Int. J. Appl. Math. Comput. Sci., 2000, Vol.10, No.4, 857–875

BETA FUZZY LOGIC SYSTEMS: APPROXIMATION PROPERTIES IN THE SISO CASE

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

In this paper, a Single-Input Single-Output (SISO) Sugeno fuzzy model of the zeroth order with Beta membership functions for input variables is adopted. After the introduction of Beta Fuzzy Logic Systems (BFLS) a constructive theory is developed to establish the fact that they are universal approximators. Based on this theory, an algorithm, which can actually construct a BFLS approximating a given continuous function with an arbitrary degree of accuracy, is described. We then show that BFLSs satisfy more critical properties which are the best approximation property and the interpolation property. We complete the paper with a series of numerical comparisons between the approximation performances of BFLSs and other classes of widely used fuzzy logic systems. These comparisons confirm that BFLSs perform best in all the cases studied.

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

1. Introduction

Fuzzy Logic Systems (FLS) have recently attracted considerable attention and have been successfully applied in various fields (Abe and Lan, 1995; Castro and Delgado, 1996; Dickerson and Kosko, 1996; Gorrini *et al.*, 1995; Kosko, 1992; 1993; Laukonen and Passino, 1994; Lee, 1990; Lewis *et al.*, 1995; Mendel, 1995; Nguyen *et al.*, 1996; Wang, 1992; Wang and Mendel, 1992; Zadeh, 1965). One of the main advantages of FLSs is that they can be designed on the basis of incomplete and approximate information, thus providing simple and fast approximations of the unknown or very complicated models.

There are two major types of FLSs: Mamdani fuzzy systems (Mamdani and Assilian, 1975) and Takagi-Sugeno fuzzy systems (Sugeno and Kang, 1988; Takagi and

* REGIM: Research Group on Intelligent Machines, University of Sfax, National School of Engineers of Sfax, Department of Electrical Engineering, BP W 3038, Sfax, Tunisia, e-mail: adel.alimi@enis.rnu.tn

** Département de Mathématique, Faculté des Sciences de Monastir, Boulevard de l'environnement, Monastir 5000, Tunisia, e-mail: Radhia.Hassine@fsm.rnu.tn

*** Unité de Physique-Mathématique, University of Sfax, Faculty of Sciences of Sfax, Route de Soukra 3038, Sfax, Tunisia, e-mail: Mohamed.Selmi@fss.rnu.tn

Sugeno, 1985). The main difference between the two families lies in the consequence of the fuzzy rules which is a fuzzy set for the former and a crisp value for the latter. Many researchers have shown that Mamdani fuzzy systems are universal approximators (e.g., Kosko, 1992; 1993; Nguyen and Kreinovich, 1992; Wang, 1992; Wang and Mendel, 1991; 1992; Wang *et al.*, 1997; Zeng and Singh, 1994; 1995), but few were interested in the Sugeno fuzzy model. Recently, Ying (1998a) has proved that the Sugeno fuzzy model with a linear rule consequent is a universal approximator. In this paper, we are interested in a SISO Sugeno fuzzy model of the zeroth order. The advantage of such a model is that it is very simple: the consequence of each fuzzy rule is a constant and we do not need a defuzzification step to construct such a system. The Sugeno fuzzy model of the zeroth order is also simpler than that with a linear rule consequent. In fact, its number of design parameters is relatively small and there is no need for a simplified model that reduces the number of these parameters (Ying, 1998a; 1998b; 1998c; Ying and Sheppard, 1997). Besides, another important point which affects the behaviour of FLSs are the membership functions for the input variables. Various types of membership functions were proposed, e.g. triangular functions (Pedrycz, 1994), pseudo trapezoidal functions (Zeng and Singh, 1994; 1995), functions using translations and dilations of one fixed function (Mao *et al.*, 1997), normal peak functions (Wang *et al.*, 1997), etc. In this paper, we consider Beta Fuzzy Logic Systems (BFLS), i.e. the FLSs in which Beta functions (Johnson, 1970) are used as the membership functions of the input variables.

The paper is organised as follows: in Section 2, we recall some definitions which will be useful. Next, we introduce Beta fuzzy sets, give their essential properties and define Beta Fuzzy Logic Systems (BFLS). In Section 4, we deal with the property of universal approximation and prove by a constructive theory that BFLSs satisfy this property. Based on this theory, we describe an algorithm that can actually construct a BFLS approximating arbitrarily well a given continuous function. A more critical property which is the best approximation property is introduced in Section 5 and BFLSs are shown to satisfy this property. We also show that BFLSs satisfy the interpolation property. We complete the paper with a series of numerical simulations comparing the approximation performances of BFLSs with other widely used classes of FLSs. These comparisons confirm that BFLSs perform best in all the cases studied.

2. Preliminaries

In the remainder of this paper we adopt the SISO Sugeno fuzzy model of the zeroth order with multiplication as a t-norm. Such a fuzzy system is modelled by the function

$$f: \begin{array}{rcl} U \subset \mathbb{R} & \longrightarrow & V \subset \mathbb{R}, \\ x & \longmapsto & \displaystyle\sum_{i=1}^{N} \frac{A_i(x)}{\sum_{j=1}^{N} A_j(x)} y_i, \end{array} \tag{1}$$

where:

- x is the input variable,
- N is the number of fuzzy rules of the form

$$R_i: \text{ if } \big(x \text{ is } A_i\big) \text{ then } \big(y = y_i\big),$$

- y_i is a constant in V which represents the consequent of the fuzzy rule R_i,
- $\vec{A_i} = (A_1, A_2, \ldots, A_N)$ is a linguistic term characterised by its membership function $\mu_{A_i}(x)$, $i = \ldots, N$.

It is clear that the Sugeno fuzzy model of the zeroth order is a special case of the Takagi-Sugeno fuzzy model with a linear rule consequent. For the latter each fuzzy rule is of the form

$$R_i: \text{ if } \big(x \text{ is } A_i\big) \text{ then } \big(y = a_i x + b_i\big).$$

We will now recall some useful definitions which can be found in (Wang *et al.*, 1997; Zeng and Singh, 1994; 1995), etc.

Definitions 1. U always denotes the universe of discourse which is a subset of $\mathbb{R}$.

1. Normal peak function: Let $A(x)$ be a fuzzy set defined on the universe of discourse U. A is called a *normal peak function* if there exists a unique point $x_0 \in U$ such that $A(x) \leq A(x_0)$ for any $x \in U$ and $A(x_0) = 1$; x_0 is the *peak point* of A.

2. Normal basis set: A collection of fuzzy sets $(A_i)_{1 \leq i \leq N}$ defined on U is said to be a *normal basis set* if, for each $i \in \{1, 2, \ldots, N\}$, A_i is a normal peak function and $\sum_{i=1}^{N} A_i(x) = 1$ for any $x \in U$.

3. A 2-phase normal basis set: Let $(A_i)_{1 \leq i \leq N}$ be a collection of fuzzy sets defined on U. We suppose that $(A_i)_{1 \leq i \leq N}$ is a normal basis set. Then $(A_i)_{1 \leq i \leq N}$ is said to be *2-phase* if, for any $x \in U$, there are at most two consecutive normal peak functions A_i and A_{i+1} such that $A_i(x) \neq 0$ and $A_{i+1}(x) \neq 0$.

4. Pseudo Trapezoid-Shaped Function: Let U be a bounded interval of $\mathbb{R}$; a *pseudo trapezoid-shaped function* $A(x; a, b, c, d, h)$ is every continuous function on U given by

$$A(x; a, b, c, d, h) = \begin{cases} I(x), & x \in [a, b[, \\ h, & x \in [b, c[, \\ D(x), & x \in]c, d], \\ 0, & x \in U \backslash [a, d], \end{cases}$$

where a, b, c and d are points of U such that $a \leq b \leq c \leq d$, $a < d$ and h is a positive real number. I is a strictly increasing positive (or null) function on $[a, d[$ and D is a strictly decreasing positive (or null) function on $]c, d]$. While $h = 1$, we write $A(x; a, b, c, d)$ instead of $A(x; a, b, c, d, 1)$. In this case, A is said to be a *normal pseudo-trapezoid shaped function.*

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

$$M(A) = \{x\ ;\ x \in U \text{ and } A(x) = 1\}. \tag{2}$$

6. Order between fuzzy sets: Let A and B be two fuzzy sets defined on a subset U of $\mathbb{R}$. We say that $A > B$ if and only if $M(A) > M(B)$.

   Recall that $M(A) > M(B) \Longleftrightarrow \forall x \in M(A),\ \forall y \in M(B),\ x > y$.

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

8. Consistency: Fuzzy sets $(A_i)_{1 \le i \le N}$ are said to be *consistent* in U if the following property is satisfied: If $A_i(x_0) = 1$, for $x_0 \in U$, then $A_j(x_0) = 0$, for every $i \neq j$.

3. Beta Fuzzy Logic Systems

Beta functions have been proposed to be used as membership functions of the input variables (Alimi, 1998; Alimi, 1997a; 1997b; Johnson, 1970). This subsection is devoted to the introduction of the Beta function and its main properties.

Definition 2. Let $a, b \in \mathbb{R}$ satisfy $a < b$ and let $p, q > 0$. A *Beta function* is defined over $\mathbb{R}$ by

$$\beta(x) = \begin{cases} \left(\dfrac{x-a}{c-a}\right)^p \left(\dfrac{b-x}{b-c}\right)^q & \text{if } x \in]a, b[, \\ 0 & \text{otherwise,} \end{cases} \tag{3}$$

where $c = (pb + qa)/(p + q)$.

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

$$\beta(x) = \beta(x; p, q, a, b). \tag{4}$$

Proposition 1. *Beta functions have the following properties:*

1. *Every Beta function is continuous on* $\mathbb{R}$ *and has compact support.*

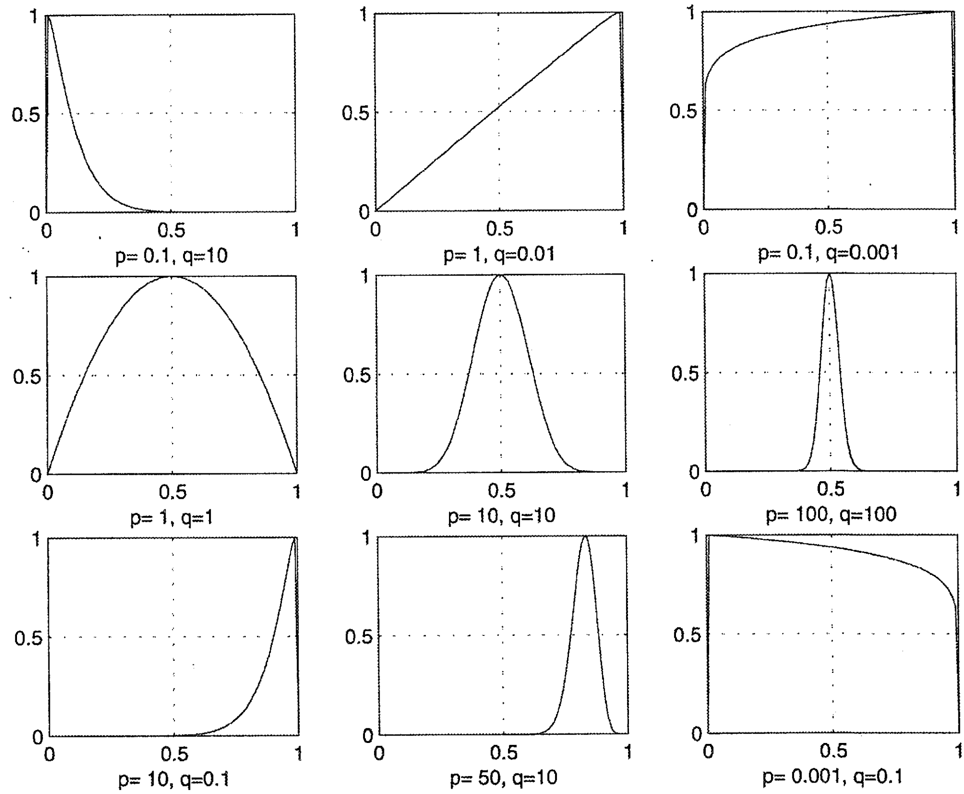

Fig. 1. Flexibility of Beta functions on [0,1].

2. *Every Beta function is a normal peak function, the peak point being*

$$c = \frac{pb + qa}{p + q}. \tag{5}$$

It is clear that $c \in]a, b[$.

3. *The restriction of a Beta function to* $[a, b]$ *is a pseudo-trapezoidal-shaped function.*

Proof.

1. This result is trivial.

2. β is differentiable over $]a, b[$ and

$$\beta'(x) = \left[\frac{pb + qa - (p + q)x}{(x - a)(b - x)}\right] \beta(x) \tag{6}$$

for every $x \in]a,b[$. Moreover, $\beta'(x) = 0$ if and only if $x = c$, $\beta'(x) > 0$ over $]a,c[$ and $\beta'(x) < 0$ over $]c,b[$, so that the Beta function achieves its maximum at c which is in $]a,b[$, and one can easily check that $\beta(c) = 1$. In fact, the term $(c-a)^{-p}(b-c)^{-q}$ has been chosen to normalise the Beta function: due to this term we have $0 \le \beta(x) \le 1$.

3. We have $\beta'(x) > 0$ over $]a,c[$ and $\beta'(x) < 0$ over $]c,b[$, so β is strictly increasing over $]a,c[$ and strictly decreasing over $]c,b[$. In consequence, it is a pseudo-trapezoidal-shaped function. ∎

Definition 3. A *SISO Beta Fuzzy Logic System* (*BFLS*) is every FLS given by (1), where a Beta function is chosen as the membership function of the input variables.

Definition 4. The *Beta Basis Functions* (*BBF's*) $(B_i)_{1\le i\le N}$ are given by

$$B_i(x) = \frac{\beta_i(x)}{\sum\limits_{j=1}^{N} \beta_j(x)}. \tag{7}$$

Proposition 2. *$(B_i)_{1\le i\le N}$ is a normal basis set.*

Proof. One can easily verify that each β_i is a normal peak function and that $\sum_{i=1}^{N} B_i(x) = 1$. ∎

4. BFLSs are Universal Approximators

4.1. Universal Approximation Property

Let $U = [A,D]$ be a compact set of $\mathbb{R}$ and $(\mathcal{C}(U), \|\cdot\|_\infty)$ the set of all continuous functions from U to $\mathbb{R}$, endowed with the uniform metric (i.e. the metric given by $\|f\|_\infty = \sup_{x\in U} |f(x)|$ for every f in $\mathcal{C}(U)$).

Definition 5. A subset $\mathcal{A}$ of $\mathcal{C}(U)$ satisfies 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 $(\mathcal{C}(U), \|\cdot\|_\infty)$.

In this section, we show by a constructive theory that BFLSs satisfy this property. For this purpose, we suppose in the remainder of this section that the BFLSs are given by

$$f(x) = \sum_{i=1}^{N} \frac{\beta_i(x; p_i, q_i, a_i, b_i)}{\sum\limits_{j=1}^{N} \beta_j(x; p_j, q_j, a_j, b_j)} y_i, \tag{8}$$

where $\left(\beta_i(x; p_i, q_i, a_i, b_i)\right)_{1 \leq i \leq N}$ is a family of Beta functions satisfying the following hypotheses:

(H_1) $A = c_1$,

(H_2) $c_i \leq a_{i+1} < b_i \leq c_{i+1}$ for every $i \in \{1, \ldots, N-1\}$,

(H_3) $D = c_N$.

Recall that $c_i = (p_i b_i + q_i a_i)/(p_i + q_i)$. Let us note that if we take the restriction of f to U, this restriction remains in $\mathcal{C}(U)$ even if y_1 and y_N are not zero.

Proposition 3. *Under hypotheses (H_1)–(H_3) we have:*

(P_1) $(\beta_i)_{1 \leq i \leq N}$ *are pseudo-trapezoidal-shaped and normal.*

(P_2) $\left(\beta_i\right)_{1 \leq i \leq N}$ *are consistent and complete in the universe of discourse* U.

(P_3) $\beta_1 < \beta_2 < \cdots < \beta_N$.

(P_4) $(B_i)_{1 \leq i \leq N}$ *is a 2-phase normal basis set.*

Proof.

(P_1) It is obvious.

(P_2) $\beta_i(x) = 1$ iff $x = c_i$. Moreover, $\beta_j(c_i) = 0$ for every $i \neq j$, so $\left(\beta_i\right)_{1 \leq i \leq N}$ are consistent. In addition to that, the supports of the Beta functions overlap because $a_{i+1} < b_i$. Hence for any $x \in U$ there is $i \in \{1, \ldots, N\}$ such that $\beta_i(x) > 0$.

(P_3) We have $M(\beta_i) = \{c_i\}$ and $c_i < c_{i+1}$, so $M(\beta_i) < M(\beta_{i+1})$ for every $i \in \{1 \leq i \leq N-1\}$. In consequence, $\beta_1 < \beta_2 < \cdots < \beta_N$.

(P_4) Every $x \in U$ is at most in the support of two consecutive Beta functions. Moreover, $\sum_{i=1}^{N} B_i(x) = 1$, so $(B_i)_{1 \leq i \leq N}$ is a 2-phase normal basis set. ■

Figure 2 shows a Beta function family satisfying hypotheses (H_1)–(H_3).

Lemma 1. *Let* $U = [A, D]$ *be a compact set of* $\mathbb{R}$. *Let* g *be a continuous function defined on* U. *If* $\varepsilon_i = \sup_{x \in \mathrm{supp}(\beta_i)} |g(x) - y_i|$ *and* $\varepsilon = \max\{\varepsilon_i \; ; \; 1 \leq i \leq N\}$, *where* $\mathrm{supp}(\beta_i)$ *is the support of* β_i, *then*

$$\|g - f\|_\infty \leq \varepsilon, \tag{9}$$

where $f(x) = \sum_{i=1}^{N} B_i(x) y_i$, $B_i(x) = \beta_i(x) / \sum_{i=1}^{N} \beta_i(x)$ *and* $(\beta_i)_{1 \leq i \leq N}$ *is a Beta function family satisfying hypotheses* (H_1)–(H_3).

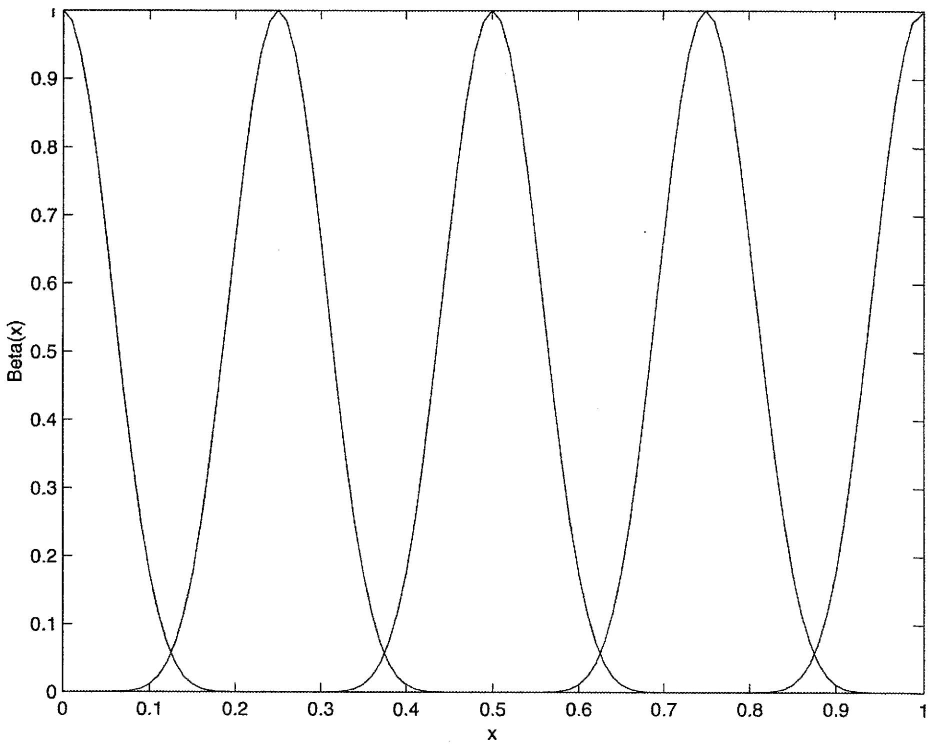

Fig. 2. A Beta function family satisfying hypotheses (H_1)–(H_3) with $N = 5$.

Proof. We can easily verify that the intervals $[b_{i-1}, a_{i+1}]$, $]a_{i+1}, b_i[$, where $i \in \{1, \ldots, N\}$, cover U (we adopt the convention $b_0 = A$ and $a_{N+1} = D$).

If $x \in U$, then $x \in [b_{i-1}, a_{i+1}]$ or $x \in]a_{i+1}, b_i[$. If $x \in [b_{i-1}, a_{i+1}]$, then $\beta_i(x) \neq 0$ and $\beta_j(x) = 0$ for every $i \neq j$, so $B_i(x) = 1$ and $B_j(x) = 0, \forall i \neq j$. In consequence,

$$|g(x) - f(x)| = |g(x) - y_i| \leq \sup_{x \in \text{supp}(\beta_i)} |g(x) - y_i| \leq \varepsilon_i. \tag{10}$$

If $x \in]a_{i+1}, b_i[$, then $\beta_i(x) \neq 0$, $\beta_{i+1}(x) \neq 0$ and $\beta_j(x) = 0$ for every $j \neq i$ and $j \neq i + 1$. Hence

$$\begin{aligned} |g(x) - f(x)| &= |g(x) - B_i(x)y_i - B_{i+1}(x)y_{i+1}| \\ &\leq \max\left\{ \sup_{x \in \text{supp}(\beta_i)} |g(x) - y_i|, \sup_{x \in \text{supp}(\beta_{i+1})} |g(x) - y_{i+1}| \right\} \\ &= \max\{\varepsilon_i, \varepsilon_{i+1}\} \leq \varepsilon. \qquad \blacksquare \end{aligned}$$

Lemma 2. *Let g be a continuous function defined on $U = [A, D]$ which is a compact set of $\mathbb{R}$ and let $\delta = \max\{b_i - a_i;\ 1 \le i \le N\}$ and $f_\delta(x) = \sum_{i=1}^{N} B_i(x) y_i$, where*

$$B_i(x) = \frac{\beta_i(x)}{\sum_{i=1}^{N} \beta_i(x)}$$

and

$$\left(\beta_i(x; p_i, q_i, a_i, b_i)\right)_{1 \le i \le N}$$

is a Beta function family satisfying the hypotheses (H_1)–(H_3). If $m_i \le y_i \le M_i$, where $m_i = \inf_{x \in [a_i, b_i]} g(x)$, and $M_i = \sup_{x \in [a_i, b_i]} g(x)$, then

$$\lim_{\delta \to 0} \|f_\delta - g\|_\infty = 0. \tag{11}$$

Proof. Given $\varepsilon > 0$, there exists $\delta(\varepsilon) > 0$ such that if $|x - x'| \le \delta(\varepsilon)$, then $|g(x) - g(x')| \le \varepsilon$, because g is continuous on $U = [A, D]$ which is a compact set so it is uniformly continuous.

For $\delta < \delta(\varepsilon)$ let us show that $\sup_{x \in U} |g(x) - f_\delta(x)| \le \varepsilon$. It is well-known that $m_i \le y_i \le M_i$. One can easily find $x_i \in [a_i, b_i]$ such that $g(x_i) = y_i$. Accordingly, $\sup_{x \in \mathrm{supp}(\beta_i)} |g(x) - y_i| = \sup_{x \in \mathrm{supp}(\beta_i)} |g(x) - g(x_i)|$,

$$\left.\begin{array}{l} m_i \le y_i \le M_i \\ m_i \le g(x) \le M_i \end{array}\right\} \Longrightarrow |g(x) - y_i| \le M_i - m_i,$$

Because $x, x_i \in [a_i, b_i]$, we have $\delta < \delta(\varepsilon)$ and $|x - x_i| < \delta(\varepsilon)$.

The difference $M_i - m_i$ is less than ε. Furthermore, g is continuous over $[a_i, b_i]$ which is a compact set, so its supremum and infumum are achieved. In consequence, $M_i - m_i = g(t_i) - g(s_i)$, s_i and t_i are two points of $[a_i, b_i]$, so $|s_i - t_i| < \delta$. Accordingly, $|g(x) - g(x_i)| \le \varepsilon \Longrightarrow \sup_{x \in \mathrm{supp}(\beta_i)} |g(x) - y_i| \le \varepsilon$. Thus, using the previous lemma, we conclude that

$$\sup_{x \in U} |g(x) - y_i| \le \varepsilon. \tag{12}$$

■

Theorem 1. *Let g be a continuous function defined on $U = [A, D]$ and $\varepsilon > 0$. Then there is a BFLS given by (8) such that*

$$\|g - f\|_\infty \le \varepsilon. \tag{13}$$

Proof. Let $\delta(N) = (D - A)/(N - 1)$ and $x_i(N) = A + (i - 1)\delta(N)$, where $i \in \{0, 1, \ldots, N + 1\}$. Then

$$A = x_1(N) < x_2(N) < \cdots < x_N(N) = D.$$

Further, we construct the fuzzy sets

$$A_1(x,N) = \beta(x;p,p,x_0,x_2) \text{ restricted to } [A,D],$$

$$A_i(x,N) = \beta(x;p,p,x_{i-1},x_{i+1}) \text{ for } i \in \{2,\ldots,N-1\},$$

$$A_N(x,N) = \beta(x;p,p,x_{N-1},x_{N+1}) \text{ restricted to } [A,D],$$

where p is a strictly positive real number.

The consequent of each fuzzy rule is $y_i(N) = g(x_i(N))$. Then $m_i \le y_i(N) \le M_i$, where $m_i = \min_{x\in[x_{i-1},x_{i+1}]} g(x)$ and $M_i = \max_{x\in[x_{i-1},x_{i+1}]} g(x)$. Recall that $\mathrm{supp}(A_i) = [x_{i-1}, x_{i+1}]$. From Lemma 2 we deduce that

$$\lim_{N\to\infty} \|g - f_N\|_\infty = \lim_{\delta\to 0} \|g - f_N\|_\infty = 0,$$

where $f_N(x) = \sum_{i=1}^{N} B_i(x,N) y_i(N)$ and $B_i(x,N) = A_i(x,N)/\sum_{i=1}^{N} A_i(x,N)$. ■

Corollary 1. *Let $L^p(U)$ denote the set of all functions from U to $\mathbb{R}$ such that $\|f\|_p < \infty$, where $\|f\|_p = (\int_U |f(t)|^p \,\mathrm{d}t)^{\frac{1}{p}}$. Then the set of BFLSs is dense in $L^p(U)$.*

Proof. The proof is immediate while using the density in $L^p(U)$ of continuous functions defined on compact supports, and the fact that U is bounded. ■

4.2. A Constructive Learning Algorithm

In this section, an algorithm that can actually construct a BFLS approximating a given continuous function with a given degree of accuracy, is described.

Algorithm. Let g be a continuous function to be approximated on $[A,D]$ and ε a strictly positive real number. We begin with $N = 2$.

Step 1. Let $\delta = (D-A)/(N-1)$ and $x_i(N) = A+(i-1)\delta$, where $i \in \{0,\ldots,N+1\}$.

Step 2. Construct Beta fuzzy sets as $\beta_i(x) = \beta(x;1,1,x_{i-1},x_{i+1})$ for $i \in \{1,\ldots,N\}$.

Step 3. Construct the BFLS as follows:

$$f(x) = \sum_{i=1}^{N} \frac{\beta_i(x;1,1,x_{i-1},x_{i+1})}{\sum_{j=1}^{N} \beta_j(x;1,1,x_{j-1},x_{j+1})} y_i,$$

where $y_i = g(x_i)$.

Step 4. If $\|f-g\|_\infty \le \varepsilon$ then STOP, otherwise set $N = N+1$ and go back to Step 1.

Remark 1. It is true that, for $i = 0$, $x_0 = A - \delta$ is outside the domain of the definition of the function g. The same happens for $i = N$: $x_N = D + \delta$ is also outside the domain of the definition of g. We only need to take the restriction of the function f_N to $[A, D]$. This restriction gives a universal approximator to g.

Theorem 2. *The foregoing algorithm converges.*

Proof. The function family $(\beta_i(x; 1, 1, x_{i-1}, x_{i+1}))_{1 \le i \le N}$ satisfies hypotheses (H_1)–(H_3). In fact

1. $c_1 = (x_0 + x_2)/2 = x_1 = A$
2. $c_i = (x_{i-1} + x_{i+1})/2 = x_i$, $a_{i+1} = x_i$ and $b_i = x_{i+1}$, so (H_2) is satisfied, and
3. $c_N = x_N = D$.

Using Theorem 1, we conclude that

$$\lim_{N \to \infty} \|f - g\|_\infty = 0. \tag{14}$$

■

5. BFLSs are Best Approximators

In this section, we deal with the essential definitions and properties concerning the best approximation property (Rudin, 1974; Yosida, 1974).

Definition 6. Let $\mathcal{A}$ be a subset of $\big(\mathcal{C}(U), \|\cdot\|_\infty\big)$, where $U \subset \mathbb{R}^n$.

1. We define the distance of an element $f \in \mathcal{C}(U)$ to $\mathcal{A}$ by
$$d(f, \mathcal{A}) = \inf_{g \in \mathcal{A}} \|f - g\|_\infty. \tag{15}$$
2. An element $f_0 \in \mathcal{C}(U)$ is said to be a *best approximation* from f to $\mathcal{A}$ if
$$d(f, \mathcal{A}) = \|f - f_0\|_\infty.$$
3. 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.*
4. A subset $\mathcal{A}$ of $\Big(\mathcal{C}(U), \|\cdot\|_\infty\Big)$ is a *Chebyshev set* if, for every $f \in \mathcal{C}(U)$, there is a unique element $f_0 \in \mathcal{A}$ such that $\|f - f_0\|_\infty = d(f, \mathcal{A})$.

Proposition 4.

1. *Let $\mathcal{A}$ be a subset of $\Big(\mathcal{C}(U), \|\cdot\|_\infty\Big)$. If $\mathcal{A}$ is an existence set, then it is closed.*
2. *Every closed and bounded subset of a finite-dimensional linear subspace is compact.*
3. *If $\mathcal{A}$ is a compact set of $\Big(\mathcal{C}(U), \|\cdot\|_\infty\Big)$, 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 perceptron neural networks of the backpropagation type do not possess the best approximation property. If we consider such a network with m hidden units, then the functions that it computes belong to the set

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

where σ is a sigmoid function. It has been proved that σ^m is not closed, so it cannot constitute an existence set (Girosi and Poggio, 1990).

On the other hand, the same authors proved that radial basis function networks are best approximators (Girosi and Poggio, 1990). We know that BFLSs and Beta neuro-fuzzy systems are functionally equivalent (Alimi, 1998). The question to raise is as follows: Do BFLSs possess the property of best approximation? The answer is positive and to prove it, we need the following lemmas.

Lemma 3. *Let $\mathcal{B}_N$ be the set of all functions defined on U and given by $f(x) = \sum_{i=1}^{N} B_i(x) y_i$, where $B_i(x) = \beta_i(x)/\sum_{j=1}^{N} \beta_j(x)$ and $y_i \in \mathbb{R}$. Suppose that $(\beta_i)_{1 \le i \le N}$ is a Beta function family satisfying hypotheses (H_1)–(H_3). Then $\mathcal{B}_N$ is a linear N-dimensional subspace of $\mathcal{C}(U)$.*

Proof. It is clear that $\mathcal{B}_N$ is a linear subspace of $\mathcal{C}(U)$. Since $(B_i)_{1 \le i \le N}$ is a generating family of $\mathcal{B}_N$, to prove that the dimension of $\mathcal{B}_N$ is equal to N, it suffices to prove that $(B_i)_{1 \le i \le N}$ are linearly independent. Suppose that $\sum_{i=1}^{N} B_i(x) y_i = 0$, for every $x \in U$. Let us show that $y_i = 0$. For $x = c_{i_0}$, we have $\sum_{i=1}^{N} B_i(c_{i_0}) y_i = 0$, so $y_{i_0} = 0$ because $B_{i_0}(c_{i_0}) = 1$ and $B_i(c_{i_0}) = 0$ for all $i \neq i_0$. ∎

Lemma 4. *Let the hypotheses of Lemma 3 be satisfied. If f is an element of $\mathcal{C}(U) \setminus \mathcal{B}_N$, then the set $\mathcal{A} = \{g \in \mathcal{B}_N; \|f - g\|_\infty \le \|f\|_\infty\}$ is compact in $(\mathcal{C}(U), \|\cdot\|_\infty)$.*

Proof. It is clear that $\mathcal{A}$ is closed and bounded and $\mathcal{B}_N$ is finite-dimensional, so $\mathcal{A}$ is compact. ∎

Now, we will give the main result of this section.

Theorem 3. *Under the same hypotheses as in Lemma 3, the set $\mathcal{B}_N$ has the property of the best approximation.*

Proof. The proof consists in showing that $\mathcal{B}_N$ is an existence set. Consider a fixed element f_0 of $\mathcal{C}(U)$. The closest point to f_0 in $\mathcal{B}_N$ is in the set

$$\{g \in \mathcal{B}_N; \|g - f_0\|_\infty \le \|f - f_0\|_\infty\},$$

where f is an arbitrary fixed element of $\mathcal{C}(U)$. By Lemma 4 the previous set is compact, and the result follows. ∎

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

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

Let us denote by $L^2(U)$ the space of all integrable functions defined from U to $\mathbb{R}$, which satisfy $\|f\|_2 = \left(\int_U |f(t)|^2\,\mathrm{d}t\right)^{\frac{1}{2}} < +\infty$, endowed with the scalar product

$$\langle f|g\rangle = \int_U f(t)g(t)\,\mathrm{d}t \tag{16}$$

It is well-known that $L^2(U)$ is a Hilbert space.

Theorem 4. *Under the same hypotheses as in Lemma 3, the set $\mathcal{B}_N$ is a Chebyshev set with respect 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*

$$\|f - f_0\|_2 = \inf_{g\in\mathcal{B}_N} \|f - g\|_2. \tag{17}$$

Proof. Let $d = \inf_{g\in\mathcal{B}_N} \|f - g\|_2$. If $d = 0$, then $f \in \mathcal{B}_N$. Since $\mathcal{B}_N$ is a linear subspace of a finite dimension N, it is closed and the result is proved.

If $d > 0$, set B_n as 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 as 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 sought element f_0 is in the set $P = \bigcap_{n\in\mathbb{G}^*} P_n$. 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 since it is convex. The parallelogram law gives

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

The 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

$$\|a - b\|_2^2 \leq \frac{4}{n}\left(\frac{1}{n} + 2d\right). \tag{18}$$

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

6. BFLSs Satisfy the Interpolation Property

In the previous sections, we have shown that for every continuous function defined on a compact set of $\mathbb{R}$, 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, given a continuous function f defined on U and taking the values $y_1, y_2, \ldots, y_N$ at N distinct points $x_1, x_2, \ldots, x_N$ of U, we are interested in finding a BFLS modelled by g that also satisfies $g(x_i) = y_i$ for every $i \in \{1, 2, \ldots, N\}$.

Theorem 5. *Let f be a continuous function defined on U such that $f(x_i) = y_i$ for all $i \in \{1, 2, \ldots, N\}$, where x_i are N distinct points of U and $y_i \in \mathbb{R}$. Then there is a BFLS $g \in B_N$ such that*

$$g(x_i) = f(x_i) \ \ \forall i \in \{1, 2, \ldots, N\}. \tag{19}$$

Proof. Because x_i are all distinct, we can arrange them such that $x_1 < \cdots < x_n$.

Let $d_i = \inf((x_{i+1} - x_i)/3, \ (x_i - x_{i-1})/3)$, $a_i = x_i - d_i$ and $b_i = x_i + d_i$. Consider the function

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

Then $\beta_i(x_i) = \beta_i((a_i + b_i)/2) = 1$ and $\beta_i(x_j) = 0$ for $i \neq j$. The function

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

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

7. Simulation Results

In order to confirm our theoretical results, we have performed two types of numerical simulations. Mitaim and Kosko (1996) compared the performances of different membership functions of the Standard Additive Model (SAM). In the one-dimensional case, they considered the following six functions:

$$f_1(x) = 3x(x-1)(x-1.9)(x-0.7)(x+1.8), \qquad -2 \le x \le 2, \tag{20}$$

$$f_2(x) = 10 \tan^{-1}\left[\frac{(x-0.2)(x-0.7)(x+0.8)}{x+1.4}\right], \qquad -1 \le x \le 1, \tag{21}$$

$$f_3(x) = \frac{100(x+.95)(x+.6)(x+.4)(x-.1)(x-.4)(x-.8)(x-.9)}{(x+1.7)(x-2)^2}, \qquad -1 \le x \le 1, \tag{22}$$

$$f_4(x) = 8\sin(10x^2 + 5x + 1), \qquad -1 \le x \le 1, \tag{23}$$

$$f_5(x) = 10 \tan^{-1}\left[\frac{(x-0.2)(x-0.7)(x+0.8)}{(x+1.4)(x-1.1)x+0.7}\right], \qquad -1 \le x \le 1, \tag{24}$$

$$f_6(x) = 10\Big(e^{-5|x|} + e^{-3|x-0.8|/10} + e^{-10|x+0.6|}\Big), \qquad -1 \le x \le 1. \tag{25}$$

They used a supervised gradient descent to tune all the parameters of the SAM in order to minimise the squared error $E(x) = \frac{1}{2}\big(f(x) - F(x)\big)^2$. Here $f(x)$ is an approximated function chosen from among $f_1, f_2, \ldots, f_6$ and $F(x)$ is the SAM. They compared popular membership functions such as triangles, trapezoids and Gaussian with sinc ($\operatorname{sinc}(x) = \sin x/x$) and concluded that the sinc function chosen as the if-part set function is the most likely to produce a quick and accurate function approximation.

In the first series of our numerical simulations, we considered the Sugeno fuzzy model of the zeroth order with 12 fuzzy rules:

$$f(x) = \sum_{i=1}^{12} \frac{A_i(x)}{\sum_{j=1}^{12} A_j(x)} y_i, \tag{26}$$

where the functions A_i were chosen as Beta, Gaussian and sinc functions. The approximated functions were the same as those chosen by Mitaim and Kosko. We used a modified version of the supervised gradient descent (Box *et al.*, 1989; Grace, 1994) to tune all the parameters of the fuzzy sets in order to minimise the relative error

$$E_r = \frac{1}{2}\left(\frac{f(x) - F(x)}{F(x)}\right)^2. \tag{27}$$

We computed the relative error between the approximated function and the output of the FLSs. The training as well as the testing sets were 200 randomly chosen points of the domain of definition for each function. We obtained the results given in Table 1. They clearly show that BFLSs are the best. This fact can be explained as follows: Beta functions depend on four parameters p, q, a, and b. Parameters a and b determine the support of the Beta function which can be translated, shrunk or dilated according to the values of a and b. Furthermore, p and q allow the Beta functions to have different shapes which are symmetric if $p = q$ and asymmetric if $p \neq q$. This flexibility in shape added to the fact that Beta functions are of compact support allows BFLSs to be very good function approximators.

Table 1. Relative approximation errors of functions f_1–f_6 by different FLSs.

function	BFLS	Gaussian FLS	sinc FLS
f_1	$5.5589 \times 10^{-5}\%$	$1.3229 \times 10^{-4}\%$	$4.6308 \times 10^{-2}\%$
f_2	$1.7984 \times 10^{-4}\%$	$1.8675 \times 10^{-4}\%$	$8.1124 \times 10^{-3}\%$
f_3	$1.3350 \times 10^{-4}\%$	$6.6626 \times 10^{-4}\%$	$2.9980 \times 10^{-4}\%$
f_4	$5.2980 \times 10^{-3}\%$	$2.6678 \times 10^{-1}\%$	$1.5692 \times 10^{-1}\%$
f_5	$2.8628 \times 10^{-4}\%$	$3.8963 \times 10^{-4}\%$	$1.0076 \times 10^{-2}\%$
f_6	$2.1288 \times 10^{-3}\%$	$1.2755 \times 10^{-2}\%$	$3.2081 \times 10^{-2}\%$

In the second series of our numerical simulations, five other functions were chosen:

$$\text{square}(x) = \text{sign}(\cos(x/3)), \quad -3\pi \le x \le 3\pi, \tag{28}$$

$$\text{parabola}(x) = x^2, \quad 0 \le x \le 3, \tag{29}$$

$$\text{sine}(x) = \sin(x), \quad -3\pi \le x \le 3\pi, \tag{30}$$

$$\text{beakedsine}(x) = \sin(x)\exp(-0.1x), \quad -3\pi \le x \le 3\pi, \tag{31}$$

$$\text{logarithm}(x) = \log(x), \quad 0.2 \le x \le 3. \tag{32}$$

For each approximated function we took a fixed number of fuzzy rules. We also tuned the parameters of Beta fuzzy sets and Gaussian fuzzy sets in order to minimise the relative error E_r. The training as well as the testing sets were 200 randomly chosen points of the domain of definition for each function. We got the results listed in Table 2 that confirm again the superiority of BFLSs.

Table 2. Relative approximation errors of different functions by BFLSs and Gaussian FLSs.

function	number of rules	BFLS	Gaussian FLS
square	3	$8.5122 \times 10^{-2}\%$	$3.6484 \times 10^{-1}\%$
parabola	3	$3.4133 \times 10^{-4}\%$	$3.4133 \times 10^{-4}\%$
sinus	6	$3.9851 \times 10^{-2}\%$	$4.0113 \times 10^{-2}\%$
beakedsine	6	$9.4938 \times 10^{-3}\%$	$9.4965 \times 10^{-3}\%$
logarithm	4	$5.3821 \times 10^{-3}\%$	$7.1865 \times 10^{-3}\%$

8. Conclusion

In this paper, we have suggested a new membership function family for the design of FLSs. The paper consists of two complementary parts: a theoretic part in which theoretic foundations of BFLSs are given, and a numeric one consisting of a numerical comparison between BFLSs and other common FLSs.

In the first part, we have proved that under certain minor hypotheses on the Beta membership functions, BFLSs satisfy the following properties:

- Universal approximation property: This property ensures that every continuous function on a compact set U can be uniformly approximated by a BFLS. Based on this property, an algorithm that can actually construct a BFLS approximating a given continuous function arbitrarily well is described. The number of fuzzy rules must be increased so as to get a better approximation.

- Best approximation property: Here, we are interested in finding a best approximator to a given function in the set $\mathcal{B}_N$ of BFLSs with N fuzzy rules. We have shown that if $\mathcal{B}_N$ is a linear subspace of $\big(\mathcal{C}(U), \|\cdot\|_\infty\big)$ of dimension N, then there is a best approximator to any continuous function in the set $\mathcal{B}_N$. Moreover, if we are looking for a best approximator to $f \in \big(L^2(U), \|\cdot\|_2\big)$, then this approximator is unique.

- Interpolation property: This property ensures that for every 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 can find a BFLS modelled by g that also satisfies $g(x_i) = y_i$ for every $i \in \{1, 2, \dots, N\}$.

In the last part of our paper, we performed two types of numerical simulations in order to confirm that BFLSs offer quick and accurate function approximation. In the first case, we compared the performances of BFLSs, Gaussian FLSs and sinc FLSs. Table 1 shows that the relative error E_r between the approximated function and the BFLS is smaller in all the cases studied. In fact, Beta functions depend on four parameters p, q, a, and b, which gives them a great flexibility. In the other series of our numerical simulations five other functions were chosen. For each approximated function we took a fixed number of fuzzy rules. The results of Table 2 confirm once again that BFLSs are the best.

Acknowledgment

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

References

Abe S. and Lan M.-S. (1995): *Fuzzy rules extraction directly from numerical data for function approximation.* — IEEE Trans. Syst. Man Cybern., Vol.25, pp.119–129.

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

Alimi M.A. (1997b): *Beta fuzzy basis functions for the design of universal robust neuro-fuzzy controllers.* — Proc. Sêminaire sur la commande robuste & ses applications: SCRA'97, Nabeul, Tunisia, pp.C1–C5.

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

Box M.J., Davies D. and Swann W.H. (1989): *Non-linear Optimization Techniques.* — New York: Oliver & Boyd.

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. Syst. 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.

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

Grace A. (1994): *Optimization Toolbox for Use with MATLAB.* — New York: MathWorks Inc.

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

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

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

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 syntheses with a fuzzy logic controller.* — Int. J. Man-Machine Stud., Vol.7, No.12, 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.

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

Maitaim S. and Kosko B. (1996): *What is the best shape for a fuzzy set in function approximation?.* — Proc. 5th IEEE Int. Conf. *Fuzzy Systems*, New Orleans, Fuzzy 96 IEEE, Vol.2, pp.1237–1243.

Nguyen H.T. and Kreinovich V. (1992): *On approximations of controls by fyzzy systems.* — Tech. Report # TR 92–93/302, LIFE Chair of Fuzzy Theory, Tokyo Institute of Technology.

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.

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.

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.

Wang L.-X. (1992): *Fuzzy systems are universal approximators.* — Proc. IEEE Int. Conf. *Fuzzy Systems*, San Diego, California, pp.1163–1169.

Wang L.-X. and Mendel J. (1991): *Generating fuzzy rules from numerical data with applications.* — Tech. Report TR USC–SIPI, # 169, Signal and Image Processing Institute, Univ. Southern California.

Wang L.-X. and Mendel J.M. (1992): *Fuzzy basis functions, universal approximation, 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. (1998a): *General SISO Takagi-Sugeno fuzzy systems with linear rule consequent are universal approximators.* — IEEE Trans. Fuzzy Syst., Vol.6, No.4, pp.582–587.

Ying H. (1998b): *The Takagi-Sugeno fuzzy controllers using the simplified linear control rules are nonlinear variable gain controllers.* — Automatica, Vol.34, pp.157–167.

Ying H. (1998c): *Constructing nonlinear variable gain controllers via the Takagi-Sugeno fuzzy control.* — IEEE Trans. Fuzzy Syst., Vol.6, pp.226–234.

Ying H. and Sheppard L.C. (1997): *Analytical structure of Takagi-Sugeno fuzzy PID controllers and their applications in control of mean arterial pressure in patients.* — Proc. 5th Europ. Congress *Intell. Tech. Soft Comput.*, Aachen, Germany, pp.1374–1378.

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

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.