) = \frac{\sum_{r=1}^{N} \bar{y}^r \cdot \mathrm{agr}_r(\bar{\mathbf{x}}, \bar{y}^r)}{\sum_{r=1}^{N} \mathrm{agr}_r(\bar{\mathbf{x}}, \bar{y}^r)}$$

c) Based on the learning sequence $(\bar{\mathbf{x}}(1), d(1)), (\bar{\mathbf{x}}(2), d(2)), \ldots$ we wish to determine all parameters (including the system type represented by value of λ) and weights of NFIS such that

$$e(t) = \frac{1}{2}[f(\bar{\mathbf{x}}(t)) - d(t)]^2 \tag{6.3}$$

is minimized. The steepest descent optimization algorithm with constraints can be applied to solve this problem.

6.3. BASIC COMPROMISE SYSTEMS

Flexible compromise neuro-fuzzy systems exhibit simultaneous behaviour of two basic fuzzy inference systems: Mamdani-type and logical-type. The quasi-implications (6.2) are compositions of two

components, and are a practical implementations of the compromise fuzzy reasoning described in Chapter 4. The value of the compromise parameter $\lambda \in [0,1]$ imposes a dominance of one fuzzy model over another; Mamdani-type if the parameter goes to 0 or logical-type if it goes to 1. The logical AND connective is used to connect the antecedents in the individual rules. This connective is expressed by a t-norm. As it was noted in [108], the combination of the Mamdani and logical methods raises the problem of aggregation in the compromise model. In the case of the Mamdani method we use an OR aggregation while in the case of the logical method an AND aggregation is applied. A combination of these two methods requires a final aggregation being a compromise between an OR and an AND aggregation. The basic neuro-fuzzy inference systems of an AND-type employ combinations of "engineering" and fuzzy implication. The systems are given by the formula:

AND Ia

$$\tau_k(\bar{\mathbf{x}}) = T\left\{\mu_{A_1^k}(\bar{x}_1),\ldots,\mu_{A_n^k}(\bar{x}_n)\right\} \tag{6.4}$$

$$I_{k,r}(\bar{\mathbf{x}},\bar{y}^r) = \begin{pmatrix} (1-\lambda)I_{eng}\left(\tau_k(\bar{\mathbf{x}}),\mu_{B^k}(\bar{y}^r)\right)+ \\ +\lambda I_{fuzzy}\left(\tau_k(\bar{\mathbf{x}}),\mu_{B^k}(\bar{y}^r)\right) \end{pmatrix} \tag{6.5}$$

$$\mathrm{agr}_r(\bar{\mathbf{x}},\bar{y}^r) = \begin{pmatrix} (1-\lambda)S\left\{I_{1,r}(\bar{\mathbf{x}},\bar{y}^r),\ldots,I_{N,r}(\bar{\mathbf{x}},\bar{y}^r)\right\}+ \\ +\lambda T\left\{I_{1,r}(\bar{\mathbf{x}},\bar{y}^r),\ldots,I_{N,r}(\bar{\mathbf{x}},\bar{y}^r)\right\} \end{pmatrix} \tag{6.6}$$

Alternatively, if we express a t-norm and a t-conorm by an H-function with parameters $\nu = 0$ or $\nu = 1$, the basic compromise NFIS can be described as follows:

AND Ib

$$\tau_k(\bar{\mathbf{x}}) = H\begin{pmatrix} \mu_{A_1^k}(\bar{x}_1),\ldots,\mu_{A_n^k}(\bar{x}_n)\ ; \\ 0 \end{pmatrix} \tag{6.7}$$

$$I_{k,r}(\bar{\mathbf{x}},\bar{y}^r) = \begin{pmatrix} (1-\lambda)H\begin{pmatrix} \tilde{N}_1(\tau_k(\bar{\mathbf{x}})),\mu_{B^k}(\bar{y}^r)\ ; \\ 0 \end{pmatrix}+ \\ +\lambda H\begin{pmatrix} \tilde{N}_0(\tau_k(\bar{\mathbf{x}})),\mu_{B^k}(\bar{y}^r)\ ; \\ 1 \end{pmatrix} \end{pmatrix} \tag{6.8}$$

$$\mathrm{agr}_r\left(\bar{\mathbf{x}},\bar{y}^r\right)=\left(\begin{array}{l}(1-\lambda)H\left(\begin{array}{c}I_{1,r}\left(\bar{\mathbf{x}},\bar{y}^r\right),\ldots,I_{N,r}\left(\bar{\mathbf{x}},\bar{y}^r\right);\\ 1\end{array}\right)+\\ +\lambda H\left(\begin{array}{c}I_{1,r}\left(\bar{\mathbf{x}},\bar{y}^r\right),\ldots,I_{N,r}\left(\bar{\mathbf{x}},\bar{y}^r\right);\\ 0\end{array}\right)\end{array}\right) \tag{6.9}$$

It is easily seen that for $\lambda=0$ the above system is of Mamdani-type, and for $\lambda=1$ it becomes of the logical-type.

The block diagram realizing a strength of rules $\tau_k\left(\bar{\mathbf{x}}\right)$, $k=1,\ldots,N$, is depicted in Fig. 5.10 (see Section 5.5). In Fig. 6.1 and 6.2 we depict block diagrams realizing the quasi-implication $I_{k,r}\left(\bar{\mathbf{x}},\bar{y}^r\right)$ given by (6.8) and the aggregation $\mathrm{agr}_r\left(\bar{\mathbf{x}},\bar{y}^r\right)$ given by (6.9), respectively.

The quasi-implication $I_{k,r}\left(\bar{\mathbf{x}},\bar{y}^r\right)$, $k=1,\ldots,N$, $r=1,\ldots,N$, following Fig. 6.1, is determined by a soft composition controlled by parameter λ of two components, i.e. fuzzy inference (logical-type) $b_{k,r}^I\left(\bar{\mathbf{x}},\bar{y}^r\right)$ and "engineering" inference (Mamdani-type) $c_{k,r}^I\left(\bar{\mathbf{x}},\bar{y}^r\right)$. The dominance of one component over another depends on the value of parameter λ.

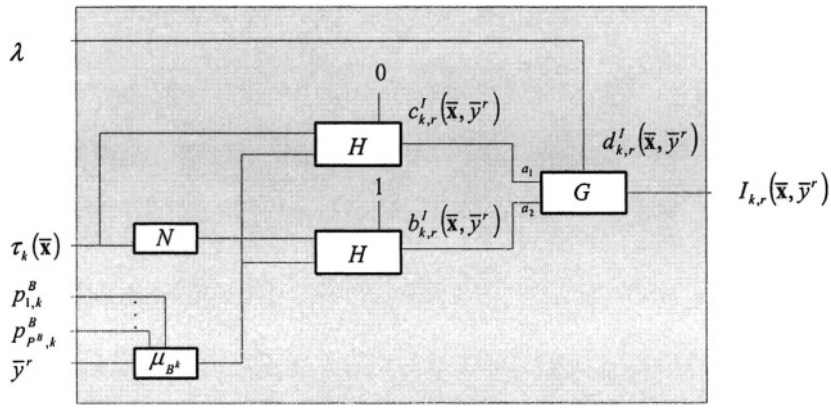

Fig. 6.1. Implication operator in basic AND-type NFIS (AND Ib)

The aggregation operator $\mathrm{agr}_r\left(\bar{\mathbf{x}},\bar{y}^r\right)$, $r=1,\ldots,N$, following Fig. 6.2, is determined by a soft composition of two components, i.e. $b_r^{\mathrm{agr}}\left(\bar{\mathbf{x}},\bar{y}^r\right)$ and $c_r^{\mathrm{agr}}\left(\bar{\mathbf{x}},\bar{y}^r\right)$. Component $b_r^{\mathrm{agr}}\left(\bar{\mathbf{x}},\bar{y}^r\right)$ describes the aggregation in the case of the logical method while component $c_r^{\mathrm{agr}}\left(\bar{\mathbf{x}},\bar{y}^r\right)$ in the case of the Mamdani

method. Observe that component $b_r^{agr}(\bar{\mathbf{x}}, \bar{y}^r)$ is realized by an H-function parameterized by $\nu = 0$ (equivalent to a t-norm) whereas component $c_r^{agr}(\bar{\mathbf{x}}, \bar{y}^r)$ is realized by an H-function parameterized by $\nu = 1$ (equivalent to a t-conorm). The dominance of one component over another depends on the value of parameter λ.

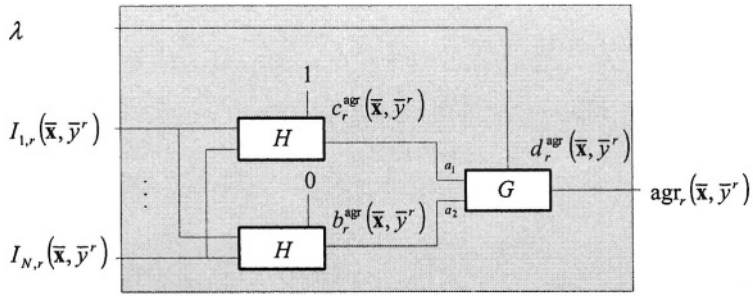

Fig. 6.2. Aggregation of rules in basic AND-type NFIS (AND Ib)

In Fig. 6.3 we depict the shape of the hyperplanes described by formulas (6.7)-(6.9) for $n = 2$, $N = 2$ and the varying parameter λ. We applied the algebraic triangular norms.

6.4. SOFT COMPROMISE SYSTEMS

In this section we extend the basic compromise NFIS described in Section 6.3. The flexibility in that section was based on a compromise fuzzy reasoning being a composition of the Mamdani-type and the logical-type fuzzy inference. The dominance of one inference model over another depends on the compromise operator λ. In this section we assume that fuzzy norms (and the H-function) in the antecedents connection, implication and aggregation are parameterised by parameters p^{τ}, p^{I}, p^{agr} respectively. An additional flexibility is obtained by incorporating soft fuzzy norms described in Section 4.3. The flexibility parameters α^{τ}, α^{I}, α^{agr} allow to balance between operators (6.4), (6.5), (6.6) and corresponding arithmetic averages being a kind of competitive solution to those given in Section 6.3. Another flexibility can be incorporated by an application of parameterized triangular norms described in Section 4.4.

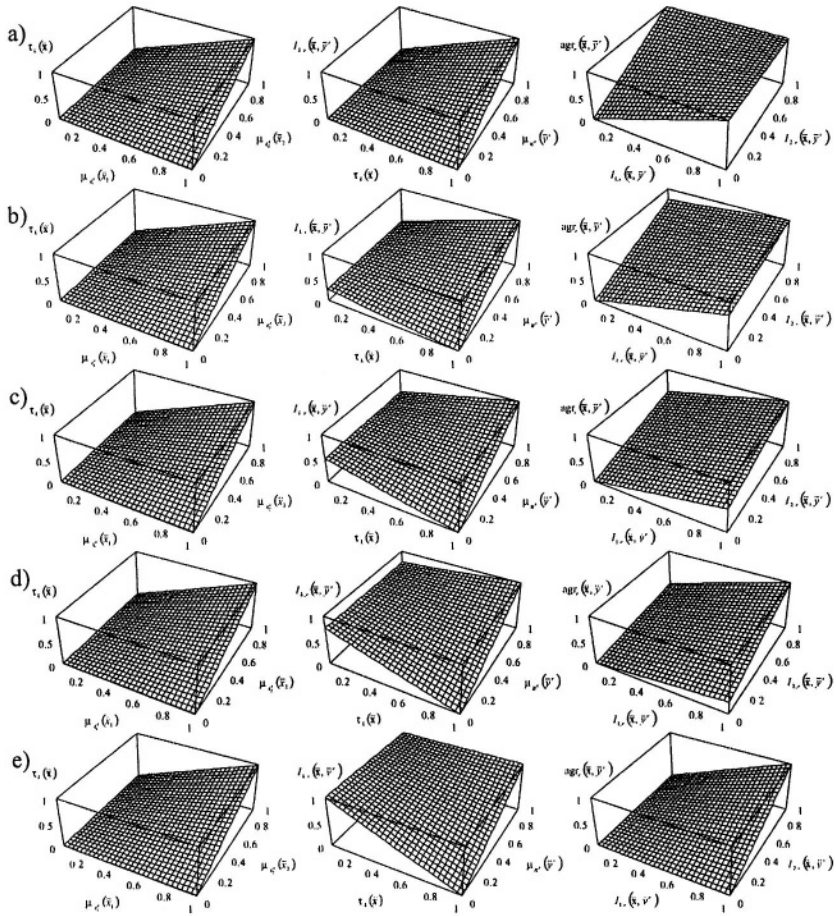

Fig. 6.3. 3D plots of basic AND-type NFIS (AND Ib)
for $n = 2$, $N = 2$ and a) $\lambda = 0.00$, b) $\lambda = 0.25$, c) $\lambda = 0.50$, d) $\lambda = 0.75$, e) $\lambda = 1.00$

The soft compromise NFIS of an AND-type are given by:

AND IIa

$$\tau_k(\bar{\mathbf{x}}) = \begin{pmatrix} (1-\alpha^{\tau})\text{avg}\left(\mu_{A_1^k}(\bar{x}_1),\ldots,\mu_{A_n^k}(\bar{x}_n)\right)+ \\ +\alpha^{\tau}\vec{T}\left\{\begin{matrix}\mu_{A_1^k}(\bar{x}_1),\ldots,\mu_{A_n^k}(\bar{x}_n)\ ; \\ p^{\tau}\end{matrix}\right\} \end{pmatrix} \tag{6.10}$$

$$I_{k,r}(\bar{\mathbf{x}},\bar{y}^r) = \begin{pmatrix} (1-\alpha^{I})\text{avg}\left(\tilde{N}_{1-\lambda}(\tau_k(\bar{\mathbf{x}})),\mu_{B^k}(\bar{y}^r)\right)+ \\ +\alpha^{I}\left(\begin{matrix}(1-\lambda)\vec{I}_{eng}\left(\begin{matrix}\tau_k(\bar{\mathbf{x}}),\mu_{B^k}(\bar{y}^r)\ ; \\ p^{I}\end{matrix}\right)+ \\ +\lambda\vec{I}_{fuzzy}\left(\begin{matrix}\tau_k(\bar{\mathbf{x}}),\mu_{B^k}(\bar{y}^r)\ ; \\ p^{I}\end{matrix}\right)\end{matrix}\right) \end{pmatrix} \tag{6.11}$$

$$\text{agr}_r(\bar{\mathbf{x}},\bar{y}^r) = \begin{pmatrix} (1-\alpha^{\text{agr}})\text{avg}\left(I_{1,r}(\bar{\mathbf{x}},\bar{y}^r),\ldots,I_{N,r}(\bar{\mathbf{x}},\bar{y}^r)\right)+ \\ +\alpha^{\text{agr}}\left(\begin{matrix}(1-\lambda)\vec{S}\left\{\begin{matrix}I_{1,r}(\bar{\mathbf{x}},\bar{y}^r),\ldots,I_{N,r}(\bar{\mathbf{x}},\bar{y}^r)\ ; \\ p^{\text{agr}}\end{matrix}\right\}+ \\ +\lambda\vec{T}\left\{\begin{matrix}I_{1,r}(\bar{\mathbf{x}},\bar{y}^r),\ldots,I_{N,r}(\bar{\mathbf{x}},\bar{y}^r)\ ; \\ p^{\text{agr}}\end{matrix}\right\}\end{matrix}\right) \end{pmatrix} \tag{6.12}$$

Alternatively, if we express a t-norm and t-conorm by an H-function with parameters $\nu = 0$ or $\nu = 1$, the soft compromise systems can be described as follows:

AND IIb

$$\tau_k(\bar{\mathbf{x}}) = \begin{pmatrix} (1-\alpha^{\tau})\text{avg}\left(\mu_{A_1^k}(\bar{x}_1),\ldots,\mu_{A_n^k}(\bar{x}_n)\right)+ \\ +\alpha^{\tau}\vec{H}\left(\begin{matrix}\mu_{A_1^k}(\bar{x}_1),\ldots,\mu_{A_n^k}(\bar{x}_n)\ ; \\ p^{\tau},0\end{matrix}\right) \end{pmatrix} \tag{6.13}$$

$$I_{k,r}(\bar{\mathbf{x}},\bar{y}^r)=\left(\begin{array}{c}(1-\alpha^I)\text{avg}\left(\tilde{N}_{1-\lambda}(\tau_k(\bar{\mathbf{x}})),\mu_{B^k}(\bar{y}^r)\right)+\\ +\alpha^I\left(\begin{array}{c}(1-\lambda)\ddot{H}\left(\begin{array}{c}\tilde{N}_1(\tau_k(\bar{\mathbf{x}})),\mu_{B^k}(\bar{y}^r)\,;\\ p^I,0\end{array}\right)+\\ +\lambda\ddot{H}\left(\begin{array}{c}\tilde{N}_0(\tau_k(\bar{\mathbf{x}})),\mu_{B^k}(\bar{y}^r)\,;\\ p^I,1\end{array}\right)\end{array}\right)\end{array}\right) \tag{6.14}$$

$$\text{agr}_r(\bar{\mathbf{x}},\bar{y}^r)=\left(\begin{array}{c}(1-\alpha^{\text{agr}})\text{avg}\left(I_{1,r}(\bar{\mathbf{x}},\bar{y}^r),\ldots,I_{N,r}(\bar{\mathbf{x}},\bar{y}^r)\right)+\\ +\alpha^{\text{agr}}\left(\begin{array}{c}(1-\lambda)\ddot{H}\left(\begin{array}{c}I_{1,r}(\bar{\mathbf{x}},\bar{y}^r),\ldots,I_{N,r}(\bar{\mathbf{x}},\bar{y}^r)\,;\\ p^{\text{agr}},1\end{array}\right)+\\ +\lambda\ddot{H}\left(\begin{array}{c}I_{1,r}(\bar{\mathbf{x}},\bar{y}^r),\ldots,I_{N,r}(\bar{\mathbf{x}},\bar{y}^r)\,;\\ p^{\text{agr}},0\end{array}\right)\end{array}\right)\end{array}\right) \tag{6.15}$$

It is easily seen that the above system reduces to a basic compromise system described in Section 6.3 if $\alpha^\tau=\alpha^I=\alpha^{\text{agr}}=1$ and parameterized triangular norms are replaced by standard triangular norms.

In Fig. 6.4 and 6.5 we depict block diagrams realizing the quasi-implication $I_{k,r}(\bar{\mathbf{x}},\bar{y}^r)$ given by (6.14) and the aggregation $\text{agr}_r(\bar{\mathbf{x}},\bar{y}^r)$ given by (6.15), respectively.

The block diagram realizing a strength of rules $\tau_k(\bar{\mathbf{x}})$, $k=1,\ldots,N$, is depicted in Fig. 5.14 (see Section 5.6). The quasi-implication $I_{k,r}(\bar{\mathbf{x}},\bar{y}^r)$, $k=1,\ldots,N$, $r=1,\ldots,N$, following Fig. 6.4, is determined by a soft composition, controlled by parameter α^I, of two components, i.e. $a^I_{k,r}(\bar{\mathbf{x}},\bar{y}^r)$ and $d^I_{k,r}(\bar{\mathbf{x}},\bar{y}^r)$. Component $d^I_{k,r}(\bar{\mathbf{x}},\bar{y}^r)$ is a soft composition, controlled by parameter λ of two components $b^I_{k,r}(\bar{\mathbf{x}},\bar{y}^r)$ and $c^I_{k,r}(\bar{\mathbf{x}},\bar{y}^r)$. The composition $d^I_{k,r}(\bar{\mathbf{x}},\bar{y}^r)$ is determined analogously as in basic compromise systems described in Section 6.3. Component $a^I_{k,r}(\bar{\mathbf{x}},\bar{y}^r)$ is the arithmetic average of the compromise operator (5.5), operating on the firing rule strength $\tau_k(\bar{\mathbf{x}})$, $k=1,\ldots,N$, and the value of consequence membership functions $\mu_{B^k}(\bar{y}^r)$, $k=1,...,N$, $r=1,\ldots,N$, determined in point $\bar{y}^r$, $r=1,\ldots,N$, at which functions μ_{B^k} attain their maximum equal to 1.

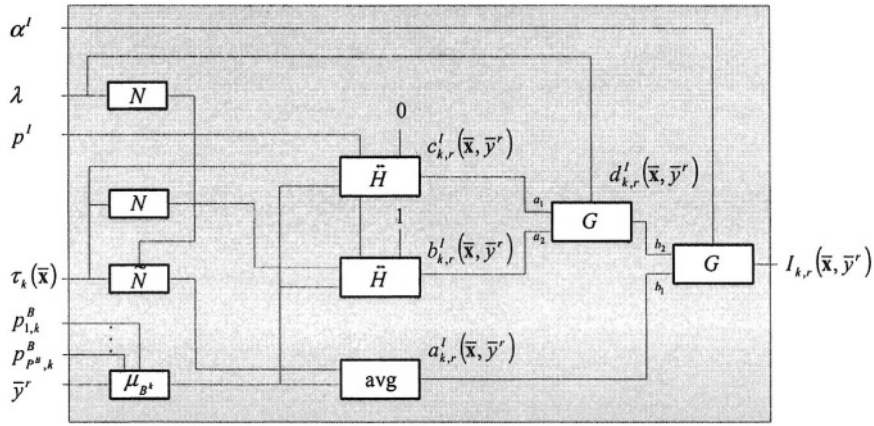

Fig. 6.4. Implication operator in soft AND-type NFIS (AND IIb)

The aggregation operator $\mathbf{agr}_r(\bar{\mathbf{x}},\bar{y}^r)$, $r=1,\ldots,N$, following in Fig. 6.5, is determined by a soft composition, controlled by parameter α^{agr}, of two components, i.e. $a_r^{\mathrm{agr}}(\bar{\mathbf{x}},\bar{y}^r)$ and $d_r^{\mathrm{agr}}(\bar{\mathbf{x}},\bar{y}^r)$. Component $a_r^{\mathrm{agr}}(\bar{\mathbf{x}},\bar{y}^r)$ is the average of quasi-implications $I_{k,r}(\bar{\mathbf{x}},\bar{y}^r)$, $k=1,\ldots,N$, $r=1,\ldots,N$, determined in points $\bar{y}^r$, $r=1,\ldots,N$, at which consequent membership functions μ_{B^k}, $k=1,\ldots,N$, attain their maximum equal to 1. The composition $d_r^{\mathrm{agr}}(\bar{\mathbf{x}},\bar{y}^r)$ of components $b_r^{\mathrm{agr}}(\bar{\mathbf{x}},\bar{y}^r)$ and $c_r^{\mathrm{agr}}(\bar{\mathbf{x}},\bar{y}^r)$ is determined analogously as in basic compromise systems described in Section 6.3. The dominance of one component over another depends on the value of parameter λ.

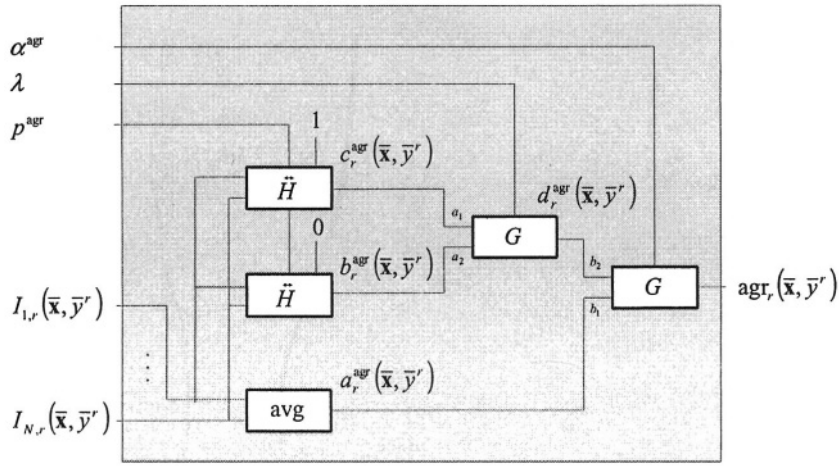

Fig. 6.5. Aggregation of rules in soft AND-type NFIS (AND IIb)

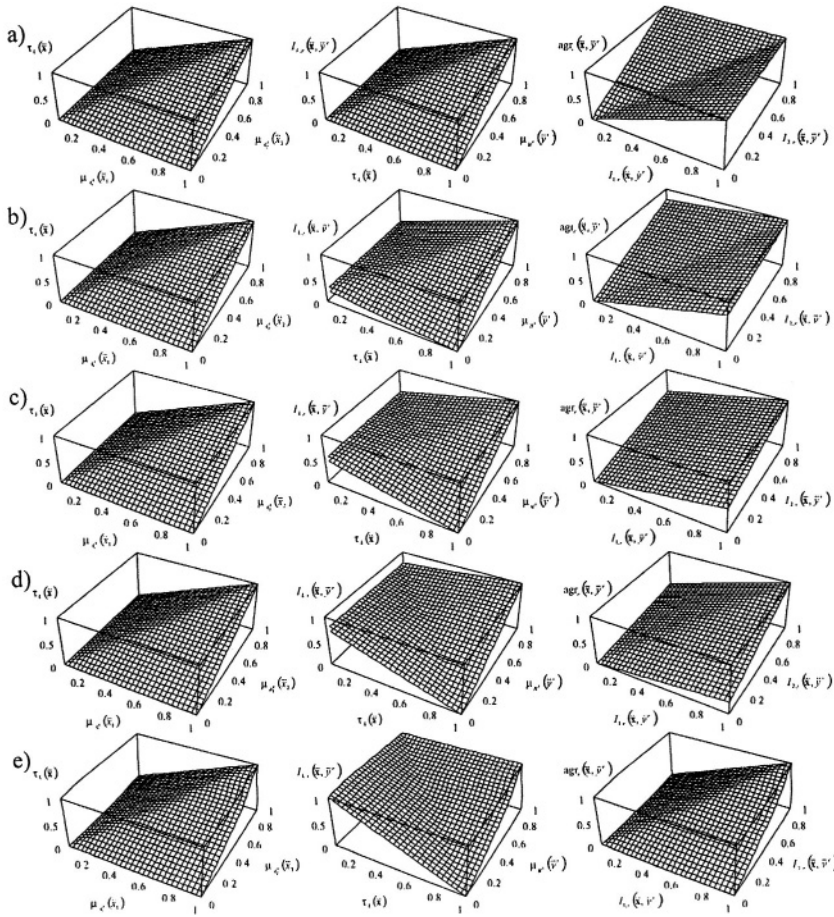

Fig. 6.6. 3D plots of soft AND-type NFIS (AND IIb)
for $n = 2$, $N = 2$, $\alpha^{\tau} = \alpha^{I} = \alpha^{\text{agr}} = 1.0$, $p^{\tau} = p^{I} = p^{\text{agr}} = 10$
and a) $\lambda = 0.00$, b) $\lambda = 0.25$, c) $\lambda = 0.50$, d) $\lambda = 0.75$, e) $\lambda = 1.00$

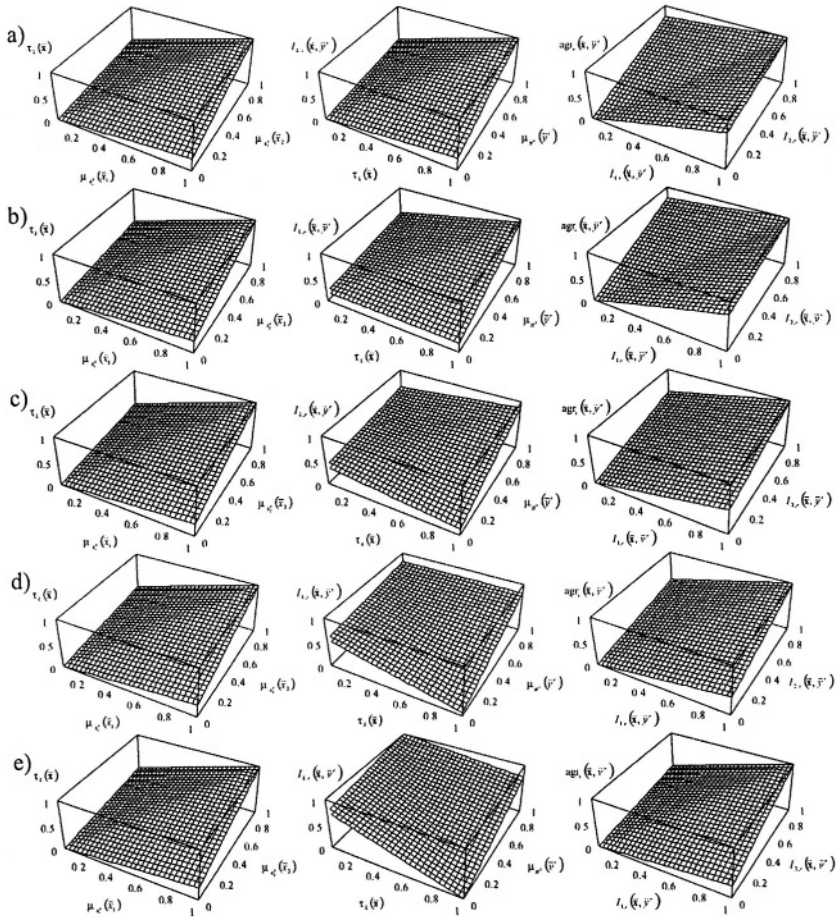

Fig. 6.7. 3D plots of soft AND-type NFIS (AND IIb)
for $n = 2$, $N = 2$, $\alpha^{\tau} = \alpha^{I} = \alpha^{agr} = 0.5$, $p^{\tau} = p^{I} = p^{agr} = 10$
and a) $\lambda = 0.00$, b) $\lambda = 0.25$, c) $\lambda = 0.50$, d) $\lambda = 0.75$, e) $\lambda = 1.00$

In Fig. 6,6 and Fig. 6.7 we depict the shape of the hyperplanes described by formulas (6.13)-(6.15) for $n = 2$ and $N = 2$. We applied the Dombi triangular norms.

6.5. WEIGHTED COMPROMISE SYSTEMS

Analogously to the weighted flexible systems of an OR-type presented in Section 5.7, we incorporate weights to the soft compromise system given in Section 6.4. Introducing weights to that system we get the weighted soft NFIS of an AND-type:

AND IIIa

$$\tau_k(\bar{\mathbf{x}}) = \begin{pmatrix} (1-\alpha^{\tau})\text{avg}\left(\mu_{A_1^k}(\bar{x}_1),\ldots,\mu_{A_n^k}(\bar{x}_n)\right)+ \\ +\alpha^{\tau}\ddot{T}^{*}\left\{\begin{matrix}\mu_{A_1^k}(\bar{x}_1),\ldots,\mu_{A_n^k}(\bar{x}_n)\ ; \\ w_{1,k}^{\tau},\ldots,w_{n,k}^{\tau},p^{\tau}\end{matrix}\right\} \end{pmatrix} \tag{6.16}$$

$$I_{k,r}(\bar{\mathbf{x}},\bar{y}^r) = \begin{pmatrix} (1-\alpha^{I})\text{avg}\left(\tilde{N}_{1-\lambda}(\tau_k(\bar{\mathbf{x}})),\mu_{B^k}(\bar{y}^r)\right)+ \\ +\alpha^{I}\begin{pmatrix} (1-\lambda)\ddot{I}_{eng}\begin{pmatrix}\tau_k(\bar{\mathbf{x}}),\mu_{B^k}(\bar{y}^r)\ ; \\ p^{I}\end{pmatrix}+ \\ +\lambda\ddot{I}_{fuzzy}\begin{pmatrix}\tau_k(\bar{\mathbf{x}}),\mu_{B^k}(\bar{y}^r)\ ; \\ p^{I}\end{pmatrix}\end{pmatrix} \end{pmatrix} \tag{6.17}$$

$$\text{agr}_r(\bar{\mathbf{x}},\bar{y}^r) = \begin{pmatrix} (1-\alpha^{\text{agr}})\text{avg}\left(I_{1,r}(\bar{\mathbf{x}},\bar{y}^r),\ldots,I_{N,r}(\bar{\mathbf{x}},\bar{y}^r)\right)+ \\ +\alpha^{\text{agr}}\begin{pmatrix} (1-\lambda)\ddot{S}^{*}\left\{\begin{matrix}I_{1,r}(\bar{\mathbf{x}},\bar{y}^r),\ldots,I_{N,r}(\bar{\mathbf{x}},\bar{y}^r)\ ; \\ w_1^{\text{agr}},\ldots,w_N^{\text{agr}},p^{\text{agr}}\end{matrix}\right\}+ \\ +\lambda\ddot{T}^{*}\left\{\begin{matrix}I_{1,r}(\bar{\mathbf{x}},\bar{y}^r),\ldots,I_{N,r}(\bar{\mathbf{x}},\bar{y}^r)\ ; \\ w_1^{\text{agr}},\ldots,w_N^{\text{agr}},p^{\text{agr}}\end{matrix}\right\}\end{pmatrix} \end{pmatrix} \tag{6.18}$$

Alternatively, if we express a t-norm and a t-conorm by an H-function with parameters $\nu = 0$ or $\nu = 1$, the weighted compromise systems NFIS can be described as follows:

AND IIIb

$$\tau_k(\bar{\mathbf{x}}) = \begin{pmatrix} (1-\alpha^\tau)\text{avg}\left(\mu_{A_1^k}(\bar{x}_1),\ldots,\mu_{A_n^k}(\bar{x}_n)\right)+ \\ +\alpha^\tau \ddot{H}^*\begin{pmatrix} \mu_{A_1^k}(\bar{x}_1),\ldots,\mu_{A_n^k}(\bar{x}_n)\ ; \\ w_{1,k}^\tau,\ldots,w_{n,k}^\tau, p^\tau, 0 \end{pmatrix} \end{pmatrix} \tag{6.19}$$

$$I_{k,r}(\bar{\mathbf{x}},\bar{y}^r) = \begin{pmatrix} (1-\alpha^I)\text{avg}\left(\tilde{N}_{1-\lambda}(\tau_k(\bar{\mathbf{x}})),\mu_{B^k}(\bar{y}^r)\right)+ \\ +\alpha^I \begin{pmatrix} (1-\lambda)\ddot{H}\begin{pmatrix} \tilde{N}_1(\tau_k(\bar{\mathbf{x}})),\mu_{B^k}(\bar{y}^r)\ ; \\ p^I, 0 \end{pmatrix}+ \\ +\lambda\ddot{H}\begin{pmatrix} \tilde{N}_0(\tau_k(\bar{\mathbf{x}})),\mu_{B^k}(\bar{y}^r)\ ; \\ p^I, 1 \end{pmatrix} \end{pmatrix} \end{pmatrix} \tag{6.20}$$

$$\text{agr}_r(\bar{\mathbf{x}},\bar{y}^r) = \begin{pmatrix} (1-\alpha^{\text{agr}})\text{avg}\left(I_{1,r}(\bar{\mathbf{x}},\bar{y}^r),\ldots,I_{N,r}(\bar{\mathbf{x}},\bar{y}^r)\right)+ \\ +\alpha^{\text{agr}} \begin{pmatrix} (1-\lambda)\ddot{H}^*\begin{pmatrix} I_{1,r}(\bar{\mathbf{x}},\bar{y}^r),\ldots,I_{N,r}(\bar{\mathbf{x}},\bar{y}^r)\ ; \\ w_1^{\text{agr}},\ldots,w_N^{\text{agr}}, p^{\text{agr}}, 1 \end{pmatrix}+ \\ +\lambda\ddot{H}^*\begin{pmatrix} I_{1,r}(\bar{\mathbf{x}},\bar{y}^r),\ldots,I_{N,r}(\bar{\mathbf{x}},\bar{y}^r)\ ; \\ w_1^{\text{agr}},\ldots,w_N^{\text{agr}}, p^{\text{agr}}, 0 \end{pmatrix} \end{pmatrix} \end{pmatrix} \tag{6.21}$$

It is easily seen that the above system reduces to a soft compromise system described in Section 6.4 if $w_{i,k}^\tau = 1$, $i = 1,\ldots,n$, $k = 1,\ldots,N$ and $w_k^{\text{agr}} = 1$, $k = 1,\ldots,N$.

The block diagrams realizing a strength of rules $\tau_k(\bar{\mathbf{x}})$, $k = 1,\ldots,N$, and quasi-implication $I_{k,r}(\bar{\mathbf{x}},\bar{y}^r)$, $k = 1,\ldots,N$, $r = 1,\ldots,N$, are depicted in Fig. 5.23 and Fig. 6.4, respectively. In Fig. 6.8 we depict the block diagram realizing the aggregation $\text{agr}_r(\bar{\mathbf{x}},\bar{y}^r)$ given by (6.21).

The aggregation operator $\text{agr}_r(\bar{\mathbf{x}},\bar{y}^r)$, $r = 1,\ldots,N$, following Fig. 6.8, is determined by a soft composition of two components, i.e. $a_r^{\text{agr}}(\bar{\mathbf{x}},\bar{y}^r)$ and $d_r^{\text{agr}}(\bar{\mathbf{x}},\bar{y}^r)$. Component $a_r^{\text{agr}}(\bar{\mathbf{x}},\bar{y}^r)$ is the average of quasi-implications $I_{k,r}(\bar{\mathbf{x}},\bar{y}^r)$, $k = 1,\ldots,N$, $r = 1,\ldots,N$, determined in points $\bar{y}^r$, $r = 1,\ldots,N$,

at which consequent membership functions μ_{B^k}, $k = 1,\ldots,N$, attain their maximum equal to 1. The composition $d_r^{\text{agr}}(\bar{\mathbf{x}}, \bar{y}^r)$ of components $b_r^{\text{agr}}(\bar{\mathbf{x}}, \bar{y}^r)$ and $c_r^{\text{agr}}(\bar{\mathbf{x}}, \bar{y}^r)$ is determined analogously as in basic compromise systems described in Section 6.4, however certainty weights w_k^{agr}, $k = 1,...,N$, are additionally incorporated into the weighted compromise system. The dominance of one component over another depends on the value of parameter α^{agr}.

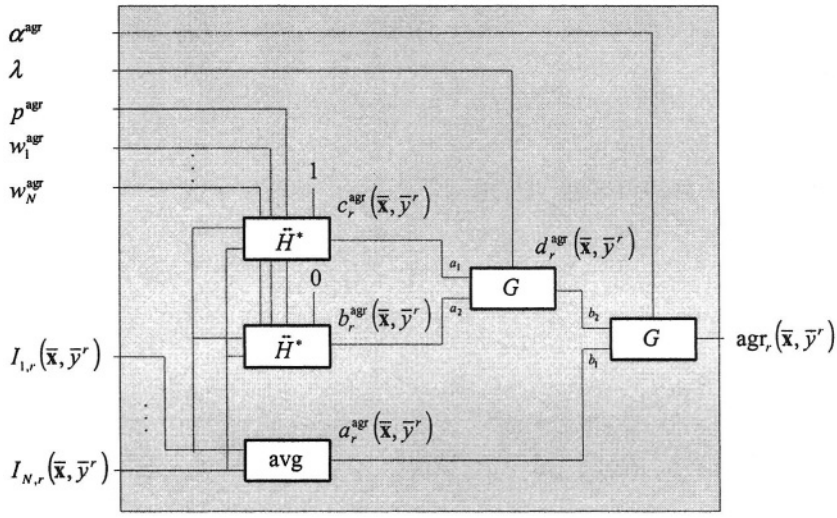

Fig. 6.8. Aggregation of rules operator in weighted soft AND-type NFIS (AND IIIb)

In Fig. 6.9 and Fig. 6.10 we depict the shape of the hyperplanes described by formulas (6.19)-(6.21) for $n = 2$ and $N = 2$. We applied the Dombi triangular norms.

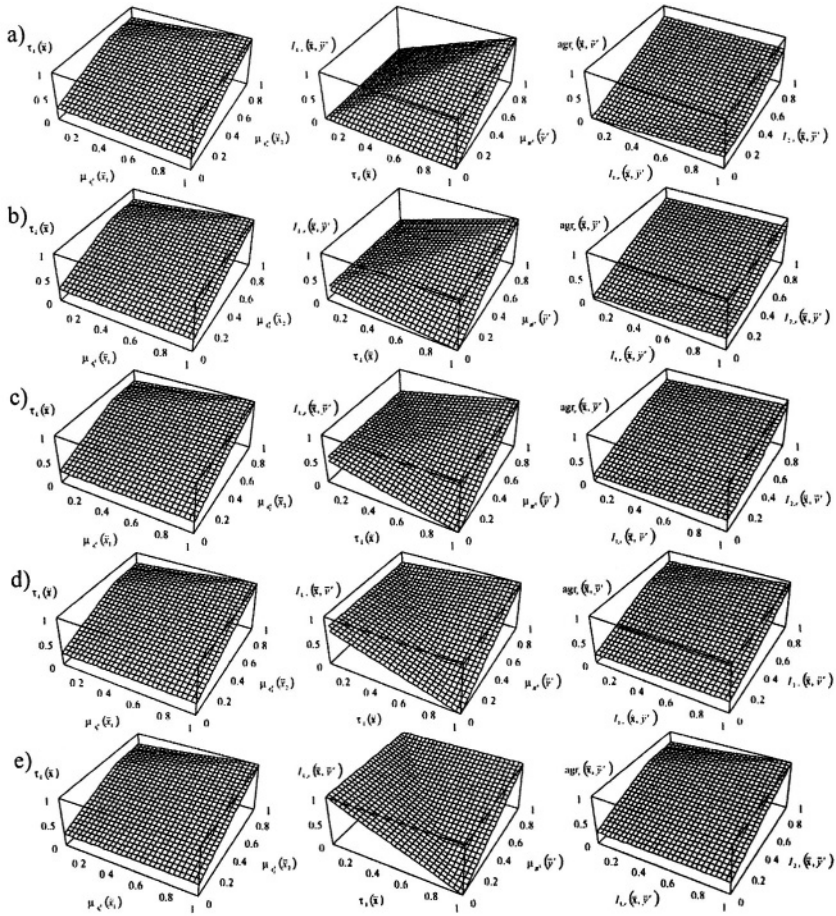

Fig. 6.9. 3D plots of weighted soft AND-type NFIS (AND IIIb) for $n = 2$, $N = 2$, $\alpha^{\tau} = \alpha^{I} = \alpha^{agr} = 1.0$, $p^{\tau} = p^{I} = p^{agr} = 10$, $w_{1,k}^{\tau} = w_1^{agr} = 0.25$, $w_{2,k}^{\tau} = w_2^{agr} = 0.75$, and a) $\lambda = 0.00$, b) $\lambda = 0.25$, c) $\lambda = 0.50$, d) $\lambda = 0.75$, e) $\lambda = 1.00$

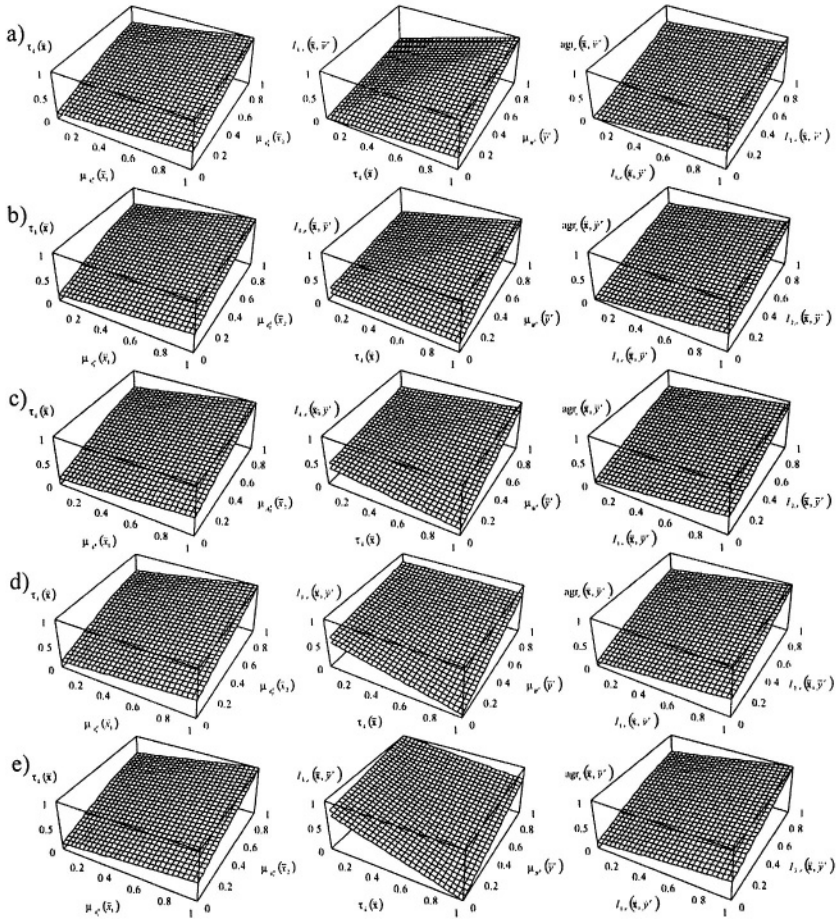

Fig. 6.10. 3D plots of weighted soft AND-type NFIS (AND IIIb) for $n = 2$, $N = 2$, $\alpha^{\tau} = \alpha^{I} = \alpha^{\text{agr}} = 0.5$, $p^{\tau} = p^{I} = p^{\text{agr}} = 10$, $w_{1,k}^{\tau} = w_1^{\text{agr}} = 0.25$, $w_{2,k}^{\tau} = w_2^{\text{agr}} = 0.75$, and a) $\lambda = 0.00$, b) $\lambda = 0.25$, c) $\lambda = 0.50$, d) $\lambda = 0.75$, e) $\lambda = 1.00$

6.6. LEARNING ALGORITHMS

In this section we follow the ideas given in Section 5.8. We will derive the learning algorithms for flexible compromise AND-type neuro-fuzzy systems having the layered structure shown in Fig. 5.29. The block of rules' activation and the block of defuzzification are the same as those shown in Fig. 5.30 and 5.33, respectively. Therefore, we will describe the block of implication and the block of aggregation which are different for AND-type and OR-type systems. We will apply the gradient optimization with constraints in order to optimize:

- $\lambda \in [0,1]$,
- $\alpha^{\tau} \in [0,1]$, $\alpha^{I} \in [0,1]$, $\alpha^{\text{agr}} \in [0,1]$,
- $p^{\tau} \in [0,\infty)$, $p^{I} \in [0,\infty)$, $p^{\text{agr}} \in [0,\infty)$,
- $w_{i,k}^{\tau} \in [0,1]$, $i = 1,...,n$, $k = 1,...,N$,
- $w_{k}^{\text{agr}} \in [0,1]$, $k = 1,...,N$.

Analogously to neuro-fuzzy systems studied in Chapter 5, we will use the constraint functions (33) and (35) in the Appendix, to perform the optimization of the above parameters and weights. Moreover, we will find in the process of learning parameters of the membership functions $\mu_{A_i^k}(x_i)$ and $\mu_{B^k}(y)$, $i = 1,\ldots,n$, $k = 1,\ldots,N$:

- $p_{u,i,k}^{A}$, $u = 1,\ldots,P^{A}$, $i = 1,...,n$, $k = 1,...,N$,
- $p_{u,k}^{B}$, $u = 2,\ldots,P^{B}$, $k = 1,...,N$,
- $p_{1,k}^{B} = \bar{y}^{k}$, $k = 1,...,N$.

a) General learning procedures

Parameter λ of NFIS is updated by the recursive procedure

$$\lambda(t+1) = \lambda(t) - \eta \Delta\lambda(t) \tag{6.22}$$

where correction Δ is given by

$$\Delta\lambda = \sum_{k=1}^{N} \sum_{r=1}^{N} \varepsilon_{k,r}^{I}\{\lambda\} + \sum_{r=1}^{N} \varepsilon_{r}^{\text{agr}}\{\lambda\} \tag{6.23}$$

Other parameters of the NFIS are updated by the recursive procedures given in Section 5.8.

b) Block of implications

The errors propagated through the block of implications (see Fig. 6.11) are determined as follows:

$$\varepsilon_{k,r}^{I}\{\lambda\}=\varepsilon_{k,r}^{I}\left(\begin{array}{l}\dfrac{\partial I_{k,r}\left(\bar{\mathbf{x}},\bar{y}^{r}\right)}{\partial d_{k,r}^{I}\left(\bar{\mathbf{x}},\bar{y}^{r}\right)}\dfrac{\partial d_{k,r}^{I}\left(\bar{\mathbf{x}},\bar{y}^{r}\right)}{\partial\lambda}+\\ +\dfrac{\partial I_{k,r}\left(\bar{\mathbf{x}},\bar{y}^{r}\right)}{\partial a_{k,r}^{I}\left(\bar{\mathbf{x}},\bar{y}^{r}\right)}\dfrac{\partial a_{k,r}^{I}\left(\bar{\mathbf{x}},\bar{y}^{r}\right)}{\partial\tilde{N}_{1-\lambda}\left(\tau_{k}\left(\bar{\mathbf{x}}\right)\right)}\cdot\\ \cdot\dfrac{\partial\tilde{N}_{1-\lambda}\left(\tau_{k}\left(\bar{\mathbf{x}}\right)\right)}{\partial\left(1-\lambda\right)}\dfrac{\partial N(\lambda)}{\partial\lambda}\end{array}\right)\tag{6.24}$$

$$\varepsilon_{k,r}^{I}\{\alpha^{I}\}=\varepsilon_{k,r}^{I}\frac{\partial I_{k,r}\left(\bar{\mathbf{x}},\bar{y}^{r}\right)}{\partial\alpha^{I}}\tag{6.25}$$

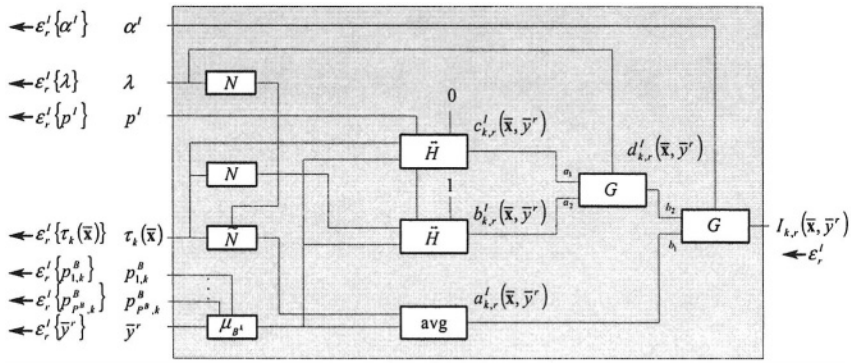

Fig. 6.11. Block of implications

$$\varepsilon_{k,r}^{I}\{p^{I}\}=\varepsilon_{k,r}^{I}\frac{\partial I_{k,r}\left(\bar{\mathbf{x}},\bar{y}^{r}\right)}{\partial d_{k,r}^{I}\left(\bar{\mathbf{x}},\bar{y}^{r}\right)}\left(\begin{array}{l}\dfrac{\partial d_{k,r}^{I}\left(\bar{\mathbf{x}},\bar{y}^{r}\right)}{\partial c_{k,r}^{I}\left(\bar{\mathbf{x}},\bar{y}^{r}\right)}\dfrac{\partial c_{k,r}^{I}\left(\bar{\mathbf{x}},\bar{y}^{r}\right)}{\partial p^{I}}+\\ +\dfrac{\partial d_{k,r}^{I}\left(\bar{\mathbf{x}},\bar{y}^{r}\right)}{\partial b_{k,r}^{I}\left(\bar{\mathbf{x}},\bar{y}^{r}\right)}\dfrac{\partial b_{k,r}^{I}\left(\bar{\mathbf{x}},\bar{y}^{r}\right)}{\partial p^{I}}\end{array}\right)\tag{6.26}$$

$$\varepsilon_{k,r}^{I}\left\{p_{u,k}^{B}\right\}=\varepsilon_{k,r}^{I}\left(\frac{\partial I_{k,r}\left(\bar{\mathbf{x}},\bar{y}^{r}\right)}{\partial d_{k,r}^{I}\left(\bar{\mathbf{x}},\bar{y}^{r}\right)}\left(\begin{array}{l}\dfrac{\partial d_{k,r}^{I}\left(\bar{\mathbf{x}},\bar{y}^{r}\right)}{\partial c_{k,r}^{I}\left(\bar{\mathbf{x}},\bar{y}^{r}\right)}\cdot\\ \cdot\dfrac{\partial c_{k,r}^{I}\left(\bar{\mathbf{x}},\bar{y}^{r}\right)}{\partial\mu_{B^{k}}\left(\bar{y}^{r}\right)}+\\ +\dfrac{\partial d_{k,r}^{I}\left(\bar{\mathbf{x}},\bar{y}^{r}\right)}{\partial b_{k,r}^{I}\left(\bar{\mathbf{x}},\bar{y}^{r}\right)}\cdot\\ \cdot\dfrac{\partial b_{k,r}^{I}\left(\bar{\mathbf{x}},\bar{y}^{r}\right)}{\partial\mu_{B^{k}}\left(\bar{y}^{r}\right)}\end{array}\right)+\right.$$
$$\left.+\frac{\partial I_{k,r}\left(\bar{\mathbf{x}},\bar{y}^{r}\right)}{\partial a_{k,r}^{I}\left(\bar{\mathbf{x}},\bar{y}^{r}\right)}\frac{\partial a_{k,r}^{I}\left(\bar{\mathbf{x}},\bar{y}^{r}\right)}{\partial\mu_{B^{k}}\left(\bar{y}^{r}\right)}\right)\frac{\partial\mu_{B^{k}}\left(\bar{y}^{r}\right)}{\partial p_{u,k}^{B}}\qquad(6.27)$$

$$\varepsilon_{k,r}^{I}\left\{\bar{y}^{r}\right\}=\varepsilon_{k,r}^{I}\left(\frac{\partial I_{k,r}\left(\bar{\mathbf{x}},\bar{y}^{r}\right)}{\partial d_{k,r}^{I}\left(\bar{\mathbf{x}},\bar{y}^{r}\right)}\left(\begin{array}{l}\dfrac{\partial d_{k,r}^{I}\left(\bar{\mathbf{x}},\bar{y}^{r}\right)}{\partial c_{k,r}^{I}\left(\bar{\mathbf{x}},\bar{y}^{r}\right)}\cdot\\ \cdot\dfrac{\partial c_{k,r}^{I}\left(\bar{\mathbf{x}},\bar{y}^{r}\right)}{\partial\mu_{B^{k}}\left(\bar{y}^{r}\right)}+\\ +\dfrac{\partial d_{k,r}^{I}\left(\bar{\mathbf{x}},\bar{y}^{r}\right)}{\partial b_{k,r}^{I}\left(\bar{\mathbf{x}},\bar{y}^{r}\right)}\cdot\\ \cdot\dfrac{\partial b_{k,r}^{I}\left(\bar{\mathbf{x}},\bar{y}^{r}\right)}{\partial\mu_{B^{k}}\left(\bar{y}^{r}\right)}\end{array}\right)+\right.$$
$$\left.+\frac{\partial I_{k,r}\left(\bar{\mathbf{x}},\bar{y}^{r}\right)}{\partial a_{k,r}^{I}\left(\bar{\mathbf{x}},\bar{y}^{r}\right)}\frac{\partial a_{k,r}^{I}\left(\bar{\mathbf{x}},\bar{y}^{r}\right)}{\partial\mu_{B^{k}}\left(\bar{y}^{r}\right)}\right)\frac{\partial\mu_{B^{k}}\left(\bar{y}^{r}\right)}{\partial\bar{y}^{r}}\qquad(6.28)$$

$$\varepsilon_{k,r}^{I}\left\{\tau_{k}\left(\bar{\mathbf{x}}\right)\right\}=\varepsilon_{k,r}^{I}\left(\frac{\partial I_{k,r}\left(\bar{\mathbf{x}},\bar{y}^{r}\right)}{\partial d_{k,r}^{I}\left(\bar{\mathbf{x}},\bar{y}^{r}\right)}\left(\begin{array}{l}\dfrac{\partial d_{k,r}^{I}\left(\bar{\mathbf{x}},\bar{y}^{r}\right)}{\partial c_{k,r}^{I}\left(\bar{\mathbf{x}},\bar{y}^{r}\right)}\dfrac{\partial c_{k,r}^{I}\left(\bar{\mathbf{x}},\bar{y}^{r}\right)}{\partial\tau_{k}\left(\bar{\mathbf{x}}\right)}+\\ +\dfrac{\partial d_{k,r}^{I}\left(\bar{\mathbf{x}},\bar{y}^{r}\right)}{\partial b_{k,r}^{I}\left(\bar{\mathbf{x}},\bar{y}^{r}\right)}\cdot\\ \cdot\dfrac{\partial b_{k,r}^{I}\left(\bar{\mathbf{x}},\bar{y}^{r}\right)}{\partial\left(1-\tau_{k}\left(\bar{\mathbf{x}}\right)\right)}\dfrac{\partial N\left(\tau_{k}\left(\bar{\mathbf{x}}\right)\right)}{\partial\tau_{k}\left(\bar{\mathbf{x}}\right)}\end{array}\right)+\right.$$
$$\left.+\frac{\partial I_{k,r}\left(\bar{\mathbf{x}},\bar{y}^{r}\right)}{\partial a_{k,r}^{I}\left(\bar{\mathbf{x}},\bar{y}^{r}\right)}\frac{\partial a_{k,r}^{I}\left(\bar{\mathbf{x}},\bar{y}^{r}\right)}{\partial\widetilde{N}_{1-\lambda}\left(\tau_{k}\left(\bar{\mathbf{x}}\right)\right)}\frac{\partial\widetilde{N}_{1-\lambda}\left(\tau_{k}\left(\bar{\mathbf{x}}\right)\right)}{\partial\tau_{k}\left(\bar{\mathbf{x}}\right)}\right)\qquad(6.29)$$

where

$$\frac{\partial I_{k,r}\left(\bar{\mathbf{x}},\bar{y}^r\right)}{\partial a^I_{k,r}\left(\bar{\mathbf{x}},\bar{y}^r\right)}=\frac{\partial}{\partial a^I_{k,r}\left(\bar{\mathbf{x}},\bar{y}^r\right)}G\left(\begin{matrix}a^I_{k,r}\left(\bar{\mathbf{x}},\bar{y}^r\right),d^I_{k,r}\left(\bar{\mathbf{x}},\bar{y}^r\right);\\ \alpha^I\end{matrix}\right) \tag{6.30}$$

$$\frac{\partial I_{k,r}\left(\bar{\mathbf{x}},\bar{y}^r\right)}{\partial d^I_{k,r}\left(\bar{\mathbf{x}},\bar{y}^r\right)}=\frac{\partial}{\partial d^I_{k,r}\left(\bar{\mathbf{x}},\bar{y}^r\right)}G\left(\begin{matrix}a^I_{k,r}\left(\bar{\mathbf{x}},\bar{y}^r\right),d^I_{k,r}\left(\bar{\mathbf{x}},\bar{y}^r\right);\\ \alpha^I\end{matrix}\right) \tag{6.31}$$

$$\frac{\partial I_{k,r}\left(\bar{\mathbf{x}},\bar{y}^r\right)}{\partial \alpha^I}=\frac{\partial}{\partial \alpha^I}G\left(\begin{matrix}a^I_{k,r}\left(\bar{\mathbf{x}},\bar{y}^r\right),d^I_{k,r}\left(\bar{\mathbf{x}},\bar{y}^r\right);\\ \alpha^I\end{matrix}\right) \tag{6.32}$$

$$\frac{\partial a^I_{k,r}\left(\bar{\mathbf{x}},\bar{y}^r\right)}{\partial \tilde{N}_{1-\lambda}\left(\tau_k\left(\bar{\mathbf{x}}\right)\right)}=\frac{\partial}{\partial \tilde{N}_{1-\lambda}\left(\tau_k\left(\bar{\mathbf{x}}\right)\right)}\text{avg}\left(\tilde{N}_{1-\lambda}\left(\tau_k\left(\bar{\mathbf{x}}\right)\right),\mu_{B^k}\left(\bar{y}^r\right)\right) \tag{6.33}$$

$$\frac{\partial a^I_{k,r}\left(\bar{\mathbf{x}},\bar{y}^r\right)}{\partial \mu_{B^k}\left(\bar{y}^r\right)}=\frac{\partial}{\partial \mu_{B^k}\left(\bar{y}^r\right)}\text{avg}\left(\tilde{N}_{1-\lambda}\left(\tau_k\left(\bar{\mathbf{x}}\right)\right),\mu_{B^k}\left(\bar{y}^r\right)\right) \tag{6.34}$$

$$\frac{\partial b^I_{k,r}\left(\bar{\mathbf{x}},\bar{y}^r\right)}{\partial p^I}=\frac{\partial}{\partial p^I}\ddot{H}\left(\begin{matrix}1-\tau_k\left(\bar{\mathbf{x}}\right),\mu_{B^k}\left(\bar{y}^r\right);\\ p^I,1\end{matrix}\right) \tag{6.35}$$

$$\frac{\partial b^I_{k,r}\left(\bar{\mathbf{x}},\bar{y}^r\right)}{\partial \mu_{B^k}\left(\bar{y}^r\right)}=\frac{\partial}{\partial \mu_{B^k}\left(\bar{y}^r\right)}\ddot{H}\left(\begin{matrix}1-\tau_k\left(\bar{\mathbf{x}}\right),\mu_{B^k}\left(\bar{y}^r\right);\\ p^I,1\end{matrix}\right) \tag{6.36}$$

$$\frac{\partial b^I_{k,r}\left(\bar{\mathbf{x}},\bar{y}^r\right)}{\partial \left(1-\tau_k\left(\bar{\mathbf{x}}\right)\right)}=\frac{\partial}{\partial \left(1-\tau_k\left(\bar{\mathbf{x}}\right)\right)}\ddot{H}\left(\begin{matrix}1-\tau_k\left(\bar{\mathbf{x}}\right),\mu_{B^k}\left(\bar{y}^r\right);\\ p^I,1\end{matrix}\right) \tag{6.37}$$

$$\frac{\partial c^I_{k,r}\left(\bar{\mathbf{x}},\bar{y}^r\right)}{\partial p^I}=\frac{\partial}{\partial p^I}\ddot{H}\left(\begin{matrix}\tau_k\left(\bar{\mathbf{x}}\right),\mu_{B^k}\left(\bar{y}^r\right);\\ p^I,0\end{matrix}\right) \tag{6.38}$$

$$\frac{\partial c^I_{k,r}\left(\bar{\mathbf{x}},\bar{y}^r\right)}{\partial \mu_{B^k}\left(\bar{y}^r\right)}=\frac{\partial}{\partial \mu_{B^k}\left(\bar{y}^r\right)}\ddot{H}\left(\begin{matrix}\tau_k\left(\bar{\mathbf{x}}\right),\mu_{B^k}\left(\bar{y}^r\right);\\ p^I,0\end{matrix}\right) \tag{6.39}$$

$$\frac{\partial c^I_{k,r}\left(\bar{\mathbf{x}},\bar{y}^r\right)}{\partial \tau_k\left(\bar{\mathbf{x}}\right)}=\frac{\partial}{\partial \tau_k\left(\bar{\mathbf{x}}\right)}\ddot{H}\left(\begin{matrix}\tau_k\left(\bar{\mathbf{x}}\right),\mu_{B^k}\left(\bar{y}^r\right);\\ p^I,0\end{matrix}\right) \tag{6.40}$$

$$\frac{\partial d_{k,r}^{I}\left(\bar{\mathbf{x}},\bar{y}^{r}\right)}{\partial b_{k,r}^{I}\left(\bar{\mathbf{x}},\bar{y}^{r}\right)}=\frac{\partial}{\partial b_{k,r}^{I}\left(\bar{\mathbf{x}},\bar{y}^{r}\right)}G\left(\begin{matrix}c_{k,r}^{I}\left(\bar{\mathbf{x}},\bar{y}^{r}\right),b_{k,r}^{I}\left(\bar{\mathbf{x}},\bar{y}^{r}\right);\\ \lambda\end{matrix}\right) \tag{6.41}$$

$$\frac{\partial d_{k,r}^{I}\left(\bar{\mathbf{x}},\bar{y}^{r}\right)}{\partial c_{k,r}^{I}\left(\bar{\mathbf{x}},\bar{y}^{r}\right)}=\frac{\partial}{\partial c_{k,r}^{I}\left(\bar{\mathbf{x}},\bar{y}^{r}\right)}G\left(\begin{matrix}c_{k,r}^{I}\left(\bar{\mathbf{x}},\bar{y}^{r}\right),b_{k,r}^{I}\left(\bar{\mathbf{x}},\bar{y}^{r}\right);\\ \lambda\end{matrix}\right) \tag{6.42}$$

$$\frac{\partial d_{k,r}^{I}\left(\bar{\mathbf{x}},\bar{y}^{r}\right)}{\partial \lambda}=\frac{\partial}{\partial \lambda}G\left(\begin{matrix}c_{k,r}^{I}\left(\bar{\mathbf{x}},\bar{y}^{r}\right),b_{k,r}^{I}\left(\bar{\mathbf{x}},\bar{y}^{r}\right);\\ \lambda\end{matrix}\right) \tag{6.43}$$

c) Block of aggregation

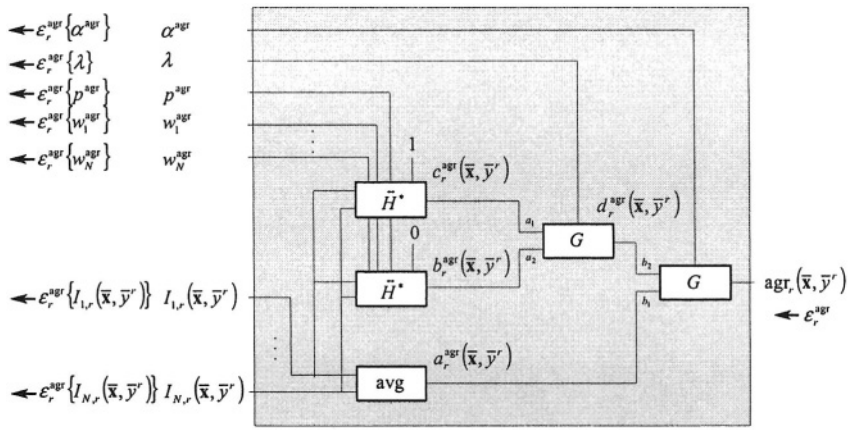

Fig. 6.12. Block of aggregation

The errors propagated through the block of aggregation (see Fig. 6.12) are given by:

$$\varepsilon_r^{\mathrm{agr}}\{\lambda\}=\varepsilon_r^{\mathrm{agr}}\frac{\partial \mathrm{agr}_r\left(\bar{\mathbf{x}},\bar{y}^r\right)}{\partial d_r^{\mathrm{agr}}\left(\bar{\mathbf{x}},\bar{y}^r\right)}\frac{\partial d_r^{\mathrm{agr}}\left(\bar{\mathbf{x}},\bar{y}^r\right)}{\partial \lambda} \tag{6.44}$$

$$\varepsilon_r^{\mathrm{agr}}\{\alpha^{\mathrm{agr}}\}=\varepsilon_r^{\mathrm{agr}}\frac{\partial \mathrm{agr}_r\left(\bar{\mathbf{x}},\bar{y}^r\right)}{\partial \alpha^{\mathrm{agr}}} \tag{6.45}$$

$$\varepsilon_r^{\text{agr}}\left\{p^{\text{agr}}\right\}=\varepsilon_r^{\text{agr}}\frac{\partial\,\text{agr}_r\left(\bar{\mathbf{x}},\bar{y}^r\right)}{\partial d_r^{\text{agr}}\left(\bar{\mathbf{x}},\bar{y}^r\right)}\left(\begin{array}{l}\dfrac{\partial d_r^{\text{agr}}\left(\bar{\mathbf{x}},\bar{y}^r\right)}{\partial c_r^{\text{agr}}\left(\bar{\mathbf{x}},\bar{y}^r\right)}\dfrac{\partial c_r^{\text{agr}}\left(\bar{\mathbf{x}},\bar{y}^r\right)}{\partial p^{\text{agr}}}+\\ +\dfrac{\partial d_r^{\text{agr}}\left(\bar{\mathbf{x}},\bar{y}^r\right)}{\partial b_r^{\text{agr}}\left(\bar{\mathbf{x}},\bar{y}^r\right)}\dfrac{\partial b_r^{\text{agr}}\left(\bar{\mathbf{x}},\bar{y}^r\right)}{\partial p^{\text{agr}}}\end{array}\right) \tag{6.46}$$

$$\varepsilon_r^{\text{agr}}\left\{w_r^{\text{agr}}\right\}=\varepsilon_r^{\text{agr}}\frac{\partial\,\text{agr}_r\left(\bar{\mathbf{x}},\bar{y}^r\right)}{\partial d_r^{\text{agr}}\left(\bar{\mathbf{x}},\bar{y}^r\right)}\left(\begin{array}{l}\dfrac{\partial d_r^{\text{agr}}\left(\bar{\mathbf{x}},\bar{y}^r\right)}{\partial c_r^{\text{agr}}\left(\bar{\mathbf{x}},\bar{y}^r\right)}\dfrac{\partial c_r^{\text{agr}}\left(\bar{\mathbf{x}},\bar{y}^r\right)}{\partial w_r^{\text{agr}}}+\\ +\dfrac{\partial d_r^{\text{agr}}\left(\bar{\mathbf{x}},\bar{y}^r\right)}{\partial b_r^{\text{agr}}\left(\bar{\mathbf{x}},\bar{y}^r\right)}\dfrac{\partial b_r^{\text{agr}}\left(\bar{\mathbf{x}},\bar{y}^r\right)}{\partial w_r^{\text{agr}}}\end{array}\right) \tag{6.47}$$

$$\varepsilon_r^{\text{agr}}\left\{I_{k,r}\left(\bar{\mathbf{x}},\bar{y}^r\right)\right\}=\varepsilon_r^{\text{agr}}\left(\begin{array}{l}\dfrac{\partial\,\text{agr}_r\left(\bar{\mathbf{x}},\bar{y}^r\right)}{\partial d_r^{\text{agr}}\left(\bar{\mathbf{x}},\bar{y}^r\right)}\left(\begin{array}{l}\dfrac{\partial d_r^{\text{agr}}\left(\bar{\mathbf{x}},\bar{y}^r\right)}{\partial c_r^{\text{agr}}\left(\bar{\mathbf{x}},\bar{y}^r\right)}\cdot\\ \cdot\dfrac{\partial c_r^{\text{agr}}\left(\bar{\mathbf{x}},\bar{y}^r\right)}{\partial I_{k,r}\left(\bar{\mathbf{x}},\bar{y}^r\right)}+\\ +\dfrac{\partial d_r^{\text{agr}}\left(\bar{\mathbf{x}},\bar{y}^r\right)}{\partial b_r^{\text{agr}}\left(\bar{\mathbf{x}},\bar{y}^r\right)}\cdot\\ \cdot\dfrac{\partial b_r^{\text{agr}}\left(\bar{\mathbf{x}},\bar{y}^r\right)}{\partial I_{k,r}\left(\bar{\mathbf{x}},\bar{y}^r\right)}\end{array}\right)+\\ +\dfrac{\partial\,\text{agr}_r\left(\bar{\mathbf{x}},\bar{y}^r\right)}{\partial a_r^{\text{agr}}\left(\bar{\mathbf{x}},\bar{y}^r\right)}\dfrac{\partial a_r^{\text{agr}}\left(\bar{\mathbf{x}},\bar{y}^r\right)}{\partial I_{k,r}\left(\bar{\mathbf{x}},\bar{y}^r\right)}\end{array}\right) \tag{6.48}$$

where

$$\frac{\partial\,\text{agr}_r\left(\bar{\mathbf{x}},\bar{y}^r\right)}{\partial a_r^{\text{agr}}\left(\bar{\mathbf{x}},\bar{y}^r\right)}=\frac{\partial}{\partial a_r^{\text{agr}}\left(\bar{\mathbf{x}},\bar{y}^r\right)}G\left(\begin{array}{c}a_r^{\text{agr}}\left(\bar{\mathbf{x}},\bar{y}^r\right),d_r^{\text{agr}}\left(\bar{\mathbf{x}},\bar{y}^r\right);\\ \alpha^{\text{agr}}\end{array}\right) \tag{6.49}$$

$$\frac{\partial\,\text{agr}_r\left(\bar{\mathbf{x}},\bar{y}^r\right)}{\partial d_r^{\text{agr}}\left(\bar{\mathbf{x}},\bar{y}^r\right)}=\frac{\partial}{\partial d_r^{\text{agr}}\left(\bar{\mathbf{x}},\bar{y}^r\right)}G\left(\begin{array}{c}a_r^{\text{agr}}\left(\bar{\mathbf{x}},\bar{y}^r\right),d_r^{\text{agr}}\left(\bar{\mathbf{x}},\bar{y}^r\right);\\ \alpha^{\text{agr}}\end{array}\right) \tag{6.50}$$

$$\frac{\partial\,\text{agr}_r\left(\bar{\mathbf{x}},\bar{y}^r\right)}{\partial\alpha^{\text{agr}}}=\frac{\partial}{\partial\alpha^{\text{agr}}}G\left(\begin{array}{c}a_r^{\text{agr}}\left(\bar{\mathbf{x}},\bar{y}^r\right),d_r^{\text{agr}}\left(\bar{\mathbf{x}},\bar{y}^r\right);\\ \alpha^{\text{agr}}\end{array}\right) \tag{6.51}$$

$$\frac{\partial a_r^{\text{agr}}\left(\bar{\mathbf{x}},\bar{y}^r\right)}{\partial I_{k,r}\left(\bar{\mathbf{x}},\bar{y}^r\right)}=\frac{\partial}{\partial I_{k,r}\left(\bar{\mathbf{x}},\bar{y}^r\right)}\text{avg}\left(I_{1,r}\left(\bar{\mathbf{x}},\bar{y}^r\right),\ldots,I_{N,r}\left(\bar{\mathbf{x}},\bar{y}^r\right)\right) \tag{6.52}$$

$$\frac{\partial b_r^{\text{agr}}\left(\overline{\mathbf{x}},\overline{y}^r\right)}{\partial p^{\text{agr}}}=\frac{\partial}{\partial p^{\text{agr}}}\ddot{H}^*\left(\begin{array}{c} I_{1,r}\left(\overline{\mathbf{x}},\overline{y}^r\right),\ldots,I_{N,r}\left(\overline{\mathbf{x}},\overline{y}^r\right);\\ w_1^{\text{agr}},\ldots,w_N^{\text{agr}},p^{\text{agr}},0\end{array}\right) \tag{6.53}$$

$$\frac{\partial b_r^{\text{agr}}\left(\overline{\mathbf{x}},\overline{y}^r\right)}{\partial w_r^{\text{agr}}}=\frac{\partial}{\partial w_r^{\text{agr}}}\ddot{H}^*\left(\begin{array}{c} I_{1,r}\left(\overline{\mathbf{x}},\overline{y}^r\right),\ldots,I_{N,r}\left(\overline{\mathbf{x}},\overline{y}^r\right);\\ w_1^{\text{agr}},\ldots,w_N^{\text{agr}},p^{\text{agr}},0\end{array}\right) \tag{6.54}$$

$$\frac{\partial b_r^{\text{agr}}\left(\overline{\mathbf{x}},\overline{y}^r\right)}{\partial I_{k,r}\left(\overline{\mathbf{x}},\overline{y}^r\right)}=\frac{\partial}{\partial I_{k,r}\left(\overline{\mathbf{x}},\overline{y}^r\right)}\ddot{H}^*\left(\begin{array}{c} I_{1,r}\left(\overline{\mathbf{x}},\overline{y}^r\right),\ldots,I_{N,r}\left(\overline{\mathbf{x}},\overline{y}^r\right);\\ w_1^{\text{agr}},\ldots,w_N^{\text{agr}},p^{\text{agr}},0\end{array}\right) \tag{6.55}$$

$$\frac{\partial c_r^{\text{agr}}\left(\overline{\mathbf{x}},\overline{y}^r\right)}{\partial p^{\text{agr}}}=\frac{\partial}{\partial p^{\text{agr}}}\ddot{H}^*\left(\begin{array}{c} I_{1,r}\left(\overline{\mathbf{x}},\overline{y}^r\right),\ldots,I_{N,r}\left(\overline{\mathbf{x}},\overline{y}^r\right);\\ w_1^{\text{agr}},\ldots,w_N^{\text{agr}},p^{\text{agr}},1\end{array}\right) \tag{6.56}$$

$$\frac{\partial c_r^{\text{agr}}\left(\overline{\mathbf{x}},\overline{y}^r\right)}{\partial w_r^{\text{agr}}}=\frac{\partial}{\partial w_r^{\text{agr}}}\ddot{H}^*\left(\begin{array}{c} I_{1,r}\left(\overline{\mathbf{x}},\overline{y}^r\right),\ldots,I_{N,r}\left(\overline{\mathbf{x}},\overline{y}^r\right);\\ w_1^{\text{agr}},\ldots,w_N^{\text{agr}},p^{\text{agr}},1\end{array}\right) \tag{6.57}$$

$$\frac{\partial c_r^{\text{agr}}\left(\overline{\mathbf{x}},\overline{y}^r\right)}{\partial I_{k,r}\left(\overline{\mathbf{x}},\overline{y}^r\right)}=\frac{\partial}{\partial I_{k,r}\left(\overline{\mathbf{x}},\overline{y}^r\right)}\ddot{H}^*\left(\begin{array}{c} I_{1,r}\left(\overline{\mathbf{x}},\overline{y}^r\right),\ldots,I_{N,r}\left(\overline{\mathbf{x}},\overline{y}^r\right);\\ w_1^{\text{agr}},\ldots,w_N^{\text{agr}},p^{\text{agr}},1\end{array}\right) \tag{6.58}$$

$$\frac{\partial d_r^{\text{agr}}\left(\overline{\mathbf{x}},\overline{y}^r\right)}{\partial c_r^{\text{agr}}\left(\overline{\mathbf{x}},\overline{y}^r\right)}=\frac{\partial}{\partial c_r^{\text{agr}}\left(\overline{\mathbf{x}},\overline{y}^r\right)}G\left(\begin{array}{c} c_r^{\text{agr}}\left(\overline{\mathbf{x}},\overline{y}^r\right),b_r^{\text{agr}}\left(\overline{\mathbf{x}},\overline{y}^r\right);\\ \lambda\end{array}\right) \tag{6.59}$$

$$\frac{\partial d_r^{\text{agr}}\left(\overline{\mathbf{x}},\overline{y}^r\right)}{\partial b_r^{\text{agr}}\left(\overline{\mathbf{x}},\overline{y}^r\right)}=\frac{\partial}{\partial b_r^{\text{agr}}\left(\overline{\mathbf{x}},\overline{y}^r\right)}G\left(\begin{array}{c} c_r^{\text{agr}}\left(\overline{\mathbf{x}},\overline{y}^r\right),b_r^{\text{agr}}\left(\overline{\mathbf{x}},\overline{y}^r\right);\\ \lambda\end{array}\right) \tag{6.60}$$

$$\frac{\partial d_r^{\text{agr}}\left(\overline{\mathbf{x}},\overline{y}^r\right)}{\partial \lambda}=\frac{\partial}{\partial \lambda}G\left(\begin{array}{c} c_r^{\text{agr}}\left(\overline{\mathbf{x}},\overline{y}^r\right),b_r^{\text{agr}}\left(\overline{\mathbf{x}},\overline{y}^r\right);\\ \lambda\end{array}\right) \tag{6.61}$$

6.7. SIMULATION RESULTS

In this section we present four simulations of the AND-type neuro-fuzzy systems. We use benchmarks described in Sections 3.6. Each of the four simulations is designed in the same fashion:

(i) In the first experiment, based on the input-output data, we learn the parameters of the membership functions and a system type $\lambda \in [0,1]$

of the basic compromise system described in Section 6.3. It will be seen that the optimal values of λ, determined by the gradient procedure, are either zero or one.

(ii) In the second experiment, we learn the parameters of the membership functions of the basic compromise system described in Section 6.3 choosing values of λ as opposite to that obtained in experiment (i). Obviously, we expect a worse performance of the neuro-fuzzy system comparing with experiment (i).

(iii) In the third experiment, we learn the parameters of the membership functions, system type $\lambda \in [0,1]$ and soft parameters $\alpha^{\tau} \in [0,1]$, $\alpha^{I} \in [0,1]$, $\alpha^{\text{agr}} \in [0,1]$ of the soft compromise system described in Section 6.4 assuming that classical (not-parameterised) triangular norms are applied.

(iv) In the fourth experiment, we learn the same parameters as in the third experiment and, moreover, the weights $w_{i,k}^{\tau} \in [0,1]$, $i = 1,\ldots,n$, $k = 1,...,N$, in the antecedents of rules and weights $w_k^{\text{agr}} \in [0,1]$, $k = 1,\ldots,N$, of the aggregation operator of the rules. In all diagrams (weights representation) we separate $w_{i,k}^{\tau} \in [0,1]$, $i = 1,\ldots,n$, $k = 1,...,N$, from $w_k^{\text{agr}} \in [0,1]$, $k = 1,\ldots,N$, by a vertical dashed line.

(v) In each of the above simulations we apply the Zadeh H-implication and the algebraic H-implication described in examples 5.5 and 5.6, respectively. In separate experiments we repeat simulations (i)-(iv) replacing the Zadeh H-implication and the algebraic H-implication by quasi-implications generated by parameterised triangular norms: the Dombi H-implication and the Yager H-implication. In these simulations we additionally incorporate parameters $p^{\tau} \in [0,\infty)$, $p^{I} \in [0,\infty)$, $p^{\text{agr}} \in [0,\infty)$.

Box and Jenkins Gas Furnace problem

The experimental results for the Box and Jenkins Gas Furnace problem are depicted in Tables 6.1 and 6.2 for the not-parameterised (Zadeh and algebraic) and parameterised (Dombi and Yager) H-functions, respectively. For experiment (iv) the final values (after learning) of weights $w_{i,k}^{r} \in [0,1]$ and $w_k^{agr} \in [0,1]$, $i = 1,...,6$, $k = 1,...,4$, are shown in Fig. 6.13 (Zadeh and algebraic H-functions) and Fig. 6.14 (Dombi and Yager H-functions).

Table 6.1 Experimental results

AND-TYPE NFIS WITH NON-PARAMETRISED H-FUNCTIONS (BOX AND JENKINS GAS FURNACE PROBLEM)						
Experiment number	Name of flexibility parameter	Initial values	Final values after learning		RMSE (learning sequence)	
			Zadeh H-function	Algebraic H-function	Zadeh H-function	Algebraic H-function
i	λ	0.5	0.0000	0.0000	0.5029	0.5012
ii	λ	1	-	-	0.6389	0.6324
iii	λ	0.5	0.0000	0.0000	0.4125	0.4013
	α^r	1	0.9847	0.9741		
	α^l	1	0.9546	0.9558		
	α^{agr}	1	0.9938	0.9349		
iv	λ	0.5	0.0000	0.0000	0.2477	0.2442
	α^r	1	0.9808	0.9437		
	α^l	1	0.9994	0.9507		
	α^{agr}	1	0.9683	0.9504		
	$\mathbf{w}^r$	**1**	Fig. 6.13-a	Fig. 6.13-b		
	$\mathbf{w}^{agr}$	**1**	Fig. 6.13-a	Fig. 6.13-b		

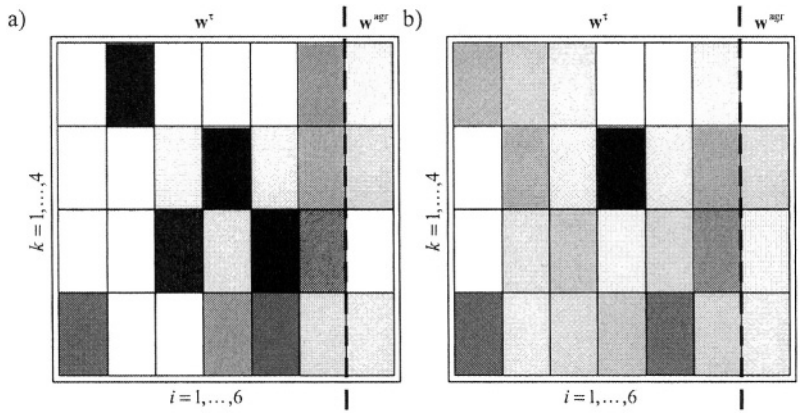

Fig. 6.13. Weights representation in the Modeling of Box and Jenkins Gas Furnace problem for AND-type NFIS and a) Zadeh H-function, b) algebraic H-function

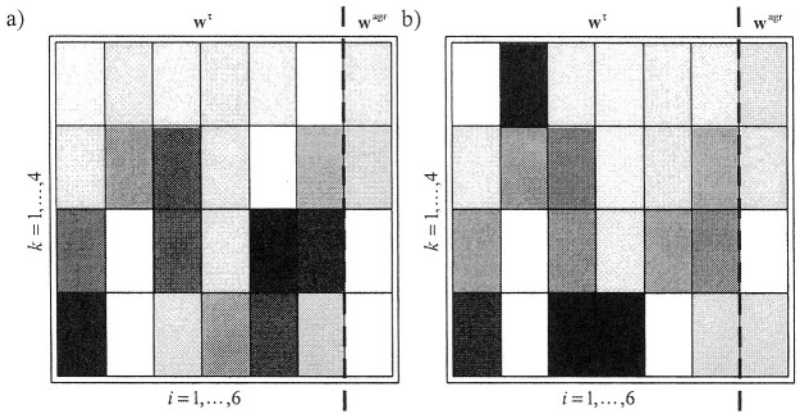

Fig. 6.14. Weights representation in the Modeling of Box and Jenkins Gas Furnace problem for AND-type NFIS and a) Dombi H-function, b) Yager H-function

Table 6.2 Experimental results

AND-TYPE NFIS WITH PARAMETRISED H-FUNCTIONS (BOX AND JENKINS GAS FURNACE PROBLEM)						
Experiment number	**Name of flexibility parameter**	**Initial values**	**Final values after learning**		**RMSE (learning sequence)**	
			Dombi H-function	**Yager H-function**	**Dombi H-function**	**Yager H-function**
i	λ	0.5	0.0000	0.0000	0.5169	0.5059
ii	λ	1	-	-	0.6523	0.6588
iii	λ	0.5	0.0000	0.0000	0.4352	0.4236
	p^r	10	9.8567	21.5916		
	p^l	10	11.7979	22.3708		
	p^{agr}	10	9.8144	4.4304		
	α^r	1	0.9563	0.9957		
	α^l	1	0.9942	0.9388		
	α^{agr}	1	0.9407	0.9665		
iv	λ	0.5	0.0000	0.0000	0.2481	0.2487
	p^r	10	11.2171	24.2793		
	p^l	10	9.6944	23.6784		
	p^{agr}	10	11.2106	2.9272		
	α^r	1	0.9592	0.9604		
	α^l	1	0.9958	0.9500		
	α^{agr}	1	0.9532	0.9623		
	$\mathbf{w}^r$	**1**	Fig. 6.14-a	Fig. 6.14-b		
	$\mathbf{w}^{agr}$	**1**	Fig. 6.14-a	Fig. 6.14-b		

Glass Identification problem

The experimental results for the Glass Identification problem are depicted in Tables 6.3 and 6.4 for the not-parameterised (Zadeh and algebraic) and parameterised (Dombi and Yager) H-functions, respectively. For experiment (iv) the final values (after learning) of weights $w_{i,k}^r \in [0,1]$ and $w_k^{agr} \in [0,1]$, $i = 1,...,9$, $k = 1,\ldots,2$, are shown in Fig. 6.15 (Zadeh and algebraic H-functions) and Fig. 6.16 (Dombi and Yager H-functions).

Table 6.3 Experimental results

AND-TYPE NFIS WITH NON-PARAMETRISED H-FUNCTIONS (GLASS IDENTIFICATION PROBLEM)								
Experiment number	Name of flexibility parameter	Initial values	Final values after learning		Mistakes [%] (learning sequence)		Mistakes [%] (testing sequence)	
			Zadeh H-function	Algebraic H-function	Zadeh H-function	Algebraic H-function	Zadeh H-function	Algebraic H-function
i	λ	0.5	1.0000	1.0000	4.00	3.33	3.13	3.13
ii	λ	0	-	-	4.00	4.67	3.13	3.13
iii	λ	0.5	1.0000	1.0000	2.67	2.67	1.56	3.13
	α^r	1	0.0047	0.0564				
	α^l	1	0.9493	0.9717				
	α^{agr}	1	0.9201	0.9480				
iv	λ	0.5	1.0000	1.0000	2.00	2.00	1.56	1.56
	α^r	1	0.0691	0.0498				
	α^l	1	0.9612	0.9170				
	α^{agr}	1	0.9434	0.9791				
	$\mathbf{w}^r$	**1**	Fig. 6.15-a	Fig. 6.15-b				
	$\mathbf{w}^{agr}$	**1**	Fig. 6.15-a	Fig. 6.15-b				

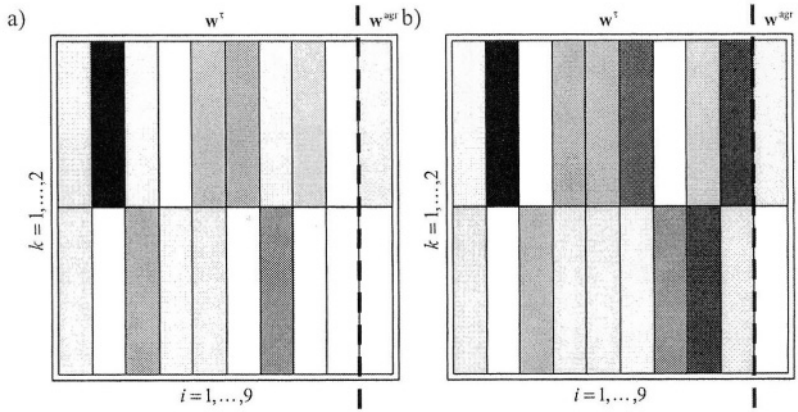

Fig. 6.15. Weights representation in the Glass Identification problem for AND-type NFIS and a) Zadeh H-function, b) algebraic H-function

Table 6.4 Experimental results

AND-TYPE NFIS WITH PARAMETRISED H-FUNCTIONS (GLASS IDENTIFICATION PROBLEM)								
Experiment number	Name of flexibility parameter	Initial values	Final values after learning		Mistakes [%] (learning sequence)		Mistakes [%] (testing sequence)	
			Dombi H-function	Yager H-function	Dombi H-function	Yager H-function	Dombi H-function	Yager H-function
i	λ	0.5	1.0000	1.0000	3.33	3.33	3.13	3.13
ii	λ	0	-	-	4.00	4.00	3.13	3.13
iii	λ	0.5	1.0000	1.0000	2.67	2.67	1.56	1.56
	p^r	10	9.8816	13.3516				
	p^l	10	9.7060	9.9069				
	p^{agr}	10	9.2553	10.8202				
	α^r	1	0.0166	0.1211				
	α^l	1	0.8357	0.9804				
	α^{agr}	1	0.9498	0.9079				
iv	λ	0.5	1.0000	1.0000	1.33	1.33	1.56	1.56
	p^r	10	9.7505	12.3408				
	p^l	10	9.9503	10.2859				
	p^{agr}	10	9.7660	9.3689				
	α^r	1	0.0074	0.0951				
	α^l	1	0.8551	0.9202				
	α^{agr}	1	0.9337	0.9428				
	$\mathbf{w}^r$	**1**	Fig. 6.16-a	Fig. 6.16-b				
	$\mathbf{w}^{agr}$	**1**	Fig. 6.16-a	Fig. 6.16-b				

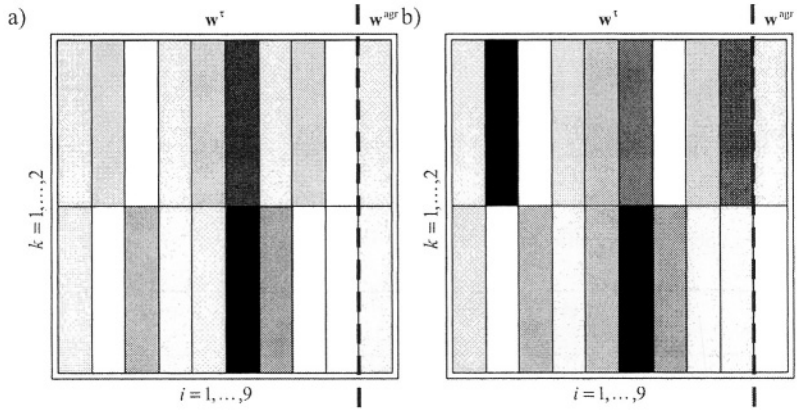

Fig. 6.16. Weights representation in the Glass Identification problem for AND-type NFIS and a) Dombi H-function, b) Yager H-function

Modeling of Static Nonlinear Function (HANG) problem

The experimental results for the Modeling of the Static Nonlinear Function problem are shown in Tables 6.5 and 6.6 for the not-parameterised (Zadeh and algebraic) and parameterised (Dombi and Yager) H-functions, respectively. For experiment (iv) the final values (after learning) of weights $w_{i,k}^{\tau} \in [0,1]$ and $w_k^{agr} \in [0,1]$, $i = 1,...,2$, $k = 1,...,5$, are depicted in Fig. 6.17 (Zadeh and algebraic H-functions) and Fig. 6.18 (Dombi and Yager H-functions).

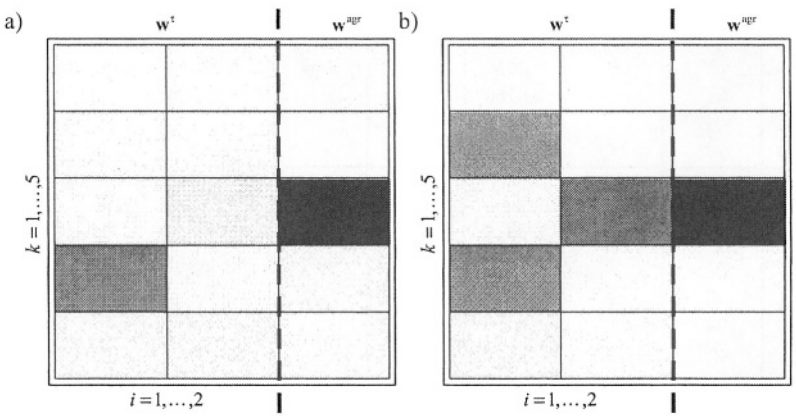

Fig. 6.17. Weights representation in the HANG problem for AND-type NFIS and a) Zadeh H-function, b) algebraic H-function

Table 6.5 Experimental results

AND-TYPE NFIS WITH NON-PARAMETRISED H-FUNCTIONS (MODELING OF STATIC NONLINEAR FUNCTION PROBLEM)						
Experiment number	**Name of flexibility parameter**	**Initial values**	**Final values after learning**		**RMSE (learning sequence)**	
			Zadeh H-function	**Algebraic H-function**	**Zadeh H-function**	**Algebraic H-function**
i	λ	0.5	0.0000	0.0000	0.1218	0.1028
ii	λ	1	-	-	0.1407	0.1264
iii	λ	0.5	0.0000	0.0000	0.1160	0.1073
	α^r	1	0.0324	0.9410		
	α^l	1	0.9353	0.9439		
	α^{agr}	1	0.8872	0.9628		
iv	λ	0.5	0.0000	0.0000	0.1008	0.0936
	α^r	1	0.0461	0.9850		
	α^l	1	0.9684	0.9875		
	α^{agr}	1	0.9269	0.9903		
	$\mathbf{w}^r$	**1**	Fig. 6.17-a	Fig. 6.17-b		
	$\mathbf{w}^{agr}$	**1**	Fig. 6.17-a	Fig. 6.17-b		

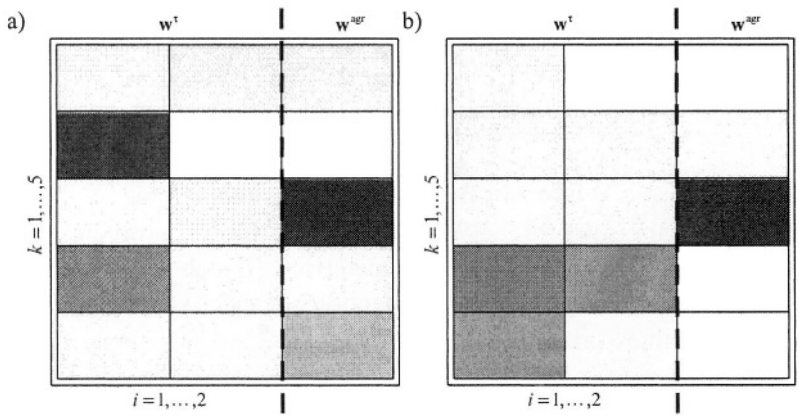

Fig. 6.18. Weights representation in the HANG problem for AND-type NFIS and a) Dombi H-function, b) Yager H-function

Table 6.6 Experimental results

AND-TYPE NFIS WITH PARAMETRISED H-FUNCTIONS (MODELING OF STATIC NONLINEAR FUNCTION PROBLEM)						
Experiment number	Name of flexibility parameter	Initial values	Final values after learning		RMSE (learning sequence)	
			Dombi H-function	Yager H-function	Dombi H-function	Yager H-function
i	λ	0.5	0.0000	0.0000	0.1386	0.1047
ii	λ	1	-	-	0.1586	0.1297
iii	λ	0.5	0.0000	0.0000	0.1247	0.1118
	p^r	10	18.2764	5.6494		
	p^l	10	10.8980	11.9890		
	p^{agr}	10	10.3142	3.2564		
	α^r	1	0.0285	0.9539		
	α^l	1	0.9774	0.9631		
	α^{agr}	1	0.9665	0.9632		
iv	λ	0.5	0.0000	0.0000	0.1096	0.0986
	p^r	10	17.9950	4.4833		
	p^l	10	9.9850	11.9089		
	p^{agr}	10	10.7893	3.4419		
	α^r	1	0.0027	0.9295		
	α^l	1	0.9315	0.9832		
	α^{agr}	1	0.9404	0.9706		
	$\mathbf{w}^r$	**1**	Fig. 6.18-a	Fig. 6.18-b		
	$\mathbf{w}^{agr}$	**1**	Fig. 6.18-a	Fig. 6.18-b		

Wisconsin Breast Cancer problem

The experimental results for the Wisconsin Breast Cancer problem are shown in Tables 6.7 and 6.8 for the not-parameterised (Zadeh and algebraic) and parameterised (Dombi and Yager) H-functions, respectively. For experiment (iv) the final values (after learning) of weights $w_{i,k}^r \in [0,1]$ and $w_k^{agr} \in [0,1]$, $i = 1,\ldots,9$, $k = 1,\ldots,2$, are depicted in Fig. 6.19 (Zadeh and algebraic H-functions) and Fig. 6.20 (Dombi and Yager H-functions).

Table 6.7 Experimental results

AND-TYPE NFIS WITH NON-PARAMETRISED H-FUNCTIONS (WISCONSIN BREAST CANCER PROBLEM)								
Experiment number	Name of flexibility parameter	Initial values	Final values after learning		Mistakes [%] (learning sequence)		Mistakes [%] (testing sequence)	
			Zadeh H-function	Algebraic H-function	Zadeh H-function	Algebraic H-function	Zadeh H-function	Algebraic H-function
i	λ	0.5	1.0000	1.0000	2.93	2.51	2.44	1.95
ii	λ	0	-	-	3.14	2.72	2.93	2.44
iii	λ	0.5	1.0000	1.0000	2.72	2.30	2.44	1.95
	α^{τ}	1	0.0537	0.0456				
	$\alpha^{\prime}$	1	0.9624	0.9739				
	α^{agr}	1	0.9690	0.9808				
iv	λ	0.5	1.0000	1.0000	1.88	1.67	1.46	1.46
	α^{τ}	1	0.0544	0.0082				
	$\alpha^{\prime}$	1	0.9890	0.8782				
	α^{agr}	1	0.9920	0.9992				
	$\mathbf{w}^{\tau}$	**1**	Fig. 6.19-a	Fig. 6.19-b				
	$\mathbf{w}^{agr}$	**1**	Fig. 6.19-a	Fig. 6.19-b				

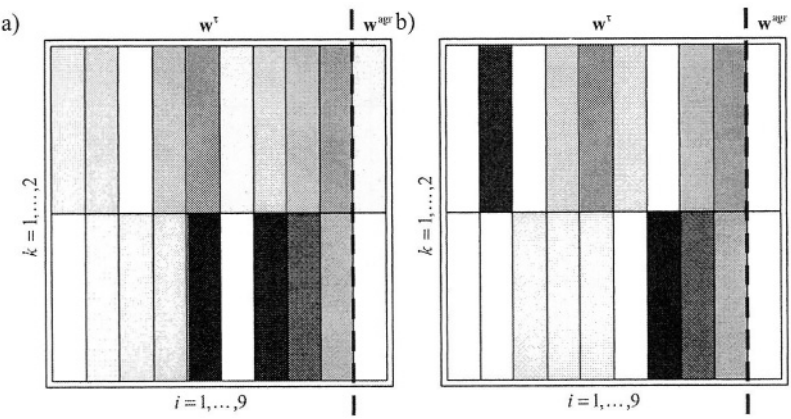

Fig. 6.19. Weights representation in the Wisconsin Breast Cancer problem for AND-type NFIS and a) Zadeh H-function, b) algebraic H-function

Table 6.8 Experimental results

AND-TYPE NFIS WITH PARAMETRISED H-FUNCTIONS (WISCONSIN BREAST CANCER PROBLEM)								
Experiment number	Name of flexibility parameter	Initial values	Final values after learning		Mistakes [%] (learning sequence)		Mistakes [%] (testing sequence)	
			Dombi H-function	Yager H-function	Dombi H-function	Yager H-function	Dombi H-function	Yager H-function
i	λ	0.5	1.0000	1.0000	2.30	2.09	1.95	1.95
ii	λ	0	-	-	2.51	2.30	1.95	1.95
iii	λ	0.5	1.0000	1.0000	2.09	1.88	1.95	1.46
	p^r	10	15.5566	5.3325				
	p^l	10	15.3728	1.0547				
	p^{agr}	10	12.8735	10.8646				
	α^r	1	0.0543	0.0444				
	α^l	1	0.9258	0.9885				
	α^{agr}	1	0.9545	0.9739				
iv	λ	0.5	1.0000	1.0000	1.46	1.46	1.46	1.46
	p^r	10	14.5891	12.5328				
	p^l	10	15.0974	1.4338				
	p^{agr}	10	5.3193	12.2100				
	α^r	1	0.0028	0.0496				
	α^l	1	0.8741	0.9746				
	α^{agr}	1	0.9098	0.9240				
	$\mathbf{w}^r$	**1**	Fig. 6.20-a	Fig. 6.20-b				
	$\mathbf{w}^{agr}$	**1**	Fig. 6.20-a	Fig. 6.20-b				

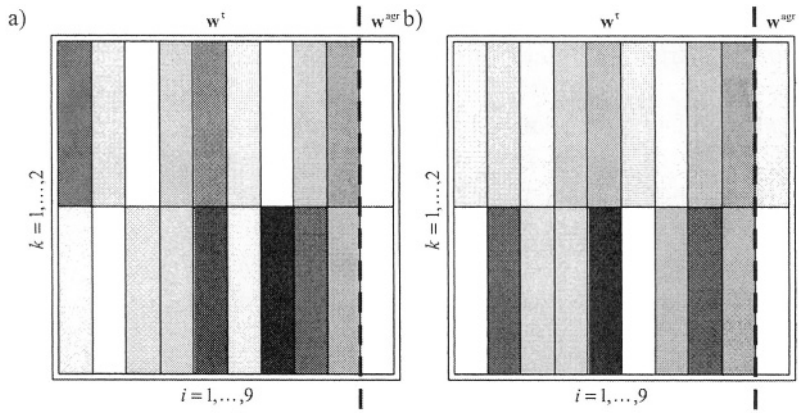

Fig. 6.20. Weights representation in the Wisconsin Breast Cancer problem for AND-type NFIS and a) Dombi H-function, b) Yager H-function

6.8. SUMMARY AND DISCUSSION

In this chapter we studied the fuzzy inference characterized by the simultaneous appearance of the Mamdani-type and the logical-type reasoning. Three neuro-fuzzy structures have been developed based on the flexibility parameters and weights described in Chapter 4. From simulations given in Section 6.7 it follows that the incorporation of soft parameters α^{τ}, α^{I} and α^{agr} improves the performance of basic neuro-fuzzy systems. The best results are observed for neuro-fuzzy systems based on weighted triangular norms.

6.9. PROBLEMS

Problem 6.1. Derive a compromise fuzzy implication based on the Yager triangular norm.

Problem 6.2. Incorporate the Yager triangular norm to the construction of the basic compromise neuro-fuzzy system.

Problem 6.3. Derive a compromise fuzzy implication based on the Larsen inference and the Reichenbach fuzzy implication.

Problem 6.4. Apply the compromise fuzzy implication derived in Problem 6.3 to the construction of the basic compromise neuro-fuzzy system.

Problem 6.5. Derive a compromise fuzzy implication based on the Larsen inference and the binary fuzzy implication. Observe that in this problem model (6.2) is not true, i.e. T and S are not dual triangular norms.

Problem 6.6. Apply the compromise fuzzy implication derived in Problem 6.5 to the construction of the soft compromise systems.

Problem 6.7. Apply the exponential operator [9, 45]

$$E_{T,S,\lambda}(a_1,\ldots,a_n) = \left(T\{a_1,\ldots,a_n\}\right)^{1-\lambda}\left(S\{a_1,\ldots,a_n\}\right)^{\lambda} \tag{6.62}$$

to construct a basic compromise NFIS.

Problem 6.8. Apply operator (6.62) to construct a soft compromise NFIS.

Problem 6.9. Apply operator (6.62) to construct a weighted compromise NFIS.

Chapter 7

FLEXIBLE MAMDANI-TYPE NEURO-FUZZY SYSTEMS

7.1. INTRODUCTION

In Chapter 5 we constructed OR-type flexible neuro-fuzzy systems with quasi-implications

$$I(a,b;v)=H\left(\tilde{N}_{1-v}(a),b;v\right) \tag{7.1}$$

whereas in Chapter 6 we developed AND-type flexible neuro-fuzzy systems with compromise reasoning

$$I(a,b;\lambda)=(1-\lambda)T\{a,b\}+\lambda S\{1-a,b\} \tag{7.2}$$

Setting $v=0$ in model (7.1) or $\lambda=0$ in model (7.2) we get the Mamdani-type system. Therefore, for $v=\lambda=0$ all the results (including learning procedures) of Chapter 5 and 6 are applicable to the Mamdani-type systems.

In this chapter we study the Mamdani-type systems from another perspective. Our goal is to reveal the connectionist nature of these systems depending on the functional form of used "engineering implications" (t-norms). We will study various Mamdani-type architectures described in terms of "engineering implications".

7.2. PROBLEM DESCRIPTION

As we mentioned in Section 5.5, the Mamdani-type system can be expressed in terms of an H function:

$$\tau_k(\bar{x}) = H\left(\mu_{A_1^k}(\bar{x}_1),\ldots,\mu_{A_n^k}(\bar{x}_n)\ ;0\right) \tag{7.3}$$

$$I_{k,r}(\bar{\mathbf{x}},\bar{y}^r) = H\left(\widetilde{N}_1(\tau_k(\bar{\mathbf{x}})),\mu_{B^k}(\bar{y}^r)\ ;0\right) \tag{7.4}$$

$$\mathrm{agr}_r(\bar{\mathbf{x}},\bar{y}^r) = H\left(I_{1,r}(\bar{\mathbf{x}},\bar{y}^r),\ldots,I_{N,r}(\bar{\mathbf{x}},\bar{y}^r)\ ;1\right) \tag{7.5}$$

The above description and corresponding neuro-fuzzy structures do not reflect the functional form of a t-norm used for the antecedent-consequent connection. Therefore our problem is to design a neuro-fuzzy system realizing a mapping $f:\mathbf{X}\to\mathbf{Y}$ such that the fuzzy inference is described by the "engineering implication" and the structure of the system explicitly depends on the form of the t-norm serving as the "engineering implication". Moreover, we will incorporate certainty weights to the aggregation of rules and connectives of antecedents. In simulations we also study the Mamdani-type systems with soft triangular norms.

7.3. NEURO-FUZZY STRUCTURES

In case of weighted fuzzy systems, the firing strength of rules is defined by

$$\mu_{A^k}(\bar{\mathbf{x}}) = \underset{i=1}{\overset{n}{T^*}}\left\{\mu_{A_i^k}(\bar{x}_i), w_{i,k}^{\tau}\right\} \tag{7.6}$$

where $w_{i,k}^{\tau}$ is the weight of the i-th component of the input vector in the k-th knowledge base rule and T^* is defined by (4.7).

On the basis of the knowledge comprised in the rule-base and formulas (3.14)-(3.16) we obtain individual fuzzy sets $\bar{B}^k$, $k = 1,\ldots,N$. Then we have to aggregate them into one fuzzy set B'. In case of the Mamdani approach, the aggregation is carried out by (3.18) and (3.19), and in case of the weighted aggregation we have

$$\mu_{B'}(y) = \underset{k=1}{\overset{N}{S^*}}\left\{\mu_{\bar{B}^k}(y), w_k^{\mathrm{agr}}\right\} \tag{7.7}$$

where w_k^{agr} is the weight of the k-th rule and S^* is defined by (4.9).

Combining (7.6) and (7.7) with (3.14) and (3.23), we obtain

$$\bar{y} = \frac{\sum_{r=1}^{N} \bar{y}^r \cdot \overset{N}{\underset{k=1}{S^*}}\left\{ I\left(\overset{n}{\underset{i=1}{T^*}}\left\{ \mu_{A_i^k}(\bar{x}_i), w_{i,k}^{\tau} \right\}, \mu_{B^k}(\bar{y}^r) \right), w_k^{\text{agr}} \right\}}{\sum_{r=1}^{N} \overset{N}{\underset{k=1}{S^*}}\left\{ I\left(\overset{n}{\underset{i=1}{T^*}}\left\{ \mu_{A_i^k}(\bar{x}_i), w_{i,k}^{\tau} \right\}, \mu_{B^k}(\bar{y}^r) \right), w_k^{\text{agr}} \right\}} \tag{7.8}$$

Since "engineering implications" (3.14) used in the Mamdani approach are described by t-norms, formula (7.8) takes the form

$$\bar{y} = \frac{\sum_{r=1}^{N} \bar{y}^r \cdot \overset{N}{\underset{k=1}{S^*}}\left\{ T\left\{ \overset{n}{\underset{i=1}{T^*}}\left\{ \mu_{A_i^k}(\bar{x}_i), w_{i,k}^{\tau} \right\}, \mu_{B^k}(\bar{y}^r) \right\}, w_k^{\text{agr}} \right\}}{\sum_{r=1}^{N} \overset{N}{\underset{k=1}{S^*}}\left\{ T\left\{ \overset{n}{\underset{i=1}{T^*}}\left\{ \mu_{A_i^k}(\bar{x}_i), w_{i,k}^{\tau} \right\}, \mu_{B^k}(\bar{y}^r) \right\}, w_k^{\text{agr}} \right\}} \tag{7.9}$$

We will now further elaborate equation (7.8). Assuming that fuzzy sets B^k are normal, i.e.

$$\mu_{B^k}(\bar{y}^k) = 1 \tag{7.10}$$

and using the boundary condition of triangular norms $T\{a,1\} = a$, we get

$$\mu_{\bar{B}^k}(\bar{y}^k) = \mu_{A^k}(\bar{\mathbf{x}}) \tag{7.11}$$

Therefore, for $y = \bar{y}^r$ we can rewrite (7.7) as follows

$$\mu_{B^r}(\bar{y}^r) = S^*\left\{ \left(\mu_{A^r}(\bar{\mathbf{x}}), w_r^{\text{agr}} \right), \overset{N}{\underset{\substack{k=1 \\ k \neq r}}{S^*}}\left\{ I\left(\mu_{A^k}(\bar{\mathbf{x}}), \mu_{B^k}(\bar{y}^r) \right), w_k^{\text{agr}} \right\} \right\} \tag{7.12}$$

Consequently, combining equations (7.12) and (7.6) with (3.14) and (3.23), equation (7.8) becomes

$$\bar{y} = \frac{\sum_{r=1}^{N} y^r \cdot S^*\left\{ \left(\overset{n}{\underset{i=1}{T^*}}\left\{ \mu_{A_i^r}(\bar{x}_i) \right\}, w_r^{\text{agr}} \right), \overset{N}{\underset{\substack{k=1 \\ k \neq r}}{S^*}}\left\{ I\left(\overset{n}{\underset{i=1}{T^*}}\left\{ \mu_{A_i^k}(\bar{x}_i) \right\}, \mu_{B^k}(\bar{y}^r) \right), w_k^{\text{agr}} \right\} \right\}}{\sum_{r=1}^{N} S^*\left\{ \left(\overset{n}{\underset{i=1}{T^*}}\left\{ \mu_{A_i^r}(\bar{x}_i) \right\}, w_r^{\text{agr}} \right), \overset{N}{\underset{\substack{k=1 \\ k \neq r}}{S^*}}\left\{ I\left(\overset{n}{\underset{i=1}{T^*}}\left\{ \mu_{A_i^k}(\bar{x}_i) \right\}, \mu_{B^k}(\bar{y}^r) \right), w_k^{\text{agr}} \right\} \right\}} \tag{7.13}$$

It will be convenient to use the following notation

$$p_{k,r} = \mu_{B^k}(\bar{y}^r) \tag{7.14}$$

and

$$\tilde{p}_{k,r} = 1 - p_{k,r} = 1 - \mu_{B^k}(\bar{y}^r) \tag{7.15}$$

Using (7.14) we can rewrite (7.13) in the form

$$\bar{y} = \frac{\sum_{r=1}^{N} \bar{y}^r \cdot S\left\{\left(\begin{array}{c} \underset{i=1}{\overset{n}{T^*}}\left\{\begin{array}{c}\mu_{A_i^r}(\bar{x}_i),\\ w_{i,r}^{\tau}\end{array}\right\},\\ w_r^{\text{agr}}\end{array}\right), \underset{\substack{k=1\\k\neq r}}{\overset{N}{S^*}}\left\{I\left(\begin{array}{c}\underset{i=1}{\overset{n}{T^*}}\left\{\begin{array}{c}\mu_{A_i^k}(\bar{x}_i),\\ w_{i,k}^{\tau}\end{array}\right\}, p_{k,r}\end{array}\right),\right.\right.\left.\left.\begin{array}{c}\\ w_k^{\text{agr}}\end{array}\right)\right\}\right\}}{\sum_{r=1}^{N} S\left\{\left(\begin{array}{c} \underset{i=1}{\overset{n}{T^*}}\left\{\begin{array}{c}\mu_{A_i^r}(\bar{x}_i),\\ w_{i,r}^{\tau}\end{array}\right\},\\ w_r^{\text{agr}}\end{array}\right), \underset{\substack{k=1\\k\neq r}}{\overset{N}{S^*}}\left\{I\left(\begin{array}{c}\underset{i=1}{\overset{n}{T^*}}\left\{\begin{array}{c}\mu_{A_i^k}(\bar{x}_i),\\ w_{i,k}^{\tau}\end{array}\right\}, p_{k,r}\end{array}\right),\right.\right.\left.\left.\begin{array}{c}\\ w_k^{\text{agr}}\end{array}\right)\right\}\right\}} \tag{7.16}$$

Mamdani (min) relation

$$I(a,b) = \min\{a,b\} \tag{7.17}$$

Putting function (7.17) directly into (7.16), we obtain a formula describing a neuro-fuzzy inference system with the Mamdani inference

$$\bar{y} = \frac{\sum_{r=1}^{N} \bar{y}^r \cdot S^*\left\{\left(\begin{array}{c} \underset{i=1}{\overset{n}{T^*}}\left\{\begin{array}{c}\mu_{A_i^r}(\bar{x}_i),\\ w_{i,r}^{\tau}\end{array}\right\},\\ w_r^{\text{agr}}\end{array}\right), \underset{\substack{k=1\\k\neq r}}{\overset{N}{S^*}}\left\{\min\left\{\begin{array}{c}\underset{i=1}{\overset{n}{T^*}}\left\{\begin{array}{c}\mu_{A_i^k}(\bar{x}_i),\\ w_{i,k}^{\tau}\end{array}\right\}, p_{k,r}\end{array}\right\},\right.\right.\left.\left.\begin{array}{c}\\ w_k^{\text{agr}}\end{array}\right)\right\}\right\}}{\sum_{r=1}^{N} S^*\left\{\left(\begin{array}{c} \underset{i=1}{\overset{n}{T^*}}\left\{\begin{array}{c}\mu_{A_i^r}(\bar{x}_i),\\ w_{i,r}^{\tau}\end{array}\right\},\\ w_r^{\text{agr}}\end{array}\right), \underset{\substack{k=1\\k\neq r}}{\overset{N}{S^*}}\left\{\min\left\{\begin{array}{c}\underset{i=1}{\overset{n}{T^*}}\left\{\begin{array}{c}\mu_{A_i^k}(\bar{x}_i),\\ w_{i,k}^{\tau}\end{array}\right\}, p_{k,r}\end{array}\right\},\right.\right.\left.\left.\begin{array}{c}\\ w_k^{\text{agr}}\end{array}\right)\right\}\right\}} \tag{7.18}$$

Figure 7.1 shows the structure of the neuro-fuzzy system described by formula (7.18).

Larsen (product) relation

$$I(a,b) = a \cdot b \tag{7.19}$$

Putting function (7.19) directly into (7.16), we obtain a formula describing a neuro-fuzzy system with the Larsen inference

$$\bar{y} = \frac{\sum_{r=1}^{N} \bar{y}^r \cdot S^*\left\{\left(\begin{array}{c} \overset{n}{\underset{i=1}{T^*}}\left\{\begin{array}{c}\mu_{A_i^r}(\bar{x}_i),\\ w_{i,r}^{\tau}\end{array}\right\},\\ w_r^{\text{agr}}\end{array}\right), \underset{\substack{k=1\\k\neq r}}{\overset{N}{S^*}}\left\{\left(\begin{array}{c} \overset{n}{\underset{i=1}{T^*}}\left\{\begin{array}{c}\mu_{A_i^k}(\bar{x}_i),\\ w_{i,k}^{\tau}\end{array}\right\}\cdot p_{k,r},\\ w_k^{\text{agr}}\end{array}\right)\right\}\right\}}{\sum_{r=1}^{N} S^*\left\{\left(\begin{array}{c} \overset{n}{\underset{i=1}{T^*}}\left\{\begin{array}{c}\mu_{A_i^r}(\bar{x}_i),\\ w_{i,r}^{\tau}\end{array}\right\},\\ w_r^{\text{agr}}\end{array}\right), \underset{\substack{k=1\\k\neq r}}{\overset{N}{S^*}}\left\{\left(\begin{array}{c} \overset{n}{\underset{i=1}{T^*}}\left\{\begin{array}{c}\mu_{A_i^k}(\bar{x}_i),\\ w_{i,k}^{\tau}\end{array}\right\}\cdot p_{k,r},\\ w_k^{\text{agr}}\end{array}\right)\right\}\right\}} \tag{7.20}$$

Figure 7.2 shows the structure of the neuro-fuzzy system described by formula (7.20).

Bounded product relation

$$I(a,b) = \max\{0, a+b-1\} \tag{7.21}$$

Putting (7.21) directly into (7.16), we obtain a formula describing a neuro-fuzzy inference system with the bounded product inference

$$\bar{y} = \frac{\sum_{r=1}^{N} \bar{y}^r \cdot S^*\left\{\begin{array}{c}\left(\overset{n}{\underset{i=1}{T^*}}\left\{\mu_{A_i^r}(\bar{x}_i), w_{i,r}^{\tau}\right\}, w_r^{\text{agr}}\right),\\ \underset{\substack{k=1\\k\neq r}}{\overset{N}{S^*}}\left\{\max\left\{\overset{n}{\underset{i=1}{T^*}}\left\{\mu_{A_i^k}(\bar{x}_i), w_{i,k}^{\tau}\right\}+p_{k,r}-1,0\right\}, w_k^{\text{agr}}\right\}\end{array}\right\}}{\sum_{r=1}^{N} S^*\left\{\begin{array}{c}\left(\overset{n}{\underset{i=1}{T^*}}\left\{\mu_{A_i^r}(\bar{x}_i), w_{i,r}^{\tau}\right\}, w_r^{\text{agr}}\right),\\ \underset{\substack{k=1\\k\neq r}}{\overset{N}{S^*}}\left\{\max\left\{\overset{n}{\underset{i=1}{T^*}}\left\{\mu_{A_i^k}(\bar{x}_i), w_{i,k}^{\tau}\right\}+p_{k,r}-1,0\right\}, w_k^{\text{agr}}\right\}\end{array}\right\}} \tag{7.22}$$

Figure 7.3 shows the structure of the neuro-fuzzy system described by formula (7.22).

Simplified Mamdani-type structures

The structures presented so far are quite complex and complicated. One possibility of their simplification is to assume, that membership functions of consequent fuzzy sets are sufficiently distant from each other, so that the following assumption holds

$$\mu_{B^k}(\bar{y}^r) \approx 0 \tag{7.23}$$

for $k, r = 1...N$ and $k \neq r$. Figure 7.4 shows examples of such fuzzy sets.

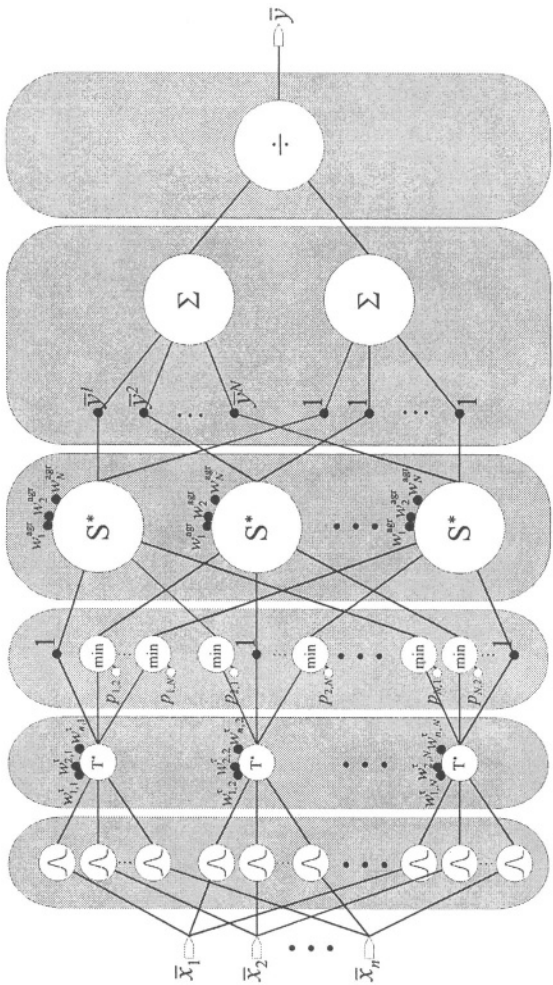

Fig. 7.1. Structure of the weighted neuro-fuzzy system with the Mamdani inference

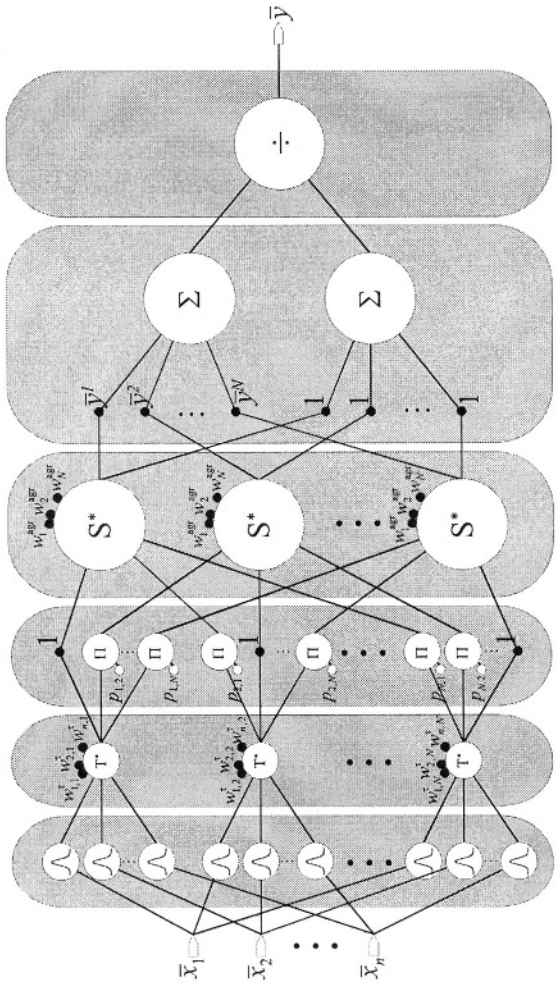

Fig. 7.2. Structure of the weighted neuro-fuzzy system with the Larsen inference

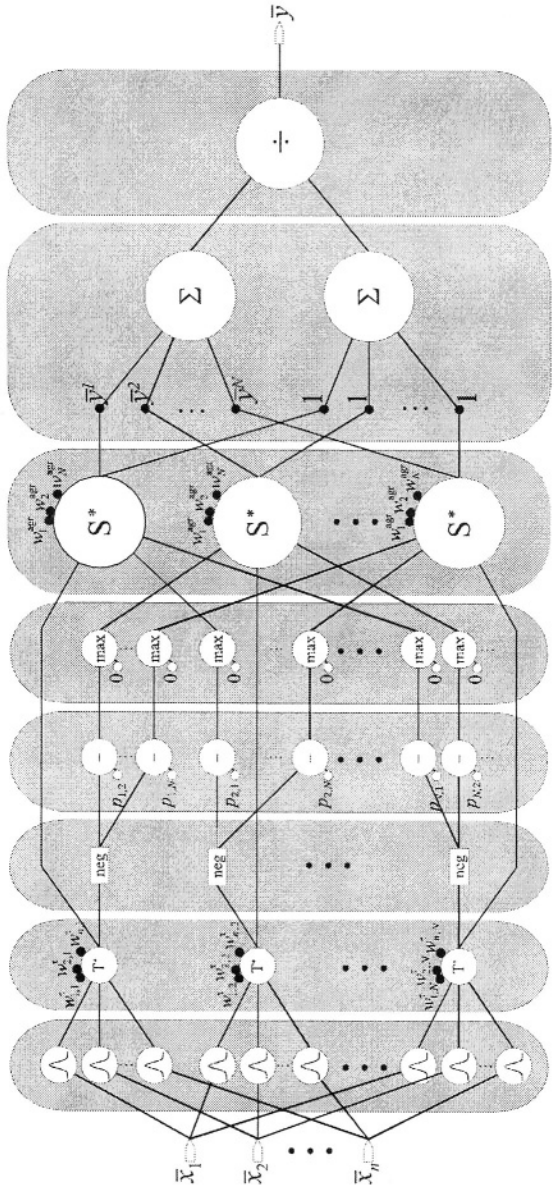

Fig. 7.3. Structure of the weighted neuro-fuzzy system with the bounded product inference

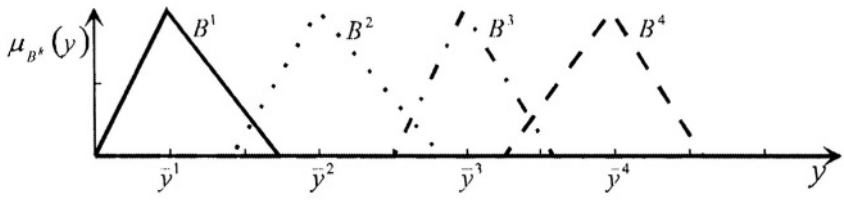

Fig. 7.4. Exemplary fuzzy sets satisfying assumption (7.23)

Combining assumption (7.23) with formula (7.13), we obtain the description of simplified neuro-fuzzy systems based on "engineering implications"

$$\bar{y} = \frac{\sum_{r=1}^{N} \bar{y}^r \cdot T^*_{i=1}\left\{\mu_{A_i^r}(x_i), w_{i,r}^{\tau}\right\} \cdot w_r^{\text{agr}}}{\sum_{r=1}^{N} T^*_{i=1}\left\{\mu_{A_i^r}(x_i), w_{i,r}^{\tau}\right\} \cdot w_r^{\text{agr}}} \tag{7.24}$$

Figure 7.5 shows the structure of a simplified neuro-fuzzy system described by formula (7.24).

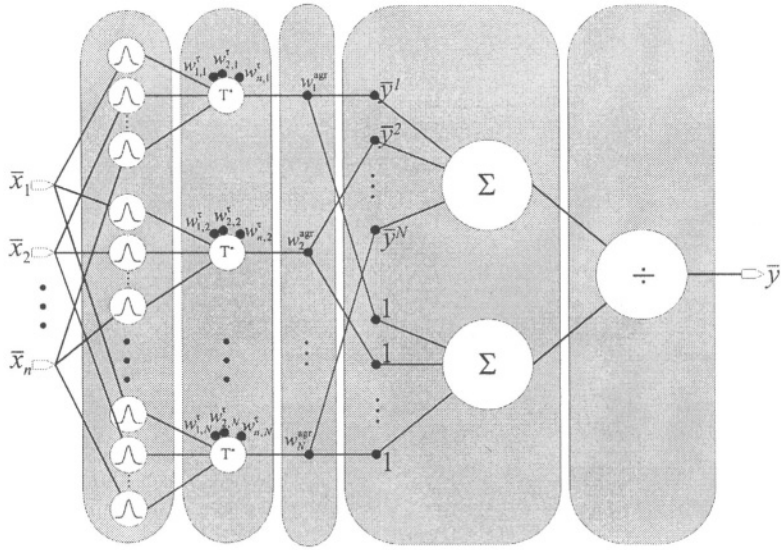

Fig. 7.5. Simplified structure of the weighted neuro-fuzzy system with a t-norm inference ("engineering implication")

7.4. SIMULATION RESULTS

In this section we present four simulations of Mamdani-type neuro-fuzzy systems. We use benchmarks described in Sections 3.6. Each of the four simulations is designed in the same fashion:

(i) In the first experiment, based on the input-output data, we learn the parameters of the membership functions.

(ii) In the second experiment, we learn the parameters of the membership functions and soft parameters $\alpha^r \in [0,1]$, $\alpha^l \in [0,1]$, $\alpha^{agr} \in [0,1]$.

(iii) In the third experiment, we learn the parameters of the membership functions and, moreover, the weights $w_{i,k}^r \in [0,1]$, $i = 1,\ldots,n$, $k = 1,\ldots,N$, in the antecedents of rules and weights $w_k^{agr} \in [0,1]$, $k = 1,\ldots,N$, of the aggregation operator of rules. In all diagrams (weights representation) we separate $w_{i,k}^r \in [0,1]$, $i = 1,\ldots,n$, $k = 1,...,N$, from $w_k^{agr} \in [0,1]$, $k = 1,..., N$, by a vertical dashed line.

(iv) In the fourth experiment, we learn the parameters of the membership functions, soft parameters $\alpha^r \in [0,1]$, $\alpha^l \in [0,1]$, $\alpha^{agr} \in [0,1]$, and the weights $w_{i,k}^r \in [0,1]$, $i = 1,\ldots,n$, $k = 1,\ldots,N$, in the antecedents of rules and weights $w_k^{agr} \in [0,1]$, $k = 1,\ldots,N$, of the aggregation operator of rules.

The parameters and weights in experiments (i)-(iv) are determined by gradient procedures presented in Section 5.8 (setting $\nu = 0$) or in Section 6.7 (setting $\lambda = 0$). In each simulation we apply two "engineering implications" (t-norms) to connect antecedents and consequents: min (Mamdani) and product (Larsen). In the case of the Mamdani connective the rules are aggregated by the max t-conorm, in the case of the Larsen connective the rules are aggregated by the algebraic t-conorm.

Box and Jenkins Gas Furnace problem

The experimental results for the Box and Jenkins Gas Furnace problem are depicted in Table 7.1. The final values (after learning) of weights $w_{i,k}^{\tau} \in [0,1]$ and $w_k^{agr} \in [0,1]$, $i = 1,...,6$, $k = 1,...,4$, are shown in Fig. 7.6. Assuming min inference, in Fig. 7.6-a.1 and 7.6-a.2 we present the results of experiments (iii) and (iv) in Table 7.1, respectively. Analogous results for the product inference are given in Fig. 7.6-b.1 and 7.6-b.2.

Table 7.1 Experimental results

MAMDANI-TYPE NFIS (BOX AND JENKINS GAS FURNACE PROBLEM)						
Experiment number	Name of flexibility parameter	Initial values	Final values after learning		RMSE (learning sequence)	
			Min inference	Algebraic inference	Min inference	Algebraic inference
i	-	-	-	-	0.4982	0.4912
ii	α^{τ}	1	0.9728	0.9834	0.4025	0.3989
	α^{j}	1	0.9997	0.9618		
	α^{agr}	1	0.9990	0.9925		
iii	$\mathbf{w}^{\tau}$	**1**	Fig. 7.6-a	Fig. 7.6-b	0.3422	0.3126
	$\mathbf{w}^{agr}$	**1**	Fig. 7.6-a	Fig. 7.6-b		
iv	α^{τ}	1	0.9294	0.9582	0.2533	0.2501
	α^{j}	1	0.9336	0.9555		
	α^{agr}	1	0.9508	0.9723		
	$\mathbf{w}^{\tau}$	**1**	Fig. 7.6-a	Fig. 7.6-b		
	$\mathbf{w}^{agr}$	**1**	Fig. 7.6-a	Fig. 7.6-b		

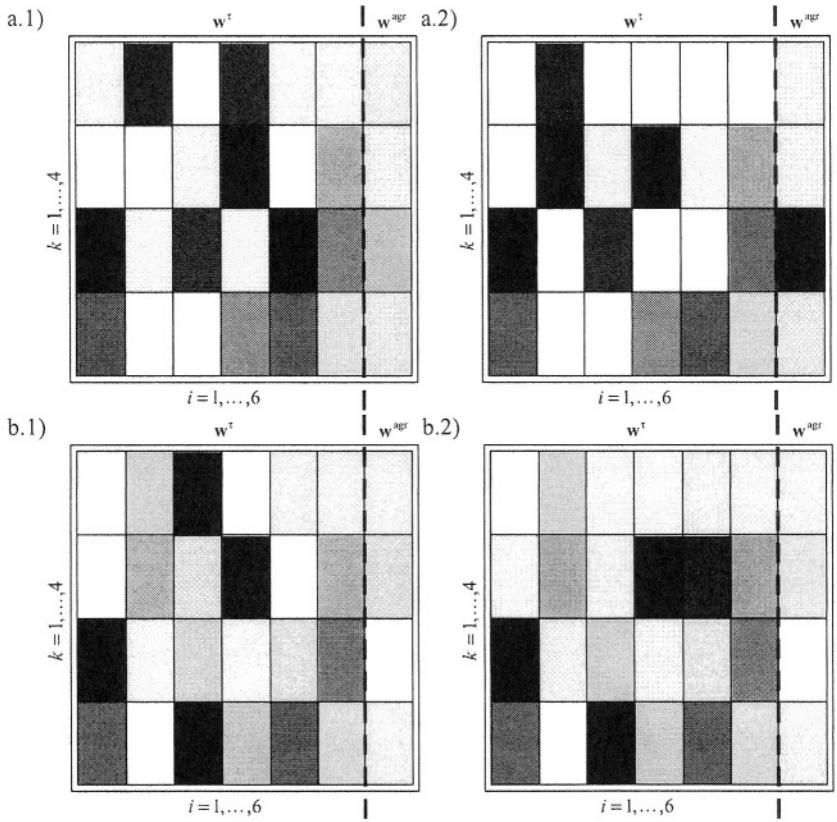

Fig. 7.6. Weights representation in the Modeling of Box and Jenkins Gas Furnace problem for Mamdani-type NFIS and a) min inference, b) product inference

Glass Identification problem

The experimental results for the Glass Identification problem are depicted in Table 7.2. The final values (after learning) of weights $w_{i,k}^{r} \in [0,1]$ and $w_k^{agr} \in [0,1]$, $i = 1,...,9$, $k = 1,...,2$, are shown in Fig. 7.7. Assuming min inference, in Fig. 7.7-a.1 and 7.7-a.2 we present the results of experiments (iii) and (iv) in Table 7.2, respectively. Analogous results for the product inference are given in Fig. 7.7-b.1 and 7.7-b.2.

Table 7.2 Experimental results

MAMDANI-TYPE NFIS (GLASS IDENTIFICATION PROBLEM)								
Experiment number	Name of flexibility parameter	Initial values	Final values after learning		Mistakes [%] (learning sequence)		Mistakes [%] (testing sequence)	
			Min inference	Algebraic inference	Min inference	Algebraic inference	Min inference	Algebraic inference
i	-	-	-	-	4.00	4.67	3.13	3.13
ii	α^r	1	0.0713	0.0943	3.33	4.00	3.13	3.13
	α^l	1	0.9552	0.9355				
	α^{agr}	1	0.0342	0.9522				
iii	$\mathbf{w}^r$	**1**	Fig. 7.7-a	Fig. 7.7-b	3.33	3.33	3.13	3.13
	$\mathbf{w}^{agr}$	**1**	Fig. 7.7-a	Fig. 7.7-b				
iv	α^r	1	0.0748	0.0427	3.33	3.33	1.56	1.56
	α^l	1	0.8636	0.9517				
	α^{agr}	1	0.9294	0.9702				
	$\mathbf{w}^r$	**1**	Fig. 7.7-a	Fig. 7.7-b				
	$\mathbf{w}^{agr}$	**1**	Fig. 7.7-a	Fig. 7.7-b				

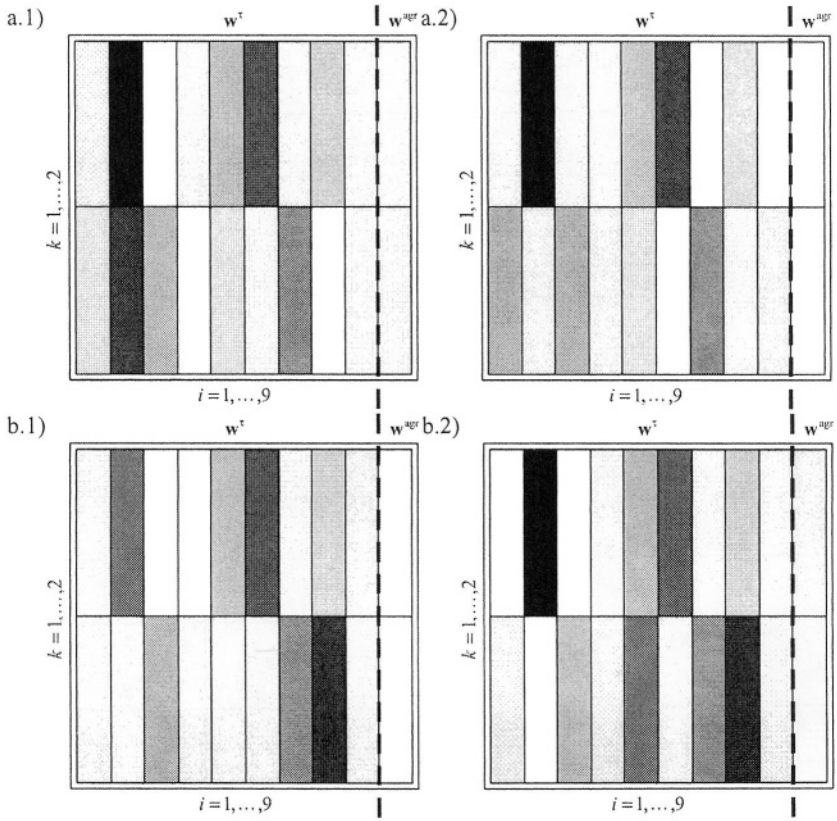

Fig. 7.7. Weights representation in the Glass Identification problem for Mamdani-type NFIS and a) min inference, b) product inference

Modeling of Static Nonlinear Function (HANG) problem

The experimental results for the Modeling of the Static Nonlinear Function problem are shown in Table 7.3. The final values (after learning) of weights $w_{i,k}^{r} \in [0,1]$ and $w_{k}^{agr} \in [0,1]$, $i = 1,\ldots,2$, $k = 1,\ldots,5$, are depicted in Fig. 7.8. Assuming min inference, in Fig. 7.8-a.1 and 7.8-a.2 we present the results of experiments (iii) and (iv) in Table 7.3, respectively. Analogous results for the product inference are given in Fig. 7.8-b.1 and 7.8-b.2.

Table 7.3 Experimental results

MAMDANI-TYPE NFIS (MODELING OF STATIC NONLINEAR FUNCTION PROBLEM)						
Experiment number	Name of flexibility parameter	Initial values	Final values after learning		RMSE (learning sequence)	
			Min inference	Algebraic inference	Min inference	Algebraic inference
i	-	-	-	-	0.1258	0.1098
ii	α^{r}	1	0.0316	0.9467	0.1205	0.6905
	α^{l}	1	0.9737	0.9683		
	α^{agr}	1	0.9724	0.9955		
iii	$\mathbf{w}^{r}$	**1**	Fig. 7.8-a	Fig. 7.8-b	0.1029	0.0671
	$\mathbf{w}^{agr}$	**1**	Fig. 7.8-a	Fig. 7.8-b		
iv	α^{r}	1	0.0345	0.9764	0.0964	0.0529
	α^{l}	1	0.9839	0.9292		
	α^{agr}	1	0.8885	0.9230		
	$\mathbf{w}^{r}$	**1**	Fig. 7.8-a	Fig. 7.8-b		
	$\mathbf{w}^{agr}$	**1**	Fig. 7.8-a	Fig. 7.8-b		

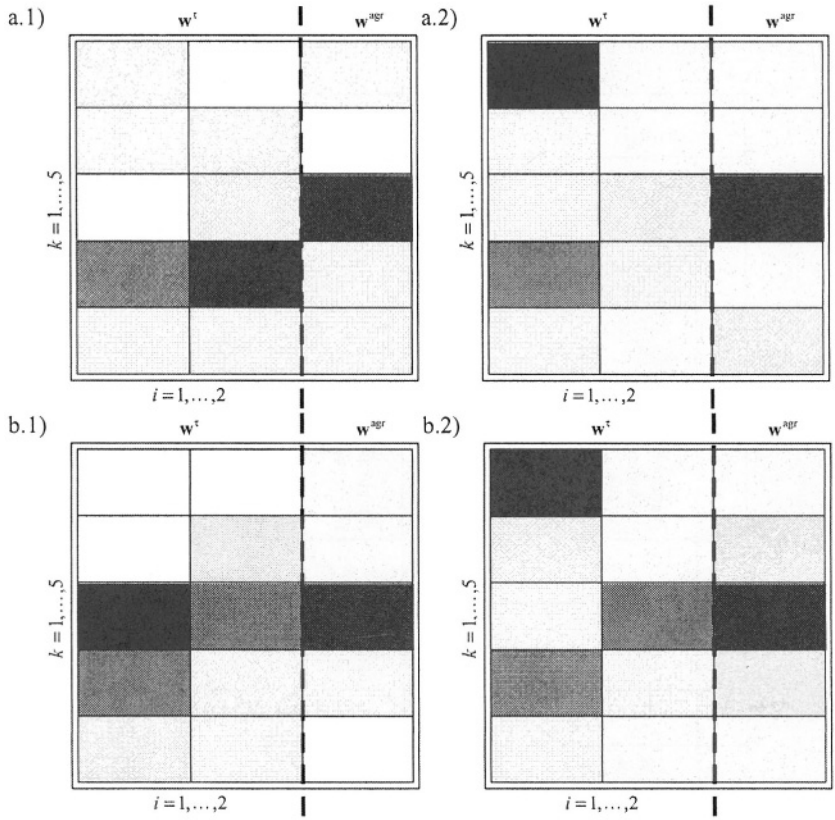

Fig. 7.8. Weights representation in the HANG problem for Mamdani-type NFIS and a) min inference, b) product inference

Wisconsin Breast Cancer problem

The experimental results for the Wisconsin Breast Cancer problem are shown in Table 7.4. The final values (after learning) of weights $w_{i,k}^{r} \in [0,1]$ and $w_k^{agr} \in [0,1]$, $i = 1,...,9$, $k = 1,...,2$, are depicted in Fig. 7.9. Assuming min inference, in Fig. 7.9-a.1 and 7.9-a.2 we present the results of experiments (iii) and (iv) in Table 7.4, respectively. Analogous results for the product inference are given in Fig. 7.9-b.1 and 7.9-b.2.

Table 7.4 Experimental results

MAMDANI-TYPE NFIS (WISCONSIN BREAST CANCER PROBLEM)								
Experiment number	Name of flexibility parameter	Initial values	Final values after learning		Mistakes [%] (learning sequence)		Mistakes [%] (testing sequence)	
			Min inference	Algebraic inference	Min inference	Algebraic inference	Min inference	Algebraic inference
i	-	-	-	-	3.14	2.51	2.44	2.44
ii	α^r	1	0.1714	0.1354	3.14	2.51	1.95	1.95
	α^l	1	0.9888	0.8677				
	α^{agr}	1	0.1867	0.8903				
iii	$\mathbf{w}^r$	**1**	Fig. 7.9-a	Fig. 7.9-b	3.14	2.30	1.46	1.46
	$\mathbf{w}^{agr}$	**1**	Fig. 7.9-a	Fig. 7.9-b				
iv	α^r	1	0.9185	0.8447	2.93	2.30	1.46	1.46
	α^l	1	0.9140	0.9238				
	α^{agr}	1	0.8346	0.8737				
	$\mathbf{w}^r$	**1**	Fig. 7.9-a	Fig. 7.9-b				
	$\mathbf{w}^{agr}$	**1**	Fig. 7.9-a	Fig. 7.9-b				

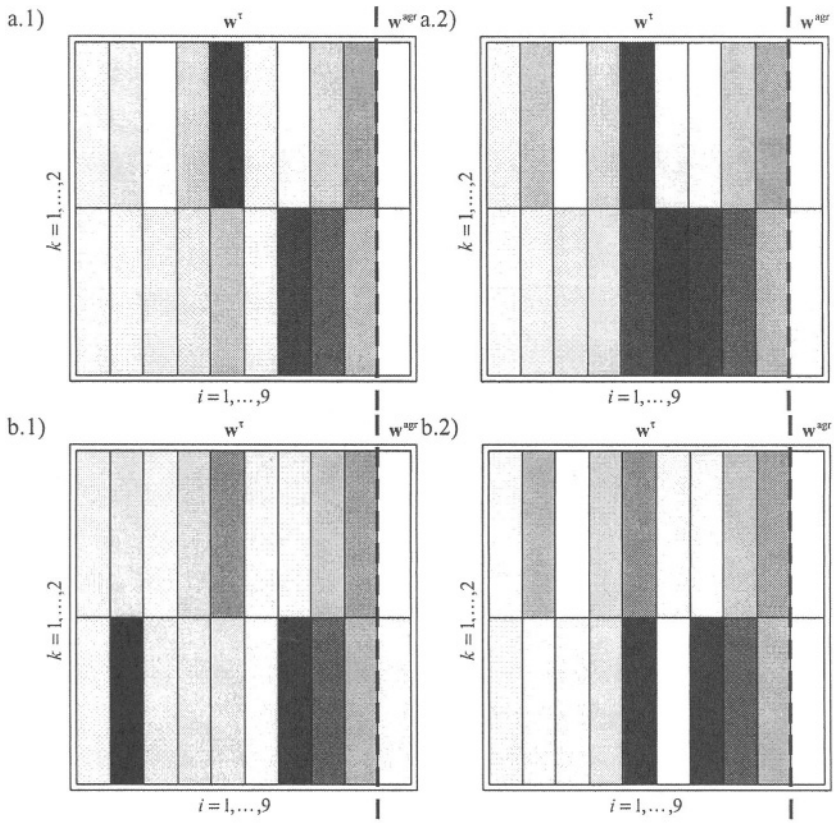

Fig. 7.9. Weights representation in the Wisconsin Breast Cancer problem for Mamdani-type NFIS and a) min inference, b) product inference

7.5. SUMMARY AND DISCUSSION

In this chapter we derived various weighted neuro-fuzzy structures based on the Mamdani inference. We have also obtained a simplified neuro-fuzzy structure assuming that membership functions of consequent fuzzy sets are sufficiently distant from each other. From simulations in Section 7.4 it follows that the incorporation of weights leads to a better performance of fuzzy systems than the incorporation of soft parameters. Their combination, i.e. weights+soft parameters, gives the best performance of Mamdani-type neuro-fuzzy systems.

7.6. PROBLEMS

Problem 7.1. Replace defuzzifier (3.23) by

$$\bar{y} = \frac{\sum_{r=1}^{N} \bar{y}^r \cdot \mu_{\bar{B}^r}(\bar{y}^r)}{\sum_{r=1}^{N} \mu_{\bar{B}^r}(\bar{y}^r)} \tag{7.25}$$

and derive the Mamdani-type system based on the Larsen relation.

Problem 7.2. Generalize solution of Problem 7.1 using a t-norm as an "engineering implication".

Problem 7.3. Show that defuzzifiers (3.23) and (7.25) lead to the same description of the Mamdani-type system provided that assumption (7.23) holds.

Problem 7.4. Incorporate weights to a fuzzy system derived in Problem 7.2 and compare the result with formula (7.24).

Problem 7.5. Incorporate softness to a fuzzy system derived in Problem 7.2.

Problem 7.6. Prove that under certain conditions fuzzy systems described by formulas (3.23), (3.49) and (7.25) are functionally equivalent.

This page intentionally left blank

Chapter 8

FLEXIBLE LOGICAL-TYPE NEURO-FUZZY SYSTEMS

8.1. INTRODUCTION

Flexible logical-type neuro-fuzzy systems can be derived from the OR-type systems (Chapter 5) or the AND-type systems (Chapter 6) setting $v = 0$ or $\lambda = 0$, respectively. However, in this way we get the logical-type system described by an S-implication only. In this chapter we do not want to restrict the design process to an S-implication. We admit other fuzzy implications listed in Table 2.1. Moreover, we will reveal the connectionist nature of logical-type systems; each structure will reflect the actual fuzzy implication used in the design process.

8.2. PROBLEM DESCRIPTION

It is easily seen that logical-type systems can be expressed in terms of an H function:

$$\tau_k(\bar{\mathbf{x}}) = H\left(\mu_{A_1^k}(\bar{x}_1),\ldots,\mu_{A_n^k}(\bar{x}_n)\ ;0\right) \tag{8.1}$$

$$I_{k,r}(\bar{\mathbf{x}},\bar{y}^r) = H\left(\tilde{N}_0(\tau_k(\bar{\mathbf{x}})),\mu_{B^k}(\bar{y}^r)\ ;1\right) \tag{8.2}$$

$$\mathrm{agr}_r(\bar{\mathbf{x}},\bar{y}^r) = H\left(I_{1,r}(\bar{\mathbf{x}},\bar{y}^r),\ldots,I_{N,r}(\bar{\mathbf{x}},\bar{y}^r)\ ;0\right) \tag{8.3}$$

The above description is applicable to systems with an S-implication only, and it does not reflect the functional form of the fuzzy implication used for the antecedent-consequent connection. Therefore our problem is to design

a neuro-fuzzy system realizing a mapping $f: \mathbf{X} \to \mathbf{Y}$ such that the fuzzy inference is described by the fuzzy implication and the structure of the system explicitly depends on the functional form of that fuzzy implication.

Moreover, we will incorporate certainty weights to the aggregation of rules and to the connectives of antecedents. In simulations we also study logical-type systems with soft triangular norms.

8.3. NEURO-FUZZY STRUCTURES

In the logical-type model, the aggregation is carried out according to formulas (3.20) and (3.21). The weighted aggregation is given by

$$\mu_{B'}(y) = \underset{k=1}{\overset{N}{T^*}} \left\{ \mu_{\bar{B}^k}(y), w_k^{\text{agr}} \right\} \tag{8.4}$$

where w_k^{agr} is the weight of the k-th rule.

Similarly to the Mamdani approach, to design a neuro-fuzzy system with a weighted aggregation we use the defuzzification method described by (3.23). Combining (8.4) and (7.6) with (3.14) and (3.23), we obtain the following description of the system

$$\bar{y} = \frac{\sum_{r=1}^{N} \bar{y}^r \cdot \underset{k=1}{\overset{N}{T^*}} \left\{ I\left(\underset{i=1}{\overset{n}{T^*}} \left\{ \mu_{A_i^k}(\bar{x}_i), w_{i,k}^r \right\}, \mu_{B^k}(\bar{y}^r) \right), w_k^{\text{agr}} \right\}}{\sum_{r=1}^{N} \underset{k=1}{\overset{N}{T^*}} \left\{ I\left(\underset{i=1}{\overset{n}{T^*}} \left\{ \mu_{A_i^k}(\bar{x}_i), w_{i,k}^r \right\}, \mu_{B^k}(\bar{y}^r) \right), w_k^{\text{agr}} \right\}} \tag{8.5}$$

The first group of implications used in the logical approach belongs to the so-called S-implications given by

$$I(a,b) = S\{1-a, b\} \tag{8.6}$$

where S is a t-conorm. The binary, Łukasiewicz, Reichenbach and Fodor implications (see Table 2.1) belong to this group. Assuming that fuzzy sets B^k are normal, see condition (7.10), and using the boundary condition of triangular norms $S\{a,1\} = 1$, from (8.6) and (3.14) we obtain

$$\mu_{\bar{B}^k}(\bar{y}^k) = 1 \tag{8.7}$$

Hence, formula (8.4) for $y = \bar{y}^r$ takes the following form

$$\begin{aligned}\mu_{B^r}(\bar{y}^r) &= T^*\left\{(1, w_r^{\text{agr}}), \underset{\substack{k=1\\k\neq r}}{\overset{N}{T^*}}\left\{I\left(\mu_{A^k}(\bar{\mathbf{x}}), \mu_{B^k}(\bar{y}^r)\right), w_k^{\text{agr}}\right\}\right\} \\ &= \underset{\substack{k=1\\k\neq r}}{\overset{N}{T^*}}\left\{I\left(\mu_{A^k}(\bar{\mathbf{x}}), \mu_{B^k}(\bar{y}^r)\right), w_k^{\text{agr}}\right\}\end{aligned} \tag{8.8}$$

Consequently, formula (8.5) becomes

$$\bar{y} = \frac{\sum_{r=1}^{N} \bar{y}^r \cdot \underset{\substack{k=1\\k\neq r}}{\overset{N}{T^*}}\left\{I\left(\underset{i=1}{\overset{n}{T^*}}\left\{\mu_{A_i^k}(\bar{x}_i), w_{i,k}^{\tau}\right\}, \mu_{B^k}(\bar{y}^r)\right), w_k^{\text{agr}}\right\}}{\sum_{r=1}^{N} \underset{\substack{k=1\\k\neq r}}{\overset{N}{T^*}}\left\{I\left(\underset{i=1}{\overset{n}{T^*}}\left\{\mu_{A_i^k}(\bar{x}_i), w_{i,k}^{\tau}\right\}, \mu_{B^k}(\bar{y}^r)\right), w_k^{\text{agr}}\right\}} \tag{8.9}$$

or

$$\bar{y} = \frac{\sum_{r=1}^{N} \bar{y}^r \cdot \underset{\substack{k=1\\k\neq r}}{\overset{N}{T^*}}\left\{I\left(\underset{i=1}{\overset{n}{T^*}}\left\{\mu_{A_i^k}(\bar{x}_i), w_{i,k}^{\tau}\right\}, p_{k,r}\right), w_k^{\text{agr}}\right\}}{\sum_{r=1}^{N} \underset{\substack{k=1\\k\neq r}}{\overset{N}{T^*}}\left\{I\left(\underset{i=1}{\overset{n}{T^*}}\left\{\mu_{A_i^k}(\bar{x}_i), w_{i,k}^{\tau}\right\}, p_{k,r}\right), w_k^{\text{agr}}\right\}} \tag{8.10}$$

with notation (7.14).

Binary implication

$$I(a,b) = \max\{1-a, b\} \tag{8.11}$$

Putting function (8.11) into (8.10), we obtain a formula describing the neuro-fuzzy system with the binary inference

$$\bar{y} = \frac{\sum_{r=1}^{N} \bar{y}^r \cdot \underset{\substack{k=1\\k\neq r}}{\overset{N}{T^*}}\left\{\max\left\{1-\underset{i=1}{\overset{n}{T^*}}\left\{\mu_{A_i^k}(\bar{x}_i), w_{i,k}^{\tau}\right\}, p_{k,r}\right\}, w_k^{\text{agr}}\right\}}{\sum_{r=1}^{N} \underset{\substack{k=1\\k\neq r}}{\overset{N}{T^*}}\left\{\max\left\{1-\underset{i=1}{\overset{n}{T^*}}\left\{\mu_{A_i^k}(\bar{x}_i), w_{i,k}^{\tau}\right\}, p_{k,r}\right\}, w_k^{\text{agr}}\right\}} \tag{8.12}$$

Figure 8.1 depicts the structure of the neuro-fuzzy system described by formula (8.12).

Łukasiewicz implication

$$I(a,b) = \min\{1, 1-a+b\} \tag{8.13}$$

Inserting function (8.13) directly into (8.10), we obtain a formula describing the neuro-fuzzy system with the Łukasiewicz inference

$$\bar{y} = \frac{\sum_{r=1}^{N} \bar{y}^r \cdot \underset{\substack{k=1 \\ k\neq r}}{\overset{N}{T^*}} \left\{ \min\left\{ 1, 1 - \overset{n}{\underset{i=1}{T^*}} \left\{ \mu_{A_i^k}(\bar{x}_i), w_{i,k}^{\tau} \right\} + p_{k,r} \right\}, w_k^{\text{agr}} \right\}}{\sum_{r=1}^{N} \underset{\substack{k=1 \\ k\neq r}}{\overset{N}{T^*}} \left\{ \min\left\{ 1, 1 - \overset{n}{\underset{i=1}{T^*}} \left\{ \mu_{A_i^k}(\bar{x}_i), w_{i,k}^{\tau} \right\} + p_{k,r} \right\}, w_k^{\text{agr}} \right\}} \tag{8.14}$$

Figure 8.2 depicts the structure of the neuro-fuzzy system described by formula (8.14).

Reichenbach implication

$$I(a,b) = 1 - a + a \cdot b \tag{8.15}$$

Putting function (8.15) directly into (8.10), we obtain a formula describing the neuro-fuzzy system with the probabilistic inference

$$\bar{y} = \frac{\sum_{r=1}^{N} \bar{y}^r \cdot \underset{\substack{k=1 \\ k\neq r}}{\overset{N}{T^*}} \left\{ 1 - \overset{n}{\underset{i=1}{T^*}} \left\{ \mu_{A_i^k}(\bar{x}_i), w_{i,k}^{\tau} \right\} \cdot \tilde{p}_{k,r}, w_k^{\text{agr}} \right\}}{\sum_{r=1}^{N} \underset{\substack{k=1 \\ k\neq r}}{\overset{N}{T^*}} \left\{ 1 - \overset{n}{\underset{i=1}{T^*}} \left\{ \mu_{A_i^k}(\bar{x}_i), w_{i,k}^{\tau} \right\} \cdot \tilde{p}_{k,r}, w_k^{\text{agr}} \right\}} \tag{8.16}$$

Figure 8.3 depicts the structure of the neuro-fuzzy system described by formula (8.16).

Fodor implication

$$I(a,b) = \begin{cases} 1 & \text{if} \quad a \le b \\ \max\{1-a, b\} & \text{if} \quad a > b \end{cases} \tag{8.17}$$

Inserting function (8.17) directly into (8.10), we obtain a formula describing the neuro-fuzzy system with the Fodor inference

$$\bar{y} = \frac{\sum_{r=1}^{N} \bar{y}^r \cdot \underset{\substack{k=1 \\ k\neq r}}{\overset{N}{T^*}} \left\{ \max\left\{ \begin{array}{c} 1 - \overset{n}{\underset{i=1}{T^*}} \left\{ \mu_{A_i^k}(\bar{x}_i), w_{i,k}^{\tau} \right\}, \\ p_{k,r}, \\ \overset{n}{\underset{i=1}{T^*}} \left\{ \mu_{A_i^k}(\bar{x}_i), w_{i,k}^{\tau} \right\} \overset{\le}{*} p_{k,r} \end{array} \right\}, w_k^{\text{agr}} \right\}}{\sum_{r=1}^{N} \underset{\substack{k=1 \\ k\neq r}}{\overset{N}{T^*}} \left\{ \max\left\{ \begin{array}{c} 1 - \overset{n}{\underset{i=1}{T^*}} \left\{ \mu_{A_i^k}(\bar{x}_i), w_{i,k}^{\tau} \right\}, \\ p_{k,r}, \\ \overset{n}{\underset{i=1}{T^*}} \left\{ \mu_{A_i^k}(\bar{x}_i), w_{i,k}^{\tau} \right\} \overset{\le}{*} p_{k,r} \end{array} \right\}, w_k^{\text{agr}} \right\}} \tag{8.18}$$

where

$$a \overset{\leq}{*} b = \begin{cases} 1 & \text{if} \quad a \leq b \\ 0 & \text{if} \quad a > b \end{cases} \tag{8.19}$$

Figure 8.4 depicts the structure of the neuro-fuzzy system described by formula (8.18).

We will now show that the Zadeh and Willmott implications (see Table 2.1) lead to the same general description of corresponding neuro-fuzzy systems. In this case assumption (7.10) regarding the normality of fuzzy sets B^k implies

$$\mu_{\bar{B}^k}(\bar{y}^k) = \max\left\{\mu_{A^k}(\bar{\mathbf{x}}), 1 - \mu_{A^k}(\bar{\mathbf{x}})\right\} \tag{8.20}$$

Therefore, formula (8.4) for $y = \bar{y}^r$ can be rewritten as follows

$$\mu_{B^r}(\bar{y}^r) = T^*\left\{\begin{array}{l}\left(\max\left\{\mu_{A^r}(\bar{\mathbf{x}}), 1 - \mu_{A^r}(\bar{\mathbf{x}})\right\}, w_r^{\text{agr}}\right), \\ \underset{\substack{k=1 \\ k\neq r}}{\overset{N}{T^*}}\left\{I\left(\mu_{A^k}(\bar{\mathbf{x}}), \mu_{B^k}(\bar{y}^r)\right), w_k^{\text{agr}}\right\}\end{array}\right\} \tag{8.21}$$

Combining formulas (8.21) and (7.6) with (3.23) we get the following general description of the neuro-fuzzy system based on the Zadeh and Willmott fuzzy implications

$$\bar{y} = \frac{\sum_{r=1}^{N} \bar{y}^r \cdot T^*\left\{\begin{array}{l}\left(\max\left\{\begin{array}{l}\underset{i=1}{\overset{n}{T^*}}\left\{\mu_{A_i^r}(\bar{x}_i), w_{i,r}^{\tau}\right\} \\ 1 - \underset{i=1}{\overset{n}{T^*}}\left\{\mu_{A_i^r}(\bar{x}_i), w_{i,r}^{\tau}\right\}\end{array}\right\}, w_r^{\text{agr}}\right), \\ \underset{\substack{k=1 \\ k\neq r}}{\overset{N}{T^*}}\left\{I\left(\underset{i=1}{\overset{n}{T^*}}\left\{\mu_{A_i^k}(\bar{x}_i), w_{i,k}^{\tau}\right\}, \mu_{B^k}(\bar{y}^r)\right), w_k^{\text{agr}}\right\}\end{array}\right\}}{\sum_{r=1}^{N} T^*\left\{\begin{array}{l}\left(\max\left\{\begin{array}{l}\underset{i=1}{\overset{n}{T^*}}\left\{\mu_{A_i^r}(\bar{x}_i), w_{i,r}^{\tau}\right\} \\ 1 - \underset{i=1}{\overset{n}{T^*}}\left\{\mu_{A_i^r}(\bar{x}_i), w_{i,r}^{\tau}\right\}\end{array}\right\}, w_r^{\text{agr}}\right), \\ \underset{\substack{k=1 \\ k\neq r}}{\overset{N}{T^*}}\left\{I\left(\underset{i=1}{\overset{n}{T^*}}\left\{\mu_{A_i^k}(\bar{x}_i), w_{i,k}^{\tau}\right\}, \mu_{B^k}(\bar{y}^r)\right), w_k^{\text{agr}}\right\}\end{array}\right\}} \tag{8.22}$$

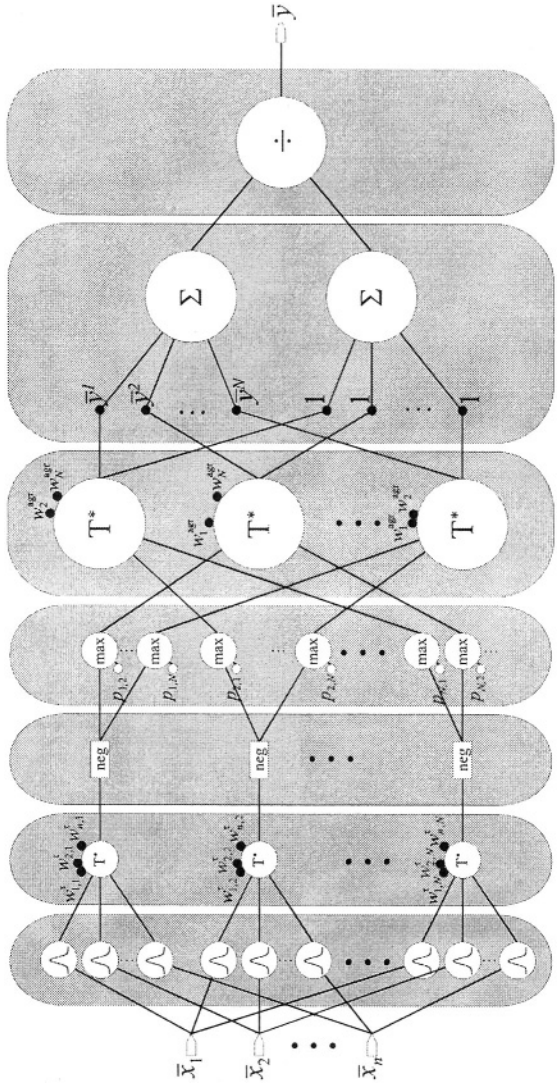

Fig. 8.1. Structure of the weighted neuro-fuzzy system with the binary implication

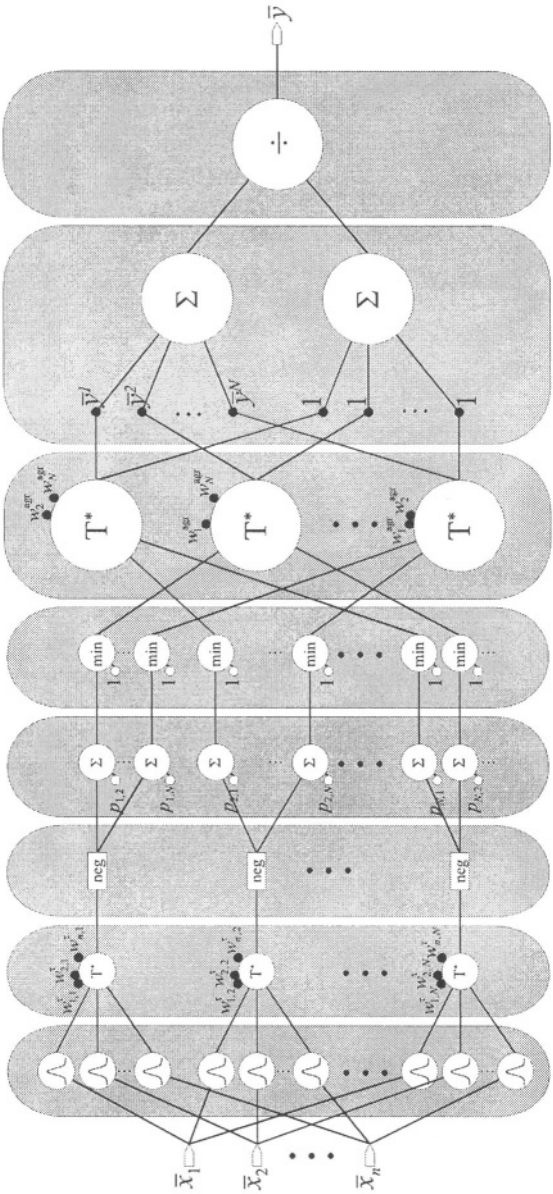

Fig. 8.2. Structure of the weighted neuro-fuzzy system with the Łukasiewicz implication

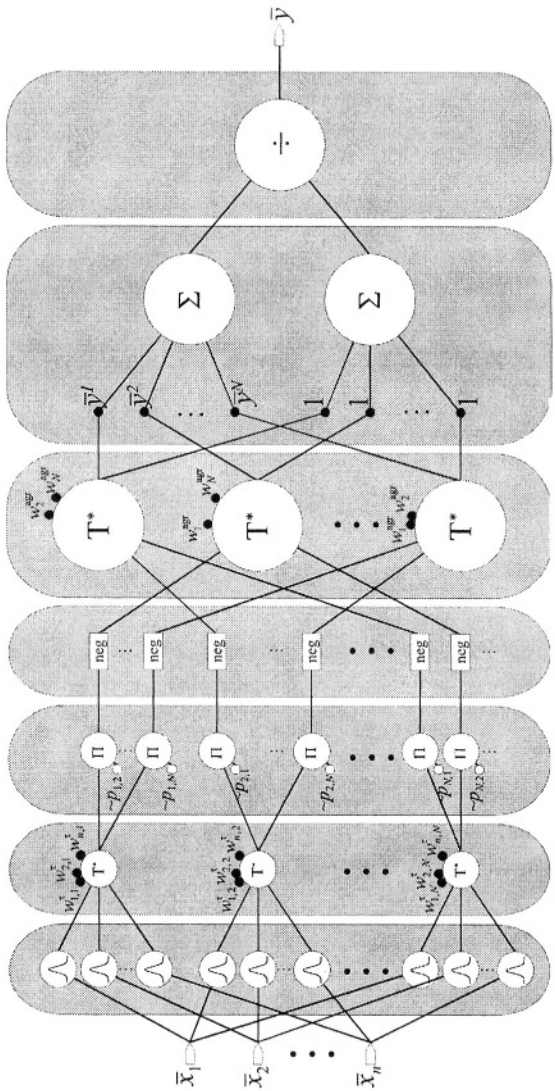

Fig. 8.3. Structure of the weighted neuro-fuzzy system with the Reichenbach implication

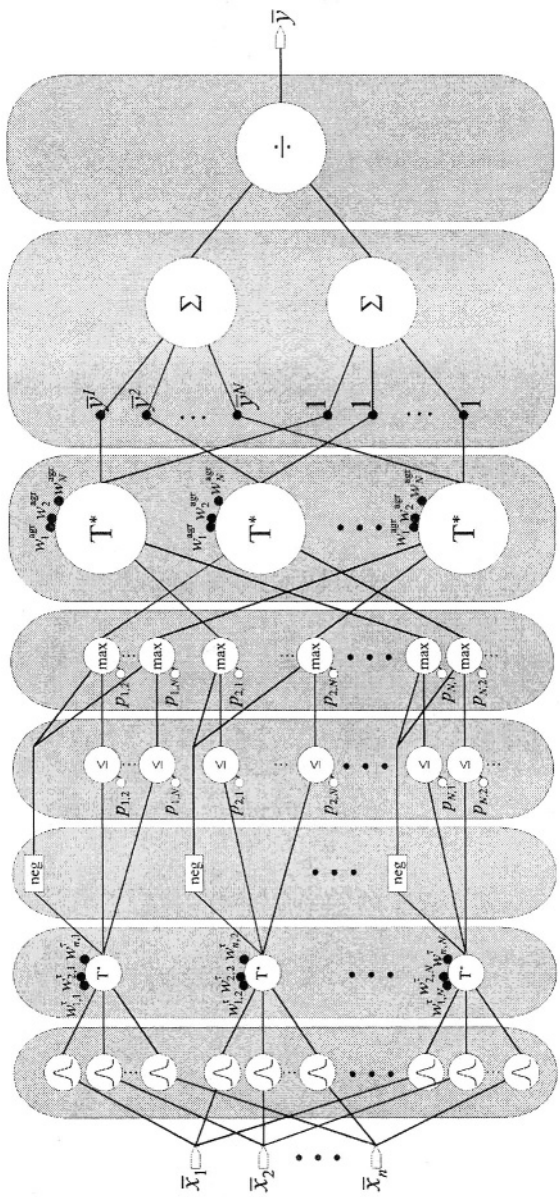

Fig. 8.4. Structure of the weighted neuro-fuzzy system with the Fodor implication

Using notation (7.14), formula (8.22) can be rewritten in the form

$$\bar{y} = \frac{\sum_{r=1}^{N} \bar{y}^r \cdot T^*\left\{\left(\max\left\{\begin{array}{l} T_{i=1}^{n*}\left\{\mu_{A_i^r}(\bar{x}_i), w_{i,r}^{\tau}\right\} \\ 1 - T_{i=1}^{n*}\left\{\mu_{A_i^r}(\bar{x}_i), w_{i,r}^{\tau}\right\} \end{array}\right\}, w_r^{\text{agr}}\right), \underset{\substack{k=1 \\ k\neq r}}{T^{N*}}\left\{I\left(T_{i=1}^{n*}\left\{\mu_{A_i^k}(\bar{x}_i), w_{i,k}^{\tau}\right\}, p_{k,r}\right), w_k^{\text{agr}}\right\}\right\}}{\sum_{r=1}^{N} T^*\left\{\left(\max\left\{\begin{array}{l} T_{i=1}^{n*}\left\{\mu_{A_i^r}(\bar{x}_i), w_{i,r}^{\tau}\right\} \\ 1 - T_{i=1}^{n*}\left\{\mu_{A_i^r}(\bar{x}_i), w_{i,r}^{\tau}\right\} \end{array}\right\}, w_r^{\text{agr}}\right), \underset{\substack{k=1 \\ k\neq r}}{T^{N*}}\left\{I\left(T_{i=1}^{n*}\left\{\mu_{A_i^k}(\bar{x}_i), w_{i,k}^{\tau}\right\}, p_{k,r}\right), w_k^{\text{agr}}\right\}\right\}} \tag{8.23}$$

A concrete form of (8.23) depends on the functional form of fuzzy implication $I(a,b)$. We will consider two cases.

Zadeh implication

$$I(a,b) = \max\{\min\{a,b\}, 1-a\} \tag{8.24}$$

Putting function (8.24) directly into (8.23), we obtain a formula describing the neuro-fuzzy system with min inference

$$\bar{y} = \frac{\sum_{r=1}^{N} \bar{y}^r \cdot T^*\left\{\left(\max\left\{\begin{array}{l} T_{i=1}^{n*}\left\{\mu_{A_i^r}(\bar{x}_i), w_{i,r}^{\tau}\right\} \\ 1 - T_{i=1}^{n*}\left\{\mu_{A_i^r}(\bar{x}_i), w_{i,r}^{\tau}\right\} \end{array}\right\}, w_r^{\text{agr}}\right), \underset{\substack{k=1 \\ k\neq r}}{T^{N*}}\left\{\max\left\{\begin{array}{l} \min\left\{T_{i=1}^{n*}\left\{\mu_{A_i^k}(\bar{x}_i), w_{i,k}^{\tau}\right\}, p_{k,r}\right\}, \\ 1 - T_{i=1}^{n*}\left\{\mu_{A_i^k}(\bar{x}_i), w_{i,k}^{\tau}\right\} \end{array}\right\}, w_k^{\text{agr}}\right\}\right\}}{\sum_{r=1}^{N} T^*\left\{\left(\max\left\{\begin{array}{l} T_{i=1}^{n*}\left\{\mu_{A_i^r}(\bar{x}_i), w_{i,r}^{\tau}\right\} \\ 1 - T_{i=1}^{n*}\left\{\mu_{A_i^r}(\bar{x}_i), w_{i,r}^{\tau}\right\} \end{array}\right\}, w_r^{\text{agr}}\right), \underset{\substack{k=1 \\ k\neq r}}{T^{N*}}\left\{\max\left\{\begin{array}{l} \min\left\{T_{i=1}^{n*}\left\{\mu_{A_i^k}(\bar{x}_i), w_{i,k}^{\tau}\right\}, p_{k,r}\right\}, \\ 1 - T_{i=1}^{n*}\left\{\mu_{A_i^k}(\bar{x}_i), w_{i,k}^{\tau}\right\} \end{array}\right\}, w_k^{\text{agr}}\right\}\right\}} \tag{8.25}$$

Figure 8.5 depicts the structure of the neuro-fuzzy system described by formula (8.25).

Willmott implication

$$I(a,b)=\min\{\max\{1-a,b\},\max\{a,1-b,\min\{b,1-a\}\}\} \tag{8.26}$$

Inserting function (8.26) directly into (8.23), we obtain a formula describing the neuro-fuzzy system with the Willmott inference

$$\bar{y}=\frac{\sum_{r=1}^{N}\bar{y}^r\cdot T^*\left\{\underset{\substack{k=1\\k\neq r}}{\overset{N}{T^*}}\left\{\min\left\{\begin{array}{l}\max\left\{1-\underset{i=1}{\overset{n}{T^*}}\left\{\mu_{A_i^k}(\bar{x}_i),w_{i,k}^{\tau}\right\},p_{k,r}\right\},\\ \max\left\{\begin{array}{l}\underset{i=1}{\overset{n}{T^*}}\left\{\mu_{A_i^k}(\bar{x}_i),w_{i,k}^{\tau}\right\},\tilde{p}_{k,r},\\ \min\left\{\begin{array}{l}p_{k,r},\\ 1-\underset{i=1}{\overset{n}{T^*}}\left\{\mu_{A_i^k}(\bar{x}_i),w_{i,k}^{\tau}\right\}\end{array}\right\}\end{array}\right\}\end{array}\right\},w_k^{\text{agr}}\right\}\right\},\left(\max\left\{\mu_{A^r}(\bar{\mathbf{x}}),1-\underset{i=1}{\overset{n}{T^*}}\left\{\mu_{A_i^r}(\bar{x}_i),w_{i,r}^{\tau}\right\}\right\},w_r^{\text{agr}}\right)\right\}}{\sum_{r=1}^{N}T^*\left\{\underset{\substack{k=1\\k\neq r}}{\overset{N}{T^*}}\left\{\min\left\{\begin{array}{l}\max\left\{1-\underset{i=1}{\overset{n}{T^*}}\left\{\mu_{A_i^k}(\bar{x}_i),w_{i,k}^{\tau}\right\},p_{k,r}\right\},\\ \max\left\{\begin{array}{l}\underset{i=1}{\overset{n}{T^*}}\left\{\mu_{A_i^k}(\bar{x}_i),w_{i,k}^{\tau}\right\},\tilde{p}_{k,r},\\ \min\left\{\begin{array}{l}p_{k,r},\\ 1-\underset{i=1}{\overset{n}{T^*}}\left\{\mu_{A_i^k}(\bar{x}_i),w_{i,k}^{\tau}\right\}\end{array}\right\}\end{array}\right\}\end{array}\right\},w_k^{\text{agr}}\right\}\right\},\left(\max\left\{\mu_{A^r}(\bar{\mathbf{x}}),1-\underset{i=1}{\overset{n}{T^*}}\left\{\mu_{A_i^r}(\bar{x}_i),w_{i,r}^{\tau}\right\}\right\},w_r^{\text{agr}}\right)\right\}} \tag{8.27}$$

Figure 8.6 depicts the structure of the neuro-fuzzy system described by formula (8.27).

We will now consider neuro-fuzzy systems based on the Goguen, Rescher, Gödel and Yager fuzzy implications given in Table 2.1. It is easily seen that for these implications equation (8.7) is true and consequently corresponding to them neuro-fuzzy systems are described by formula (8.10). A concrete form of that formula depends on the fuzzy implication.

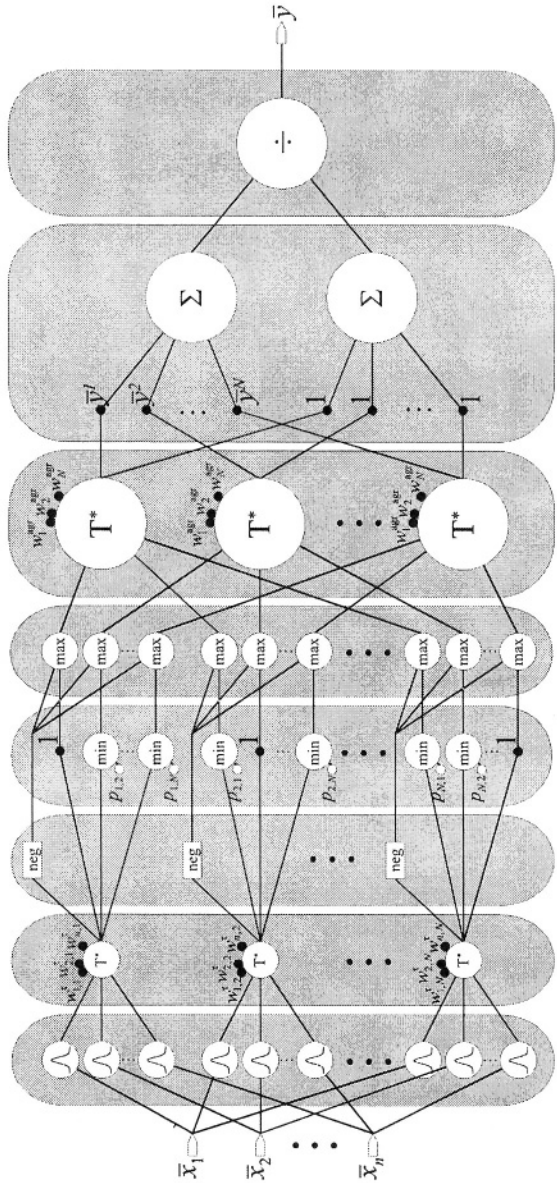

Fig. 8.5. Structure of the weighted neuro-fuzzy system with the Zadeh implication

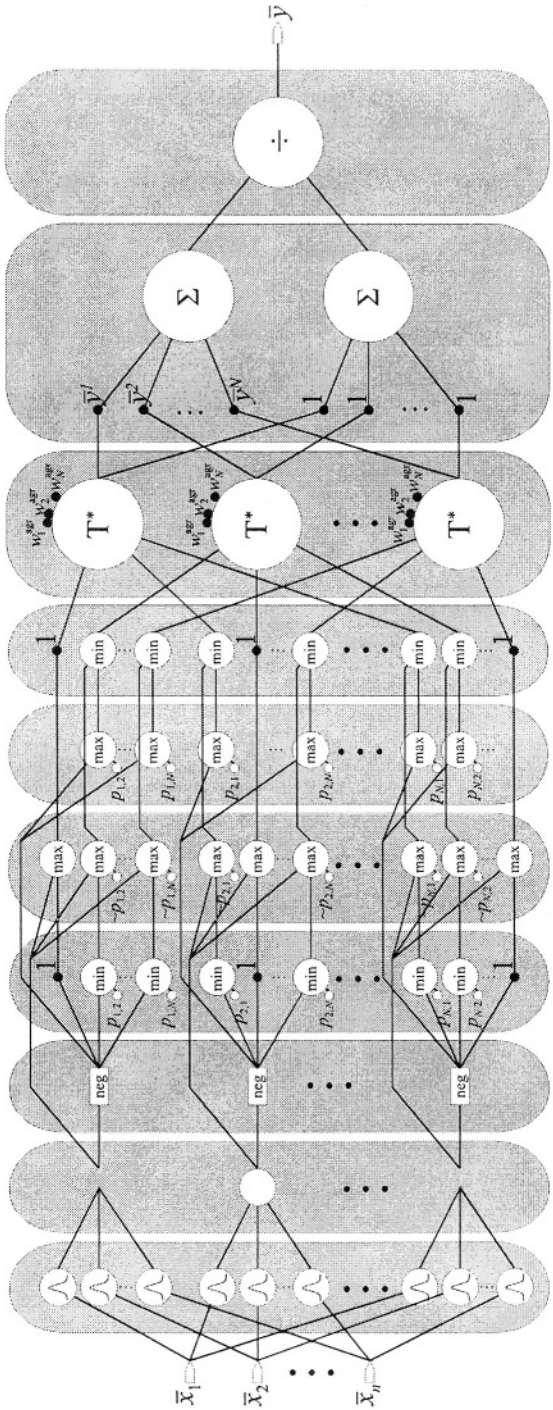

Fig. 8.6. Structure of the weighted neuro-fuzzy system with the Willmott implication

Goguen implication

$$I(a,b)=\begin{cases}\min\left\{1,\dfrac{b}{a}\right\} & \text{if} \quad a>0\\ 1 & \text{if} \quad a=0\end{cases} \tag{8.28}$$

Assuming that $a>0$ and putting function (8.28) into (8.10), we obtain a formula describing the neuro-fuzzy system with the Goguen implication

$$\bar{y}=\frac{\sum_{r=1}^{N}\bar{y}^r\cdot\underset{\substack{k=1\\k\neq r}}{\overset{N}{T}}\left\{\min\left\{1,\dfrac{p_{k,r}}{\overset{n}{\underset{i=1}{T^*}}\left\{\mu_{A_i^k}(\bar{x}_i),w_{i,k}^{\tau}\right\}}\right\},w_k^{\text{agr}}\right\}}{\sum_{r=1}^{N}\underset{\substack{k=1\\k\neq r}}{\overset{N}{T}}\left\{\min\left\{1,\dfrac{p_{k,r}}{\overset{n}{\underset{i=1}{T^*}}\left\{\mu_{A_i^k}(\bar{x}_i),w_{i,k}^{\tau}\right\}}\right\},w_k^{\text{agr}}\right\}} \tag{8.29}$$

Figure 8.7 depicts the structure of the neuro-fuzzy system described by formula (8.29).

Rescher implication

$$I(a,b)=\begin{cases}1 & \text{if} \quad a\le b\\ 0 & \text{if} \quad a>b\end{cases} \tag{8.30}$$

Using (8.19) formula (8.30) can be rewritten in the form

$$I(a,b)=a\overset{\le}{*}b \tag{8.31}$$

Inserting function (8.31) directly into (8.10), we obtain a formula describing the neuro-fuzzy system with the Rescher implication.

$$\bar{y}=\frac{\sum_{r=1}^{N}\bar{y}^r\cdot\underset{\substack{k=1\\k\neq r}}{\overset{N}{T^*}}\left\{\overset{n}{\underset{i=1}{T^*}}\left\{\mu_{A_i^k}(\bar{x}_i),w_{i,k}^{\tau}\right\}\overset{\le}{*}p_{k,r},w_k^{\text{agr}}\right\}}{\sum_{r=1}^{N}\underset{\substack{k=1\\k\neq r}}{\overset{N}{T^*}}\left\{\overset{n}{\underset{i=1}{T^*}}\left\{\mu_{A_i^k}(\bar{x}_i),w_{i,k}^{\tau}\right\}\overset{\le}{*}p_{k,r},w_k^{\text{agr}}\right\}} \tag{8.32}$$

Figure 8.8 depicts the structure of the neuro-fuzzy system described by formula (8.32).

Gödel implication

$$I(a,b)=\begin{cases}1 & \text{if} \quad a\le b\\ b & \text{if} \quad a>b\end{cases} \tag{8.33}$$

Putting function (8.33) directly into (8.10) and using (8.19), we obtain a formula describing the neuro-fuzzy system with the Gödel implication

$$\bar{y} = \frac{\sum_{r=1}^{N} \bar{y}^r \cdot \underset{\substack{k=1 \\ k\neq r}}{\overset{N}{T}} \left\{ \min\left\{ 1, \left(\overset{n}{\underset{i=1}{T^*}} \left\{ \mu_{A_i^k}(\bar{x}_i), w_{i,k}^{\tau} \right\}^{\leq}_{*} p_{k,r} \right) + p_{k,r} \right\}, w_k^{\text{agr}} \right\}}{\sum_{r=1}^{N} \underset{\substack{k=1 \\ k\neq r}}{\overset{N}{T}} \left\{ \min\left\{ 1, \left(\overset{n}{\underset{i=1}{T^*}} \left\{ \mu_{A_i^k}(\bar{x}_i), w_{i,k}^{\tau} \right\}^{\leq}_{*} p_{k,r} \right) + p_{k,r} \right\}, w_k^{\text{agr}} \right\}} \tag{8.34}$$

Figure 8.9 depicts the structure of the neuro-fuzzy system described by formula (8.34).

Yager implication

$$I(a,b) = \begin{cases} b^a & \text{if} \quad a > 0 \\ 1 & \text{if} \quad a = 0 \end{cases} \tag{8.35}$$

Inserting function (8.35) directly into (8.10), we obtain a formula describing the neuro-fuzzy system with the Yager implication

$$\bar{y} = \frac{\sum_{r=1}^{N} \bar{y}^r \cdot \underset{\substack{k=1 \\ k\neq r}}{\overset{N}{T^*}} \left\{ p_{k,r}^{\overset{n}{\underset{i=1}{T^*}} \left\{ \mu_{A_i^k}(\bar{x}_i), w_{i,k}^{\tau} \right\}}, w_k^{\text{agr}} \right\}}{\sum_{r=1}^{N} \underset{\substack{k=1 \\ k\neq r}}{\overset{N}{T^*}} \left\{ p_{k,r}^{\overset{n}{\underset{i=1}{T^*}} \left\{ \mu_{A_i^k}(\bar{x}_i), w_{i,k}^{\tau} \right\}}, w_k^{\text{agr}} \right\}} \tag{8.36}$$

Figure 8.10 depicts the structure of the neuro-fuzzy system described by formula (8.36).

Simplified structures

Similarly to the Mamdani approach, it is also possible to propose a simplification of aforementioned structures in case of the logical approach. Observe that under assumption (7.23), equation (8.6) takes the form

$$\begin{aligned} I(a,0) &= S\{1-a,0\} \\ &= 1-a \end{aligned} \tag{8.37}$$

which boils down formula (8.10) to

$$\bar{y} = \frac{\sum_{r=1}^{N} \bar{y}^r \cdot \underset{\substack{k=1 \\ k\neq r}}{\overset{N}{T^*}} \left\{ 1 - \overset{n}{\underset{i=1}{T^*}} \left\{ \mu_{A_i^k}(\bar{x}_i), w_{i,k}^{\tau} \right\}, w_k^{\text{agr}} \right\}}{\sum_{r=1}^{N} \underset{\substack{k=1 \\ k\neq r}}{\overset{N}{T^*}} \left\{ 1 - \overset{n}{\underset{i=1}{T^*}} \left\{ \mu_{A_i^k}(\bar{x}_i), w_{i,k}^{\tau} \right\}, w_k^{\text{agr}} \right\}} \tag{8.38}$$

In this way we have obtained a neuro-fuzzy system description which employs connectives being S-implications. Figure 8.11 depicts the structure of the neuro-fuzzy system described by formula (8.38).

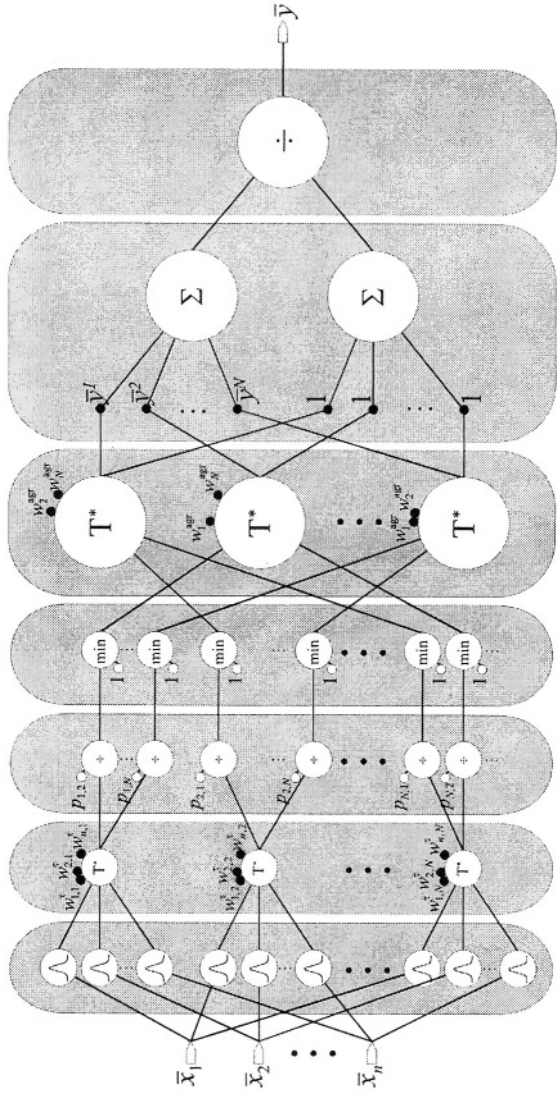

Fig. 8.7. Structure of the weighted neuro-fuzzy system with the Goguen implication

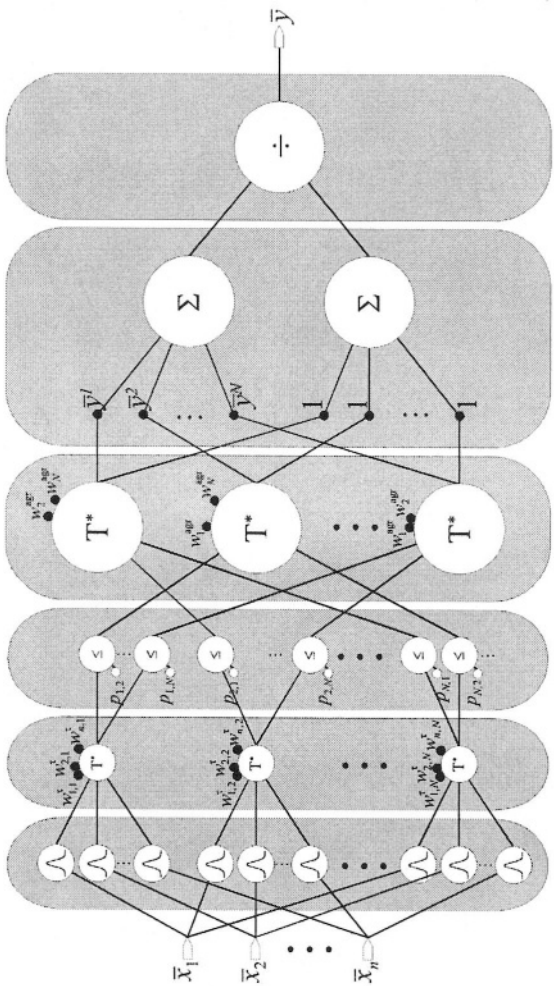

Fig. 8.8. Structure of the weighted neuro-fuzzy system with the Rescher implication

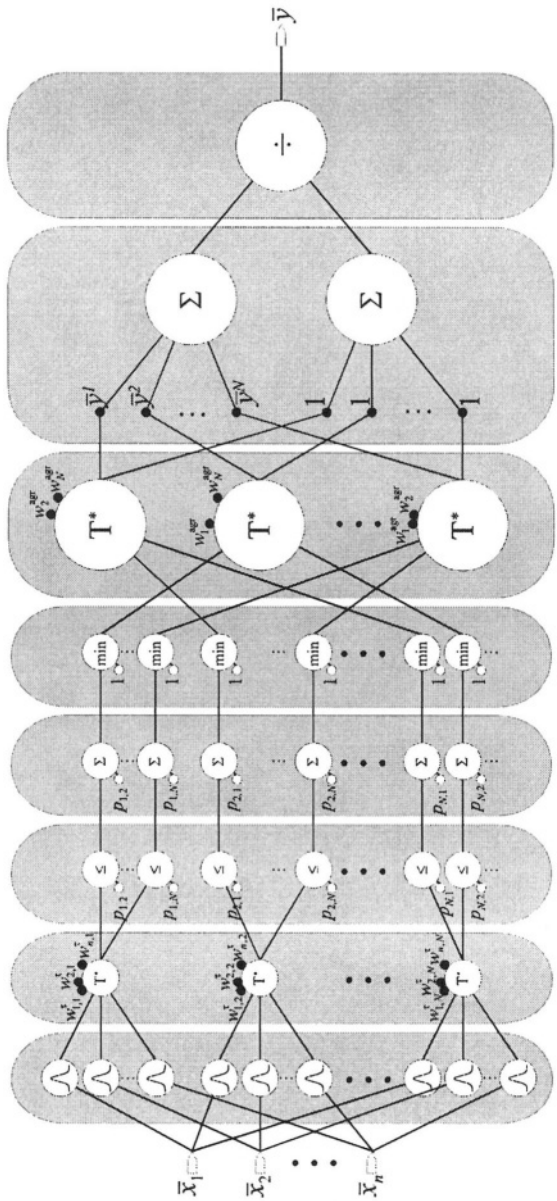

Fig. 8.9. Structure of the weighted neuro-fuzzy system with the Gödel implication

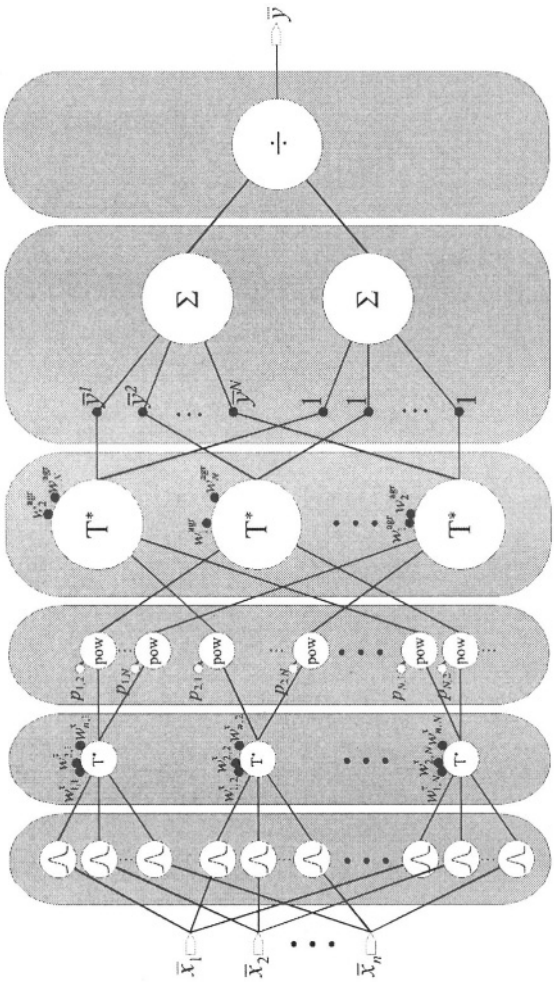

Fig. 8.10. Structure of the weighted neuro-fuzzy system with the Yager implication

In case of the Zadeh and Willmott implications, assumption (7.23) also leads to equation (8.37). Therefore, formula (8.23) takes the form

$$\bar{y} = \frac{\sum_{r=1}^{N} \bar{y}^r \cdot T^*\left\{\left(\max\left\{T_{i=1}^{n*}\left\{\mu_{A_i^r}(\bar{x}_i), w_{i,r}^{\tau}\right\}, 1 - T_{i=1}^{n*}\left\{\mu_{A_i^r}(\bar{x}_i), w_{i,r}^{\tau}\right\}\right\}, w_r^{\text{agr}}\right), T_{\substack{k=1 \\ k\neq r}}^{N*}\left\{1 - T_{i=1}^{n*}\left\{\mu_{A_i^k}(\bar{x}_i), w_{i,k}^{\tau}\right\}, w_k^{\text{agr}}\right\}\right\}}{\sum_{r=1}^{N} T^*\left\{\left(\max\left\{T_{i=1}^{n*}\left\{\mu_{A_i^r}(\bar{x}_i), w_{i,r}^{\tau}\right\}, 1 - T_{i=1}^{n*}\left\{\mu_{A_i^r}(\bar{x}_i), w_{i,r}^{\tau}\right\}\right\}, w_r^{\text{agr}}\right), T_{\substack{k=1 \\ k\neq r}}^{N*}\left\{1 - T_{i=1}^{n*}\left\{\mu_{A_i^k}(\bar{x}_i), w_{i,k}^{\tau}\right\}, w_k^{\text{agr}}\right\}\right\}} \tag{8.39}$$

for the Zadeh and Willmot implications. Figure 8.12 depicts the structure of the neuro-fuzzy system described by formula (8.39).

In case of the Goguen, Rescher, Gödel and Yager fuzzy implications, assumption (7.23) reduces their functional forms to the following

$$I(a,0) = \begin{cases} 0 & \text{if} \quad a > 0 \\ 1 & \text{if} \quad a = 0 \end{cases} \tag{8.40}$$

which is convenient to write down as

$$I(a,0) = a \overset{=}{*} 0 \tag{8.41}$$

where symbol $\overset{=}{*}$ means binary operator defined by

$$a \overset{=}{*} b = \begin{cases} 1 & \text{if} \quad a = b \\ 0 & \text{if} \quad a \neq b \end{cases} \tag{8.42}$$

Finally, using (8.10) and the above notation we get

$$\bar{y} = \frac{\sum_{r=1}^{N} \bar{y}^r \cdot T_{\substack{k=1 \\ k\neq r}}^{N*}\left\{T_{i=1}^{n*}\left\{\mu_{A_i^k}(\bar{x}_i), w_{i,k}^{\tau}\right\} \overset{=}{*} 0, w_k^{\text{agr}}\right\}}{\sum_{r=1}^{N} T_{\substack{k=1 \\ k\neq r}}^{N*}\left\{T_{i=1}^{n*}\left\{\mu_{A_i^k}(\bar{x}_i), w_{i,k}^{\tau}\right\} \overset{=}{*} 0, w_k^{\text{agr}}\right\}} \tag{8.43}$$

which is a description of the neuro-fuzzy system based on the Goguen, Rescher, Gödel and Yager fuzzy implications. Figure 8.13 depicts the structure of the neuro-fuzzy system described by formula (8.43).

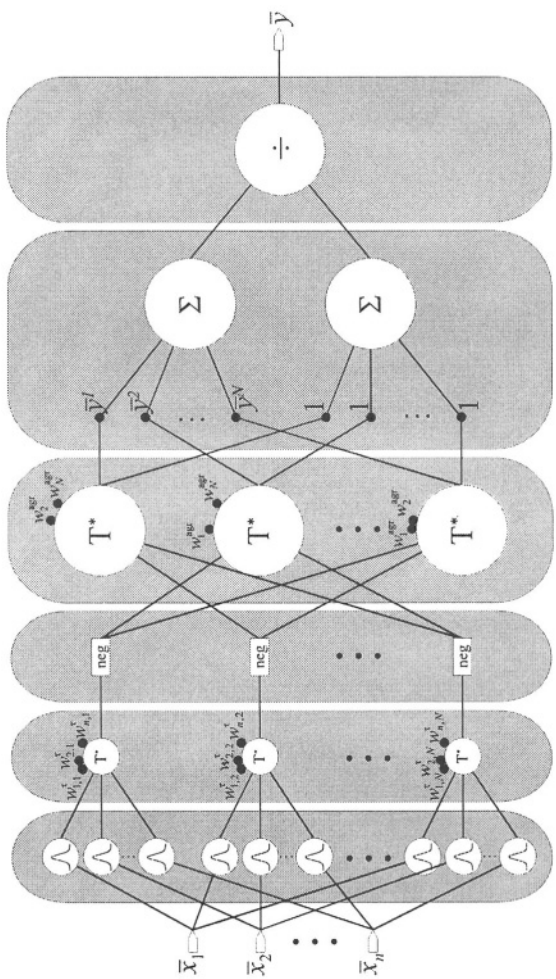

Fig. 8.11. Structure of the weighted neuro-fuzzy system with binary, Łukasiewicz, Reichenbach or the Fodor implication

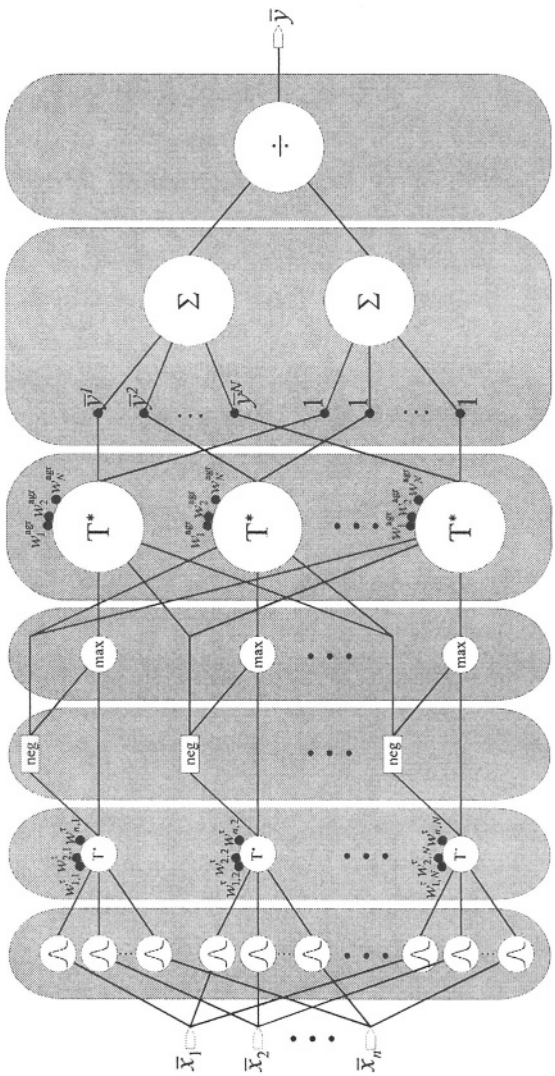

Fig. 8.12. Structure of the weighted neuro-fuzzy system with Zadeh or the Willmott implication

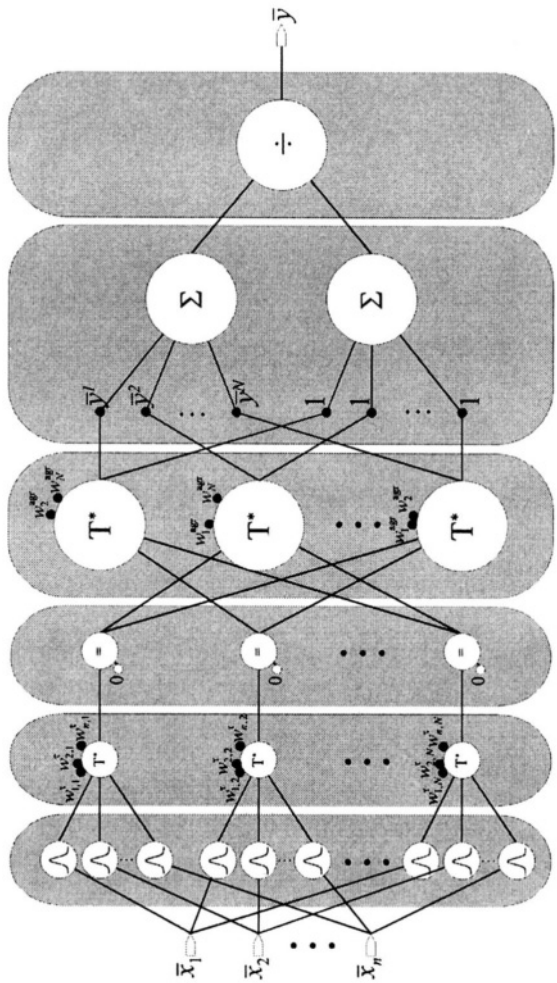

Fig. 8.13. Structure of the weighted neuro-fuzzy system with the Goguen, Rescher, Gödel or Yager implications

8.4. SIMULATION RESULTS

In this section we present four simulations of logical-type neuro-fuzzy systems. We use benchmarks described in Sections 3.6. Each of the four simulations is designed in the same fashion:

(i) In the first experiment, based on the input-output data, we learn the parameters of the membership functions.

(ii) In the second experiment, we learn the parameters of the membership functions, and soft parameters $\alpha^{\tau} \in [0,1]$, $\alpha^{t} \in [0,1]$, $\alpha^{\text{agr}} \in [0,1]$.

(iii) In the third experiment, we learn the parameters of the membership functions and, moreover, the weights $w_{i,k}^{\tau} \in [0,1]$, $i = 1,\ldots, n$, $k = 1,\ldots, N$, in the antecedents of rules and weights $w_k^{\text{agr}} \in [0,1]$, $k = 1,\ldots, N$, of the aggregation operator of the rules. In all diagrams (weights representation) we separate $w_{i,k}^{\tau} \in [0,1]$, $i = 1,\ldots, n$, $k = 1,\ldots, N$, from $w_k^{\text{agr}} \in [0,1]$, $k = 1,\ldots, N$, by a vertical dashed line.

(iv) In the fourth experiment, we learn the parameters of the membership functions, soft parameters $\alpha^{\tau} \in [0,1]$, $\alpha^{t} \in [0,1]$, $\alpha^{\text{agr}} \in [0,1]$, and the weights $w_{i,k}^{\tau} \in [0,1]$, $i = 1,\ldots, n$, $k = 1,\ldots, N$, in the antecedents of rules and weights $w_k^{\text{agr}} \in [0,1]$, $k = 1,\ldots, N$, of the aggregation operator of the rules.

In each simulation we apply six fuzzy implications to connect antecedents and consequents:

a) two S-implications: Kleene-Dienes and Reichenbach;

b) two R-implications: Gödel and Goguen;

c) two Q-implications: Zadeh and Q-algebraic given by (2.89) with min and algebraic t-norms, respectively.

In the case of the Kleene-Dienes, Gödel and Zadeh implications the rules are aggregated by the min t-norm, in the case of the Reichenbach, Goguen and Q-algebraic implications the rules are aggregated by the algebraic t-norm.

The parameters and weights in experiments (i)-(iv) based on S-implications are determined by gradient procedures presented in Section 5.8 (setting $\nu = 1$) or in Section 6.7 (setting $\lambda = 1$). In case of the R-implications and the Q-implications a slight modification of those procedures is required (see Problem 8.7). Moreover, in case of the R-implications the softness idea is incorporated heuristically as a combination (controlled by parameter α^{t}) of the Goguen or Gödel fuzzy implications with the average of arguments $1 - a$ and b of the Kleene-Dienes implication.

Box and Jenkins Gas Furnace problem

The experimental results for the Box and Jenkins Gas Furnace problem are depicted in Tables 8.1, 8.2, and 8.3 for the S, R, and Q fuzzy implications, respectively. The final values (after learning) of weights $w_{i,k}^{r} \in [0,1]$ and $w_k^{agr} \in [0,1]$, $i = 1,\ldots,6$, $k = 1,\ldots,4$, are shown in Fig. 8.14, 8.15, and 8.16 for the S, R, and Q fuzzy implications, respectively. Assuming the Kleene-Dienes implication, in Fig. 8.14-a.1 and 8.14-a.2 we present the results of experiments (iii) and (iv) in Table 8.1, respectively. Analogous results for the Reichenbach inference are given in Fig. 8.14-b.1 and 8.14-b.2. In Fig. 8.15 and Table 8.2 and Fig. 8.16 and Table 8.3 we show simulation results for the R and Q implications, respectively.

Table 8.1 Experimental results

LOGICAL-TYPE NFIS WITH S-IMPLICATIONS (BOX AND JENKINS GAS FURNACE PROBLEM)						
Experiment number	Name of flexibility parameter	Initial values	Final values after learning		RMSE (learning sequence)	
			Kleene-Dienes implication	Reichenbach implication	Kleene-Dienes implication	Reichenbach implication
i	-	-	-	-	0.6352	0.6188
ii	α^r	1	0.9326	0.0908	0.5912	0.5725
	α^l	1	0.9467	0.9006		
	α^{agr}	1	0.9836	0.9610		
iii	$\mathbf{w}^r$	**1**	Fig. 8.14-a	Fig. 8.14-b	0.3594	0.3352
	$\mathbf{w}^{agr}$	**1**	Fig. 8.14-a	Fig. 8.14-b		
iv	α^r	1	0.1082	0.9858	0.3328	0.3254
	α^l	1	0.9772	0.9815		
	α^{agr}	1	0.8746	0.9288		
	$\mathbf{w}^r$	**1**	Fig. 8.14-a	Fig. 8.14-b		
	$\mathbf{w}^{agr}$	**1**	Fig. 8.14-a	Fig. 8.14-b		

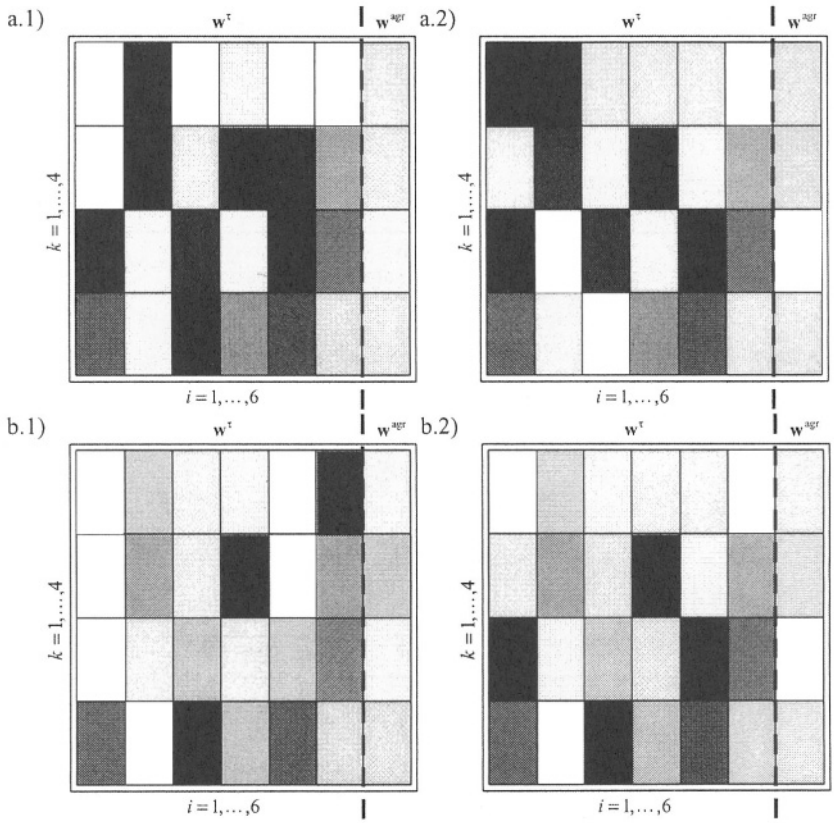

Fig. 8.14. Weights representation in the Modeling of Box and Jenkins Gas Furnace problem for logical-type NFIS and a) Kleene-Dienes S-implication, b) Reichenbach S-implication

Table 8.2 Experimental results

LOGICAL-TYPE NFIS WITH R-IMPLICATION (BOX AND JENKINS GAS FURNACE PROBLEM)						
Experiment number	Name of flexibility parameter	Initial values	Final values after learning		RMSE (learning sequence)	
			Gödel implication	Goguen implication	Gödel implication	Goguen implication
i	-	-	-	-	3.6130	3.2330
ii	α^r	1	0.0378	0.0519	3.3210	3.2290
	α^l	1	0.9012	0.9775		
	α^{agr}	1	0.9627	0.9481		
iii	$\mathbf{w}^r$	**1**	Fig. 8.15-a	Fig. 8.15-b	3.2290	3.0990
	$\mathbf{w}^{agr}$	**1**	Fig. 8.15-a	Fig. 8.15-b		
iv	α^r	1	0.1388	0.0299	2.7850	1.0990
	α^l	1	0.9275	0.9639		
	α^{agr}	1	0.0587	0.9493		
	$\mathbf{w}^r$	**1**	Fig. 8.15-a	Fig. 8.15-b		
	$\mathbf{w}^{agr}$	**1**	Fig. 8.15-a	Fig. 8.15-b		

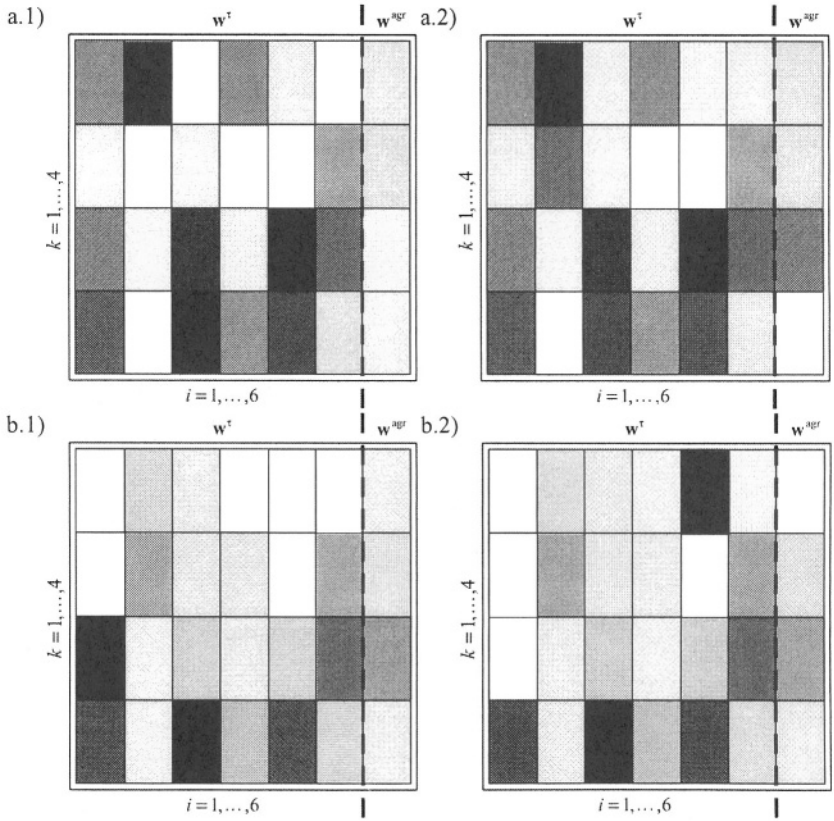

Fig. 8.15. Weights representation in the Modeling of Box and Jenkins Gas Furnace problem for logical-type NFIS and a) Gödel R-implication, b) Goguen R-implication

Table 8.3 Experimental results

LOGICAL-TYPE NFIS WITH Q-IMPLICATION (BOX AND JENKINS GAS FURNACE PROBLEM)						
Experiment number	Name of flexibility parameter	Initial values	Final values after learning		RMSE (learning sequence)	
			Zadeh implication	Algebraic Q-implication	Zadeh implication	Algebraic Q-implication
i	-	-	-	-	1.1800	2.8035
ii	α^{r}	1	0.9382	0.0065	0.9501	0.8754
	α^{l}	1	0.9526	0.9529		
	α^{agr}	1	0.0952	0.9706		
iii	$\mathbf{w}^{r}$	**1**	Fig. 8.16-a	Fig. 8.16-b	0.8118	0.8253
	$\mathbf{w}^{agr}$	**1**	Fig. 8.16-a	Fig. 8.16-b		
iv	α^{r}	1	0.9873	0.0553	0.7585	0.3250
	α^{l}	1	0.9805	0.9521		
	α^{agr}	1	0.9465	0.9332		
	$\mathbf{w}^{r}$	**1**	Fig. 8.16-a	Fig. 8.16-b		
	$\mathbf{w}^{agr}$	**1**	Fig. 8.16-a	Fig. 8.16-b		

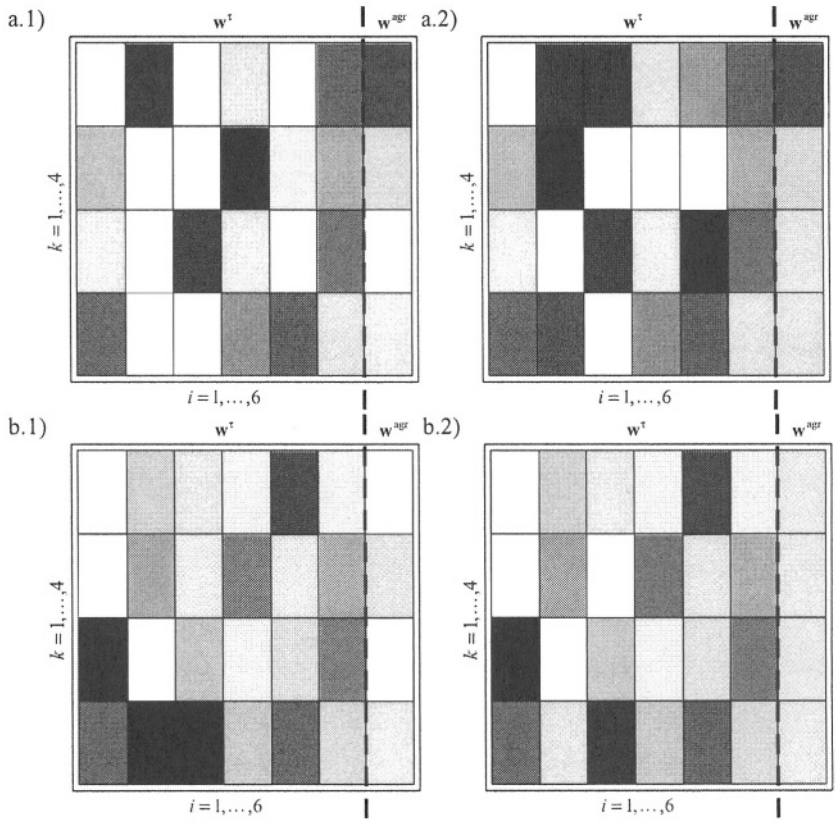

Fig. 8.16. Weights representation in the Modeling of Box and Jenkins Gas Furnace problem for logical-type NFIS and a) Zadeh Q-implication, b) algebraic Q-implication

Glass Identification problem

The experimental results for the Glass Identification problem are depicted in Tables 8.4, 8.5, and 8.6 for the S, R, and Q fuzzy implications, respectively. The final values (after learning) of weights $w_{i,k}^{r} \in [0,1]$ and $w_{k}^{agr} \in [0,1]$, $i = 1,\ldots,9$, $k = 1,\ldots,2$, are shown in Fig. 8.17, 8.18, and 8.19 for the S, R, and Q fuzzy implications, respectively. Assuming the Kleene-Dienes implication, in Fig. 8.17-a.1 and 8.17-a.2 we present the results of experiments (iii) and (iv) in Table 8.4, respectively. Analogous results for the Reichenbach inference are given in Fig. 8.17-b.1 and 8.17-b.2. In Fig. 8.18 and Table 8.5 and Fig. 8.19 and Table 8.6 we illustrate simulation results for the R and Q implications, respectively.

Table 8.4 Experimental results

LOGICAL-TYPE NFIS WITH S-IMPLICATIONS (GLASS IDENTIFICATION PROBLEM)								
Experiment number	Name of flexibility parameter	Initial values	Final values after learning		Mistakes [%] (learning sequence)		Mistakes [%] (testing sequence)	
			Kleene-Dienes implication	Reichenbach implication	Kleene-Dienes implication	Reichenbach implication	Kleene-Dienes implication	Reichenbach implication
i	-	-	-	-	3.33	3.33	3.13	3.13
ii	α^{r} α^{l} α^{agr}	1 1 1	0.0298 0.9641 0.9783	0.0060 0.9520 0.9769	2.67	3.33	3.13	1.56
iii	$\mathbf{w}^{r}$ $\mathbf{w}^{agr}$	**1** **1**	Fig. 8.17-a Fig. 8.17-a	Fig. 8.17-b Fig. 8.17-b	2.67	3.33	1.56	1.56
iv	α^{r} α^{l} α^{agr} $\mathbf{w}^{r}$ $\mathbf{w}^{agr}$	1 1 1 **1** **1**	0.0660 0.9722 0.9808 Fig. 8.17-a Fig. 8.17-a	0.0545 0.9471 0.9842 Fig. 8.17-b Fig. 8.17-b	2.00	2.67	1.56	1.56

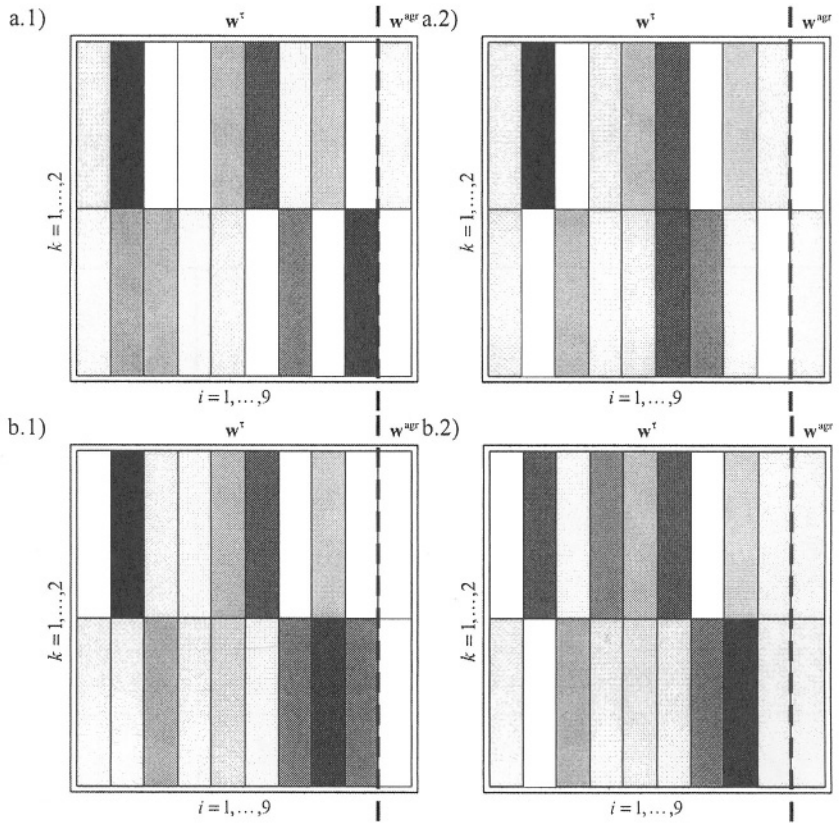

Fig. 8.17. Weights representation in the Glass Identification problem for logical-type NFIS and a) Kleene-Dienes S-implication, b) Reichenbach S-implication

Table 8.5 Experimental results

LOGICAL-TYPE NFIS WITH R-IMPLICATION (GLASS IDENTIFICATION PROBLEM)								
Experiment number	Name of flexibility parameter	Initial values	Final values after learning		Mistakes [%] (learning sequence)		Mistakes [%] (testing sequence)	
			Gödel implication	Goguen implication	Gödel implication	Goguen implication	Gödel implication	Goguen implication
i	-	-	-	-	28.67	20.67	32.81	28.13
ii	α^r α^l α^{agr}	1 1 1	0.0563 0.9733 0.9767	0.9999 0.9414 0.0041	23.33	8.67	25.00	7.81
iii	$\mathbf{w}^r$ $\mathbf{w}^{agr}$	**1** **1**	Fig. 8.18-a Fig. 8.18-a	Fig. 8.18-b Fig. 8.18-b	23.33	7.99	25.00	6.25
iv	α^r α^l α^{agr} $\mathbf{w}^r$ $\mathbf{w}^{agr}$	1 1 1 **1** **1**	0.0060 0.9850 0.9335 Fig. 8.18-a Fig. 8.18-a	0.0898 0.8903 0.9672 Fig. 8.18-b Fig. 8.18-b	11.33	6.00	14.06	6.25

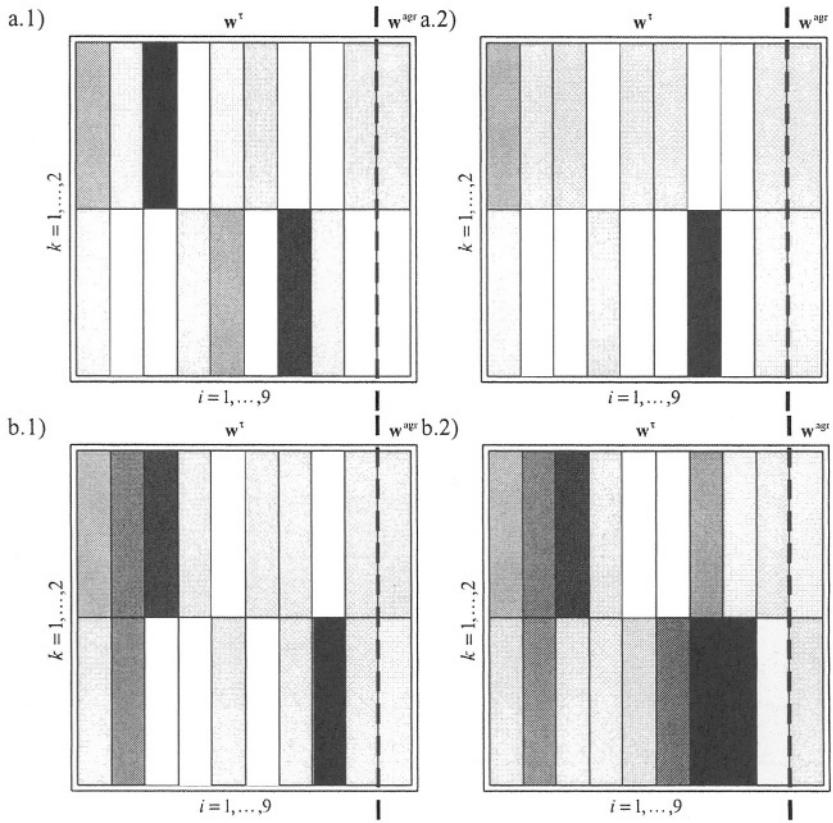

Fig. 8.18. Weights representation in the Glass Identification problem for logical-type NFIS and a) Gödel R-implication, b) Goguen R-implication

Table 8.6 Experimental results

LOGICAL-TYPE NFIS WITH Q-IMPLICATION (GLASS IDENTIFICATION PROBLEM)								
Experiment number	**Name of flexibility parameter**	**Initial values**	**Final values after learning**		**Mistakes [%] (learning sequence)**		**Mistakes [%] (testing sequence)**	
			Zadeh implication	**Algebraic Q-implication**	**Zadeh implication**	**Algebraic Q-implication**	**Zadeh implication**	**Algebraic Q-implication**
i	-	-	-	-	23.33	7.33	25.00	9.38
ii	α^r	1	0.0780	0.9978	5.33	6.67	6.25	7.81
	α^j	1	0.9066	0.9183				
	α^{agr}	1	0.9528	0.9694				
iii	$\mathbf{w}^r$	**1**	Fig. 8.19-a	Fig. 8.19-b	5.33	5.33	4.69	6.25
	$\mathbf{w}^{agr}$	**1**	Fig. 8.19-a	Fig. 8.19-b				
iv	α^r	1	0.0832	0.0063	3.33	4.00	3.13	4.69
	α^j	1	0.9845	0.9094				
	α^{agr}	1	0.9689	0.9643				
	$\mathbf{w}^r$	**1**	Fig. 8.19-a	Fig. 8.19-b				
	$\mathbf{w}^{agr}$	**1**	Fig. 8.19-a	Fig. 8.19-b				

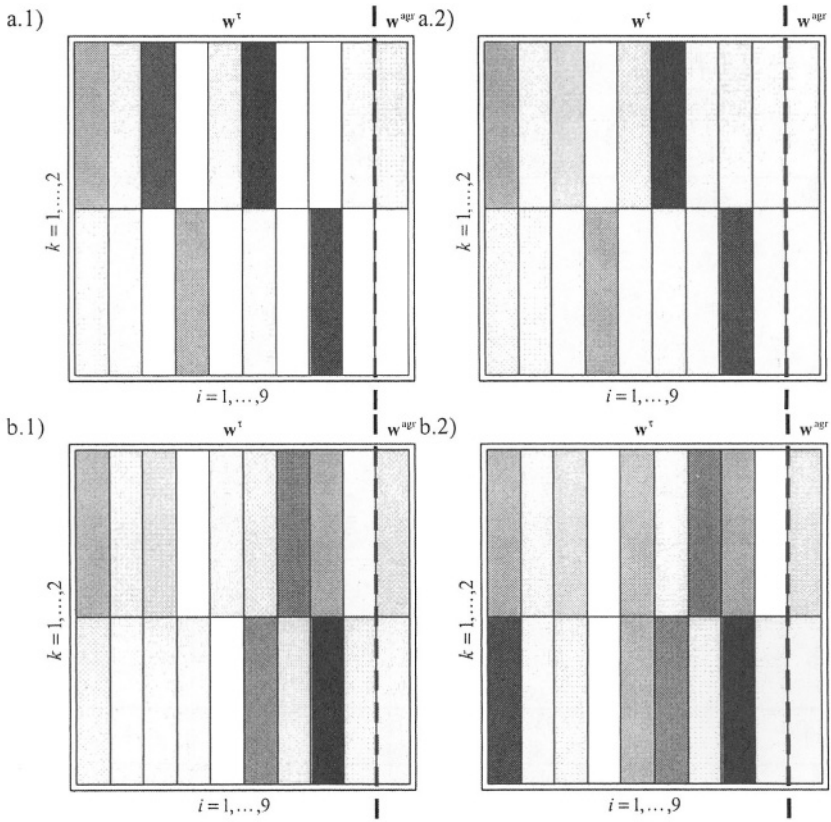

Fig. 8.19. Weights representation in the Glass Identification problem for logical-type NFIS and a) Zadeh Q-implication, b) algebraic Q-implication

Modeling of Static Nonlinear Function (HANG) problem

The experimental results for the Modeling of the Static Nonlinear Function problem are shown in Tables 8.7, 8.8, and 8.9 for the S, R, and Q fuzzy implications, respectively. The final values (after learning) of weights $w_{i,k}^{r} \in [0,1]$ and $w_{k}^{agr} \in [0,1]$, $i = 1,\ldots,2$, $k = 1,\ldots,5$, are depicted in Fig. 8.20, 8.21, and 8.22 for the S, R, and Q fuzzy implications, respectively. Assuming the Kleene-Dienes implication, in Fig. 8.20-a.1 and 8.20-a.2 we present the results of experiments (iii) and (iv) in Table 8.7, respectively. Analogous results for the Reichenbach inference are given in Fig. 8.20-b.1 and 8.20-b.2. In Fig. 8.21 and Table 8.8 and Fig. 8.22 and Table 8.9 we illustrate simulation results for the R and Q implications, respectively.

Table 8.7 Experimental results

LOGICAL-TYPE NFIS WITH S-IMPLICATIONS (MODELING OF STATIC NONLINEAR FUNCTION PROBLEM)						
Experiment number	Name of flexibility parameter	Initial values	Final values after learning		RMSE (learning sequence)	
			Kleene-Dienes implication	Reichenbach implication	Kleene-Dienes implication	Reichenbach implication
i	-	-	-	-	0.1516	0.1205
ii	α^{r}	1	0.1165	0.0418	0.1459	0.1081
	α^{l}	1	0.9207	0.8546		
	α^{agr}	1	0.9060	0.8663		
iii	$\mathbf{w}^{r}$	**1**	Fig. 8.20-a	Fig. 8.20-b	0.1312	0.1123
	$\mathbf{w}^{agr}$	**1**	Fig. 8.20-a	Fig. 8.20-b		
iv	α^{r}	1	0.0647	0.1280	0.1296	0.0983
	α^{l}	1	0.9063	0.8749		
	α^{agr}	1	0.9577	0.9809		
	$\mathbf{w}^{r}$	**1**	Fig. 8.20-a	Fig. 8.20-b		
	$\mathbf{w}^{agr}$	**1**	Fig. 8.20-a	Fig. 8.20-b		

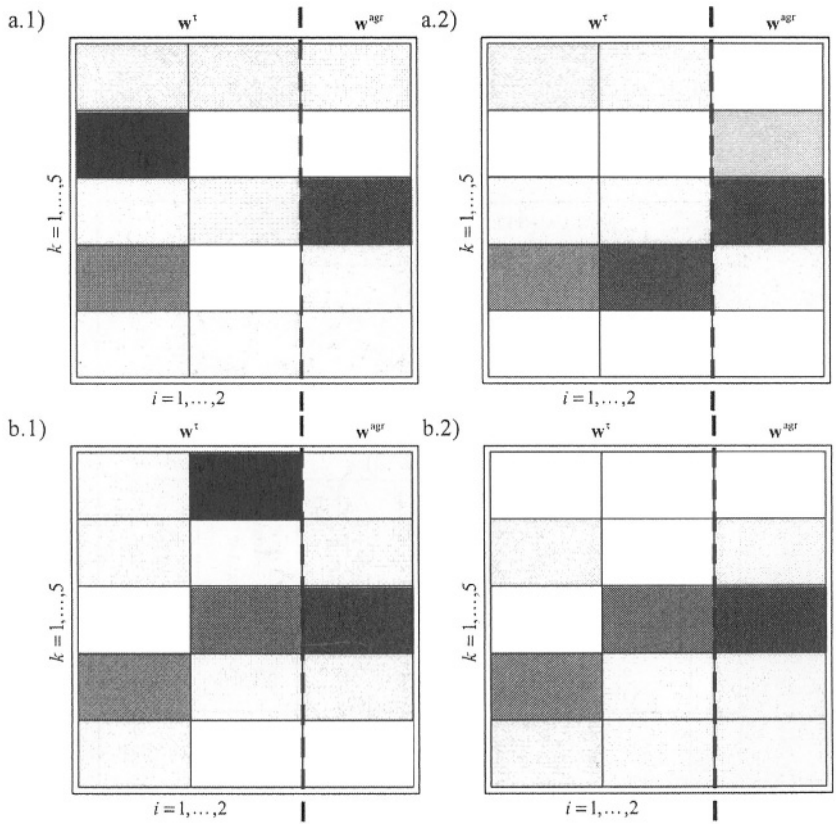

Fig. 8.20. Weights representation in the HANG problem
for logical-type NFIS and a) Kleene-Dienes S-implication, b) Reichenbach S-implication

Table 8.8 Experimental results

LOGICAL-TYPE NFIS WITH R-IMPLICATION (MODELING OF STATIC NONLINEAR FUNCTION PROBLEM)						
Experiment number	Name of flexibility parameter	Initial values	Final values after learning		RMSE (learning sequence)	
			Gödel implication	Goguen implication	Gödel implication	Goguen implication
i	-	-	-	-	0.5627	1.0540
ii	α^r	1	0.0220	0.1296	0.3074	1.0340
	α^l	1	0.8811	0.9506		
	α^{agr}	1	0.9053	0.0142		
iii	$\mathbf{w}^r$	**1**	Fig. 8.21-a	Fig. 8.21-b	0.2324	1.0130
	$\mathbf{w}^{agr}$	**1**	Fig. 8.21-a	Fig. 8.21-b		
iv	α^r	1	0.0039	0.9721	0.2062	0.0948
	α^l	1	0.9282	0.9398		
	α^{agr}	1	0.8606	0.9586		
	$\mathbf{w}^r$	**1**	Fig. 8.21-a	Fig. 8.21-b		
	$\mathbf{w}^{agr}$	**1**	Fig. 8.21-a	Fig. 8.21-b		

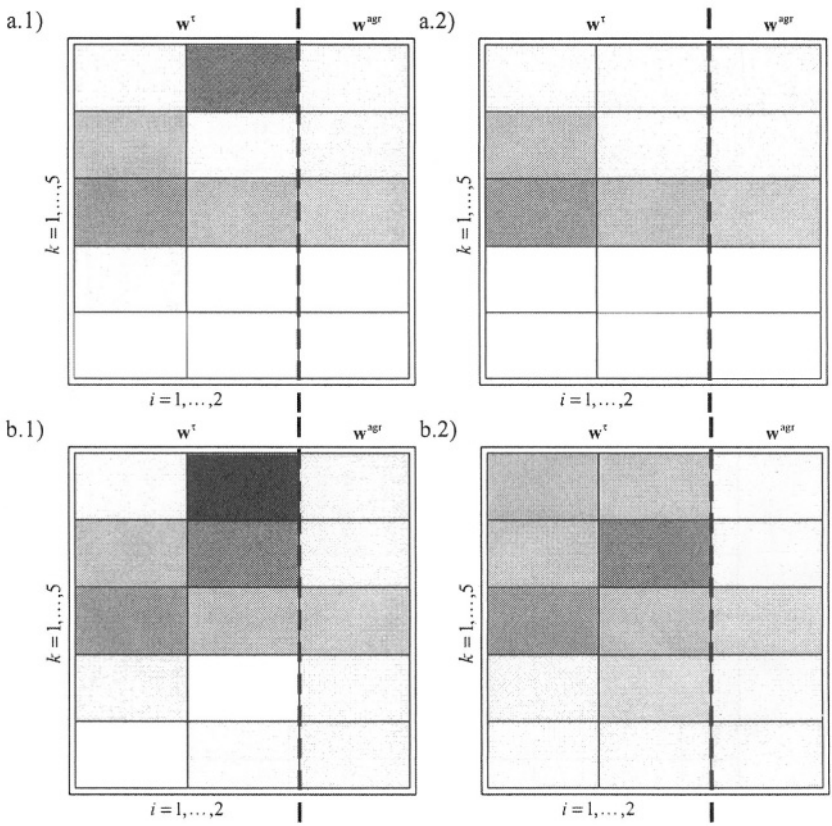

Fig. 8.21. Weights representation in the HANG problem
for logical-type NFIS and a) Gödel R-implication, b) Goguen R-implication

Table 8.9 Experimental results

LOGICAL-TYPE NFIS WITH Q-IMPLICATION (MODELING OF STATIC NONLINEAR FUNCTION PROBLEM)						
Experiment number	Name of flexibility parameter	Initial values	Final values after learning		RMSE (learning sequence)	
			Zadeh implication	Algebraic Q-implication	Zadeh implication	Algebraic Q-implication
i	-	-	-	-	1.0660	0.7355
ii	α^r	1	0.1384	0.9436	1.0330	0.3361
	α^l	1	0.9551	0.9264		
	α^{agr}	1	0.0879	0.9857		
iii	$\mathbf{w}^r$	**1**	Fig. 8.22-a	Fig. 8.22-b	0.7553	0.1617
	$\mathbf{w}^{agr}$	**1**	Fig. 8.22-a	Fig. 8.22-b		
iv	α^r	1	0.9286	0.0393	0.5396	0.0989
	α^l	1	0.9635	0.9799		
	α^{agr}	1	0.8959	0.9817		
	$\mathbf{w}^r$	**1**	Fig. 8.22-a	Fig. 8.22-b		
	$\mathbf{w}^{agr}$	**1**	Fig. 8.22-a	Fig. 8.22-b		

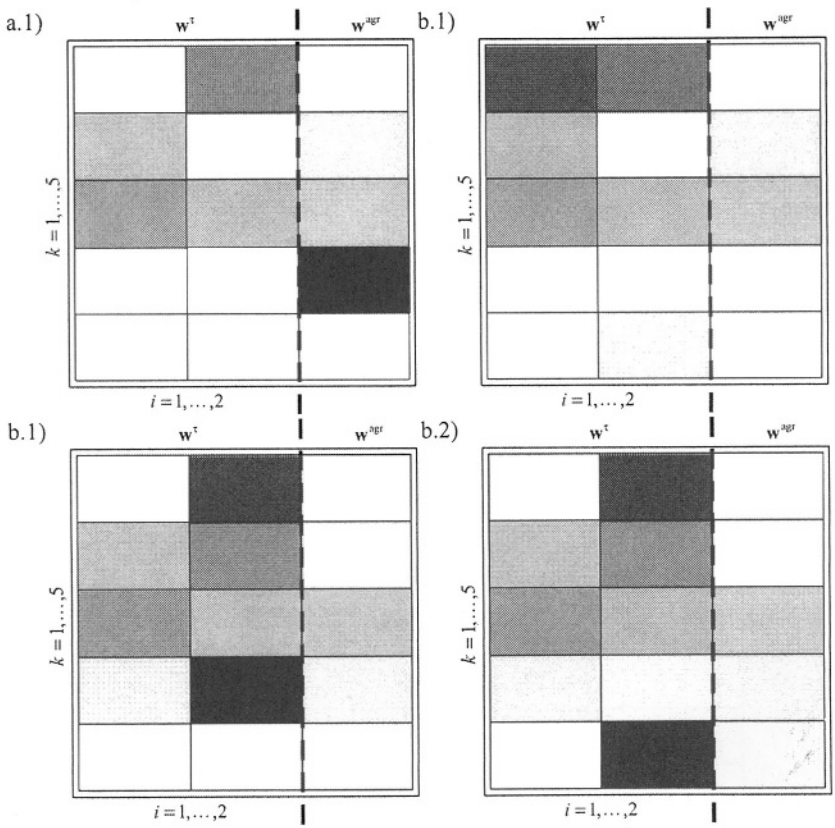

Fig. 8.22. Weights representation in the HANG problem for logical-type NFIS and a) Zadeh Q-implication, b) algebraic Q-implication

Wisconsin Breast Cancer problem

The experimental results for the Wisconsin Breast Cancer problem are shown in Tables 8.10, 8.11, and 8.12 for the S, R, and Q fuzzy implications, respectively. The final values (after learning) of weights $w_{i,k}^{r} \in [0,1]$ and $w_{k}^{agr} \in [0,1]$, $i = 1,\ldots,9$, $k = 1,\ldots,2$, are depicted in Fig. 8.23, 8.24, and 8.25 for the S, R, and Q fuzzy implications, respectively. Assuming the Kleene-Dienes implication, in Fig. 8.23-a.1 and 8.23-a.2 we present the results of experiments (iii) and (iv) in Table 8.10, respectively. Analogous results for the Reichenbach inference are given in Fig. 8.23-b.1 and 8.23-b.2. In Fig. 8.24 and Table 8.11 and Fig. 8.25 and Table 8.12 we depict simulation results for the R and Q implications, respectively.

Table 8.10 Experimental results

LOGICAL-TYPE NFIS WITH S-IMPLICATIONS (WISCONSIN BREAST CANCER PROBLEM)								
Experiment number	Name of flexibility parameter	Initial values	Final values after learning		Mistakes [%] (learning sequence)		Mistakes [%] (testing sequence)	
			Kleene-Dienes implication	Reichenbach implication	Kleene-Dienes implication	Reichenbach implication	Kleene-Dienes implication	Reichenbach implication
i	-	-	-	-	2.93	2.51	1.95	1.95
ii	α^{r} α^{t} α^{agr}	1 1 1	0.0177 0.9819 0.9773	0.0384 0.9380 0.9541	2.51	2.09	1.95	1.46
iii	$\mathbf{w}^{r}$ $\mathbf{w}^{agr}$	**1** **1**	Fig. 8.23-a Fig. 8.23-a	Fig. 8.23-b Fig. 8.23-b	2.30	2.30	1.46	1.46
iv	α^{r} α^{t} α^{agr} $\mathbf{w}^{r}$ $\mathbf{w}^{agr}$	1 1 1 **1** **1**	0.0147 0.9963 0.9489 Fig. 8.23-a Fig. 8.23-a	0.0197 0.9439 0.9456 Fig. 8.23-b Fig. 8.23-b	2.09	1.88	1.46	1.46

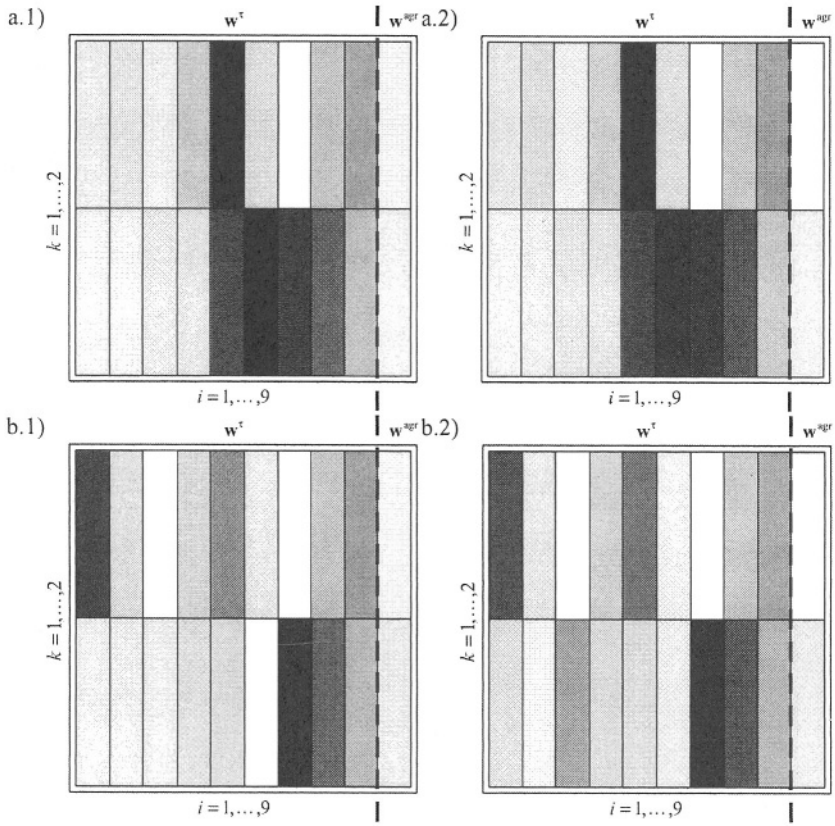

Fig. 8.23. Weights representation in the Wisconsin Breast Cancer problem for logical-type NFIS and a) Kleene-Dienes S-implication, b) Reichenbach S-implication

Table 8.11 Experimental results

LOGICAL-TYPE NFIS WITH R-IMPLICATION (WISCONSIN BREAST CANCER PROBLEM)								
Experiment number	**Name of flexibility parameter**	**Initial values**	**Final values after learning**		**Mistakes [%] (learning sequence)**		**Mistakes [%] (testing sequence)**	
			Gödel implication	**Goguen implication**	**Gödel implication**	**Goguen implication**	**Gödel implication**	**Goguen implication**
i	-	-	-	-	7.74	7.74	4.39	4.39
ii	α^r	1	0.0690	0.9809	6.90	2.30	3.90	2.93
	α^l	1	0.9007	0.8154				
	α^{agr}	1	0.0219	0.1802				
iii	$\mathbf{w}^r$	**1**	Fig. 8.24-a	Fig. 8.24-b	5.86	2.09	3.90	2.44
	$\mathbf{w}^{agr}$	**1**	Fig. 8.24-a	Fig. 8.24-b				
iv	α^r	1	0.0900	0.9347	5.23	2.09	3.41	1.95
	α^l	1	0.1456	0.8704				
	α^{agr}	1	0.9350	0.9632				
	$\mathbf{w}^r$	**1**	Fig. 8.24-a	Fig. 8.24-b				
	$\mathbf{w}^{agr}$	**1**	Fig. 8.24-a	Fig. 8.24-b				

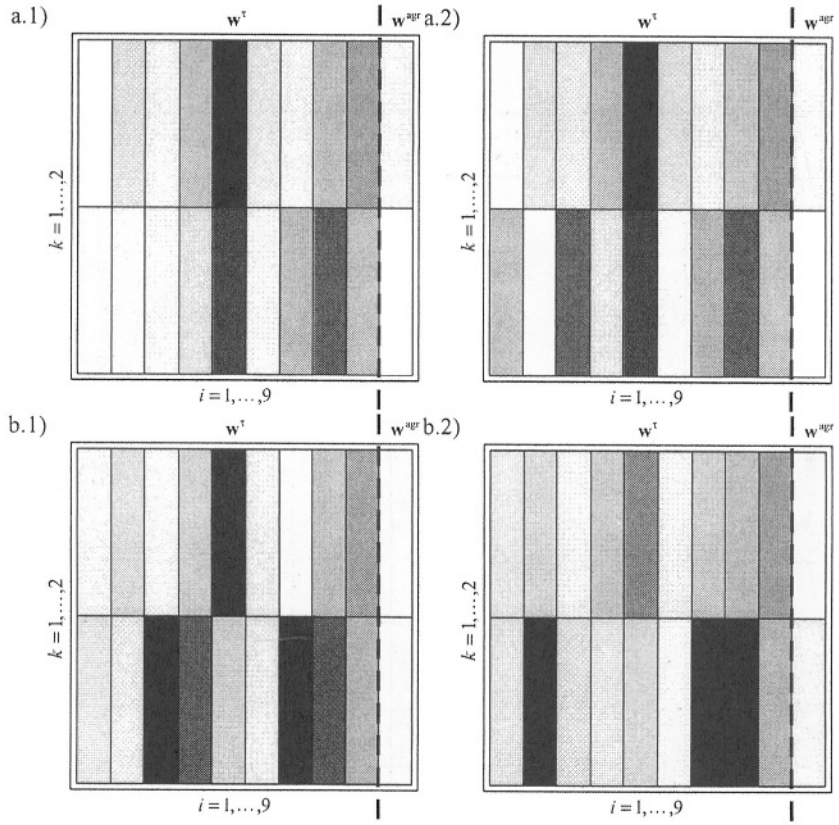

Fig. 8.24. Weights representation in the Wisconsin Breast Cancer problem for logical-type NFIS and a) Gödel R-implication, b) Goguen R-implication

Table 8.12 Experimental results

LOGICAL-TYPE NFIS WITH Q-IMPLICATION (WISCONSIN BREAST CANCER PROBLEM)								
Experiment number	Name of flexibility parameter	Initial values	Final values after learning		Mistakes [%] (learning sequence)		Mistakes [%] (testing sequence)	
			Zadeh implication	Algebraic Q-implication	Zadeh implication	Algebraic Q-implication	Zadeh implication	Algebraic Q-implication
i	-	-	-	-	2.72	2.51	3.90	3.90
ii	α^r	1	0.0354	0.9793	2.09	2.51	2.44	3.90
	α^l	1	0.9115	0.8118				
	α^{agr}	1	0.8422	0.1664				
iii	$\mathbf{w}^r$	**1**	Fig. 8.25-a	Fig. 8.25-b	2.09	2.30	2.44	2.44
	$\mathbf{w}^{agr}$	**1**	Fig. 8.25-a	Fig. 8.25-b				
iv	α^r	1	0.8105	0.1644	2.09	2.09	2.44	2.44
	α^l	1	0.9107	0.9793				
	α^{agr}	1	0.8592	0.9989				
	$\mathbf{w}^r$	**1**	Fig. 8.25-a	Fig. 8.25-b				
	$\mathbf{w}^{agr}$	**1**	Fig. 8.25-a	Fig. 8.25-b				

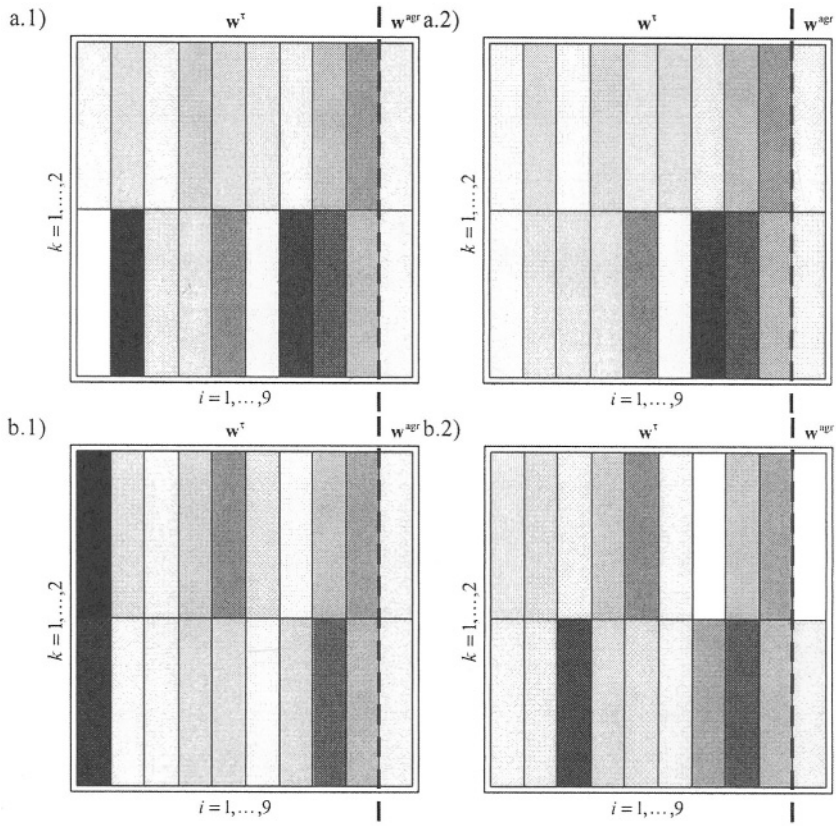

Fig. 8.25. Weights representation in the Wisconsin Breast Cancer problem for logical-type NFIS and a) Zadeh Q-implication, b) algebraic Q-implication

8.5. SUMMARY AND DISCUSSION

In this chapter we derived various weighted neuro-fuzzy logical-type structures based on the binary, Łukasiewicz, Reichenbach, Fodor, Zadeh, Willmott, Goguen, Rescher, Gödel and Yager fuzzy implications. Analogously to the Mamdani approach, we have also obtained simplified neuro-fuzzy structures assuming that condition (7.23) holds. The membership function parameters and weights are determined in the process of learning. In simulations we have also incorporated and learned soft parameters which slightly improve the performance of neuro-fuzzy systems. Again the incorporation of weights and soft parameters gives the best accuracy of all systems.

8.6. PROBLEMS

Problem 8.1. Replace defuzzifier (3.23) by (3.56) and derive a logical fuzzy system based on the binary implication (see [15] for details).

Problem 8.2. Solve Problem 8.1 assuming that a logical fuzzy system is based on the Reichenbach implication.

Problem 8.3. Solve Problem 8.1 assuming that fuzzy sets B^k, $k = 1,\ldots, N$,in rule (3.9) depend on crisp inputs $\bar{\mathbf{x}} = [\bar{x}_1,\ldots,\bar{x}_N] \in \mathbf{X}$ (see [15] for details).

Problem 8.4. Incorporate weights to a fuzzy system derived in Problem 8.1.

Problem 8.5. Derive a flexible logical-type neuro-fuzzy system based on the Dubois-Prade fuzzy implication given by

$$I(a,b) = \begin{cases} 1-a & \text{if} \quad b = 0 \\ b & \text{if} \quad a = 1 \\ 1 & \textit{otherwise} \end{cases} \tag{8.44}$$

Problem 8.6. Derive a simplified neuro-fuzzy structure based on fuzzy implication (8.44).

Problem 8.7. Derive learning algorithms for logical systems based on R and Q-implications.

This page intentionally left blank

Chapter 9

PERFORMANCE COMPARISON OF NEURO-FUZZY SYSTEMS

9.1. INTRODUCTION

In the book we have studied various flexible neuro-fuzzy systems. Their performance was evaluated based on 14 typical benchmarks given in Table 3.2 (see Section 3.6). However, because of the space limitation, we have presented in Chapters 5-8 the results of only four simulations, namely:

- Box and Jenkins Gas Furnace (approximation problem),
- Glass Identification (classification problem),
- Modeling of Static Nonlinear Function (approximation problem),
- Wisconsin Breast Cancer (classification problem).

In this chapter we show comparison charts for all the 14 simulations. The following neuro-fuzzy structures have been examined and compared:

a) Mamdani-type systems with min and product inferences,
b) Logical-type systems with S-implications,
c) Logical-type systems with R-implications,
d) Logical-type systems with Q-implications,
e) OR-type systems with standard triangular norms,
f) OR-type systems with parameterized triangular norms,
g) AND-type systems with standard triangular norms,
h) AND-type systems with parameterized triangular norms.

In each case weights $w_{i,k}^{r}$ and w_{k}^{agr}, $i = 1,...,n$, $k = 1,...,N$, and/or softening parameters α^{r}, α^{l} and α^{agr}, have been incorporated into the construction of

neuro-fuzzy systems. The initial values of parameters of membership functions have been determined by making use of the Fuzzy C-Means algorithm (see e.g. [15, 74]).

9.2. COMPARISON CHARTS

In this section we compare the performance of various neuro-fuzzy structures. For each of the 14 benchmarks given in Table 3.2 we present comparison charts (Fig. 9.1-9.14) and comparison tables (Tables 9.1-9.14). In Figures 9.1-9.14 the best results (the minimum of the RMSE or the minimum number of mistakes) of applied neuro-fuzzy structures are indicated.

Box and Jenkins Gas Furnace problem

The comparison chart and the comparison table for the Box and Jenkins Gas Furnace problem are shown in Fig. 9.1 and Table 9.1, respectively. The best result 0.2407 is achieved for the weighted flexible OR-type neuro-fuzzy system (Section 5.7) with the H-function generated by the algebraic triangular norms.

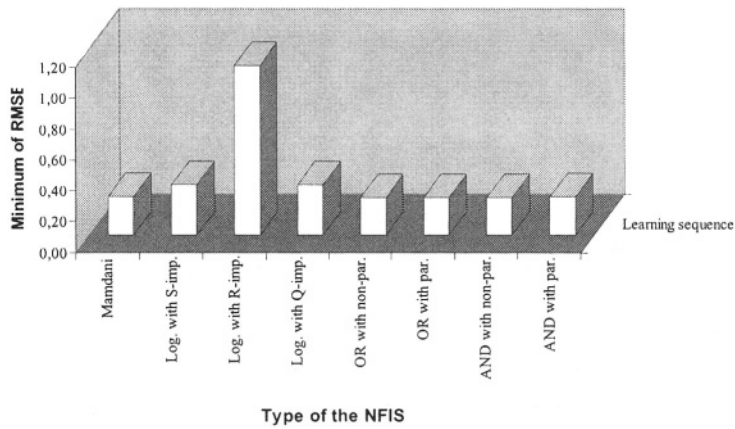

Fig. 9.1. Comparison charts in the Modeling of Box and Jenkins Gas Furnace problem for the minimum of RMSE

Table 9.1 Comparison table (Modeling of Box and Jenkins Gas Furnace problem)

MODELING OF BOX AND JENKINS GAS FURNACE PROBLEM		
Method	**No of inputs / rules**	**Training RMSE**
Tong [96]	2/19	0.6848
Pedrycz [71]	2/81	0.5656
Xu and Lu [105]	2/25	0.5727
Box and Jenkins [5]	6/-	0.4494
Sugeno and Yasukawa [92]	3/6	0.4358
Wang and Langari [101]	6/2	0.2569
Sugeno and Tanaka [91]	6/2	0.2607
Lin and Cunningham [55]	5/4	0.2664
Kim et al [43]	6/2	0.2345
Kim et al [44]	6/2	0.2190
Delgado et al [16]	2/4	0.4100
Yoshinari [112]	2/6	0.5460
Rutkowski and Cpałka [85]	6/4	0.2416
our result (N=4)	**6/4**	**0.2407**

Chemical Plant problem

The comparison chart and the comparison table for the Chemical Plant problem are shown in Fig. 9.2 and Table 9.2, respectively. The best result 0.0042 is obtained for the weighted compromise system (Section 6.5) with the algebraic t-norm. A very similar results 0.0044 is achieved for the weighted flexible systems (Section 5.7) with the H-function generated by the algebraic triangular norms.

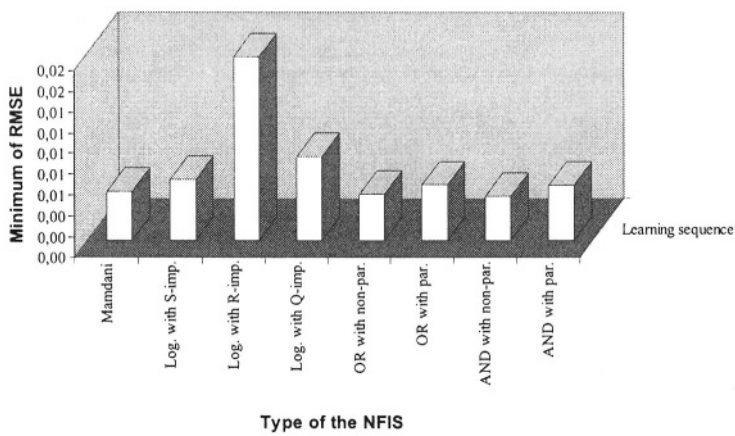

Fig. 9.2. Comparison charts in the Chemical Plant problem for the minimum of RMSE

Table 9.2 Comparison table (Chemical Plant problem)

CHEMICAL PLANT PROBLEM	
Method	**Training RMSE**
Lin and Cunningham [55]	0.0079
Pal and Chakraborty [68]	0.0092
our result (N=6)	**0.0042**

Glass Identification problem

The comparison chart and the comparison table for the Glass Identification problem are shown in Fig. 9.3 and Table 9.3, respectively. The best result 98.44% is achieved for the following systems:

- weighted compromise system (Section 6.5) with the Zadeh, algebraic, Dombi and Yager triangular norms,
- soft compromise system (Section 6.4) with the Zadeh, Dombi and Yager triangular norms,
- weighted flexible system (Section 5.7) with the H-function generated by the Zadeh, algebraic, Yager and Dombi triangular norms,
- soft flexible system (Section 5.6) with the H-function generated by the Dombi and Yager triangular norms,
- Mamdani-type system (Section 7.3) with the min and product "engineering implications",
- Logical-type system (Section 8.3) with the Kleene-Dienes and Reichenbach fuzzy implications.

Table 9.3 Comparison table (Glass Identification problem)

GLASS IDENTIFICATION PROBLEM	
Method	**Testing Acc.**
Dong and Kothari (IG) [17]	92.86
Dong and Kothari (IG+LA) [17]	93.09
Dong and Kothari (GR) [17]	92.86
Dong and Kothari (GR+LA) [17]	93.10
Rutkowski and Cpałka [85]	93.75
our result (N=2)	**98.44**

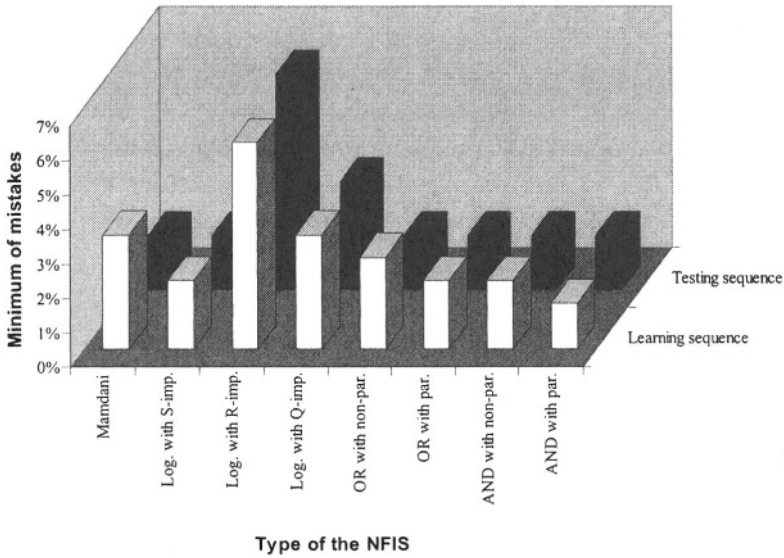

Fig. 9.3. Comparison charts in the Glass Identification problem for minimum of mistakes

Ionosphere problem

The comparison chart and the comparison table for the Ionosphere problem are shown in Fig. 9.4 and Table 9.4, respectively. The best result 94.29 is achieved for the following neuro-fuzzy systems:

- weighted flexible system (Section 5.7) with the H-function generated by algebraic triangular norms,
- weighted compromise system (Section 6.5) with the algebraic triangular norms.

Table 9.4 Comparison table (Ionosphere problem)

IONOSPHERE PROBLEM	
Method	**Testing Acc.**
González and Pérez (CN2) [27]	83.96
González and Pérez (C4.5) [27]	86.50
González and Pérez (LVQ) [27]	87.90
González and Pérez (SLAVE) [27]	82.40
González and Pérez (SLAVE with OR and AND) [27]	85.40
González and Pérez (SLAVE with OR, AND and Generalization) [27]	90.90
González and Pérez (SLAVE with all operators) [27]	91.80
Rutkowski and Cpałka [85]	92.38
our result (N=2)	**94.29**

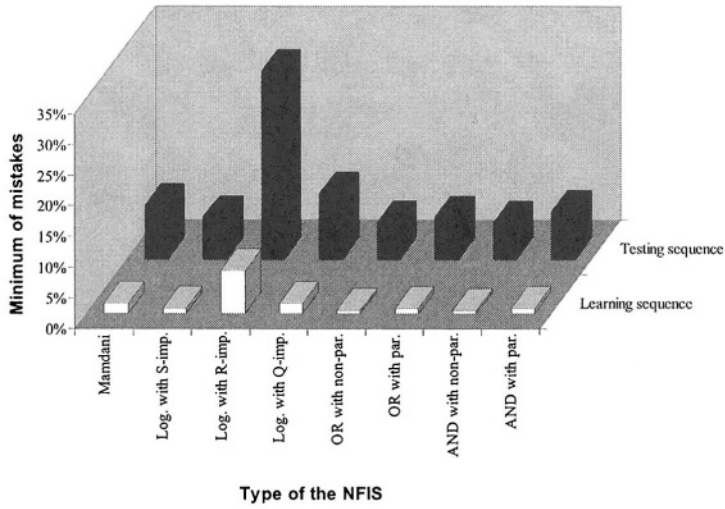

Fig. 9.4. Comparison charts in the Ionosphere problem for minimum of mistakes

Iris problem

The comparison chart and the comparison table for the Iris problem are shown in Fig. 9.5 and Table 9.5, respectively. The best result 97.78% is obtained for the following systems:

- weighted flexible system (Section 5.7) with the H-function generated by the Dombi and Yager triangular norms,
- weighted compromise system (Section 6.5) with the Dombi and Yager triangular norms.

Table 9.5 Comparison table (Iris problem)

IRIS PROBLEM	
Method	**Testing Acc.**
González and Pérez (CN2) [27]	94.16
González and Pérez (C4.5) [27]	92.70
González and Pérez (LVQ) [27]	95.70
González and Pérez (SLAVE) [27]	95.70
González and Pérez (SLAVE with OR and AND) [27]	95.70
González and Pérez (SLAVE with OR, AND and Generalization) [27]	95.70
González and Pérez (SLAVE with all operators) [27]	95.70
Rutkowski and Cpałka [85]	97.78
our result (N=2)	**97.78**

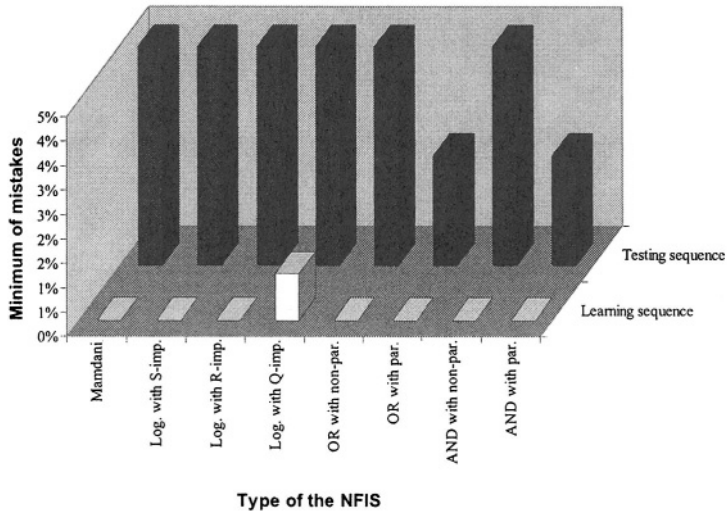

Fig. 9.5. Comparison charts in the Iris problem for minimum of mistakes

Modeling of Static Nonlinear Function (HANG) problem

The comparison chart and the comparison table for the Modeling of Static Nonlinear Function problem are shown in Fig. 9.6 and Table 9.6, respectively. The best result 0.0485 is achieved for the weighted flexible system (Section 5.7) with the H-function generated by the algebraic triangular norms.

Table 9.6 Comparison table (Modeling of Static Nonlinear Function problem)

MODELING OF STATIC NONLINEAR FUNCTION (HANG) PROBLEM		
Method	**No of rules**	**Training RMSE**
Kim et al (SI) [43]	3	0.0935
Sugeno and Yasukawa [92]	6	0.2810
Kim et al (SI with FMAIUP) [44]	2	0.1403
Rutkowski and Cpałka [85]	4	0.0739
our result (N=5)	**5**	**0.0485**

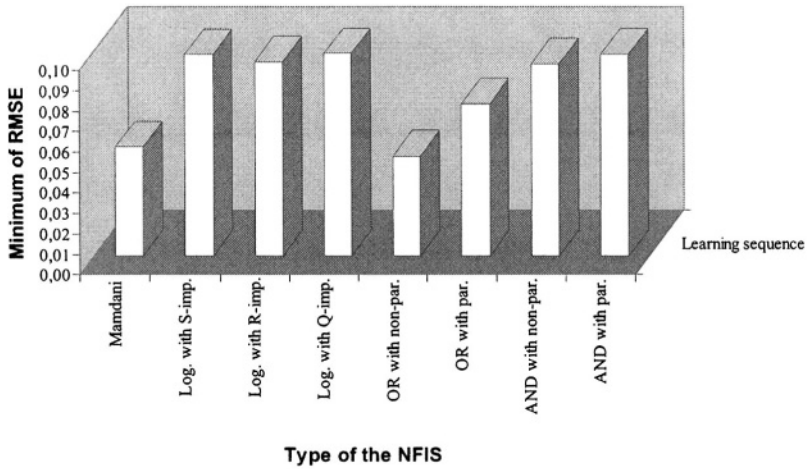

Fig. 9.6. Comparison charts in the Modeling of Static Nonlinear Function (HANG) problem for the minimum of RMSE

The Three Monk's problems

The comparison chart and the comparison table for the Three Monk's problems are shown in Fig. 9.8-9.10 and Tables 9.8-9.10, respectively. For the Monk1 problem the best result 100% is achieved for the following systems:

- weighted flexible system (Section 5.7) with the H-function generated by the Dombi, Yager and algebraic triangular norms,
- weighted compromise system (Section 6.5) with the Dombi, Yager and algebraic triangular norms,
- Mamdani-type system (Section 7.3) with the algebraic triangular norms.

For the Monk2 problem the best result 100% is obtained for the following systems:

- weighted flexible system (Section 5.7) with the H-function generated by the Dombi and Yager triangular norms,
- weighted compromise system (Section 6.5) with the Dombi and Yager triangular norms.

For the Monk3 problem the best result 100% is obtained for the same systems as in the case of the Monk1 problem.

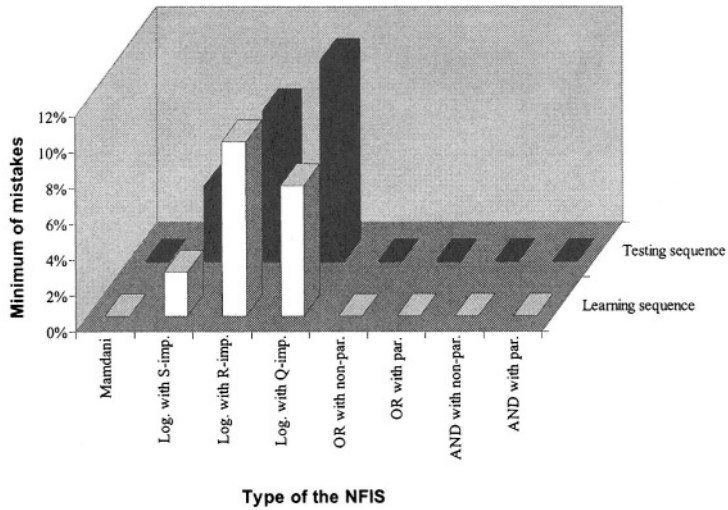

Fig. 9.7. Comparison charts in the Monk1 problem for the minimum of mistakes

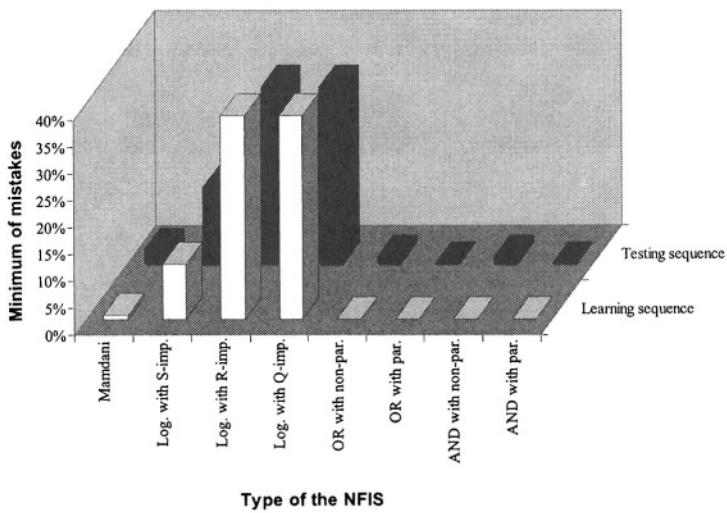

Fig. 9.8. Comparison charts in the Monk2 problem for the minimum of mistakes

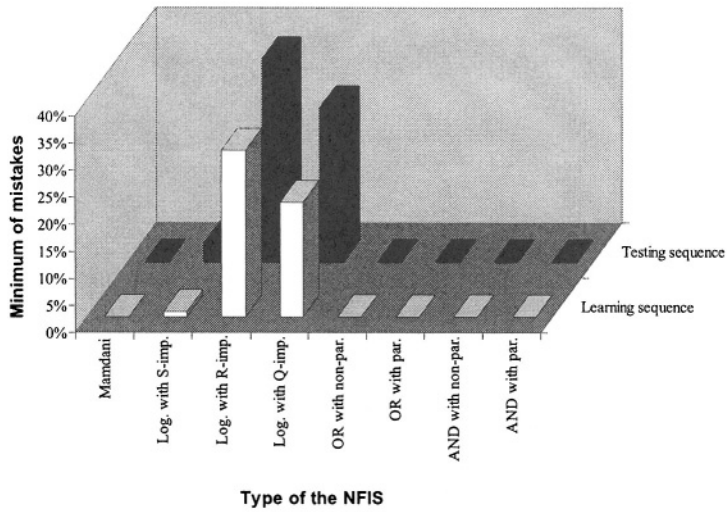

Fig. 9.9. Comparison charts in the Monk3 problem for the minimum of mistakes

Table 9.7 Comparison table (Monk1 problem)

MONK1 PROBLEM		
Method	**Training Acc.**	**Testing Acc.**
Dong and Kothari (IG) [17]	100.00	86.75
Dong and Kothari (IG+LA) [17]	100.00	100.00
Dong and Kothari (GR) [17]	100.00	84.72
Dong and Kothari (GR+LA) [17]	100.00	100.00
Rutkowski and Cpałka [85]	100.00	100.00
our result (N=4)	**100.00**	**100.00**

Table 9.8 Comparison table (Monk2 problem)

MONK2 PROBLEM		
Method	**Training Acc.**	**Testing Acc.**
Dong and Kothari (IG) [17]	80.47	63.42
Dong and Kothari (IG+LA) [17]	82.25	63.65
Dong and Kothari (GR) [17]	81.06	64.81
Dong and Kothari (GR+LA) [17]	81.66	66.51
Rutkowski and Cpałka [85]	95.86	88.19
our result (N=16)	**100.00**	**100.00**

Table 9.9 Comparison table (Monk3 problem)

MONK3 PROBLEM		
Method	**Training Acc.**	**Testing Acc.**
Dong and Kothari (IG) [17]	100.00	94.40
Dong and Kothari (IG+LA) [17]	100.00	94.40
Dong and Kothari (GR) [17]	100.00	90.28
Dong and Kothari (GR+LA) [17]	100.00	90.74
Rutkowski and Cpałka [85]	100.00	100.00
our result (*N*=3)	**100.00**	**100.00**

Nonlinear Dynamic Plant problem

The comparison chart and the comparison table for the Nonlinear Dynamic Plant problem are shown in Fig. 9.10 and Table 9.10, respectively. The best result 0.0101 is achieved for the weighted compromise system (Section 6.5) with the algebraic triangular norms. A similar result 0.0107 is obtained for the weighted flexible system (Section 5.7) with the H-function generated by the algebraic triangular norms.

Table 9.10 Comparison table (Nonlinear Dynamic Plant problem)

NONLINEAR DYNAMIC PLANT PROBLEM			
Method	**No of rules**	**Training RMSE**	**Testing RMSE**
Wang and Yen [104]	40	0.0182	0.0263
Wang and Yen [104]	28	0.0182	0.0245
Wang and Yen [103]	36	0.0053	0.0714
Wang and Yen [103]	23	0.0057	0.0436
Wang and Yen [103]	36	0.0014	0.0539
Wang and Yen [103]	24	0.0014	0.0253
Yen and Wang [110]	25	0.0152	0.0202
Yen and Wang [110]	20	0.0261	0.0155
Setnes and Roubos [89]	7	0.1265	0.0346
Setnes and Roubos [89]	7	0.0548	0.0221
Setnes and Roubos [89]	5	0.0762	0.0500
Setnes and Roubos [89]	5	0.0274	0.0187
Setnes and Roubos [89]	4	0.0346	0.0217
Roubos and Setnes [75]	5	0.0700	0.0539
Roubos and Setnes [75]	5	0.0374	0.0243
Roubos and Setnes [75]	5	0.0288	0.0187
Rutkowski and Cpałka [85]	5	0.0328	0.0211
our result (*N*=6)	**6**	**0.0193**	**0.0101**

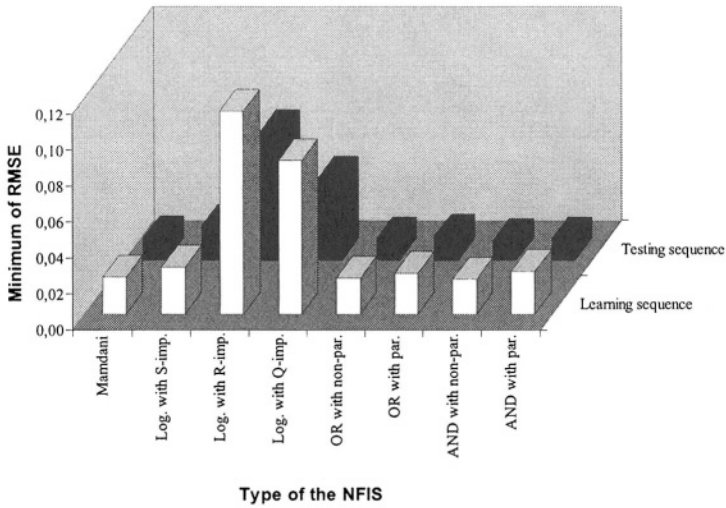

Fig. 9.10. Comparison charts in the Nonlinear Dynamic Plant problem for the minimum of RMSE

Pima Indians Diabetes problem

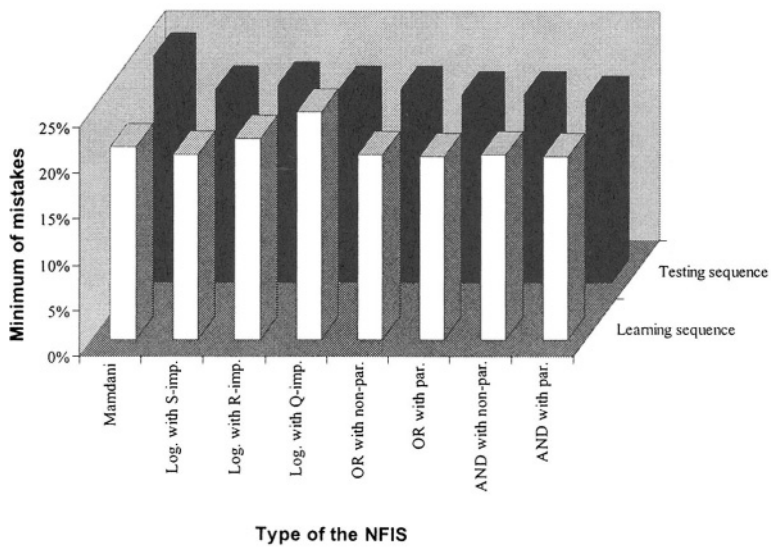

Fig. 9.11. Comparison charts in the Pima Indians Diabetes problem for the minimum of mistakes

The comparison chart and the comparison table for the Pima Indians Diabetes problem are shown in Fig. 9.11 and Table 9.11, respectively. The best result 80.2% is achieved for the weighted compromise system (Section 6.5) with the Yager triangular norms.

Table 9.11 Comparison table (Pima Indians Diabetes problem)

PIMA INDIANS DIABETES PROBLEM	
Method	**Testing Acc.**
Smith et al [98]	76.0
Ster and Dobnikar (Logdisc) [90]	77.7
Ster and Dobnikar (IncNet) [90]	77.6
Ster and Dobnikar (DIPOL92) [90]	77.6
Ster and Dobnikar (LDA) [90]	77.5
Ster and Dobnikar (SMART) [90]	76.8
Ster and Dobnikar (ASI) [90]	76.6
Ster and Dobnikar (FDA) [90]	76.5
Ster and Dobnikar (BP) [90]	76.4
Ster and Dobnikar (LVQ) [90]	75.8
Ster and Dobnikar (RBF) [90]	75.7
Ster and Dobnikar (LFC) [90]	75.8
Ster and Dobnikar (NB) [90]	75.3
Ster and Dobnikar (SNB) [90]	75.4
Ster and Dobnikar (DB-CART) [90]	74.4
Ster and Dobnikar (ASR) [90]	74.3
Ster and Dobnikar (CART: 11nodes) [90]	73.7
Ster and Dobnikar (C4.5) [90]	73.0
Ster and Dobnikar (CART) [90]	72.8
Ster and Dobnikar (Kohonen SOM) [90]	72.2
Ster and Dobnikar (kNN) [90]	71.9
Rutkowski and Cpałka [85]	78.6
our result (N=2)	**80.2**

Rice Taste problem

The comparison chart and the comparison table for the Rice Taste problem are shown in Fig. 9.12 and Table 9.12, respectively. The best result 0.0369 is obtained for the weighted compromise system (Section 6.5) with the Zadeh triangular norms.

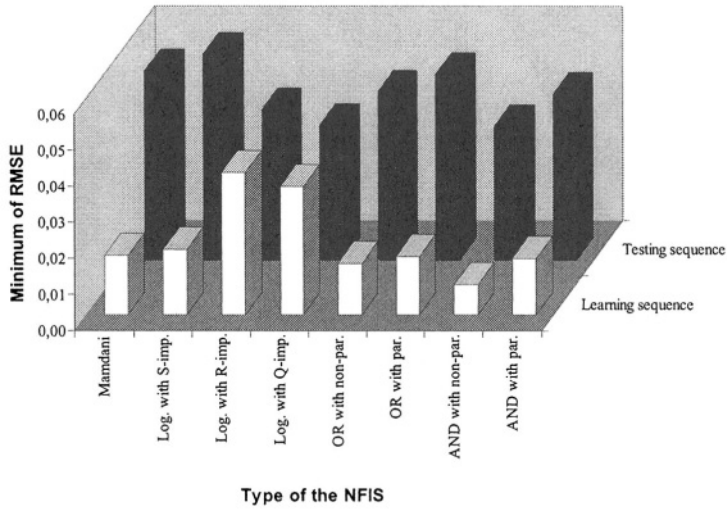

Fig. 9.12. Comparison charts in the Rice Taste problem for the minimum of RMSE

Table 9.12 Comparison table (Rice Taste problem)

RICE TASTE PROBLEM		
Method	**Training RMSE**	**Testing RMSE**
Ishibuchi et al (Heuristic. K=4) [31]	0.0800	0.0624
Ishibuchi et al (Heuristic. K=5) [31]	0.0414	0.0764
Ishibuchi et al (Learning. K=3) [31]	0.0339	0.0518
Ishibuchi et al (Learning. K=5) [31]	0.0205	0.1887
Ishibuchi et al (Hybrid. K=2) [31]	0.0434	0.0563
Ishibuchi et al (Hybrid. K=5) [31]	0.0154	0.0812
Ishibuchi et al (NN 5 20 1) [31]	0.0330	0.0547
Ishibuchi et al (NN 5 30 1) [31]	0.0333	0.0542
Nozaki et al (Heuristic. α=50. K=2) [66]	0.0752	0.0872
Nozaki et al (Heuristic. α=10. K=3) [66]	0.0542	0.0698
Nozaki et al (Heuristic. α=5. K=4) [66]	0.0429	0.0600
Nozaki et al (Heuristic. α=10. K=4) [66]	0.0423	0.0606
Nozaki et al (Heuristic (α=5. K=5) [66]	0.0372	0.0762
Nozaki et al (Learning. e=200. K=2) [66]	0.0420	0.0560
Nozaki et al (Learning. e=500. K=2) [66]	0.0400	0.0611
Nozaki et al (Learning. e=100. K=3) [66]	0.0339	0.0516
Nozaki et al (Learning. e=500. K=3) [66]	0.0283	0.0658
Nozaki et al (Learning. e=200. K=4) [66]	0.0248	0.1183
Nozaki et al (Learning. e=500. K=4) [66]	0.0179	0.1200
Nozaki et al (Learning. e=200. K=5) [66]	0.0163	0.1883
Nozaki et al (Learning. e=400. K=5) [66]	0.0126	0.1887
Rutkowski and Cpalka [85]	0.0155	0.0427
our result (*N*=5)	**0.0081**	**0.0369**

Wine Recognition problem

The comparison chart and the comparison table for the Wine Recognition problem are shown in Fig. 9.13 and Table 9.13, respectively. The best result 100% is achieved for the following systems:

- weighted flexible system (Section 5.7) with the H-function generated by the Zadeh, Yager, Dombi and algebraic triangular norms,
- soft flexible system (Section 5.6) with the Zadeh, Yager, Dombi and algebraic triangular norms,
- weighted compromise system (Section 6.5) with the Zadeh, Yager, Dombi and algebraic triangular norms,
- soft compromise system (Section 6.4) with the Zadeh, Yager, Dombi and algebraic triangular norms,
- logical-type system (Section 8.3) with the Reichenbach implication.

Table 9.13 Comparison table (Wine Recognition problem)

WINE RECOGNITION PROBLEM	
Method	**Testing Acc.**
Corcoran and Sen [13]	100.0
Ishibuchi et al [30]	99.4
González and Pérez (SLAVE) [27]	89.8
González and Pérez (SLAVE with OR and AND) [27]	90.4
González and Pérez (SLAVE with OR, AND and Generalization) [27]	90.4
González and Pérez (SLAVE with all operators) [27]	93.8
Rutkowski and Cpałka [85]	100.0
our result (N=2)	**100.0**

The comparison chart and the comparison table for the Wisconsin Breast Cancer problem are shown in Fig. 9.14 and Table 9.14, respectively. The best result 98.5% is obtained for the following systems:

- weighted flexible system (Section 5.7) with the H-function generated by the Zadeh, Yager, Dombi and algebraic triangular norms,
- weighted compromise system (Section 6.5) with the Zadeh, Yager, Dombi and algebraic triangular norms,
- Mamdani-type system (Section 7.3) with the min and product "engineering implications",
- logical-type system (Section 8.3) with the Kleene-Dienes and Reichenbach implications.

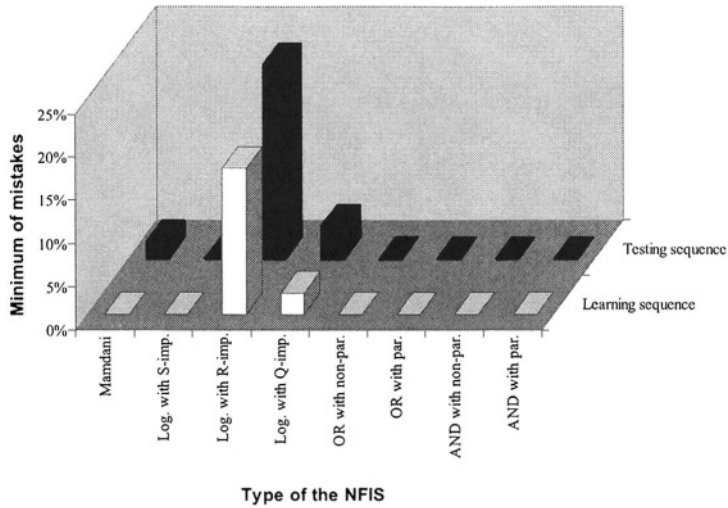

Fig. 9.13. Comparison charts in the Wine Recognition problem for the minimum of mistakes

Wisconsin Breast Cancer problem

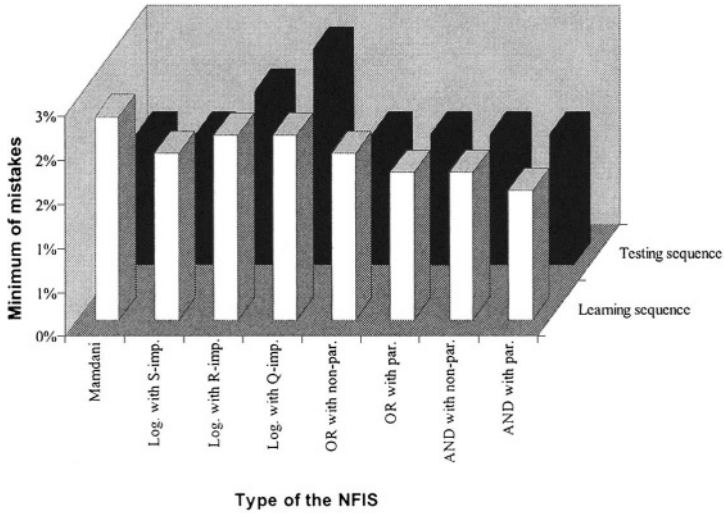

Fig. 9.14. Comparison charts in the Wisconsin Breast Cancer problem for minimum of mistakes

Table 9.14 Comparison table (Wisconsin Breast Cancer problem)

WISCONSIN BREAST CANCER PROBLEM	
Method	**Testing Acc.**
Dong and Kothari (IG) [17]	94.3
Dong and Kothari (IG+LA) [17]	94.1
Dong and Kothari (GR) [17]	94.7
Dong and Kothari (GR+LA) 17]	94.8
Ster and Dobnikar (Fisher LDA) [90]	96.8
Ster and Dobnikar (MLP+BP) [09]	96.7
Ster and Dobnikar (LVQ) [90]	96.6
Ster and Dobnikar (Bayes) [90]	96.6
Ster and Dobnikar (Naive Bayes) [90]	96.4
Ster and Dobnikar (LDA) [90]	96.0
Ster and Dobnikar (LFC, ASI, ASR) [90]	94.4-95.6
Ster and Dobnikar (Quadratic DA) [90]	34.5
Rutkowski and Cpałka [85]	96.6
our result (*N*=2)	**98.5**

9.3. SUMMARY AND DISCUSSION

In this book we proposed a new class of operators called quasi-triangular norms. The operators have been applied to design a new class of neuro-fuzzy systems. At the design stage we do not assume a concrete form of fuzzy inference. Therefore, we applied H-operators and trained flexible neuro-fuzzy systems. For the first time in literature the inference of neuro-fuzzy systems has been established in the process of learning: either the Mamdani-type represented by a t-norm or the logical-type represented by an S-implication. When the learning process is finished, the trained neuro-fuzzy systems based on H-operators take simpler forms based on t-norms or t-conorms. In Chapters 5-8 we derived and tested various new neuro-fuzzy structures. They are characterized as follows:

(i) The OR-type system is "more Mamdani" $(0 < \nu < 0.5)$ or "more logical" $(0.5 < \nu < 1)$. In the process of learning only one type of system is established ($\nu = 0$ or $\nu = 1$).

(ii) The AND-type system is a combination, controlled by parameter $\lambda \in [0,1]$, of Mamdani-type and logical-type systems. In the process of learning only one type of system is established ($\lambda = 0$ or $\lambda = 1$).

(iii) The OR-type is equivalent to the AND-type system if $\nu = \lambda = 0$ (Mamdani-type) or $\nu = \lambda = 1$ (logical-type).

(iv) Parameters ν and λ after learning do not take any value in the interval (0,1) because only for $\nu = \lambda = 0$ or $\nu = \lambda = 1$ systems presented in Chapters 5 and 6 are defined by triangular norms.

(v) Comparing various neuro-fuzzy systems presented in Chapters 5, 7 and 8 we see that when parameter ν in the OR-type system (Chapter 5) takes the final value equal 0 (the system is of a Mamdani-type) then the performance (RMSE or mistakes [%]) of the Mamdani-type system presented in Chapter 7 is better than the performance of the logical-type system presented in Chapter 8. This fact is illustrated in Table 9.15 on the HANG problem.

Table 9.15 Performance comparison for the HANG problem

	OR-type neuro-fuzzy system with algebraic triangular norms and with ν=0 after learning (Chapter 5)	Mamdani-type neuro-fuzzy system with product implication (Chapter 7)	Logical-type neuro-fuzzy system with Goguen implication (Chapter 8)
Performance (RMSE)	0.0485	0.0529	0.0948

Similarly, when parameter ν in the OR-type system (Chapter 5) takes the final value equal 1 (the system is of a logical-type) then the performance (RMSE or mistakes [%]) of the logical-type system presented in Chapter 8 is better than the performance of the Mamdani-type system given in Chapter 7. This fact is illustrated in Table 9.16 on the Ionosphere problem.

Table 9.16 Performance comparison for the Ionosphere problem

	OR-type neuro-fuzzy system with algebraic triangular norms and with ν=1 after learning (Chapter 5)	Mamdani-type neuro-fuzzy system with product implication (Chapter 7)	Logical-type neuro-fuzzy system with Reichenbach implication (Chapter 8)
Performance (%)	5.71	8.57	6.67

This confirms that the OR-type neuro-fuzzy systems produce in the process of learning a correct system type. The same conclusion can be drawn analysing the performance of the AND-type system.

(vi) The most influential parameters are certainty weights $w_{i,k}^{\tau} \in [0,1]$, $i = 1,\ldots,n$, $k = 1,\ldots,N$ and $w_k^{agr} \in [0,1]$, $k = 1,\ldots,N$. They significantly improve the performance of the system in the process of learning.

(vii) The influence of soft parameters $\alpha^{\tau} \in [0,1]$, $\alpha^{I} \in [0,1]$, $\alpha^{agr} \in [0,1]$ on the performance of the system varies depending on the problem. In many cases soft parameters improve the performance of neuro-fuzzy systems even if certainty weights are not simultaneously used.

(viii) The Dombi and Yager parameterised triangular norms sometimes improve the performance of neuro-fuzzy systems because they have additional parameters to be tuned during the learning process.

(ix) The best results are obtained for neuro-fuzzy systems based on all the flexibility parameters described in Chapter 4. The weighted flexible systems (Section 5.7) and the weighted compromise systems (Section 6.5) are superior to other systems in most cases.

(x) In most cases neuro-fuzzy systems based on S-implications give better results than R and Q-implications.

As we have mentioned, the OR-type systems, presented in Chapter 5, and the AND-type systems, presented in Chapter 6, produce the same type of inference (Mamdani or logical) in the process of learning. From simulations it follows that (see Table 9.17):

a) Neuro-fuzzy systems presented in Chapters 5 and 6 become of the Mamdani-type ($\nu = \lambda \simeq 0$) for examples 1, 2, 6, 7, 8, 9, 10, 12.

b) Neuro-fuzzy systems presented in Chapters 5 and 6 become of the logical-type ($\nu = \lambda = 1$) for examples 3, 4, 5, 11, 13, 14.

Table 9.17 The results of simulations

No	Name of simulation	Type of inference model
1	Box and Jenkins Gas Furnace	Mamdani ($\nu = \lambda = 0$)
2	Chemical Plant	Mamdani ($\nu = \lambda = 0$)
3	Glass Identification	logical ($\nu = \lambda = 1$)
4	Ionosphere	logical ($\nu = \lambda = 1$)
5	Iris	logical ($\nu = \lambda = 1$)
6	Modeling of Static Nonlinear Function (HANG)	Mamdani ($\nu = \lambda = 0$)
7	Monk 1	Mamdani ($\nu = \lambda = 0$)
8	Monk 2	Mamdani ($\nu = \lambda = 0$)
9	Monk 3	Mamdani ($\nu = \lambda = 0$)
10	Nonlinear Dynamic Plant	Mamdani ($\nu = \lambda = 0$)
11	Pima Indians Diabetes	logical ($\nu = \lambda = 1$)
12	Rice Taste	Mamdani ($\nu = \lambda = 0$)
13	Wine Recognition	logical ($\nu = \lambda = 1$)
14	Wisconsin Breast Cancer	logical ($\nu = \lambda = 1$)

We conclude that the Mamdani-type systems are more suitable to approximation problems, whereas the logical-type systems may be preferred for classification problems.

It would be interesting to verify if flexible neuro-fuzzy systems studied in this book are universal approximators (see [46, 47, 100, 102]) which are capable of uniformly approximating any nonlinear function over $\mathbf{U} \subset \mathbf{R}^n$ to any degree of accuracy if $\mathbf{U}$ is a compact set. It is well known that the Mamdani-type systems are universal approximators (see e.g. [100]). On the other hand, it has been stated (see [47], page 140) that generalizations of the standard fuzzy rule based modeling methodology have a universal approximation property. This explains a very good performance of flexible Mamdani-type neuro-fuzzy systems studied in this book. The universal approximation property for logical neuro-fuzzy systems remains an open problem.

Another convergence property is possessed by probabilistic neural networks [77, 78, 79, 86, 87, 88] applied to system modeling and classification. In the future research it would be also interesting to investigate relations between flexible neuro-fuzzy systems and probabilistic neural networks. It has been observed in [100] that Mamdani-type neuro-fuzzy systems are functionally very similar to probabilistic neural networks.

Appendix

In this Appendix we show derivative operators necessary to derive learning procedures for neuro-fuzzy structures presented in Chapters 5 and 6.

a) Basic operators

Addition operator

$$y = \sum_{i=1}^{n} x_i \tag{1}$$

$$\frac{\partial y}{\partial x_i} = 1 \tag{2}$$

Multiplication operator

$$y = \prod_{i=1}^{n} x_i \tag{3}$$

$$\frac{\partial y}{\partial x_i} = \prod_{\substack{j=1 \\ j \neq i}}^{n} x_j \tag{4}$$

Division operator

$$y = \frac{a}{b} \tag{5}$$

$$\frac{\partial y}{\partial a} = \frac{1}{b} \tag{6}$$

$$\frac{\partial y}{\partial b} = -\frac{a}{b^2} \tag{7}$$

Minimum operator

$$y = \min_{i=1\ldots n}\{x_i\} \tag{8}$$

$$\frac{\partial y}{\partial x_i} = \begin{cases} 1 & \text{for} \quad x_i = y \\ 0 & \text{for} \quad x_i \neq y \end{cases} \tag{9}$$

Maximum operator

$$y = \max_{i=1\ldots n}\{x_i\} \tag{10}$$

$$\frac{\partial y}{\partial x_i} = \begin{cases} 1 & \text{for} \quad x_i = y \\ 0 & \text{for} \quad x_i \neq y \end{cases} \tag{11}$$

Compromise operator

$$\tilde{N}_v(a) = (1 - f_z(v))(1 - a) + f_z(v)a \tag{12}$$

$$\frac{\partial \tilde{N}_v(a)}{\partial a} = 2f_z(v) - 1 \tag{13}$$

$$\frac{\partial \tilde{N}_v(a)}{\partial v} = (2a - 1)\frac{\partial f_z(v)}{\partial v} \tag{14}$$

Arithmetic average operator

$$\text{avg}(a_1, a_2, \ldots, a_n) = \frac{1}{n}\sum_{i=1}^{n} a_i \tag{15}$$

$$\frac{\partial \,\text{avg}(a_1, a_2, \ldots, a_n)}{\partial a_i} = \frac{1}{n} \tag{16}$$

Aggregation operator

$$G(a_1, a_2\ ; \phi) = (1 - f_z(\phi))a_1 + f_z(\phi)a_2 \tag{17}$$

$$\frac{\partial G(a_1, a_2\ ; \phi)}{\partial a_1} = 1 - f_z(\phi) \tag{18}$$

$$\frac{\partial G(a_1, a_2\ ; \phi)}{\partial a_2} = f_z(\phi) \tag{19}$$

$$\frac{\partial G(a_1,a_2;\phi)}{\partial \phi} = -(a_1 - a_2)\frac{\partial f_z(\phi)}{\partial \phi} \tag{20}$$

Defuzzifier operator

$$\mathrm{def}(a_1,a_2,\ldots,a_n;w_1,w_2,\ldots,w_n) = \mathrm{def}(\mathbf{a};\mathbf{w}) = \frac{\sum_{i=1}^{n} w_i a_i}{\sum_{i=1}^{n} a_i} \tag{21}$$

$$\frac{\partial\, \mathrm{def}(\mathbf{a};\mathbf{w})}{\partial a_j} = (w_j - \mathrm{def}(\mathbf{a};\mathbf{w}))\frac{1}{\sum_{i=1}^{n} a_i} \tag{22}$$

$$\frac{\partial\, \mathrm{def}(\mathbf{a};\mathbf{w})}{\partial w_j} = \left(a_j - \mathrm{def}(\mathbf{a};\mathbf{w})\frac{\partial a_j}{\partial w_j}\right)\frac{1}{\sum_{i=1}^{n} a_i} \tag{23}$$

b) Membership functions

Gaussian membership function

$$\mu_A(x) = \exp\left(-\left(\frac{x-\bar{x}}{\sigma}\right)^2\right) \tag{24}$$

$$\frac{\partial \mu_A(x)}{\partial x} = -\mu_A(x)\frac{2(x-\bar{x})}{\sigma^2} \tag{25}$$

$$\frac{\partial \mu_A(x)}{\partial \bar{x}} = \mu_A(x)\frac{2(x-\bar{x})}{\sigma^2} \tag{26}$$

$$\frac{\partial \mu_A(x)}{\partial \sigma} = \mu_A(x)\frac{2(x-\bar{x})^2}{\sigma^3} \tag{27}$$

Triangular membership function

$$\mu_A(x) = \begin{cases} 0 & \text{for} \quad x \le a \text{ or } x \ge c \\ \dfrac{x-a}{b-a} & \text{for} \quad a \le x \le b \\ \dfrac{c-x}{c-b} & \text{for} \quad b \le x \le c \end{cases} \tag{28}$$

$$\frac{\partial \mu_A(x)}{\partial x} = \begin{cases} 0 & \text{for} \quad x < a \text{ or } x > c \\ \frac{1}{2(b-a)}\varepsilon & \text{for} \quad x = a \\ \frac{1}{b-a}\varepsilon & \text{for} \quad a < x < b \\ \frac{c-2b+a}{2(c-b)(b-a)}\varepsilon & \text{for} \quad x = b \\ -\frac{1}{c-b}\varepsilon & \text{for} \quad b < x < c \\ -\frac{1}{2(c-b)}\varepsilon & \text{for} \quad x = c \end{cases} \tag{29}$$

$$\frac{\partial \mu_A(x)}{\partial a} = \begin{cases} 0 & \text{for} \quad x \le a \text{ or } x > b \\ \frac{1}{2(b-a)}\varepsilon & \text{for} \quad x = b \\ \frac{x-a}{(b-a)^2}\varepsilon & \text{for} \quad a \le x < b \end{cases} \tag{30}$$

$$\frac{\partial \mu_A(x)}{\partial b} = \begin{cases} 0 & \text{for} \quad x \le a \text{ or } x \ge c \\ \frac{a-x}{(b-a)^2}\varepsilon & \text{for} \quad a \le x < b \\ \frac{a-2b+c}{2(c-b)(b-a)}\varepsilon & \text{for} \quad x = b \\ \frac{c-x}{(c-b)^2}\varepsilon & \text{for} \quad b < x \le c \end{cases} \tag{31}$$

$$\frac{\partial \mu_A(x)}{\partial c} = \begin{cases} 0 & \text{for} \quad x \le b \text{ or } x > c \\ \frac{1}{2(c-b)}\varepsilon & \text{for} \quad x = c \\ \frac{x-b}{(c-b)^2}\varepsilon & \text{for} \quad b < x < c \end{cases} \tag{32}$$

c) Constraints

Constraint for parameters $\nu \in [0,1]$, $\lambda \in [0,1]$, $\alpha^{\tau} \in [0,1]$, $\alpha^{l} \in [0,1]$, $\alpha^{\text{agr}} \in [0,1]$, $w_{i,k}^{\tau} \in [0,1]$, $i = 1,\ldots,n$, $k = 1,\ldots,N$, $w_k^{\text{agr}} \in [0,1]$, $k = 1,\ldots,N$,

$$f_z(x) = \frac{p_{z3}}{1 + \exp(-(p_{z1}x - p_{z2}))} + p_{z4} \tag{33}$$

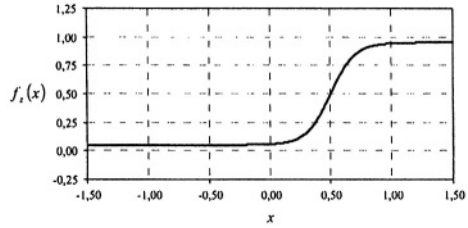

Fig. 15. Plot of function (33)
for $p_{z1}=10$, $p_{z2}=5$, $p_{z3}=0.9$, $p_{z4}=0.05$

$$\frac{\partial f_z(x)}{\partial x}=-\frac{p_{z1}}{p_{z3}}(p_{z3}+p_{z4}-f_z(x))(p_{z4}-f_z(x)) \tag{34}$$

Constraint for parameters $p^{\tau}\in[0,\infty)$, $p^{l}\in[0,\infty)$, $p^{\mathrm{agr}}\in[0,\infty)$

$$f_z(x)=\frac{x}{1+\exp(-(p_{z1}x-p_{z2}))}+p_{z3} \tag{35}$$

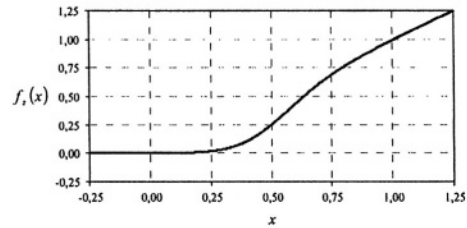

Fig. 16. Plot of function (35)
for $p_{z1}=10$, $p_{z2}=5$, $p_{z3}=0$

$$\frac{\partial f_z(x)}{\partial x}=\frac{-p_{z3}+f_z(x)}{x}(1+p_{z1}(p_{z3}+x-f_z(x))) \tag{36}$$

d) H-functions

Argument of the H-function (given by (5.48))

$$\arg_i(a_i,w_i,v)=G\begin{pmatrix}N(f_{z2}(w_i)N(a_i)),f_{z2}(w_i)a_i\,;\\ f_z(v)\end{pmatrix} \tag{37}$$

$$\frac{\partial \arg_i(a_i,w_i,v)}{\partial a_i}=f_{z2}(w_i) \tag{38}$$

$$\frac{\partial \arg_i(a_i,w_i,v)}{\partial w_i} = (a + f_z(v) - 1)\frac{\partial f_{z2}(w_i)}{\partial w_i} \tag{39}$$

$$\frac{\partial \arg_i(a_i,w_i,v)}{\partial v} = (f_{z2}(w_i) - 1)\frac{\partial f_z(v)}{\partial v} \tag{40}$$

The Zadeh H-function with weighted arguments

$$H^*(\mathbf{a};\mathbf{w},v) = \tilde{N}_v\left(\max_{i=1,\ldots,n}\left\{\tilde{N}_v(\arg_i(a_i,w_i,v))\right\}\right) \tag{41}$$

$$H^*(\mathbf{a};\mathbf{w},v) = \tilde{N}_v(h^*(\mathbf{a};\mathbf{w},v)) \tag{42}$$

where

$$h^*(\mathbf{a};\mathbf{w},v) = \max_{i=1,\ldots,n}\left\{\tilde{N}_v(\arg_i(a_i,w_i,v))\right\} \tag{43}$$

$$\frac{\partial H^*(\mathbf{a};\mathbf{w},v)}{\partial a_i} = \begin{cases} (2f_z(v)-1)^2 \cdot \dfrac{\partial \arg_i(a_i,w_i,v)}{\partial a_i} & \text{for } h^*(\mathbf{a};\mathbf{w},v) = \tilde{N}_v(\arg_i(a_i,w_i,v)) \\ 0 & \text{for } h^*(\mathbf{a};\mathbf{w},v) \neq \tilde{N}_v(\arg_i(a_i,w_i,v)) \end{cases} \tag{44}$$

$$\frac{\partial H^*(\mathbf{a};\mathbf{w},v)}{\partial w_i} = \begin{cases} (2f_z(v)-1)^2 \cdot \dfrac{\partial \arg_i(a_i,w_i,v)}{\partial w_i} & \text{for } h^*(\mathbf{a};\mathbf{w},v) = \tilde{N}_v(\arg_i(a_i,w_i,v)) \\ 0 & \text{for } h^*(\mathbf{a};\mathbf{w},v) \neq \tilde{N}_v(\arg_i(a_i,w_i,v)) \end{cases} \tag{45}$$

$$\begin{aligned} \frac{\partial H^*(\mathbf{a};\mathbf{w},v)}{\partial v} &= \frac{\partial f_z(v)}{\partial v}(2h^*(\mathbf{a};\mathbf{w},v) - 1) + \\ &+ (2f_z(v)-1)\max_{i=1,\ldots,n}\left\{\begin{array}{l}(2f_z(v)-1)\dfrac{\partial \arg_i(a_i,w_i,v)}{\partial v} + \\ + (2\arg_i(a_i,w_i,v)-1)\dfrac{\partial f_z(v)}{\partial v}\end{array}\right\} \end{aligned} \tag{46}$$

The algebraic H-function with weighted arguments

$$H^*(\mathbf{a};\mathbf{w},v) = \tilde{N}_v\left(1 - \prod_{i=1}^{n}(1 - \tilde{N}_v(\arg_i(a_i,w_i,v)))\right) \tag{47}$$

$$H^*(\mathbf{a};\mathbf{w},v) = \tilde{N}_v(h^*(\mathbf{a};\mathbf{w},v)) \tag{48}$$

where

$$h^*(\mathbf{a};\mathbf{w},v) = 1 - \prod_{i=1}^{n}(1 - \tilde{N}_v(\arg_i(a_i,w_i,v))) \tag{49}$$

$$\frac{\partial H^*(\mathbf{a};\mathbf{w},v)}{\partial a_i} = (2f_z(v)-1)^2\frac{\partial \arg_i(a_i,w_i,v)}{\partial a_i}\prod_{\substack{u=1 \\ u\neq i}}^{n}(1 - \tilde{N}_v(\arg_u(a_u,w_u,v))) \tag{50}$$

$$\frac{\partial H^*(\mathbf{a};\mathbf{w},v)}{\partial w_i}=(2f_z(v)-1)^2\frac{\partial \arg_i(a_i,w_i,v)}{\partial w_i}\prod_{\substack{u=1\\u\neq i}}^{n}\left(1-\tilde{N}_v(\arg_u(a_u,w_u,v))\right) \tag{51}$$

$$\begin{aligned}\frac{\partial H^*(\mathbf{a};\mathbf{w},v)}{\partial v}&=(2h^*(\mathbf{a};\mathbf{w},v)-1)\frac{\partial f_z(v)}{\partial v}+\\&+(2f_z(v)-1)\sum_{i=1}^{n}\left(\begin{array}{l}\left(\begin{array}{l}(2\arg_i(a_i,w_i,v)-1)\frac{\partial f_z(v)}{\partial v}+\\+(2f_z(v)-1)\frac{\partial \arg_i(a_i,w_i,v)}{\partial v}\end{array}\right)\cdot\\\cdot\prod_{\substack{u=1\\u\neq i}}^{n}\left(1-\tilde{N}_v(\arg_u(a_u,w_u,v))\right)\end{array}\right)\end{aligned} \tag{52}$$

The Dombi H-function with weighted arguments

$$\ddot{H}^*(\mathbf{a};\mathbf{w},p,v)=\tilde{N}_v\left(1-\left(1+\left(\sum_{i=1}^{n}\left(\tilde{N}_v(\arg_i(a_i,w_i,v))^{-1}-1\right)^{-f_{z1}(p)}\right)^{\frac{1}{f_{z1}(p)}}\right)^{-1}\right) \tag{53}$$

$$\ddot{H}^*(\mathbf{a};\mathbf{w},p,v)=\tilde{N}_v\left(1-\ddot{h}^*(\mathbf{a};\mathbf{w},p,v)\right) \tag{54}$$

$$p\in(0,\infty) \tag{55}$$

where

$$\ddot{h}^*(\mathbf{a};\mathbf{w},p,v)=\left(1+\left(\sum_{i=1}^{n}\left(\tilde{N}_v(\arg_i(a_i,w_i,v))^{-1}-1\right)^{-f_{z1}(p)}\right)^{\frac{1}{f_{z1}(p)}}\right)^{-1} \tag{56}$$

$$\begin{aligned}\frac{\partial \ddot{H}^*(\mathbf{a};\mathbf{w},p,v)}{\partial a_i}&=(2f_z(v)-1)^2\frac{\left(\ddot{h}^*(\mathbf{a};\mathbf{w},p,v)^{-1}-1\right)^{1-f_{z1}(p)}}{\ddot{h}^*(\mathbf{a};\mathbf{w},p,v)^{-2}}\cdot\\&\cdot\frac{\left(\tilde{N}_v(\arg_i(a_i,w_i,v))^{-1}-1\right)^{-f_{z1}(p)-1}}{\tilde{N}_v(\arg_i(a_i,w_i,v))^2}\cdot\\&\cdot\frac{\partial \arg_i(a_i,w_i,v)}{\partial a_i}\end{aligned} \tag{57}$$

$$\begin{aligned}\frac{\partial \ddot{H}^*(\mathbf{a};\mathbf{w},p,v)}{\partial w_i}&=(2f_z(v)-1)^2\frac{\left(\ddot{h}^*(\mathbf{a};\mathbf{w},p,v)^{-1}-1\right)^{1-f_{z1}(p)}}{\ddot{h}^*(\mathbf{a};\mathbf{w},p,v)^{-2}}\cdot\\&\cdot\frac{\left(\tilde{N}_v(\arg_i(a_i,w_i,v))^{-1}-1\right)^{-f_{z1}(p)-1}}{\tilde{N}_v(\arg_i(a_i,w_i,v))^2}\cdot\\&\cdot\frac{\partial \arg_i(a_i,w_i,v)}{\partial w_i}\end{aligned} \tag{58}$$

$$\frac{\partial \ddot{H}^*(\mathbf{a};\mathbf{w},p,\nu)}{\partial p} = \frac{2f_z(\nu)-1}{f_{z1}(p)} \frac{\left(\ddot{h}^*(\mathbf{a};\mathbf{w},p,\nu)^{-1}-1\right)^{1-f_{z1}(p)}}{\ddot{h}^*(\mathbf{a};\mathbf{w},p,\nu)^{-2}} \cdot \left(\sum_{i=1}^{n} \frac{-\ln\left(\tilde{N}_\nu(\arg_i(a_i,w_i,\nu))^{-1}-1\right)}{\left(\tilde{N}_\nu(\arg_i(a_i,w_i,\nu))^{-1}-1\right)^{f_{z1}(p)}} + \frac{\ln\left(\ddot{h}^*(\mathbf{a};\mathbf{w},p,\nu)^{-1}-1\right)}{\left(\ddot{h}^*(\mathbf{a};\mathbf{w},p,\nu)^{-1}-1\right)^{f_{z1}(p)}} \right) \cdot \frac{\partial f_{z1}(p)}{\partial p} \tag{59}$$

$$\frac{\partial \ddot{H}^*(\mathbf{a};\mathbf{w},p,\nu)}{\partial \nu} = \left(1-2\ddot{h}^*(\mathbf{a};\mathbf{w},p,\nu)\right)\frac{\partial f_z(\nu)}{\partial \nu} + (2f_z(\nu)-1)\frac{\left(\ddot{h}^*(\mathbf{a};\mathbf{w},p,\nu)^{-1}-1\right)^{1-f_{z1}(p)}}{\ddot{h}^*(\mathbf{a};\mathbf{w},p,\nu)^{-2}} \cdot \sum_{i=1}^{n} \left(\frac{\left(\tilde{N}_\nu(\arg_i(a_i,w_i,\nu))^{-1}-1\right)^{-f_{z1}(p)-1}}{\tilde{N}_\nu(\arg_i(a_i,w_i,\nu))^2} \cdot \left((2f_z(\nu)-1)\frac{\partial \arg_i(a_i,w_i,\nu)}{\partial \nu} + (2\arg_i(a_i,w_i,\nu)-1)\frac{\partial f_z(\nu)}{\partial \nu} \right) \right) \tag{60}$$

The Yager H-function with weighted arguments

$$\ddot{H}^*(\mathbf{a};\mathbf{w},p,\nu) = \tilde{N}_\nu\left(\min\left\{ 1, \left(\sum_{i=1}^{n} \tilde{N}_\nu(\arg_i(a_i,w_i,\nu))^{f_{z1}(p)} \right)^{\frac{1}{f_{z1}(p)}} \right\} \right) \tag{61}$$

$$\ddot{H}^*(\mathbf{a};\mathbf{w},p,\nu) = \tilde{N}_\nu\left(\min\left\{1, \ddot{h}^*(\mathbf{a};\mathbf{w},p,\nu)\right\}\right) \tag{62}$$

$$\ddot{H}^*(\mathbf{a};\mathbf{w},p,\nu) = \begin{cases} \tilde{N}_\nu\left(\ddot{h}^*(\mathbf{a};\mathbf{w},p,\nu)\right) & \text{for} \quad \ddot{h}^*(\mathbf{a};\mathbf{w},p,\nu) \le 1 \\ \tilde{N}_\nu(1) & \text{for} \quad \ddot{h}^*(\mathbf{a};\mathbf{w},p,\nu) > 1 \end{cases} \tag{63}$$

$$p \in (0,\infty) \tag{64}$$

where

$$\ddot{h}^*(\mathbf{a};\mathbf{w},p,\nu) = \left(\sum_{i=1}^{n} \tilde{N}_\nu(\arg_i(a_i,w_i,\nu))^{f_{z1}(p)} \right)^{\frac{1}{f_{z1}(p)}} \tag{65}$$

$$\frac{\partial \vec{H}^*(\mathbf{a};\mathbf{w},p,\nu)}{\partial a_i}=\begin{cases}(2f_z(\nu)-1)^2\cdot \\ \cdot \vec{h}^*(\mathbf{a};\mathbf{w},p,\nu)^{1-f_{z1}(p)}\cdot \\ \cdot \tilde{N}_\nu(\arg_i(a_i,w_i,\nu))^{f_{z1}(p)-1}\cdot \\ \cdot \dfrac{\partial \arg_i(a_i,w_i,\nu)}{\partial a_i} & \text{for } \vec{h}^*(\mathbf{a};\mathbf{w},p,\nu)\le 1 \\ 0 & \text{for } \vec{h}^*(\mathbf{a};\mathbf{w},p,\nu)>1\end{cases} \tag{66}$$

$$\frac{\partial \vec{H}^*(\mathbf{a};\mathbf{w},p,\nu)}{\partial w_i}=\begin{cases}(2f_z(\nu)-1)^2\cdot \\ \cdot \vec{h}^*(\mathbf{a};\mathbf{w},p,\nu)^{1-f_{z1}(p)}\cdot \\ \cdot \tilde{N}_\nu(\arg_i(a_i,w_i,\nu))^{f_{z1}(p)-1}\cdot \\ \cdot \dfrac{\partial \arg_i(a_i,w_i,\nu)}{\partial w_i} & \text{for } \vec{h}^*(\mathbf{a};\mathbf{w},p,\nu)\le 1 \\ 0 & \text{for } \vec{h}^*(\mathbf{a};\mathbf{w},p,\nu)>1\end{cases} \tag{67}$$

$$\frac{\partial \vec{H}^*(\mathbf{a};\mathbf{w},p,\nu)}{\partial p}=\begin{cases}\dfrac{2f_z(\nu)-1}{f_{z1}(p)}\vec{h}^*(\mathbf{a};\mathbf{w},p,\nu)^{1-f_{z1}(p)}\cdot \\ \cdot\left(\sum_{i=1}^{n}\dfrac{\ln\left(\tilde{N}_\nu(\arg_i(a_i,w_i,\nu))\right)}{\tilde{N}_\nu(\arg_i(a_i,w_i,\nu))^{-f_{z1}(p)}}+\right. \\ \left.-\dfrac{\ln\left(\vec{h}^*(\mathbf{a};\mathbf{w},p,\nu)\right)}{\vec{h}^*(\mathbf{a};\mathbf{w},p,\nu)^{-f_{z1}(p)}}\right)\cdot \\ \cdot\dfrac{\partial f_{z1}(p)}{\partial p} & \text{for } \vec{h}^*(\mathbf{a};\mathbf{w},p,\nu)\le 1 \\ 0 & \text{for } \vec{h}^*(\mathbf{a};\mathbf{w},p,\nu)>1\end{cases} \tag{68}$$

$$\frac{\partial \vec{H}^*(\mathbf{a};\mathbf{w},p,\nu)}{\partial \nu}=\begin{cases}\left(2\vec{h}^*(\mathbf{a};\mathbf{w},p,\nu)-1\right)\dfrac{\partial f_z(\nu)}{\partial \nu}+ \\ +(2f_z(\nu)-1)\dfrac{1}{\vec{h}^*(\mathbf{a};\mathbf{w},p,\nu)^{f_{z1}(p)-1}}\cdot \\ \cdot\sum_{i=1}^{n}\left(\begin{array}{l}\tilde{N}_\nu(\arg_i(a_i,w_i,\nu))^{f_{z1}(p)-1}\cdot \\ \cdot\left((2f_z(\nu)-1)\dfrac{\partial \arg_i(a_i,w_i,\nu)}{\partial \nu}+\right. \\ \left.+(2\arg_i(a_i,w_i,\nu)-1)\dfrac{\partial f_z(\nu)}{\partial \nu}\right)\end{array}\right) & \text{for } \vec{h}^*\begin{pmatrix}\mathbf{a};\\ \mathbf{w},p,\nu\end{pmatrix}\le 1 \\ \dfrac{\partial f_z(\nu)}{\partial \nu} & \text{for } \vec{h}^*\begin{pmatrix}\mathbf{a};\\ \mathbf{w},p,\nu\end{pmatrix}>1\end{cases} \tag{69}$$

This page intentionally left blank

Bibliography

[1] R. A. Aliev and R. R. Aliev. *Soft Computing and its Applications. World Scientific Publishing,* Singapore-New Jersey-London-Hong Kong, 2001.

[2] R. Babuška and M. Setnes. *Data-driven Construction of Transparent Fuzzy Models,* In: H. B. Verbruggen, H.-J. Zimmermann and R. Babuška (Eds.), Fuzzy Algorithms for Control, Kluwer Academic Publishers, Boston 1999.

[3] H. R. Berenji and P. Khedkar. Learning and tuning fuzzy logic controllers through reinforcements, *IEEE Trans. Neural Networks,* vol. 3. pp. 724-740, October 1992.

[4] B. Bouchon-Meunier (Ed.). *Aggregation and Fusion of Imperfect Information,* Physica-Verlag, New York, 1998.

[5] G. E. P. Box and G. M. Jenkins. *Time Series Analysis, Forecasting and Control,* San Francisco, CA: Holden Day, 1970.

[6] Z. Bubnicki. Uncertain variables and their application to decision making, *IEEE Trans. on SMC, Part A: Systems and Humans,* vol. 31, pp. 587-596, 2001.

[7] Z. Bubnicki, *Uncertain Logics, Variables and Systems,* Springer-Verlag, Berlin-London-New York, 2002.

[8] Z. Bubnicki. A unified approach to descriptive and prescriptive concepts in uncertain decision systems, *Systems Analysis Modeling Simulation,* vol. 42, pp. 331-342, 2002.

[9] T. Calvo, G. Mayor and R. Mesiar (Eds.), *Aggregation Operators, New Trends and Applications,* Physica-Verlag, New York, 2002.

[10] T. Calvo and R. Mesiar. Continuous generated associative aggregation operators, *Fuzzy Sets and Systems,* vol. 126, pp. 191-199, 2002.

[11] M. Y. Chen and D. A. Linkens. A systematic neuro-fuzzy modeling framework with application to material property prediction, *IEEE Trans. on Fuzzy Systems,* vol. 31, pp. 781-790, October 2001.

[12] C. C. Chuang, S. F. Su and S. S. Chen. Robust TSK fuzzy modeling for function approximation with outliers, *IEEE Trans. on Fuzzy Systems,* vol. 9, pp. 810-821, December 2001.

[13] A. L. Corcoran and S. Sen. Using real-valued genetic algorithms to evolve rule sets for classification, *Proc. of the 1st IEEE Conf. Evolut. Computat.,* Orlando, FL, June 1994, pp. 120-124.

[14] O. Cordòn, F. Herrera and A. Peregrin. Applicability of the fuzzy operators in the design of fuzzy logic controllers, *Fuzzy Sets and Systems,* vol. 86, No. 1, pp. 15-41, 1997.

[15] E. Czogała and J. Łęski. *Fuzzy and Neuro-Fuzzy Intelligent Systems,* Physica-Verlag Company, Heidelberg, New York, 2000.

[16] M. Delgado, A. F. Gómez-Skarmeta and F. Martin. Fuzzy clustering-based rapid prototyping for fuzzy rule-based modeling, *IEEE Trans. on Fuzzy Systems,* vol. 5, pp. 223-233, May 1997.

[17] M. Dong and R. Kothari. Look-ahead based fuzzy decision tree induction, *IEEE Trans. on Fuzzy Systems,* vol. 9, pp. 461-468, June 2001.

[18] J. C. Duan and F. L. Chung. Cascaded fuzzy neural network based on syllogistic fuzzy reasoning, *IEEE Trans. on Fuzzy Systems,* vol. 9, pp. 293-306, April 2001.

[19] D. Dubois and H. Prade. Weighted minimum and maximum operations in fuzzy set theory, *Information Sciences,* vol. 39, pp. 205-210, 1986.

[20] R. L. Eubank. *Nonparametric Regression and Spline Smoothing,* Marcel Dekker, New York, 1999.

[21] J. C. Fodor. On fuzzy implication operators, *Fuzzy Sets and Systems,* vol. 42, pp. 293-300, 1991.

[22] J. C. Fodor, R. R. Yager and A. Rybalov. Structure of uniforms, *Int. J. Uncertainty Fuzziness Knowledge-based Systems,* vol. 5 (4), pp. 411-427, 1997.

[23] D. B. Fogel. *Evolutionary Computation: Towards a New Philosophy of Machine Intelligence,* IEEE Press, New York, 1995.

[24] R. Fuller. *Introduction to Neuro-Fuzzy Systems, Advances in Soft Computing,* Physica-Verlag, New York, 2000.

[25] A. E. Gaweda and J. M. Żurada. Data-driven design of fuzzy system with relational input partition, *Proc. of the Intern. Conference on Fuzzy Systems, FUZZ-IEEE'2001,* Melbourne, Australia, December 2-5, 2001.

[26] A. E. Gaweda and J. M. Żurada. Fuzzy neural network with relational fuzzy rules," *Proc. of the Intern. Joint Conference on Neural Networks, IJCNN'2000,* vol. 5, pp. 3-8, Como, Italy, July 23-27, 2000.

[27] A. González and R. Pérez. SLAVE: a genetic learning system based on an iterative approach, *IEEE Trans. on Fuzzy Systems,* vol. 7, pp. 176-191, April 1999.

[28] M. B. Gorzałczany. *Computational Intelligence Systems and Applications, Neuro-Fuzzy and Fuzzy Neural Synergisms,* Springer-Verlag, New York, 2002.

[29] K. Hirota. *Industrial Applications of Fuzzy Technology,* Springer Verlag, Tokyo, Berlin, Heidelberg, New York, 1993.

[30] H. Ishibuchi, T. Nakashima and T. Murata. Performance Evaluation of fuzzy classifier systems for multidimensional pattern classification problems, *IEEE Trans. Syst., Man, Cybern., pt. B,* vol. 29, pp. 601-618, October 1999.

[31] H. Ishibuchi and T. Nakashima. Effect of rule weights in fuzzy rule-based classification systems, *IEEE Trans. on Fuzzy Systems,* vol. 9, pp. 506-515, August 2001.

[32] R. Jager. *Fuzzy Logic in Control. Ph.D. Dissertation,* T.U. Delft, 1995.

[33] J. S. R. Jang. ANFIS: Adaptive-network-based fuzzy inference system, *IEEE Trans. Syst., Man, Cybern.,* vol. 23, pp. 665-685, June 1993.

[34] J. S. R. Jang and C. T. Sun. Neuro-fuzzy modeling and control, *The Proc. of the IEEE,* vol. 83, pp. 378-406, March 1995.

[35] J. S. Jang, C. T. Sun and E. Mizutani. *Neuro-Fuzzy and Soft Computing,* Prentice Hall, Englewood Cliffs, 1997.

[36] S. Jensei. A note on the ordinal sum theorem and its consequence for the construction of triangular norms, *Fuzzy Sets and Systems,* vol. 126, pp. 199-207, 2002.

[37] C.-F. Juang and C.-T. Lin. An on-line self-constructing neural fuzzy inference network and its applications, *IEEE Trans. on Fuzzy Systems,* vol. 6, pp. 12-32, February 1998.

[38] J. Kacprzyk. *Multistage Fuzzy Control,* John Wiley & Sons, Chichester, 1997.

[39] N. Kasabov. *Foundations of Neural Networks,* Fuzzy Systems and Knowledge Engineering. The MIT Press, CA, MA, 1996.

[40] N. Kasabov. Learning fuzzy rules and approximate reasoning in fuzzy neural networks and hybrid systems, *Fuzzy Sets and Systems,* vol. 82, pp. 135-149, 1996.

[41] N. Kasabov. DENFIS: dynamic evolving neural-fuzzy inference system and its application for time-series prediction, *IEEE Trans. on Fuzzy Systems,* vol. 10, pp. 144-154, April 2002.

[42] V. Kecman. *Learning and Soft Computing,* MIT, Cambridge, 2001.

[43] E. Kim, M. Park, S. Ji and M. Park. A new approach to fuzzy modeling, *IEEE Trans. on Fuzzy Systems,* vol. 5, pp. 328-337, August 1997.

[44] E. Kim, M. Park, S. Kimand and M. Park. A transformed input-domain approach to fuzzy modeling, *IEEE Trans. on Fuzzy Systems,* vol. 6, pp. 596-604, November 1998.

[45] E. P. Klement, R. Mesiar and E. Pap. *Triangular Norms,* Kluwer Academic Publishers, Netherlands 2000.

[46] B. Kosko. *Fuzzy Engineering,* Prentice Hall, 1997.

[47] V. Kreinovich, G. C. Mouzouris and H. T. Nguyen. *Fuzzy Rule Based Modeling As a Universal Approximation Tool,* in Fuzzy Systems. Modeling and Control, H. T. Nguyen and M. Sugeno (Eds.), Kluwer Academic Publishers, 1998.

[48] L. I. Kuncheva. *Fuzzy Classifier Design,* Physica-Verlag, New York, 2000.

[49] K.-M. Lee, D.-H. Kwak and H. Lee-Kwang. A fuzzy neural network model for fuzzy inference and rule tuning, *Intern. Journal of Uncertainty, Fuzziness and Knowledge-Based Systems,* vol. 2, no. 3, pp. 265-277, 1994.

[50] K.-M. Lee, D.-H. Kwak and H. Lee-Kwang. Fuzzy inference neural network for fuzzy model tuning, *IEEE Trans. on Systems, Man, and Cybernetics. Part B.,* Vol. 26, No. 4, pp. 637-645, 1996.

[51] C. T. Lin and C. S. G. Lee. Neural-network-based fuzzy logic control and decision system, *IEEE Trans. Comput.,* vol. 40, pp. 1320-1336, December 1991.

[52] C. T. Lin. *Neural Fuzzy Control Systems with Structure and Parameter Learning,* World Scientific, Singapore, 1994.

[53] C. T. Lin and Y. C. Lu. A neural fuzzy systems with linguistic teaching signals, *IEEE Trans. on Fuzzy Systems,* vol. 3, pp. 169-189, May 1995.

[54] C. T. Lin and G. C. S. Lee. *Neural Fuzzy Systems,* A Neuro-Fuzzy Synergism to Intelligent Systems, Englewood Cliffs, NJ: Prentice Hall, 1997.

[55] Y. Lin and G. A. Cunningham III. A New Approach To Fuzzy-Neural System Modeling, *IEEE Trans. on Fuzzy Systems,* vol. 3, pp. 190-198, May 1995.

[56] R. Lowen. *Fuzzy Set Theory. Basic Concepts, Techniques and Bibliography,* Kluwer Academic Publishers, 1996.

[57] J. M. Mendel. Fuzzy logic systems for engineering: a tutorial, *Proc. of the IEEE,* 83(3):345-377, March 1995.

[58] J. M. Mendel. *Uncertain Rule-Based Fuzzy Logic Systems: Introduction and New Directions,* Prentice Hall PTR, Upper Saddle River, NJ, 2001.

[59] Z. Michalewicz. *Genetic Algorithms + Data Structures = Evolution Programs,* Springer-Verlag, Berlin, Heidelberg, New York, 1992.

[60] G. C. Mouzouris and J. M. Mendel. Nonsingleton fuzzy logic systems: Theory and application, *IEEE Trans. on Fuzzy Systems,* vol. 5, No. 1, pp. 56-71, 1997.

[61] D. Nauck and R. Kruse. *Designing Neuro-Fuzzy Systems Through Back-Propagation,* in: Pedrycz W. (Ed.), Fuzzy Modeling: Paradigms and Practice, Kluwer Academic Publishers, Boston, pp. 203-228, 1996.

[62] D. Nauck, F. Klawon and R. Kruse. *Foundations of Neuro-Fuzzy Systems,* Chichester, U.K.: Wiley, 1997.

[63] D. Nauck and R. Kruse. How the learning of rule weights affects the interpretability of fuzzy systems, *WCCI, Alaska,* pp. 1235-1240, 1998.

[64] D. Nauck and R. Kruse. Neuro-fuzzy systems for function approximation, *Fuzzy Sets and Systems,* vol. 101, pp. 261-271, 1999.

[65] J. Nie and D. Linkens. *Fuzzy-Neural Control. Principles, Algorithms and Applications,* Prentice Hall, New York, London, 1995.

[66] K. Nozaki, H. Ishibuchi and K. Tanaka. A simple but powerful heuristic method for generating fuzzy rules from numerical data, *Fuzzy Sets and Systems,* vol. 86, pp. 251-270, 1995.

[67] A. Pagan and A. Ullah. *Nonparametric Econometrics,* Cambridge Univ. Press., London, 1999.

[68] N. R. Pal and D. Chakraborty. *Simultaneous Feature Analysis and SI,* In: Bunke H. and Kandel A. (Eds.), Neuro-Fuzzy Pattern Recognition, World Scientific, Singapore 2000.

[69] Z. Pawlak. Rough sets, *Intern. Journal of Information and Computer Science,* vol. 11, no. 341, 1982.

[70] Z. Pawlak. *Rough Sets. Theoretical Aspects of Reasoning about Data,* Kluwer Academic Publishers, Dordrecht, 1991.

[71] W. Pedrycz. A identification algorithm in fuzzy relational systems, *Fuzzy Sets and Systems,* vol. 13, pp. 153-167, 1984.

[72] W. Pedrycz. Fuzzy neural networks with reference neurons as pattern classifiers, *IEEE Trans. Neural Networks,* vol. 3, no. 5, pp. 770-775, 1992.

[73] W. Pedrycz and F. Gomide. *An Introduction to Fuzzy Sets,* Cambridge, Massachusetts: The MIT Press, 1998.

[74] A. Piegat. *Modeling and Fuzzy Control,* Springer-Verlag, New York, 1999.

[75] H. Roubos and M. Setnes. Compact and transparent fuzzy models and classifiers through iterative complexity reduction, *IEEE Trans. on Fuzzy Systems,* vol. 9, pp. 516-524, August 2001.

[76] D. Rutkowska. *Neuro-Fuzzy Architectures and Hybrid Learning,* Springer-Verlag 2002.

[77] L. Rutkowski and E. Rafajłowicz. On global rate of convergence of some nonparametric identification procedures, *IEEE Trans. on Automatic Control,* vol. AC-34, no.10, pp. 1089-1091, 1989.

[78] L. Rutkowski. Identification of MISO nonlinear regressions in the presence of a wide class of disturbances, *IEEE Trans. on Information Theory,* IT-37, pp. 214-216, 1991.

[79] L. Rutkowski. Multiple Fourier series procedures for extraction of nonlinear regressions from noisy data, *IEEE Trans. on Signal Processing,* vol 41, pp. 3062-3065, 1993.

[80] L. Rutkowski and K. Cpałka. *Flexible Structures of Neuro-Fuzzy Systems, Quo Vadis Computational Intelligence,* Studies in Fuzziness and Soft Computing, vol. 54, Springer-Verlag, pp. 479-484, 2000.

[81] L. Rutkowski and K. Cpałka. A general approach to neuro-fuzzy systems, *The 10th IEEE Intern. Conference on Fuzzy Systems*, Melbourne 2001.

[82] L. Rutkowski and K. Cpałka. Designing and learning of adjustable quasi-triangular norms with applications to neuro-fuzzy systems, *Technical Report, Department of Computer Engineering, Technical University of Czestochowa,* 2002.

[83] L. Rutkowski and K. Cpałka. A neuro-fuzzy controller with a compromise fuzzy reasoning, *Control and Cybernetics,* vol. 31, no. 2, pp. 297-308, 2002.

[84] L. Rutkowski and K. Cpałka. Flexible weighted neuro-fuzzy systems, *9th Intern. Conference on Neural Information Processing (ICONIP'02),* Orchid Country Club, Singapore, November 18-22, 2002.

[85] L. Rutkowski and K. Cpałka. Flexible neuro-fuzzy systems, *IEEE Trans. Neural Networks,* vol. 14, pp. 554-574, May 2003.

[86] L. Rutkowski. Adaptive probabilistic neural networks for pattern classification in time-varying environment, *IEEE Trans. on Neural Networks,* vol. 15, 2004.

[87] L. Rutkowski. Generalized regression neural networks in time-varying environment, *IEEE Trans. on Neural Networks,* vol. 15, 2004.

[88] L. Rutkowski. *New Soft Computing Techniques For System Modeling, Pattern Classification and Image Processing,* Springer-Verlag, 2004.

[89] M. Setnes and H. Roubos. GA-fuzzy modeling and classification complexity and performance, *IEEE Trans. on Fuzzy Systems,* vol. 8, pp. 509-521, October 2000.

[90] B. Ster and A. Dobnikar. Neural networks in medical diagnosis: Comparison with other methods, *Proc. Int. Conf. EANN'96,* A. Bulsari (Ed.), 1996, pp. 427-430.

[91] M. Sugeno and K. Tanaka. Successive identification on a fuzzy model and its applications to prediction of a complex system, *Fuzzy Sets and Systems,* vol. 42, pp. 315-334, 1991.

[92] M. Sugeno and T. Yasukawa. A fuzzy-logic based approach to qualitative modeling, *IEEE Trans. on Fuzzy Systems,* vol. 1, pp. 7-31, February 1993.

[93] R. Tadeusiewicz. *Neural Networks,* RM Academic Publishing House, Warsaw (in Polish), 1993.

[94] R. Tadeusiewicz. *Elementary Introduction to Neural Networks with Computer Programs,* Academic Publishing House, Warsaw (in Polish), 1998.

[95] T. Takagi and M. Sugeno. Fuzzy identification of systems and its application to modeling and control, *IEEE Trans. Syst., Man, Cybern.,* vol. 15, pp. 116-132, 1985.

[96] R. M. Tong. The evaluation of fuzzy models derived from experimental data, *Fuzzy Sets and Systems,* vol. 4, pp. 1-12, 1980.

[97] E. C. C. Tsang, John W. T. Lee and D. S. Yeung. Tuning certainty factor and local weight of fuzzy production rules using fuzzy neural network, *IEEE Trans. of Systems, Man and Cybernetics, Part B,* vol. 32, no.1, pp. 91-98, February 2002.

[98] UCI respository of machine learning databases, C. J. Mertz, P. M. Murphy. Available online: http://www.ics.uci.edu/pub/machine -learning-databases.

[99] J. S. Wang and C. S. G. Lee. Self-adaptive neuro-fuzzy inference systems for classification applications, *IEEE Trans. on Fuzzy Systems,* vol. 10, pp. 790-802, December 2002.

[100] L. X. Wang. *Adaptive Fuzzy Systems and Control.* PTR Prentice Hall, Englewood Cliffs, 1994.

[101] L. X. Wang and R. Langari. Building Sugeno-type models using fuzzy discretization and orthogonal parameter estimation techniques, *IEEE Trans. on Fuzzy Systems,* vol. 3, pp. 454-458, November 1995.

[102] L. X. Wang and C. Wei. Approximation accuracy of some neuro-fuzzy approaches, *IEEE Trans. on Fuzzy Systems,* vol. 8, pp. 470-478, August 2000.

[103] L. Wang and J. Yen. Application of statistical information criteria for optimal fuzzy model construction, *IEEE Trans. on Fuzzy Systems,* vol. 6, pp. 362-371, August 1998.

[104] L. Wang and J. Yen. Extracting fuzzy rules for system modeling using a hybrid of genetic algorithms and Kalman filter, *Fuzzy Sets and Systems,* vol. 101, pp. 353-362, 1999.

[105] W. Xu and Y. Z. Lu. Fuzzy model identification and self-learning for dynamic systems, *IEEE Trans. Syst., Man, Cybern.,* vol. 17, pp. 683-689, July 1987.

[106] R. R. Yager. Fuzzy logic controller structures, *Proc. SPIE Symp. Laser Sci. Optics Appl.* 368-378, 1990.

[107] R. R. Yager. A general approach to rule aggregation in fuzzy logic control, *Appl. Intelligence,* vol.2, pp. 333-351, 1992.

[108] R. R. Yager and D. P. Filev. *Essentials of Fuzzy Modeling and Control.* John Wiley & Sons, 1994.

[109] R. R. Yager and L.A. Zadeh (Eds.). *Fuzzy Sets, Neural Networks and Soft Computing,* Van Nostrand Reinhold, New York, 1994.

[110] J. Yen and L. Wang. Simplifying fuzzy rule based models using orthogonal transformation methods, *IEEE Trans. Syst. Man Cybern.,* vol. 29, pp. 13-24, February 1999.

[111] D. S. Yeung and C. C. Tsang. A multi-level weighted fuzzy reasoning algorithm for expert systems, *IEEE Trans. on Systems, Man & Cybernetics, Part A,* vol. 28, no.2, pp.149-158, March 1998.

[112] Y. Yoshinari, W. Pedrycz and K. Hirota. Construction of fuzzy models through clustering techniques, *Fuzzy Sets and Systems,* vol. 54, pp. 157-165, 1993.

[113] L. A. Zadeh. Fuzzy sets, *Information and Control,* vol. 8, no. 3, pp. 338-353, 1965.

[114] L. A. Zadeh. Similarity relations and fuzzy orderings, *Information Science,* vol. 3, pp. 177-200, 1971.

[115] L. A. Zadeh. *Towards a Theory of Fuzzy Systems,* In: Kalman R.E. and DeClaris N. (Eds.), Aspects of Network and System Theory, Holt, Rinehart and Winston, New York, 1971.

[116] L. A. Zadeh. A fuzzy-set theoretic interpretation of linguistic hedges, *Journal of Cybernetics,* vol. 2, pp. 4-34, 1972.

[117] L. A. Zadeh. Outline of a new approach to the analysis of complex systems and decision processes, *IEEE Trans. on Systems, Man, and Cybernetics,* vol. SMC-3, no. 1, pp. 28-44, 1973.

[118] L. A. Zadeh. On the analysis of large scale systems, In: Gottinger H. (Ed.), *Systems Approaches and Environment Problems,* Vandenhoeck and Ruprecht, pp. 23-37, 1974.

[119] L. A. Zadeh. Fuzzy logic and its application to approximate reasoning, *Information Proc.,* vol. 74, pp.591-594, 1974.

[120] L. A. Zadeh. Calculus of fuzzy restrictions, In: Zadeh L.A., Fu K.-S., Tanaka K. and Shimura M. (Eds.), *Fuzzy Sets and Their Applications to Cognitive and Decision Processes,* Academic Press, New York, pp.1-39, 1975.

[121] L. A. Zadeh. The concept of a linguistic variable and its application to approximate reasoning, *Information Science, Part I,* vol. 8, pp. 199-249, Part II, vol. 8, pp. 301-357, Part III, vol. 9, pp. 43-80, 1975.

[122] L. A. Zadeh. A fuzzy-algorithmic approach to the definition of complex or imprecise concepts, *Intern. Journal of Man-Machine Studies,* vol. 8, pp. 246-291, 1976.

[123] L. A. Zadeh. Fuzzy sets and information granularity, In: Gupta M., Ragade R., and Yager R. (Eds.), *Advances in Fuzzy Set Theory and Applications,* North Holland, Amsterdam, pp. 3-18, 1979.

[124] L. A. Zadeh. Test-score semantics for natural languages and meaning representation via PRUF, In: Rieger B. (Ed.), *Empirical Semantics,* Germany, pp. 281-349, 1981.

[125] L. A. Zadeh. The role of fuzzy logic in the management of uncertainty in expert systems, *Fuzzy Sets and Systems,* vol. 11, pp. 199-227, 1983.

[126] L. A. Zadeh. Outline of a computational approach to mean in and knowledge representation based on a concept of a generalized assignment statement", In: Thoma M. and Wyner A. (Eds.), *Proceed ings of the Intern. Seminar on Artificial Inteligence and Man-Machine Systems,* Springer-Verlag, Heidelberg, pp. 198-211, 1986.

[127] L. A. Zadeh. Fuzzy logic, neural networks and soft computing, *Commun. of the ACM,* vol. 37, no. 3, pp. 77-84, 1994.

[128] L. A. Zadeh. Fuzzy logic and the calculi of fuzzy rules and fuzzy graphs: a precis, *Multiple Valued Logic,* vol. 1, pp. 1-38, 1996.

[129] L. A. Zadeh. Fuzzy logic = computing with words, *IEEE Trans. on Fuzzy Systems,* vol. 4, pp. 103-111, 1996.

[130] L. A. Zadeh. Toward a theory of fuzzy information granulation and its centrality in human reasoning and fuzzy logic, *Fuzzy Sets and Systems,* vol. 90, pp. 111-127, 1997.

[131] L. A. Zadeh. New Frontiers in Fuzzy Logic and Soft Computing, *Proc. of Fourth Conference, Neural Networks and Their Application,* Zakopane, pp. 1-4, 1999.

[132] L. A. Zadeh. From computing with numbers to computing with words-from manipulation of measurements to manipulation of perceptions, *IEEE Trans. on Circuits and Systems-I: Fundamental Theory and Applications,* vol. 45, no. 1, pp. 105-119, 1999.

[133] L. A. Zadeh. Outline of a computational theory of perceptions based on computing with words, In: Sinha N.K. and Gupta M.M. (Eds.), *Soft Computing and Intelligent Systems: Theory and Applications,* Academic Press, San Diego, New York, Tokio, pp. 3-22, 2000.

[134] L. A. Zadeh. A new direction in AI. Toward a computational theory of perceptions, *AI Magazine,* vol. 22, no. 1, pp. 73-84, 2001.

[135] J. M. Żurada. *Introduction to Artificial Neural Systems,* West Publishing Company, 1992.

Index

G

H

I

K

L

Ł

M

N

O

P

Q

R

S

T

U

W

Y

Z