ELSEVIER

Fuzzy Sets and Systems 130 (2002) 159–174

FUZZY sets and systems

www.elsevier.com/locate/fss

Approximation theory of fuzzy systems based upon genuine many-valued implications — MIMO cases ☆

Yong-Ming Li[a,b,*], Zhong-Ke Shi[b], Zhi-Hui Li[a]

[a] *Department of Mathematics, Shaanxi Normal University, Xi'an 710062, People's Republic of China*
[b] *Department of Automatic Control, Northwestern Polytechnical University, Xi'an 710072, People's Republic of China*

Received 30 December 1999; received in revised form 20 March 2001; accepted 3 April 2001

Abstract

It is constructively proved that the multi-input–multi-output fuzzy systems based upon genuine many-valued implications are universal approximators (they are called Boolean type fuzzy systems in this paper). The general approach to construct such fuzzy systems is given, that is, through the partition of the output region (by the given accuracy). Two examples are provided to demonstrate the way in which fuzzy systems are designed to approximate given functions with a given required approximation accuracy. © 2002 Elsevier Science B.V. All rights reserved.

Keywords: Fuzzy system; Genuine many-valued implication; Universal approximator

1. Introduction

In recent years, there have been a number of applications of fuzzy systems theory in various fields, for example, in control systems. In most of these applications, the main design objective is to construct a fuzzy system to approximate a desired control or decision (often experts). From a mathematical point of view, fuzzy systems are just functions mapping their input to output. In the context of control, the question is whether a fuzzy controller can always be constructed to approximate any desired continuous and nonlinear control solution with enough accuracy. For fuzzy systems used as models, the issue is whether a fuzzy model can always be established which is capable of approximating any continuous and nonlinear physical system arbitrarily well. The questions are of both theoretical and practical importance. If the fuzzy systems were proved to be universal approximators, then one would feel more comfortable to utilize them as controllers and models. If not, the fuzzy systems should be used to solve only those control and modeling problems that they are capable of.

Due to its importance, the issue of fuzzy systems as universal approximators has drawn significant attention in the past few years and progress has been made [1,5,6,11,13,15,17,19–24]. Consider a fuzzy system that comprises four principle components: fuzzifier, fuzzy rule base, fuzzy inference engine, and

☆ This work is supported by National Science Foundation of China (Grant No. 19901028) and National 973 Foundational Research Program G1998030417.

* Corresponding author. Fax: +86-29-530-7025.
E-mail address: liyongm@snnu.edu.cn (Y.-M. Li).

0165-0114/02/$ - see front matter
PII: S0165-0114(01)00172-5

defuzzifier. Assume that the fuzzifier is the most commonly used singleton fuzzifier, defuzzifier is the center of areas (COA) or averaging of maximums (MOMs), and the fuzzy rule base for an multi-input–single-output (MISO) system consists of rules in the following forms:

$$\begin{aligned}\text{(MISO)}\ \ R_k:\ &\text{IF } x_1 \textit{ is } A_k^1 \text{ AND } x_2 \textit{ is } A_k^2 \text{ AND } \cdots \\ &\text{AND } x_m \textit{ is } A_k^m, \\ &\text{THEN } y \textit{ is } B_k \quad (k=1,2,\ldots,n),\end{aligned} \tag{1}$$

where $A_k^j(B_k)$ in U_j (V) about variable x_k (y) are linguistic terms characterized by fuzzy membership function $A_k^j(x^j)$ $(B_k(y))$. We generally consider normal membership functions, such as triangular, trapezoidal or Gaussian functions. Each R_k can be viewed as a fuzzy implication (relation) $A_k=A_k^1\times A_k^2\times\cdots\times A_k^m\to B_k$ over $U\times V=(U_1\times U_2\times\cdots\times U_m)\times V$. The choice of fuzzy inference is flexible. This flexibility is determined by the implication operators chosen. For the previous approximation work in [1,5,6,11,13,15,17,19–24], mainly Mamdani type fuzzy systems and Takagi–Sugeno (TS) type fuzzy systems, the implication operator $\to$ is always chosen as conjunctive type implication, that is, choosing $\to$ as a t-norm, such as min ($\wedge$) and product operator. We call the corresponding fuzzy systems the conjunctive type fuzzy systems as in [6]. For the choice of implication operators as genuine many-valued implications (i.e. the generalization of that of classic implications, mainly contain S-implications, R-implications and QL-implications [7–9]), at present, there do not exist any approximation results in the literature for such fuzzy systems. Although J.L. Castro has done some work in this respect as genuine many-valued implications chosen as R-implications (with some restrictions), we shall see later that his inference approach corresponding to R-implications is not appropriate. However, as the results of the experiment show [2–4,10], the Boolean type fuzzy systems (we call the fuzzy system based on genuine many-valued implications the Boolean type fuzzy system as in [7]) have good control capability compared to other implication operators such as conjunctive type implication operators. Furthermore, since the fuzzy logic based on genuine many-valued implications is a generalization of classic logic, it has been studied as multi-valued logic systems [8,9] and thus has a strict logic foundation [18], it also has a widespread use in expert systems [2–4,10]. The question "*Are they also universal approximators and how is their approximation mechanisms*" still remains unanswered. However, we have answered this question for SISO Boolean type fuzzy systems in [14]. In this paper, these results will be extended to multi-input–multi-output (MIMO) fuzzy systems based on genuine many-valued implications.

The structure of the rest of the paper is as follows. In Section 2, we review the inference methods of multi-rules fuzzy systems for MIMO fuzzy systems. We discuss the approximation properties of MIMO fuzzy systems based on R-implications in Section 3 and those of MIMO fuzzy systems based on S- and QL-implications in Section 4. In Section 5, we shall illustratively design two fuzzy systems based on genuine many-valued implication for two given functions with a required approximation accuracy and compare the optimal fuzzy rules to those conjunctive fuzzy systems. Conclusions are made in the last section. One appendix for the proofs of the propositions in the paper is included.

2. The choice of multi-rules fuzzy inference methods

We first review the (multi-rules) fuzzy inference method presented in the previous paper [14]. Since the MIMO fuzzy systems can always be separated into a group of MISO fuzzy systems [12], without loss of generality, we assume in this paper that fuzzy systems are MISO systems $f: U\subseteq R^m \to V\subseteq R$ as form (1), where $U=U_1\times U_2\times\cdots\times U_m\subseteq R^m$ is the input space and $V\subseteq R$ is the output space. The MIMO versions of all results in this paper can be easily obtained by doing a few simple manipulations.

For the rule R_k, its antecedent fuzzy set $A_k=A_k^1\times A_k^2\times\cdots\times A_k^m$ on $U=U_1\times U_2\times\cdots\times U_m$ with membership function $A_k(x)=A_k^1(x_1)*A_k^2(x_2)*\cdots*A_k^m(x_m)$, where $*$ is the T-norm [8,9], $x=(x_1,x_2,\ldots,x_m)\in U$. Then each R_k can be viewed as fuzzy implication relation $A_k=A_k^1\times A_k^2\times\cdots\times A_k^m\to B_k$, which is a fuzzy set in $U\times V$ with a membership function $R_k(x,y)=A_k(x)\to B_k(y)$.

For fuzzy system (1) and a given input A', inference approaches are presented as follows:

$$B' = A' \circ \bigcup_{k=1}^{m} (A_k \to B_k), \tag{2}$$

$$B' = \bigcup_{k=1}^{m} A' \circ (A_k \to B_k), \tag{3}$$

$$B' = A' \circ \bigcap_{k=1}^{m} (A_k \to B_k), \tag{4}$$

$$B' = \bigcap_{k=1}^{m} A' \circ (A_k \to B_k), \tag{5}$$

where $\circ$ denotes the relation composition operator, it is the just generalized Zadeh's max-$*$ composition rules, where $*$ is a certain t-norm operator.

Assume that the input is a singleton $x = x_0$, then in the above algorithm, (2) and (3) are equivalent, and (4) and (5) are equivalent.

Assume that the requirement of the membership function $C_k(x)$ $(x \in X \subseteq R)$, for each input variable and for each output variable in (1), is the same as that of paper [23]. We assume that $\{C_k\}$ is continuous, normal, consistent and complete, that is, for any k, there exists $x \in X$ such that $C_k(x) = 1$; and for any $x \in X$, there exists k such that $C_k(x) > 0$; and if $C_k(x) = 1$, then $C_j(x) = 0$ for any $j \neq k$. In particular, if $\sum C_k(x) = 1$ holds for any $x \in X$, then $\{C_k\}$ is called a normal base set. For the background of this requirement, we refer to [23]. We assume that $C_k(x)$ is a pseudo-trapezoid-shaped membership function for any k, that is, $C_k(x)$ has the following forms:

$$C_k(x, x_k, b_k, c_k, x_{k+1}) = \begin{cases} M(x), & x \in [x_k, b_k), \\ 1, & x \in [b_k, c_k), \\ D(x), & x \in [c_k, x_{k+1}], \\ 0, & x \in U - [x_k, x_{k+1}], \end{cases} \tag{6}$$

where $x_k \leqslant b_k \leqslant c_k \leqslant x_{k+1}$, $x_k < x_{k+1}$, $M(x) \geqslant 0$ is strictly increasing in $[x_k, b_k)$ and $D(x) \geqslant 0$ is strictly decreasing in $[c_k, x_{k+1}]$. In particular, if $b_k = c_k$, then $C_k(x)$ is pseudo-triangle-shaped membership function. We assume that the membership function is pseudo-triangle-shaped in the following, the discussion for that of pseudo-trapezoid-shaped membership function is similar.

We have the following general result.

Lemma 2.1 (Li et al. [14]). *If the implication operator is chosen as a t-norm, inference algorithm is assumed as* (4), *then* $B' \equiv 0$.

Hence, if we use inference method (4) and (5) for conjunctive type implication operator, then the output is always 0 for any input. In this case, the control is impossible for any processes; the inference method (2) and (3) should be used. This is also the theoretical explanation of the related experiment results in [3,4,10].

On the other hand, if $\to$ is chosen as genuine many-valued implication, then $I(0,x) \equiv 1$ holds for any $x \in [0,1]$, in this case, we should use inference method (4) and (5) instead of (2) and (3), some reason is presented in the following lemma.

Lemma 2.2 (Li et al. [14] and Mántaras [16]). *If the implication operator is chosen as a genuine many-valued implication, inference method is assumed as* (2), *then* $B' \equiv 1$.

Remark 2.3. From the formal logic and matching points of view, we can also give some explanation as follows:

From the view of semantics of formal logic, fuzzy inference based on genuine many-valued implications assume that, if the truth-values of the antecedents of the rules are false, then the truth-values of the rules are true, and truth-values of the rules are non-increasing about the antecedents of the rules. However, in the actions of control, or in the common-sense inference, we always assume that, if the truth-values of the antecedents of the rules are false or nearly false, then the corresponding rules shall not be fired or have little effect on the final control. In this case, it is reasonable to use the inference form (5) instead of (3) so as to diminish the effect of those rules with small matching degree of the antecedents.

On the other hand, from the matching point of view, fuzzy inference based on conjunctive implications assumes that, the bigger the matching degrees (or the truth-values) of the antecedents of the rules, the larger is the contribution of the rules to the final control. So, if the matching degrees of the antecedents of the rules are zeros (false), then the corresponding rules shall not be fired or have little affect on the final control action. In this case, it is reasonable to use the inference form (3) instead of (5).

Remark 2.4. From Remark 2.3, we give some notes about the constructions of the fuzzy systems based on R-implications by Castro in [5].

For R-implications $I: [0,1]\times[0,1]\to[0,1]$ satisfying the condition $I(a,0)=0$ if $a\neq 0$, given input $x=x_0$, then Castro assumed that the output of the ith rules was

$$B_i'(y)=\begin{cases}0, & A_i(x_0)=0,\\ I(A_i(x_0),B_i(y)), & A_i(x_0)\neq 0\end{cases} \tag{7}$$

and the final fuzzy output was

$$B'(y)=\bigvee_i B_i(y). \tag{8}$$

In fact, Castro used the inference method (3) instead of (5). As mentioned in the above remark, this is only reasonable for the case I being a conjunctive implication, so the choice of inference approach (8) using R-implications is not appropriate, even though there has been some restriction on each consequent part of fuzzy rules, as (7). From the matching point of view, the required principle: for input $x=x_0$ is the smaller the value (not necessary 0) $A_i(x_0)$, the smaller will be the consequent value $B_i'(y)$ of the corresponding rule R_i. However, choice (7) as a consequent value $B_i'(y)$ of the corresponding rule R_i is contrary to the required principle. In fact, since I is an R-implication, $I(a,b)$ is nonincreasing about the first variable a, given an input $x=x_0$, the smaller the value (not necessary 0) of the antecedent $A_i(x_0)$, the larger is the value of the consequent $B_i'(y)$ (if $B_i'(y)\neq 0$) of the corresponding rule R_i. This forces $B'(y)$, which is the supremum of all $B_i'(y)$, to be large (enough) at y as long as $B_i'(y)\neq 0$, even though $A_i(x_0)$ is extremely small. This points out the inference approach (7) and (8) used by Castro in [5], for R-implications is not appropriate.

3. Approximation capability of MISO fuzzy systems based on R-implication

Suppose that I is an R-implication in this section, T is the corresponding t-norm. Then $I:[0,1]\times[0,1]\to[0,1]$ is defined as follows:

$$I(a,b)=\sup\{c\in[0,1]\mid T(a,c)\leqslant b\}. \tag{9}$$

The defuzzification method is always assumed as the averaging of MOA.

For MISO fuzzy system (1), let $x=x_0=(x_0^1,\ldots,x_0^m)$ be a singleton input, then the fuzzy output based on I is

$$B'(y)=\bigwedge_{k=1}^{n}(A_k(x_0)\to B_k(y)). \tag{10}$$

Suppose that only $A_{k(1)}(x_0),\ldots,A_{k(l)}(x_0)$ are not zero in the following, then Eq. (10) can be simply rewritten as follows:

$$B'(y)=\bigwedge_{i=1}^{l}(A_{k(i)}(x_0)\to B_{k(i)}(y)). \tag{11}$$

Since we require $\{B_i\}$ to be consistent, for any $y\in V$, there are at most two adjacent elements B_k, B_{k+1} such that $B_k(y)>0$, $B_{k+1}(y)>0$. In this case, Eq. (11) can be calculated as follows:

$$B'(y)=\bigwedge_{i\in T_1}(A_{k(i)}(x_0)\to B_k(y))\wedge\bigwedge_{i\in T_2}(A_{k(i)}(x_0)\to B_{k+1}(y))\bigwedge_{i\in T_3}(A_{k(i)}(x_0)\to 0). \tag{12}$$

For input $x=x_0$, there is one determined output $y=y_0$, for system (1). For this output $y=y_0$, it is reasonable to require that the value $B'(y_0)$ is mainly determined by the first two terms of Eq. (12), that is,

$$B'(y)=\bigwedge_{i\in T_1}(A_{k(i)}(x_0)\to B_k(y))\wedge\bigwedge_{i\in T_2}(A_{k(i)}(x_0)\to B_{k+1}(y)). \tag{13}$$

If we have the following assumption on fuzzy system (1), then Eqs. (12) and (13) are equivalent.

Assumption. For any singleton input $x=x_0$, if the rules $R_{k(1)},\ldots,R_{k(l)}$ are fired, that is, the membership values $A_{k(1)}(x_0),\ldots,A_{k(l)}(x_0)$ of the antecedents of these rules are not zero, then the membership functions $B_{k(1)}(y),\dots,B_{k(l)}(y)$ *of the consequent part of these rules have at most two adjacent different elements* $B_k(y), B_{k+1}(y)$.

In fact, for some R-implications, this assumption is necessary.

Lemma 3.1. *If the membership functions* $\{B_i\}$ *are required to be consistent, and the implication operator* I *satisfies the condition* $I(a,0)\neq 0$ *whenever* $a\neq 0$, *then for the rule base* (1), *if we use inference method* (10) *as above, and the consequent part* $B_{k(1)}(y),\dots,B_{k(l)}(y)$ *have at least* 3 *elements, then* $B'(y)\equiv 0$.

In fact, the above assumption is also reasonable, because many rules based on the practical fuzzy control have the forms, and we can indeed construct such kinds of fuzzy systems (in Theorem 3.9) below. This requirement justifies the following principle: *If the inputs or antecedent parts intersect, then the outputs or consequent parts must intersect.*

In this paper, we always make this assumption on fuzzy systems. In this case, Eq. (13) always holds. For a Boolean implication $\to$, since $a\to b$ is non-increasing to first variable a and T_1, T_2 are finite sets of $\{1,2,\dots,l\}$ in Eq. (13), then there exists $k_1\in T_1$ such that $\bigwedge_{i\in T_1}(A_{k(i)}(x_0)\to B_k(y)) = (\bigvee_{i\in T_1} A_{k(i)}(x_0))\to B_k(y) = A_{k_1}(x_0)\to B_k(y)$ and $k_2\in T_2$ such that $\bigwedge_{i\in T_2}(A_{k(i)}(x_0)\to B_{k+1}(y)) = A_{k_2}(x_0)\to B_{k+1}(y)$. Without loss of generality, we can assume that $k_1=k$, $k_2=k+1$, then Eq. (13) can be simply rewritten as follows:

$$B'(y) = (A_k(x_0)\to B_k(y)) \wedge (A_{k+1}(x_0)\to B_{k+1}(y)). \tag{14}$$

Let $Y_0=\{y_0 \mid B'(y_0)=\max_{y\in V} B'(y)\}$, then the control output y_0 is one chosen point of Y_0 (such as the middle point of Y_0, but we do not make this restriction). We thus obtain a function $y_0=G(x_0)$, $x_0\in U$; it is called the system function of (1) in this paper.

We first discuss the form of $G(x)$ in the following propositions.

Lemma 3.2. *For any continuous t-norm* T, *the following equation has one solution for any* $a,b\in(0,1]$

$$T(a,z) = T(b,1-z). \tag{15}$$

Remark 3.3. (1) If $f(z)=T(a,z): [0,1]\to[0,a]$ is strictly the increasing mapping for any $a\in(0,1]$, then the solution of Eq. (15) is unique.

(2) As for some examples, we give the solution of (15) for some t-norms:

(i) $T=\wedge$, *min* operator: If $a\leqslant b$, then the solution of Eq. (15) is

$$z=\begin{cases} 1/2, & a>1/2,\\ 1-a, & a\leqslant 1/2\end{cases}$$

if $a>b$, then the solution of (15) is

$$z=\begin{cases} 1/2, & b>1/2,\\ 1-b, & b\leqslant 1/2\end{cases}$$

(ii) $T=\bullet$, *product* operator: $z=b/(a+b)$.

(iii) $T(a,b)=max\{0,a+b-1\}$, *bounded sum*: $z=(1+|b-a|)/2$.

Corollary 3.4. *For* $a,b\in(0,1]$, *if* $a+b\leqslant 1$, *then Eq.* (15) *has one solution* z *satisfying* $a\leqslant 1-z$ *and* $b\leqslant z$.

Theorem 3.5. *Suppose that* T *is a continuous t-norm,* I *is the corresponding R-implication, then the system function of MISO fuzzy systems* (1) *based on* I *is defined as follows: for any input* $x=x_0$, *there are two rules* R_k, R_{k+1} *such that Eq.* (14) *holds, then*

$$y_0 = G(x_0) = B_k^{-1}(z), \tag{16}$$

where z *is just the solution of Eq.* (15) *for* $a=A_{k+1}(x_0)$, $b=A_k(x_0)$ *and* $y_0=B_k^{-1}(z)$ *satisfies the condition* $B_k(y_0)+B_{k+1}(y_0)=1$.

Corollary 3.6. *If* I *is the Goguen implication, that is,* $I(a,b)=1\wedge b/a$, *then the system function of* (1) *based on* I *has the following form*:

$$y_0 = G(x_0) = B_k^{-1}\left(\frac{A_k(x_0)}{A_k(x_0)+A_{k+1}(x_0)}\right).$$

This is from Theorem 3.5 and Remark 3.3(2)(ii).

Corollary 3.7. *If I is Gŏdel-implication, then the system function of* (1) *based on I is as follows*:

$$y_0 = G(x_0) = \begin{cases} B_k^{-1}(1/2), & \min\{A_k(x_0), A_{k+1}(x_0)\} > 1/2, \\ B_k^{-1}(1 - A_{k+1}(x_0)), & A_{k+1}(x_0) \leqslant 1/2 \text{ and } A_{k+1}(x_0) \leqslant A_k(x_0), \\ B_k^{-1}(1 - A_k(x_0)), & A_k(x_0) \leqslant 1/2 \text{ and } A_{k+1}(x_0) \geqslant A_k(x_0). \end{cases}$$

This is from Theorem 3.5 and Remark 3.3(2)(i).

Corollary 3.8. *If I is Lukasitwicz implication, that is*, $I(a,b) = (1 - a + b) \wedge 1$, *then the system function of* (1) *has the following form*:

$$y_0 = G(x_0) = B_k^{-1}(1/2 + |A_k(x_0) - A_{k+1}(x_0)|/2).$$

This is from Theorem 3.5 and Remark 3.3(2)(iii).

For fuzzy systems based on R-implication operators, the results in [3,4,10] showed their good control capability compared to other implication operators such as conjunctive type implication operators. Then, whether or not the fuzzy systems based upon these implications are universal approximators, the following theorem answers this problem.

Theorem 3.9. *The fuzzy systems based on R-implications with MOA defuzzifier are universal approximators. That is, for any continuous function* $f: U \to R$ *over a compact subset* $U \subseteq R^m$ *and an arbitrary given positive number* ε, *there is a fuzzy system, its corresponding system function* $y = G(x)$ (*as given in Theorem* 3.5) *based on a given R-implication with MOA defuzzifier satisfies the inequality relation*

$$\max_{x \in U} |f(x) - G(x)| < \varepsilon.$$

Proof. For simplicity and better presentation, we will prove the case of two variables (i.e., $m = 2$). The proof for more variables is similar.

Since f is continuous, the image of f is also a closed interval assumed as $[c, d]$. For the given ε, there exists a natural number N such that $(d - c)/N < \varepsilon/2$. Let $e = (d - c)/N$, suppose that $y_1 = c, y_2 = c + e, \dots, y_{k+1} = c + ke, \dots, y_{N+1} = d$. Constructing membership functions $B_1, B_2, \dots, B_{N+1}$ such that the support of B_1 is $[y_1, y_2)$, the support of B_{N+1} is $(y_N, y_{N+1}]$, the support of B_k is (y_{k-1}, y_{k+1}) for $1 < k < N + 1$, and $\{B_k\}_{k=1}^{N+1}$ is a normal base set (such as the symmetric triangle-shaped membership functions), then $B_k(y_k) = 1$ for any $k = 1, 2, \dots, N+1$. Let

$$U_1 = f^{-1}([y_1, y_1 + \tfrac{2}{3}e)), \dots, U_k = f^{-1}((y_k - \tfrac{2}{3}e, y_k + \tfrac{2}{3}e)), \dots,$$
$$U_{N+1} = f^{-1}((y_{N+1} - \tfrac{2}{3}e, y_{N+1}])$$

then $\{U_1, U_2, \dots, U_{N+1}\}$ forms an open cover of U and $U_{i_1} \cap U_{i_2} \neq \emptyset$ if and only if i_1, i_2 are the adjacent numbers.

For the given e, since f is uniform continuous over compact set U, there exists a positive number δ, such that $|f(x) - f(x')| < e/2$ whenever $d(x, x') < \delta$, where $d(x, x') = \max\{|x_1 - x_1'|, |y_1 - y_1'|\}$, $x = (x_1, y_1)$, $x' = (x_1', y_1')$. Therefore, there exists a finite set $F = \{x^1 = (x_1^1, x_2^1), \dots, x^s = (x_1^s, x_2^s)\}$ such that the $\delta/2$-neighborhood of F forms another open cover of U, that is, $\{(x_1^i - \delta/2, x_1^i + \delta/2) \times (x_2^i - \delta/2, x_2^i + \delta/2) \mid i = 1, \dots, s\}$ forms an open cover of U, and the following appropriate condition holds.

Appropriate condition. *Write the partition points of* $[a_1, b_1]$ *and* $[a_2, b_2]$ *formed by F by* $x_1^1 < \cdots < x_l^1$, $x_1^2 < \cdots < x_d^2$ (*then* $|x_i^1 - x_{i+1}^1| < \delta$ *and* $|x_i^2 - x_{i+1}^2| < \delta$), *then for any open rectangle* $(x_i^1, x_{i+2}^1) \times (x_j^2, x_{j+2}^2)$, *it meets* $\{U_k\}$ *at most two adjacent elements.*

Constructing the normal base set $\{A_i^1\}$ *for* $[a_1, b_1]$ and $\{A_i^2\}$ for $[a_2, b_2]$ such that the support of A_k^1 is (x_{k-1}^1, x_{k+1}^1) and the support of A_k^2 is (x_{k-1}^2, x_{k+1}^2), then $A_k^1(x_k^1) = 1$ and $A_k^2(x_k^2) = 1$ for any k. Fuzzy rules base is designed as follows:

R_{ij}: IF x_1 *is* A_i^1 AND x_2 *is* A_j^2,

THEN y $C_{ij}.(i = 1, 2, \dots, l; j = 1, 2, \dots, d))$,

where C_{ij} is chosen as there exists k such that $x_i \in U_k$ (there are at most two U_k such that $x_i \in U_k$ as the construction of U_k), then C_{ij} is chosen as B_k.

As the appropriate condition and the properties of the open cover $\{U_k\}$, the assumption of this paper for the rule base holds. This is because, for any singleton input $x = x_0$, if the rules $R_{k(1)}, \ldots, R_{k(l)}$ are fired, that is, the membership values $A_{k(1)}(x_0), \ldots, A_{k(l)}(x_0)$ of the antecedents of these rules are not zero, then the membership functions $B_{k(1)}(y), \ldots, B_{k(l)}(y)$ of the consequent part of these rules have at most two adjacent different elements B_k, B_{k+1}.

For any input $x_0 \in [a_1, b_1] \times [a_2, b_2]$, if $x_0 \in [x_i^1, x_{i+1}^1] \times [x_j^2, x_{j+1}^2]$, then

$$B'(y) = A_i(x_0) \to C_i(y) \wedge A_{i+1}(x_0) \to C_{i+1}(y).$$

Suppose that $C_i = B_k$, then $C_{i+1} = B_{k+1}$, B_k or B_{k-1} as the appropriate condition. In this case, $y_0 = C_i^{-1}(A_i(x_0))$ (as in Theorem 3.5) is in $[y_{k-1}, y_k]$ or $[y_k, y_{k+1}]$. It is no hurt to assume that $y_0 \in [y_k, y_{k+1}]$, then $\max\{|y_0 - y_k|, |y_0 - y_{k+1}|\} \leqslant e$. Noting that x_0 belongs to at most two adjacent open sets U_k or U_{k+1}, it follows that $|f(x_0) - f(y_k)| < \frac{2}{3}e$ or $|f(x_0) - f(y_{k+1})| < \frac{2}{3}e$, and thus

$$\begin{aligned}
&|G(x_0) - f(x_0)| \\
&\quad = |y_0 - f(x_0)| \\
&\quad \leqslant \min\{|y_0 - y_k| + |y_k - f(x_0)|, |y_0 - y_{k+1}| \\
&\qquad + |y_{k+1} - f(x_0)|\} \\
&\quad \leqslant e + \tfrac{2}{3}e = \tfrac{5}{3}e < \tfrac{5}{6}\varepsilon < \varepsilon.
\end{aligned}$$

Hence, $\max_{x \in U}\{|G(x) - f(x)|\} < \varepsilon$.

Furthermore, if all the consequent membership functions $\{B_i\}$ are assumed as symmetric triangle-shaped, then the system function $y = G(x)$ defined by (16) has the simple form that we calculated in the following. As an example, only two inputs fuzzy systems are considered. R-implication is chosen as Goguen-implication, t-norm is just a product operation.

In this case, for any singleton input $x_0 = (x_0^1, x_0^2)$, there are at most four rules fired, and Eq. (10) has the following form:

$$\begin{aligned}
B'(y) = {} & A_i^1(x_0^1) A_j^2(x_0^2) \to B_{ij}(y) \\
& \wedge A_i^1(x_0^1) A_{j+1}^2(x_0^2) \to B_{i,j+1}(y) \\
& \wedge A_{i+1}^1(x_0^1) A_j^2(x_0^2) \to B_{i+1,j}(y) \\
& \wedge A_{i+1}^1(x_0^1) A_{j+1}^2(x_0^2) \to B_{i+1,j+1}(y).
\end{aligned}$$

As the assumption for the fuzzy rule base in this paper, there are three cases for consequent membership functions:

(1) $B_{i,j}(y) = B_{i+1,j}(y) = B_{i,j+1}(y) = B_{i+1,j+1}(y) = B_k$. In this case, since B_k is symmetrically triangle-shaped and $B_k(y_k) = 1$, then there exists non-negative real number δ such that for any $y \in [y_k - \delta, y_k + \delta]$, $B'(y)$ attains its maximum value 1. It follows that

$$y_0 = G(x_0) = y_k. \tag{17a}$$

(2) $B_{i,j}(y) = B_{i+1,j}(y) = B_{i,j+1}(y) = B_k, B_{i+1,j+1}(y) = B_{k+1}$. Let $a = A_i^1(x_0^1)$, $b = A_j^2(x_0^2)$, as B_k, B_{k+1} are triangle-shaped, it follows that

$$\begin{aligned}
B'(y) = {} & \frac{y - y_{k+1}}{ab(y_k - y_{k+1})} \wedge \frac{y - y_{k+1}}{a(1-b)(y_k - y_{k+1})} \\
& \wedge \frac{y - y_{k+1}}{(1-a)b(y_k - y_{k+1})} \\
& \wedge \frac{y - y_k}{(1-a)(1-b)(y_{k+1} - y_k)} \wedge 1 \\
= {} & \frac{y - y_{k+1}}{\max\{ab, a(1-b), b(1-a)\}(y_k - y_{k+1})} \\
& \wedge \frac{y - y_k}{(1-a)(1-b)(y_{k+1} - y_k)} \wedge 1.
\end{aligned}$$

There are three cases

(2a) $b \geqslant a$ and $a \leqslant 1/2$. In this case,

$$\begin{aligned}
B'(y) = {} & \frac{y - y_{k+1}}{(1-a)b(y_k - y_{k+1})} \\
& \wedge \frac{y - y_k}{(1-a)(1-b)(y_{k+1} - y_k)} \wedge 1
\end{aligned}$$

and $y_0 = G(x_0)$ satisfies the following equation:

$$\begin{aligned}
& \frac{y_0 - y_{k+1}}{(1-a)b(y_k - y_{k+1})} \\
& \quad = \frac{y_0 - y_k}{(1-a)(1-b)(y_{k+1} - y_k)} \\
& \quad \Rightarrow (1-b)y_0 - (1-b)y_{k+1} = -by_0 + by_k \\
& \quad \Rightarrow y_0 = by_k + (1-b)y_{k+1}.
\end{aligned} \tag{17b}$$

In this case, the four fired rules $R_{i,j}, R_{i+1,j}, R_{i,j+1}, R_{i+1,j+1}$ reduce to two simple rules

IF x_2 *is* A_j^2 THEN y *is* B_k,

IF x_2 *is* A_{j+1}^2 THEN y *is* B_{k+1}.

(2b) $b \geqslant 0.5$ and $a \geqslant 0.5$. In this case,

$$B'(y) = \frac{y - y_{k+1}}{ab(y_k - y_{k+1})} \wedge \frac{y - y_k}{(1-a)(1-b)(y_{k+1} - y_k)} \wedge 1$$

and $y_0 = G(x_0)$ satisfies the following equation:

$$\frac{y_0 - y_{k+1}}{ab(y_k - y_{k+1})} = \frac{y_0 - y_k}{(1-a)(1-b)(y_{k+1} - y_k)} \Rightarrow (1-a)(1-b)(y_0 - y_{k+1}) = -ab(y_0 - y_k) \Rightarrow y_0 = \frac{1-a-b+ab}{1-a-b+2ab} y_k + \frac{ab}{1-a-b+2ab} y_{k+1}. \tag{17c}$$

(2c) $a \geqslant b$ and $a \leqslant 0.5$. In this case, as in case (2a), we can obtain that

$$y_0 = a y_k + (1-a) y_{k+1}. \tag{17d}$$

The four fired rules were reduced to two simple rules as in case (2b).

Case (2) can be seen as three of the four fired consequent membership functions that are same, while the other is different.

(3) $B_{i,j}(y) = B_{i+1,j}(y) = B_k, B_{i,j+1}(y) = B_{i+1,j+1}(y) = B_{k+1}$, two of four fired consequent membership functions are the same and the other two are the other similar ones. In this case,

$$B'(y) = \frac{y - y_{k+1}}{\max\{ab, a(1-b)\}(y_k - y_{k+1})} \wedge \frac{y - y_k}{\max\{b(1-a), (1-a)(1-b)\}(y_{k+1} - y_k)} \wedge 1.$$

There are two cases:

(3a) $b \geqslant 0.5$. In this case, $y_0 = G(x_0)$ satisfies the following equation:

$$\frac{y_0 - y_{k+1}}{ab(y_k - y_{k+1})} = \frac{y_0 - y_k}{b(1-a)(y_{k+1} - y_k)} \Rightarrow (1-a)(y_0 - y_{k+1}) = -a(y_0 - y_k) \Rightarrow y_0 = a y_k + (1-a) y_{k+1}. \tag{17e}$$

(3b) $b \leqslant 0.5$. In this case, $y_0 = G(x_0)$ satisfies the following equation:

$$\frac{y_0 - y_{k+1}}{a(1-b)(y_k - y_{k+1})} = \frac{y_0 - y_k}{(1-a)(1-b)(y_{k+1} - y_k)} \Rightarrow (1-a)(y_0 - y_{k+1}) = -a(y_0 - y_k) \Rightarrow y_0 = a y_k + (1-a) y_{k+1}. \tag{17f}$$

As in cases (2b) and (2c), the four fired rules are reduced to two simple rules in case (3).

The combination of formulas (17a)–(17f) is just the analytic representation of systems function $y = G(x_1, x_2)$ for this kind of fuzzy system. Example 2 in Section 5 illustrates the formula.

In fact, the above discussion to two-input fuzzy systems (1) reflects certain decoupling properties of fuzzy systems based on R-implications, and this kind of properties should be researched in the future.

Remark 3.10. The results of Theorem 3.9 show the approximation capability of MISO Fuzzy systems based on R-implications, and the proof of Theorem 3.9 gives an approach to explicitly construct such a fuzzy system, in particular, its consequent part is constructive. It is not like other approaches [1,5,6,11,13,15,17,19–24], where their membership functions in the consequent part cover each other not satisfying the consistent conditions or consequent part that is simply chosen as a real number. We actually provide an approach to construct normal, complete and consistent membership functions in the consequent part.

The approximation mechanism of MISO fuzzy systems based on R-implications is similar to that of conjunctive type implications, that is, the nearer the distances of the input $x = x_0$ and the antecedent membership A_i, the nearer the distances of the output

$y_0 = B_k^{-1}(z)$ and the corresponding consequent membership B_i will be, and the output y_0 is only related with the 2^m adjacent base elements of the input $x = x_0$.

4. Approximation capability of MISO fuzzy systems based on S-implications and QL-implications

S-implication and QL-implication are defined, respectively as follows:

$$I(a,b) = S(N(a),b),\ I(a,b) = S(N(a),T(a,b)), \tag{18}$$

where S is a t-conorm, T is a t-norm, N is a negation, S and T are dual through N [8–10]. They both generalize classical genuine many-valued logic and satisfy the following condition

$$I(0,b) = 1. \tag{19}$$

They generally do not satisfy the condition $I(a,b) = 1$ whenever $a \leqslant b$, so the discussion of Section 3 does not hold for S- and QL-implications. We shall give other analyses in the following.

The requirements of membership functions are the same as those of Theorem 3.9. In this case, inference formula (14) also holds, that is,

$$B'(y) = S(N(A_i(x_0)), B_i(y)) \wedge S(N(A_{i+1}(x_0)), B_{i+1}(y)), \tag{20}$$

$$B'(y) = S(N(A_i(x_0)), T(A_i(x_0), B_i(y))) \wedge S(N(A_{i+1}(x_0)), T(A_i(x_0), B_{i+1}(y))). \tag{21}$$

The defuzzifier is also assumed as MOA.

Lemma 4.1. *Let $I(a,b) = (1-a) \vee b$ be Kleene–Dienes implication, $I(a,b) = (1-a) \vee (a \wedge b)$ be Zadeh implication, then the system function corresponding to fuzzy system* (1) *is*

$$y = G(x) = \begin{cases} y_i, & A_i(x_0) > A_{i+1}(x_0), \\ y_{i+1}, & A_i(x_0) < A_{i+1}(x_0), \\ (y_i + y_{i+1})/2, & A_i(x_0) = A_{i+1}(x_0), \end{cases} \tag{22}$$

which lies in the interval $[y_i, y_{i+1}]$, the intersection supports of B_i and B_{i+1}. Generally, $G(x)$ is not continuous.

Lemma 4.2. *For S-implication $I(a,b) = 1 - a + ab$, Reichenbach implication, the system function of fuzzy system* (1) *is*

$$y = B_i^{-1}\left(\frac{A_i(x_0)}{A_i(x_0) + A_{i+1}(x_0)}\right), \tag{23}$$

which lies in the interval $[y_i, y_{i+1}]$, the intersection supports of B_i and B_{i+1}.

Lemma 4.3. *For QL-implication $I(a,b) = 1 - a + a^2 b$ and fuzzy system* (1), *its system function is*

$$y = B_i^{-1}\left(\frac{A_i(x_0) - A_{i+1}(x_0) + A_{i+1}^2(x)}{A_i^2(x) + A_{i+1}^2(x)}\right), \tag{24}$$

which lies in the interval $[y_i, y_{i+1}]$, the intersection supports of B_i and B_{i+1}.

Although the corresponding system functions of fuzzy system (1) based on different S-implications and QL-implications are not the same, the following approximation theorem always holds:

Theorem 4.4. *MISO fuzzy systems based on S-implications and QL-implications with MOA defuzzifier are also universal approximators.*

Remark 4.5. Since the formulas of fuzzy systems based on different kinds of S-implications and QL-implications are different, the corresponding system functions of fuzzy system (1) are not the same. Therefore, their corresponding control capabilities are different. Furthermore, from the results of Lemmas 4.2 and 4.3, we can implement continuous control action with different weights through the S-implication $I(a,b) = 1 - a + ab$ and QL-implication $I(a,b) = 1 - a + a^2 b$. In this case, we have a better control capability. These results give a theoretic explanation of related experiment results in [3,4,7,10].

5. Examples

In the following examples, we shall use our approach to approximate the given continuous functions. The comparison of optimal fuzzy rules by our approach and those by Ying in [21] is presented.

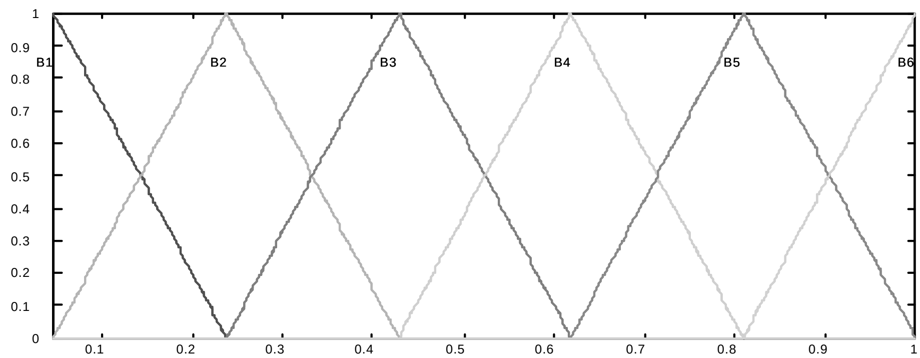

Fig. 1. The membership functions of the output variable in Example 1.

Example 1. Design a fuzzy system based on genuine many-valued implication to approximate the continuous function $F(x) = \sin(x)/x$ defined on $[-3, 3]$ with $\varepsilon = 0.2$.

The image set of $F(x)$ on $[-3, 3]$ is $[0.0470, 1]$. Design the fuzzy system through dividing the image set $[0.0470, 1]$ into $[(1 - 0.0470)/0.2] + 1 = 5$ parts, and then decide the input fuzzy membership functions.

The membership functions of the consequent part are $B_1, B_2, B_3, B_4, B_5, B_6$, which are depicted as follows (Fig. 1).

The membership functions of input variables are $A_1, A_2, \ldots, A_{11}$, which are depicted as follows (Fig. 2).

Fuzzy rule base is

R_1: IF x *is* A_1 *or* A_{11}, THEN y *is* B_1;

R_2: IF x *is* A_2 *or* A_{10}, THEN y *is* B_2;

R_3: IF x *is* A_3 *or* A_9, THEN y *is* B_3;

R_4: IF x *is* A_4 *or* A_8, THEN y *is* B_4;

R_5: IF x *is* A_5 *or* A_7, THEN y *is* B_5;

R_6: IF x *is* A_6, THEN y *is* B_6.

Using R-implication as implication operator, the systems function of this fuzzy system is depicted in the following graph (Fig. 3). A comparison of the graph of $F(x) = \sin x/x$ is also depicted in the same figure.

The approximation error is $\varepsilon = 0.0864 < 0.2$.

In this example, we use only 6 rules to approximate $F(x) = \sin x/x$ with an accuracy $\varepsilon = 0.2$, while Ying in [21] used 207 rules to approximate the same function.

Example 2. Design a fuzzy system based on genuine many-valued implication to uniformly approximate the polynomial $z = P(x_1, x_2) = 0.52 + 0.1x_1 + 0.38x_2 - 0.06x_1x_2$ defined on $[-1, 1]^2$ with $\varepsilon = 0.1$.

The image set of P over $[-1, 1]^2$ is $[-0.02, 0.94]$. Design the fuzzy system by dividing the image set $[-0.02, 0.94]$ into $[(0.94 - (-0.02))/0.1] + 1 = 10$ parts, and then decide the input fuzzy membership functions.

The membership functions of the consequent part are $B_1, B_2, \ldots, B_{11}$, which are depicted as follows (Fig. 4).

The membership functions of input variables x_1 and x_2 are $A_1^1, A_2^1, \ldots, A_7^1$ and $A_1^2, A_2^2, \ldots, A_{19}^2$, they are depicted as follows (Figs. 5 and 6), respectively. Table 1 gives us Fuzzy rule base as follows:

In this table, the natural number $1, 2, 3, \ldots$ represents the index number of the membership function for the related variable x_1, x_2 or z.

Using Goguen implication, the figure of system function $G(x_1, x_2)$ is depicted as follows (Fig. 7). A comparison of the figure of $P(x_1, x_2)$ and the figure of the errors of the system function (or approximation function) and the origin function $P(x_1, x_2)$ are also depicted in Figs. 8 and 9.

The approximation error is $\varepsilon = 0.084 < 0.1$.

In this example, we use only 133 rules (not considering decoupling) to approximate $P(x_1, x_2) = 0.52 + 0.1x_1 + 0.38x_2 - 0.06x_1x_2$ with an accuracy $\varepsilon = 0.1$, while Ying in [21] used 225 rules to approximate the same function.

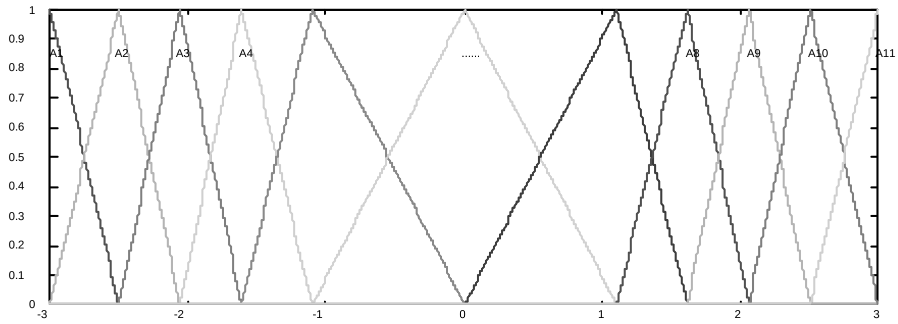

Fig. 2. The membership functions of the input variable in Example 1.

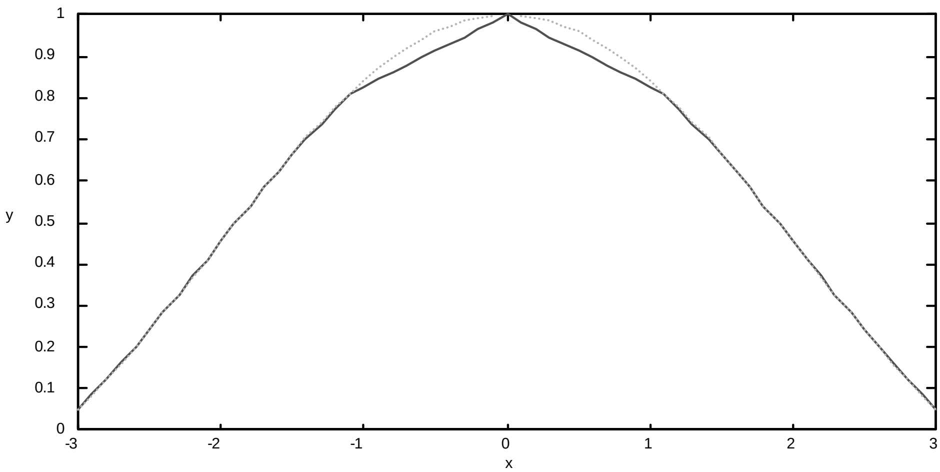

Fig. 3. The comparison of system function $y = G(x)$ (solid-line) and the origin function $y = F(x)$ (dotted-line).

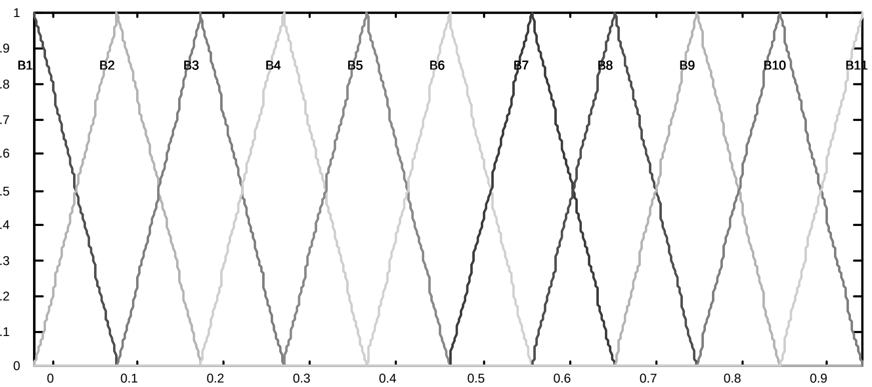

Fig. 4. The membership functions of the output variable in Example 2.

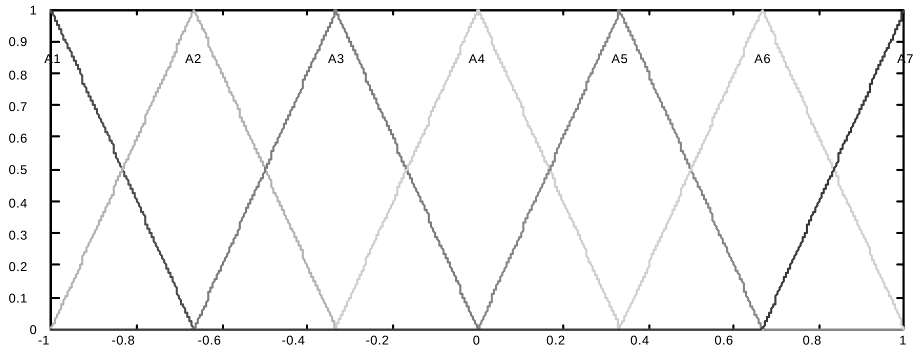

Fig. 5. The membership functions of the first input variable in Example 2, where A_i represents A_i^1 $(i=1,\ldots,7)$.

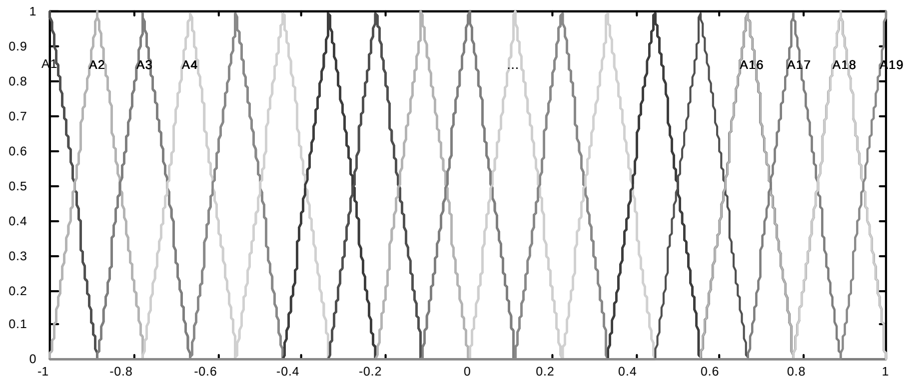

Fig. 6. The membership functions of the second input variable in Example 2, where A_i represents A_i^2 $(i=1,\ldots,19)$.

Table 1
Z values

	x_2																		
x_1	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
1	1	2	2	3	3	4	4	5	5	6	6	7	7	8	8	9	9	10	10
2	2	2	2	3	3	4	4	5	5	6	6	7	7	8	8	9	9	10	10
3	2	3	3	3	3	4	5	5	6	6	7	7	8	8	9	9	9	10	11
4	3	3	3	4	4	4	5	6	6	7	7	8	8	9	9	9	10	10	11
5	4	4	4	4	5	5	5	6	7	7	8	8	8	9	9	9	10	10	11
6	4	4	5	5	5	6	6	6	7	7	8	8	9	9	9	10	10	11	11
7	5	5	5	5	6	6	6	7	7	7	8	8	9	9	9	10	10	11	11

6. Conclusion

We discuss the approximation theory of MIMO Boolean type fuzzy systems and obtain the following results:

(1) The fuzzy systems based upon R-, S-, QL-implications are universal approximators, their

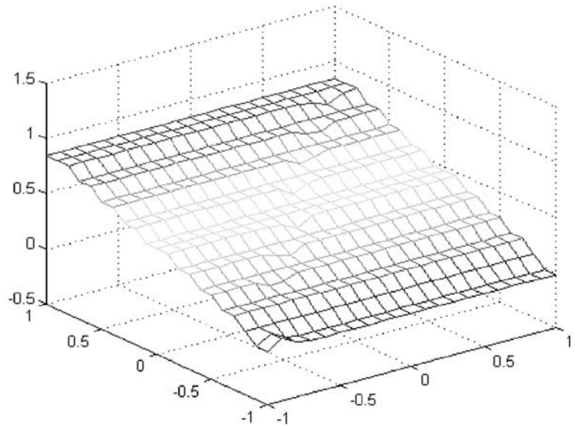

Fig. 7. The system function $z=G(x_1,x_2)$ of the fuzzy system in Example 2 based on Goguen-implication.

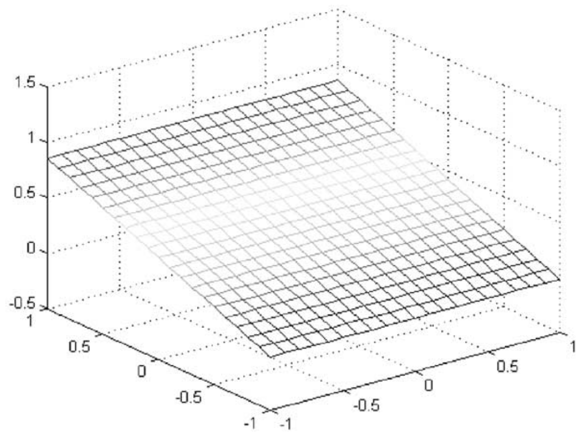

Fig. 8. The figure of origin function $z=P(x_1,x_2)$ in Example 2.

corresponding formulas of system functions are given.

(2) General approaches to construct fuzzy systems are given; two examples illustrate the approaches.

(3) Defuzzification methods and inference methods are given for Boolean type fuzzy systems.

The optimal fuzzy rules and decoupling properties of MIMO (i.e. (1)) Boolean type fuzzy systems should be further researched in the future.

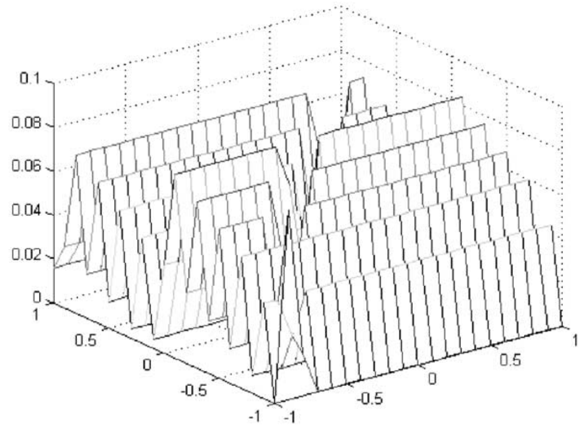

Fig. 9. Errors of the origin function and approximation function (system function) in Example 2.

Appendix

Proof of Lemma 3.1. As the choice of implication I and $\{B_i\}$, it follows that, if $B_{k(1)}(y),\ldots,B_{k(l)}(y)$ have at least three elements, then for any $y\in V$, there exists $i(y), 1\leqslant i(y)\leqslant l$ such that $B_{k(i)}(y)=0$. In this case, $A_{k(i)}(x_0)\neq 0$ and $A_{k(i)}(x_0)\rightarrow B_{k(i)}(y)=A_{k(i)}(x_0)\rightarrow 0=0$. It follows that $B'(y)=\bigwedge_{i=1}^{l}(A_{k(i)}(x_0)\rightarrow B_{k(i)}(y))=0$. Hence, $B'(y)\equiv 0$. $\square$

Proof of Lemma 3.2. Since T is a continuous t-norm, the functions $f:[0,1]\rightarrow[0,a]$ and $g:[0,1]\rightarrow[0,b]$ defined by $f(z)=T(a,z)$, $g(z)=T(b,1-z)$ are continuous surjections, $f(z)$ is nonincreasing and $g(z)$ is nondecreasing. Let $h(z)=f(z)-g(z)$, then $h(0)=f(0)-g(0)=T(a,0)-T(b,1)=-b$ and $h(1)=f(1)-g(1)=T(a,1)-T(b,0)=a$, and thus $h:[0,1]\rightarrow[-b,a]$ is a nonincreasing continuous surjection. Since $a,b\geqslant 0$ and $h(0)=-b\leqslant 0$, $h(1)=a\geqslant 0$, there is at least one point $z\in[0,1]$ such that $h(z)=0$. This z is just the solution of Eq. (15). $\square$

Proof of Corollary 3.4. Since $a+b\leqslant 1$, it follows that $a\leqslant 1-b$, $b\leqslant 1-a$, then $[b,1-a]\neq\emptyset$. For any $z\in[b,1-a]$, $h(z)=T(a,z)-T(b,1-z)$ has the following properties: $h(b)=T(a,b)-T(b,1-z)=T(b,a)-T(b,1-b)\leqslant 0$ and $h(1-a)=T(a,1-a)-T(b,a)=T(a,1-a)-T(a,b)\geqslant 0$. Since $h(z)$ is continuous, it follows that there exists one $z\in[b,1-a]$

such that $h(z)=0$, this is just to say that Eq. (15) has one solution z satisfying $a \leqslant 1-z$ and $b \leqslant z$. □

Proof of Theorem 3.5. For any input $x=x_0$, there are two rules R_k, R_{k+1} to make Eq. (14) hold, that is, $B'(y)=(A_k(x_0)\to B_k(y))\wedge(A_{k+1}(x_0)\to B_{k+1}(y))$. Let $F(x_0,y)=B'(y)$, since $\to, T$ and $\wedge$ are continuous, $F(x_0,y)$ is continuous with respect to x_0 and y. Thus, for any x_0, there exists one $y_0\in V$ such that $B'(y_0)=\max_{y\in V} B'(y)$, and thus $y_0=G(x_0)$ is also continuous w.r.t. x_0.

Since $A_k(x_0)\to B_k(y)$ and $A_{k+1}(x_0)\to B_{k+1}(y)$ are both continuous membership functions with nonempty intersections, so the y that $B'(y)$ gets its maximum value from satisfies the following equation:

$$(A_k(x_0)\to B_k(y)) = (A_{k+1}(x_0)\to B_{k+1}(y)). \quad (A.1)$$

Let $Y_1=\{y\in V \mid A_k(x_0)\leqslant B_k(y)$ and $A_{k+1}(x_0)\leqslant B_{k+1}(y)\}$, if $Y_1\neq\emptyset$, then $A_k(x_0)+A_{k+1}(x_0)\leqslant B_k(y)+B_{k+1}(y)=1$, from Corollary 3.4, there exists a solution z of Eq. (15) such that $b=A_k(x_0)\leqslant z$ and $a=A_{k+1}(x_0)\leqslant 1-z$. Let $z=B_k(y)$ and $1-z=B_{k+1}(y)$, since $B_k(y)$ is pseudo-triangle-shaped, there exists one unique point $y_0=B_k^{-1}(z)$ satisfying $B_k(y_0)+B_{k+1}(y_0)=1$ and $y_0\in Y_1$.

If $Y_1=\emptyset$, then we must have $A_k(x_0)>B_k(y)$ and $A_{k+1}(x_0)>B_{k+1}(y)$. Since T is continuous, for any $a\in(0,1]$, the map $T(a,\_):[0,1]\to[0,a]$ is always surjective, then we should have

$$T(A_k(x_0), A_k(x_0)\to B_k(y)) = B_k(y),$$

$$T(A_{k+1}(x_0), A_{k+1}(x_0)\to B_{k+1}(y)) = B_{k+1}(y). \quad (A.2)$$

From Eqs. (A.1) and (A.2), we can obtain the following equation:

$$T(A_{k+1}(x_0), B_k(y)) = T(A_k(x_0), B_{k+1}(y)). \quad (A.3)$$

Let $z=B_k(y)$, then z is just the solution of Eq. (15) for $a=A_{k+1}(x_0)$, $b=A_k(x_0)$. In this case, $(A.3)$ has only one solution $y_0=B_k^{-1}(z)$ satisfying $B_k(y_0)+B_{k+1}(y_0)=1$. □

Proof of (Appropriate Condition) of Theorem 3.9. Otherwise, there exist at least three elements $U_{j_1}, U_{j_2}, U_{j_3}$ that intersect with $E=(x_i^1, x_{i+2}^1)\times x_j^2, x_{j+2}^2)$. Without loss of generality, it is no hurt to assume that $j_1<j_2<j_3$, then $U_{j_1}\cap U_{j_3}=\emptyset$. Choosing $a_l=(a_l^1,a_l^2)\in U_{j_l}\cap(x_i,x_{i+2})$ $(l=1,2,3)$, then it follows that

$$f(a_l)\in(y_{j_l}-\tfrac{2}{3}e, y_{j_l}+\tfrac{2}{3}e)$$

and thus, $|f(a_1)-f(a_3)|>(y_{j_3}-\frac{2}{3}e)-(y_{j_1}+\frac{2}{3}e)=(y_{j_3}-y_{j_1})-\frac{4}{3}e$. Being the choice of y_{j_3}, y_{j_1}, it follows that $y_{j_3}-y_{j_1}\geqslant 2e$, and thus,

$$|f(a_1)-f(a_3)| > 2e-\tfrac{4}{3}e=\tfrac{2}{3}e>\tfrac{1}{2}e. \quad (A.4)$$

On the other hand, since $a_1,a_3\in E$ and $d((x_{i+2}^1, x_{j+2}^2),(x_i^1,x_j^2))\leqslant d((x_{i+2}^1,x_{j+2}^2),(x_{i+1}^1,x_{j+1}^2))+d((x_{i+1}^1, x_{j+1}^2),(x_i^1,x_j^2))\leqslant\delta/2+\delta/2=\delta$, it follows that $|a_1-a_3|<\delta$, and thus, $|f(a_1)-f(a_3)|<e/2$, this inequality contradicts with the inequality (A.4). This proves that the appropriate condition holds.

To prove Lemma 4.1, we need the following lemma.

Lemma A. *For two different membership functions $A_i=A_i^1*\cdots*A_i^m$ and $A_j=A_j^1*\cdots*A_j^m$ in MISO fuzzy system* (1), *where $*$ is any t-norm, for any input $x_0\in U=U_1\times\cdots\times U_m$, we always have $A_i(x_0)+A_j(x_0)\leqslant 1$.*

Proof. If $A_i(x_0)=0$ or $A_j(x_0)=0$, then $A_i(x_0)+A_j(x_0)\leqslant 1$ is obvious.

If $A_i(x_0)\neq 0$ or $A_j(x_0)\neq 0$, then for all $1\leqslant i\leqslant m$, $A_i^k(x_0)\neq 0$ and $A_j^k(x_0)\neq 0$. Since $\{A_i^k\}$ is a norm base for variable $x_k\in U_k$, there is at least one k such that $A_i^k(x_0)+A_j^k(x_0)=1$, then it follows that $A_i(x_0)+A_j(x_0)=A_i^1*\cdots*A_i^m(x_0)+A_j^1*\cdots*A_j^m(x_0)\leqslant A_i^k(x_0)+A_j^k(x_0)=1$. □

Proof of Lemma 4.1. In this case, Eqs. (20) and (21) can be rewritten as follows:

$$B'(y)=[(1-A_i(x_0))\vee B_i(y)] \wedge[(1-A_{i+1}(x_0))\vee B_{i+1}(y)], \quad (A.5)$$

$$B'(y)=[(1-A_i(x_0))\vee(A_i(x_0)\wedge B_i(y))] \wedge[(1-A_{i+1}(x_0))\vee(A_{i+1}(x_0)\wedge B_{i+1}(y))], \quad (A.6)$$

where $x_0 \in (x_i, x_{i+1})$. Write $a = A_i(x_0)$, $b = A_{i+1}(x_0)$ and $z = B_i(y)$, since $a + b = A_i(x_0) + A_{i+1}(x_0) \leqslant 1$ (Lemma A), Eqs. (A.5) and (A.6) are as follows:

$$B'(y) = [1 - (a \vee b)] \vee [(1-b) \wedge z] \vee [(1-a) \wedge (1-z)] \vee [z \wedge (1-z)], \quad \text{(A.7)}$$

$$B'(y) = [1 - (a \vee b)] \vee [(1-b) \wedge a \wedge z] \vee [(1-a) \wedge b \wedge (1-z)] \vee [a \wedge b \wedge z \wedge (1-z)] \quad \text{(A.8)}$$

then $B'(y)$ can be calculated as follows:

$$B'(y) = \begin{cases} (1-a) \vee [(1-b) \wedge z], & a > b, \\ (1-b) \vee [(1-a) \wedge (1-z)], & a < b, \\ 1-a, & a = b, \end{cases} \quad \text{(A.9)}$$

$$B'(y) = \begin{cases} (1-a) \vee [a \wedge z], & a > b, \\ (1-b) \vee [b \wedge (1-z)], & a < b, \\ 1-a, & a = b, \end{cases} \quad \text{(A.10)}$$

Since $B_i(y)$ is symmetric, the point that $B'(y)$ gets its maximum value for the case $a > b$ must satisfy the relation $B_{i+1}(y) \geqslant 1-b \geqslant a$, in this case $y_0 = y_i$; and similarly, $y_0 = y_{i+1}$ for the case $a < b$; for the case $a = b$, the maximum value $1-a$ of $B'(y)$ gets at all $y \in [c, d]$, in this case, we obtain that $y_0 = (c+d)/2$, this value will affect the approximation capability of the corresponding fuzzy system, so we make some corrections in this case: as the fired consequent membership functions are only B_i and B_{i+1}, then it is reasonable that y_0 lies in (y_i, y_{i+1}) (we assume that the support set of $B_i(y)$ and $B_{i+1}(y)$ is (y_{i-1}, y_{i+1}) and (y_i, y_{i+2}), respectively), then we get $y_0 = (y_i + y_{i+1})/2$. This completes the proof. □

Proof of Lemma 4.2. In this case, Eq. (22) can be rewritten as follows:

$$B'(y) = [1 - A_i(x_0) + A_i(x_0)B_i(y)] \wedge [1 - A_{i+1}(x_0) + A_{i+1}(x_0)B_{i+1}(y)]. \quad \text{(A.11)}$$

It is no hurt to assume that the support set of $B_i(y)$ and $B_{i+1}(y)$ is (y_{i-1}, y_{i+1}) and (y_i, y_{i+2}), respectively, in the following. We declare that the unique point that $B'(y)$ gets its maximum value lies in $[y_i, y_{i+1}]$. First, if $y \notin [y_{i-1}, y_{i+2}]$, then $B'(y) = [1 - A_i(x_0)] \wedge [1 - A_{i+1}(x_0)] = 1 - (a \vee b)$, where $A_i(x_0) = a$, $A_{i+1}(x_0) = b$, it follows that if there exists y such that $B'(y) > 1 - (a \vee b)$, then $y \in (y_{i-1}, y_{i+2})$. We prove that there exists y such that $B'(y) > 1 - (a \vee b)$ in the following. Let $z = B_i(y)$, then

$$\begin{aligned} B'(y) &= (1 - a + az) \wedge [1 - b + b(1-z)] \\ &= (1 - a + az) \wedge (1 - bz). \end{aligned} \quad \text{(A.12)}$$

In order to demand $B'(y) > 1 - (a \vee b)$, it suffices to require that the following inequality has a solution z for any $a \in [0, 1]$

$$\begin{aligned} 1 - a + az &> 1 - (a \vee b), \\ 1 - bz &> 1 - (a \vee b). \end{aligned} \quad \text{(A.13)}$$

The above inequality can be rewritten as follows:

$$\begin{aligned} a(1-z) &< a \vee b, \\ bz &< a \vee b. \end{aligned} \quad \text{(A.14)}$$

The above equality holds for any $0 < z < 1$.

For a fixed $a \in [0, 1]$, because $z_1 = 1 - a + az$ and $z_2 = 1 - bz$ represent two direct lines, its slope is a and $(-b)$, respectively, then z_1 is strictly increasing about z and z_2 is strictly decreasing about z. The point that $B'(y)$ gets its maximum value is the intersection point of these two direct lines, then it satisfies the equation $z = a/(a+b)$, that is

$$B_i(y_0) = \frac{A_i(x_0)}{A_i(x_0) + A_{i+1}(x_0)} \quad \text{and}$$

$$B_{i+1}(y_0) = \frac{A_{i+1}(x_0)}{A_i(x_0) + A_{i+1}(x_0)}. \quad \text{(A.15)}$$

For the input $x_0 \in [x_i, x_{i+1}]$, $B_i(y)$ and $B_{i+1}(y)$ are the only two fired consequent parts, then y_0 must lie in the intersections of the supports of $B_i(y)$ and $B_{i+1}(y)$, that is, $y_0 \in [y_i, y_{i+1}]$. Since $B_i(y)$ is strictly monotone in the interval $[y_i, y_{i+1}]$, there exists a unique point

$$y_0 = B_i^{-1}\left(\frac{A_i(x_0)}{A_i(x_0) + A_{i+1}(x_0)}\right)$$

to make Eq. $(A.15)$ hold, this is just the output of the fuzzy system.

The proof of Lemma 4.3 is similar to that of Lemma 4.2.

The proof of Theorem 4.4 is similar to that of Theorem 3.9.

References

[1] J.J. Buckley, Universal fuzzy controllers, Automatica 18 (6) (1992) 1245–1248.

[2] A.J. Bugarín, S. Barro, Reasoning with truth values on compacted fuzzy chained rules, IEEE Trans. Systems, Man Cybernet. — Part B 28 (1) (1998) 34–46.

[3] Z. Cao, A. Kandel, Applicability of some fuzzy implication operators, Fuzzy Sets and Systems 31 (1989) 151–186.

[4] Z. Cao, D. Park, A. Kandel, Investigations on the applicability of fuzzy inference, Fuzzy Sets and Systems 49 (1992) 151–169.

[5] J.L. Castro, Fuzzy logic controllers are universal approximators, IEEE Trans. Systems, Man Cybernet. 25 (4) (1995) 629–635.

[6] J.L. Castro, M. Delgado, Fuzzy systems with defuzzification are universal approximators, IEEE Trans. Systems, Man Cybernet. — Part B 26 (1) (1996) 149–152.

[7] O. Cordón, F. Herrera, A. Peregrín, Applicability of the fuzzy operators in the design of fuzzy logic controllers, Fuzzy Sets and Systems 86 (1997) 15–41.

[8] D. Dubois, H. Prade, Fuzzy sets in approximate reasoning, Part 1: inference with possibility distribution, Fuzzy Sets and Systems 40 (1991) 143–202.

[9] D. Dubois, J. Lang, H. Prade, Fuzzy sets in approximate reasoning, Part 2: logical approaches, Fuzzy Sets and Systems 40 (1991) 203–244.

[10] J.B. Kiszka, M.E. Kochanska, D.S. Sliwinska, The influence of some fuzzy implication operators on the accuracy of a fuzzy model — Part 1, Part 2. Fuzzy Sets and Systems 15 (1985) 111–128, 223–240.

[11] B. Kosko, Fuzzy systems as universal approximators, IEEE Trans. Comput. 43 (11) (1994) 1329–1333.

[12] C.C. Lee, Fuzzy logic in control systems: fuzzy logic controller, IEEE Trans. Systems Man Cybernet. 20 (1990) 404–435.

[13] Y.M. Li, Z.K. Shi, Approximation capability of fuzzy I/O systems based on similarity measures, WCICA'2000, Hefei, China, June 28 – July 2, 2000, pp. 1790–1794.

[14] Y.M. Li, Z.K. Shi, Z.H. Li, Approximation theory of fuzzy systems based upon genuine many-valued implications—SISO cases, Fuzzy Sets and Systems, in this issue.

[15] D. Lissin, M.A. Gennert, Proceedings NAFIPS 1999, pp. 184–188.

[16] L. Mántaras, Approximate reasoning models, Ellis Horwood Series in Artificial Intelligence, Wiley, 1990, 68p.

[17] H.T. Nguyen, V. Kreinovich, O. Sirisaugtakin, Fuzzy control as a universal control tool, Fuzzy Sets and Systems 80 (1996) 71–86.

[18] G.J. Wang, On the logic foundation of fuzzy reasoning, Inform. Sci. 117 (1) (1999) 47–88.

[19] L.X. Wang, J.M. Mendel, Fuzzy basic functions, universal approximation, and orthogonal least-square learning, IEEE Trans. Neural Networks 3 (5) (1992) 807–814.

[20] P.Z. Wang, S.H. Tan, F.M. Song, P. Liang, Constructive theory for fuzzy systems, Fuzzy Sets and Systems 88 (1997) 195–203.

[21] H. Ying, Sufficient conditions on general fuzzy systems as function approximators, Automatic 30 (3) (1994) 521–525.

[22] H. Ying, Sufficient conditions on uniform approximation of multivariate functions by Takagi–Sugeno fuzzy systems with linear rule consequent, IEEE Trans. Systems, Man Cybernet. — Part A 28 (4) (1998) 515–520.

[23] X.J. Zeng, M.G. Singh, Approximation theory of fuzzy systems — SISO case, IEEE Trans Fuzzy Systems 2 (2) (1994) 162–176.

[24] X.J. Zeng, M.G. Singh, Approximation theory of fuzzy systems — MIMO case, IEEE Trans Fuzzy Systems 3 (2) (1995) 219–235.