Letters

General SISO Takagi–Sugeno Fuzzy Systems with Linear Rule Consequent Are Universal Approximators

Hao Ying

Abstract— **Takagi–Sugeno (TS) fuzzy systems have been employed as fuzzy controllers and fuzzy models in successfully solving difficult control and modeling problems in practice. Virtually all the TS fuzzy systems use linear rule consequent. At present, there exist no results (qualitative or quantitative) to answer the fundamentally important question that is especially critical to TS fuzzy systems as fuzzy controllers and models, "Are TS fuzzy systems with linear rule consequent universal approximators?" If the answer is yes, then how can they be constructed to achieve prespecified approximation accuracy and what are the sufficient conditions on systems configuration? In this paper, we provide answers to these questions for a general class of single-input single-output (SISO) fuzzy systems that use any type of continuous input fuzzy sets, TS fuzzy rules with linear consequent and a generalized defuzzifier containing the widely used centroid defuzzifier as a special case. We first constructively prove that this general class of SISO TS fuzzy systems can uniformly approximate any polynomial arbitrarily well and then prove, by utilizing the Weierstrass approximation theorem, that the general TS fuzzy systems can uniformly approximate any continuous function with arbitrarily high precision. Furthermore, we have derived a formula as part of sufficient conditions for the fuzzy approximation that can compute the minimal upper bound on the number of input fuzzy sets and rules needed for any given continuous function and prespecified approximation error bound. An illustrative numerical example is provided.**

Index Terms— **Functional approximation, fuzzy control, fuzzy modeling, fuzzy systems, Takagi–Sugeno fuzzy systems, and universal approximators.**

I. Introduction

THE issue of fuzzy systems as universal approximators has drawn a considerable amount of attention in the past few years (e.g., [1], [4], [5], [7], [9]–[11], [13], [15], [16], [23]) and is currently an active research area. The basic question is, "Can fuzzy systems approximate any real continuous function to any degree of accuracy on a compact domain (closed and bounded in a finite dimensional space)?" The answer to this question is indeed of both theoretical and practical importance and will have significant implications for the applications of fuzzy systems, especially as fuzzy controllers and fuzzy models. From a mathematical point of view, fuzzy systems are functions mapping their inputs to outputs. In the context of fuzzy control, this approximation question asks whether or not a fuzzy controller can be constructed to approximate any continuous nonlinear control solution, whereas in the context of modeling, the question of interest is whether or not a fuzzy model is capable of approximating any physical dynamic model that is continuous and nonlinear.

There are two major types of fuzzy systems: Mamdani fuzzy systems and Takagi–Sugeno (TS) fuzzy systems. The main difference lies in the consequent of fuzzy rules. Mamdani fuzzy systems use fuzzy sets as rule consequent whereas TS fuzzy systems employ linear functions of input variables as rule consequent [14]. All the existing results on fuzzy systems as universal approximators deal with Mamdani fuzzy systems only and no result is available for TS fuzzy systems with linear rule consequent. The existence result in [3] is interesting, but it is for an uncommon two-input single-output TS fuzzy system. The system uses a linear defuzzifier (i.e., the defuzzifier does not have a denominator) instead of the widely used centroid defuzzifier and the rule consequent is nonlinear. Furthermore, the majority of the existing fuzzy approximation results are proved by using the Stone–Weierstrass theorem and, thus, are just existence results on some specific types of Mamdani fuzzy systems. So far, only two papers (i.e., [18], [22]) are on necessary conditions for fuzzy approximation and very few papers (e.g., [17], [24]) are about quantitative aspects of fuzzy approximation, with [17] being the first one.

The important questions regarding TS fuzzy systems as universal approximators include: 1) are TS fuzzy systems with *linear* rule consequent universal approximators and 2) what are the sufficient conditions for TS fuzzy systems with linear rule consequent as universal approximators? The second question can also be asked in another way, "Given a continuous function and a prespecified approximation accuracy, how many input fuzzy sets and rules are needed to guarantee the desired approximation accuracy?"

In this paper, we provide answers to both the questions for a general class of SISO TS fuzzy systems with linear rule consequent. The TS fuzzy systems in our study are general because they use any type of continuous input fuzzy sets, TS rules with linear consequent and the generalized defuzzifier containing the centroid defuzzifier as a special case. Our proof approach is constructive and we complete the proof in two steps. In the first step, we prove that this general class of SISO fuzzy systems can uniformly approximate any polynomial

Manuscript received March 7, 1997; revised May 15, 1998. This work was supported in part by a Biomedical Engineering Research Grant from the Whitaker Foundation and by Grant 004952-054 from the Texas Higher Education Coordinating Board.

The author is with the Department of Physiology and Biophysics, Biomedical Engineering Center, The University of Texas Medical Branch, Galveston, TX 77555 USA.

Publisher Item Identifier S 1063-6706(98)08258-7.

arbitrarily well. In the second step, we utilize the Weierstrass approximation theorem and prove that the general TS fuzzy systems can uniformly approximate any continuous function with arbitrarily high precision. We have also derived a formula that can be used to compute the minimal upper bound on the number of input fuzzy sets and rules needed for any given continuous function and prespecified approximation accuracy. Finally, we give a numerical example to illustrate how to use the formula.

II. General SISO TS Fuzzy Systems with Linear Rule Consequent

The general TS fuzzy systems in this investigation uses a continuous-time or discrete-time input variable $x(t)$, where t is time. Throughout this paper all functions, variables, and parameters are real and all the functions are defined on real closed intervals. For notational simplicity, from now on we use x instead of $x(t)$ and, without loss of generality, we assume $-1 \le x \le 1$.

To fuzzify x, $N = 2n + 1$ $(n \ge 1)$ fuzzy sets are used. Each fuzzy set, denoted by A_j $(j = 0, \pm 1, \cdots, \pm n)$, has a continuous membership function μ_j, which can be any shape. By "any shape," we mean all reasonable continuous membership functions, including the popular ones such as triangles, trapezoids, and normal distributions. Of course, by definition, irregular or irrational membership functions, as long as they are continuous, are also inevitably included. But this does not adversely affect the results in this paper at all. We can assume arbitrary membership functions here because the results that we have established, as will be shown later, are independent of the shape of the membership functions. The positive, negative, and zero subscripts j may be interpreted as linguistic description of "positive," "negative," and "zero" (e.g., A_1 stands for "positive very small" and A_{-5} means "negative large"). We partition $[-1, 1]$ into $2n$ equal intervals, each of which is $[j/n, (j+1)/n]$. Of the $2n+1$ fuzzy sets, one is defined over $[-1, -(n-1)/n]$, another over $[(n-1)/n, 1]$ and each of the remaining ones over $[(k-1)/n, (k+1)/n]$ where $-(n-1) \le k \le n-1$.

Thus, N TS fuzzy rules are used to cover all A_j. Recall that a TS fuzzy rule with linear consequent [14] is expressed as

$$\text{Rule \#m:} \quad \text{IF } x \text{ is } A_{p_m} \text{ THEN } y = a_m + b_m x \tag{1}$$

where y is output of the fuzzy systems, p_m is an integer satisfying $-n \le p_m \le n$, and a_m and b_m are design parameters whose values are determined by the fuzzy system developer. There are a total of $2N$ design parameters for N rules. In different rules, the values of these parameters may be different.

To significantly reduce the number of design parameters, we have recently proposed a simplified linear TS rule consequent [19]–[21], which is expressed as

$$\begin{array}{ll} \text{Rule \#1:} & \text{IF } x \text{ is } A_{p_1} \text{ THEN } y = k_1(a + bx) \\ & \vdots \\ \text{Rule \#}m\text{:} & \text{IF } x \text{ is } A_{p_m} \text{ THEN } y = k_m(a + bx) \\ & \vdots \\ \text{Rule \#}N\text{:} & \text{IF } x \text{ is } A_{p_N} \text{ THEN } y = k_N(a + bx) \end{array} \tag{2}$$

where a, b, $k_1, \cdots, k_N$ are $N + 2$ design parameters that can be any value chosen by the developer. One sees that all the rule consequent use the same linear function $a+bx$ and all the rules are proportional to each other. The rule proportionality represented by k_m for rule #m can be different from one rule to another.

The major advantage of the simplified TS rule consequent over the original one is the significant reduction in the number of design parameters. The reduction is by a factor of $(N - 2)/2N$ which is almost 50% for larger N. For TS fuzzy systems with more than one input variable, the reduction is even greater; the more the number of input variables, the greater the reduction. To illustrate the extent of the reduction, let us take a SISO fuzzy system that uses ten fuzzy rules (i.e., $N = 10$) as an example. The original TS rule consequent will require 20 parameters whereas the simplified consequent only 12 (i.e., a 40% reduction). This kind of parameter reduction is not only desirable, but also necessary in many practical applications, especially control applications (recall that the PID controller is so popular, dominant, and practically useful primarily because it has only three parameters to tune and, hence, is very easy to use). We have shown, through rigorous mathematical analysis, that the TS fuzzy PID controllers using the simplified consequent not only have a far smaller number of design parameters, but also possess some desirable variable gain characteristics [19]–[21]. Since the TS fuzzy systems with the simplified rule consequent are universal approximators (as will be demonstrated later), TS fuzzy systems using them are more practically useful.

We should point out that the simplified linear TS rule consequent is a special case of the original linear TS rule consequent. According to (1) and (2), the former becomes the latter when for $m = 1, \cdots, N$,

$$a_m = k_m a$$
$$b_m = k_m b.$$

To make our approximation results general in the rest of the paper, we will focus on the fuzzy systems using the original linear TS rule consequent and establish approximation conditions for them. Apparently, the results will hold for the TS fuzzy systems using the simplified linear consequent as a special case.

The generalized defuzzifier [6] is used to calculate y, which actually is a mapping F_n: $[-1, 1] \to (-\infty, \infty)$

$$F_n(x) \stackrel{\text{def}}{=} y = \frac{\sum_{m=1}^{N} (\mu_m)^\alpha \cdot (a_m + b_m x)}{\sum_{m=1}^{N} (\mu_m)^\alpha} \tag{3}$$

where μ_m is the membership for the rule consequent in rule #m. Different defuzzification results can be obtained by using different α values where $0 \le \alpha < +\infty$. The most widely used centroid defuzzifier is a special case of this generalized defuzzifier when $\alpha = 1$, and the popular mean of maximum defuzzifier is another special case when $\alpha = \infty$. $F_n(x)$ is a function sequence with respect to n (e.g., [12]). From now on, $F_n(x)$ will always be used to represent the general SISO TS fuzzy systems.

III. The General TS Fuzzy Systems Are Universal Approximators

In this section, we will first constructively prove that the general fuzzy systems with linear TS rule consequent can uniformly approximate any polynomial to any degree of accuracy. We will then utilize the fact that any continuous function defined on a closed interval can always be approximated by a polynomial arbitrarily well to prove that the general TS fuzzy systems can uniformly approximate any continuous function with arbitrary precision.

Theorem 1: $F_n(x)$ can uniformly approximate, with arbitrarily high precision, any polynomial $P_h(x)$ defined on $[-1,1]$

$$P_h(x) = \sum_{i=0}^{h} \beta_i x^i$$

where h is the order of the polynomial.

Proof: We will use the constructive proof approach. Since $P_h(x)$ is explicitly given, we can use it to construct N TS fuzzy rules as follows:

$$\begin{aligned}
&\text{Rule \#1:} \quad \text{IF } x \text{ is } A_{p_1} \text{ THEN} \\
&\qquad y = P_h\left(\frac{p_1}{n}\right) - \beta_1 \frac{p_1}{n} + \beta_1 x \\
&\qquad \vdots \\
&\text{Rule \#}m\text{:} \quad \text{IF } x \text{ is } A_{p_m} \text{ THEN} \\
&\qquad y = P_h\left(\frac{p_m}{n}\right) - \beta_1 \frac{p_m}{n} + \beta_1 x \\
&\qquad \vdots \\
&\text{Rule \#}N\text{:} \quad \text{IF } x \text{ is } A_{p_N} \text{ THEN} \\
&\qquad y = P_h\left(\frac{p_N}{n}\right) - \beta_1 \frac{p_N}{n} + \beta_1 x.
\end{aligned} \tag{4}$$

In other words, in the fuzzy rules of the form (1), we let

$$\begin{aligned}
a_m &= P_h\left(\frac{p_m}{n}\right) - \beta_1 \frac{p_m}{n} \\
b_m &= \beta_1
\end{aligned}$$

where $m = 1, \cdots, N$. Note that a_m and b_m are still constants. Using (4), we obtain

$$F_n(x) = \frac{\sum_{m=1}^{N} (\mu_m)^\alpha \left[P_h\left(\frac{p_m}{n}\right) - \beta_1 \frac{p_m}{n} + \beta_1 x\right]}{\sum_{m=1}^{N} (\mu_m)^\alpha}.$$

Because at any time

$$\frac{p_m}{n} \le x \le \frac{p_m + 1}{n} \tag{5}$$

the following inequality holds:

$$\lim_{n\to\infty} \frac{p_m}{n} \le x \le \lim_{n\to\infty} \frac{p_m + 1}{n}.$$

Thus,

$$\lim_{n\to\infty} \frac{p_m}{n} = \lim_{n\to\infty} \frac{p_m + 1}{n} = x$$

and, consequently,

$$\lim_{n\to\infty} P_h\left(\frac{p_m}{n}\right) = P_h(x).$$

Therefore,

$$\begin{aligned}
\lim_{n\to\infty} F_n(x) &= \lim_{n\to\infty} \frac{\sum_{m=1}^{N} (\mu_m)^\alpha \left[P_h\left(\frac{p_m}{n}\right) - \beta_1 \frac{p_m}{n} + \beta_1 x\right]}{\sum_{m=1}^{N} (\mu_m)^\alpha} \\
&= P_h(x)
\end{aligned}$$

which means $F_n(x)$ approaches $P_h(x)$ as $n \to \infty$.

Now we need to prove that the approximation is uniform. To accomplish this, we will derive a formula that can be used to calculate a positive integer n^* based on a given function and approximation error bound $\varepsilon > 0$ such that $\forall n > n^*$

$$\|F_n - P_h\|_{[-1,1]} = \max_{x\in[-1,1]} |F_n(x) - P_h(x)| < \varepsilon.$$

We want

$$\begin{aligned}
&\max_{x\in[-1,1]} \left| \frac{\sum_{m=1}^{N} (\mu_m)^\alpha \left[P_h\left(\frac{p_m}{n}\right) - \beta_1 \frac{p_m}{n} + \beta_1 x\right]}{\sum_{m=1}^{N} (\mu_m)^\alpha} - P_h(x) \right| \\
&= \max_{x\in[-1,1]} \left| \frac{\sum_{m=1}^{N} (\mu_m)^\alpha \left[P_h\left(\frac{p_m}{n}\right) - \beta_1 \frac{p_m}{n} + \beta_1 x - P_h(x)\right]}{\sum_{m=1}^{N} (\mu_m)^\alpha} \right| \\
&< \varepsilon
\end{aligned}$$

which implies that we want

$$\begin{aligned}
&\left| P_h\left(\frac{p_m}{n}\right) - \beta_1 \frac{p_m}{n} + \beta_1 x - P_h(x) \right| \\
&\quad \le \left| P_h\left(\frac{p_m}{n}\right) - P_h(x) \right| + \left| \beta_1 \frac{p_m}{n} - \beta_1 x \right| \\
&\quad = \left| \sum_{i=0}^{h} \beta_i \left[\left(\frac{p_m}{n}\right)^i - x^i \right] \right| + |\beta_1| \left| \frac{p_m}{n} - x \right| \\
&\quad \le \sum_{i=0}^{h} |\beta_i| \left| \left(\frac{p_m}{n}\right)^i - x^i \right| + |\beta_1| \left| \frac{p_m}{n} - x \right| < \varepsilon.
\end{aligned} \tag{6}$$

Because of (5),

$$\left| \frac{p_m}{n} - x \right| \le \frac{1}{n}. \tag{7}$$

Hence, for the second term in the last part of inequality (6),

$$|\beta_1| \left| \frac{p_m}{n} - x \right| \le \frac{|\beta_1|}{n}.$$

For the first term in the last part of inequality (6), we have (e.g., [8])

$$\begin{aligned}
&\sum_{i=0}^{h} |\beta_i| \left| \left(\frac{p_m}{n}\right)^i - x^i \right| \\
&\quad = \sum_{i=0}^{h} |\beta_i| \left| \frac{p_m}{n} - x \right| \left| \binom{i}{1} \left(\frac{p_m}{n}\right) x^{i-1} + \cdots \right. \\
&\qquad \left. + \binom{i}{i-1} \left(\frac{p_m}{n}\right)^{i-1} x + \left(\frac{p_m}{n}\right)^i \right| \\
&\quad \le \frac{1}{n} \sum_{i=0}^{h} |\beta_i| \left[\binom{i}{1} + \cdots + \binom{i}{i-1} + 1 \right] \\
&\quad = \frac{1}{n} \sum_{i=0}^{h} |\beta_i| (2^i - 1) = \frac{1}{n} \sum_{i=1}^{h} |\beta_i| (2^i - 1).
\end{aligned}$$

In the above derivation, we utilized (7) as well as the following:

$$|x| \le 1, \quad \left| \frac{p_m}{n} \right| \le 1 \quad \text{and} \quad \sum_{k=0}^{i} \binom{i}{k} = 2^i.$$

So, the last part of inequality (6) becomes

$$\frac{1}{n^*}\sum_{i=1}^{h}|\beta_i|(2^i-1)+\frac{|\beta_1|}{n^*}<\varepsilon$$

we obtain

$$n^* > \frac{|\beta_1|+\sum_{i=1}^{h}|\beta_i|(2^i-1)}{\varepsilon}. \tag{8}$$

Having derived this formula, we have actually completed the task of proving that the general TS fuzzy systems can approximate any polynomials uniformly with arbitrarily high accuracy. □

Utilizing the Weierstrass approximation theorem [2], we now prove that the general TS fuzzy systems are universal approximators. The Weierstrass approximation theorem in essence states: To any continuous function $G(x)$ on a closed interval, given approximation error bound $\varepsilon > 0$, there always exists a polynomial that can approximate $G(x)$ uniformly with the desired accuracy. Generally, the smaller the ε, the higher the polynomial degree.

Theorem 2 (Universal Approximation Theorem for the General SISO TS Fuzzy Systems): The general TS fuzzy systems with linear rule consequent can uniformly approximate any continuous function on a closed interval to any degree of accuracy.

Proof: According to the Weierstrass approximation theorem, polynomial $P_h(x)$ can uniformly approximate $G(x)$ with arbitrary accuracy. That is, $\forall\varepsilon_1 > 0$, $\|P_h - G\| < \varepsilon_1$. Further, on the basis of Theorem 1, $\forall\varepsilon_2 > 0$, $\|F_n - P_h\| < \varepsilon_2$. Hence,

$$\|F_n - G\| \le \|F_n - P_h\| + \|P_h - G\| = \varepsilon, \quad \varepsilon = \varepsilon_1 + \varepsilon_2$$

meaning $F_n(x)$ can uniformly approximate $G(x)$ arbitrarily well. □

Based on the two theorems just proved, one sees that the general TS fuzzy systems with linear rule consequent can uniformly approximate any continuous function because they have the ability to uniformly approximate any polynomial. Combining these two theorems, we get the following quantitative theorem concerning sufficient conditions for the general TS fuzzy systems as universal approximators.

Theorem 3 (Sufficient Conditions for the Universal Approximation): Given an approximation error bound $\varepsilon > 0$, let $\|P_h - G\| < \varepsilon_1$ and $\|F_n - P_h\| < \varepsilon_2$, where $\varepsilon = \varepsilon_1 + \varepsilon_2$. $\|F_n - G\| < \varepsilon$ if n is such chosen that $n > n^*$ where

$$n^* > \frac{|\beta_1|+\sum_{i=1}^{h}|\beta_i|(2^i-1)}{\varepsilon_2}. \tag{9}$$

Proof: This formula was already derived in the proof of Theorem 1 [i.e., (8)]. One just needs to replace ε in (8) with ε_2 to yield this formula. □

Note that formula (9) holds for both the fuzzy systems with the original linear TS rule consequent and the fuzzy systems with the simplified linear TS rule consequent. For a given function and approximation accuracy, the same n^* will be computed by the formula and so is N. This fact further verifies the parameter reduction (i.e., $N+2$ vs. $2N$) of our simplified rule consequent, as described earlier.

In theory, we can use any integer larger than the calculated n^* as n^* to achieve the uniform approximation. But in order to use as few fuzzy sets and rules as possible in practice, one should always use the computed n^* if it is an integer or the smallest integer larger than the calculated n^* if it is not an integer. Accordingly, the number of fuzzy sets and rules is $N = 2n^* + 1$.

We should point out that formula (9) represents sufficient conditions rather than necessary conditions or necessary and sufficient conditions. As such, the n^* calculated is likely to be somewhat conservative (i.e., too large.) Another factor causing the conservatism is the independence of the formula on both the membership functions of the fuzzy sets and the defuzzifier type (i.e., value of α.) The formula can always be used to determine the number of the fuzzy sets and rules needed, regardless of the shape of the membership functions and type of the defuzzifier used. Given the general applicability of the formula, overestimation is only natural, logical, and inevitable.

The formula is particularly of theoretical value because in traditional function approximation theory quantitative results are the norm. As always, numerical results are far more precise and important than qualitative statements. Prior to the results in the present paper, there did not even exist any qualitative approximation results on TS fuzzy systems with linear rule consequent, let alone quantitative ones.

The formula can be useful when the minimal upper bound of the number of fuzzy sets and rules is needed during the design of fuzzy systems. Theoretically speaking, any integer larger than n^* is an upper bound and, thus, there are infinite number of upper bounds. Practically, through, the computed n^* is the only sensible upper bound because it is the smallest. We call it the minimal upper bound to stress that it is the smallest upper bound relative to all the other upper bounds. At present, it is our intention to use the term "minimal upper bound" only in the context of the present paper.

One may wonder whether a lower bound on the number of fuzzy sets and rules needed can be estimated. Indeed, this can be done and the reader is referred to our recent paper [22] in which we have established necessary conditions for the general MISO (multiple-input single-output) fuzzy systems with linear TS rule consequent as universal approximators. The general MISO fuzzy systems contain the general SISO fuzzy systems studied in present paper as a special case. The number of fuzzy sets and rules really needed to approximate a given function with a prespecified accuracy is between the lower bound and upper bound and depends on the function and the accuracy.

According to (9), it appears to be desirable to use a larger ε_2 and a smaller h in order to yield a smaller n^*. However, in general, these goals cannot be achieved simultaneously. This is because for a fixed ε, a larger ε_2 means a smaller ε_1, which, in turn, results in a larger h. When ε is very small, both ε_1 and ε_2 are very small while h is very large, making n^* extremely large. As an extreme, if $\varepsilon = 0$, n^* approaches ∞.

In a fuzzy control problem, $G(x)$ represents a desired nonlinear control solution whereas in a fuzzy modeling problem $G(x)$ represents the true nonlinear dynamic model of the system to be modeled. In practice, $G(x)$ is unknown to the fuzzy system developer. The task of the developer is to realize

$G(x)$ or, when this is impossible, approximate it with enough accuracy by a fuzzy system through manipulation of different components of the fuzzy system.

According to (3), the general TS fuzzy systems can be expressed as

$$F_n(x) = \varphi(x)x + \xi(x)$$

where

$$\varphi(x) = \frac{\sum_{m=1}^{N} (\mu_m)^\alpha \cdot b_m}{\sum_{m=1}^{N} (\mu_m)^\alpha}$$

and

$$\xi(x) = \frac{\sum_{m=1}^{N} (\mu_m)^\alpha \cdot a_m}{\sum_{m=1}^{N} (\mu_m)^\alpha}.$$

Hence, when used as a fuzzy controller, $F_n(x)$ is actually a nonlinear controller with time-varying gains changing with x. The gain is $\varphi(x)$ for x whereas $\xi(x)$ may be regarded as a time-varying control offset. Apparently, the general fuzzy systems with our simplified linear TS rule consequent are also nonlinear controllers with variable gains. Thus, the simplified linear rule consequent is not only general because the general TS fuzzy systems using it are universal approximators, but also desirable from the viewpoint of nonlinear control theory. For more detailed analysis of the fuzzy controllers using the simplified rule consequent, the reader is referred to our recent papers [19]–[21].

IV. Illustrative Example

To demonstrate how to use (9) to compute the minimal upper bound on the number of fuzzy sets and rules, we provide the following numerical example.

Example: How many fuzzy sets and rules are needed for a general TS fuzzy system using linear rule consequent to uniformly approximate function $G(x) = e^x$ defined on $[-1, 1]$ with (1) $\varepsilon = 0.1$ and (2) $\varepsilon = 0.01$?

Solution: It can easily be shown that the following polynomial of degree five can uniformly approximate the given $G(x)$ with truncation error less than 0.0038 (i.e., we can choose $\varepsilon_1 = 0.0038$):

$$e^x \approx P_5(x) = 1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \frac{x^4}{24} + \frac{x^5}{120}.$$

Hence, $h = 5$, $\beta_1 = 1$, $\beta_2 = 1/2$, $\beta_3 = 1/6$, $\beta_4 = 1/24$, and $\beta_5 = 1/120$. For $\varepsilon = 0.1$, $\varepsilon_2 = 0.1 - \varepsilon_1 = 0.0962$ whereas for $\varepsilon = 0.01$, $\varepsilon_2 = 0.01 - \varepsilon_1 = 0.0062$. Using (9), one can easily compute $n^* = 57.7$ for the larger ε and $n^* = 895.2$ for the smaller ε. Hence, $n^* = 58$ and 896, respectively. Consequently, 117 and 1793 fuzzy sets/rules are needed, respectively. As expected, the smaller the approximation error bound, the greater the number of fuzzy sets and rules required.

We point out again that the example serves the purpose of showing how to use (9). The computed n^* may be conservative because its calculation is unrelated to the membership functions of the fuzzy sets and the defuzzifier type. Nevertheless, if 58 and 896 fuzzy sets are used, the respective desired approximation accuracy of 0.01 and 0.001 will be guaranteed. It would be a mistake if one would conclude, from this numerical example, that a large number of fuzzy sets and rules are *necessary* for the fuzzy systems to approximate simple functions like e^x. This is not true, as the computed n^* only represents sufficient conditions, not necessary ones or necessary and sufficient ones, as we have pointed out earlier.

According to the necessary conditions that we have recently established for the general MISO TS fuzzy systems [22], only a few fuzzy sets and rules *may* be sufficient to approximate simple functions such as e^x with high approximation accuracy. In other words, the lower bound is a small number. The number really required depends on the membership functions of the fuzzy sets and the defuzzifier type and is between a few and the n^* computed above. At present, no theory exists that could produce an exact number and the trial-and-error method seems to be the only means available for experimental determination of its value.

V. Conclusion

We have constructively proved that the general SISO fuzzy systems using linear TS rule consequent, including our simplified one as a special case, can uniformly approximate any polynomial on a closed interval arbitrarily well. We have also proved, by utilizing the Weierstrass approximation theorem that the general TS fuzzy systems can uniformly approximate any continuous function on a closed interval as accurately as one desires. We have derived a formula that can be used to compute the minimal upper bound on the number of input fuzzy sets and rules needed for any given continuous function and approximation accuracy. Since the formula is unrelated to the membership functions of the fuzzy sets and defuzzifier type, the minimal upper bound may be conservative. Deriving a formula that would relate the upper bound to the membership functions of the fuzzy sets and the defuzzifier type is desirable as it could provide a less conservative (i.e., smaller) upper bound. This task is technically challenging, however.

The qualitative and quantitative results in the present paper are important not only because they are the first ones in existence, but also because they provide much needed theoretical support for using TS fuzzy systems in various fields, especially as controllers and models.

Acknowledgment

The author would like to thank the anonymous referees for their helpful comments that have led to the better presentation of this paper.

References

[1] S. Abe and M.-S. Lan, “Fuzzy rules extraction directly from numerical data for function approximation,” *IEEE Trans. Syst., Man, Cybern.*, vol. 25, pp. 119–129, 1995.
[2] I. N. Bronshtein and K. A. Semendyayev, *Handbook of Mathematics*. New York: Van Nostrand Reinhold, 1985.
[3] J. J. Buckley, “Sugeno type controllers are universal controllers,” *Fuzzy Sets Syst.*, vol. 53, pp. 299–303, 1993.
[4] J. L. Castro and M. Delgado, “Fuzzy systems with defuzzification are universal approximators,” *IEEE Trans. Syst., Man, Cybern.*, vol. 26, pp. 149–152, 1996.
[5] J. A. Dickerson and B. Kosko, “Fuzzy function approximation with ellipsoidal rules,” *IEEE Trans. Syst., Man, Cybern.*, vol. 26, pp. 542–560, 1996.

[6] D. P. Filev and R. R. Yager, "A generalized defuzzification method via BAD distributions," *Int. J. Intell. Syst.*, vol. 6, pp. 687–697, 1991.

[7] V. Gorrini, T. Salome, and H. Bersini, "Self-structuring fuzzy systems for function approximation," in *Proc. 1995 IEEE Int. Conf. Fuzzy Syst.*, Yokohama, Japan, Mar. 1995, pp. 919–926.

[8] K. Knopp, *Theory and Applications of Infinite Series*. New York: Dover, 1990.

[9] B. Kosko, "Fuzzy systems as universal approximators," in *Proc. IEEE Int. Conf. Fuzzy Syst.*, San Diego, CA, Mar. 1992, pp. 1153–1162.

[10] E. G. Laukonen and K. M. Passino, "Fuzzy systems for function approximation with applications to failure estimation," in *Proc. 1994 IEEE Int. Symp. Intell. Contr.*, Columbus, OH, Aug. 1994, pp. 184–189.

[11] F. L. Lewis, S.-Q. Zhu, and K. Liu, "Function approximation by fuzzy systems," in *Proc. 1995 Amer. Contr. Conf.*, Seattle, WA, June 1995, pp. 3760–3764.

[12] A. Moser, "Extending the domain of definition of functional series for nonlinear systems," *Automatica*, vol. 32, pp. 1233–1234, 1996.

[13] H. T. Nguyen, V. Kreinovich, and O. Sirisaengtaksin, "Fuzzy control as a universal control tool," *Fuzzy Sets Syst.*, vol. 80, pp. 71–86, 1996.

[14] T. Takagi and M. Sugeno, "Fuzzy identification of systems and its applications to modeling and control," *IEEE Trans. Syst., Man, Cybern.*, vol. 15, pp. 116–132, 1985.

[15] L.-X. Wang, "Fuzzy systems are universal approximators," in *Proc. IEEE Int. Conf. Fuzzy Syst.*, San Diego, CA, Mar. 1992.

[16] L.-X. Wang and J. M. Mendel, "Fuzzy basis functions, universal approximation, and orthogonal least-squares learning," *IEEE Trans. Neural Networks*, vol. 3, pp. 807-814, 1992.

[17] H. Ying, "Sufficient conditions on general fuzzy systems as function approximators," *Automatica*, vol. 30, pp. 521–525, 1994.

[18] H. Ying and G.-R. Chen, "Necessary conditions for some typical fuzzy systems as universal approximators," *Automatica*, vol. 33, pp. 1333–1338, 1997.

[19] H. Ying, "The Takagi-Sugeno fuzzy controllers using the simplified linear control rules are nonlinear variable gain controllers," *Automatica*, vol. 34, pp. 157–167, 1998.

[20] H. Ying, "Constructing nonlinear variable gain controllers via the Takagi-Sugeno fuzzy control," *IEEE Trans. Fuzzy Syst.*, vol. 6, pp. 226–234, 1998.

[21] H. Ying and L. C. Sheppard, "Analytical structure of Takagi-Sugeno fuzzy PID controllers and their applications in control of mean arterial pressure in patients," in *Proc. 5th Eur. Congress Intell. Tech. Soft Comput.*, Aachen, Germany, Sept. 1997, pp. 1374–1378.

[22] H. Ying, Y.-S. Ding, S.-K. Li, and S.-H. Shao, "Comparison of necessary conditions for typical Takagi–Sugeno and Mamdani fuzzy systems as universal approximators," to be published.

[23] X.-J. Zeng and M. G. Singh, "Approximation theory of fuzzy systems—MIMO case," *IEEE Trans. Fuzzy Syst.*, vol. 3, pp. 219–235, 1995.

[24] X.-J. Zeng and M. G. Singh, "Approximation accuracy analysis of fuzzy systems as function approximators," *IEEE Trans. Fuzzy Syst.*, vol. 4, pp. 44–63, 1996.