BETA FUZZY LOGIC SYSTEMS: APPROXIMATION PROPERTIES IN THE MIMO CASE

ADEL M. ALIMI[*], RADHIA HASSINE[**], MOHAMED SELMI[***]

[*] REGIM: REsearch Group on Intelligent Machines, Department of Electrical Engineering
University of Sfax, ENIS, BP W, Sfax 3038, Tunisia
e-mail: Adel.Alimi@ieee.org

[**] Department of Mathematics, Faculty of Sciences of Monastir
Boulevard de l'environnement, Monastir 5000, Tunisia
e-mail: Radhia.Hassine@fsm.rnu.tn

[***] Laboratory of Physics and Mathematics, Department of Mathematics
Faculty of Sciences of Sfax, Sfax 3038, Tunisia
e-mail: Mohamed.Selmi@fss.rnu.tn

Many researches have been interested in the approximation properties of Fuzzy Logic Systems (FLS), which, like neural networks, can be seen as approximation schemes. Almost all of them tackled the Mamdani fuzzy model, which was shown to have many interesting approximation features. However, only in few cases the Sugeno fuzzy model was considered. In this paper, we are interested in the zero-order Multi-Input–Multi-Output (MIMO) Sugeno fuzzy model with Beta membership functions. This leads to Beta Fuzzy Logic Systems (BFLS). We show that BFLSs are universal approximators. We also prove that they possess the best approximation property and the interpolation characteristic.

Keywords: Beta function, universal approximation property, best approximation property, interpolation property, Sugeno fuzzy model, MIMO systems

1. Introduction

Fuzzy logic systems (FLSs) were introduced in order to approximate a decision or a control function with a given accuracy (Bouchon-Meunier, 1995; Kosko, 1993; Mamdani and Assilian, 1975; Mendel, 1995; Sugeno and Kang, 1988; Terano *et al.*, 1992; Zadeh, 1965). In fact, when the system to be controlled is too complex, it is difficult and often impossible to model its behaviour using mathematical equations (Jang, 1993; Jang and Sun, 1995; Nguyen and Kreinovich, 1992; Nguyen *et al.*, 1996; Takagi and Sugeno, 1985; Yen *et al.*, 1995). In this case, it is easier to describe system behaviour via fuzzy linguistic fuzzy rules. With these fuzzy rules and fuzzy logic concepts, one can construct a function $f : U \subset \mathbb{R}^n \longrightarrow V \subset \mathbb{R}^p$ that models the system behaviour so it is natural to relate the construction of FLSs to the theory of function approximation (Kosko, 1992; Laukonen and Passino, 1994; Lewis *et al.*, 1995).

As we all know, FLSs comprise four main components, which are, the fuzzifier, the fuzzy rule base, the fuzzy inference engine and the defuzzifier. The main difference between the Mamdani and Sugeno fuzzy systems lies in the consequents of fuzzy rules, which are fuzzy sets for the former and crisp values for the latter. Defuzzification is defined as the step which produces a crisp output for our FLS from the fuzzy set that is the output of the inference block. As was mentioned in (Mendel, 1995), many defuzzifiers were proposed in the literature. However, there is no scientific base for any of them (i.e. no defuzzifier was derived from a first principle such as maximization of fuzzy information or entropy). Consequently, defuzzification is an art rather than a science. Because we are interested in engineering applications of FL, one criterion for the choice of a defuzzifier is computational simplicity. For a Sugeno fuzzy model, the consequence of each fuzzy rule is a constant, and defuzzification in such a model is made using the centre-of-gravity method, i.e. the gravity centre of all singletons is calculated.

Note that the main candidates for defuzzifiers are the following:

- the maximum defuzzifier,
- the mean-of-maxima defuzzifier,
- the centroid defuzzifier,
- the height defuzzifier, and
- the modified height defuzzifier.

However, with this big choice of defuzzifiers, we see that there are many options of Mamdani fuzzy logic systems to choose from. This demonstrates the richness of Mamdani FLSs.

In this paper, we consider the zero-order Sugeno model, which can be seen as a Mamdani model with a singleton consequence. Many researchers proved that Mamdani fuzzy systems are universal approximators (Castro and Delgado, 1996; Dickerson and Kosko, 1996; Gorrini *et al.*, 1995; Hartani *et al.*, 1996; Wang, 1992; Wang and Mendel, 1992; Wang *et al.*, 1997; Zeng and Singh, 1994; 1995), but few of them were interested in the Sugeno fuzzy model. Recently, Ying (1998) proved that the Sugeno fuzzy model with a linear rule consequence is a universal approximator. In this paper, we are interested in the Sugeno fuzzy model of the zeroth order. The advantage of such a model is that it is simpler than the one considered by Ying (1998): the consequence of each fuzzy rule is a constant and there is no need for a defuzzification step to construct such a system. Another important point which affects the behaviour of FLSs is the type of membership functions for input variables. Different types of membership functions were proposed (Alimi, 1997b), such as triangular functions (Pedrycz, 1994), normal peak functions (Wang *et al.*, 1997), pseudo trapezoid functions (Zeng and Singh, 1994; 1995), or functions using translations and dilations of one fixed function (Mao *et al.*, 1997), etc.

In this paper, we consider MIMO Beta Fuzzy Logic Systems (BFLS) (Alimi, 2000; Alimi *et al.*, 2000), which are FLSs in which Beta functions are used as membership functions of the input variables. BFLSs were actively studied in the few last years (Alimi, 1997a; 1997c; 1997d; 1998a; 1998c; 2000; 2002; Alimi *et al.*, 2000; Hassine *et al.*, 2000; Masmoudi *et al.*, 2000) and they showed robust and interesting properties compared with other FLSs (Alimi, 1998b). The results of this paper are extensions of our previous work on SISO FLSs to the MIMO case (Alimi, 2000; Alimi *et al.*, 2000).

The organization of this paper is as follows: in the second section, we introduce Beta fuzzy sets. In Section 3, we deal with the property of universal approximation and give the essential definitions and properties needed for the study of this property. Multi-Input-Multi-Output (MIMO) BFLSs are shown in Section 4 to have the following properties:

1. basic approximation,
2. uniform approximation,
3. uniform convergence, and
4. universal approximation.

The best approximation property that seems more practical is introduced in Section 5. We will prove that BFLSs satisfy this property. Finally, in Section 6 we show that the BFLSs possess the interpolation property.

2. Beta Fuzzy Logic Systems

2.1. Mathematical Model of an FLS

A Multi-Input-Single-Output (MISO) FLS can be seen as a function $f : U \subset \mathbb{R}^n \longrightarrow V \subset \mathbb{R}$, where U is the input space, V is the output space, and $n > 1$. As was shown by Lee (1990), a MIMO fuzzy system can always be separated into a group of MISO fuzzy ones, so it is sufficient to study MISO fuzzy systems and the results concerning MIMO ones can be easily deduced.

In this paper, we adopt the zero-order Sugeno fuzzy model with multiplication as a t-norm. Then a fuzzy system is given by

$$f : \begin{array}{l} U \subset \mathbb{R}^n \longrightarrow V \subset \mathbb{R}, \\ \quad \vec{x} \longmapsto \\ \displaystyle\sum_{(i_1,i_2,\dots,i_n)\in I} \frac{\prod_{j=1}^{n} \mu_{A^j_{i_j}}(x_j)}{\sum_{(k_1,k_2,\dots,k_n)\in I} \prod_{l=1}^{n} \mu_{A^l_{k_l}}(x_l)} y_{(i_1,i_2,\dots,i_n)}, \end{array} \tag{1}$$

where

- $\vec{x} = (x_1, x_2, \dots, x_n)$ is the input variable,
- I is the set $\{(i_1, i_2, \dots, i_n) \mid 1 \le i_j \le N_j; 1 \le j \le n\}$.
- $N = \prod_{j=1}^{n} N_j$ is the number of fuzzy rules of the form

  $$R_{(i_1,i_2,\dots,i_n)} : \text{ if } \big(\vec{x} \text{ is } \vec{A}_i\big) \text{ then } \big(y = y_{(i_1,i_2,\dots,i_n)}\big),$$

- $y_{(i_1,i_2,\dots,i_n)}$ are constants in V which represent the consequences of the fuzzy rules $R_{(i_1,i_2,\dots,i_n)}$, and
- $\vec{A}_i = (A^1_{i_1}, A^2_{i_2}, \dots, A^n_{i_n})$ are linguistic terms characterized by their membership functions $\mu_{A^j_{i_j}}(x_j)$.

From (1) we see that FLSs can be considered as linear combinations of the functions

$$B_{(i_1,i_2,\dots,i_n)}(\vec{x}) = \frac{\prod_{j=1}^{n} \mu_{A^j_{i_j}}(x_j)}{\sum_{(k_1,k_2,\dots,k_n)\in I} \prod_{l=1}^{n} \mu_{A^l_{k_l}}(x_l)}, \tag{2}$$

so we can introduce the following definition:

Definition 1. *Fuzzy Basis Functions* (FBFs) are defined by

$$B_{(i_1,i_2,\dots,i_n)}(\vec{x}) = \frac{\mu_{A_{(i_1,i_2,\dots,i_n)}}(\vec{x})}{\sum_{(k_1,k_2,\dots,k_n)\in I} \mu_{A_{(k_1,k_2,\dots,k_n)}}(\vec{x})}, \tag{3}$$

where $(i_1, i_2, \dots, i_n) \in I$, and

$$\mu_{A_{(i_1,i_2,\dots,i_n)}}(\vec{x}) = \prod_{j=1}^{n} \mu_{A^j_{i_j}}(x_j), \tag{4}$$

$\vec{x} = (x_1, x_2, \dots, x_n)$.

With the use of this notation, the output of an FLS is given by

$$f(\vec{x}) = \sum_{(i_1,i_2,\dots,i_n)\in I} B_{(i_1,i_2,\dots,i_n)}(\vec{x})\, y_{(i_1,i_2,\dots,i_n)}. \tag{5}$$

2.2. Beta Functions

Beta functions (Johnson, 1970) were proposed as membership functions of the input variables (Alimi, 1997e; 1998b; 2000; 2002; Alimi *et al.*, 2000). This subsection is devoted to the introduction of Beta functions and their main properties.

Definition 2. (*Beta functions in the one-dimensional case*) Consider $a, b \in \mathbb{R}$ satisfying $a < b$, and let $p, q > 0$. In the one-dimensional case, a Beta function is given by

$$\beta(x) = \begin{cases} \left(\dfrac{x-a}{c-a}\right)^p \left(\dfrac{b-x}{b-c}\right)^q & \text{if } x \in]a,b[, \\ 0 & \text{otherwise,} \end{cases} \tag{6}$$

where

$$c = \frac{pb + qa}{p+q}. \tag{7}$$

We can see that a Beta function depends on four parameters, which gives it a great flexibility, permitting to reproduce most common shapes of membership functions (see Fig. 1). In the remainder of this paper, we shall write

$$\beta(x) = \beta(x; p, q, a, b). \tag{8}$$

Any Beta function $\beta(x)$ is characterized by the following properties:

1. $\beta(x)$ is continuous on $\mathbb{R}$.
2. $\beta(a) = \beta(b) = 0$, $c \in]a,b[$ and $\beta(c) = 1$.
3. For all $x \in]a,b[$, we get

$$\beta'(x) = \left[\frac{pb + qa - (p+q)x}{(x-a)(b-x)}\right] \beta(x). \tag{9}$$

4. We have the following relationship between p, q, a, b, and c:

$$\frac{p}{q} = \frac{c-a}{b-c}. \tag{10}$$

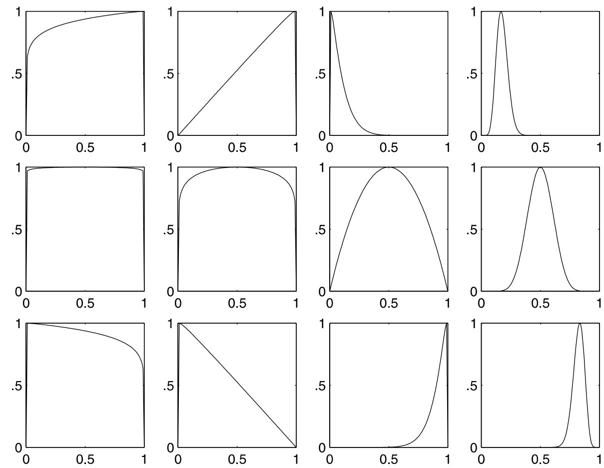

Fig. 1. Examples of Beta functions in one dimension.

Definition 3. (*Beta functions in the multidimensional case*) In the multidimensional case, a Beta function is given by

$$\beta(\vec{x}) = \begin{cases} \displaystyle\prod_{i=1}^{n} \beta_i(x_i) & \text{if } \vec{x} \in \displaystyle\prod_{i=1}^{n}]a_i, b_i[, \\ 0 & \text{otherwise,} \end{cases} \tag{11}$$

where $\vec{x} = (x_1, x_2, \dots, x_n)$ and $\beta_i(x_i) = \beta_i(x_i; p_i, q_i, a_i, b_i)$ is a one-dimensional Beta function.

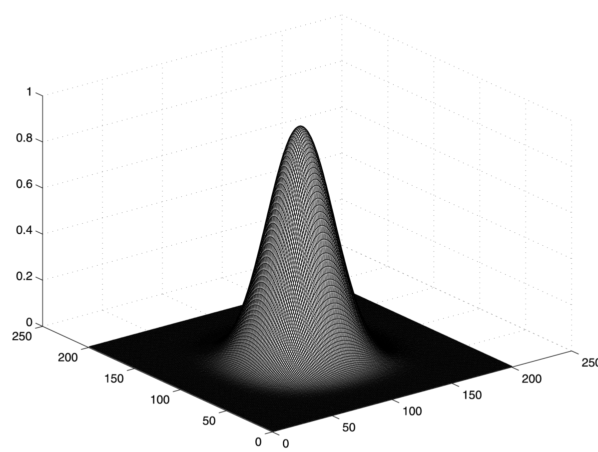

Fig. 2. Bivariate Beta function.

Definition 4. A *Beta Fuzzy Logic System* is an FLS given by (1) where the Beta functions are chosen as membership functions of the input variables.

3. Universal Approximation

Let U be a bounded set of $\mathbb{R}^n$ and $(\mathcal{C}(U), \|\cdot\|_\infty)$ be the set of all functions from U to $\mathbb{R}$, which are continuous with respect to the uniform norm (i.e. the norm given by $\|f\|_\infty = \sup_{x\in U} |f(x)|$ for every f in $\mathcal{C}(U)$).

Definition 5. A subset $\mathcal{A}$ of $\mathcal{C}(U)$ has the *universal approximation property* with respect to the norm $\|\cdot\|_\infty$ if for every $\varepsilon > 0$ and for every f in $\mathcal{C}(U)$ there exists g in $\mathcal{A}$ such that $\|f - g\|_\infty < \varepsilon$. In other words, $\mathcal{A}$ is dense in $(\mathcal{C}(U), \|\cdot\|_\infty)$.

Recently, Alimi (1997e; 1998b) proved that if U is a compact set of $\mathbb{R}^n$, then the family of functions from U to $\mathbb{R}$ such as

$$f(\vec{x}) = \sum_{i=1}^{N} f_i(\vec{x})\beta_i(\vec{x}) \tag{12}$$

is dense in $(\mathcal{C}(U), \|\cdot\|_\infty)$, i.e. for every continuous function g on a compact set there exists a function f given by (12) that approximates g arbitrarily well, where N is an arbitrary integer, the f_i's are polynomials in $x_1, x_2, \dots, x_n$ and the β_i's are N multidimensional beta functions. The proof, based on the Stone-Weierstrass theorem (Stone, 1937; 1948), consists in showing that this family is a non-empty subalgebra of $\mathcal{C}(U)$ which separates points and contains the identity function $f(x) = 1$. However, this result is not always useful, because in practice we need to design an FLS explicitly, i.e. to determine the number of fuzzy rules, to know the membership functions of the input variables and to fix the consequence of each fuzzy rule, etc.

In this paper, we propose a constructive approach to the design of BFLSs, and we need to recall some definitions and properties that can be found in (Glorennee, 1996; Zeng and Singh, 1994; 1995).

Definition 6. Let U be a bounded interval of $\mathbb{R}$. A *pseudo-trapezoid-shaped function* $PT(x; a, b, c, d, h)$ is a continuous function on U given by

$$PT(x; a, b, c, d, h) = \begin{cases} I(x) & \text{if} \quad x \in [a, b[, \\ h & \text{if} \quad x \in [b, c[, \\ D(x) & \text{if} \quad x \in]c, d], \\ 0 & \text{if} \quad x \in U \backslash [a, d], \end{cases} \tag{13}$$

where a, b, c and d are points of U such that $a \le b \le c \le d$, $a < d$ and h is a positive real number. I is a strictly increasing function on $[a, d[$, which is greater than or equal to zero, and D is a strictly decreasing function on $]c, d]$, which is also greater than or equal to zero.

While $h = 1$, instead of $PT(x; a, b, c, d, 1)$, we shall write $PT(x; a, b, c, d)$. In this case PT is said to be a normal pseudo-trapezoid-shaped function. Figure 3 shows three examples of pseudo-trapezoid-shaped functions, which are a triangular function, a trapezoid function and a Beta function.

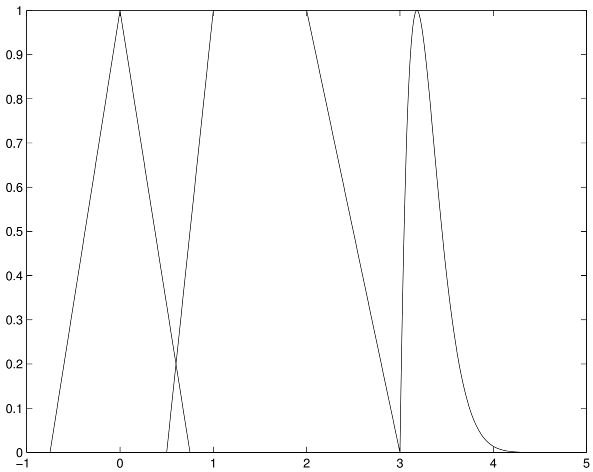

Fig. 3. Three examples of pseudo-trapezoid-shaped functions.

Definition 7. Let A be a fuzzy set defined on $U \subset \mathbb{R}^n$. A *normal subset* of A is the set

$$M(A) = \{\vec{x} \mid \vec{x} \in U \text{ and } A(\vec{x}) = 1\}. \tag{14}$$

Definition 8. A fuzzy set A defined on the universe of discourse U, is said to be *normal* if $0 \le A(x) \le 1$ for every $x \in U$.

Definition 9. (*The order between normal fuzzy sets*) Let A and B be two normal fuzzy sets defined on $U \subset \mathbb{R}$. We write $A > B$ if and only if $M(A) > M(B)$. Recall that $M(A) > M(B) \iff \forall x \in M(A),\ \forall y \in M(B) : x > y$.

Figure 4 shows an example of two triangular functions A and B satisfying $A < B$.

Definition 10. A function f defined on a subset $U = \prod_{j=1}^{n} U_j$ of $\mathbb{R}^n$ is said to be a *pseudo-trapezoid-shaped product* function if $f(\vec{x}) = \prod_{j=1}^{n} PT_j(x_j)$, where each PT_j is a pseudo-trapezoid-shaped function defined on U_j.

Definition 11. Fuzzy sets $(A_i)_{1\le i\le N}$ are said to form a *complete partition* of U if for every $\vec{x} \in U$ there exists $i \in \{1, \dots, N\}$ such that $A_i(\vec{x}) > 0$.

Definition 12. Fuzzy sets $(A_i)_{1\le i\le N}$ are said to be *consistent* in U if the following condition is satisfied: If $A_i(\vec{x}_0) = 1$ for $\vec{x}_0 \in U$, then $A_j(\vec{x}_0) = 0$ for every $i \ne j$.

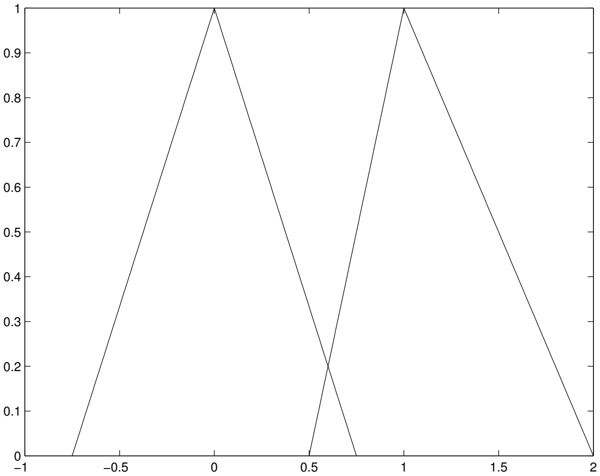

Fig. 4. Two fuzzy sets A and B such that $A < B$.

4. Approximation Properties of MIMO BFLSs

We suppose that the universe of discourse is $U = U_1 \times U_2 \times \cdots \times U_n$ where each U_j is a compact interval of $\mathbb{R}$. Multi-input–multi-output Beta fuzzy logic systems have the following approximation properties (for proofs, see Appendix):

Proposition 1. *Let $[A_{(i_1,i_2,\ldots,i_n)}]_{(i_1,i_2,\ldots,i_n)\in I}$ be fuzzy sets defined on $U = \prod_{j=1}^{n} U_j$. Suppose that their membership functions are given by*

$$A_{(i_1,i_2,\ldots,i_n)}(\vec{x}) = \prod_{j=1}^{n} A^j_{i_j}(x_j), \tag{15}$$

which are pseudo-trapezoid-shaped product functions. Then the functions $(A_{(i_1,i_2,\ldots,i_n)})_{(i_1,i_2,\ldots,i_n)\in I}$ are normal, consistent and complete in U if and only if A^j_1, $A^j_2, \ldots, A^j_{N_j}$ are in U_j for every $j \in \{1, 2, \ldots, n\}$.

Proposition 2. *Let $(B_{(i_1,i_2,\ldots,i_n)})_{(i_1,i_2,\ldots,i_n)\in I}$ be fuzzy basis functions given by*

$$B_{(i_1,i_2,\ldots,i_n)}(\vec{x}) = \frac{\beta_{(i_1,i_2,\ldots,i_n)}(\vec{x})}{\sum\limits_{(k_1,k_2,\ldots,k_n)\in I} \beta_{(k_1,k_2,\ldots,k_n)}(\vec{x})}. \tag{16}$$

Then $B_{(i_1,i_2,\ldots,i_n)}(\vec{x}) = \prod_{j=1}^{n} B^j_{i_j}(x_j)$, where

$$B^j_{i_j}(x_j) = \frac{\beta^j_{i_j}(x_j)}{\sum\limits_{i_j=1}^{N_j} \beta^j_{i_j}(x_j)}, \tag{17}$$

$i_j = 1, \ldots, N_j$ and $j = 1, 2, \ldots, n$.

Theorem 1. *Let $U = [A, D]$ be the universe of discourse, and let $[\beta_i(x; p_i, q_i, a_i, b_i)]_{1\le i\le N}$ be a family of Beta functions such that $A = c_1$, $c_i \le a_{i+1} < b_i \le c_{i+1}$, for every $i \in \{1, \ldots, N-1\}$ and $D = c_N$, where $c_i = (p_i b_i + q_i a_i)/(p_i + q_i)$. Then this family satisfies the following conditions:*

- *$P_1 : (\beta_i)_{1\le i\le N}$ are pseudo-trapezoid-shaped,*
- *$P_2 : (\beta_i)_{1\le i\le N}$ are normal,*
- *$P_3 : (\beta_i)_{1\le i\le N}$ are consistent in the universe of discourse U,*
- *$P_4 : (\beta_i)_{1\le i\le N}$ are complete,*
- *$P_5 : \beta_1 < \beta_2 < \cdots < \beta_N$.*

Figure 5 shows a Beta function family satisfying Properties P_1–P_5.

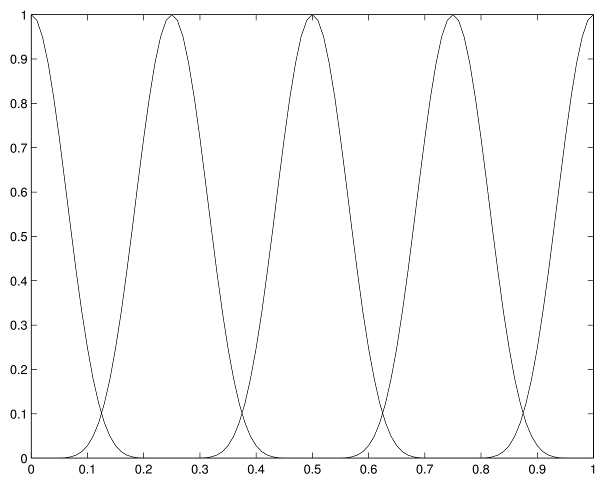

Fig. 5. Beta function family satisfying the conditions of Theorem 1.

Theorem 2. *Let $U = [A_1, D_1] \times [A_2, D_2] \cdots \times [A_n, D_n]$ and $[\beta_{(i_1,i_2,\ldots,i_n)}]_{1\le i_j\le N_j; 1\le j\le n}$ be a Beta multidimensional function family such that*

$$\beta_{(i_1,i_2,\ldots,i_n)}(\vec{x}) = \prod_{j=1}^{n} \beta^j_{i_j}(x_j), \tag{18}$$

where $\beta^j_{i_j}(x_j) = \beta(x_j; p^j_{i_j}, q^j_{i_j}, a^j_{i_j}, b^j_{i_j})$, and each family $(\beta^j_{i_j})_{1\le i_j\le N_j}$ satisfies the following conditions:

- *$A_j = c^j_1$,*
- *$c^j_{i_j} \le a^j_{i_j+1} < b^j_{i_j} \le c^j_{i_j+1}$ for all $i_j \in \{1, \ldots, N_j - 1\}$,*
- *$D_j = c^j_{N_j}$.*

Then $(\beta_{(i_1,i_2,\ldots,i_n)})_{1\le i_j\le N_j;1\le j\le n}$ *are product pseudo-trapezoid-shaped, normal, consistent and complete, and satisfy* $\beta_1^j < \cdots < \beta_{N_j}^j$ *for all* $1 \le j \le n$.

Remark 1. In the following, $g(\vec{x})$ will denote the control or the decision function to be approximated on U and $f(\vec{x})$ will be the function representing the BFLS. In what follows, we shall also assume that Properties P_1–P_5 are satisfied.

Let us now introduce the notation and definitions to give compact forms to our formulae. Thus, set

$$U_{i_j}^j = \begin{cases} [a_1^j, b_1^j[& \text{if} \quad i_j = 1, \\ ]a_{i_j}^j, b_{i_j}^j[& \text{if} \quad 2 \le i_j \le N_j - 1, \\ ]a_{N_j}^j, b_{N_j}^j] & \text{if} \quad i_j = N_j. \end{cases} \tag{19}$$

Let $I_{i_j}^j$ be the function defined from $U_{i_j}^j$ to I_{N_j}, where I_{N_j} is the set of subsets of $\{0, 1, \ldots, N_j + 1\}$ such that

$$I_{i_j}^j(x_j) = \begin{cases} \{i_j - 1, i_j\} & \text{if} \quad x_j \in]a_{i_j}^j, b_{i_j-1}^j[, \\ \{i_j\} & \text{if} \quad x_j \in [b_{i_j-1}^j, a_{i_j+1}^j], \\ \{i_j, i_j + 1\} & \text{if} \quad x_j \in]a_{i_j+1}^j, b_{i_j}^j[, \end{cases} \tag{20}$$

for all $i_j = 1, \ldots, N_j$ and $j = 1, \ldots, n$.

Lemma 1. *Under the assumptions of Theorem 2, we have*

$$U_j = \bigcup_{i_j=1}^{N_j} U_{i_j}^j \tag{21}$$

and

$$U = \bigcup_{(i_1,\ldots,i_n)\in I} U_{(i_1,i_2,\ldots,i_n)}, \tag{22}$$

where

$$U_{(i_1,i_2,\ldots,i_n)} = U_{i_1}^1 \times U_{i_2}^2 \times \cdots \times U_{i_n}^n. \tag{23}$$

Lemma 2. *Let* I_{U_j} *be the function defined from* U_j *to* I_{N_j} *such that*

$$I_{U_j}(x_j) = I_{i_j}^j(x_j) \ \text{ if } \ x_j \in U_{i_j}^j. \tag{24}$$

Then I_{U_j} *is a well-defined function, and so is the function* I_U *defined from* U *to* $I_{N_1} \times \cdots \times I_{N_n}$ *by*

$$I_U(\vec{x}) = I_{U_1}(x_1) \times \cdots \times I_{U_n}(x_n). \tag{25}$$

Theorem 3. (The basic approximation property) *Under the assumptions of Theorem 2 we have that for every* $\vec{x} \in U$

$$\begin{aligned} &|g(\vec{x}) - f(\vec{x})| \\ &= \Big| g(\vec{x}) - \sum_{(i_1,i_2,\ldots,i_n)\in I_U(\vec{x})} B_{(i_1,i_2,\ldots,i_n)}(\vec{x}) y_{(i_1,i_2,\ldots,i_n)} \Big| \\ &= \Big| g(\vec{x}) - \sum_{i_1\in I_{U_1}(x_1)} \sum_{i_2\in I_{U_2}(x_2)} \\ &\qquad \cdots \sum_{i_n\in I_{U_n}(x_n)} \Big(\prod_{j=1}^n B_{i_j}^j(x_j) \Big) y_{(i_1,i_2,\ldots,i_n)} \Big| \\ &\le \max_{(i_1,i_2,\ldots,i_n)\in I_U(\vec{x})} \Big\{ |g(\vec{x}) - y_{(i_1,i_2,\ldots,i_n)}| \Big\}. \end{aligned} \tag{26}$$

Theorem 4. (The uniform approximation properties) *Under the assumptions of Theorem 2, if we write*

$$\varepsilon_{(i_1,i_2,\ldots,i_n)} = \sup_{\vec{x}\in \operatorname{supp}(\beta_{(i_1,i_2,\ldots,i_n)})} \Big| g(\vec{x}) - y_{(i_1,i_2,\ldots,i_n)} \Big| \tag{27}$$

where $\{(i_1, i_2, \ldots, i_n)\} \in I$, $\operatorname{supp}(\beta_{(i_1,i_2,\ldots,i_n)})$ *being the support of* $\beta_{(i_1,i_2,\ldots,i_n)}$ *and* $\varepsilon = \max_{(i_1,i_2,\ldots,i_n)\in I} \varepsilon_{(i_1,i_2,\ldots,i_n)}$, *then*

$$\|g - f\|_\infty \le \varepsilon. \tag{28}$$

Theorem 5. (The uniform convergence property) *Let* $\vec{a}_{(i_1,i_2,\ldots,i_n)} = (a_{i_1}^1, a_{i_2}^2, \ldots, a_{i_n}^n)$, $\vec{b}_{(i_1,i_2,\ldots,i_n)} = (b_{i_1}^1, b_{i_2}^2, \ldots, b_{i_n}^n)$ *and*

$$\delta_{(i_1,i_2,\ldots,i_n)} = \big\| \vec{b}_{(i_1,i_2,\ldots,i_n)} - \vec{a}_{(i_1,i_2,\ldots,i_n)} \big\|, \tag{29}$$

where $\|\cdot\|$ *is any norm on* $\mathbb{R}^n$, *and*

$$\delta = \max_{(i_1,i_2,\ldots,i_n)\in I} \delta_{(i_1,i_2,\ldots,i_n)}. \tag{30}$$

For every $(i_1, i_2, \ldots, i_n) \in I$, *if*

$$\begin{aligned} \inf_{\vec{x}\in \operatorname{supp}\beta_{(i_1,i_2,\ldots,i_n)}} g(\vec{x}) &\le y_{(i_1,i_2,\ldots,i_n)} \\ &\le \sup_{\vec{x}\in \operatorname{supp}\beta_{(i_1,i_2,\ldots,i_n)}} g(\vec{x}), \end{aligned} \tag{31}$$

then

$$\lim_{\delta\to 0} \|g - f\|_\infty = 0, \tag{32}$$

where

$$f(\vec{x}) = \sum_{(i_1,i_2,\ldots,i_n)\in I} B_{(i_1,i_2,\ldots,i_n)}(\vec{x}) y_{(i_1,i_2,\ldots,i_n)}. \tag{33}$$

Theorem 6. (The universal approximation property) *Let $g(\vec{x})$ be a continuous function defined on U and let $\varepsilon > 0$ be a fixed number. Then there is a BFLS given by the function f, such that*

$$\|g - f\|_\infty \leq \varepsilon. \tag{34}$$

5. Best Approximation Property

In this section, we shall deal with the essential definitions needed for the study of the best approximation property (Rudin, 1974; Yosida, 1974), and then we shall prove that BFLSs possess this property.

Definition 13. Let $\mathcal{A}$ be a subset of $(\mathcal{C}(U), \|\cdot\|_\infty)$, where $U \subset \mathbb{R}^n$.

- We define the distance of an element $f \in \mathcal{C}(U)$ to $\mathcal{A}$ by
$$d(f, \mathcal{A}) = \inf_{g \in \mathcal{A}} \|f - g\|_\infty. \tag{35}$$
- An element $f_0 \in \mathcal{C}(U)$ is said to be a *best approximation* from f to $\mathcal{A}$ if
$$d(f, \mathcal{A}) = \|f - f_0\|_\infty. \tag{36}$$
- A subset $\mathcal{A}$ of $\mathcal{C}(U)$ is said to be an *existence set* if for every $f \in \mathcal{C}(U)$ there is an element $f_0 \in \mathcal{A}$ such that $\|f - f_0\|_\infty = d(f, \mathcal{A})$. In this case we say that $\mathcal{A}$ has the *best approximation property*.
- A subset $\mathcal{A}$ of $(\mathcal{C}(U), \|\cdot\|_\infty)$ is a *Tchebycheff set* if for every $f \in \mathcal{C}(U)$ there is a unique element $f_0 \in \mathcal{A}$ such that $\|f - f_0\|_\infty = d(f, \mathcal{A})$.

To study the best approximation property, we will use the following characterizations, which can be found in (Rudin, 1974).

Lemma 3. *Let $\mathcal{A}$ be a subset of $(\mathcal{C}(U), \|\cdot\|_\infty)$. If $\mathcal{A}$ is an existence set, then it is closed.*

Lemma 4. *Every closed, bounded subset of a finite dimensional linear subspace is compact.*

Lemma 5. *If $\mathcal{A}$ is a compact set of $(\mathcal{C}(U), \|\cdot\|_\infty)$, then $\mathcal{A}$ is an existence set.*

5.1. BFLSs Are Best Approximators with Respect to $\|\cdot\|_\infty$

Poggio and Girosi (1990) proved that multilayer perceptrons of the backpropagation type do not have the best approximation property. If we consider such a network with m hidden units, then the functions that it can compute belong to the following set with σ being a sigmoid function:

$$\sigma^m = \Big\{ f \in \mathcal{C}(U) \mid f(\vec{x}) = \sum_{i=1}^{m} c_i \sigma(\vec{x} \cdots \vec{w}_i + \theta_i); \; c_i, \theta_i \in \mathbb{R} \text{ and } \vec{w}_i \in \mathbb{R}^n \Big\}. \tag{37}$$

It was proved that σ^m is not closed so it cannot be an existence set (Girosi and Poggio, 1990). On the other hand, the same authors proved that RBF neural networks are best approximators (Girosi and Poggio, 1990). The principal question is as follows: Do BFLSs satisfy the property of best approximation? The answer is positive and to prove it, we need the following lemmas.

Lemma 6. *Under the assumptions of Theorem 2, the set $\mathcal{B}_N$ of functions from U to $\mathbb{R}$ such that*

$$f(\vec{x}) = \sum_{(i_1, i_2, \ldots, i_n) \in I} B_{(i_1, i_2, \ldots, i_n)}(\vec{x}) y_{(i_1, i_2, \ldots, i_n)} \tag{38}$$

where $y_{(i_1, i_2, \ldots, i_n)} \in \mathbb{R}$ is an N-dimensional linear subspace of $\mathcal{C}(U)$, where $N = \prod_{j=1}^{n} N_j$.

Lemma 7. *Let f be an element of $\mathcal{C}(U) \setminus \mathcal{B}_N$. Then the set*

$$A = \big\{ g \in \mathcal{B}_N \mid \|f - g\|_\infty \leq \|f\|_\infty \big\} \tag{39}$$

is a compact set of $(\mathcal{C}(U), \|\cdot\|_\infty)$.

Now we will outline the main result of this section.

Theorem 7. *The set $\mathcal{B}_N$ of BFLSs satisfying the assumptions of Theorem 2 has the best approximation property.*

In the next section, we will see that if we are looking for the best approximation in a Hilbert space, then it is unique.

5.2. Existence and Unicity of the Best Approximation in $L^2(U)$

$L^2(U)$ is the space of all functions defined from U to $\mathbb{R}$, which satisfy

$$\|f\|_2 = \Big(\int_U |f(t)|^2 \, \mathrm{d}t \Big)^{\frac{1}{2}} < +\infty, \tag{40}$$

endowed with the scalar product

$$\langle f | g \rangle = \int_U f(t) g(t) \, \mathrm{d}t. \tag{41}$$

$L^2(U)$ is a Hilbert space.

Theorem 8. *The set $\mathcal{B}_N$ of BFLSs satisfying the assumptions of Theorem 2 is a Tchebycheff set with to the norm $\|\cdot\|_2$, i.e. for every $f \in L^2(U)$ there is a unique $f_0 \in \mathcal{B}_N$ such that*

$$\|f - f_0\|_2 = \inf_{g \in \mathcal{B}_N} \|f - g\|_2. \tag{42}$$

6. Interpolation Property

In the previous sections we have shown that for every continuous function defined on a compact set of $\mathbb{R}^n$, we can construct a BFLS approximating it arbitrarily well. We have also proved that there is a best approximator to any continuous function in the set B_N of BFLSs with N fuzzy rules. In this section, we consider a continuous function f defined on U and taking the values $y_1, y_2, \dots, y_N$ at N distinct points $x_1, x_2, \dots, x_N$ of U. We are interested in finding a BFLS modelled by g that also satisfies $g(x_i) = y_i$ for every $i \in \{1, 2, \dots, N\}$.

Case 1: n=1

Because the x_i's are all distinct, we can arrange them so that $x_1 < \dots < x_n$. Let

$$d_i = \min\Big(\frac{x_{i+1} - x_i}{3}, \frac{x_i - x_{i-1}}{3}\Big), \tag{43}$$

$a_i = x_i - d_i$ and $b_i = x_i + d_i$. Consider the function

$$\beta_i(x) = \begin{cases} \dfrac{4}{(b_i - a_i)^2}(x - a_i)(b_i - x) & \text{if } x \in [a_i, b_i], \\ 0 & \text{otherwise.} \end{cases} \tag{44}$$

Then $\beta_i(x_i) = \beta_i(\frac{a_i+b_i}{2}) = 1$ and $\beta_i(x_j) = 0$ for all $j \neq i$.

The function

$$g(x) = \sum_{i=1}^{N} y_i \frac{\beta_i(x)}{\sum_{j=1}^{N} \beta_j(x)} \tag{45}$$

satisfies $g(x_i) = y_i$ for all $i = 1, 2, \dots, N$.

Case 2: $n \geq 2$

Let $\vec{x}^j = (x_1^j, \dots, x_n^j)$, $1 \leq j \leq N$ be N distinct vectors of $\mathbb{R}^n$. For each x_i^j define the one-dimensional Beta functions (β_i^j) satisfying the following hypothesis:

- If $x_i^j = x_i^k$, then $\beta_i^j = \beta_i^k$ and $\beta_i^j(x_i^j) = \beta_i^j(x_i^k) = 1$.
- If $x_i^j \neq x_i^k$, then $\beta_i^j(x_i^k) = 0$.

Let

$$\beta_j(\vec{x}) = \prod_{i=1}^{n} \beta_i^j(x_i) \tag{46}$$

for every $\vec{x} = (x_1, \dots, x_n) \in \mathbb{R}^n$. Then $\beta_j(\vec{x}^j) = 1$ and $\beta_j(\vec{x}^k) = 0$ for every $k \neq j$. Because we know that $\vec{x}^k \neq \vec{x}^j$, we can find $i_0 \in \{1, \dots, n\}$ such that $x_{i_0}^k \neq x_{i_0}^j$ and $\beta_{i_0}^j(x_{i_0}^k) = 0$. In consequence, $\beta_j(\vec{x}^k) = \prod_{i=1}^n \beta_i^j(x_i^k) = 0$.

The BFLS modelled by

$$g(\vec{x}) = \sum_{i=1}^{N} y_i \frac{\beta_i(\vec{x})}{\sum_{j=1}^{N} \beta_j(\vec{x})} \tag{47}$$

satisfies $g(x_i) = y_i$ for all $i \in \{1, \dots, N\}$. This leads to the following result:

Theorem 9. (BFLSs possess the interpolation property) *Let f be a continuous function defined on U, ($U = \prod_{j=1}^n [A_j, D_j]$ is a compact set of $\mathbb{R}^n$) such that $f(x_i) = y_i$ for all $i \in \{1, 2, \dots, N\}$ where the x_i's are N distinct points of U and $y_i \in \mathbb{R}$. Then there is a BFLS $g \in B_N$ such that*

$$g(x_i) = f(x_i) \text{ for all } i \in \{1, 2, \dots, N\}. \tag{48}$$

7. Conclusion

The study of the approximation properties of Beta fuzzy logic systems (BFLS) is an indispensable theoretic foundation for the users of such systems. In this paper we have proved that BFLSs have the following properties:

1. Basic approximation property. It gives an idea of the approximation mechanism by BFLSs.
2. Uniform approximation property. This property enables us to check if the designed BFLS has the desired approximation accuracy and suggests an idea to improve the approximation accuracy of our BFLS.
3. Uniform convergence property. It shows that we can improve the approximation accuracy of BFLSs by dividing the input space into finer fuzzy regions, which can be achieved by increasing the number of membership functions of the input variables.
4. Universal approximation property. From this property we conclude that for every function g which is continuous on a compact set, there is an explicitly designed BFLS which approximates g with an arbitrary given degree of accuracy.

5. Best approximation property. This property is of paramount interest, because in practice we fix the number of rules N and we look for a best approximator to the control function in the set of BFLSs with N fuzzy rules.

6. Interpolation property. It ensures that we can interpolate any continuous function g defined on a compact set with a BFLS. The number of fuzzy rules is equal to the number of points at which the values of g are known.

Our future work will concern the development of efficient learning algorithms for BFLS.

Acknowledgement

The authors wish to thank Professor Nabil Derbel for fruitful discussions. They also would like to acknowledge the financial support of this work by grants from the General Direction of Scientific Research (DGRST), Tunisia, under the ARUB program.

References

Alimi A.M. (2002): *Evolutionary computation for the recognition of on-line cursive handwriting*. — IETE J. Research, Special Issue on “Evolutionary Computation in Engineering Sciences” (S.K. Pal *et al.*, Eds.), in press.

Alimi M.A. (2000): *The Beta system: Toward a change in our use of neuro-fuzzy systems*. — Int. J. Manag., pp. 15–19.

Alimi A., Hassine R. and Selmi M. (2000): *Beta fuzzy logic systems: Approximation properties in the SISO case*. — Int. J. Appl. Math. Comput. Sci., Vol. 10, No. 4, pp. 857–875.

Alimi M.A. (1998a): *Segmentation of on-line cursive handwriting based on genetic algorithms*. — Proc. IEEE/IMACS Multiconf. *Computational Engineering in Systems Applications, CESA'98*, Hammamet, Tunisia, Vol. 4, pp. 435–438.

Alimi M.A. (1998b): *What are the advantages of using the Beta neuro-fuzzy system?* — Proc. IEEE/IMACS Multiconf. *Computational Engineering in Systems Applications, CESA'98*, Hammamet, Tunisia, Vol. 2, pp. 339–344.

Alimi M.A. (1998c): *Recognition of on-line handwritten characters with the Beta fuzzy neural network*. — Proc. IEEE/IMACS Multiconf. *Computational Engineering in Systems Applications, CESA'98*, Hammamet, Tunisia, Vol. 2, pp. 335–338.

Alimi M.A. (1997a): *The recognition of arabic handwritten characters with the Beta neuro-fuzzy network*. — Proc. 17éme *Journées Tunisiennes d'Électrotechnique et d'Automatique, JTEA'97*, Nabeul, Tunisia, Vol. 1, pp. 349–356.

Alimi M.A. (1997b): *The Beta fuzzy system: Approximation of standard membership ffunctions*. — Proc. 17éme *Journées Tunisiennes d'Électrotechnique et d'Automatique, JTEA'97*, Nabeul, Tunisia, Vol. 1, pp. 108–112.

Alimi M.A. (1997c): *An evolutionary neuro-fuzzy approach to recognize on-line arabic handwriting*. — Proc. Int. Conf. *Document Analysis and Recognition, ICDAR'97*, Ulm, Germany, pp. 382–386.

Alimi M.A. (1997d): *A neuro-fuzzy approach to recognize arabic hand written characters*. — Proc. Int. Conf. *Neural Networks, ICNN'97*, Houston, TX, USA, pp. 1397–1400.

Alimi M.A. (1997e): *Beta fuzzy basis functions for the design of universal robust neuro-fuzzy controllers*. — Proc. *Séminaire sur la Commande Robuste & Ses Applications, SCRA'97*, Nabeul, Tunisia, pp. C1–C5.

Bouchon-Meunier B. (1995): *La logique floue et ses applications*. — Paris: Addison-Wesley.

Castro J.L. and Delgado M. (1996): *Fuzzy systems with defuzzification are universal approximators*. — IEEE Trans. Syst. Man Cybern., Vol. 26, pp. 149–152.

Dickerson J.A. and Kosko B. (1996): *Fuzzy function approximation with ellipsoidal rules*. — IEEE Trans. Sys. Man Cybern., Vol. 26, pp. 542–560.

Girosi F. and Poggio T. (1990): *Networks and best approximation property*. — Biol. Cybern., Vol. 63, pp. 169–176

Glorennee P.-Y. (1996): *Quelques aspects analytiques des systèmes d'inférence floue*. — RAIRO – APII – JESA, Vol. 30, Nos. 2–3, pp. 231–254.

Gorrini V., Salome T. and Bersini H. (1995): *Self-structering fuzzy systems for function approximation*. — Proc. IEEE Int. Conf. *Fuzzy Syst.*, Yokohama, Japan, pp. 919–926.

Hartani R., Nguyen T.H. and Bouchon-Meunier B.(1996): *Sur l'approximation universelle des systèmes flous*. — RAIRO – APII – JESA, Vol. 30, No. 5, pp. 645–663.

Hassine R., Alimi M.A. and Selmi M. (2000): *What about the best approximation property of Beta fuzzy logic systems?* In: New Frontiers in Computational Intelligence and Its Applications (M. Mohammadian, Ed.). — Amsterdam: IOS Press, The Netherlands, pp. 62–67.

Jang J.-S-R. (1993): *ANFIS: Adaptive network based fuzzy inference systems*. — IEEE Trans. Syst. Man Cybern., Vol. 23, No. 3, pp. 665–685.

Jang J.-S-R. and Sun C.-T. (1995): *Neuro-fuzzy modeling and control*. — Proc. IEEE, Vol. 83, No. 3, pp. 378–406.

Johnson N.I. (1970): *Continuous Univariate Distributions*. — Boston: Houghton Mifflin Co.

Kosko B. (1993): *Fuzzy Thinking: The New Science of Fuzzy Logic*. — New York: Hyperion.

Kosko B. (1992): *Fuzzy systems as universal approximators*. — Proc. IEEE Int. Conf. *Fuzzy Syst.*, San Diego, CA, pp. 1153–1162.

Laukonen E.G. and Passino K.M. (1994): *Fuzzy systems for function approximation with applications to failure estimation*. — Proc. IEEE Int. Symp. *Intell. Contr.*, Columbus, OH, pp. 184–189.

Lee C.C. (1990): *Fuzzy logic in control systems: Fuzzy logic control – Part I*. — IEEE Trans. Syst. Man Cybern., Vol. 20, No. 2, pp. 404–418.

Lewis F.L., Zhu S.-Q. and Liu K. (1995): *Function approximation by fuzzy systems*. — Proc. *Amer. Contr. Conf.*, Seattle, WA, pp. 3760–3764.

Mamdani E.H. and Assilian S. (1975): *An experiment in linguistic synthesis with a fuzzy logic controller*. — Int. J. Man-Mach. Stud., Vol. 7, No. 1, pp. 1–13.

Mao Z.H., Li Y.D. and Zhang X.F. (1997): *Approximation capability of fuzzy systems using translations and dilations of one fixed function as membership functions*. — IEEE Trans. Fuzzy Syst., Vol. 5, No. 3. pp. 468–473.

Masmoudi M., Samet M., and Alimi M.A. (2000): *A bipolar implementation of the Beta neuron*. — Int. J. Electron., Vol. 87, No. 6, pp. 675–682.

Mendel J.M. (1995): *Fuzzy logic systems for engineering: A tutorial*. — Proc. IEEE, Vol. 83, No. 3, pp. 345–377.

Nguyen H.T., Kreinovich V. and Sirisaengtaksin O. (1996): *Fuzzy control as a universal control tool*. — Fuzzy Sets Syst., Vol. 80, No. 1, pp. 71–86.

Nguyen H.T. and Kreinovich V. (1992): *On approximations of controls by fuzzy systems*. — Tech. Rep., No. TR 92-93/302, LIFE Chair of Fuzzy Theory, Tokyo Institute of Technology.

Pedrycz W. (1994): *Why triangular membership functions?* — Fuzzy Sets Syst., Vol. 64, No. 1, pp. 21–30.

Rudin W. (1974): *Real and Complex Analysis*. — New York: McGraw-Hill.

Stone M.H. (1937): *Applications of the theory of Boolean rings to general topology*. — AMS Trans., Vol. 41, pp. 375–481.

Stone M.H. (1948): *The generalized weierstrass approximation theorem*. — Math. Mag., Vol. 21, pp. 167–183, 237–254.

Sugeno M. and Kang G.T. (1988): *Structure identification of fuzzy models*. — Fuzzy Sets Syst., Vol. 28, No. 1, pp. 15–33.

Takagi T. and Sugeno M.(1985): *Fuzzy identification of systems and its applications to modeling and control*. — IEEE Trans. Syst. Man Cybern., Vol. 15, pp. 116–132.

Terano T., Asai K. and Sugeno M. (1992): *Fuzzy Systems Theory and its Applications*. — New York: Academic Press.

Yen J., Langari R. and Zadeh L. (1995): *Industrial Applications of Fuzzy Logic and Intelligent Systems*. — New York: IEEE Press.

Yosida K. (1974): *Functional Analysis*. — Berlin: Springer.

Wang L.-X. (1992): *Fuzzy systems are universal approximators*. — Proc. IEEE Int. Conf. *Fuzzy Systems*, San Diego, CA. pp. 1163–1170.

Wang L.-X. and Mendel J.M. (1992): *Fuzzy basis functions, universal approximations, and orthogonal least-squares learning*. — IEEE Trans. Neural Netw., Vol. 3, No. 5, pp. 807–814.

Wang P.-Z., Tan S., Song F. and Liang P. (1997): *Constructive theory for fuzzy systems*. — Fuzzy Sets Syst., Vol. 88, No. 2, pp. 195–203.

Ying H. (1998): *General SISO Takagi–Sugeno fuzzy systems with linear rule consequent are universal approximators*. — IEEE Trans. Fuzzy Syst., Vol. 6, No. 4. pp. 582–587.

Zadeh L.A. (1965): *Fuzzy sets*. — Inf. Contr., Vol. 8, pp. 338–353.

Zeng X.-J.and Singh M.G.(1994): *Approximation theory of fuzzy systems—SISO Case*. — IEEE Trans. Fuzzy Syst., Vol. 2, No. 2, pp. 162–176.

Zeng X.-J.and Singh M.G.(1995): *Approximation theory of fuzzy systems—MIMO Case*. — IEEE Trans. Fuzzy Syst., Vol. 3, No. 2, pp. 219–235.

Appendix

In this appendix we give the proofs of all lemmas, propositions and theorems except Theorem 9, for which the proof is included in the paper.

Proof of Proposition 1. We have

$$\max_{\vec{x}\in U} A_{(i_1,i_2,\ldots,i_n)}(\vec{x}) = \prod_{j=1}^{n} \big(\max_{x_j\in U_j} A^j_{i_j}(x_j)\big). \tag{49}$$

Thus fuzzy sets $(A_{(i_1,i_2,\ldots,i_n)})_{(i_1,i_2,\ldots,i_n)\in I}$ are normal if and only if $A^j_{i_j}$ are also normal for all $i_j = 1,2,\ldots,N_j$ and $j = 1,2,\ldots,n$. From the equation

$$A_{(i_1,i_2,\ldots,i_n)}(\vec{x}) = \prod_{j=1}^{n} A^j_{i_j}(x_j) \tag{50}$$

we have $A_{(i_1,i_2,\ldots,i_n)}(\vec{x}) > 0$ if and only if $A^j_{i_j}(x_j) > 0$ for every $j \in \{1,2,\ldots,n\}$.

We can easily see that the completeness of the fuzzy sets $(A_{(i_1,i_2,\ldots,i_n)})_{(i_1,i_2,\ldots,i_n)\in I}$ is equivalent to that of $A^j_1, A^j_2, \ldots, A^j_{N_j}$ for every $j \in \{1,2,\ldots,n\}$.

Let us show that $A_{(i_1,i_2,\ldots,i_n)}, (i_1,i_2,\ldots,i_n) \in I$ are consistent in U if and only if $A^j_1, A^j_2, \ldots, A^j_{N_j}$ are consistent in U_j for every $j \in \{1,2,\ldots,n\}$.

First suppose that $A^j_1, A^j_2, \ldots, A^j_{N_j}$ are consistent in U_j for every $j \in \{1,2,\ldots,n\}$. Let $\vec{x}_0 = (x^0_1, x^0_2, \ldots, x^0_n)$ be a fixed element of U such that $A_{(i_1,i_2,\ldots,i_n)}(\vec{x}_0) = 1$. Then $A^j_{i_j}(x^0_j) = 1$ for every $j \in \{1,2,\ldots,n\}$. From the consistency of $A^j_1, A^j_2, \ldots, A^j_{N_j}$; $j \in \{1,2,\ldots,n\}$ we deduce that $A^j_{k_j}(x^0_j) = 0$ for all $k_j \neq i_j$, $j \in \{1,2,\ldots,n\}$. In consequence,

$$A_{(k_1,k_2,\ldots,k_n)}(\vec{x}_0) = \prod_{j=1}^{n} A^j_{k_j}(x^0_j) = 0 \tag{51}$$

for every $(k_1,k_2,\ldots,k_n) \neq (i_1,i_2,\ldots,i_n)$.

Conversely, suppose that $A^j_{i_0}(x^0_j) = 1$ for a fixed j. We know that $(A_{(i_1,i_2,\dots,i_n)})_{(i_1,i_2,\dots,i_n)\in I}$ are normal if and only if so are $A^j_1, A^j_2, \dots, A^j_{N_j}$ for all $j \in \{1,2,\dots,n\}$. Thus we can find x^0_k such that $A^k_1(x^0_k) = 1$ for $k = 1,\dots,j-1,j+1,\dots,n$.

Let $\vec{x}_0 = (x^0_1,\dots,x^0_j,\dots,x^0_n)$. Then we have $A_{(1,\dots,1,i_0,1\dots,1)}(\vec{x}_0) = 1$. We deduce that $A_{(1,\dots,1,i_j,1\dots,1)}(\vec{x}_0) = 0$ for all $i_j \neq i_0$ and, consequently, $A^j_{i_j}(x^0_j) = 0$ for all $i_j \neq i_0$. ■

Proof of Proposition 2. We can easily verify that

$$\sum_{(i_1,i_2,\dots,i_n)\in I} \prod_{j=1}^{n} \beta^j_{i_j}(x_j) = \prod_{j=1}^{n} \sum_{i_j=1}^{N_j} \beta^j_{i_j}(x_j). \quad (52)$$

Then

$$\begin{aligned} B_{(i_1,i_2,\dots,i_n)}(\vec{x}) &= \frac{\prod_{j=1}^{n} \beta^j_{i_j}(x_j)}{\sum_{(i_1,i_2,\dots,i_n)\in I} \beta_{(i_1,i_2,\dots,i_n)}(\vec{x})} \\ &= \frac{\prod_{j=1}^{n} \beta^j_{i_j}(x_j)}{\prod_{j=1}^{n} \Big(\sum_{i_j=1}^{N_j} \beta^j_{i_j}(x_j) \Big)} \\ &= \prod_{j=1}^{n} \left(\frac{\beta^j_{i_j}(x_j)}{\sum_{i_j=1}^{N_j} \beta_{i_j}(x_j)} \right) \\ &= \prod_{j=1}^{n} B_{i_j}(x_j). \end{aligned} \quad (53)$$

■

Proof of Theorem 1. $[P_1 + P_2]$: β_i is pseudo-trapezoid-shaped and normal. Let $x \in]a_i, b_i[$. We have $\beta'_i(x) = 0 \iff x = c_i = \frac{p_i b_i + q_i a_i}{p_i + q_i}$ and $c_i = \frac{p_i b_i + q_i a_i}{p_i + q_i} \in]a_i, b_i[$. Moreover, β_i is monotonically increasing on $]a_i, c_i[$ and monotonically decreasing on $]c_i, b_i[$.

We know that if $\beta_i(\frac{p_i b_i + q_i a_i}{p_i + q_i}) = 1$, then β_i is pseudo-trapezoid-shaped, normal, and $\beta_i(x; p_i, q_i, a_i, b_i) = PT(x; a_i, c_i, c_i, b_i)$

$[P_3]$: Consistency: $\beta_i(x) = 1 \iff x = c_i$ and $b_{i-1} \leq c_i \leq a_{i+1}$, so that $c_i \in \operatorname{supp}(\beta_i)$ and $c_i \notin \operatorname{supp}(\beta_j)$ for every $i \neq j$. In consequence, $\beta_j(x) = 0$ for every $i \neq j$.

$[P_4]$: Completeness: Let $x \in [A, D]$. If $x \in [A, b_1[$ then $\beta_1(x) > 0$ and $\beta_2(b_1) > 0$ because we know that $a_2 < b_1$. If $x \in]a_i, b_i[$ then $\beta_i(x) > 0$. If $x \in]a_N, D]$ then $\beta_N(x) > 0$ and $\beta_{N-1}(a_N) > 0$.

$[P_5]$: $M(\beta_i) = \{c_i\}$, since we know that $c_i < c_{i+1}$, so $\beta_i < \beta_{i+1}$. ■

Proof of Theorem 2. The proof of this theorem is evident while using Proposition 1 and Theorem 1. ■

Proof of Lemma 1. Let us prove that $U_j \subset \cup_{i_j=1}^{N_j} U^j_{i_j}$. The other inclusion is trivially satisfied.

Let x_j be an element of U_j. Since the fuzzy sets $\beta^j_1, \beta^j_2, \dots, \beta^j_{N_j}$ are complete in U_j, we can find $i_j \in \{1,\dots,N_j\}$ such that $\beta^j_{i_j}(x_j) > 0$, so $x_j \in \operatorname{supp}(\beta^j_{i_j}) = U^j_{i_j}$. In consequence, $U_j = \cup_{i_j=1}^{N_j} U^j_{i_j}$.

We will now prove that $U \subset \cup_{(i_1,\dots,i_n)\in I} U_{(i_1,i_2,\dots,i_n)}$. Let $\vec{x} = (x_1,x_2,\dots,x_n) \in U = U_1 \times \dots \times U_n$. Then $x_j \in U_j$ for all $j \in \{1,\dots,n\}$. By the above result, there exists i_j such that $x_j \in U^j_{i_j}$, i.e. $\vec{x} \in U_{(i_1,i_2,\dots,i_n)}$. In consequence, $U \subset \cup_{(i_1,\dots,i_n)\in I} U_{(i_1,i_2,\dots,i_n)}$. ■

Proof of Lemma 2. If $x_j \in U^j_{i_j}$ and $x_j \notin U^j_k$ for all $i_j \neq k$, then $I_{U_j}(x_j)$ is well defined. If $x_j \in U^j_{i_j} \cap U^j_k$ for some $k \neq i_j$ then, due to the inequalities $a^j_{i_j} < b^j_{i_j-1} < a^j_{i_j+1} < b^j_{i_j}$, we have $k = i_j + 1$ or $k = i_j - 1$. Hence the value of $I^j_k(x_j)$ is the same as the value of $I^j_{i_j}(x_j)$ in the two cases. In consequence, I_U is a well-defined function. ■

Proof of Theorem 3. We have

$$\begin{aligned} &|g(\vec{x}) - f(\vec{x})| \\ &= \Big| g(\vec{x}) - \sum_{(i_1,i_2,\dots,i_n)\in I} B_{(i_1,i_2,\dots,i_n)}(\vec{x}) y_{(i_1,i_2,\dots,i_n)} \Big| \\ &= \Big| g(\vec{x}) - \sum_{i_1=1}^{N_1} \cdots \sum_{i_n=1}^{N_n} B_{(i_1,i_2,\dots,i_n)}(\vec{x}) y_{(i_1,i_2,\dots,i_n)} \Big| \\ &= \Big| g(\vec{x}) \\ &\quad - \sum_{i_1=1}^{N_1} \cdots \sum_{i_n=1}^{N_n} B^1_{i_1}(x_1) \cdots B^n_{i_n}(x_n) y_{(i_1,i_2,\cdot,i_n)} \Big|, \end{aligned} \quad (54)$$

where $B^j_1(x_j), B^j_2(x_j), \dots, B^j_{N_j}(x_j)$ are given by (17).

For every $\vec{x} = (x_1,x_2,\dots,x_n) \in U$ there is i_1 such that $x_1 \in U_{i_1}$, so we have one of the following three cases:

1. $\boldsymbol{x_1 \in [a_{i_1}, b_{i_1-1}]}$
   We have $B^1_{i_1}(x_1) + B^1_{i_1-1}(x_1) = 1$ and $B^1_{j_1}(x_1) = 0$ for every $j_1 \neq i_1$ and $j_1 \neq i_1 - 1$, $j_1 \in$

$\{1, 2, \ldots, n_1\}$, so

$$\sum_{i_1=1}^{N_1} \cdots \sum_{i_n=1}^{N_n} B^1_{i_1}(x_1) \cdots B^n_{i_n}(x_n)$$
$$= \sum_{j_1\in\{i_1-1,i_1\}} \sum_{i_2=1}^{N_2} \cdots \sum_{i_n=1}^{N_n} B^1_{i_1}(x_1) \cdots B^n_{i_n}(x_n). \quad (55)$$

2. $\boldsymbol{x_1} \in [\boldsymbol{b_{i_1-1}}, \boldsymbol{a_{i_1+1}}]$.

   In this case $B^1_{i_1}(x_1) = 1$ and $B^1_{j_1}(x_1) = 1$ for every $j_1 \neq i_1$, so

$$\sum_{i_1=1}^{N_1} \sum_{i_2=1}^{N_2} \cdots \sum_{i_n=1}^{N_n} B^1_{i_1}(x_1) \cdots B^n_{i_n}(x_n)$$
$$= \sum_{i_2=1}^{N_2} \cdots \sum_{i_n=1}^{N_n} B^1_{i_1}(x_1) \cdots B^n_{i_n}(x_n). \quad (56)$$

3. $\boldsymbol{x_1} \in [\boldsymbol{a_{i_1+1}}, \boldsymbol{b_{i_1}}]$.

   For $B^1_{i_1}(x_1) + B^1_{i_1+1}(x_1) = 1$ and $B^1_{j_1}(x_1) = 0$ for every $j_1 \neq i_1$ and $j_1 \neq i_1 - 1$, $j_1 \in \{1, 2, \ldots, n_1\}$, we have

$$\sum_{i_1=1}^{N_1} \cdots \sum_{i_n=1}^{N_n} B^1_{i_1}(x_1) \cdots B^n_{i_n}(x_n)$$
$$= \sum_{j_1\in\{i_1,i_1+1\}1} \sum_{i_2=1}^{N_2} \cdots \sum_{i_n=1}^{N_n} B_{i_1}(x_1) \cdots B_{i_n}(x_n). \quad (57)$$

Accordingly,

$$\sum_{i_1=1}^{N_1} \cdots \sum_{i_n=1}^{N_n} B^1_{i_1}(x_1) \cdots B^n_{i_n}(x_n)$$
$$= \sum_{j_1\in I_{U_1}(x_1)} \sum_{i_2=1}^{N_2} \cdots \sum_{i_n=1}^{N_n} B^1_{j_1}(x_1) B^2_{i_2}(x_2) \cdots B^n_{i_n}(x_n). \quad (58)$$

Using the same method, we can also prove that

$$\sum_{i_2=1}^{N_2} \cdots \sum_{i_n=1}^{N_n} B^2_{i_2}(x_2) \cdots B^n_{i_n}(x_n)$$
$$= \sum_{j_2\in I_{U_2}(x_2)} \sum_{i_3=1}^{N_3} \cdots \sum_{i_n=1}^{N_n} B^2_{j_2}(x_2) \cdots B^n_{i_n}(x_n). \quad (59)$$

All these equalities give

$$|g(\vec{x}) - f(\vec{x})|$$
$$= \Big| g(\vec{x}) - \sum_{i_1=1}^{N_1} \cdots \sum_{i_n=1}^{N_n} B_{(i_1,\ldots,i_n)}(\vec{x}) y_{(i_1,\ldots,i_n)} \Big|$$
$$= \Big| g(\vec{x}) - \sum_{i_1\in I_{U_1(x_1)}} \cdots \sum_{i_n\in I_{U_n}} \Big(\prod_{j=1}^{n}\Big) B^j_{i_j}(x_j) y_{(i_1,\ldots,i_n)} \Big|$$
$$= \Big| \sum_{i_1\in I_{U_1(x_1)}} \cdots \sum_{i_n\in I_{U_n}(x_n)} \Big(\prod_{j=1}^{n} B^j_{i_j}(x_j)\Big)$$
$$\times \left(g(x_1, x_2, \ldots, x_n) - y_{(i_1,\ldots,i_n)}\right) \Big|$$
$$= \Big| \sum_{(i_1,i_2,\ldots,i_n)\in I_{U(\vec{x})}} B_{(i_1,i_2,\ldots,i_n)}(\vec{x}) \big[g(x_1, x_2, \ldots, x_n)$$
$$- y_{(i_1,\ldots,i_n)}\big] \Big|$$
$$\leq \max \big\{ |g(\vec{x}) - y_{(i_1,i_2,\ldots,i_n)}| \mid (i_1, i_2, \ldots, i_n)$$
$$\in I_U(\vec{x}) \big\}. \quad (60)$$

■

Proof of Theorem 4. We know that $\mathrm{supp}_{\beta_{(i_1,i_2,\ldots,i_n)}} \subset U$, which yields

$$\sup_{\vec{x}\in U} |g(\vec{x}) - f(\vec{x})|$$
$$= \sup_{\vec{x}\in U} \Big| g(\vec{x}) - \sum_{(i_1,i_2,\ldots,i_n)\in I_{U(\vec{x})}} B_{(i_1,i_2,\ldots,i_n)}(\vec{x}) y_{(i_1,\ldots,i_n)} \Big|$$
$$\geq \sup_{\mathrm{supp}_{\beta_{(i_1,i_2,\ldots,i_n)}}} \Big| g(\vec{x})$$
$$- \sum_{(i_1,i_2,\ldots,i_n)\in I_{U(\vec{x})}} B_{(i_1,i_2,\ldots,i_n)}(\vec{x}) y_{(i_1,\ldots,i_n)} \Big|$$
$$= \varepsilon_{(i_1,i_2,\ldots,i_n)}. \quad (61)$$

Thus $\sup_{\vec{x}\in U} |g(\vec{x}) - f(\vec{x})| \geq \varepsilon$.

On the other hand, we know that for every $\vec{x} \in U$ there is $(i_1, i_2, \ldots, i_n) \in I$ such that $\vec{x} \in \mathrm{supp}_{\beta_{(i_1,i_2,\ldots,i_n)}}$. This implies

$$|g(\vec{x}) - f(\vec{x})|$$
$$= \Big| g(\vec{x}) - \sum_{(i_1,i_2,\ldots,i_n)\in I_U(\vec{x})} B_{(i_1,i_2,\ldots,i_n)}(\vec{x}) y_{(i_1,\ldots,i_n)} \Big|$$
$$\leq \varepsilon_{(i_1,\ldots,i_n)} \leq \varepsilon. \quad (62)$$

■

Proof of Theorem 5. All the norms on $\mathbb{R}^n$ are equivalent, so we can use the infinity norm $\|\cdot\|_\infty$ i.e. the norm given by $\|\vec{x}\|_\infty = \max\{|x_i| \mid i = 1, 2, \ldots, n\}$ for every $\vec{x} \in \mathbb{R}^n$.

Let $\delta_{(i_1,\ldots,i_n)} = \|\vec{b}_{(i_1,\ldots,i_n)} - \vec{a}_{(i_1,\ldots,i_n)}\|_\infty$ and $\delta = \max\{\delta_{(i_1,\ldots,i_n)} \mid (i_1,\ldots,i_n) \in I\}$. Here g is continuous on the compact set U, so it is uniformly continuous. For every $\varepsilon > 0$ there is $\delta(\varepsilon) > 0$ such that $|g(\vec{x}) - g(\vec{x}')| < \varepsilon$ for every $\vec{x}$ and $\vec{x}'$ satisfying $\|\vec{x} - \vec{x}'\|_\infty < \delta(\varepsilon)$.

Let $m_{(i_1,i_2,\ldots,i_n)} = \inf_{\vec{x} \in \mathrm{supp}_{\beta_{(i_1,i_2,\ldots,i_n)}}} g(\vec{x})$ and $M_{(i_1,i_2,\ldots,i_n)} = \sup_{\vec{x} \in \mathrm{supp}_{\beta_{(i_1,i_2,\ldots,i_n)}}} g(\vec{x})$. We know that g is continuous on $\mathrm{supp}_{\beta_{(i_1,i_2,\ldots,i_n)}}$. Hence we can find $\vec{x}_m = (x_1^m, x_2^m, \ldots, x_n^m)$ and $\vec{x}_M = (x_1^M, x_2^M, \ldots, x_n^M) \in \mathrm{supp}_{\beta_{(i_1,i_2,\ldots i_n)}}$ such that $g(\vec{x}_m) = m_{(i_1,i_2,\ldots,i_n)}$ and $g(\vec{x}_M) = M_{(i_1,i_2,\ldots,i_n)}$.

On the other hand, $\mathrm{supp}_{\beta_{(i_1,i_2,\ldots i_n)}} = [a^1_{i_1}, b^1_{i_1}] \times [a^2_{i_2}, b^2_{i_2}] \times \cdots \times [a^n_{i_n}, b^n_{i_n}]$, so $x_j^m, x_j^M \in [a^j_{i_j}, b^j_{i_j}]$ for every $j = 1, 2, \ldots, n$. We have

$$\begin{aligned} \|\vec{x}_m - \vec{x}_M\|_\infty &= \max_{1 \le j \le n} \{|x_j^m - x_j^M|\} \\ &\le \max_{1 \le j \le n} \{|b_{i_j} - a_{i_j}|\} \\ &= \|\vec{b}_{(i_1,\ldots,i_n)} - \vec{a}_{(i_1,\ldots,i_n)}\|_\infty. \end{aligned} \tag{63}$$

Let $\varepsilon > 0$ be a fixed real number, and let $\delta < \delta(\varepsilon)$. We also know that $m_{(i_1,i_2,\ldots,i_n)} \le g(\vec{x}) \le M_{(i_1,i_2,\ldots,i_n)}$ and $m_{(i_1,i_2,\ldots,i_n)} \le y_{(i_1,i_2,\ldots,i_n)} \le M_{(i_1,i_2,\ldots,i_n)}$. Then $|g(\vec{x}) - y_{(i_1,i_2,\ldots,i_n)}| \le M_{(i_1,i_2,\ldots,i_n)} - m_{(i_1,i_2,\ldots,i_n)}$ for every $\vec{x} \in \mathrm{supp}_{\beta_{(i_1,i_2,\ldots,i_n)}}$.

From the uniform continuity of g we deduce that $\|\vec{x}_m - \vec{x}_M\|_\infty \le \delta < \delta(\varepsilon)$ and $|g(\vec{x}) - y_{(i_1,i_2,\ldots,i_n)}| \le g(\vec{x}_M) - g(\vec{x}_m) < \varepsilon$ for every $x \in \mathrm{supp}_{\beta_{(i_1,i_2,\ldots,i_n)}}$. In other words, $\sup_{x \in \mathrm{supp}_{\beta_{(i_1,i_2,\ldots,i_n)}}} |g(\vec{x}) - y_{(i_1,i_2,\ldots,i_n)}| \le \varepsilon$.

From Theorem 4 we conclude that

$$\sup_{x \in U} |g(\vec{x}) - f(\vec{x})| \le \varepsilon. \tag{64}$$

■

Proof of Theorem 6. $U = [A_1, D_1] \times [A_2, D_2] \times \cdots \times [A_n, D_n]$ is the universe of discourse. Let $\delta_j(N) = (D_j - A_j)/(N-1)$ and $x_{i_j}(N) = A_j + (i_j - 1)\delta_j(N)$, where $i_j \in \{0, 1, \ldots, N+1\}$. Then $A_j = x_{1_j}(N) < x_{2_j}(N) < \cdots < x_{N_j}(N) = D_j$.

We construct the following membership functions:

- $\beta_{1_j}(x, N) = \beta(x; p, p, x_{0_j}, x_{2_j})$ restricted to $[A_j, D_j]$,
- $\beta_{i_j}(x, N) = \beta(x; p, p, x_{i_j-1}, x_{i_j+1})$ for every $i_j \in \{2, \ldots, N-1\}$,
- $\beta_{N_j}(x, N) = \beta(x; p, p, x_{N_j-1}, x_{N_j+1})$ restricted to $[A_j, D_j]$,

where p is a strictly positive real number.

The consequent of each fuzzy rule is

$$y_{(i_1,i_2,\ldots,i_n)}(N) = g\Big(x_{i_1}(N), x_{i_2}(N), \ldots, x_{i_n}(N)\Big). \tag{65}$$

Then $m_{(i_1,i_2,\ldots,i_n)} \le y_{(i_1,i_2,\ldots,i_n)}(N) \le M_{(i_1,i_2,\ldots,i_n)}$, where

$$m_{(i_1,i_2,\ldots,i_n)} = \inf_{\vec{x} \in \mathrm{supp}(\beta_{(i_1,i_2,\ldots,i_n)})} g(\vec{x}), \tag{66}$$

$$M_{(i_1,i_2,\ldots,i_n)} = \sup_{\vec{x} \in \mathrm{supp}(\beta_{(i_1,i_2,\ldots,i_n)})} g(\vec{x}), \tag{67}$$

$$\beta_{(i_1,i_2,\ldots,i_n)}(\vec{x}) = \prod_{j=1}^{n} \beta^j_{i_j}(x_j). \tag{68}$$

Using Theorem 5, we deduce that

$$\lim_{N \to \infty} \sup_{x \in U} |g(\vec{x}) - f_N(\vec{x})| = 0, \tag{69}$$

where

$$f_N(x) = \sum_{(i_1,i_2,\ldots,i_n) \in I} B_{(i_1,i_2,\ldots,i_n)}(x, N) y_{(i_1,i_2,\ldots,i_n)}(N), \tag{70}$$

$$B_{(i_1,i_2,\ldots,i_n)}(x, N) = \frac{\beta_{(i_1,i_2,\ldots,i_n)}(x, N)}{\sum\limits_{(i_1,i_2,\ldots,i_n) \in I} \beta_{(i_1,i_2,\ldots,i_n)}(x, N)}. \tag{71}$$

Proof of Lemma 3. It is clear that $\mathcal{B}_N$ is a sublinear space of $\mathcal{C}(U)$. $\big(B_{(i_1,i_2,\ldots,i_n)}\big)_{(i_1,i_2,\ldots,i_n) \in I}$ is a generating family of $\mathcal{B}_N$. To prove that $\dim \mathcal{B}_N = N$, we will only prove that $\big(B_{(i_1,i_2,\ldots,i_n)}\big)_{(i_1,i_2,\ldots,i_n) \in I}$ are linearly independent.

Suppose that

$$\sum_{(i_1,i_2,\ldots,i_n) \in I} B_{(i_1,i_2,\ldots,i_n)}(\vec{x}) y_{(i_1,i_2,\ldots,i_n)} = 0. \tag{72}$$

For every $\vec{x} \in U$ let us show that $y_{(i_1,i_2,\ldots,i_n)} = 0$. Let $(k_1, k_2, \ldots, k_n)$ be a fixed index in I.

From Proposition 2 we have

$$B_{(k_1,k_2,\ldots,k_n)}(c_{k_1}, c_{k_2}, \ldots, c_{k_n}) = \prod_{j=1}^{n} B^j_{k_j}(c_{k_j}), \tag{73}$$

where

$$B^j_{k_j}(c_{k_j}) = \frac{\beta^j_{k_j}(c_{k_j})}{\sum\limits_{i_j=1}^{N_j} \beta^j_{i_j}(c_{i_j})}. \tag{74}$$

Then $B_{(k_1,k_2,\ldots,k_n)}(c_{k_1}, c_{k_2}, \ldots, c_{k_n}) \neq 0$ because $\beta^j_{k_j}(c_{k_j}) = 1$ and $B_{(i_1,i_2,\ldots,i_n)}(c_{k_1}, c_{k_2}, \ldots, c_{k_n}) = 0$ for every $(i_1, i_2, \ldots, i_n) \neq (k_1, k_2, \ldots, k_n)$. So $y_{(k_1,k_2,\ldots,k_n)} = 0$. ■

Proof of Lemma 4. It is clear that $\mathcal{A}$ is closed and bounded. Since $\mathcal{B}_N$ is finite dimensional, $\mathcal{A}$ is compact. ■

Proof of Theorem 7. The proof consists in showing that $\mathcal{B}_N$ is an existence set. Consider a fixed element f_0 of $\mathcal{C}(U)$. Then the closest point to f_0 in $\mathcal{B}_N$ is in the set $\{g \in \mathcal{B}_N \mid \|g - f_0\|_\infty \leq \|f - f_0\|_\infty\}$, where f is an arbitrary fixed element of $\mathcal{C}(U)$, which is a compact set by Lemma 4, and the result follows. ■

Proof of Theorem 8. Let $d = \inf_{g\in\mathcal{B}_N} \|f - g\|_2$. If $d = 0$ then $f \in \mathcal{B}_N$. In fact, $\mathcal{B}_N$ is a linear subspace of a finite dimension N, so it is closed and the result is then proved. If $d > 0$, then let B_n be the closed ball with centre f and radius $d + 1/n$, where n is a non-negative integer. The set $P_n = B_n \cap \mathcal{B}_N$ is convex and closed, because it is the intersection of two sets which are convex and closed ($\mathcal{B}_N$ is finite dimensional, so it is convex and closed). Moreover, P_n is non-empty. The element f_0 looked for is in the set $P = \cap_{n\in\mathbb{N}^\star} P_n$.

We will show that P is non-empty and reduced to one point. Let a and b be two elements of P_n and $m = (a + b)/2$. Then m is also an element of P_n because it is convex.

The parallelogram equality gives

$$2\|f - m\|_2^2 + \|b - a\|_2^2 = 2\Big(\|b - f\|_2^2 + \|f - a\|_2^2\Big). \quad (75)$$

The three quantities $\|f - m\|_2^2$, $\|f - a\|_2^2$ and $\|f - b\|_2^2$ are between d^2 and $d^2 + 1/n^2$, so

$$\|a - b\|_2^2 \leq \frac{4}{n}\Big(\frac{1}{n} + 2d\Big). \quad (76)$$

We then conclude that the diameter of P_n tends to 0 as $n \to \infty$. Then $(P_n)_n$ is a sequence of closed embedded, non-empty sets whose diameters tend to 0. Moreover, it is in $\mathcal{B}_N$, which is complete, so their intersection is non-empty and reduced to a unique point f_0. ■

Received: 26 November 2001
Revised: 4 July 2002
Re-revised: 28 November 2002