Fuzzy Sets, Fuzzy Controllers, and Neural Networks

Detlef Nauck Frank Klawonn Rudolf Kruse
Department of Computer Science
Technical University of Braunschweig
Bueltenweg 74–75, W-3300 Braunschweig, Germany
Tel. +49.531.391.3155, Fax: +49.531.391.5936

Abstract

This paper gives a short introduction into Fuzzy Set Theory, presents an overview on fuzzy controllers, and discusses possible combinations between fuzzy controllers and neural networks.

Fuzzy Sets suggested by L.A. Zadeh [32] offer a possibility to formally describe linguistic expressions like *tall, fast, medium,* etc., and to operate on them.

Fuzzy controllers use fuzzy sets to represent linguistic values of the input and output variables of a physical system, and describe their relations by fuzzy if–then rules. The idea of fuzzy control is to simulate a human expert who is able to control the system by translation of his or her linguistic inference rules into a control function.

Artificial neural networks are highly parallel architectures consisting of simple processing elements which communicate through weighted connections. They are able to approximate functions or to solve certain tasks by learning from examples. Combinations of neural networks and fuzzy controllers can help to overcome problems in the design and tuning processes of fuzzy controllers.

1 Introduction

This paper collects three talks given on the workshop on *Physiological Aspects of Regulation Between Chaos and Reflex*, an introduction to Fuzzy Set Theory, an overview to fuzzy control and a talk about possible combinations of neural networks and fuzzy controllers.

The paper wants to serve as an introduction to the above mentioned topics, and is intended for the reader who is not familiar with the subject. It cannot be understood as a comprehensive presentation, and the reader is referred to the references mentioned within the text.

The first and main part of the paper concentrates on fuzzy set theory, and shows how to formally describe linguistic expressions like *tall, fast, medium,* etc., and how to operate on them. Semantical aspects of fuzzy sets are also considered.

The second part presents fuzzy controllers that use fuzzy sets to represent linguistic values of the input and output variables of a physical system, and describe their relations by fuzzy if–then rules. The Mamdani controller, the most widely used approach in fuzzy control applications, is presented as an example. The idea of fuzzy control is to simulate a human expert who is able to control the system by translation of his or her linguistic inference rules into a control function.

The last part of the paper deals with combinations of fuzzy controllers and neural networks. It is shown that those combinations can help to overcome problems in the design and tuning process of fuzzy controllers.

In: Wissenschaftliche Zeitschrift der Humboldt-Universität zu Berlin, Reihe Medizin **41**, Nr. 4 (1992), 99–120 (Scientific Journal of the Humboldt-University of Berlin, Series Medicine **41**, no. 4 (1992), 99–120)

2 Foundations of Fuzzy Set Theory

This section deals with the foundations of fuzzy set theory. We first present the notion of a fuzzy set, and discuss forms of possible representations. Then we show how to operate on fuzzy sets, and discuss some semantical aspects of fuzzy sets.

2.1 The Notion of a Fuzzy Set

If we consider a certain universe, for example the set $\mathbb{N}$ of all integers, we generally describe data by defining subsets of the given universe. The feature *less then 10* relative to $\mathbb{N}$ can be characterized by the set

$$A = \{1, 2, 3, 4, 5, 6, 7, 8, 9\} \subseteq \mathbb{N}.$$

In addition to this representation as a list, we can also understand *less then 10* as a predicate, i.e. considering an integer $n \in \mathbb{N}$ *less_than_10*(n) is true only if n is a number from 1 to 9:

$$A = \{n \in \mathbb{N} \mid less\_than\_10(n)\}.$$

A third kind of representation of the datum *integer that is less than 10* is the definition of the respective characteristic function

$$\begin{aligned} \mathbb{I}_A \quad &: \quad \mathbb{N} \rightarrow \{0, 1\} \\ \mathbb{I}_A(n) \quad &= \quad \begin{cases} 1, \text{if } less\_than\_10(n) \\ 0, \text{otherwise} \end{cases} \end{aligned}$$

that yields the value 1 for each element of the universe that belongs to A, and the value 0 for each element that does not.

This concept of a *crisp* set can be used to characterize e.g. the property "at least 45 years old" considering the universe of all living people. The characteristic function of this set is depicted in Figure 1a. A problem arises if the linguistic term "old" has to be characterized. The use of any characteristic function results in the unsatisfactory situation that there is an age x where a person is considered to be "old" whereas a $x - 1$ year old person is said to be "not old".

A possible solution to this problem is to generalize the definition of the characteristic function in a way that it yields values from the interval $[0, 1]$ and not just the two values of the set $\{0, 1\}$. This leads to the notion of a *fuzzy set* [32, 17, 20].

Definition 2.1 *A* **fuzzy set** μ *of* X *is a function that maps from the universe* X *into the unit interval, i.e.* $\mu : X \rightarrow [0, 1]$*.* $F(X)$ *denotes the set of all fuzzy sets of* X*.*

The value $\mu(x)$ denotes the *membership degree* of x to the fuzzy set μ. Figure 1b shows a (subjectively defined) *membership function* of the fuzzy set μ describing the linguistic term "old". In this example an age of 70 is considered to belong to the concept "old" with certainty ($\mu(70) = 1$), whereas an age of 20 does certainly not have this property ($\mu(20) = 0$). The age of 45 is considered to be "old" to some extent only, its membership degree is 0.5.

The use of fuzzy sets to formally represent vague data is often done in an intuitive way, because in many applications there is no model that provides a clear interpretation of the membership degrees. Furthermore it is a general problem that concepts like "old" can only be determined subjectively.

Example 2.2 Let x be the number of rotations per minute of a computer harddisk. Obiously this kind of information is afflicted with vagueness, so it is more realistic to use the linguistic datum

"The number of rotations per second is nearly x",

and to interpret it accordingly.

The interval $[x - \varepsilon, x + \varepsilon]$ can be used for modelling, but this causes the already mentioned problem of counter-intuitive crisp boundaries at $x - \varepsilon$ and $x + \varepsilon$. If statistical data are not available, so that probabilistic methods cannot be used, fuzzy sets are a possible solution, particularly since experts can often specify them intuitively in a very direct manner. In this example let the expert choose the fuzzy

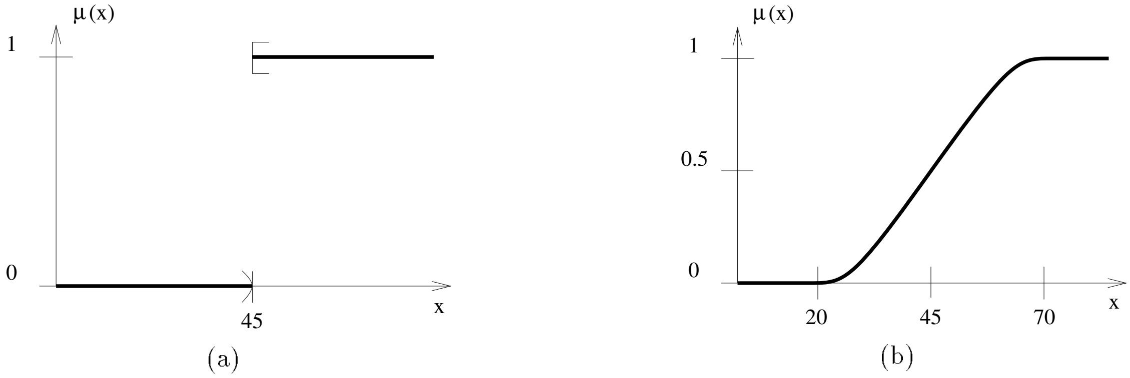


Figure 1: (a) crisp set characterizing the term "at least 45 years old", (b) fuzzy set characterizing the term "old"

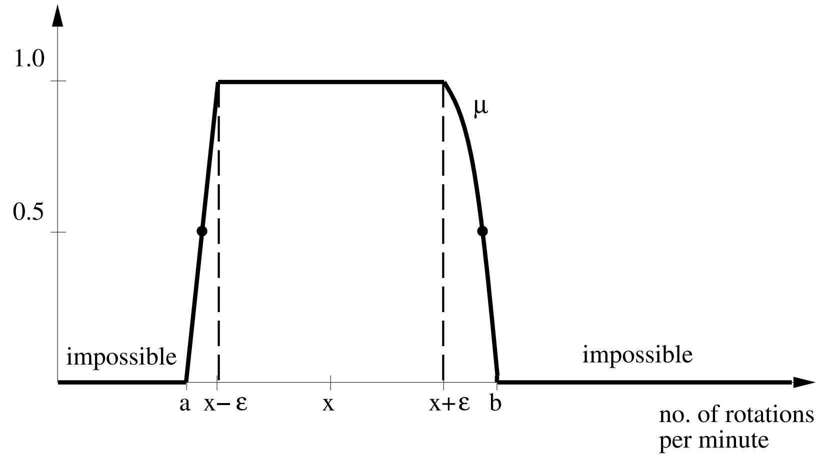


Figure 2: Fuzzy set characterizing the number of rotations of a computer harddisk

set of figure 2. Values for x that are less than a or greater than b are considered as impossible, whereas values between $x - \varepsilon$ and $x + \varepsilon$ occur with certainty. For the remaining areas the expert chooses an monotonously increasing or decreasing course of the function, respectively. $\diamond$

Example 2.3 A digital representation of an X-ray photograph can be obtained by dividing it into an $m \times n$ pixel field, and attributing a grey level to each pixel. Let $P \stackrel{def}{=} \{x_1, \dots, x_m\} \times \{y_1, \dots, y_n\}$ be the set of pixels and $G \stackrel{def}{=} \{g_0, \dots, g_{\max}\}$ be the set of grey levels ordered by their darkness such that g_0 corresponds to "white", and $g_{\max}$ corresponds to "black".
The X-ray photograph can be represented as a grey level image by the function

$$\rho : P \to G.$$

The function

$$\tau : G \to [0,1]$$

yields a value from the unit interval for each grey level, with $\tau(g_0) = 0$, $\tau(g_{\max}) = 1$, $\forall g, g' \in G : g \sqsubset g' \iff \tau(g) < \tau(g')$, where $\sqsubset$ denotes the linear ordering on G. The mapping

$$\mu \stackrel{def}{=} \tau \circ \rho$$

can be understood as the characteristic function of the grey level image. That means that in addition to the interpretation of linguistic data fuzzy sets can also be used for the description of grey level images. $\diamond$

Using fuzzy sets one has to consider what kind of objects are being described by them. Under the *physical interpretation* a fuzzy set characterizes a vague object, e.g. a grey level image that is visible on a physically existing sheet of paper.

As in the case of the X-ray photograph mentioned in Example 2.3 mostly the *epistemic interpretation* is considered, where a fuzzy set describes the vague observation of an existing crisp object of interest. So the X-ray photograph has to be regarded as a vague (two-dimensional) observation of an exsiting three-dimensional crisp object.

2.2 Representations of Fuzzy Sets

There are different possibilities to determine and represent a membership functions characterizing a fuzzy set. If the universe X contains only a finite number of elements, a fuzzy set μ of X will be defined by specifying for each element $x \in X$ its membership degree $\mu(x)$. If the number of elements is very large or a continuum is chosen for X than $\mu(x)$ can be better defined by a function that can use parameters which are adapted to the actual modelling problem.

Let $X = \mathbb{R}$, the linguistic expression "big" can be interpreted by linear functions like

$$\mu_{a,b}(x) \stackrel{def}{=} \begin{cases} 0 & \text{if} \quad x \le a \\ \frac{x-a}{b-a} & \text{if} \quad a \le x \le b \\ 1 & \text{if} \quad x \ge b, \end{cases}$$

where $a < b$ is assumed.

The function depicted in Figure 3 is an alternative interpretation, using the coordinates of the defined points as parameters.

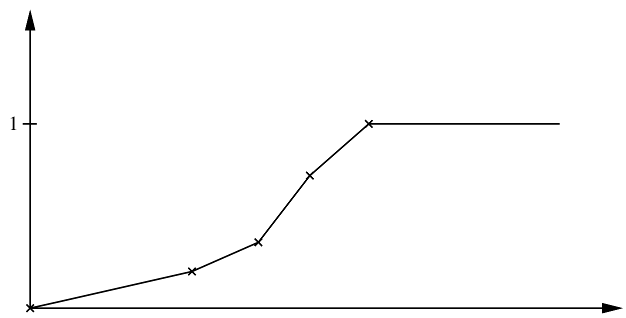


Figure 3: Interpretation of the concept "big" by a fuzzy set

Exponentially increasing functions like

$$\mu_{a,b}(x) \stackrel{def}{=} \begin{cases} 1 - e^{-a(x-b)} & \text{if} \quad b \le x \\ 0 & \text{if} \quad b > x \end{cases}$$

with $a > 0$ and $b \in \mathbb{R}$ are also possible.

To interpret linguistic expressions like "roughly (10)", that are often called *fuzzy numbers,* simply symmetrical triangles

$$\mu_{m,d}(x) \stackrel{def}{=} \begin{cases} 1 - \left|\frac{m-x}{d}\right| & \text{if} \quad m - d \le x \le m + d \\ 0 & \text{if} \quad x < m - d \ \vee \ x > m + d \end{cases}$$

with $d > 0$ or Gaussian functions like

$$\mu_{a,b}(x) \stackrel{def}{=} \exp(-a(x-b)^2), \qquad a > 0,\ b \in \mathbb{R},$$

are used.

The expression "roughly between a and b" can be simply characterized by the trapezoidal function

$$\mu_{a,b,c,d}(x) \stackrel{def}{=} \begin{cases} \frac{x-a}{b-a} & \text{if } a \leq x < b \\ 1 & \text{if } b \leq x \leq c \\ \frac{x-d}{c-d} & \text{if } c < x \leq d \\ 0 & \text{if } x < a \vee x > d \end{cases}$$

where $a < b < c < d$ holds.

Representing a fuzzy set by using the characterizing membership function is called *vertical representation,* considering the graphical representation of the function. Often an expert defines a fuzzy set by specifying for all membership degrees α of a chosen (finite) subset of $[0,1]$ those elements of X that have at least the membership degree α. This is the *horizontal representation* of fuzzy sets by using their *α-cuts.*

Definition 2.4 *Let $\mu \in F(X)$ and $\alpha \in [0,1]$. The set*

$$[\mu]_\alpha \stackrel{def}{=} \{x \in X \mid \mu(x) \geq \alpha\}$$

is called the α-**cut** *of μ.*

Example 2.5 Let μ be the triangluar function on $\mathbb{R}$ depicted in Figure 4. The α-cuts of μ are in this

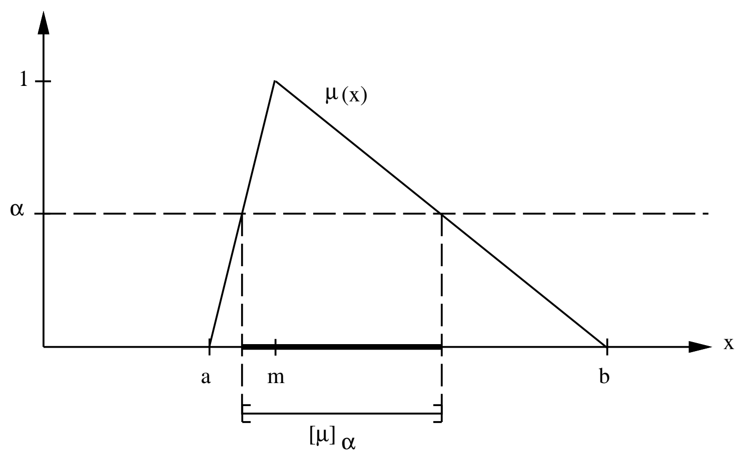


Figure 4: α-cuts of a fuzzy set

case defined by

$$[\mu]_\alpha = \begin{cases} [a + \alpha(m-a), b - \alpha(b-m)] & \text{if } 0 < \alpha \leq 1 \\ \mathbb{R} & \text{if } \alpha = 0. \end{cases}$$

$\diamond$

The following important properties of α-cuts can be directly taken from Figure 4.

Theorem 2.6 *With $\mu \in F(X)$ and $\alpha, \beta \in [0,1]$ the following holds:*

$$\begin{array}{ll} (a) & [\mu]_0 = X \\ (b) & \alpha < \beta \Longrightarrow [\mu]_\alpha \supseteq [\mu]_\beta \\ (c) & \bigcap\limits_{\alpha:\alpha<\beta} [\mu]_\alpha = [\mu]_\beta \end{array}$$

Each fuzzy set can be represented by its α-cuts. The following *representation theorem* holds. It allows to operate on the α-cuts of a fuzzy set which is often less complicated than to operate on the membership functions.

Theorem 2.7 *Let $\mu \in F(X)$. Then*

$$\mu(x) = \sup_{\alpha \in [0,1]} \left\{ \min(\alpha, \mathbb{I}_{[\mu]_\alpha}(x)) \right\}.$$

2.3 Operations on Fuzzy Sets

The basic set theoretical operations intersection, union and complement can also be defined for fuzzy sets. It is expected that the intersection $\mu \cap \mu'$ of two fuzzy sets $\mu, \mu' \in F(X)$ can be calculated element by element, i.e. there is a function $\sqcap : [0,1]^2 \to [0,1]$ such that $(\mu \cap \mu')(x) = \sqcap(\mu(x), \mu'(x))$ holds. To accept $\sqcap$ as an intersection operator it has to fulfill certain axioms, which lead to the definition of a t-norm, known from many-valued logics.

Definition 2.8 *A function $\top : [0,1]^2 \to [0,1]$ is called* **t-norm**, *if the following conditions hold:*

(i)	$\top(a,1) = a$	*(unit 1)*
(ii)	$a \le b \Longrightarrow \top(a,c) \le \top(b,c)$	*(monotonicity)*
(iii)	$\top(a,b) = \top(b,a)$	*(commutativity)*
(iv)	$\top(a,\top(b,c)) = \top(\top(a,b),c)$	*(associativity)*

Obviously $\top$ is non-decreasing in both arguments, and $\top(a,0) = 0$ holds. The most often used t-norms are

$$\begin{aligned} \top_{\min}(a,b) &= \min(a,b), \\ \top_{\text{Luka}}(a,b) &= \max(0, a+b-1) \quad \text{and} \\ \top_{\text{prod}}(a,b) &= a \cdot b. \end{aligned}$$

Dual to the concept of the t-norm is the notion of the t-conorm, that can be used for the definiton of generalized union operators.

Definition 2.9 *A function $\perp : [0,1]^2 \to [0,1]$ is called* **t-conorm**, *iff $\perp$ is commutative, associative, non-decreasing in both arguments, and 0 is its unit.*

Examples of t-conorms are

$$\begin{aligned} \perp_{\min}(a,b) &= \max(a,b), \\ \perp_{\text{Luka}}(a,b) &= \min(a+b,1) \quad \text{and} \\ \perp_{\text{prod}}(a,b) &= a+b-ab. \end{aligned}$$

t-norms and t-conorms induce connectives for fuzzy sets by

$$\begin{aligned} (\mu \cap_\top \mu')(x) &\stackrel{def}{=} \top(\mu(x), \mu'(x)) \quad \text{and} \\ (\mu \cup_\perp \mu')(x) &\stackrel{def}{=} \perp(\mu(x), \mu'(x)). \end{aligned}$$

Example 2.10 In his paper "Fuzzy Sets" from 1965 L.A. Zadeh uses the operators

$$\begin{array}{rcll} (\mu \cap \mu')(x) & \stackrel{def}{=} & \min(\mu(x), \mu'(x)) & \text{intersection} \\ (\mu \cup \mu')(x) & \stackrel{d}{=} & \max(\mu(x), \mu'(x)) & \text{union} \\ \overline{\mu}(x) & \stackrel{def}{=} & 1 - \mu(x) & \text{complement} \end{array}$$

which have beneficial algebraic properties, and are computationally and graphically easy to handle from a practical point of view (see Figure 5. ◇

In addition to the set-theoretic operations it is possible to extend mappings of the form $\phi : X^n \to Y$ to mappings $\hat{\phi} : F(X)^n \to F(Y)$. The transition from a function ϕ to its extension $\hat{\phi}$ is called *extension principle* [33, 20].

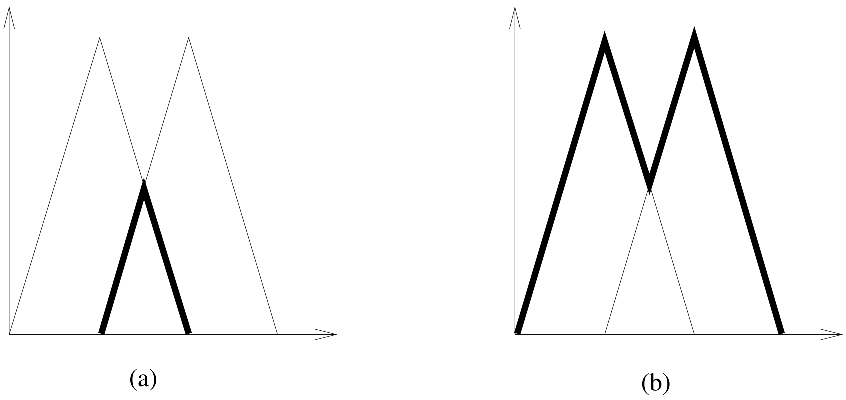


Figure 5: (a) intersection of two fuzzy sets using the min-operator, (b) union of two fuzzy sets using the max-operator

Definition 2.11 *Let $\phi : X^n \to Y$ be a mapping. The* **extension** *of ϕ is defined by*

$$\begin{aligned} &\hat{\phi} : F(X)^n \to F(Y) \quad \textit{with} \\ &\hat{\phi}(\mu_1, \ldots, \mu_n)(y) \stackrel{def}{=} \sup\{\min\{\mu_1(x_1), \ldots, \mu_n(x_n)\} \mid (x_1, \ldots, x_n) \in X^n \\ &\qquad\qquad\qquad\qquad\qquad\qquad \wedge\ y = \phi(x_1, \ldots, x_n)\}, \end{aligned}$$

where $\sup \emptyset \stackrel{def}{=} 0$ *is assumed.*

With the help of the extension principle we obtain for the addition $\mu \oplus \mu'$ of arbitrary fuzzy sets $\mu, \mu' \in F(\mathbb{R})$:

$$(\mu \oplus \mu')(t) \;=\; \sup\{\min\{\mu(x_1), \mu'(x_2)\} \mid x_1, x_2 \in \mathbb{R} \wedge x_1 + x_2 = t\}.$$

In Figure 6 the addition of the two simple fuzzy sets "approximately 2" and "approximately between 2 and 3" is shown. If, as in this case, fuzzy sets are *fuzzy convex,* i.e. all their α-cuts are convex, it is easy to operate on them. In general operations on fuzzy sets are much more complicated, especially if the vertical representation is used. Thus it is preferred to use the horizontal representation, and to operate on the α-cuts, so principles of interval arithmetics can be used [20].

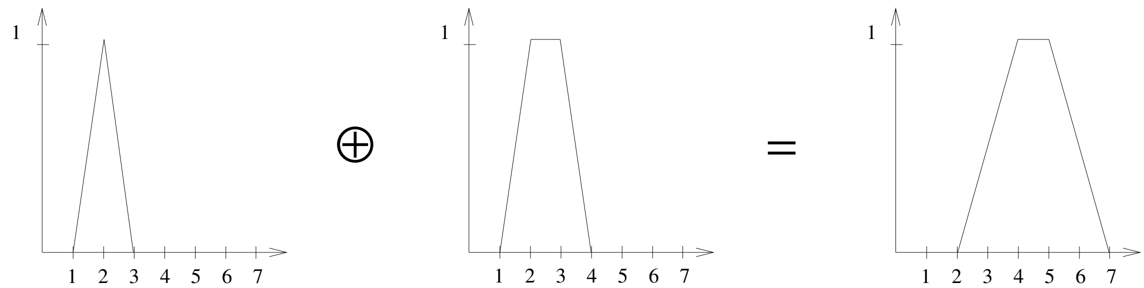


Figure 6: Addition of two simple fuzzy sets

The extension principle enables the fuzzification of whole theories like e.g. the statistical analysis of vague data [19].

2.4 On Semantics of Fuzzy Sets

The existence of different interpretations of probabilities [7], and the difficult acquisition of subjective probabilities [8] show the importance of clarifying the semantics of theoretical concepts. Similar approaches for elucidating the semantics of fuzzy sets are seldom found in the literature. Often fuzzy sets are just used in an heuristical, ad hoc way. Therefore we will briefly consider two different semantical approaches, where the first one regards fuzzy sets as compressed imprecise conflicting information, and the second approach views fuzzy sets as crisp points in vague environments. For a thorough treatment see [20].

The first approach uses fuzzy sets for a condensed representation of imprecise, generally inconsistent information about an unknown, but existing value. This is an epistemic interpretation of fuzzy sets, and in physics an error in measurement is assumed in this cases.

The basic assumption of this model is that vagueness of data is caused by the existence of set-valued information in a system C of different concurrent observation contexts [9, 10]. The contexts can be weighted by a probability measure to express preferences between them.

Definition 2.12 *Let $(C, 2^C, P_C)$ be a finite probability space, X an arbitrary non-empty universe, and $\Gamma : C \to 2^X$ a set-valued mapping. The pair (P_C, Γ) is called* **random set**.

Example 2.13 Four persons are asked, at which height a woman is considered to be tall. Their answers are used to represent the linguistic value "tall" in a condensed form. The first person considers all heights over 160 cm as tall, the second person all heights over 165 cm, the third person all heights over 170 cm, and the fourth person all heights over 180 cm. Each of the four persons forms a context,i.e. we have $C \stackrel{def}{=} \{c_1, c_2, c_3, c_4\}$. Because there are no preferences we choose $P_C(\{c_i\}) = 0.25$, $i = 1, 2, 3, 4$. The mapping $\Gamma : C \to 2^{\mathbb{R}}$ is determined by $c_1 \mapsto [160, \infty)$, $c_2 \mapsto [165, \infty)$, $c_3 \mapsto [170, \infty)$, $c_4 \mapsto [180, \infty)$.

The question whether a woman with a height of 168 cm is to be called tall, has to be answered with "yes" in the contexts c_1 and c_2, and with "no" in the contexts c_3 and c_4. The condensed answer "In two out of four contexts she is tall" is sufficient. This information compression can be performed for each height by writing down for each value $x \in \mathbb{R}$ the weight

$$\frac{|\{c \in C | x \in \Gamma(c)\}|}{|C|} = \sum_{c \in C : x \in \Gamma(c)} P_C(\{c\})$$

of the supporting contexts, and presenting it in the form of a generalized histograph.

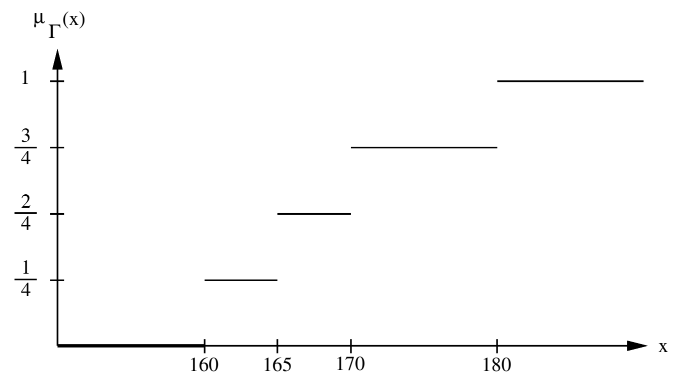


Figure 7: Fuzzy set induced by the random set (P_C, Γ)

Formally we receive a fuzzy set $\mu_\Gamma : \mathbb{R} \to [0, 1]$, but beyond that μ_Γ has a well defined meaning. $\diamond$

The second model we want to consider assumes that vagueness is due to the indistinguishability of certain objects within a defined observation area.

Definition 2.14 *Let $\top$ be a t-norm. A mapping $E : X \times X \to [0, 1]$ is called* **equality relation** *with respect to $\top$, if the properties*

$$\begin{array}{lll} (a) & E(x, x) = 1 & (reflexivity) \\ (b) & E(x, x') = E(x', x) & (symmetry) \\ (c) & \top(E(x, x'), E(x', x'')) \leq E(x, x'') & (transitivity) \end{array}$$

hold.

The value $E(x, x')$ is interpreted as the degree of acceptance to which x and x' are considered to be equal or not distinguishable, respectively. The most simple example for an equality relation is the usual equality on X, i.e.

$$E(x, x') = \begin{cases} 1, & \text{if } x = x' \\ 0, & \text{otherwise.} \end{cases}$$

Example 2.15 Let X be a set of objects observed through panes of milky glass. Each pane has a thickness measured by values between 0 and 1. We call two objects x and x' α-identical if we are not able to distinguish them when observed through a pane of thickness $\alpha \in [0, 1]$.

$$E_1 \stackrel{def}{=} \sup\{1 - \alpha | x \text{ and } x' \text{ are } \alpha\text{−identical}\}$$

is an equality relation with respect to the t-norm $\top_{\min}$.

Let $x, y \in \mathbb{R}$,

$$E_2 \stackrel{def}{=} 1 - \min\{|x - y|, 1\}$$

is an equality relation with respect to the t-norm $\top_{\text{Luka}}$. $\diamond$

Definition 2.16 *Let E be an equality relation on X with respect to the t-norm $\top$. A fuzzy set μ of X that shows the property*

$$\top(\mu(x), E(x, y)) \leq \mu(y) \qquad \textit{(extensionality)}$$

is called **extensional**.

Only those fuzzy sets of X are extensional that respect the equality relation in a way that two objects which are similar according to the equality relation must not have very different membership degrees with respect to the elements of X.

Especially simple extensional fuzzy sets are *singletons* μ_{x_0} for a given $x_0 \in X$ and

$$\mu_{x_0} = E(x, x_0), \quad x \in X.$$

The fuzzy set μ_{x_0} is the smallest extensional fuzzy set for that $\mu_{x_0}(x_0) = 1$ holds, if $\top$ is continuous.

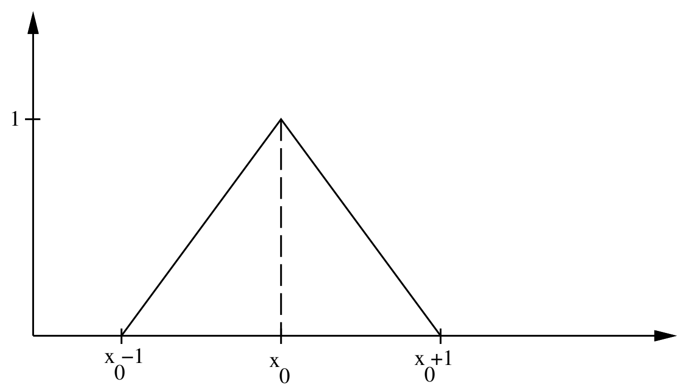


Figure 8: Membership function of μ_{x_0}

Example 2.17 Let $X = \mathbb{R}, \delta : \mathbb{R} \times \mathbb{R} \to \mathbb{R}, (x, y) \mapsto |x - y|$ be the standard metric on $\mathbb{R}$, and E_δ the equality relation induced by $\top_{\text{Luka}}$

$$E_\delta = 1 - \min\{\delta(x, y), 1\}.$$

Then the singleton μ_{x_0} has the triangular membership function

$$\mu_{x_0} = 1 - \min\{|x_0 - x|, 1\}$$

that is depicted in Figure 8 $\diamond$

Not only single elements of X induce fuzzy sets with respect to an equality relation. To each subset of X the smallest extensional fuzzy set can determined.

Definition 2.18
Let E be an equality relation on X with respect to the continous t-norm $\top$. The **extensional hull** *of the set $M \subseteq X$ is the fuzzy set*

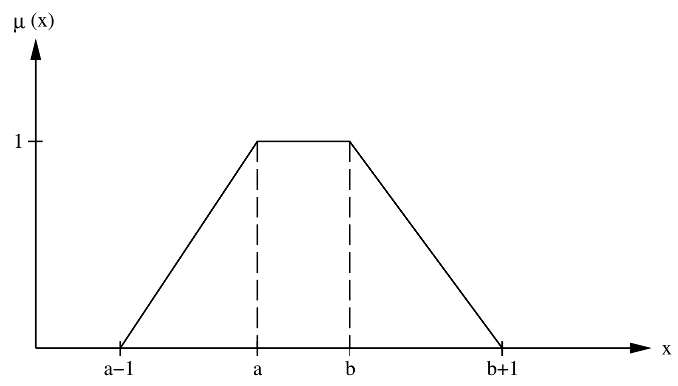
$$\mu_M : X \to [0,1], \qquad x \mapsto \sup\{E(x,x') \mid x' \in M\}.$$

μ_M is the smallest extensional fuzzy set for that $\forall x \in M : \mu_M(x) = 1$ holds.

Example 2.19 We use the same equality relation on $\mathbb{R}$ as in Example 2.17. Let $M = [a,b]$ be an interval. The extensional hull of M is given by a trapezoidal function.

$$\mu_M = \begin{cases} 1, & \text{if} \quad a \leq x \leq b \\ x-a+1, & \text{if} \quad a-1 \leq x \leq a \\ b-x+1, & \text{if} \quad b \leq x \leq b+1 \\ 0, & \text{otherwise} \end{cases}$$

$\diamond$



The consideration of extensional fuzzy sets respects the indistinguishability of objects induced by an equality relation. For a more thorough treatment of this subject and other interpretations used in the literature see [15, 20].

3 Fuzzy Control

In the last years fuzzy control has become one of the most active areas for research in the application of fuzzy set theory. The idea was first presented by Mamdani [22] in 1974, and since then many applications have been developed, most of them in Japan. Known applications are e.g. the train operation system of the subway of Sendai, and Japanese consumer electronics like camcorders, or washing machines.

The main idea of fuzzy control is to simulate the behaviour of an human expert who is in the position to specify the most important properties of a process to be controlled. The specification is done in linguistic rules like

if the car is **fast** *and the wall is* **near**, *then one has to brake* **strongly**.

The terms “fast”, “near”, and “strong” are represented as fuzzy sets, and are vague descriptions of certain values of the input variables, and the output variable. These linguistic control rules are interpreted by fuzzy relations, and each of them specifies a relationship between vague input values and vague output values.

3.1 Fuzzy Logic

The main idea of fuzzy set theory is that the relation "is element of" ($\in$) is not just true or false, but can be expressed by any value of the unit interval. If $\in$ is regarded as a predicate, it is not required in fuzzy set theory that it is just true or false, whereas all other predicates, e.g the equality, remain two-valued. In fuzzy logic the set of possible truth values is extended from the set $\{0, 1\}$ to the interval $[0, 1]$ for all predicates.

Each logical connective is associated with a mapping $[0, 1]^2 \to [0, 1]$ (or $[0, 1] \to [0, 1]$ for the negation) that allows to determine the truth value of logical expressions. To valuate conjunction and disjunction, t-norms and t-conorms are used. The negation is usually defined as $[\![\neg\varphi]\!] = 1 - [\![\varphi]\!]$, where $[\![\varphi]\!] \in [0, 1]$ denotes the truth value of the expression φ. In table 1 several examples for implications are shown.

name	$[\![\varphi \to \psi]\!]$
Lukasiewicz	$\min\{1 - [\![\varphi]\!] + [\![\psi]\!], 1\}$
Gödel	$\begin{cases} 1 & \text{if } [\![\varphi]\!] \leq [\![\psi]\!] \\ [\![\psi]\!] & \text{otherwise} \end{cases}$
Goguen	$\begin{cases} 1 & \text{if } [\![\varphi]\!] = 0 \\ \max\{1, \frac{[\![\psi]\!]}{[\![\varphi]\!]}\} & \text{otherwise} \end{cases}$
Kleene-Dienes	$\max\{1 - [\![\varphi]\!], [\![\psi]\!]\}$
Zadeh	$\max\{1 - [\![\varphi]\!], \min\{[\![\varphi]\!], [\![\psi]\!]\}\}$
Reichenbach	$1 - [\![\varphi]\!] + [\![\varphi]\!] \cdot [\![\psi]\!]$

Table 1: : Implications with different valuation functions

The truth values of simple fuzzy predicates like x is A where A is a linguistic expression and is represented by a fuzzy set μ is determined by the membership degree $\mu(x)$.

3.2 Fuzzy Controllers

In most cases classical methods of control engineering are related to *physical modelling*, where the considered control task is characterized by a mathematical model which often consists of a system of differential equations to be (numerically) solved. If a solution has been found, then well–known techniques of approximation, linearization, validation, and stability analysis are applied to obtain an appropriate control function which quantifies the relationships between the input values and the corresponding output values of the given control system.

Whenever physical modelling is difficult to realize because of complexity problems, or the available information turns out to be rather vague and uncertain than crisp, then *cognitive modelling* seems to be reasonable, if experience has shown that the underlying control task is feasible by human experts.

Cognitive modelling develops a semiformal description of the control strategy and formalizes it by a cognitive model instead of the unknown physical model. Similar to the physical model the iterative procedure of tuning, validation, and stability analysis has to be executed to get an appropriate control function. Using a cognitive approach like fuzzy control it is therefore not intended to get a model of the process, but rather a model of the expert who is in the position to specify the most important properties of the process.

The usual approach in fuzzy control is to define a number of concurrent *fuzzy–if–then rules* like

if X *is* **positive large** *and* Y *is* **positive small**, *then* C *is* **positive medium**,

when we consider a controller with two input and one output variable. The terms "positive large", "positive medium", and "positive small" are represented as fuzzy sets, and are vague descriptions of certain values of the input variables X and Y, and the output variable C. These linguistic control rules are interpreted by fuzzy relations, and each of them specifies a relationship between vague input values and vague output values.

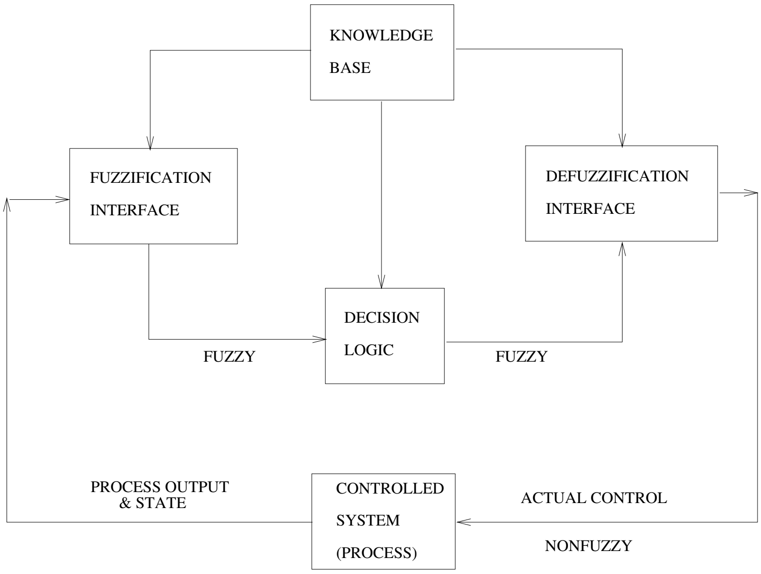


Figure 9: Architecture of a fuzzy controller

The parts of a fuzzy controller (figure 9) carry out the following four steps.

- The *fuzzification interface* receives the current values of the input variables and eventually transforms them into a suitable domain (e.g. $[-1, 1]$). Additionally, it can convert the measured values in linguistic terms or fuzzy sets. The crisp value x_0 usually is transformed into the fuzzy set $\mu^{(x_0)}$ with

$$\mu^{(x_0)} = \begin{cases} 1 & \text{if } x = x_0 \\ 0 & \text{otherwise} \end{cases} .$$

  If information about the exactness is available, or the measured values themself are vague, other fuzzy sets may occur.

- The *knowledge base* contains information about the domains of the variables, kinds of normalization, and the fuzzy sets associated with the linguistic terms. Also a rule base in form of linguistic control rules is stored in the knowledge base.

- The *decision logic* determines the information about the control variables with the help of the measured input values and the knowledge base.

- The task of the *defuzzification interface* is to create a crisp control value out of the information about the control variable of the decision logic by using a suitable transformation.

We will now look at the fuzzy control approach of Mamdani [22], which is an ad hoc approach that is used in most of the practical fuzzy control applications.

3.3 The Mamdani controller

We consider a system with to n measured values $x_1, \ldots, x_n$ with domains $X_1, \ldots, X_n$ and one control value y with the domain Y. To create suitable rules for the knowledge base of the fuzzy controller the sets $X_1, \ldots, X_n$ and Y are "fuzzy-partioned" with the help of fuzzy sets, i.e. on the set X_i p_i different fuzzy sets $\mu_1^{(i)}, \ldots, \mu_{p_1}^{(i)}$ are defined and associated with different linguistic terms like e.g. *positive small*. Usually the sets X_i are real intervals, and the fuzzy sets are determined by triangular functions

$$\mu_{x_0,\varepsilon} : X_1 \to [0, 1], \; x \mapsto 1 - \min\{\varepsilon \cdot |x - x_0|, 1\},$$

where x_0 determines the top, and $\varepsilon > 0$ the width of the triangle. At the "borders" of the interval

$$\mu_1^{(1)} : [a, b] \to [0, 1], \; x \mapsto \begin{cases} 1 & \text{if } x \le x_0 \\ 1 - \min\{\varepsilon \cdot (x - x_0), 1\} & \text{otherwise,} \end{cases}$$

and

$$\mu_{p_1}^{(1)} : [a,b] \to [0,1],\ x \mapsto \begin{cases} 1 & \text{if } x_{p_1} \le x \\ 1 - \min\{\varepsilon \cdot (x_{p_1} - x), 1\} & \text{otherwise,} \end{cases}$$

respectively, are used. These kinds of functions are chosen, because the calculations within the fuzzy controller are much simpler, if piecewise linear functions are used. In Figure 10 a fuzzy-partion with 7 fuzzy sets is shown that are associated with the linguistic terms *negative large (nl), negative medium (nm), negative small (ns), approximately zero (z), positive small (ps), positive medium (pm),* and *positive large (pl).*

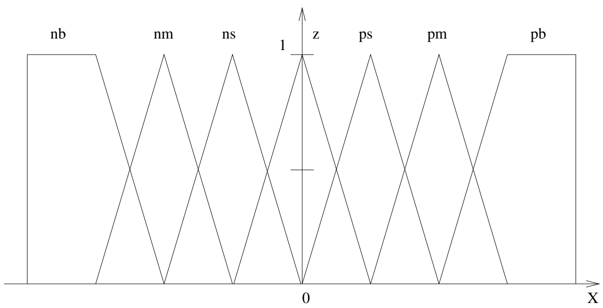


Figure 10: Fuzzy-partion with 7 fuzzy sets

The rule base consists of rules that have the form

$$R_r : \textbf{if } \xi_1 \textbf{ is } A_{i_{1,r}}^{(1)} \textbf{ and } \ldots \textbf{and } \xi_n \textbf{ is } A_{i_{n,r}}^{(n)} \textbf{ then } \eta \textbf{ is } B_{i_r} \quad (r = 1, \ldots, k),$$

where the $A_{i_{1,r}}^{(1)}, \ldots, A_{i_{n,r}}^{(n)}, B_{i_r}$ are linguistic terms attributed to the fuzzy sets $\mu_{i_{1,r}}^{(1)}, \ldots, \mu_{i_{n,r}}^{(n)}, \mu_{i_r}$ according to the fuzzy-partions of the sets $X_1, \ldots, X_n$ and Y, respectively. These control rule have not to be understood as implications in the case of Mamdani's approach, but in the sense of a piecewise defined function, i.e. the k rules correspond to the "function definition" $\eta = \phi(\xi_1, \ldots, \xi_n)$ with

$$\eta \stackrel{def}{=} \begin{cases} B_{i_1} & \text{if} \quad \xi_1 \hat{=} A_{i_{1,1}}^{(i)} \quad \text{and} \ldots \text{and} \quad \xi_n \hat{=} A_{i_{n,1}}^{(i)}, \\ \vdots & \vdots \qquad\qquad\qquad \vdots \\ B_{i_k} & \text{if} \quad \xi_1 \hat{=} A_{i_{1,k}}^{(i)} \quad \text{and} \ldots \text{and} \quad \xi_n \hat{=} A_{i_{n,k}}^{(i)}. \end{cases}$$

Example 3.1 For the problem of balancing an inverted pendulum we use for each the of sets X_1, X_2 and Y a fuzzy-partion as depicted in Figure 10. The sets are the domains of the angle θ, the angle velocity $\dot{\theta}$ and the force F needed to drive the pendulum to a vertical state, respectively. As the rule base we use the 19 rules presented in table 2, where e.g. the entry in the second row of the table reads as the rule

if θ **is** *approximately zero* **and** $\dot{\theta}$ **is** *negative small* **then** F **is** *positive medium.*

◇

A tuple $(x_1, \ldots, x_n) \in X_1 \times \cdots \times X_n$ of currently measured values received by the fuzzification interface is passed directly to the decision logic (after a transformation, eventually).

The decision logic evaluates each rule independently, at first. For each rule R_i the degree to which its antecedent is fulfilled given the measured values is determined. This is done by calculating the value $\mu_{i_{\nu,r}}^{(\nu)}(x_\nu)$ for $\nu = 1, \ldots, n$ the degree to which x_i corresponds to the linguistic term associated with the fuzzy set $\mu_{i_{\nu,r}}^{(\nu)}$. Because the $x_i, \ldots, x_n$ have to fulfill the linguistic terms $A_{i_{1,r}}^{(1)}, \ldots A_{i_{n,r}}^{(n)}$ the values $\mu_{i_{\nu,r}}^{(\nu)} (\nu = 1, \ldots, n)$ have to conjunctively combined. This is done by determining

$$\alpha_r \stackrel{def}{=} \min\{\mu_{i_{1,r}}^{(1)}(x_1), \ldots, \mu_{i_{n,r}}^{(n)}(x_n)\}.$$

$\dot{\Theta}$ \ Θ	nb	nm	ns	z	ps	pm	pb
nb			ps	pb			
nm				pm			
ns	nm		ns	ps			
z	nb	nm	ns	z	ps	pm	pb
ps				ns	ps		pm
pm				nm			
pb				nb	ns		

Table 2: The rule base to balance the inverted pendulum

Instead of min also any other (continous) t-norm can be used. The value α_r is the degree to which the antecedent of rule R_r is fulfilled. The output of rule R_r is the fuzzy set of control variables that is obtained by "cutting off" the "output" fuzzy set μ_{i_r} of rule R_r at the membership degree determined by α_r. Formally the rule R_r induces the fuzzy set

$$\mu_{x_1,\ldots,x_n}^{\text{output}(R_r)} : Y \to [0,1], y \mapsto \min\{\mu_{i_{1,r}}^{(1)}(x_1), \ldots, \mu_{i_{n,r}}^{(n)}(x_n), \mu_{i_r}^{(n)}(y)\}.$$

After all the rules have been evaluated the fuzzy sets have to be combined to a single fuzzy set by using the max-operator (union). Altogether the decision logic supplies the fuzzy set

$$\mu_{x_1,\ldots,x_n}^{\text{output}(R_r)} : Y \to [0,1], \eta \mapsto \max_{r\in\{1,\ldots,k\}} \{\min\{\mu_{i_{1,r}}^{(1)}(x_1), \ldots, \mu_{i_{n,r}}^{(n)}(x_n), \mu_{i_r}(\eta)\}\}.$$

This fuzzy set has to be transformed into a crisp control value by the defuzzification interface. This is usually done by one of the following three methods.

- The *max criterion method* chooses an arbitrary value $\eta \in Y$ where the fuzzy set $\mu_{x_1,\ldots,x_n}^{\text{output}}$ has a maximal membership degree.
- The *Mean of Maximum method (MOM)* needs the set Y to be an intervall an the set $\text{MAX}(\mu_{x_1,\ldots,x_n}^{\text{output}}) \stackrel{def}{=} \{y \in Y \mid \forall y' \in Y : \mu_{x_1,\ldots,x_n}^{\text{output}}(y') \leq \mu_{x_1,\ldots,x_n}^{\text{output}}(y)\}$ to be non-empty and (Lebesgue-) measurable. The control action is the mean value of the set $\text{MAX}(\mu_{x_1,\ldots,x_n}^{\text{output}})$. If the set $\text{MAX}(\mu_{x_1,\ldots,x_n}^{\text{output}})$ is finite, the control action is given by

  $$\eta \stackrel{def}{=} \frac{1}{\left|\text{MAX}(\mu_{x_1,\ldots,x_n}^{\text{output}})\right|} \sum_{y\in\text{MAX}(\mu_{x_1,\ldots,x_n}^{\text{output}})} y,$$

  if not, by

  $$\eta \stackrel{def}{=} \frac{1}{\displaystyle\int_{y\in\text{MAX}(\mu_{x_1,\ldots,x_n}^{\text{output}})} dy} \cdot \int_{y\in\text{MAX}(\mu_{x_1,\ldots,x_n}^{\text{output}})} y dy.$$

- The *center of area method (COA)* is the most widely used defuzzification method. It demands the same condtions as the MOM method. As control action the value under the center of gravity of the area determined by $\mu_{x_1,\ldots,x_n}^{\text{output}}$ is chosen. The value is calculated by

  $$\eta \stackrel{def}{=} \frac{1}{\displaystyle\int_{y\in Y} \mu_{x_1,\ldots,x_n}^{\text{output}}(y) dy} \cdot \int_{y\in Y} y \cdot \mu_{x_1,\ldots,x_n}^{\text{output}}(y) dy.$$

In Figure 11 all of the above described steps that are carried out by the Mamdani max–min controller are graphically presented using two rules.

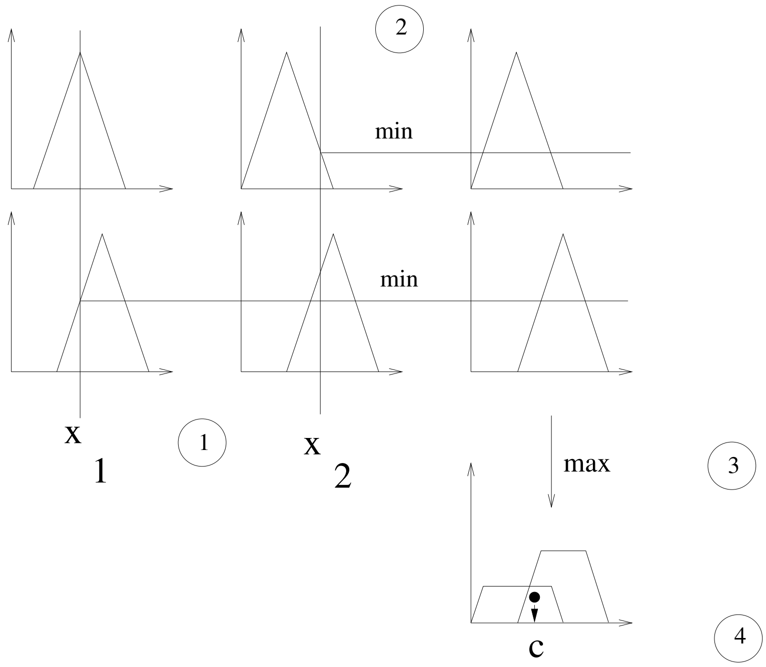


Figure 11: Steps of a Mamdani max–min controller

The COA method for defuzzification guarantees a smooth control behaviour, contrarily to the other two presented methods. The disadvantages of it are the complex calculations that have to be carried out, the lack of theoretical approval and anomalies in certain situations.

Example 3.2 Consider a fuzzy controller steering a car in a way to avoid obstacles. If an obstacle occurs right ahead, the fuzzy set depicted in Figure 12 could by determined by the decision logic, which could be interpreted as "swerve to the right or to the left", a plausible control action. Both methods,

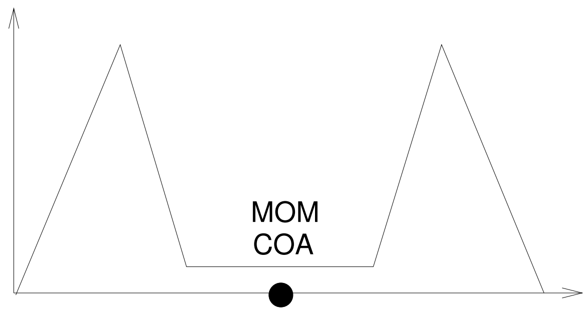


Figure 12: Undesired defuzzification result by COA and MOM

COA and MOM, deliver the value 0 as control action, i.e. the car continues to drive ahead until it collides with the obstacle. ◇

The described anomaly does not occur with convex fuzzy sets, which can be interpreted as a representation of a single (vague) value or interval. A suitable defuzzification method would have to choose between different control actions (choose one of the two triangles in Figure 12) and then to transform the fuzzy set into a crisp value, or vice versa. For a thorough treatment of defuzzification methods see [20].

Some of the concepts applied in fuzzy control are based on a rather intuitive understanding of fuzzy set theory without having a clear model which motivates and justifies these concepts. The classical version of the max–min controller tries to handle the above described steps in different mathematical structures.

While the inference mechanism may be clarified by logics, the center of area defuzzification method cannot be justified in a logical calculus. For a formal justification of the max–min controller with the help of equality relations see [15, 20].

4 Combining Fuzzy Controllers and Neural Networks

To design a fuzzy controller linguistic rules and membership functions to represent linguistic values have to be determined. The specification of good linguistic rules depends on the knowledge of the control expert. But the translation of these into fuzzy set theory is not formalized and arbitrary choices concerning for example the shape of membership functions have to be made. The quality of the fuzzy controller can be drastically influenced by changing shapes of membership functions. Thus methods for tuning fuzzy controllers are necessary.

Artificial neural networks are highly parallel architectures consisting of simple processing elements which communicate through weighted connections. They are able to approximate functions or to solve certain tasks by learning from examples [16]. Neural networks can also be used for control problems, and they are able to learn (supervised or unsupervised) from given data. The problem is that the learning process takes a lot of time and is not guaranteed to be successful. Furthermore it is not possible to integrate prior knowledge to simplify the learning process, or to extract knowledge e.g. in form of rules from the trained network. Foundations of artificial neural networks are not considered in this paper. For an introduction see e.g. [1, 5, 16]. The term neural net used in the following has no relation to biological models. We only consider neural networks as artificial information processing systems.

A combination of both approaches offer the possibility to the solve the tuning and design problem of fuzzy control. The combination of neural networks and fuzzy controllers assembles the advantages of both approaches and avoids the drawbacks of them.

There exist cooperative approaches that use neural networks to optimize certain parameteres of an ordinary fuzzy controller, or to preprocess data and extract fuzzy control rules from data. Hybrid architectures integrate the concepts of a fuzzy controller into a neural network or understand a fuzzy controller as a neural network.

4.1 Tuning Fuzzy Control Parameters by Neural Networks

The choice of a particular membership function representing a linguistic term is more or less arbitrary. For example, consider the linguistic term *approximately zero*. Obviously, the corresponding fuzzy set should be a unimodal function reaching its maximum at the value zero. Neither the shape, which could be triangular or gaussian, nor the range, i.e. the support of the membership function is uniquely determined by *approximately zero*. Generally, the control expert has some idea about the range of the membership function, but he would not be able to argue about small changes of his specified range.

Therefore, the tuning of membership functions becomes an import issue in fuzzy control. Since this tuning task can be viewed as an optimization problem neural networks offer a possibilty to solve this problem. A straightforward approach is to assume a certain shape for the membership functions which depends on different parameters that can be learned by a neural network. This idea was carried out in [26] where the membership functions are assumed to be symmetrical triangular functions depending on two parameters, one of them determining where the function reaches its maximum, the other giving the length of the support. Gaussian membership functions were used in [13].

Both approaches require a set of training data in the form correct input–output tuples and a specification of the rules including a preliminary definition of the corresponding membership functions.

A method which can cope with arbitrary membership functions for the input variables is proposed in [11, 30, 31]. The training data have to be divided in r disjoint clusters $R_1, \ldots, R_r$. Each cluster R_i corresponds to a control rule $\mathcal{R}_i$. Elements of the clusters are tuples of input–output values of the form (x, y) where x can be a vector $x = (x_1, \ldots, x_n)$ of n input variables.

This means that the rules are not specified in terms of linguistic variables, but in the form of crisp input–output tuples.

A multilayer perceptron with n input units, some hidden layers, and r output units can be used to learn these clusters. The input data for this learning task are the input vectors of all clusters, i.e. the set $\{x \mid \exists i\ \exists y : (x,y) \in R_i\}$. The target output $t_{u_i}(x)$ for input x at output unit u_i is defined as

$$t_{u_i}(x) = \begin{cases} 1 & \text{if there exists } y \text{ s.t. } (x,y) \in R_i \\ 0 & \text{otherwise.} \end{cases}$$

After the network has learned its weights, arbitrary values for x can be taken as inputs. Then the output at output unit u_i can be interpreted as the degree to which x matches the antecedent of rule $\mathcal{R}_i$, i.e. the function

$$x \mapsto o_{u_i}$$

is the membership function for the fuzzy set representing the linguistic term on the left–hand side of rule $\mathcal{R}_i$.

In case of a Mamdani type fuzzy controller the same technique can be applied to the output variable, resulting in a neural network which determines the fuzzy sets for the right–hand sides of the rules.

4.2 Neural Networks for Extracting Fuzzy Rules

If a set of input–output tuples for a given control problem is available, we can try to extract fuzzy (control) rules from this set. This can either be done by fuzzy clustering methods (for an overview see [3]) or by using neural networks.

The input vectors of the input–output tuples can be taken as inputs for a Kohonen self–organizing map, which can be interpreted in terms of linguistic variables [27]. The main idea for this interpretation is to refrain from the winner–take–all principle after the weights for the self–organizing map are learned. Thus each output unit corresponds to a antecedent of a fuzzy control rule. Depending on how far output unit u_i is from being the 'winner' given input vector x, a matching degree $\mu_i(x)$ can be specified, yielding the degree to which x satisfies the antecedent of the corresponding rule.

Another way to obtain directly a fuzzy clustering is to apply the modified Kohonen network proposed in [4].

Kosko uses another approach to generate fuzzy–if–then rules from existing data [18]. Kosko shows that fuzzy sets can be viewed as points in a multidimensional unit hypercube. This makes it possible to use *fuzzy associative memories (FAM)* to represent fuzzy rules. Special adaptive clustering algorithms allow to learn these representations (AFAM).

4.3 Fuzzy Control Oriented Neural Networks

Another approach to combine fuzzy control and neural networks is to create a special architecture out of standard feed–forward nets that can be interpreted as a fuzzy controller. In [2] such a system is described.

Berenji calls this architecture *ARIC (approximate reasoning based intelligent control).* It is a neural network model of a fuzzy controller and learns by updating its prediction of the physical system's behaviour and fine tunes a predefined control knowledge base [2].

This kind of architecture allows to combine the advantages of neural networks and fuzzy controllers. The system is able to learn, and the knowledge used within the system has the form of fuzzy–if–then rules. By predefining these rules with the help of an experienced operator the system has not to learn from scratch, so it learns faster than a standard neural control system.

The ARIC architecture consists of two special feed–forward neural networks, the *action–state evaluation network (AEN)* and the *action selection network (ASN).* The ASN is a multilayer neural network representation of a fuzzy controller. In fact, it consists of two separated nets, where the first one is the fuzzy inference part and the second one is a neural network that calculates $p[t, t+1]$, a *measure of confidence associated with the fuzzy inference value* $u(t+1)$, using the weights of time t and the system state of time $t+1$. Berenji calls this value "probability". A so called *stochastic modifier* combines the recommended control value $u(t)$ of the fuzzy inference part and the probability value p and determines the final output value $u'(t) = o(U(t), p[t, t+1])$ of the ASN. The hidden units z_i of the fuzzy inference network represent the fuzzy rules, the input units x_j the rule antecedents, and the output unit u represents the control action, that is the defuzzified combination of the conclusions of all rules (output of hidden units).

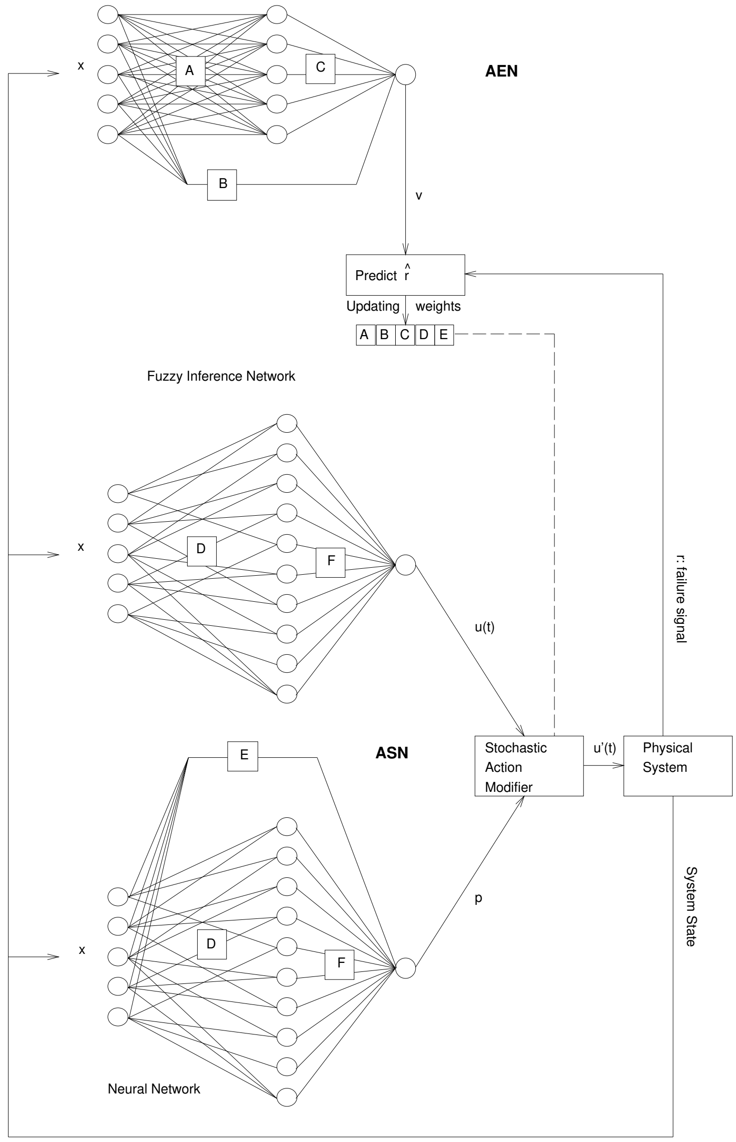


Figure 13: Berenji's ARIC architecture

The AEN is a feed–forward neural network with one hidden layer, that receives the system state as its input and an error signal r from the physical system as additional information. It evaluate the current state and outputs a reinforcement signal used to calculate the weight changes in the system.

Berenji has applied his ARIC architecture to cart–pole balancing and has shown that the system is able to solve this task. In Figure 13 an outline of ARIC is shown.

We call the Berenji's system *a fuzzy control oriented neural network.* It is a system that consists of more or less standard neural networks, where one network employs a fuzzy aspect in using membership functions within its units and a fuzzy inference scheme.

4.4 Neural Network Oriented Fuzzy Control

In this section we present our approach to a combination of neural networks and fuzzy controllers (see also [6]). Our main concern is to keep the structure of the fuzzy controller that is determined by the fuzzy rules. We think of those rules as a piece of structural knowledge that gives us a roughly correct representation of the system to be controlled. We consider it a natural approach to define a fuzzy error for our system, which, according to its structure, we may call neural fuzzy controller. This enables us to define a non-supervised learning algorithm that is described at length in [23]. Training a fuzzy controller with such a learning procedure allows us to keep track of the changes and to interpret the modified rules.

An example for the structure of our neural fuzzy controller is depicted in Figure 14. The modules X_1 and X_2 represent the input variables, and deliver their crisp values to their μ-modules which contain the respective membership functions. The μ-modules are connected to the R-modules which represent the fuzzy if-then rules. Each μ-module gives to all of its connected R-modules, the membership value $\mu_{ij}(x_i)$ of its input variable X_i. The R-modules use a t-norm (min-operation in this case) to calculate the conjunction of their inputs and pass this value forward to one of the ν-modules, which contain the membership functions representing the output variable. By passing through the ν-modules these values are changed to the conclusion of the respective rule. This means the implication (min-implication in this case) is carried out to obtain the value of the conclusion, and because monotonic membership functions are used, the ν-modules pass pairs $(r_l, \nu_k^{-1}(r_l))$ to the C-module, where the final output value is calculated by

$$c = \frac{\sum_{i=1}^{n} r_i \nu_{R_i}^{-1}(r_i)}{\sum_{i=1}^{n} r_i},$$

where n is the number of rules, and r_i is the degree to which rule R_i has fired.

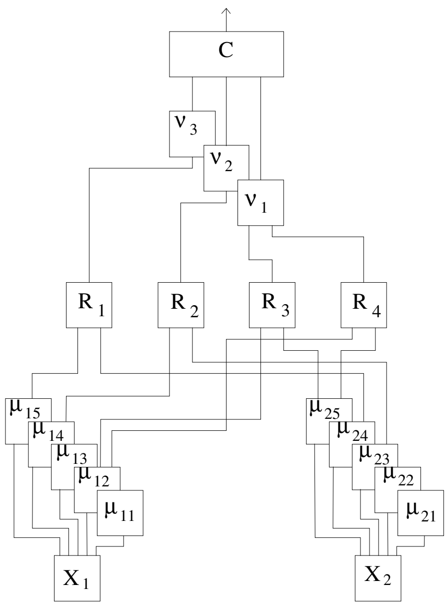


Figure 14: The structure of the neural fuzzy controller

As one can easily see, the system in Figure 14 resembles a feedforward neural network. The X-, R-, and C-modules can be viewed as the neurons and the μ- and ν-units as the adaptable weights of the network. The fact that one μ-module can be connected to more than one R-module is equivalent to connections in a neural network that share a common weight. This is very important, because we want each linguistic value to be represented by only one membership function that is valid for all rules. By this restriction we retain the structural knowledge that we put into the system by defining the rules. In other neural fuzzy systems this fact is not recognized [14, 2] and it is possible that one linguistic value is represented by different membership functions.

Our goal is to tune the membership functions of the controller by a learning algorithm according to a measure that adequately describes the state of the plant under consideration. The desired optimal state of the plant can be described by a vector of state variable values. But usually we are content with the current state if the variables have roughly taken these values. And so it is natural to define the goodness of the current state by a membership function from which we can derive a fuzzy error [23] that characterizes the performance of our neural fuzzy controller.

Consider a system with n state variables $X_1, \ldots, X_n$. We define the fuzzy-goodness G_1 as

$$G_1 = \min(\mu_{X_1}^{optimal}, \ldots, \mu_{X_n}^{optimal}),$$

where the membership functions $\mu_{X_i}^{optimal}$ have to be defined according to the requirements of the plant under consideration.

In addition of a near optimal state we also consider states as good, where the incorrect values of the state variables compensate each other in a way, that the plant is driven towards its optimal state. We define the fuzzy-goodness G_2 as

$$G_2 = \min(\mu^{compensate_1}(X_1, \ldots, X_n), \ldots, \mu^{compensate_k}(X_1, \ldots, X_n))$$

where the membership functions $\mu^{compensate_j}$ again have to be defined according to the requirements of the plant. There may be more than one $\mu^{compensate_j}$ and they may depend on two or more of the state variables.

The overall fuzzy-goodness is defined as

$$G = g(G_1, G_2),$$

where the operation g has to be specified according to the actual application [23]. The fuzzy-error that is made by our neural fuzzy controller is defined as

$$E = 1 - G.$$

The learning algorithm depending on this fuzzy error is able to change the membership functions. The *fuzzy error propagation learning algorithm* is an unsupervised learning algorithm where each R-module changes the membership functions of its antecedents and conclusion according to the extent it has fired under the given circumstances and according to an error describing the performance of the control system. The neural fuzzy controller does not have to learn from scratch, but knowledge in the form of fuzzy if-then rules is coded into the system. The learning procedure does not change this structural knowledge. It tunes the membership functions in an obvious way, and the semantics of the rules is not blurred by any semantically suspicious factors or weights attached to the rules.

The algorithm finds the most appropriate membership functions for the current control situation, while the meaning of each rule is not changed. If the learning is successful, this means that the rule base is suitable for the control task, and the changes in the fuzzy sets are due to an unsuitable modelling of the linguistic values by the a priori membership functions. We can interpret the new membership functions to give us the best fuzzy sets for the current task.

An extension of this model that is able to learn rules by changing the number of hidden nodes, is also possible.

5 Conclusions

In this paper we presented an introduction to fuzzy set theory and fuzzy controllers. We described the Mamdani controller used in most fuzzy control applications and discussed the problem of defuzzification. As we have pointed out, it is necessary to always consider semantical aspects, when fuzzy sets are used, and we have shown that the concept of an equality relation is a possible approach to do so.

We discussed four possible combinations of neural networks and fuzzy controllers. The first two approaches use stand–alone neural networks to tune membership functions or to create fuzzy–if–then rules. The remaining concepts both try to build homogeneous systems that benefit from the advantages of both worlds.

References

[1] Aleksander, I., Morton, H.: An Introduction to Neural Computing. Chapman and Hall, London, 1990.

[2] Berenji, H.B.: A Reinforcement Learning-Based Architecture for Fuzzy Logic Control. Int. J. Approximate Reasoning 6 (1992), 267-292.

[3] Bezdek, J.C., Pal, S.K. (eds.): Fuzzy Models for Pattern Recognition. IEEE Press, New York, 1992.

[4] Bezdek, J.C., Tsao, E.C., Pal, N.K.: Fuzzy Kohonen Clustering Networks. Proc. IEEE International Conference on Fuzzy Systems 1992, San Diego (1992), 1035–1043.

[5] Brause, R.: Neuronale Netze. Teubner, Stuttgart, 1991.

[6] Eklund, P., Klawonn, F.,Nauck, D.: Distributing Errors in Neural Fuzzy Control. Proc. 2nd Int. Conf. on Fuzzy Logic and Neural Networks (IIZUKA-92), Iizuka, Juli 17-22, (1992), 1139-1142.

[7] Fine, T.: Theories of Probability. Academic Press, New York, 1973.

[8] Fishburn, P.C.: The Axioms of Subjective Probability. Statistical Science **1**, (1986) 3, 335–358.

[9] Gebhardt, J., Kruse, R.: The Context Model — An Integrating View of Vagueness and Uncertainty. International Journal of Approximate Reasoning, Special Issue on Fuzzy Expert Systems (to be published).

[10] Gebhardt, J.: Ein integrierender Ansatz zur possibilistischen Inferenz in wissensbasierten Systemen. Dissertationsschrift. Technische Universität Braunschweig (1992).

[11] Hayashi, I., Nomura, H., Yamasaki, H., Wakami, N.: Construction of Fuzzy Inference Rules by NFD and NDFL. Int. J. Approximate Reasoning **6** (1992), 241–266.

[12] Hirota, K.: Survey of Industrial Applications of Fuzzy Control in Japan. Proc. IJCAI–91 Workshop on Fuzzy Control, Sydney (1992), 18–20.

[13] Ichihashi, H.: Iterative Fuzzy Modelling and a Hierarchical Network. In: Lowen, R., Roubens, M. (eds.): Proc. 4th IFSA Congress, Engineering, Brussels (1992), 49–52.

[14] Jang, J.-S.R.: Fuzzy Modelling Using Generalized Neural Networks and Kalman Filter Algorithm. Proc. AAAI-91, vol. 2 (1991), 762-767.

[15] Klawonn, F., Kruse, R.: Equality Relations as a Basis for Fuzzy Control. Fuzzy Sets and Systems, **54** (March 1993) 2, 147–156.

[16] Klawonn, F., Kruse, R., Nauck, D.: Neuronale Netze in der KI, Vieweg, Braunschweig, 1993 (to be published).

[17] Klir, G.J., Folger, T.A.: Fuzzy Sets, Uncertainty, and Information. Prentice Hall, 1988.

[18] Kosko, B.: Neural Networks and Fuzzy Systems. Prentice–Hall, Englewood Cliffs, 1992.

[19] Kruse, R., Meyer, K.D.: Statistics on Vague Data. Reidel, Dordrecht, 1987.

[20] Kruse, R., Gebhardt, J., Klawonn, F.: Fuzzy Systeme. Teubner, Stuttgart, 1993 (to be published).

[21] Lee, C.C.: Fuzzy Logic in Control Systems: Fuzzy Logic Controller. IEEE Trans. Systems, Man, Cybernetics 20 (1990), 404-418 (part I), 419–435 (part II).

[22] Mamdani, E.H.: Applications of Fuzzy Algorithms for a Simple Dynamic Plant. Proc. IEE **121** (1974), 1585–1588.

[23] Nauck, D., Kruse, R.: A Neural Fuzzy Controller Learning by Fuzzy Error Propagation. Proc. Conf. of North American Fuzzy Information Processing Society (NAFIPS'92), Puerto Vallarta, Mexico, Dec. 14-17 (1992), 388–397.

[24] Nauck, D.,Kruse, R.: Interpreting Changes in the Fuzzy Sets of a Self–Adaptive Neural Fuzzy Controller. Proc. Workshop on Industrial Applications of Fuzzy Control and Intelligent Systems, College Station, Texas, Dec. 2–4 (1992), 146–152.

[25] Nauck, D., Kruse, R.: A Fuzzy Neural Network Learning Fuzzy Rules and Membership Functions by Fuzzy Error Backpropagation. Proc. IEEE Int. Conf. on Neural Networks. San Francisco, Mar. 28 - Apr. 1 (1993), 1022-1027.

[26] Nomura, H., Hayashi, I., Wakami, N.: A Learning Method of Fuzzy Inference Rules by Descent Method. Proc. IEEE International Conference on Fuzzy Systems, San Diego 1992, 203–210.

[27] Pedrycz, W., Card, H.C.: Linguistic Interpretation of Self–Organizing Maps. Proc. IEEE International Conference on Fuzzy Systems, San Diego 1992, 371–378.

[28] Rumelhart, D.E., McClelland, J.L.: Parallel Distributed Processing, Vol. 1. MIT Press, Cambridge, Massachusetts, 1986.

[29] Shao, S.: Fuzzy Self-Organizing Controller and its Application for Dynamic Processes. Fuzzy Sets and Systems 26 (1988), 151-164.

[30] Takagi, H., Hayashi, I.: NN–Driven Fuzzy Reasoning. Int. J. Approximate Reasoning 5 (1991), 191–212.

[31] Takagi, H., Kouda, T., Kojima, Y.: Neural Networks Based on Approximate Reasoning Architecture. Japanese Journal of Fuzzy Theory and Systems 3 (1991), 63-74.

[32] Zadeh, L.A.: Fuzzy Sets. Information and Control **8** (1965), 338–353.

[33] L.A. Zadeh: Outline of a New Approach to the Ananlysis of Complex Systems and Decision Processes. IEEE Trans. Syst. Man Cybern. 3 (1973), 28–44.