Fuzzy Sets and Vague Environments

Frank Klawonn
Department of Computer Science
Technical University of Braunschweig
D–38106 Braunschweig, Germany

Abstract

In this paper we propose a natural approach to handle imprecise numbers as they arise for example from measurements. Fuzzy sets turn out to be a canonical representation for such imprecise numbers that are induced by taking different tolerance or error bounds into account. Fuzzy sets are induced by scaling factors that describe the magnitude of the imprecision. On the other, the scaling factors can be derived from given fuzzy sets so that we have a correspondence between scaling factors and fuzzy sets.

When these concepts are applied to control problems, the max–min rule is rediscovered as an interpolations technique. Viewing fuzzy control as an interpolation technique in vague environments enables us to validate various concepts for the design and tuning of fuzzy controllers and suggests new also new methods based on clear semantics.

Keywords: Vague environment; interpolation; fuzzy control.

1 Introduction

The formal definition of a fuzzy set as a mapping from an underlying universe to the unit interval generalizes the notion of a characteristic function by allowing membership degrees between 0 and 1. In applications fuzzy sets are used to represent vague concepts like *small* or *approximately zero*. Intermediate membership values are intuitively appealing for such concepts, but often the question of assigning a membership degree between 0 and 1 a concrete interpretation is not even considered. There are, of course, approaches to the interpretation of grades of membership for example as in possibility theory [1] or based on probabilistic concepts [3].

It should be emphasized that the various interpretations for membership degrees are not viewed as competing models but as approaches to handle different phenomena like uncertainty or imprecision. In this sense, our interpretation of fuzzy sets must not be understood in a dogmatic way. It is one possible approach that suits well when imprecise numbers have to be processed.

The basic concept of our model is to admit different tolerance or error bounds for the numbers to be processed and to keep track of these bounds. In this way we can introduce an appropriate interpolation technique which is applicable to control problems.

The general idea of cognitive control is to model the behaviour of a control expert instead of building a mathematical or physical model of the considered process. Fuzzy control as a cognitive control method is based on vaguely specified control rules. The use of fuzzy sets as a representation form for vague knowledge is intuitively appealing, but a clear semantics for the fuzzy sets and the applied operators is often not provided.

There is still confusion caused by misinterpretations as modus ponens of the inference mechanism applied to the linguistic control rules [13]. From a purely logical viewpoint fuzzy controllers would not use the max–min rule [7].

We propose to see fuzzy control as an example of fuzzy interpolation as it was already mentioned in a different context in [2]. In this paper we introduce a very simple model of cognitive control based on two concepts: a partially known crisp control function and the notion of a vague environment which reflects the idea of identifying values whose difference is small. More formal, mathematical approaches based on these ideas can be found in [5, 10, 12].

In this paper we discuss a more simple approach which provides a semantical background for Mamdani's fuzzy control model [14]. We concentrate on the consequences this interpretation implies for the design and tuning of a fuzzy controller.

Section 2 introduces the concept of vague environments. In Sections 3 we apply this notion to cognitive control and discuss the connections to Mamdani's model. Section 4 is devoted to the consequences that are implied by our interpretation of Mamdani's model.

2 Vague Environments

In engineering applications we have in general to deal with real–valued measurements and control actions in quantified form. Also in medical diagnosis so called reference values are very common and the medical doctor tries to find the best fitting reference values in a look–up table in order to determine the optimal dose of a medicine for a certain patient. We should be aware of the fact that the involved real numbers can never be exact. Of course, in many applications the inexactness is small enough so that we do not have to worry about it.

In the following we will provide a model that is able to represent this inexactness in order to handle problems connected to this phenomenon. Two different forms of inexactness can be distinguished:

- enforced inexactness of measurement and control values which is caused by the limited precision of measuring or other instruments or by properties of the physical environment which make an exact measurement impossible.

- intended imprecision where we are not interested in arbitrary exactness or where it even does not make sense. As an example consider the room temperature. A difference of $0.0000001°$C of the temperature is neither for a human being of interest nor should it influence the air conditioning system.

A straight forward approach to model the above mentioned phenomena of inexactness is to identify values whose distance is less than an error– or tolerance bound $\varepsilon > 0$. This identification can lead to problems since it does not satisfy the law of transitivity, i.e. although x_1 and x_2 as well as x_2 and x_3 are identified according to $|x_1 - x_2| \leq \varepsilon$ and $|x_2 - x_3| \leq \varepsilon$, it is possible that x_1 and x_3 should not be identified due to $|x_1 - x_3| > \varepsilon$. This phenomenon is also known as the Poincaré paradox. A typical consequence from this non–transitivity can be experienced for example in the following situation. The decision to buy a certain luxurious car does in general not depend on an increase of the price of 1\$. But it is of course not allowed to iterate this argument, otherwise we would accept any price which is not true.

A consequence of this non–transitivity is that it is impossible to define adequate equivalence classes of indistinguishable or identifiable numbers. The most common approach simply partitions the real numbers into disjoint intervals of a certain length and identifies values that fall in the same interval. This leads automatically to incoherent treatment of values that are near the boundary of an interval.

The choice of an appropriate ε can be a very crucial point in applications. Therefore, we will consider a whole set of such error bounds, namely the unit interval. At first, this choice might look a little bit arbitrary. But as we will see, it can cover the most general case.

It is of course not sufficient, simply to say that two values are indistinguishable or similar if their distance is less than ε, or in other words, they are distinguishable if their distance is greater than ε. This would lead to the paradoxical situation that two temperatures are indistinguishable if they are measured in Celsius, whereas the same temperatures are distinguishable when measured in Fahrenheit according to the greater scaling factor for Fahrenheit. For this reason we allow to introduce a scaling factor $c \geq 0$ and consider two values as ε–distinguishable if their distance times c is greater than ε. Let us assume that the set of numbers we are dealing with (f.e. a set of possible temperatures) is the interval $X = [a, b]$. Introducing a scaling factor for ε–distinguishability corresponds to a transformation of the interval $[a, b]$ to the interval $[0, c \cdot (b - a)] \subseteq [0, \infty)$. The transformation is given by

$$t_c : [a, b] \to [0, \infty), \qquad x \mapsto c \cdot (x - a). \tag{1}$$

In order to decide whether two values $x_1, x_2 \in [a, b]$ are indistinguishable we can, instead of measuring their distance directly in $[a, b]$ and multiplying it by c, take the distance of their transformed values, i.e. $|t_c(x_1) - t_c(x_2)|$.

Although by using such a scaling factor c we can overcome the problem of different scalings as in the case of Fahrenheit and Celsius, we are not able to model the fact that a measurement instrument might provide quite precise values in a certain

temperature (in °C)	scaling factor	interpretation
< 15	0	don't care (much too cold)
15–19	0.25	too cold but nearly o.k., not too sensitive
19–23	1.5	very sensitive, near optimum
23–27	0.25	too hot but nearly o.k., not too sensitive
> 27	0	don't care (much too hot)

Table 1: Scaling factors for the room temperature.

range whereas out of this range the measured values are less reliable. Also in the case of intended inexactness we might wish to distinguish between values in a certain range very carefully, but for other ranges we are not interested in precise values. To solve this problem of differing precision for different ranges we introduce varying scaling factors for the ranges. A scaling factor $c > 1$ implies a weak indistinguishability (strong distinguishability) for values in the corresponding range, whereas a scaling factor $c < 1$ leads to a strong indistinguishability. Consider for example the case of the room temperature. If we are interested in keeping the room temperature at a comfortable value, we might choose the scaling factors shown in table 1.

The great scaling factor 1.5 for the range of temperatures between 19°C and 23°C indicates that these temperatures are distinguished very sensitively and we are able to carry out a fine adjustment in order to meet the optimal temperature. For temperatures lower than 19°C but higher than 15°C we are not so interested in the exact temperature since such temperatures are considers as too cold. For temperatures even below 15°C we are not at all interested in the concrete value because for these frosty values we simply have to heat as much as possible independent from the concrete value of the temperature. Temperatures above 23°C are treated analogously. Let us assume that $X = [0, 35]$ is the set of possible temperatures. The function $c : X \to [0, \infty)$, assigning to each temperature the corresponding scaling factor, is shown figure 1. The corresponding transformation induced by these scaling factors is illustrated in Figure 2.

It is easy to check that the piecewise linear transformation in figure 2 from $X = [a, b] = [0, 35]$ to $[0, \infty)$ can be computed by

$$t_c : [a, b] \to [0, \infty), \qquad x \mapsto \int_a^x c(s)ds \tag{2}$$

where the function c is given by

$$c : X \to [0, \infty), \qquad s \mapsto \begin{cases} 0 & \text{if} \quad 0 \leq s < 15 \\ 0.25 & \text{if} \quad 15 \leq s < 19 \\ 1.5 & \text{if} \quad 19 \leq s < 23 \\ 0.25 & \text{if} \quad 23 \leq s < 27 \\ 0 & \text{if} \quad 27 \leq s < 35. \end{cases} \tag{3}$$

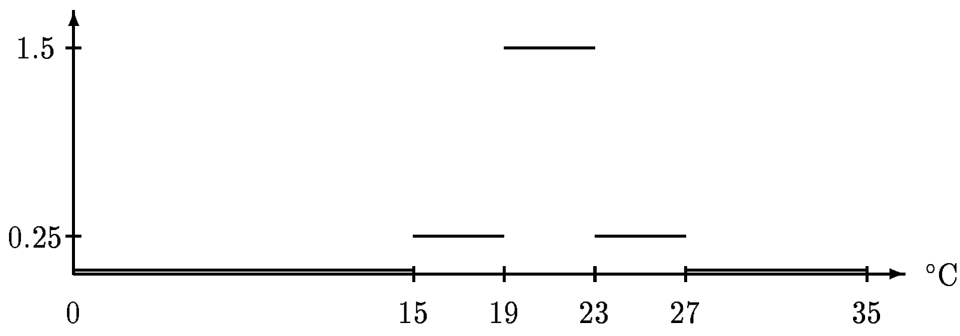

Figure 1: The scaling factor function for the room temperature example.

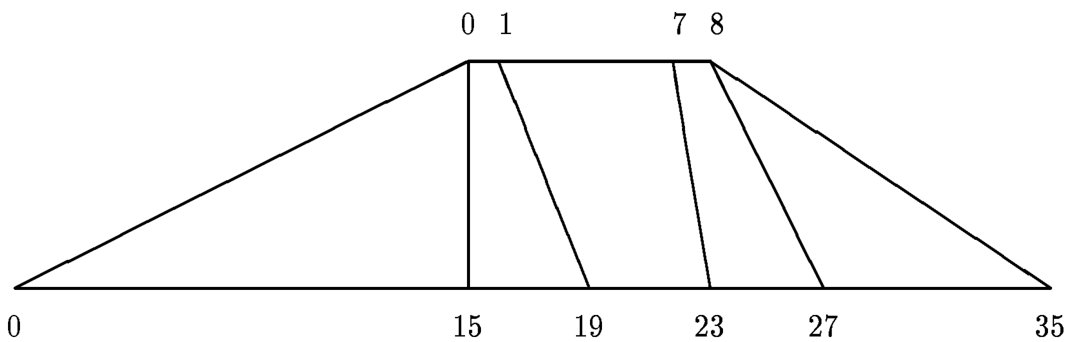

Figure 2: The transformation induced by the scaling function in figure 1.

If we choose $\varepsilon = 0.5$ then the temperatures 20.1°C and 20.5°C are ε–distinguishable whereas 16°C and 12°C are ε–indistinguishable. By equation (2) the transformed values for 12, 16, 20.1, and 20.5 are 0, 0.25, 2.65, and 3.25, respectively.

Note that equation (2) coincides with equation (1) when we choose a constant scaling function c.

The scaling function in figure 1 reflects the idea that we distinguish values near the optimal temperature very carefully, whereas the distinguishability decreases the farther away we go from the optimal value. The piecewise linear function was only chosen to elucidate the principle of different scaling factors and to have a simple transformation function. In the most general case we associate with each value x of our set $X = [a, b]$ a scaling factor $c(x) \geq 0$. The function c has not to be piecewise linear. All we have to assume is that c is integrable. For the transformation induced by such a general scaling function equation (2) is still valid. The distance $\delta_c(x_1, x_2)$ of the transformed values of x_1 and x_2 is given by

$$\delta_c(x_1, x_2) = \left| \int_a^{x_1} c(s)ds - \int_a^{x_2} c(s)ds \right| = \left| \int_{x_1}^{x_2} c(s)ds \right|. \tag{4}$$

x_1 and x_2 are considered to be ε–distinguishable with respect to the scaling function c if their 'transformed distance' $\delta_c(x_1, x_2)$ is greater than ε.

We now turn to the problem of representing a *vague environment* that is characterized by a distance function δ_c of the above mentioned type. We do not consider only one fixed value ε, but a whole set of values for ε, each of them leading to a different ε–distinguishability. We consider all numbers from the unit interval as possible values for ε. If one would prefer to have a smaller or larger interval as possible values for ε, this can be amended by an appropriate choice of the scaling function c. If for example the scaling function c is replaced by the scaling function $\hat{c} = \lambda \cdot c$ then ε–distinguishability with respect to c corresponds to (ε/λ)–distinguishability with respect to $\hat{c}$. In this sense allowing all values from the unit interval for ε covers already the most general case.

For each $\varepsilon \in [0, 1]$ we associate with the value $x_0 \in X$ all values $x \in X$ which are not ε–distinguishable from x_0 (with respect to the scaling function c), i.e. the set

$$S_{x_0,\varepsilon} = \{x \in X \mid \delta_c(x, x_0) \leq \varepsilon\}. \tag{5}$$

A more convenient representation of this family of sets is described by the mapping

$$\mu_{x_0} : X \to [0, 1], \quad x \mapsto 1 - \min\{\delta_c(x, x_0), 1\}, \tag{6}$$

so that we have

$$S_{x_0,\varepsilon} = \{x \in X \mid \mu_{x_0}(x) \geq 1 - \varepsilon\}.$$

$\mu_{x_0}(x)$ can be interpreted intuitively as the degree to which x can be identified with x_0. Therefore we can understand μ_{x_0} as the fuzzy set of values that are indistinguishable to x_0. Note that in the most simple case where we have the same scaling factor $c > 0$ for all $x \in X$, i.e. a constant scaling function, we obtain a triangular

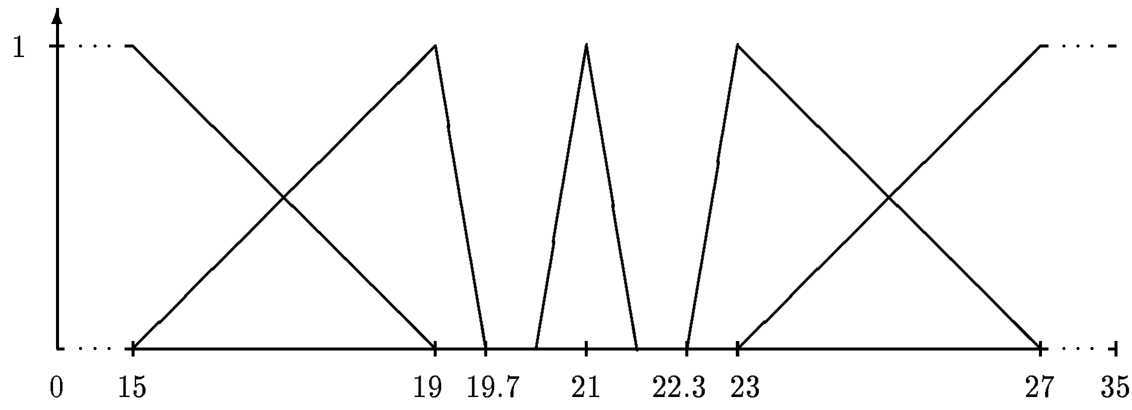

Figure 3: The fuzzy sets μ_{x_0} associated with the values $x_0 = 15, 19, 21, 23, 27$ in the vague environment induced by the scaling function in equation (3).

membership function with slope c taking its maximum at x_0 as the fuzzy set μ_{x_0} which represents the value x_0 in the vague environment X.

Note that the α–cut $\{x \in X \mid \mu_{x_0} \geq \alpha\}$ of the fuzzy set μ_{x_0} is equal to the set $S_{x_0,1-\alpha}$ of elements that are $(1-\alpha)$–indistinguishable from x_0.

Let us return to the vague environment for the room temperature example characterized by the scaling factors in equation (3). Figure 3 illustrates the fuzzy sets μ_{x_0} that are associated with the values $x_0 = 15, 19, 21, 23, 27$ in this vague environment.

The fuzzy sets in figure 3 are all of triangular or trapezoidal type. This is not necessarily the case as figure 4 illustrates where the fuzzy sets μ_{x_0} associated with the values $x_0 = 18$ and $x_0 = 22.5$ are shown. Since the scaling function in equation (3) is piecewise constant, the fuzzy set μ_{x_0} associated with a value x_0 will always be piecewise linear.

To obtain other shapes for the fuzzy set associated with the value x_0, an appropriate scaling function has to be defined. As an example let us consider a bell shaped fuzzy set of the form

$$\mu : \mathbb{R} \to [0,1], \qquad x \mapsto \exp\left(-\frac{1}{2}\left(\frac{x-x_0}{\sigma}\right)^2\right). \tag{7}$$

Choosing

$$c : \mathbb{R} \to [0,\infty), \qquad x \mapsto \frac{|x-x_0|}{\sigma^2} \cdot \exp\left(-\frac{1}{2}\left(\frac{x-x_0}{\sigma}\right)^2\right) \tag{8}$$

as the scaling function, we obtain $\mu = \mu_{x_0}$, i.e. the fuzzy set μ represents the value x_0 in the vague environment induced by the scaling function c. This is a direct consequence of equation (4) for the transformed distance, since (8) is simply the absolute value of the first derivate of (7).

More generally we can state the following theorem.

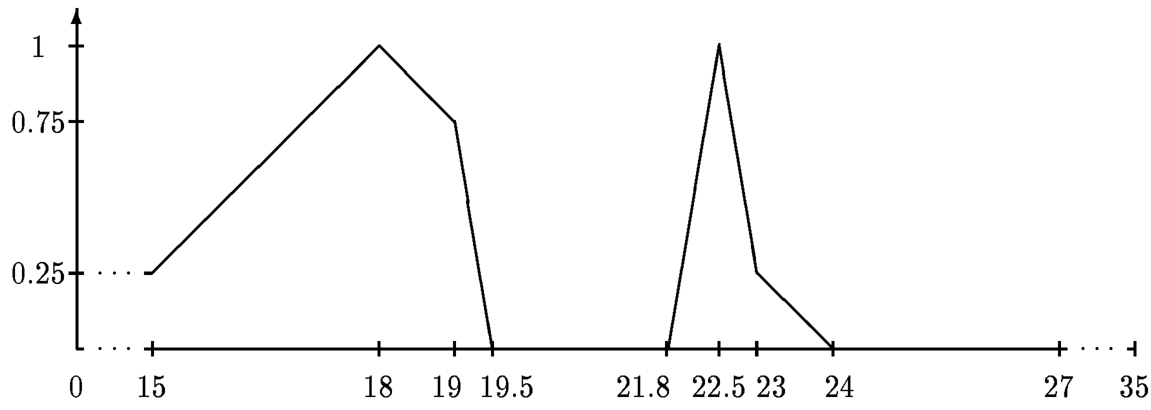

Figure 4: The fuzzy sets μ_{x_0} associated with the values $x_0 = 18$ and $x_0 = 22.5$ in the vague environment induced by the scaling factors in equation (3).

Theorem 2.1 *Let* $\mu : \mathbb{R} \to [0,1]$ *be a fuzzy set such that there exists* $x_0 \in \mathbb{R}$ *with*

(i) $\mu(x_0) = 1$,

(ii) μ *is a non–decreasing function on* $(-\infty, x_0]$,

(iii) μ *is a non–increasing function on* $[x_0, \infty)$,

(iv) μ *is continuous,*

(v) μ *is almost everywhere differentiable.*

Then there exists a scaling function $c : \mathbb{R} \to [0, \infty)$ *such that* μ *coincides with the fuzzy set* μ_{x_0} *which is associated with the value* x_0 *in the vague environment induced by* c.

Proof. Choose $c(x) = |\mu'(x)| = \left|\frac{d\mu}{dx}\right|$ as the scaling function. □

It is obvious, that the reverse of theorem 2.1 also holds, which means that, given a scaling function $c : \mathbb{R} \to [0, \infty)$ and a value x_0, then the fuzzy μ_{x_0} associated with x_0 in the vague environment induced by c satisfies conditions (i) – (v) of theorem 2.1.

Conditions (i) – (iii) guarantee that the fuzzy set is fuzzy convex (i.e. all its α–cuts are convex), so that it can be considered as the representation of a single value in a vague environment. Non–fuzzy convex sets cannot appear in vague environments when fuzzy sets stand only for single values. It is, of course, possible not just to associate with a single value a fuzzy set in a vague environment, but to associate

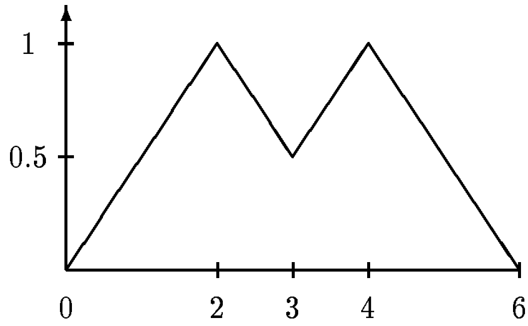

Figure 5: The fuzzy set $\mu_{\{2,4\}}$ in the vague environment induced by the constant scaling function $c = 0.5$.

with any set of values a corresponding fuzzy set by generalizing equations (5) and (6) for a set $M \subseteq X$ by

$$S_{M,\varepsilon} = \{x \in X \mid \exists x_0 \in M : \ \delta_c(x, x_0) \leq \varepsilon\}.$$

and

$$\mu_M : X \to [0,1], \quad x \mapsto 1 - \min\left\{\inf_{x_0 \in M}\{\delta_c(x, x_0)\}, 1\right\},$$

respectively. Figure 5 illustrates an example for such a non–fuzzy convex fuzzy set associated with the set $M = \{2,4\}$ in the vague environment induced by the constant scaling function $c = 0.5$.

This example shows that also non–fuzzy convex fuzzy sets can appear in vague environments, when sets of values instead of single values are considered. However, in this paper we will not pursue this topic in detail.

The concept of scaling factors for the transformation enforces conditions (iv) and (v) of theorem 2.1. If we do not insist on transformations induced by scaling functions, we may allow as a generalization of equation (2) any non–decreasing function $t : [a,b] \to [0,\infty)$ as transformation, where t is not necessarily continuous. Such transformation functions are discussed in connection with equality relations, which we will relate to vague environments at the end of this section to our approach, in [6]. An example of a non–continuous transformation is illustrated in figure 6. The

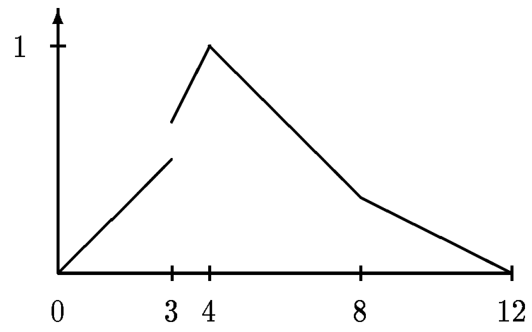

Figure 6: A non–continuous fuzzy set representing the value $x_0 = 4$ in the vague environment induced by the transformation function (9).

underlying transformation is

$$t : [0, 12] \to [0, \infty), \qquad x \mapsto \begin{cases} x/6 & \text{if } \; 0 \leq s < 3 \\ 2/3 + (x-3)/3 & \text{if } \; 3 \leq s < 4 \\ 1 + (x-4)/6 & \text{if } \; 4 \leq s < 8 \\ 5/3 + (x-8)/12 & \text{if } \; 8 \leq s < 12 \end{cases} \tag{9}$$

so that the transformed distance $\delta(x_1, x_2)$ of x_1 and x_2 is given by $\delta(x_1, x_2) = |t(x_1) - t(x_2)|$. Based on equation (6) the fuzzy set shown in figure 6 is induced by the value $x_0 = 4$.

Up to now we have only considered fuzzy sets as representations of crisp values in vague environments that were described by scaling functions. In this way a fuzzy partition as in figure 3 is induced by a set of crisp values together with a scaling function. We now turn to the question whether we can provide a vague environment for a given fuzzy partition such that the corresponding fuzzy sets can be interpreted as representations of crisp values in this vague environment. The following theorem answers this questions.

Theorem 2.2 *Let $(\mu_i)_{i \in I}$ be an at most countable family of fuzzy sets on $\mathbb{R}$ and let $(x_0^{(i)})_{i \in I}$ be a family of real numbers such that $\mu_i(x_0^{(i)}) = 1$ holds and the conditions (i) – (v) of theorem 2.1 are satisfied for all $i \in I$. There exists a scaling function $c : \mathbb{R} \to [0, \infty)$ such that μ_i coincides with the fuzzy set $\mu_{x_0^{(i)}}$ (for each $i \in I$), which is associated with the value $x_0^{(i)}$ in the vague environment induced by c, if and only if*

$$\min\{\mu_i(x), \mu_j(x)\} > 0 \quad \Rightarrow \quad |\mu_i'(x)| = \left|\mu_j'(x)\right| \tag{10}$$

holds almost everywhere for all $i, j \in I$.

Proof. Assume that (10) is satisfied. Define the scaling function

$$c : \mathbb{R} \to [0, \infty), \qquad x \mapsto \begin{cases} |\mu_i'(x)| & \text{if } \mu_i(x) > 0 \\ 0 & \text{otherwise.} \end{cases}$$

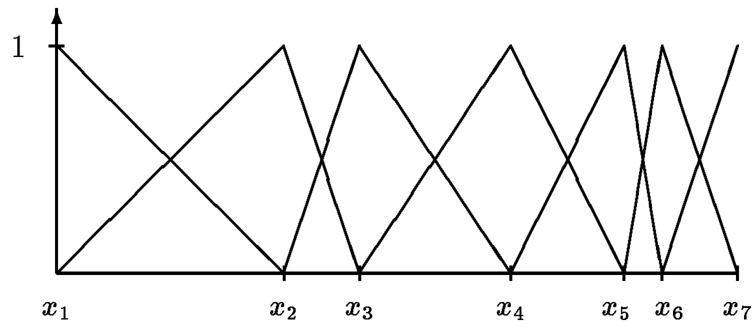

Figure 7: A typical fuzzy partition for which a scaling function can be defined easily.

(2.2) guarantees that c is well defined almost everywhere. Theorem 2.1 yields that $\mu_i = \mu_{x_0^{(i)}}$ holds for all $i \in I$. Note that it is sufficient for the proof of theorem 2.1 that the scaling function coincides with the derivate of the fuzzy set only on the support of the fuzzy set.

In order to prove the reverse implication, we assume now that there is a scaling function $c : \mathbb{R} \to [0, \infty)$ such that $\mu_i = \mu_{x_0^{(i)}}$ holds for all $i \in I$. Let $i, j \in I$ and let $x \in \mathbb{R}$ with $\min\{\mu_i(x), \mu_j(x)\} > 0$. By definition we have

$$\mu_k(x) \;=\; 1 - \left| \int_{x_0^{(k)}}^{x} c(s)ds \right|$$

for $k \in \{i, j\}$, which implies

$$|\mu_k'(x)| \;=\; c(x)$$

if μ_k is differentiable at x. Since μ_i and μ_j are almost everywhere differentiable, we obtain

$$|\mu_i'(x)| \;=\; c(x) \;=\; |\mu_j'(x)|$$

almost everywhere. □

Theorem 2.2 simply states that we can find a corresponding scaling function for a given fuzzy partition if for each real number $x \in \mathbb{R}$ the absolute value of the slope at x is the same for all fuzzy sets in the fuzzy partition whenever x belongs to the support of the fuzzy set.

A very common type of fuzzy partition is obtained by choosing crisp values $x_1 < x_2 < \ldots < x_n$ and defining the fuzzy set μ_i (for $1 < i < n$) by a triangular membership function which takes its maximum at x_i and reaches the value zero at x_{i-1} and x_{i+1}, respectively. Such a fuzzy partition is illustrated in figure 7.

For such fuzzy partitions the corresponding scaling function can be defined as the piecewise constant function

$$c(x) \;=\; \frac{1}{x_{i+1} - x_i} \qquad \text{if } x_i < x < x_{i+1},$$

so that the fuzzy sets μ_i represent the values x_i in the vague environment induced by c.

This section has introduced fuzzy sets as representations of crisp values in vague environments. A vague environment is characterized by a scaling function $c : X \to [0, \infty)$. The value $c(x)$ specifies how sensitive we have to distinguish between values in the neighbourhood of x.

As theorem 2.2 shows, it is not only possible to generate fuzzy sets by a scaling function together with crisp values, but we can also derive in many cases a corresponding scaling function from a given fuzzy partition. A fuzzy set determines implicitly a corresponding scaling function by its first derivate.

The membership degrees in the unit interval are connected to error– or tolerance bounds where the membership degree $\alpha \in [0, 1]$ is associated with the tolerance bound $\varepsilon = 1 - \alpha$. In this sense a fuzzy set μ_{x_0} with $\mu_{x_0}(x_0) = 1$ represents the value x_0 in different consideration contexts where each context is associated with a tolerance bound. The α–cut of the fuzzy set μ_{x_0} stands for the value x_0 in the context with tolerance bound $\varepsilon = 1 - \alpha$ in which we identify all values with x_0 whose (transformed) distance to x_0 is not greater than ε. In the following section we will make use of this context view to develop an interpolation technique in vague environments which can be applied to fuzzy control.

The principal behind the concept of a scaling function c is the definition of a transformation which induces a (pseudo–)metric or modified distance measure δ_c as in equation (4). ε–distinguishability is determined on the basis of this distance measure δ_c. For a fixed value $x_0 \in \mathbb{R}$ the membership degree of $x \in \mathbb{R}$ to the fuzzy set μ_{x_0} is 1 minus the distance between x_0 and x with respect to δ_c. $\mu_{x_0}(x)$ can also be interpreted as the degree to which x_0 and x are equal or similar.

The concept of scaling functions is used in this paper to have an intuitively appealing and easy to carry out approach for defining transformed distances. In principal it is possible to generalize this idea and to specify the transformed distance directly in the form of a pseudo–metric δ which induces a similarity or equality relation E_δ by $E(x_1, x_2) = 1 - \min\{\delta(x_1, x_2), 1\}$. An equality relation is a mapping $E : X \times X \to [0, 1]$ satisfying

(i)	$E(x, x) = 1$	(total existence)
(ii)	$E(x, y) = E(y, x)$	(symmetry)
(iii)	$E(x, y) * E(y, z) \leq E(x, z)$	(transitivity)

where $*$ stands for the Łukasiewicz t–norm given by $\alpha * \beta = \max\{\alpha + \beta - 1, 0\}$. $E(x, y)$ is interpreted as the degree to which x and y are equal. With respect to such an equality relation the fuzzy set μ_{x_0} represents the extensional hull of the set $\{x_0\}$, i.e. the smallest extensional fuzzy set with $\mu_{x_0}(x_0) = 1$. Extensionality means that the fuzzy set respects the equality relation in the sense $\mu_{x_0}(x) * E(x, y) \leq \mu_{x_0}(y)$, i.e. whenever x has a non–zero membership grade to the fuzzy set μ_{x_0} and x and y

are equal to some degree, then also y should belong to μ_{x_0} to a certain extend. For details on equality relations we refer to [4, 10, 16, 17].

Considering arbitrary equality relations that are not necessarily induced by scaling functions makes it possible to loosen the condition in theorem 2.1. In this way equality relations for less restrictive types of fuzzy partitions can be provided, so that the fuzzy sets of the fuzzy partitions represent crisp values as their extensional hulls. In [6] it is shown that having only conditions (i) – (iii) of theorem 2.1 for the fuzzy sets of the fuzzy partition we can derive a corresponding equality relation which is representable as the infimum of equality relations that are induced by arbitrary, possibly non–continuous transformations. Other types of transformations are discussed in [15].

Equality relations that do not have to be induced by transformations are considered in [5] with respect to the Łukasiewicz t–norm and for arbitrary t–norms in [12]. The main result for the general case of arbitrary equality relations is that a fuzzy partition is induced by an equality relation and a set of crisp points if and only if for any two fuzzy sets of the fuzzy partition their non–disjointness degree is less than or equal to the degree that these two fuzzy sets are equal. However, even in the case that there exists a corresponding scaling function for a given fuzzy partition, the equality relation derived from a fuzzy partition in [5, 11, 12] is in general coarser than the one induced by the corresponding scaling function [8].

3 Interpolation in Vague Environments

Now, after having introduced the concept of a vague environment in the previous section and having discussed the connections to fuzzy sets, we can apply these notions to interpolation in vague environments. At the end of this section we will apply this interpolation technique to control an we will rediscover the well known Mamdani type fuzzy control method.

Interpolation in vague environments is based on the following ideas. We are looking for a function $\varphi : X_1 \times \ldots \times X_n \to Y$ that associates with each input tuple $(x_1, \ldots, x_n) \in X_1 \times \ldots \times X_n$ an appropriate output value $y = \varphi(x_1, \ldots, x_n)$. The domains $X_1, \ldots, X_n$, and Y are considered as vague environments, i.e. we have to specify corresponding scaling factors as proposed in the previous section.

Generally, we are not able to define the function φ (otherwise we would have solved the control problem). But we might be able to provide the output value not for all but for certain input tuples, i.e. we know the output value

$$y^{(i)} = \varphi(x_1^{(i)}, \ldots, x_n^{(i)}) \qquad (i = 1, \ldots, r)$$

for the r input tuples $(x_1^{(1)}, \ldots, x_n^{(1)}), \ldots, (x_1^{(r)}, \ldots, x_n^{(r)})$.

But we are still in trouble if we have to define an output value y for an input tuple $(x_1, \ldots, x_n)$ for which the output value is not specified. We somehow have to interpolate the function φ by using the partially specified function and the properties of the vague environments. Thus we have to face the question, what can we do when we are given the input $(x_1, \ldots, x_n)$ that is different from the tuples $(x_1^{(i)}, \ldots, x_n^{(i)})$ $(i = 1, \ldots, r)$?

We make the following assumption. If x_1 and $x_1^{(i)}$, as well as x_2 and $x_2^{(i)}$ as well as ..., as well as x_n and $x_n^{(i)}$ are not ε–distinguishable, then it is reasonable to choose an output value y that is not ε–distinguishable from the output value $y^{(i)}$ for the tuple $(x_1^{(i)}, \ldots, x_n^{(i)})$, where $i \in \{1, \ldots, r\}$.

To illustrate this, we consider the input tuple $(x_1, \ldots, x_n)$ and restrict at first to only one tuple $(x_1^{(i)}, \ldots, x_n^{(i)})$. Let $\varepsilon \in [0,1]$. If x_1 and $x_1^{(i)}$, as well as x_2 and $x_2^{(i)}$ as well as ..., as well as x_n and $x_n^{(i)}$ are not ε–distinguishable, then we obtain the set of elements that are not ε–distinguishable from $y^{(i)}$ as reasonable output values. But if at least one of the pairs $(x_1, x_1^{(i)})$, ..., $(x_n, x_n^{(i)})$ is ε–distinguishable, we gain no information about the output value by the above assumption.

In other words, we consider an output value y as appropriate on the ε–distinguishability level or context (with respect to the specified input–output tuple $((x_1^{(i)}, \ldots, x_n^{(i)}), y^{(i)}))$ if y and $y^{(i)}$, as well as x_1 and $x_1^{(i)}$, as well as ..., as well as x_n and $x_n^{(i)}$ are not ε–distinguishable. We can characterize this information by making use of the representation of the elements that are indistinguishable to a certain value x_0 in the form of a fuzzy set as it was introduced in section 2. We associate with y the value

$$\mu^{\text{output}}_{(x_1,\ldots,x_n);i}(y) = 1 - \inf\{\varepsilon \in [0,1] \mid y \text{ is appropriate on the } \varepsilon\text{-distinguishability level}\}$$

which can be understood as the maximal appropriateness degree of y with respect to the input–output tuple $((x_1^{(i)}, \ldots, x_n^{(i)}), y^{(i)})$. In terms of the fuzzy set representation of the values $x_1, \ldots, x_n, y$ in the vague environments $X_1, \ldots, X_n, Y$, respectively, we obtain the equation

$$\mu^{\text{output}}_{(x_1,\ldots,x_n);i}(y) = \min\{\mu_{x_1^{(i)}}(x_1), \ldots, \mu_{x_n^{(i)}}(x_n), \mu_{y^{(i)}}(y)\}.$$

Since we have to take into account all specified input–output tuples in our model, we consider an output value y as appropriate on the ε–distinguishability level if y is appropriate on the ε–distinguishability level with respect to at least one of the specified input–output tuples. If we characterize the maximal appropriateness degree of y as a fuzzy set in the same way as we did it for one input–output tuple, we obtain the fuzzy set

$$\mu^{\text{output}}_{(x_1,\ldots,x_n)}(y) = \max_{i\in\{1,\ldots,r\}}\{\mu^{\text{output}}_{(x_1,\ldots,x_n);i}(y)\}.$$

This interpolation technique does not provide a unique function φ. It does only enforce constraints on φ. The basic idea behind this interpolation technique is the following. For a given input tuple $(x_1, \ldots, x_n) \in X_1 \times \ldots \times X_n$ choose one of the tuples $(x_1^{(i)}, \ldots, x_n^{(i)})$ for which the output $y^{(i)}$ was specified and choose an appropriate ε such that $(x_1, \ldots, x_n)$ and $(x_1^{(i)}, \ldots, x_n^{(i)})$ are ε–indistinguishable. Then you may take $y \in Y$ as an output value for $(x_1, \ldots, x_n)$ if y is ε–indistinguishable from $y^{(i)}$.

Now we can apply this interpolation technique to control. We consider a control problem in the following way. We take measurements of n input variables $\xi_1, \ldots, \xi_n$ with $X_1, \ldots, X_n$ as underlying domains and and we have to determine the value of one output or control variable η with underlying domain Y. We are looking for a control function $\varphi : X_1 \times \ldots \times X_n \to Y$ that assigns to each input tuple an appropriate control value.

In order to apply our interpolation technique, scaling functions for the domains $X_1, \ldots, X_n, Y$ have to be specified. In addition we need a partial control function that determines the control value $y^{(i)}$ for some input tuples $(x_1^{(i)}, \ldots, x_n^{(i)})$, $(i = 1, \ldots, r)$. This partial control function may be given by a control expert or derived from experimental data.

We can now establish the connection between our technique and Mamdani's fuzzy control model. Each value $x_k^{(i)}$ and $y^{(i)}$ of the specified tuples is associated with a fuzzy set in the corresponding vague environment as it is explained in section 2. In this way we obtain a fuzzy partition of the domain X_k by the fuzzy sets that are associated with the values $x_k^{(i)}$ $(i = 1, \ldots, r)$. Note that some of the $x_k^{(i)}$ may be equal. The $y^{(i)}$ induce a fuzzy partition of Y.

For a specified input–output tuple $\left((x_1^{(i)}, \ldots, x_n^{(i)}), y^{(i)}\right)$ we consider the linguistic control rule

> *If ξ_1 is (approximately) $x_1^{(i)}$ and ... and ξ_n is (approximately) $x_n^{(i)}$ then η is (approximately) $y^{(i)}$. $(i = 1, \ldots, r)$*

where we associate with the linguistic terms (approximately) $x_1^{(i)}$, ..., (approximately) $x_n^{(i)}$, and (approximately) $y^{(i)}$ the fuzzy sets $\mu_{x_1^{(i)}}$, ..., $\mu_{x_n^{(i)}}$, and $\mu_{y^{(i)}}$, respectively. $\xi_1, \ldots, \xi_n$ are the input variables, η is the output variable.

The important observation that we can make now is that the fuzzy set $\mu^{\text{output}}_{(x_1,\ldots,x_n);i}$ derived by interpolation in vague environments corresponds exactly to the output of the above mentioned single control rule. Moreover, the fuzzy set $\mu^{\text{output}}_{(x_1,\ldots,x_n)}$ obtained by interpolation in vague environments is exactly the output fuzzy set of the max–min fuzzy controller based on the r control rules of the above mentioned type.

Example 3.1 To elucidate the connection between the max–min fuzzy controller and the concept of interpolation in vague environments, we consider a simple fuzzy controller with two input variables ξ_1 and ξ_2 and one output variable η. The underlying domains are $X_1 = [0, 11]$, $X_2 = [0, 9]$, and $Y = [0, 4]$. The fuzzy partitions of X_1, X_2, Y are shown in figures 8 – 10, respectively. Table 2 shows the rule base of the fuzzy controller.

From the fuzzy partitions we can derive the following scaling functions.

$$c_{X_1} : [0, 11] \to [0, \infty), \qquad x \mapsto \begin{cases} 0 & \text{if} \quad 0 \le x < 1 \\ 1/3 & \text{if} \quad 1 \le x < 10 \\ 0 & \text{if} \quad 10 \le x < 11 \end{cases}$$

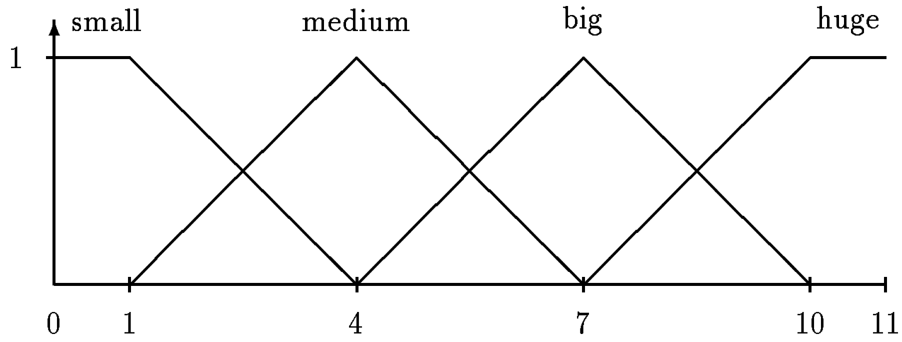

Figure 8: The fuzzy partition of X_1

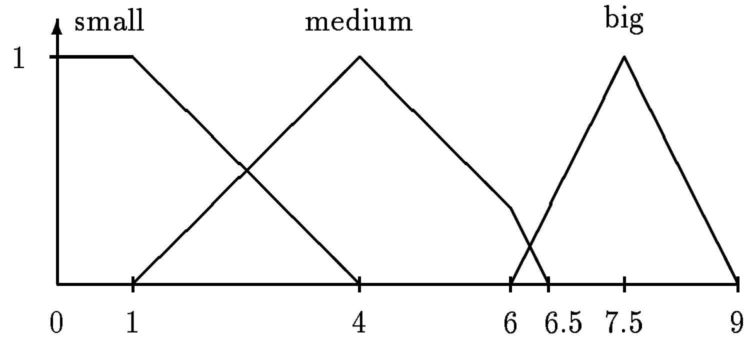

Figure 9: The fuzzy partition of X_2.

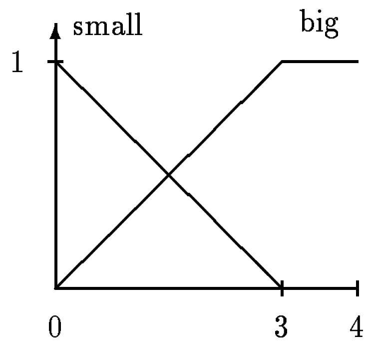

Figure 10: The fuzzy partition of Y.

	small	medium	big	huge
small	small	small		big
medium	small		big	big
big		big	big	big

Table 2: The rule base of a fuzzy controller.

	small	medium	big	huge
X_1	1	4	7	10
X_2	1	4	7.5	
Y	0		3	

Table 3: The values associated with the fuzzy sets.

$$c_{X_2} : [0,9] \to [0,\infty), \qquad x \mapsto \begin{cases} 0 & \text{if} \quad 0 \le x < 1 \\ 1/3 & \text{if} \quad 1 \le x < 6 \\ 2/3 & \text{if} \quad 6 \le x < 9 \end{cases}$$

$$c_Y : [0,4] \to [0,\infty), \qquad y \mapsto \begin{cases} 1/3 & \text{if} \quad 0 \le y < 3 \\ 0 & \text{if} \quad 3 \le y < 4 \end{cases}$$

The values associated with the fuzzy sets are shown in table 3. The fuzzy sets represent the corresponding values in the vague environment induced by the respective scaling function.

From tables 2 and 3 we obtain the partial control function described in table 4.

This example will also be used to illustrate the results in section 4.

Let us summarize the results of this section. We have introduced an intuitively appealing and plausible approach to control based on the concept of interpolation in

φ	1	4	7	10
1	0	0		3
4	0		3	3
7.5		3	3	3

Table 4: The partial control function.

vague environments. Although at first glance there is no connection to Mamdani's fuzzy control model, it turns out that our method leads to the same computations as the max–min rule (before defuzzification). For reasons of simplicity we will not discuss defuzzification strategies here.

In Mamdani's model fuzzy partitions and linguistic control rules have to be specified. In our approach we need characterizations of the vague environments in the form of appropriate scaling functions and a partial control mapping. We showed that we can always interpret a controller based on interpolation in vague environments as a max–min fuzzy controller. In most cases it also possible to translate a max–min fuzzy controller to a controller based on interpolation in vague environments. For this translation it is necessary to derive appropriate scaling functions from the specified fuzzy partitions and to transform the rule base into a corresponding partial control function.

Corresponding scaling functions for a vague environment can be derived from a fuzzy partition by applying theorem 2.2, which requires the fuzzy partition to fulfill certain constraints, that are satisfied in many applications.

The consequences of this equivalence between Mamdani's and our control model for the design and tuning of a fuzzy controller will be discussed in the following section.

4 Consequences for Fuzzy Control

After the first design of a fuzzy controller it is in general necessary to tune this controller in order to obtain satisfactory or optimal control actions. Apart from drastic changes like altering operations (for example taking other t–norms or t–conorms than min or max, respectively), varying the defuzzification strategy or a total redesign of the controller, usually only small changes are considered. We will discuss in this paper the impacts of tuning membership functions and changing the rule base in the view of interpolation in vague environments, where only the partial control mapping and the scaling functions can be tuned.

Let us first turn to the problem of tuning membership functions. If we fix the level 1 of the membership function and only change the width of fuzzy sets, this corresponds in our model to a change in the vague environments, i.e. a variation of the scaling functions. In order to maintain the possibility of translating the Mamdani fuzzy controller to vague environments, it is necessary to change neighbouring fuzzy sets of a fuzzy partition accordingly, which means for example to guarantee that the slopes have the same absolute values. Figure 11 illustrates a consistent change of two fuzzy sets that corresponds to an increase of the scaling function between 4 and 7 by the factor 1.5. The dotted lines indicate the changed fuzzy sets.

A change of the support of the membership functions and maintaining their shapes does not effect the vague environments, but leads to a different partial control function. If a fuzzy set represents the value $x_k^{(i)}$ in the vague environment and it is shifted one unit to the right, it then represents the value $(x_k^{(i)} + 1)$. A simple shift

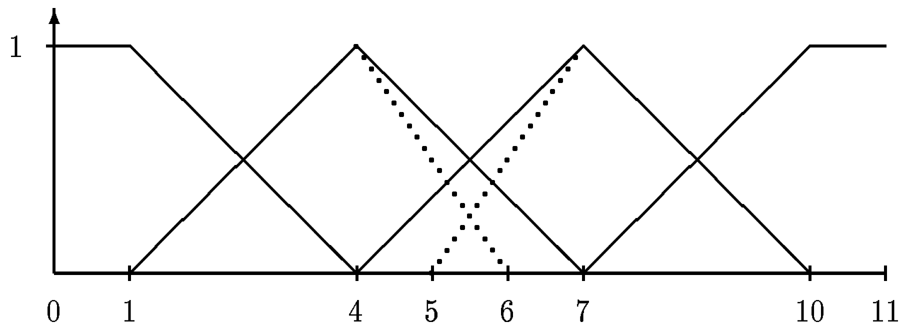

Figure 11: This tuning of the two fuzzy sets representing the linguistic terms medium and big in X_1 corresponds to multiplying the scaling factor between 4 and 7 by 1.5.

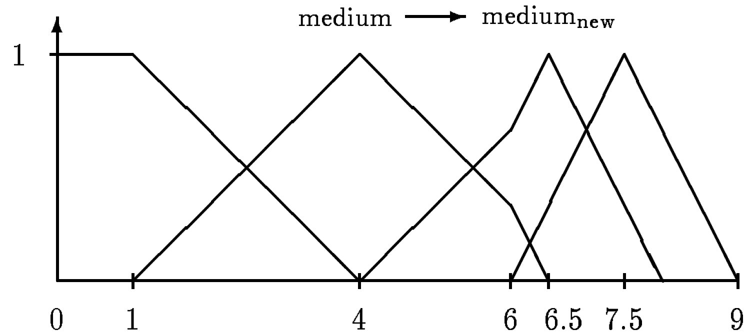

Figure 12: Shifting the fuzzy set for the linguistic term medium in X_2 2.5 units to the right and changing it accordingly to the greater scaling factor in the range between 6 and 9.

of a fuzzy set leads only to a consistent fuzzy partition for which we can derive a scaling function, when condition (10) of theorem 2.2 is still satisfied. Too guarantee this consistency condition, in addition to the shift of the membership function its slopes should be changed accordingly in order to maintain the scaling function. An example is shown in figure 12 where the fuzzy set μ is shifted to the right and it is changed according to the greater scaling factor induced by the fuzzy set ν in the range between 6 and 9.

We should also take into account that due the fuzzy partitions moving a fuzzy set to the left or right does have multiple effects on the partial control function since all rules that contain the the linguistic term corresponding to the fuzzy set are influenced. A Mamdani fuzzy controller is always based on a partial control function which is a subset of a grid in the space $X_1 \times \ldots \times X_n \times Y$. It is not possible to change a single point of the grid, we can on only move whole lines of the grid. If, for example, it turned out that the value we obtain from the rule

If ξ_1 huge and ξ_2 is big then η is big

in example 3.1 is too high, and we shift the fuzzy set for the linguistic term big one unit to the left so that it represents the value two instead of three, we would replace all three's by two's in table 4 for the partial control function, instead of changing only the value in the lower left corner.

All together we can see that tuning membership functions can in general be interpreted as a simultaneous change of the vague environments and the partial control function. Tuning the rule base means to alter the partial control function without influencing the vague environments.

The advantage of translating Mamdani's fuzzy control model to our approach is that we can interpret the changes caused by tuning. We should be aware of the fact that an arbitrary tuning of membership functions leads to a simultaneous change of the vague environments and the partial control function which is in general not a desired effect. When tuning membership functions we should also take the consistency condition into account that is needed in order to be able to derive scaling functions (compare theorem 2.2 and figure 12).

There are also other consequences induced by our approach. For example, we can explain, why it is reasonable to use fuzzy partitions in which the support of of each fuzzy set is chosen in such a way that it exactly covers the range between the points where its neighbouring fuzzy sets reach their maximum. For triangular membership functions this guarantees always the existence of a corresponding scaling function. This strategy also implies that the fuzzy sets are more dense in ranges where they have smaller supports. Since smaller supports induce greater scaling factors, i.e. higher distinguishability, this leads to having more points for interpolation where the distinguishability is high.

Although we did not discuss the problem of defuzzification in this paper, our approach may also be used to find constraints for defuzzification operators. It might be reasonable to require from the control function φ induced by the max–min rule together with the defuzzification strategy that φ does not map values that are ε–indistinguishable to values that are ε–distinguishable. This constraint is expressed by the following inequality.

$$\min\{\delta_{c_{X_1}}(x_1, \tilde{x}_1), \ldots, \delta_{c_{X_n}}(x_n, \tilde{x}_n)\} \;\geq\; \delta_{c_Y}(\varphi(x_1, \ldots, x_n), \varphi(\tilde{x}_1, \ldots, \tilde{x}_n))$$

This is a very restrictive condition and it corresponds to the extensionality of φ with respect to the equality relations induced by the distance functions $\delta_{c_{X_1}}, \ldots, \delta_{c_{X_n}}$, and δ_{c_Y}. For a discussion of this constraint we refer to [11, 12].

5 Conclusions

We have provided a very simple but clear interpretation of Mamdani's fuzzy control method. Making use of this interpretation we can explain the consequences of tuning a fuzzy controller and can judge which tuning strategies are reasonable or might lead to undesired results. It is also possible to see our approach as a stand–alone model and to apply it directly. (Indeed, we have applied it successfully to idle speed control of the Volkswagen GTI engine [9].) Then we can rely on the clear semantics of our model and nevertheless use standard fuzzy hard– and software tools, since we can translate our controller based on interpolation in vague environments to a max–min fuzzy controller.

References

[1] D. Dubois, H. Prade, Possibility Theory. Plenum Press, New York (1988).

[2] D. Dubois, H. Prade, Basic Issues on Fuzzy Rules and their Application to Fuzzy Control. Proc. IJCAI–91 Workshop on Fuzzy Control (1991), 5–17.

[3] E. Hisdal, Are Grades of Membership Probabilities? Fuzzy Sets and Systems 25 (1988), 325–348.

[4] U. Höhle, M–Valued Sets and Sheaves over Integral Commutative CL–Monoids. In: S.E. Rodabaugh, E.P. Klement, U. Höhle (eds.), Applications of Category Theory to Fuzzy Subsets. Kluwer, Dordrecht (1992), 33–72.

[5] U. Höhle, F. Klawonn, Fuzzy Control und Ununterscheidbarkeit. Proc. VDE–Fachtagung Technische Anwendungen von Fuzzy–Systemen. Dortmund (1992), 3–9.

[6] J. Jacas, J. Recasens, Fuzzy Numbers and Equality Relations. Proc. 2nd IEEE International Conference on Fuzzy Systems 1993, IEEE, San Francisco (1993), 1298–1301.

[7] F. Klawonn, On a Łukasiewicz Logic Based Controller. In: MEPP'92 International Seminar on Fuzzy Control through Neural Interpretations of Fuzzy Sets. Åbo Akademi, Reports on Computer Science & Mathematics, Ser. B. No 14, Turku (1992), 53–56.

[8] F. Klawonn, Mamdani's Model in the View of Equality Relations. Proc. First European Congress on Fuzzy and Intelligent Technologies (EUFIT'93), Aachen (1993), 364–369.

[9] F. Klawonn, J. Gebhardt, R. Kruse, Fuzzy Control on the Basis of Equality Relations – with an Example from Idle Speed Control. (submitted for publication).

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

[11] F. Klawonn, R. Kruse, Fuzzy Control as Interpolation on the Basis of Equality Relations. Proc. 2nd IEEE International Conference on Fuzzy Systems 1993, IEEE, San Francisco (1993), 1125–1130.

[12] R. Kruse, J. Gebhardt, F. Klawonn, Fuzzy Systems. Wiley, New York (1994).

[13] C.C. Lee, Fuzzy Logic in Control Systems: Fuzzy Logic Controller, Part II. IEEE Transactions on Systems, Man, and Cybernetics 20 (1990), 419–43.

[14] E.H. Mamdani, Application of Fuzzy Logic to Approximate Reasoning Using Linguistic Systems. IEEE Trans. Comput. 26 (1977), 1182–1191.

[15] T. Sudkamp, Similarity, Interpolation, and Fuzzy Rule Construction. Fuzzy Sets and Systems 58 (1993), 73–86.

[16] E. Trillas, L. Valverde, An Inquiry into Indistinguishability Operators. In: H.J. Skala, S. Termini, E. Trillas (eds.), Aspects of Vagueness. Reidel, Dordrecht (1984), 231–256.

[17] L.A. Zadeh, Similarity Relations and Fuzzy Orderings. Information Sciences 3 (1971), 177–200.