

Fuzzy Controllers With Conditionally Firing Rules

Bernhard Moser and Mirko Navara

Abstract—Mamdani controller was successfully used in many applications. One of its interpretations is that it uses a fuzzy relation as an approximation of the desirable input–output correspondence. We analyze mathematical properties of Mamdani controller and notice that it has lower computational complexity when compared to the residuum-based controller. However, we show that in standard situations, both these fuzzy controllers do not represent the rule base properly in the sense of finding a solution to the related system of fuzzy relational equations. First, we consider the premises and consequents as typical inputs and outputs and we want their correspondence to be kept. Second, we require that each normal input produces an output that bears nontrivial information. These two conditions appear to be almost contradictory for the previous controllers. We suggest a generalization of Mamdani controller which allows us to satisfy these requirements. Theory and experiments suggest that it performs better without any change of rule base and without a substantial increase of complexity.

Index Terms—Fuzzy control, fuzzy interpolation, fuzzy relational equation.

I. MOTIVATION

THE concept of approximate reasoning as it was conceived by Zadeh [12], [13] provides a framework which allows us to model and process vague linguistic information. The idea is to model linguistic terms by fuzzy sets, their (logical) relationship by fuzzy relations and their composition by the so-called *compositional rule of inference* [12]. As an important field of applications we refer to control processes for which linguistic information of a human expert about the required input–output behavior of the controller is available (for an introduction see, e.g., [4]).

Let X and Y denote the input and the output space, respectively. The spaces X and Y are supposed to be convex subsets of finite-dimensional real vector spaces. Then usually the expert's knowledge can be expressed by means of a rule base $\Theta = (X_i, Y_i)_{i=1}^n$ of if–then rules having the form

$$\text{if } x \in X_i \text{ then } y \in Y_i$$

where $i \in \{1, \dots, n\}$ and X_i, Y_i are fuzzy subsets of X, Y , respectively. For a universe of discourse Z , let $\mathcal{F}(Z)$ denote the set of all fuzzy subsets of Z . According to the paradigm of

approximate reasoning, the knowledge from the rule base Θ can be represented by a fuzzy relation $R \in \mathcal{F}(X \times Y)$. Applying the compositional rule of inference to a fuzzy input $X^* \in \mathcal{F}(X)$ and the relation R , a fuzzy output $Y^* \in \mathcal{F}(Y)$ is derived via

$$Y^* = X^* \circ_T R \quad (1)$$

i.e.,

$$\forall y \in Y: Y^*(y) = \sup_{x \in X} T(X^*(x), R(x, y))$$

where T is a t-norm modeling a fuzzy conjunction [3].

It is interesting to point out that the first successful practical applications of fuzzy sets were realized by means of the so-called Mamdani (or Mamdani–Assilian) inference [9] which results from (1) for

$$R(x, y) = \max_{i \leq n} T(X_i(x), Y_i(y)). \quad (2)$$

We denote the induced mapping by $Mam_{\Theta}: \mathcal{F}(X) \rightarrow \mathcal{F}(Y)$. Its explicit expression is

$$Mam_{\Theta}(X^*)(y) = \sup_{x \in X} T\left(X^*(x), \max_{i \leq n} T(X_i(x), Y_i(y))\right). \quad (3)$$

Note that Mamdani's approach is not fully coherent with the paradigm of approximate reasoning. For example, by employing a fuzzy conjunction instead of a fuzzy implication, Mamdani's approach does not represent the logical meaning of a linguistic if–then rule.

A fuzzy set is called *convex* if all its α -cuts are convex sets. For fuzzy subsets A_1, \dots, A_n of the same universe, Z , their *convex hull* is the smallest (w.r.t. the pointwise ordering) convex fuzzy set C satisfying $A_i(z) \leq C(z)$ for all $z \in Z, i = 1, \dots, n$. A fuzzy set is called *normal* if it attains the value 1 at some point. The *support* of a fuzzy set $A \in \mathcal{F}(Z)$ is $\text{Supp}A = \{z \in Z: A(z) > 0\}$. For $x \in Z$, let $\chi_x \in \mathcal{F}(Z)$ denote a Singleton, i.e.,

$$\chi_x(z) = \begin{cases} 1, & \text{if } z = x \\ 0, & \text{if } z \neq x \end{cases}$$

for all $z \in Z$.

In this paper, we are interested in Mamdani's formula and other inference methods as *fuzzy interpolation* $\text{Int}_{\Theta}: \mathcal{F}(X) \rightarrow \mathcal{F}(Y)$ with respect to the rule base $\Theta = (X_i, Y_i)_{i=1}^n$. The following axioms will be considered.

- Int1) If the input coincides with one of the premises, then the resulting output coincides with the corresponding consequent, i.e., $\forall i \in \{1, \dots, n\}: \text{Int}_{\Theta}(X_i) = Y_i$.
- Int2) For each normal input $X^* \in \mathcal{F}(X)$, the output $\text{Int}_{\Theta}(X^*)$ is not contained in all consequents, i.e., there is an index $i \in \{1, \dots, n\}$ with $\text{Int}_{\Theta}(X^*) \not\leq Y_i$.
- Int3) The output $\text{Int}_{\Theta}(X^*)$ belongs to the convex hull of $Y_i, i \in F$, where $F = \{i: \text{Supp}X_i \cap \text{Supp}X^* \neq \emptyset\}$.

Manuscript received August 29, 2000; revised July 26, 2001 and October 23, 2001. This work was supported by the Czech Ministry of Education under Research Programme MSM 212300013 "Decision Making and Control in Manufacturing," Projects Aktion Österreich - Tschechien 16p12 and 23p16, and Grant 201/97/0437 of the Grant Agency of the Czech Republic.

B. Moser is with CYBERhouse, GmbH, A-4020 Linz, Austria (e-mail: moser@cyberhouse.at).

M. Navara is with the Center for Machine Perception, Department of Cybernetics, Faculty of Electrical Engineering, Czech Technical University, CZ-166, Czech Republic (e-mail: navara@cmp.felk.cvut.cz).

Publisher Item Identifier S 1063-6706(02)04834-8.

Axiom Int1) states that the “typical inputs” used in the rules produce the corresponding outputs. Axiom Int2) ensures the “significance” of each output set. It also implies that empty outputs cannot occur for normal premises. Axiom Int3) is a “weak monotonicity” of outputs provided that the rules are locally monotonic (i.e., in some interval the rules describe a monotonic relation). Another interpretation (see [12]) understands the pairs (X_i, Y_i) as fuzzy points and the relation R as a fuzzy function from X to Y . Axiom Int1) states that this function goes through the given points, i.e., it is an interpolation. Axiom Int2) expresses a “connectedness” of the fuzzy graph and the condition that for each normal input the fuzzy function has a significant value (giving nontrivial information about the output). Axiom Int3) says that the interpolation preserves monotonicity in a weak sense.

In Section II, we investigate under which conditions on the rule base Θ the Mamdani inference (3) meets the axioms Int1)–Int3). It turns out that these natural axioms of a fuzzy interpolation lead to rather restrictive conditions, especially for the standard choice of the t-norm T and the shapes of the membership functions of the fuzzy sets X_i and Y_i . A similar analysis is done in Section III for relations R in (1) made up by the residuum induced by the t-norm T . In Section IV, we use Mamdani’s formula (3) and generalize it by changing the scales of membership degrees of premises and consequents. This allows us to satisfy Int1) under quite general conditions.

In contrast to the approach given by (1), in Section V we propose a construction for a fuzzy interpolation that does not rely on the compositional rule of inference any more. We modify the Mamdani’s formula in a way that the rules are firing conditionally depending on the degree of overlapping with some of the premises. As a result we obtain a formula which maps a fuzzy input to a fuzzy output and performs a desirable fuzzy interpolation under very general conditions.

Assumption 1: Throughout the paper we assume that X and Y are nonempty compact convex subsets of finite-dimensional real vector spaces. The rule base is $\Theta = (X_i, Y_i)_{i=1}^n$, where $n \geq 2$, $X_i \in \mathcal{F}(X)$ are normal and $Y_i \in \mathcal{F}(Y)$, $i = 1, \dots, n$. We denote by T a fixed continuous t-norm.

To make the examples simpler, we demonstrate the ideas on rule bases with a single input and single output (i.e., X and Y are sets of reals). Nevertheless, all results in this paper are formulated and proved for the multidimensional case.

II. MAMDANI’S APPROACH

In this section, we want to analyze under which conditions on the rule base Θ and the t-norm T the mapping Mam_{Θ} meets the conditions Int1)–Int3) for a fuzzy interpolation.

Axiom Int1) leads to the question under which conditions (2) satisfies the system of relational equations

$$Y_i = X_i \circ_T R, \quad i \in \{1, \dots, n\}. \quad (4)$$

A complete answer is given by de Baets in [1] (see Theorem 1). It provides a necessary and sufficient condition in terms of an inequality between the “degree of overlapping” of the premises and the “degree of indistinguishability” of the consequents. To

formulate it, we use the *residuum* induced by T , i.e., the operation $I_T: [0, 1] \times [0, 1] \rightarrow [0, 1]$ such that

$$I_T(a, b) = \sup\{c \in [0, 1]: T(a, c) \leq b\}.$$

(Various notations for this operation can be found in literature, e.g., $a \Rightarrow b$ in [5], $a \rightarrow_T b$ in [8], or $\hat{T}(a|b)$ in [2]. We use the notation of [1] from which the following notions are taken). Notice that, for T continuous, we may take \max instead of \sup in the latter formula. Recall that the residuum I_T is characterized by the so-called *adjointness* (see, e.g., [5])

$$\forall a, b, c \in [0, 1]: (T(a, b) \leq c \Leftrightarrow a \leq I_T(b, c)). \quad (5)$$

The corresponding *bimplication* (also called *symmetrized biresiduum*, see [2]) is

$$E_T(a, b) = \min\{I_T(a, b), I_T(b, a)\}.$$

Let A, B be fuzzy subsets of a universe Z . We define their *degree of overlapping* $\mathcal{D}_T(A, B)$ by

$$\mathcal{D}_T(A, B) = \sup_{z \in Z} T(A(z), B(z)) \quad (6)$$

and the *degree of indistinguishability* $\mathcal{E}_T(A, B)$ by

$$\mathcal{E}_T(A, B) = \inf_{z \in Z} E_T(A(z), B(z)).$$

Recall that the degree of indistinguishability $\mathcal{E}_T: \mathcal{F}(Z) \times \mathcal{F}(Z) \rightarrow [0, 1]$ is a fuzzy equality relation on $\mathcal{F}(Z)$ with respect to T [7]. Notice that all these notions are dependent on the choice of the t-norm T .

Due to the associativity and the continuity of T , the Mamdani inference (3) can be expressed by means of the degree of overlapping of the input X^* and the premises

$$\begin{aligned} Mam_{\Theta}(X^*)(y) &= \max_{i \leq n} T \left(\sup_{x \in X} T(X^*(x), X_i(x)), Y_i(y) \right) \\ &= \max_{i \leq n} T(\mathcal{D}_T(X^*, X_i), Y_i(y)). \end{aligned} \quad (7)$$

Let us point out that (7) provides an algorithm that avoids complex computations in the compound space $X \times Y$. The degree of overlapping can be computed in one cycle, then another cycle gives the output. In this sense Mamdani controller is well designed, while many other formulas require nested cycles for input and output space. Now, we are ready to formulate the theorem.

Theorem 1 [1]: Let $\Theta = (X_i, Y_i)_{i=1}^n$ be a rule base with $X_i(x_i) = 1$ for pairwise different elements $x_1, \dots, x_n \in X$. Then, (2) satisfies the system (4) iff all $i, j \in \{1, \dots, n\}$ satisfy

$$\mathcal{D}_T(X_i, X_j) \leq \mathcal{E}_T(Y_i, Y_j). \quad (8)$$

The criterion (8) can be interpreted as a many-valued model of the assertion

“If the premises X_i and X_j are not disjoint, then the corresponding images Y_i and Y_j are equal.”

In particular, (8) ensures that there is at most one image corresponding to a premise. For Singletons, this property characterizes a relation to be a function.

Inequality (8) yields a rather restrictive condition on the rule base Θ for t-norms without (nontrivial) zero divisors. This case includes the t-norms most commonly used in the field of fuzzy

control applications, the minimum $T_M(a, b) = \min(a, b)$ and the algebraic product $T_P(a, b) = a \cdot b$. Indeed, suppose that $\mathcal{E}_T(Y_i, Y_j) = 0$ for some i, j . [This is not an unusual situation for the shapes of membership functions commonly used in fuzzy control. Note that for t-norms without zero divisors we have $\mathcal{E}_T(Y_i, Y_j) > 0$ iff $\text{Supp}Y_i = \text{Supp}Y_j$, and this usually is not the case]. Then, (8) implies $\mathcal{D}_T(X_i, X_j) = 0$, which means that the fuzzy premises X_1, \dots, X_n must not overlap at all (their supports have to be disjoint). When this condition is satisfied, axiom Int2) is usually violated. For example, if the premises are continuous, then their supports are open in X . As they are nonempty and mutually disjoint, there is an $x^* \in X$ not belonging to any support of a premise. The normal input χ_{x^*} gives the empty output, violating Int2).

As a consequence of this we get the following proposition.

Proposition 1: Let T be a t-norm without zero divisors. Let $X_1, \dots, X_n \in \mathcal{F}(X)$, be continuous normal premises and let $Y_1, \dots, Y_n \in \mathcal{F}(Y)$ be consequents with mutually different supports. Then the Mamdani inference (3) does not satisfy the axioms Int1) and Int2) simultaneously.

This means that, for the standard choice of fuzzy sets and a t-norm T without zero divisors, the Mamdani algorithm cannot be considered as a good fuzzy interpolation in the sense of our axioms.

Proposition 2: The mapping Mam_Θ violates axiom Int2) iff there is an element $x^* \in X$ such that

$$\forall i \in \{1, \dots, n\}: X_i(x^*) \leq \min_{j \leq n} \inf_{y \in Y} I_T(Y_i(y), Y_j(y)). \quad (9)$$

Proof: Axiom Int2) is violated iff there is some normal input $X^* \in \mathcal{F}(X)$ with $X^*(x^*) = 1$, $x^* \in X$, such that for all $j \in \{1, \dots, n\}$ we have $X^* \circ_T R \leq Y_j$, where R is the Mamdani relation defined by (2). Observe that this inequality is equivalent to any of the following conditions:

$$\begin{aligned} \forall y \in Y \forall x \in X \forall j: T(X^*(x), R(x, y)) \\ \leq Y_j(y) \\ \forall y \in Y \forall x \in X \forall j \forall i: T(X^*(x), T(X_i(x), Y_i(y))) \\ \leq Y_j(y) \\ \forall y \in Y \forall x \in X \forall j \forall i: T(X^*(x), X_i(x)) \\ \leq I_T(Y_i(y), Y_j(y)) \end{aligned}$$

where the latter reformulation was obtained by the adjointness (5). For $x = x^*$, this implies

$$\forall y \in Y \forall j \forall i: X_i(x^*) \leq I_T(Y_i(y), Y_j(y))$$

and (9) follows.

Except for the choice $x = x^*$, all the transformations were equivalent. To prove that condition (9) implies that axiom Int2) cannot be satisfied, consider the input $X^* = \chi_{x^*}$ and follow the arguments just outlined, proceeding in the reverse direction. ■

Note that axiom Int3) is trivially satisfied for the Mamdani inference. Indeed, for an input set X^* , the corresponding output set is obtained as the maximum of the fuzzy subsets of Y

$$Mam_\Theta(X^*)(y) = \max_{i \in F} T(\mathcal{D}_T(X^*, X_i), Y_i(y)) \leq \max_{i \in F} Y_i(y).$$

The maximum $\max_{i \in F} Y_i(y)$ is contained in the convex hull of Y_i , $i \in F$.

Concluding this section, we may state that under reasonable conditions the Mamdani inference meets the requirements of a fuzzy interpolation only if the t-norm T has zero divisors. However, also for t-norms having zero divisors (e.g., the Łukasiewicz t-norm $T_L(a, b) = \max\{a + b - 1, 0\}$) inequality (8) for fuzzy sets Y_1, \dots, Y_n with pairwise vanishing degrees of indistinguishability leads to the restriction that $X_i(x) = 1$ implies $X_j(x) = 0$ whenever $i \neq j$.

III. FUZZY CONTROLLER BASED ON THE RESIDUUM

In this section, we consider (1) with the fuzzy relation $R \in \mathcal{F}(X \times Y)$ of the form

$$R(x, y) = \min_{i \leq n} I_T(X_i(x), Y_i(y)). \quad (10)$$

We denote the induced mapping by $\text{Res}_\Theta: \mathcal{F}(X) \rightarrow \mathcal{F}(Y)$. Axiom Int1) again leads to the question under which conditions on the rule base Θ (10) satisfies the system of relational equations (4). A characterization similar to the inequality (8) is not known in literature. Nevertheless, it can be shown that if the system (4) has a solution at all then (10) is a solution (for details, see, e.g., [3]). Indeed, if there is a solution \tilde{R} satisfying (4), then we have, for all $y \in Y$ and all indexes $i \in \{1, \dots, n\}$, the equation

$$Y_i(y) = \sup_{x \in X} T(X_i(x), \tilde{R}(x, y)).$$

By applying the adjointness (5), we obtain

$$\tilde{R}(x, y) \leq I_T(X_i(x), Y_i(y)) = R(x, y)$$

for all $x \in X$, $y \in Y$, $i \in \{1, \dots, n\}$. As this holds for all i , we get

$$\tilde{R}(x, y) \leq \min_{i \leq n} I_T(X_i(x), Y_i(y)) = R(x, y).$$

This inequality implies

$$\begin{aligned} Y_i(y) &= X_i \circ_T \tilde{R} \leq X_i \circ_T R \\ &= \sup_{x \in X} T \left(X_i(x), \min_{j \leq n} I_T(X_j(x), Y_j(y)) \right) \\ &\leq \sup_{x \in X} T(X_i(x), I_T(X_i(x), Y_i(y))) \leq Y_i(y) \end{aligned}$$

stating that (10) is the biggest (w.r.t. the pointwise ordering) solution of (4) if there is a solution at all. Provided that a compositional rule of inference of the form (1) is used, the restrictions for the rule base Θ to satisfy axiom Int1) are as weak as possible. In particular, inequality (8) is a sufficient criterion for (10) to satisfy axiom Int1).

Contrary to the case of Mam_Θ , no nice necessary and sufficient property for Res_Θ to satisfy axiom Int1) is known. (A detailed analysis of solvability of such systems of fuzzy relational equations is given in [10]. Also a degree of solvability is defined and studied there. Nevertheless, we think that [10] does not possess a method of testing solvability with an effort much smaller than necessary for finding the solution.) The reason of this nonsymmetry is the supremum in the compositional rule (1) as the following arguments show. We have always

$\text{Mam}_\Theta(X_i) \geq Y_i$. Whenever condition (8) is violated at some point $(x, y) \in X \times Y$ and for some j , i.e.,

$$T(X_i(x), X_j(x)) > E_T(Y_i(y), Y_j(y))$$

then the i th relation from (4) is violated, too. For Res_Θ , we have always the reverse inequality $\text{Res}_\Theta(X_i) \leq Y_i$. In order to have the equality, it is sufficient that

$$\forall i \forall y \in Y \exists x \in X: \min_{j \leq n} T(X_i(x) I_T(X_j(x), Y_j(y))) = Y_i(y).$$

This minimum can be achieved for $j = i$, giving the following sufficient condition:

$$\forall i \forall y \in Y \exists x \in X: (X_i(x) \geq Y_i(y) \text{ and } \forall j: I_T(X_j(x), Y_j(y)) \geq I_T(X_i(x), Y_i(y))). \quad (11)$$

In this sense, the necessary and sufficient conditions for Res_Θ to satisfy Int1) may be independent of local changes of X_i , $i = 1, \dots, n$. In particular, if we take for x an x_i satisfying $X_i(x_i) = 1$, then the condition (11) reduces to a stronger sufficient condition [obtained by a double use of the adjointness (5)]

$$\forall i \forall j \forall y \in Y: X_j(x_i) \leq I_T(Y_i(y), Y_j(y)). \quad (12)$$

As $X_j(x_i) \leq \mathcal{D}_T(X_i, X_j)$ and $I_T(Y_i(y), Y_j(y)) \geq \mathcal{E}_T(Y_i, Y_j)$, the condition (12) is apparently weaker than (8) [but still not necessary for Res_Θ to satisfy Int1)]. In comparison to (11), (12) is stronger, but easier to verify.

Let us give examples of rule bases for which (10) does not satisfy the axioms Int1)–Int3) simultaneously. The examples are drastically reduced in order to demonstrate the mathematical problems in the simplest cases. However, we use shapes of membership functions which are typical for practical applications.

Example 1: Let us consider a rule base consisting of three rules with $X = [0, 2]$, $X_1(x) = \max(1 - |x|, 0)$, $X_2(x) = \max(1 - |x - 1|, 0)$, $X_3(x) = \max(1 - |x - 2|, 0)$ and $Y = [1, 3]$, $Y_i = \chi_i$ for $i \in \{1, 2, 3\}$. Assume the input $X^* = \chi_{0.5}$. The output of a residuum-based fuzzy controller at $y \in Y$ is

$$\begin{aligned} (X^* \circ_T R)(y) &= \min_{i \leq 3} I_T(X_i(0.5), Y_i(y)) \\ &= \min\{I_T(0.5, \chi_1(y)), I_T(0.5, \chi_2(y)), \underbrace{I_T(0, \chi_3(y))}_1\} \\ &= I_T(0.5, 0). \end{aligned}$$

Depending on the choice of the t-norm T , we have two possibilities: In the case of $I_T(0.5, 0) = 0$, axiom Int2) is violated. In the case of $I_T(0.5, 0) > 0$, the output $X^* \circ_T R$ is not contained in the convex hull of Y_1 and Y_2 , hence axiom Int3) is not satisfied.

In the following example, we show a situation in which the axiom Int1) cannot be satisfied by means of any compositional rule of inference of (1).

Example 2: Let us consider a rule base Θ of two rules with $X = [0, 2]$ and $Y = [0, 1]$ and the membership functions $X_1(x) = 1 - 0.45x$, $X_2(x) = \min\{x, 1\}$, $Y_1(y) = 1 - y$ and $Y_2(y) = y$ (see Fig. 1). Next, we show that axiom Int1)

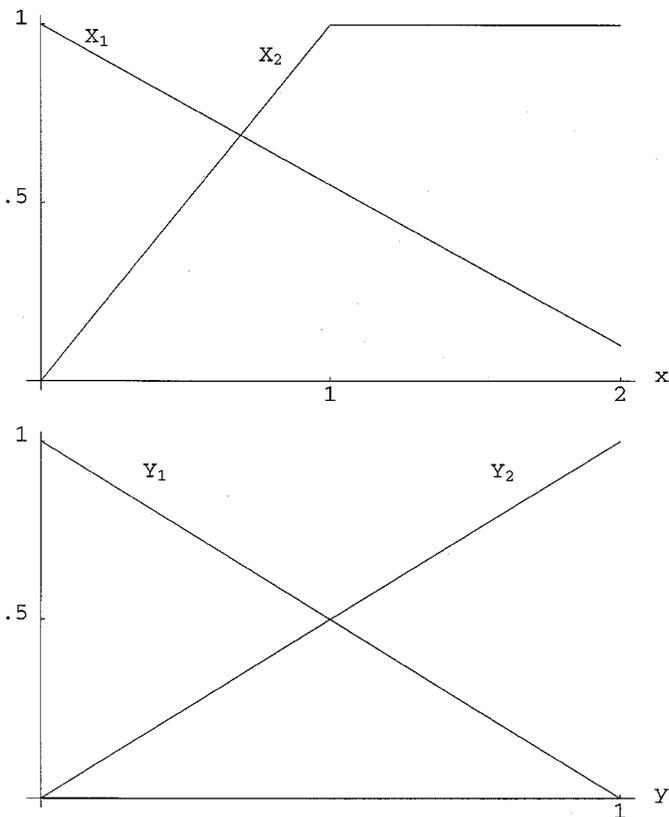


Fig. 1. Rule base of Example 2.

cannot be satisfied by a mapping of the form (1) for any t-norm T and any relation R . It suffices to prove that axiom Int1) is not satisfied for any residuum-based relation of the form (10). Consider an arbitrary continuous t-norm T and

$$R(x, y) = \min\{I_T(X_1(x), Y_1(y)), I_T(X_2(x), Y_2(y))\}.$$

Observe that

$$\begin{aligned} (X_2 \circ_T R)(1) &= \sup_{x \in X} \min\{T(X_2(x), I_T(X_1(x), Y_1(1))) \\ &\quad T(X_2(x), I_T(X_2(x), Y_2(1)))\} \\ &= \sup_{x \in X} \min\{T(X_2(x), I_T(X_1(x), 0)) \\ &\quad T(X_2(x), I_T(X_2(x), 1))\} \\ &= \sup_{x \in X} \min\{T(X_2(x), I_T(X_1(x), 0)), X_2(x)\} \\ &= \sup_{x \in X} T(X_2(x), I_T(X_1(x), 0)) \\ &\leq \sup_{x \in X} I_T(X_1(x), 0) \\ &\leq I_T\left(\inf_{x \in X} X_1(x), 0\right) \\ &= I_T(0.1, 0). \end{aligned}$$

Note that $I_T(a, b) = 1$ iff $a \leq b$, thus $I_T(0.1, 0) < 1$. Therefore, we obtain $(X_2 \circ_T R)(1) < Y_2(1)$ and $X_2 \circ_T R \neq Y_2$.

It is interesting that, in contrast to the Mamdani inference (3), here t-norms without zero divisors are the better choice in the sense that the axioms Int1)–Int3) become less restrictive.

Indeed, let us define $\gamma_T = \inf\{a \in [0, 1]: I_T(a, 0) = 0\}$. Then $\gamma_T > 0$ iff the t-norm T has zero divisors. Observe that if

$$\inf_{x \in X} \max_{i \leq n} X_i(x) < \gamma_T$$

then there is a constant $\gamma > 0$ and a normal input $X^* \in \mathcal{F}(X)$ such that $\text{Res}_\Theta(X^*)(y) \geq \gamma$ for all $y \in Y$. Indeed, take $x^* \in X$ such that $\max_{i \leq n} X_i(x^*) < \gamma_T$ and a normal input $X^* \in \mathcal{F}(X)$ with $X^*(x^*) = 1$. We obtain, for all $y \in Y$, the inequality

$$\begin{aligned} \text{Res}_\Theta(X^*)(y) &= \sup_{x \in X} T\left(X^*(x), \min_{i \leq n} I_T(X_i(x), Y_i(y))\right) \\ &\geq T\left(X^*(x^*), \min_{i \leq n} I_T(X_i(x^*), Y_i(y))\right) \\ &\geq \min_{i \leq n} I_T(X_i(x^*), 0) \\ &= I_T\left(\underbrace{\max_{i \leq n} X_i(x^*)}_{\gamma}, 0\right) > I_T(\gamma_T, 0) \end{aligned}$$

where γ (given by the above expression) is strictly positive. This means that the axiom Int3) will not be satisfied in the case of t-norms having zero divisors. For example, let X_1, \dots, X_n be normal premises and $x^* \in X$ such that $\max_{i \leq n} X_i(x^*) < \gamma_T$ and the convex hull of X_1, X_2 attains 1 at x^* . Then the input $X^* = \chi_{x^*}$ is contained in the convex hull of X_1, X_2 and it induces an output $\text{Res}_\Theta(X^*)(y)$ satisfying $\text{Res}_\Theta(X^*)(y) \geq \gamma$ for all $y \in Y$. We obtain a necessary condition that the convex hull of Y_1, Y_2 is at least γ everywhere; this is a too restrictive requirement on Y_1, Y_2 .

IV. GENERALIZED MAMDANI CONTROLLER

In this and the following section, we are going to propose a construction modifying Mamdani's formula (3) in order to get a fuzzy interpolation under as weak as possible conditions on the rule base Θ . In view of Example 2, the formula obtained is never more a composition rule, but it is based on similar principles.

First, we change the membership degrees of the premises by an increasing bijection (automorphism) $\rho: [0, 1] \rightarrow [0, 1]$. This allows us to change the degree of overlapping $\mathcal{D}_T(X_i, X_j)$ to a (possibly smaller) value $\mathcal{D}_T(X_i \circ \rho, X_j \circ \rho)$.

Remark 1: Another motivation of this step uses the t-norm T_ρ defined by

$$T_\rho(a, b) = \rho^{-1}(T(\rho(a), \rho(b))).$$

The new degree of overlapping is

$$\mathcal{D}_T(X_i \circ \rho, X_j \circ \rho) = \rho(\mathcal{D}_{T_\rho}(X_i, X_j)).$$

Further, we change the membership degrees of the consequents by an increasing bijection (isomorphism) $\sigma: [0, 1] \rightarrow [c, 1]$, where $0 \leq c < 1$. This allows us to increase the degree of indistinguishability of the consequents in order to satisfy (8). To extend the inverse of σ to the whole interval $[0, 1]$, we define the *pseudoinverse* $\sigma^{[-1]}$ as

$$\sigma^{[-1]}(t) = \begin{cases} \sigma^{-1}(t), & \text{if } t \geq c \\ 0, & \text{otherwise.} \end{cases}$$

We obtain a generalized inference rule $\text{Gen}_\Theta: \mathcal{F}(X) \rightarrow \mathcal{F}(Y)$ in the following form:

$$\begin{aligned} \text{Gen}_\Theta(X^*)(y) &= \sigma^{[-1]}\left(\sup_{x \in X} T\left(\rho(X^*(x)), \max_{i \leq n} (\rho(X_i(x)), \sigma(Y_i(y)))\right)\right) \\ &= \sigma^{[-1]}\left(\max_{i \leq n} T(\mathcal{D}_T(X^* \circ \rho, X_i \circ \rho), \sigma(Y_i(y)))\right). \end{aligned} \quad (13)$$

Theorem 2: Let $\Theta = (X_i, Y_i)_{i=1}^n$ be a rule base with $X_i(x_i) = 1$ for pairwise different elements $x_1, \dots, x_n \in X$. Let $\text{Gen}_\Theta: \mathcal{F}(X) \rightarrow \mathcal{F}(Y)$ be the mapping defined by (13). Then Gen_Θ satisfies the system of relational equations (4) iff all $i, j \in \{1, \dots, n\}$ satisfy

$$\mathcal{D}_T(X_i \circ \rho, X_j \circ \rho) \leq \mathcal{E}_T(Y_i \circ \sigma, Y_j \circ \sigma). \quad (14)$$

Proof: The proof of Theorem 2 is a routine generalization of Theorem 1. Indeed, due to the transformations ρ, σ , we work with the modified rule base $(X_i \circ \rho, Y_i \circ \sigma)_{i=1}^n$. We apply the classical Mamdani inference to the modified input $X^* \circ \rho$, and we obtain the modified output $Y^* \circ \sigma \in [c, 1]^Y$. Then, the pseudoinverse $\sigma^{[-1]}$ is used to get the outputs in the original scale $[0, 1]$. ■

If the degree of overlapping $\mathcal{D}_T(X_i, X_j)$ is less than one, but not sufficiently small, we may reduce it to an arbitrary small positive value by the choice of ρ , e.g., $\rho(t) = t^r$ for a sufficiently large r . This may allow us to satisfy axiom Int1). Nevertheless, we get a smaller and "less significant" output and we may have problems with axiom Int2). Analogously, the choice of the mapping σ , especially of the value $\sigma(0)$, helps us to satisfy Int1). Although the output set increases, it is again "less significant" if compared to $\sigma(Y_i), i = 1, \dots, n$. If we want to satisfy Int1) and Int2) simultaneously, the techniques of this section do not help much. In the next section, we shall improve them in this direction.

V. CONDITIONALLY FIRING RULES

Although we generalized the Mamdani inference substantially in the preceding section, we can go on and obtain a formula that satisfies Int1)–Int3) under very weak assumptions. In (13), we replace the degree of overlapping with the *degrees of conditional firing* of the i th rule defined as

$$\mathcal{C}_{T,i}(X^*) = \frac{\mathcal{D}_T(X^* \circ \rho, X_i \circ \rho)}{\max_{j \leq n} \mathcal{D}_T(X^* \circ \rho, X_j \circ \rho)}. \quad (15)$$

(In fact, this value is dependent not only on T and i , but also on ρ and on all premises $X_j, j = 1, \dots, n$.) For each normal input X^* , the values $\mathcal{C}_{T,i}(X^*), i = 1, \dots, n$, belong to $[0, 1]$ and at least one of them is 1. According to definition (15), the value $\mathcal{C}_{T,i}(X^*)$ can be interpreted as the quasi-truth value of the implication

"If the input X^* overlaps with the i th premise, then activate the i th rule."

This means that the rules are activated conditionally in the sense that the activation takes place provided there is a premise which overlaps with the input.

The explicit formula of the inference rule $\text{Con}_\Theta: \mathcal{F}(X) \rightarrow \mathcal{F}(Y)$ based on conditional firing is

$$\text{Con}_\Theta(X^*)(y) = \sigma^{[-1]} \left(\max_{i \leq n} T(\mathcal{C}_{T,i}(X^*), \sigma(Y_i(y))) \right). \quad (16)$$

The fuzzy inference rule based on the degree of conditional firing allows us to satisfy the axioms Int1)–Int3) under conditions that are relatively weak and easy to verify.

Theorem 3: Let $\Theta = (X_i, Y_i)_{i=1}^n$ be a rule base with normal premises and let $\rho: [0, 1] \rightarrow [0, 1]$ be any automorphism satisfying the conditions

- Con1) “covering of premises”: $\inf_{x \in X} \max_{i \leq n} X_i(x) > 0$;
- Con2) “disjointness of premises”: there is a constant $c < 1$ with $\mathcal{D}_T(X_i \circ \rho, X_j \circ \rho) \leq c$ whenever $i \neq j$;
- Con3) “significance of consequents”: for each $i \in \{1, \dots, n\}$, there is an element $y_i \in Y$ satisfying $Y_i(y_i) > \min_{j \neq i} Y_j(y_i)$.

Then for any isomorphism $\sigma: [0, 1] \rightarrow [c, 1]$ the mapping Con_Θ from (16) satisfies the axioms Int1)–Int3).

Proof: Int1): As X_i is normal, $\mathcal{D}_T(X_i \circ \rho, X_i \circ \rho) = 1$ and $\mathcal{C}_{T,j}(X_i) = \mathcal{D}_T(X_i \circ \rho, X_j \circ \rho)$. The output corresponding to the input set X_i is

$$\text{Con}_\Theta(X_i)(y) = \sigma^{[-1]} \left(\max_{j \leq n} T(\mathcal{C}_{T,j}(X_i), \sigma(Y_j(y))) \right). \quad (17)$$

For $j \neq i$, Con2) implies

$$T(\mathcal{C}_{T,j}(X_i), \sigma(Y_j(y))) \leq T(c, \sigma(Y_j(y))) \leq c \leq \sigma(Y_i(y)) \\ = T(\mathcal{C}_{T,i}(X_i), \sigma(Y_i(y))).$$

Hence, the maximum in (17) is achieved for $j = i$ and the right-hand side is $Y_i(y)$.

Int2): It is sufficient to consider the (worst) case when the input set X^* is the characteristic function of a Singleton $\{x^*\} \subseteq X$. According to Con1), always at least one rule fires totally, i.e., $\mathcal{C}_{T,i}(X^*) = 1$ for some i . The corresponding output set satisfies

$$\text{Con}_\Theta(X^*)(y) = \sigma^{[-1]} \left(\max_{j \leq n} T(\mathcal{C}_{T,j}(X^*), \sigma(Y_j(y))) \right) \\ \geq \sigma^{[-1]} (\sigma(Y_i(y))) = Y_i(y)$$

and Con3) implies Int2).

Int3): The same argument as in Section II works here, too. \blacksquare

The assumptions of Theorem 3 are very weak and natural. The condition Con1) says that the input space X is covered by the premises so that for each input value there is at least one firing rule. This is an obvious requirement which also ensures that the denominator in (15) is nonzero. To satisfy Con2), it is sufficient that $\mathcal{D}_T(X_i, X_j) < 1$ whenever $i \neq j$. Then we may choose ρ such that Con2) is fulfilled. The condition Con3) requires that each consequent is significant (i.e., not covered by the other consequents). Also, this condition is usually satisfied by the rule base of a fuzzy controller.

Under very weak assumptions Con1)–Con3), the inference rule based on conditional firing satisfies our axioms Int1)–Int3) and hence possesses a desirable fuzzy controller. We point out the following.

- The conditions of Theorem 3 are easy to check.
- The numerical effort for computing the output (16) is of the same order as for the Mamdani inference (3). To keep the numerical complexity as low as possible we may choose for example $\sigma(t) = (1 - c) \cdot t + c$ and $\rho(t) = t^r$, $r \in \mathbb{N}$.

Moreover, we get the following desirable behavior if one rule fires totally.

Proposition 3: Let us make the same assumptions as in Theorem 3. Let $X^* = \chi_{x^*}$ for $x^* \in X$ such that $X_i(x^*) = 1$ (for a fixed i). Then, $\text{Con}_\Theta(X^*) = Y_i$.

Proof: Let $j \neq i$. As

$$\mathcal{D}_T(X^* \circ \rho, X_j \circ \rho) \\ = \rho(X_j(x^*)) = T(\rho(X_i(x^*)), \rho(X_j(x^*))) \\ \leq \mathcal{D}_T(X_i \circ \rho, X_j \circ \rho) \leq c$$

and $\mathcal{D}_T(X^* \circ \rho, X_i \circ \rho) = 1$, the degrees of conditional firing are $\mathcal{C}_{T,i}(X^*) = 1$, $\mathcal{C}_{T,j}(X^*) \leq c$ for $j \neq i$. We obtain, for all $y \in Y$ and $j \neq i$

$$\sigma^{[-1]} (T(\mathcal{C}_{T,j}(X^*), \sigma(Y_j(y)))) \leq \sigma^{[-1]} (\mathcal{C}_{T,j}(X^*)) \\ \leq \sigma^{[-1]} (c) = 0 \\ \sigma^{[-1]} (T(\mathcal{C}_{T,i}(X^*), \sigma(Y_i(y)))) \leq \sigma^{[-1]} (\sigma(Y_i(y))) = Y_i(y)$$

and, as the maximum of these fuzzy sets, $\text{Con}_\Theta(X^*) = Y_i$.

Remark 2: The latter proposition means that if one rule fires totally, then the other rules are irrelevant. (Condition Con2) ensures that this cannot happen for more than one rule). This allows the output to attain all its extreme values described by the consequents. In a Mamdani controller, we have this property if and only if there are elements $x_i \in X$, $i = 1, \dots, n$, such that $X_i(x_i) = 1$, $X_j(x_i) = 0$ for all $j \neq i$.

Our method preserves the continuity of the input–output correspondence in the following sense.

Theorem 4: Let $\Theta = (X_i, Y_i)_{i=1}^n$ be a rule base satisfying Con1) and such that all X_i are continuous. Then the mapping $f: X \rightarrow \mathcal{F}(Y)$ defined by $f(x) = \text{Con}_\Theta(\chi_x)$ is continuous with respect to the pointwise (=weak) convergence in $\mathcal{F}(Y)$.

Proof: As X_i , $i = 1, \dots, n$, are continuous functions of x and also ρ is continuous, so are the degrees of overlapping $\mathcal{D}_T(\chi_x \circ \rho, X_i \circ \rho) = \rho(X_i(x))$. Also their maximum is continuous and, according to Con1), strictly positive, hence, the degrees of conditional firing are continuous functions of x . The continuity of T , σ and $\sigma^{[-1]}$ implies the pointwise convergence of the corresponding output sets. \blacksquare

VI. EXAMPLES

In this section, we demonstrate the properties of a controller based on conditional firing. We present several simple examples showing how our axioms can be satisfied. Then we show practical results confirming that this tool is not only theoretically, but also practically advantageous.

First of all, note that the conditions of Theorem 3 are satisfied by Example 2 (with the identity taken for ρ and with $c \geq 20/29$, Con2) is satisfied for any t-norm T). Thus, there is a fuzzy controller with conditionally firing rules which satisfies Int1)–Int3).

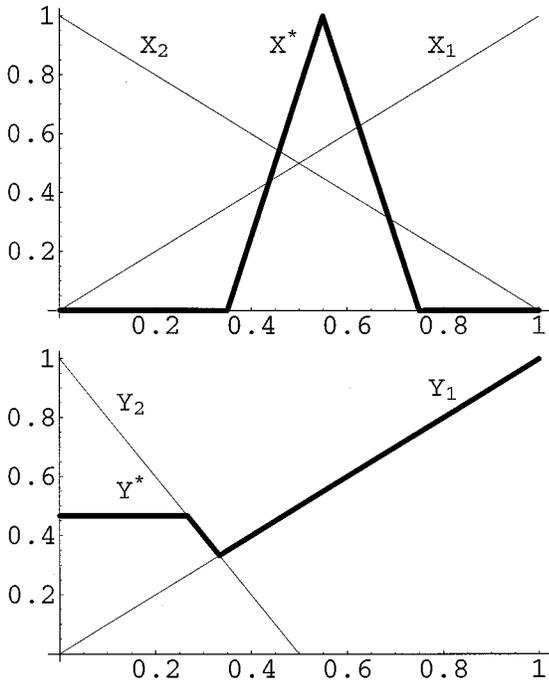


Fig. 2. Rule base of Example 3, a sample input set (X^*) and the corresponding output set of the controller with conditionally firing rules (Y^*).

This was not possible for any controller based on the compositional rule of inference, including the Mamdani controller.

Next, we present a typical fuzzy output in Example 3 and the interpolation behavior when defuzzifying the output in Example 4.

Example 3: Let us consider a rule base Θ of two rules with $X = [0, 1]$, $Y = [0, 1]$ and $X_1(x) = x$, $X_2(x) = 1 - x$, $Y_1(y) = y$ and $Y_2(y) = \max\{1 - 2y, 0\}$. We take T_M for T and the identity for ρ . The degree of overlapping is $\mathcal{D}_T(X_1, X_2) = 1/2$ and we may choose $\sigma(t) = (1 + t)/2$. Then Fig. 2 illustrates the output $Y^* = \text{Con}_\Theta(X^*)$ for $X^*(x) = \max(1 - 5 \cdot |x - 0.55|, 0)$. The degrees of overlapping are $\mathcal{D}_T(X^*, X_1) = 5/8 = 15/24$, $\mathcal{D}_T(X^*, X_2) = 11/24$, the maximum value is achieved for X_1 . Hence, the degrees of conditional firing are $\mathcal{C}_{T,1}(X^*) = 1$ and

$$\mathcal{C}_{T,2}(X^*) = \frac{\mathcal{D}_T(X^*, X_2)}{\mathcal{D}_T(X^*, X_1)} = \frac{11}{15}.$$

The output is

$$\begin{aligned} Y^*(y) &= \text{Con}_\Theta(X^*)(y) \\ &= \sigma^{[-1]}(\max\{\sigma(Y_1(y)), T_M(\frac{11}{15}, \sigma(Y_2(y)))\}) \\ &= \max\{Y_1(y), \min\{\frac{7}{15}, Y_2(y)\}\}. \end{aligned}$$

Example 4: Let us consider a rule base Θ of two rules with $X = [0, 1]$, $Y = [1, 2]$, and $X_1(x) = x$, $X_2(x) = 1 - x$, $Y_1 = \chi_2$, $Y_2 = \chi_1$. Then $\mathcal{D}_T(X_1, X_2) = T(1/2, 1/2) \leq 1/2$. We choose c such that $\mathcal{D}_T(X_1, X_2) \leq c < 1$. Let us set $\sigma(t) = (1 - c) \cdot t + c$ and $\rho(t) = t$ for all $t \in [0, 1]$. We derive the

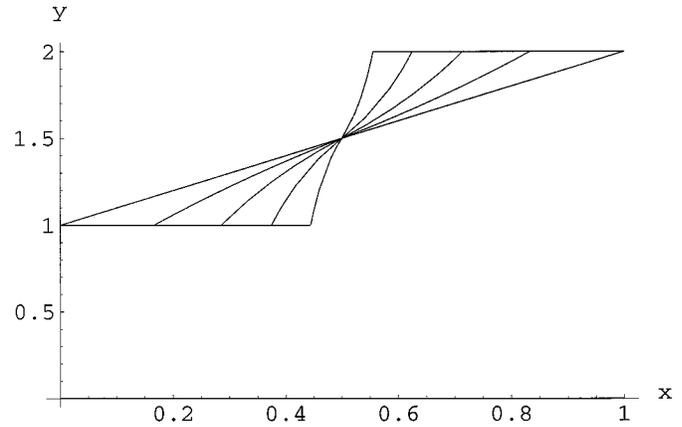


Fig. 3. Graphs of the input-output functions of Example 4.

output corresponding to a crisp input χ_x , $x \in X$. We obtain the following formula, independently of the choice of the t-norm T :

$$\text{Con}_\Theta(\chi_x) = \begin{cases} \chi_2 + \sigma^{[-1]} \left(\frac{1-x}{x} \right) \cdot \chi_1, & \text{if } x \geq 0.5 \\ \sigma^{[-1]} \left(\frac{x}{1-x} \right) \cdot \chi_2 + \chi_1, & \text{if } x < 0.5. \end{cases}$$

Defuzzifying the output $\text{Con}_\Theta(\chi_x)$ by computing the center of gravity [4], we obtain the input-output function $f: X \rightarrow Y$

$$f(x) = \begin{cases} 2, & \text{if } \frac{1}{1+c} \leq x \leq 1 \\ \frac{1+x-3cx}{1-2cx}, & \text{if } 0.5 \leq x < \frac{1}{1+c} \\ \frac{1+x-3c(1-x)}{1-2c(1-x)}, & \text{if } \frac{c}{1+c} \leq x < 0.5 \\ 1, & \text{if } 0 \leq x < \frac{c}{1+c}. \end{cases} \quad (18)$$

Observe that for the Łukasiewicz t-norm T_L the degree of overlapping vanishes, i.e., $\mathcal{D}_{T_L}(X_1, X_2) = 0$. Therefore, we can choose $c = 0$ and σ equal to the identity. Then, (18) yields $f(x) = 1 + x$ for all $x \in X$. We obtain the line segment interpolating the points $(0, 1)$ and $(1, 2)$. This result coincides with the input-output function induced by the Mamdani inference. For $0 < c < 1$ the resulting function f is 1 on the interval $[0, (c/1+c)]$ and 2 on the interval $[(1/1+c), 1]$. Fig. 3 illustrates the induced input-output functions for $T = T_L$ and $c \in \{0, 0.2, 0.4, 0.6, 0.8\}$. In this simplified example the input-output function for given $c \in [0, 1]$ from Theorem 3 is independent of the choice of the t-norm T . However, the possible values of c satisfying Con2) are limited by T . For $T = T_L$, we can choose $c \in [0, 1]$ arbitrarily. For $T = T_P$, resp. $T = T_M$, we have to take $c \geq 0.25$, resp. $c \geq 0.5$.

Until now, we presented theoretical arguments why the controller based on conditional firing should be recommended. Also practical experiments confirm this opinion. To demonstrate this, let us mention a recent simple experiment with balancing a ball on a plate. (This classical control task is also called ‘‘ball on beam,’’ see [4].) The aim was to stabilize the ball in zero position. We implemented a simple Mamdani controller solving this task. Then we used a simplified version of the controller based

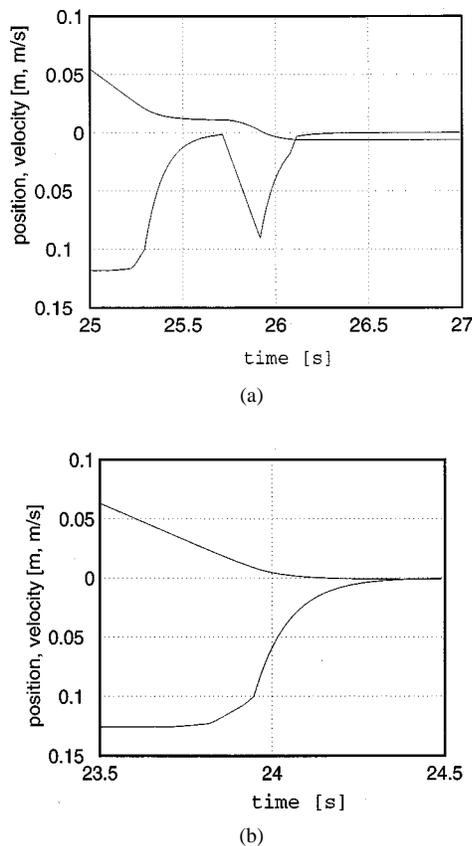


Fig. 4. (a) Results of Mamdani controller. (b) Controller with conditionally firing rules (bottom).

on conditional firing. A complete description of this experiment can be found in [11]. The position and velocity before reaching the final state can be seen in Fig. 4. One can see the following advantages of our controller:

- the speed varies monotonically and smoothly;
- the asymptotic value is smaller;
- the transient time is shorter.

What is important is that these improvements were obtained without an additional tuning of the rule base, just by switching to another type of controller. It only interprets better the knowledge already expressed by the rule base. This tool is ready for further testing on more complex tasks.

VII. CONCLUSION

We investigated the suitability of the compositional rule of inference for the purpose of fuzzy interpolation and fuzzy control. It turned out that the standard inference mechanism due to Mamdani seems not to be always a satisfactory fuzzy interpolation under practical conditions. Therefore, we modified the Mamdani inference by changing the scales of membership functions and by replacing the degree of overlapping with the truth-value of a corresponding (many-valued) conditional statement.

As a result, we obtained a suitable fuzzy interpolation under rather weak restrictions of the rule base. Finally, we illustrated its input–output behavior by means of numerical examples. As we study the fuzzy outputs in full generality, our results are applicable independently of the choice of the subsequent defuzzification.

ACKNOWLEDGMENT

The authors would like to thank L. Zadeh for fruitful discussions on the topic of the paper.

REFERENCES

- [1] B. de Baets, “A note on Mamdani controllers,” in *Intelligent Systems and Soft Computing for Nuclear Science and Industry*, D. Ruan, P. D’hondt, P. Govaerts, and E. Kerre, Eds, Singapore: World Scientific, 1996, pp. 22–28.
- [2] D. Boixader, J. Jacas, and J. Recasens, “Transitive closure and betweenness relations,” *Fuzzy Sets Syst.*, vol. 120, no. 3, pp. 415–422, 2001.
- [3] S. Gottwald, *Fuzzy Sets and Fuzzy Logic, Foundations of Applications—From a Mathematical Point of View*. Wiesbaden, Germany: Vieweg, 1994.
- [4] D. Driankov, H. Hellendoorn, and M. Reinfrank, *An Introduction to Fuzzy Control*. Berlin, Germany: Springer-Verlag, 1993.
- [5] P. Hájek, *Metamathematics of Fuzzy Logic*. Dordrecht, The Netherlands: Kluwer, 1998.
- [6] U. Höhle, “Fuzzy equalities and indistinguishability,” in *Proc. EUFIT’93*, Aachen, Germany, 1993, pp. 358–363.
- [7] F. Klawonn and R. Kruse, “Equality relations as a basis for fuzzy control,” *Fuzzy Sets Syst.*, vol. 54, pp. 147–156, 1993.
- [8] E. P. Klement, R. Mesiar, and E. Pap, *Triangular Norms*. Norwell, MA: Kluwer, 2000.
- [9] E. H. Mamdani and S. Assilian, “An experiment in linguistic synthesis with a fuzzy logic controller,” *J. Man Machine Stud.*, vol. 7, pp. 1–13, 1975.
- [10] I. Perfilieva and A. Tonis, “Criterion of solvability of fuzzy relational equations system,” in *Proc. IFSA’97*, M. Mareš, R. Mesiar, V. Novák, J. Ramík, and A. Stupňanová, Eds., 1997, pp. 90–95.
- [11] J. Št’astný. (2001) Comparison of Mamdani and CFR controller. Center for Machine Perception, Czech Tech. Univ., Res. Rep. CTU-CMP-2001-04, Prague, Czech Republic. [Online]. Available: ftp://cmp.felk.cvut.cz/pub/cmp/articles/navara/TR_Stastny.ps.gz
- [12] L. A. Zadeh, “Outline of a new approach to the analysis of complex systems and decision processes,” *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-3, pp. 28–44, 1973.
- [13] —, “A theory of approximate reasoning,” in *Machine Intelligence*, J. Hayes, D. Mitchie, and L. Mikulich, Eds. New York: Wiley, 1979, pp. 149–194.



Bernhard Moser received the M.Sc. degree in natural sciences from the University of Salzburg, Salzburg, Austria, and the Ph.D. degree in mathematics from the Johannes Kepler University of Linz, Linz, Austria, in 1991 and 1996, respectively.

From 1991 to 1998, he worked at the University of Salzburg in the Department of Computer Science, and the Johannes Kepler University of Linz in the Department of Mathematics as Research Assistant. Together with mathematical research in fuzzy logic and fuzzy control, his work included software engineering activities for industrial applications. He is the author of about 15 refereed papers. Recently, he started a new professional career within private industry, where his work includes databases, internet applications, and cross-media publication.



Mirko Navara received the M.Sc. degree in technical cybernetics and the Ph.D. degree in mathematical analysis from the Czech Technical University (CTU), Prague, Czech Republic, in 1983 and 1988, respectively, and in 2001, he received the highest National Research Degree of Doctor of Science.

From 1987 to 1996, he was a Staff Member with the Department of Mathematics, CTU, Prague, where he lectured on fuzzy logic, numerical analysis, and computer algebra systems. In 1996, he was promoted to Associate Professor. He has published over 90 refereed articles. Of his numerous research stays abroad, fellowships at Université Claude Bernard Lyon I, France, and Johannes Kepler University of Linz, should be mentioned as those which contributed a considerable amount to his mathematical orientation. In 1996, he moved to a newly open research unit Center for Machine Perception at CTU, Prague. His main research interest has been nonstandard logics (especially quantum logics and fuzzy logics) and their algebraic and measure-theoretic questions. His research orientation remained nonstandard logics with certain accent on the questions of image processing.

Dr. Navara was awarded the "Prize for Scientific Achievement" by the International Quantum Structures Association in 1996.