ELSEVIER Journal of Information Sciences 108 (1998) 91–107

INFORMATION SCIENCES
AN INTERNATIONAL JOURNAL

General Takagi–Sugeno fuzzy systems with simplified linear rule consequent are universal controllers, models and filters

Hao Ying [1]

Department of Physiology and Biophysics, Biomedical Engineering Center, Room 621, Jennie Sealy Hospital, D56, The University of Texas Medical Branch, Galveston, TX 77555-0456, USA

Received 1 February 1997; received in revised form 29 July 1997; accepted 23 October 1997

Abstract

Takagi–Sugeno (TS) fuzzy systems have successfully been employed, mainly in a trial-and-error manner, to solve many control and modeling problems; but their applications as signal filters remain to be fully explored. Compared to their nonfuzzy counterparts, TS fuzzy controllers and models are difficult to be efficiently constructed because there is a large number of design parameters in the rule consequent. The number grows dramatically with the increase of the number of input fuzzy sets and input variables. Furthermore, there exists little published result on relationship between TS fuzzy controllers/models/filters and their nonfuzzy counterparts. In this paper, we investigate, in relation to some popular nonfuzzy controllers, models and filters, analytical structure of a general class of multi-input single-output (MISO) TS fuzzy systems that use arbitrary fuzzy rules with our recently introduced simplified linear rule consequent. Other components of the fuzzy systems in this study are general: arbitrary continuous input fuzzy sets, any type of fuzzy logic AND and the generalized defuzzifier containing the widely used centroid defuzzifier as a special case. We prove that the general MISO TS fuzzy systems are: (1) nonlinear variable gain controllers when implemented as controllers, or (2) nonlinear time-varying auto-regressive with the extra input (ARX) models when implemented as models, or (3) nonlinear infinite impulse response (IIR) or finite impulse response (FIR) filters when implemented as filters. Furthermore, we constructively prove that the general TS fuzzy systems with the simplified linear rule consequent are universal approximators and can approximate any continuous function in closed do-

[1] Tel.: +1 409 722 8415; fax: +1 409 772 0751; e-mail: hying@utmb.edu.

PII: S0020-0255(97)10058-5

main arbitrarily well. The practical implication of these results is that these fuzzy systems, with much less design parameters, are always able to produce solutions to various control, modeling and filtering problems. We also establish sufficient conditions that can be used to calculate the number of input fuzzy sets and rules needed for achieving prespecified approximation accuracy.

1. Introduction

The main difference between Mamdani fuzzy systems and Takagi–Sugeno (TS) fuzzy systems lies in the consequent of the fuzzy rules: the former uses fuzzy sets whereas the latter employs (linear) functions of input variable [17]. Both types of fuzzy systems have been used, mainly in a trial-and-error fashion, as effective tools to solve various practical problems mainly in the fields of control and modeling. Their applications as signal filters, nevertheless, remain to be fully explored. At present, there exists little result in the literature on relationship between TS fuzzy controllers/models/filters and their nonfuzzy counterparts.

Regardless of the nature of the applications, from mathematics standpoint, fuzzy systems produce function mapping between system input and system output. As such, the issue of fuzzy systems as universal approximators is important and has become an intensive research subject in the past few years. Most of the existing results in the literature are for Mamdani fuzzy systems (e.g., [1,4,5,8–11,16,18,19,26]). Some of these results are obtained by using the Stone–Weierstrass theorem, and they are existence results on some particular configurations of Mamdani fuzzy systems. We have investigated necessary as well as sufficient conditions for the general Mamdani fuzzy systems as universal approximators [6,20,21]. In [26], approximation accuracy analysis is conducted for some Mamdani fuzzy systems.

Studying approximation capability of TS fuzzy systems is equally important. At present, no result is available in the literature for typical TS fuzzy systems. By "typical", we mean those TS fuzzy systems that have been widely used in practice and studied in theory, which use linear rule consequent and the centroid defuzzifier, as originally proposed by Takagi and Sugeno [17]. We have recently proved that the general typical TS fuzzy systems are universal approximators [22]. The existence result in [3] is interesting but is derived for an uncommon two-input one-output TS fuzzy system that uses a defuzzifier without denominator and rule consequent must be (high order) polynomials of input variables.

Virtually all the practical TS fuzzy systems use linear functions in rule consequent, which is critical to their practicality and usefulness. This is because when nonlinear rule consequent are used, properly choosing the structure and parameters of the rule consequent becomes extremely difficult, if not impossible. Furthermore, compared with the well-established polynomial approximators in traditional mathematics, fuzzy systems using nonlinear rule consequent, such as high order polynomial or nonlinear functions, are greatly disadvantageous in terms of complexity and practical usefulness. For these reasons, in this paper we will concentrate on TS fuzzy systems employing linear rule consequent only. Developing approximation theory for a TS fuzzy system with linear rule consequent is technically more difficult and challenging than for the same TS fuzzy system but with nonlinear rule consequent due to the restriction on the linearity of the rule consequent.

Apart from the issue of universal approximators, compared with Mamdani fuzzy systems, designing TS fuzzy systems is (much) more difficult. This is because there is a large number of design parameters in the rule consequent. To make the matter worse, selecting and tuning of these parameters have to be done in a blind trial-and-error fashion with little intuition or analytical basis. Fuzzy systems with learning capability, achieved by using neural networks, have been proposed in the literature to learn the values of the parameters on-line or off-line. This approach however is impractical and ineffective even when the number of the parameters is moderately large. Other drawbacks of this approach include the requirement of large training data sets as well as training time.

To significantly reduce the number of design parameters in rule consequent yet still retain the advantage and spirit of the TS rule scheme, we recently proposed a simplified TS rule scheme [23,24]. The linear version of the simplified rule scheme requires that all the rule consequent use linear functions and the rule consequent be proportional to one another. The rule proportionality is the novelty and useful key of our new scheme. We will show these points in detail later in this paper.

In the present paper, we will first establish a relationship between the analytical structure of a general class of multi-input single-output (MISO) TS fuzzy systems that use our simplified linear TS rule consequent and the structures of some popular nonfuzzy controllers, models and filters. We will then constructively prove that these general TS fuzzy systems are universal approximators. Finally, we will establish some sufficient conditions that can be used to calculate the number of input fuzzy sets and fuzzy rules needed for achieving pre-specified approximation accuracy. The practical implication of these results is that these TS fuzzy systems, though with much less design parameters, are always capable of producing solutions to various control, modeling and filtering problems.

2. General MISO fuzzy systems with simplified linear TS rule consequent

2.1. Configuration of the general TS fuzzy systems

The general TS fuzzy systems in this investigation use r continuous-time or discrete-time input variables, which are represented by a vector

$$\underline{x(t)} \stackrel{\text{def}}{=} (x_1(t), x_2(t), \ldots, x_r(t)),$$

where t is the time. For notational simplicity, we use $\underline{x}$ and x_i instead of $\underline{x(t)}$ and $x_i(t)$ in the rest of this paper. Without loss of generality, we suppose that

$$0 \leqslant x_i \leqslant 1, \quad i = 1, 2, \ldots, r.$$

For x_i, we partition [0,1] into $n_i (n_i \geqslant 1)$ equal intervals, each of which is $[k/n_i, (k+1)/n_i]$ where $k = 0, 1, \ldots, n_i - 1$. Over these intervals, we define $n_i + 1$ input fuzzy sets, one for each interval. Each of the fuzzy sets is denoted by $A_{i,j} (j = 0, 1, \ldots, n_i)$ whose membership function $\mu_{A_{i,j}}$ can be any continuous function. For $i_1 \neq i_2$ and $j_1 \neq j_2, \mu_{A_{i_1,j_1}}$ and $\mu_{A_{i_2,j_2}}$ may be chosen different.

To cover all the possible combinations of $A_{i,j}$'s, which is

$$\Omega \stackrel{\text{def}}{=} \prod_{i=1}^{r} (n_i + 1),$$

Ω fuzzy rules are used. In the present paper, we do not use the original TS rule consequent, a representative of which is [17]:

$$\text{Rule}\#m \qquad \text{IF } x_1 \text{ is } A_{1,p_{1,m}} \text{ AND } x_2 \text{ is } A_{2,p_{2,m}} \text{ AND } \ldots \text{ AND } x_r \text{ is } A_{r,p_{r,m}}$$
$$\text{THEN } F_\Omega(\underline{x}) = a_{0m}x_0 + a_{1m}x_1 + a_{2m}x_2 + \cdots + a_{rm}x_r. \tag{1}$$

Here, we use a dummy variable x_0 and let $x_0 \equiv 1$ to simplify the mathematical representation of the rule consequent later. $F_\Omega(\underline{x})$ is the output of the fuzzy systems. We use the subscripts Ω to indicate the parameterization of the systems output by n_i. The subscripts $p_{i,m}$'s are integers $(0 \leqslant p_{i,m} \leqslant n_i)$ and a_{im}'s are $r+1$ design parameters which can be any values chosen by the system developer. The power of TS fuzzy systems lies in the flexibility of choosing the values of these design parameters. However, this flexibility comes with a severe tradeoff: too many parameters need to be selected/tuned. Specifically, for Ω rules, up to $\kappa \stackrel{\text{def}}{=} (r+1)\Omega$ different parameters are required and κ grows very quickly with the increase of r and/or Ω (i.e., n_i). To illustrate this problem more concretely, let us assume a rather simple TS fuzzy system with only three input variables and each of them is fuzzified by merely two input fuzzy sets (i.e., $r = 3, n_1 = n_2 = n_3 = 1$). Then, $\kappa = 32$. If we keep $r = 3$ but let $n_1 = n_2 = n_3 = 2$, κ increases dramatically to 108! Any fuzzy system using too many tunable parameters is practically useless no matter how flexible it is in theory. This shortcoming of the original TS rule scheme becomes especially apparent when the TS fuzzy systems are implemented as controllers because time has proved that the PID controller, which has only three design parameters to tune, can effectively control most industrial processes with satisfactory performance. The same can be said of the TS fuzzy systems implemented as models since traditional models, such as the auto-regressive moving-average (ARMA) model, are not only effective but also only involve a small number of design parameters.

We recently proposed a simplified TS rule scheme, which uses far less parameters in rule consequent. The linear version of the scheme is as follows [23,24]:

$$
\begin{aligned}
&\text{Rule}\ \#1 \quad && \text{IF}\, x_1\, \text{is}\, A_{1,p_{1,1}}\ \text{AND}\, x_2\, \text{is}\, A_{2,p_{2,1}}\ \text{AND} \cdots \text{AND}\, x_r\, \text{is}\, A_{r,p_{r,1}} \\
& && \text{THEN}\, F_\Omega(\underline{x}) = k_1(a_0x_0 + a_1x_1 + a_2x_2 + \cdots + a_rx_r). \\
& && \vdots \\
&\text{Rule}\ \#m \quad && \text{IF}\, x_1\, \text{is}\, A_{1,p_{1,m}}\ \text{AND}\, x_2\, \text{is}\, A_{2,p_{2,m}}\ \text{AND} \cdots \text{AND}\, x_r\, \text{is}\, A_{r,p_{r,m}} \\
& && \text{THEN}\, F_\Omega(\underline{x}) = k_m(a_0x_0 + a_1x_1 + a_2x_2 + \cdots + a_rx_r). \\
& && \vdots
\end{aligned} \tag{2}
$$

In our simplified linear TS rule scheme, all the rule consequent still use linear functions but rule consequent are required to be proportional to one another. The rule proportionality is the novelty and usefulness of our scheme [23,24]. The relationship between the parameters in the original linear TS rule consequent (1) and that in the simplified linear TS rule consequent (2) is $a_{im} = k_m a_i$. Hence, the simplified rule consequent may be regarded as a special kind of the original TS rule consequent.

For Ω rules, the simplified rule consequent have up to $\lambda \stackrel{\text{def}}{=} r + 1 + \Omega$ different design parameters. The parameter reduction rate of the simplified TS rule scheme over the original TS rule scheme is

$$\eta \stackrel{\text{def}}{=} \frac{\kappa - \lambda}{\lambda} = \frac{r(\Omega - 1) - 1}{r + 1 + \Omega}.$$

Obviously, the reduction rate increases quickly with the increase of r and/or Ω (i.e., n_i). Note that when at least one of n_i's is large, Ω is large, and to the limit,

$$\lim_{n_i \to \infty} \eta = r, \quad 1 \leqslant i \leqslant r$$

meaning the simplified scheme can reduce the number of design parameters in rule consequent by a factor close to r when at least one of the input variables is fuzzified by a large number of input fuzzy sets. For a better illustration of the significance of the parameter reduction, we calculate and tabulate in Table 1 η for different combinations of r and Ω, assuming $n_1 = \cdots = n_r$.

Table 1
Parameter reduction rate of our simplified linear TS rule consequent over the original linear TS rule consequent

Reduction rate η (%)	Number of input fuzzy sets for each input variable				
Number of input variable	2	3	4	5	6
1	0	20	33	43	50
2	71	125	153	168	177
3	167	248	276	288	293
4	281	371	390	296	398

We now introduce the following notations to describe our simplified rule consequent in a more mathematically concise fashion:

$$k_m = f(p_{1,m}, \ldots, p_{r,m}) \stackrel{\text{def}}{=} f(\underline{p_m}) \quad \text{and} \quad \underline{p_m} \stackrel{\text{def}}{=} (p_{1,m}, \ldots, p_{r,m}),$$

where f can be any function. Then, the above simplified linear TS rule $\#m$ becomes:

$$\text{IF } x_1 \text{ is } A_{1,p_{1,m}} \text{ AND } x_2 \text{ is } A_{2,p_{2,m}} \text{ AND } \ldots \text{ AND } x_r \text{ is } A_{r,p_{r,m}}$$

$$\text{THEN} F_\Omega(\underline{x}) = f(\underline{p_m}) \sum_{i=0}^{r} a_i x_i.$$

We point out that our simplified TS rule scheme is not restrictive as f can be any function.

To evaluate the ANDs in the rule, any type of fuzzy logic AND [12,13] may be used and a mixture of different types of ANDs (e.g., Zadeh AND and product AND) may be used. We use the symbol $\otimes$ to represent any type of fuzzy logic ANDs used in the rule. After the fuzzy AND calculations, the membership

$$\mu_m = \mu_{A_{1,p_{1,m}}} \otimes \mu_{A_{2,p_{2,m}}} \otimes \cdots \otimes \mu_{A_{r,p_{r,m}}}$$

is assigned to $F_\Omega(\underline{x})$ in the consequent of Rule $\#m$.

Finally, the generalized defuzzifier [7] is used to calculate the output of the fuzzy systems, which is

$$F_\Omega(\underline{x}) = \frac{\sum_{m=1}^{\Omega} \left(\mu_m^\alpha f(\underline{p_m}) \sum_{i=0}^{r} a_i x_i\right)}{\sum_{m=1}^{\Omega} \mu_m^\alpha} = \frac{\sum_{m=1}^{\Omega} \mu_m^\alpha f(\underline{p_m})}{\sum_{m=1}^{\Omega} \mu_m^\alpha} \sum_{i=0}^{r} a_i x_i. \tag{3}$$

Different defuzzification results can be obtained by using different α values, where $0 \leqslant \alpha < +\infty$. The popular centroid defuzzification method is a special case of this generalized defuzzification method when $\alpha = 1$, and the mean of maximum defuzzification method is another when $\alpha = \infty$.

Note that except the simplified linear TS rule consequent, the components of the fuzzy systems described above are general, typical and widely used. The simplified linear TS rules are not restrictive because rule proportionality between any of two rules is arbitrary and can be selected by the system developer at will.

2.2. The general MISO fuzzy systems with the simplified linear TS rule consequent as controllers, models and filters

The general MISO fuzzy systems with the simplified linear TS rule consequent become general fuzzy controllers if $\underline{x}$ consists of state variables of a process to be controlled (e.g., error of process output and rate change of error of process output). $F_\Omega(\underline{x})$ then represents output of the fuzzy controllers. Now we reveal the relationship between these fuzzy controllers and some popular nonfuzzy controllers, such as the PID controller. We express (3) in the following form:

$$F_\Omega(\underline{x}) = \sum_{i=0}^{r} b_i(\underline{x}) x_i \tag{4}$$

where

$$b_i(\underline{x}) = \frac{a_i \sum_{m=1}^{\Omega} \mu_m^{\alpha} f(\underline{p_m})}{\sum_{m=1}^{\Omega} \mu_m^{\alpha}}.$$

According to (4), the general fuzzy systems, when implemented as fuzzy controllers, are actually nonlinear time-varying controllers in the context of nonfuzzy controllers. The gain is $b_i(\underline{x})$ for state variable x_i. If error, rate change of error and rate change of error rate of process output are used as state variables, we obtain, from (4), nonlinear PID controllers with variable gains changing, during control time, with the values of these state variables. The reader is referred to [23,24] for the detail analytical analysis of such TS fuzzy controllers where we show a number of practically and theoretically desirable control characteristics and properties.

The general TS fuzzy systems with the simplified linear rule consequent can also be implemented as general fuzzy models. To do so, different input variables need to be employed. Subsequently, simplified linear TS rule $\#m$ should be

$$\begin{aligned}
&\text{IF } y(n) \text{ is } C_{0,p_{0,m}} \text{ AND } y(n-1) \text{ is } C_{1,p_{1,m}} \text{ AND} \cdots \\
&\text{AND } y(n-n_a) \text{ is } C_{n_a,p_{n_a,m}} \\
&\text{AND } u(n) \text{ is } D_{0,s_{0,m}} \text{ AND } u(n-1) \text{ is } D_{1,s_{1,m}} \text{ AND} \cdots \\
&\text{AND } u(n-n_b) \text{ is } D_{n_b,s_{n_b,m}} \\
&\text{THEN} y(n+1) = f(\underline{q_m}) \left[\sum_{i=0}^{n_a} \alpha_i y(n-i) + \sum_{j=0}^{n_b} \beta_j u(n-j) \right],
\end{aligned} \tag{5}$$

where

$$\underline{q_m} \stackrel{\text{def}}{=} (p_{0,m}, \ldots, p_{n_a,m}, s_{0,m}, \ldots, s_{n_b,m}).$$

Here, $y(n-i)$ is the model output at time $n-i$ and $u(n-j)$ the model input at time $n-j$. $Y(n-i)$ and $u(n-j)$ are state variables and there are totally $n_a + n_b + 2$ of them. In the rule, $C_{i,p_{i,m}}$ and $D_{j,s_{j,m}}$ are input fuzzy sets whose definitions are the same as that of $A_{k,p_{k,m}}$ that was given earlier, and α_i and β_j are $n_a + n_b + 2$ design parameters.

Now we reveal the link between these general fuzzy models and some popular nonfuzzy models. On the basis of (3) and (5), the general fuzzy models can be expressed as

$$y(n+1) = \frac{\sum_{m=1}^{\Omega} \mu_m^{\alpha} f(\underline{q_m}) \left(\sum_{i=0}^{n_a} \alpha_i y(n-i) + \sum_{j=0}^{n_b} \beta_j u(n-j) \right)}{\sum_{m=1}^{\Omega} \mu_m^{\alpha}}$$

$$= \sum_{i=0}^{n_a} \lambda_i(\underline{z}) y(n-i) + \sum_{j=0}^{n_b} \tau_j(\underline{z}) u(n-j) \tag{6}$$

where

$$\lambda_i(\underline{z}) = \frac{\alpha_i \sum_{m=1}^{\Omega} \mu_m^{\alpha} f(\underline{q_m})}{\sum_{m=1}^{\Omega} \mu_m^{\alpha}}, \qquad \tau_j(\underline{z}) = \frac{\beta_j \sum_{m=1}^{\Omega} \mu_m^{\alpha} f(\underline{q_m})}{\sum_{m=1}^{\Omega} \mu_m^{\alpha}},$$

$$\underline{z} \stackrel{\text{def}}{=} (y(n), y(n-1), \ldots, y(n-n_a), u(n), u(n-1), \ldots, u(n-n_b)).$$

Recall that the linear time-invariant auto-regressive with the extra input (ARX) dynamic model is [14]

$$y(n+1) + \sum_{i=0}^{n_a} c_i y(n-i) = \sum_{j=0}^{n_b} d_j u(n-j) + e(n+1), \tag{7}$$

where c_i and d_j are constant parameters and $e(n+1)$ represents random error. Comparing (6) with (7), one sees that when implemented as general fuzzy models, the general fuzzy systems with the simplified linear TS rule consequent are actually nonlinear time-varying ARX dynamic models in the context of classical modeling.

Finally, we show that the general TS fuzzy systems can also be employed as digital signal filters. Note that the widely used linear infinite impulse response (IIR) filter with input $u(n)$ and output $y(n)$ is described by

$$y(n) = \sum_{i=1}^{n_a} c_i y(n-i) + \sum_{j=0}^{n_b} d_j u(n-j). \tag{8}$$

When $d_j = 0$ for all j, the IIR filter becomes a finite impulse response (FIR) filter. On the basis of the similarity between (7) and (8), it is clear that the general fuzzy systems are nonlinear time-varying IIR or FIR filters provided that the fuzzy rules are properly constructed in a fashion similar to those in (5). It has been shown that time-varying filters can be advantageous in many applications [15].

From the above analyses, it is clear that as far as the input–output relationship is concerned, the role that the general MISO TS fuzzy systems play as controllers, models or filters is mathematically the same: they generate continuous, nonlinear and time-varying mapping between the input and output of the systems. In other words, $F_\Omega(\underline{x})$ represents nonlinear mapping $F_\Omega\colon C^r[0,1] \to [a,b]$, where $C^r[0,1]$ designates r dimensional product space. A fundamentally important question is then naturally raised: can the general TS fuzzy systems always approximate any continuous and nonlinear multivariate functional relationships with an arbitrarily high accuracy? Put the question in another way: are these general fuzzy systems with the simplified linear TS rule consequent universal approximators? If the answer is yes, then these fuzzy systems are theoretically guaranteed to always be able to produce desired control, modeling and filtering results as long as the fuzzy systems are correctly constructed. We are now going to answer this important question.

3. The general MISO fuzzy systems with the simplified linear TS rule consequent as universal approximators

We will take two steps to prove that the general MISO fuzzy systems with the simplified linear TS rule consequent are universal approximators. First, we will constructively prove that the general fuzzy systems can uniformly approximate any multivariate polynomial to any degree of accuracy. Then, we will utilize the fact that any multivariate continuous function can always be approximated by multivariate polynomial arbitrarily well (i.e., the Weierstrass approximation theorem [2]) to prove that the general fuzzy systems can uniformly approximate any multivariate continuous function with arbitrary precision. We developed and used this two-step approach previously when proving that the general Mamdani fuzzy systems are universal approximators [20].

In the rest of the paper, we will always assume that $P_M(\underline{x})$ is a multivariate polynomial of degree M defined in $C^r[0,1]$

$$P_M(\underline{x}) = \sum_{d_1=0}^{M_1}\sum_{d_2=0}^{M_2}\cdots\sum_{d_r=0}^{M_r}\beta_{d_1,\ldots,d_r}x_1^{d_1}x_2^{d_2}\cdots x_r^{d_r}, \qquad \sum_{i=1}^{r}M_i = M.$$

We now prove the following result, completing the first of the two steps.

Theorem 1. *The general MISO fuzzy systems with the simplified linear TS rule consequent can uniformly approximate $P_M(\underline{x})$ with arbitrarily small approximation error bound. That is, $\forall\varepsilon > 0$, there exists a positive integer N such that when the smallest of $n_1,\ldots,n_r$ is greater than N,*

$$\|F_\Omega - P_M\|_{C^r[0,1]} = \max_{\underline{x}\in C^r[0,1]}|F_\Omega(\underline{x}) - P_M(\underline{x})| < \varepsilon.$$

Proof. We will constructively prove this result. We first use $P_M(\underline{x})$ to form $f(\underline{p_m})$. Specifically, we let

$$f(\underline{p_m}) = \frac{P_M\left(\frac{p_{1,m}}{n_1},\ldots,\frac{p_{r,m}}{n_r}\right)}{a_0x_0 + \sum_{i=1}^{r} a_i\frac{p_{i,m}}{n_i}}. \tag{9}$$

To avoid the denominator becoming zero, we choose such a_i's that $a_0x_0 + \sum_{i=1}^{r} a_ix_i \neq 0$, which can always be achieved because a_i's are design parameters whose values are chosen by the system developer at will. Substituting (9) into (3), we obtain

$$F_\Omega(\underline{x}) = \frac{\sum_{m=1}^{\Omega} \mu_m^\alpha P_M\left(\frac{p_{1,m}}{n_1}, \ldots, \frac{p_{r,m}}{n_r}\right) \frac{\sum_{i=0}^{r} a_i x_i}{a_0 x_0 + \sum_{i=1}^{r} a_i \frac{p_{i,m}}{n_i}}}{\sum_{m=1}^{\Omega} \mu_m^\alpha}. \tag{10}$$

Because at any specific time, x_i is always in one of the n_i intervals, that is,

$$\frac{p_{i,m}}{n_i} \leqslant x_i \leqslant \frac{p_{i,m}+1}{n_i}, \tag{11}$$

we have

$$\lim_{n_i \to \infty} \frac{p_{i,m}}{n_i} \leqslant x_i \leqslant \lim_{n_i \to \infty} \frac{p_{i,m}+1}{n_i}, \quad i = 1, \ldots, r,$$

which leads to

$$\lim_{n_i \to \infty} \frac{p_{i,m}}{n_i} = \lim_{n_i \to \infty} \frac{p_{i,m}+1}{n_i} = x_i.$$

Consequently, from (10) we have

$$\lim_{n_i \to \infty} F_\Omega(\underline{x}) = \lim_{n_i \to \infty} \frac{\sum_{m=1}^{\Omega} \mu_m^\alpha P_M\left(\frac{p_{1,m}}{n_1}, \ldots, \frac{p_{r,m}}{n_r}\right) \frac{\sum_{i=0}^{r} a_i x_i}{a_0 x_0 + \sum_{i=1}^{r} a_i \frac{p_{i,m}}{n_i}}}{\sum_{m=1}^{\Omega} \mu_m^\alpha} = P_M(\underline{x}),$$

meaning $F_\Omega(\underline{x})$ can approximate $P_M(\underline{x})$ arbitrarily well if n_i's (for all i) are large enough.

Now we need to prove that this approximation is uniform. We will derive a formula that can be used to calculate a positive integer N, based on pre-specified approximation error $\varepsilon > 0$, such that when the smallest of $n_1, \ldots, n_r$ is greater than N, the following will hold:

$$\max_{\underline{x} \in C^r[0,1]} |F_\Omega(\underline{x}) - P_M(\underline{x})| < \varepsilon.$$

According to (10), this inequality is achieved if the following inequality can hold:

$$\left| \frac{\sum_{i=0}^{r} a_i x_i}{a_0 x_0 + \sum_{i=1}^{r} a_i \frac{p_{i,m}}{n_i}} P_M\left(\frac{p_{1,m}}{n_1}, \ldots, \frac{p_{r,m}}{n_r}\right) - P_M(\underline{x}) \right|$$
$$= \frac{\left| P_M\left(\frac{p_{1,m}}{n_1}, \ldots, \frac{p_{r,m}}{n_r}\right) \sum_{i=0}^{r} a_i x_i - P_M(\underline{x}) \sum_{i=0}^{r} a_i \frac{p_{i,m}}{n_i} \right|}{\left| a_0 x_0 + \sum_{i=1}^{r} a_i \frac{p_{i,m}}{n_i} \right|} < \varepsilon.$$

We make the following definition:

$$\Sigma_{\min} \stackrel{\text{def}}{=} \min_{\underline{x} \in C^r[0,1]} \left| a_0 x_0 + \sum_{i=1}^{r} a_i x_i \right|$$

and obviously

$$\left|a_0x_0+\sum_{i=1}^{r}a_i\frac{p_{i,m}}{n_i}\right|\geqslant\Sigma_{\min}.$$

Using $\Sigma_{\min}$ and continuing our inequality derivation, we now want

$$\frac{\left|P_M\left(\frac{p_{1,m}}{n_1},\ldots,\frac{p_{r,m}}{n_r}\right)\sum_{i=0}^{r}a_ix_i-P_M(\underline{x})\sum_{i=0}^{r}a_i\frac{p_{i,m}}{n_i}\right|}{\left|a_0x_0+\sum_{i=1}^{r}a_i\frac{p_{i,m}}{n_i}\right|}\leqslant\frac{\left|P_M\left(\frac{p_{1,m}}{n_1},\ldots,\frac{p_{r,m}}{n_r}\right)\sum_{i=0}^{r}a_ix_i-P_M(\underline{x})\sum_{i=0}^{r}a_i\frac{p_{i,m}}{n_i}\right|}{\Sigma_{\min}}<\varepsilon$$

or equivalently

$$\left|P_M\left(\frac{p_{1,m}}{n_1},\ldots,\frac{p_{r,m}}{n_r}\right)\sum_{i=0}^{r}a_ix_i-P_M(\underline{x})\sum_{i=0}^{r}a_i\frac{p_{i,m}}{n_i}\right|<\varepsilon\Sigma_{\min}.$$

For simplicity and better presentation, we will continue our proof and formula derivation only for the case of two variables (i.e., $r=2$). The proof for more variables is similar, though. A degree M polynomial of two variables is

$$P_M(x_1,x_2)=\sum_{d_1=0}^{M_1}\sum_{d_2=0}^{M_2}\beta_{d_1,d_2}x_1^{d_1}x_2^{d_2},$$

where $M=M_1+M_2$. From the last inequality, we have,

$$\begin{aligned}&\left|P_M\left(\frac{p_{1,m}}{n_1},\ldots,\frac{p_{2,m}}{n_r}\right)\sum_{i=0}^{2}a_ix_i-P_M(x_1,x_2)\sum_{i=0}^{2}a_i\frac{p_{i,m}}{n_i}\right|\\&=\left|\sum_{d_1=0}^{M_1}\sum_{d_2=0}^{M_2}\beta_{d_1,d_2}\left[\left(\frac{p_{1,m}}{n_1}\right)^{d_1}\left(\frac{p_{2,m}}{n_2}\right)^{d_2}\sum_{i=0}^{2}a_ix_i-x_1^{d_1}x_2^{d_2}\sum_{i=0}^{2}a_i\frac{p_{i,m}}{n_i}\right]\right|\\&\leqslant\sum_{d_1=0}^{M_1}\sum_{d_2=0}^{M_2}|\beta_{d_1,d_2}|\left|\left(\frac{p_{1,m}}{n_1}\right)^{d_1}\left(\frac{p_{2,m}}{n_2}\right)^{d_2}\sum_{i=0}^{2}a_ix_i-x_1^{d_1}x_2^{d_2}\sum_{i=0}^{2}a_i\frac{p_{i,m}}{n_i}\right|.\end{aligned}$$

Due to (11), we have

$$\frac{p_{i,m}}{n_i}=x_i-\frac{\theta_{i,m}}{n_i},$$

where $0\leqslant\theta_{i,m}\leqslant 1$. Consequently, the last expression above equals

$$\begin{aligned}&\sum_{d_1=0}^{M_1}\sum_{d_2=0}^{M_2}|\beta_{d_1,d_2}|\left|\left(x_1-\frac{\theta_{1,m}}{n_1}\right)^{d_1}\left(x_2-\frac{\theta_{2,m}}{n_2}\right)^{d_2}\sum_{i=0}^{2}a_ix_i-x_1^{d_1}x_2^{d_2}\sum_{i=0}^{2}a_i\left(x_i-\frac{\theta_{i,m}}{n_i}\right)\right|\\&=\sum_{d_1=0}^{M_1}\sum_{d_2=0}^{M_2}|\beta_{d_1,d_2}|\left|\left[\left(x_1-\frac{\theta_{1,m}}{n_1}\right)^{d_1}\left(x_2-\frac{\theta_{2,m}}{n_2}\right)^{d_2}-x_1^{d_1}x_2^{d_2}\right]\right.\end{aligned}$$

$$
\begin{aligned}
&\times \sum_{i=0}^{2} a_i x_i + x_1^{d_1} x_2^{d_2} \sum_{i=1}^{2} a_i \frac{\theta_{i,m}}{n_i} \Bigg| \\
&\leqslant \left| \sum_{i=1}^{2} a_i x_i \right| \sum_{d_1=0}^{M_1} \sum_{d_2=0}^{M_2} |\beta_{d_1,d_2}| \left| \left(x_1 - \frac{\theta_{1,m}}{n_1} \right)^{d_1} \left(x_2 - \frac{\theta_{2,m}}{n_2} \right)^{d_2} - x_1^{d_1} x_2^{d_2} \right| \\
&+ |x_1^{d_1} x_2^{d_2}| \sum_{d_1=0}^{M_1} \sum_{d_2=0}^{M_2} |\beta_{d_1,d_2}| \sum_{i=1}^{2} |a_i| \frac{\theta_{i,m}}{n_i} \\
&\leqslant \sum\nolimits_{\max} \sum_{d_1=0}^{M_1} \sum_{d_2=0}^{M_2} |\beta_{d_1,d_2}| \left| \left(x_1 - \frac{\theta_{1,m}}{n_1} \right)^{d_1} \left(x_2 - \frac{\theta_{2,m}}{n_2} \right)^{d_2} - x_1^{d_1} x_2^{d_2} \right| \\
&+ \sum_{i=1}^{2} \frac{|a_i|}{n_i} \sum_{d_1=0}^{M_1} \sum_{d_2=0}^{M_2} |\beta_{d_1,d_2}|, \qquad (12)
\end{aligned}
$$

where

$$
\sum\nolimits_{\max} \stackrel{\text{def}}{=} \max_{x \in C^r[0,1]} \left| \sum_{i=0}^{2} a_i x_i \right|.
$$

For the first part of the summation in the last expression in (12), note that

$$
\begin{aligned}
&\left| \left(x_1 - \frac{\theta_{1,m}}{n_1} \right)^{d_1} \left(x_2 - \frac{\theta_{2,m}}{n_2} \right)^{d_2} - x_1^{d_1} x_2^{d_2} \right| \\
&= \Bigg| x_2^{d_2} \left[\left(x_1 - \frac{\theta_{1,m}}{n_1} \right)^{d_1} - x_1^{d_1} \right] + x_1^{d_1} \left[\left(x_2 - \frac{\theta_{2,m}}{n_2} \right)^{d_2} - x_2^{d_2} \right] \\
&+ \left[\left(x_1 - \frac{\theta_{1,m}}{n_1} \right)^{d_1} - x_1^{d_1} \right] \left[\left(x_2 - \frac{\theta_{2,m}}{n_2} \right)^{d_2} - x_2^{d_2} \right] \Bigg| \\
&= \Bigg| x_2^{d_2} \left[- C_{d_1}^1 x_1^{d_1 - 1} \frac{\theta_{1,m}}{n_1} + \cdots + (-1)^{d_1} C_{d_1}^{d_1} \left(\frac{\theta_{1,m}}{n_1} \right)^{d_1} \right] \\
&+ x_1^{d_1} \left[- C_{d_2}^1 x_2^{d_2 - 1} \frac{\theta_{2,m}}{n_2} + \cdots + (-1)^{d_2} C_{d_2}^{d_2} \left(\frac{\theta_{2,m}}{n_2} \right)^{d_2} \right] \\
&+ \left[- C_{d_1}^1 x_1^{d_1 - 1} \frac{\theta_{1,m}}{n_1} + \cdots + (-1)^{d_1} C_{d_1}^{d_1} \left(\frac{\theta_{1,m}}{n_1} \right)^{d_1} \right] \\
&\times \left[- C_{d_2}^1 x_2^{d_2 - 1} \frac{\theta_{2,m}}{n_2} + \cdots + (-1)^{d_2} C_{d_2}^{d_2} \left(\frac{\theta_{2,m}}{n_2} \right)^{d_2} \right] \Bigg| \\
&\leqslant \frac{C_{d_1}^1}{n_1} + \frac{C_{d_1}^2}{n_1} + \cdots + \frac{C_{d_1}^{d_1}}{n_1} + \frac{C_{d_2}^1}{n_2} + \frac{C_{d_2}^2}{n_2} + \cdots + \frac{C_{d_2}^{d_2}}{n_2}
\end{aligned}
$$

$$+\left(\frac{C_{d_1}^1}{n_1}+\cdots+\frac{C_{d_1}^{d_1}}{n_1}\right)\left(\frac{C_{d_2}^1}{n_2}+\cdots+\frac{C_{d_2}^{d_2}}{n_2}\right)$$
$$=\frac{2^{d_1}-1}{n_1}+\frac{2^{d_2}-1}{n_2}+\frac{(2^{d_1}-1)(2^{d_2}-1)}{n_1 n_2}\leqslant\frac{2^{d_1+d_2}-1}{N},$$

where N is the smaller of n_1 and n_2. In the above derivation, we utilized the following relations:

$$|x_i^\eta|\leqslant 1 \quad \text{where} \quad \eta\geqslant 1, \qquad \left|\frac{\theta_{i,m}}{n_i}\right|\leqslant\frac{1}{n_i}, \qquad \frac{1}{n_i}\geqslant\frac{1}{n_i^\eta} \text{ for } \eta\geqslant 1,$$

$$\sum_{k=0}^{j} C_j^k = 2^j, \quad \text{where} \quad C_j^k=\frac{j!}{(j-k)!k!}.$$

As a result, the last expression in (12) is less than or equal to

$$\sum\nolimits_{\max}\sum_{d_1=0}^{M_1}\sum_{d_2=0}^{M_2}|\beta_{d_1,d_2}|\frac{2^{d_1+d_2}-1}{N}+\frac{|a_1|+|a_2|}{N}\sum_{d_1=0}^{M_1}\sum_{d_2=0}^{M_2}|\beta_{d_1,d_2}|<\varepsilon\sum\nolimits_{\min}.$$

Solving for N, we obtain the following formula

$$N>\frac{\sum_{\max}\sum_{d_1=0}^{M_1}\sum_{d_2=0}^{M_2}|\beta_{d_1,d_2}|(2^{d_1+d_2}-1)+(|a_1|+|a_2|)\sum_{d_1=0}^{M_1}\sum_{d_2=0}^{M_2}|\beta_{d_1,d_2}|}{\varepsilon\sum_{\min}}. \tag{13}$$

If N resulted from the formula is not an integer, we will use the integer that is larger than and closest to the computed N as n_1 and n_2. Now, we have found the N for the proof of the uniform approximation of the fuzzy systems to any continuous functions. □

Using the Weierstrass approximation theorem and Theorem 1, we now prove that the general fuzzy systems with our simplified linear TS rule consequent are universal approximators. In essence, the Weierstrass approximation theorem states that any multivariate continuous function in closed domain can always be uniformly approximated by multivariate polynomial no matter how small a pre-specified approximation error bound is.

Theorem 2 (Universal approximation theorem for the general MISO fuzzy systems with the simplified linear TS rule consequent). *The general MISO fuzzy systems with the simplified linear TS rule consequent can uniformly approximate any multivariate continuous function in closed domain to any degree of accuracy.*

Proof. Designate multivariate continuous function to be approximated as $G(\underline{x})$ and the desired uniform approximation error bound as ε. According to the

Weierstrass approximation theorem, we can always find a multivariate polynomial $P_M(\underline{x})$ that can uniformly approximate $G(\underline{x})$ with the desired accuracy ε. In order words, $\forall\varepsilon_1 > 0$, $\|P_M - G\| < \varepsilon_1$. Furthermore, on the basis of Theorem 1, $\forall\varepsilon_2 > 0$, $\|F_\Omega - P_M\| < \varepsilon_2$. Hence,

$$\|F_\Omega - G\| \leqslant \|P_M - G\| + \|F_\Omega - P_M\| < \varepsilon_1 + \varepsilon_2 = \varepsilon,$$

which means $F_\Omega(\underline{x})$ can uniformly approximate $G(\underline{x})$. □

Combining Theorems 1 and 2, we obtain the following quantitative result concerning sufficient conditions for the general TS fuzzy systems as universal approximators. Although the result is stated for two-variable functions, it obviously holds for functions with more variables. For simplicity and better presentation, we do not provide the more general results in this paper.

Theorem 3 (Sufficient Conditions). *Given a continuous function $G(x_1,x_2)$ to be approximated with uniform approximation error bound ε, inequality $|F_\Omega(x_1,x_2) - G(x_1,x_2)| < \varepsilon$ holds when the smaller of n_1 and n_2 is larger than or equal to N where*

$$N > \frac{\sum_{\max}\sum_{d_1=0}^{M_1}\sum_{d_2=0}^{M_2}|\beta_{d_1,d_2}|(2^{d_1+d_2}-1) + (|a_1|+|a_2|)\sum_{d_1=0}^{M_1}\sum_{d_2=0}^{M_2}|\beta_{d_1,d_2}|}{(\varepsilon-\varepsilon_1)\sum_{\min}}. \tag{14}$$

Here, ε_1 is so chosen that $|P_M(x_1,x_2) - G(x_1,x_2)| < \varepsilon_1$.

Proof. The formula was already derived in the proof of Theorem 1 (i.e. (13)). One just needs to replace ε in (13) with $\varepsilon - \varepsilon_1$ to yield (14). □

We point out that the values of the parameters/terms in the right side of inequality (14) are known for a given fuzzy approximation problem. Hence, formula (14) is readily to be used, at least in theory, for designing fuzzy systems as function approximators.

We now put these universal approximation results into the context of fuzzy controllers, models and filters. In a control or filtering problem, $G(\underline{x})$ represents a control or filtering solution whereas in a modeling problem, $G(\underline{x})$ is the ture mathematical representation of the physical system to be modeled. Practically speaking, in all these three types of problems, $G(\underline{x})$ is unknown and the task of the fuzzy system developer is to realize $G(\underline{x})$ or, when this is impossible, approximate it with enough high approximation accuracy by a TS fuzzy system through manipulation of its different components (input fuzzy sets, fuzzy rules, etc.) The manipulation is, in virtually all the cases, trial-and-error based, which is rather difficult if the original linear TS rule consequent is used, due to the large number of parameters in the rule consequent. (The general TS fuzzy systems with the original linear rule consequent are indeed universal approximators [22].) With the results presented in this paper, we have shown that, by using our simplified linear TS rule consequent, the manipulation process will not only be significantly easier but also any pre-specified approximation accuracy, no matter how small it is, will always be achievable in theory.

4. Conclusions

In the present paper, we have investigated a general class of MISO TS fuzzy systems that use any type of continuous input fuzzy sets, any type of fuzzy logic AND, nonlinear fuzzy rules with our newly introduced simplified linear TS rule consequent and the generalized defuzzifier containing the popular centroid defuzzifier as a special case. We have related the structure of the fuzzy systems to the structures of some popular nonfuzzy controllers, models and filters. We have proved that: (1) when implemented as controllers, the fuzzy systems are nonlinear time-varying controllers with variable gains, (2) when implemented as models, the fuzzy systems are nonlinear ARX models with time-varying model parameters, and (3) when implemented as filters, the fuzzy systems are nonlinear FIR or IIR filters with time-varying filter coefficients. We have constructively proved that the general MISO TS fuzzy systems can uniformly approximate any multivariate polynominal and then proved, by utilizing the Weierstrass approximation theorem, that the fuzzy systems can uniformly approximate any multivariate continuous function in closed domain. We have also derived a formula for calculating the number of input fuzzy sets and fuzzy rules needed to achieve pre-specified approximation accuracy.

Our results provide a solid theoretical basis for the use of our simplified linear TS rule consequent in the general fuzzy systems to form universal controllers, models and filters. The practical implication is that these TS fuzzy systems, with much less design parameters, are always able to produce solutions to various control, modeling and filtering problems.

Acknowledgements

This work was supported in part by a Biomedical Engineering Research Grant from The Whitaker Foundation and by grant no. 004952-054 from Texas Higher Education Coordinating Board.

The author is grateful for the referees' helpful comments that have led to the improvement of the presentation.

References

[1] S. Abe, M.-S. Lan, Fuzzy rules extraction directly from numerical data for function approximation, IEEE Transactions on Systems, Man, and Cybernetic 25 (1995) 119–129.
[2] I.N. Bronshtein, K.A. Semendyayev, Handbook of Mathematics, Van Nostrand Reinhold, New York, 1985.
[3] J.J. Buckley, Sugeno type controllers are universal controllers, Fuzzy Sets and Systems 53 (1993) 299–303.
[4] J.L. Castro, M. Delgado, Fuzzy systems with defuzzification are universal approximators, IEEE Transactions on Systems, Man and Cybernetic 26 (1996).
[5] J.A. Dickerson, B. Kosko, Fuzzy function approximation with ellipsoidal rules, IEEE Transactions on Systems, Man and Cybernetic 26 (4) (1996) 542–560.
[6] Y. Ding, H. Ying, S. Shao, Necessary Conditions on Minimal System Configuration for General MISO Fuzzy systems as Universal Approximators, Proceedings of 1997 IEEE SMC Conference, Orlando, FL, 1997.
[7] D.P. Filev, R.R. Yager, A generalized defuzzification method via BAD distributions, International Journal of Intelligent Systems 6 (1991) 687–697.
[8] V. Gorrini, T. Salome, H. Bersini, Self-structuring fuzzy systems for function approximation, Proceedings of 1995 IEEE International Conference on Fuzzy Systems, Yokohama, Japan, 1995, pp. 919–926.
[9] B. Kosko, Fuzzy systems as universal approximators, Proceedings of IEEE International Conference of Fuzzy Systems, San Diego, USA, 1992.
[10] E.G. Laukonen, K.M. Passino, Fuzzy systems for function approximation with applications to failure estimation, Proceedings of the 1994 IEEE International Symposium on Intelligent Control, Columbus, OH, USA, 1994, pp. 184–189.
[11] F.L. Lewis, S.-Q. Zhu, K. Liu, Function approximation by fuzzy systems, proceedings of the 1995 American Control Conference, vol. 5, Seattle, WA, USA, 1995, pp. 3760–3764.
[12] M.M. Gupta, J. Qi, Theory of T-norms and fuzzy inference methods, Fuzzy Sets and Systems 40 (1991) 431–450.
[13] M.M. Gupta, J. Qi, Design of fuzzy logic controllers based on generalized T-operators, Fuzzy Sets and Systems 40 (1991) 473–490.
[14] L. Ljung, System Identification: Theory for the User, Prentice-Hall, Englewood Cliffs, NJ, 1987.
[15] D.M. Mandelbaum. On cyclically time-varying autoregressive digital filters, Circuits, Systems and Signal Processing 16 (1997) 91–106.
[16] H.T. Nguyen, V. Kreinovich, O. Sirisaengtaksin, Fuzzy control as a universal control tool, Fuzzy Sets and Systems 80 (1) (1996) 71–86.
[17] T. Takagi, M. Sugeno, Fuzzy identification of systems and its applications to modeling and control, IEEE Transactions on Systems, Man and Cybernetics 15 (1985) 116–132.
[18] L.X. Wang, Fuzzy systems are universal approximators, Proceedings of IEEE International Conference of Fuzzy Systems, San Diego, USA, 1992.
[19] L.X. Wang, J.M. Mendel, Fuzzy basis functions, universal approximation, orthogonal least-squares learning, IEEE Transactions on Neural Networks 3 (1992) 807–814.
[20] H. Ying, Sufficient conditions on general fuzzy systems as function approximators, Automatica 30 (1994) 521–525.
[21] H. Ying, G.-R. Chen, Necessary conditions for some typical fuzzy systems as universal approximators, Automatica 33 (1997) 1333–1338.
[22] H. Ying, Sufficient conditions on uniform approximation of multivariate functions by general Takagi–Sugeno fuzzy systems with Linear rule consequent, IEEE Transactions on Systems, Man, and Cybernetic, in press.

[23] H. Ying, The Takagi–Sugeno fuzzy controllers using the simplified linear control rules are nonlinear variable gain controllers, Automatica (in press).
[24] H. Ying, Constructing nonlinear variable gain controllers via the Takagi–Sugeno fuzzy control, IEEE Transactions on Fuzzy Systems (in press).
[25] X.-J. Zeng, M.G. Singh, Approximation theory of fuzzy systems – MIMO case, IEEE Transactions of Fuzzy Systems 3 (1995) 219–235.
[26] X.-J. Zeng, M.G. Singh, Approximation accuracy analysis of fuzzy systems as function approximators, IEEE Transactions of Fuzzy Systems 4 (1996) 44–63.