motions, which would be typical for industrial settings: i) motion along 3-D straight lines, and ii) general planar motion. The proposed conjectures were shown to be reliable via experimentation with real object surfaces that were marked with artificial circular markers.

The second solution approach, for 3-D general motion, advocates the use of additional features, which are coplanar to the circular features monitored. Based on the assumption that the marked object is a rigid body, the distances from a circular feature to such additional features, though unknown a priori, can be used as invariants to solve the duality problem. Experiments also showed that these methods are effective in practice.

Ill-conditioned positions do exist for all solutions proposed in this paper. However, they can be found by solving a set of linear and quadratic equations, and normally eliminated during run-time by acquiring a third image of the moving object.

The solution methods outlined herein were successfully utilized within our active-vision based moving-object recognition scheme [25].

References

[1] S. D. Ma, "Conics-based stereo, motion estimation, and pose determination," *Int. J. Comput. Vision*, vol. 10, no. 1, pp. 7–25, 1993.

[2] M. Yamamoto and K. Ikeda, "Stereoscopic correcting pose of 3-D model of object," *Time-Varying Image Processing and Moving Object Recognition*, V. Cappelini, Ed. Amsterdam, The Netherlands: Elsevier, 1994, vol. 3, pp. 127–134.

[3] T. S. Huang and A. N. Netravali, "Motion and structure from feature correspondence: A review," *Proc. IEEE*, vol. 82, pp. 252–268, Feb. 1994.

[4] Y. C. Liu and T. S. Huang, "Estimation of rigid body motion using straight line correspondences," in *IEEE Workshop Motion*, Kiawah Island, SC, 1986, pp. 47–52.

[5] M. E. Spetsakis and J. Aloimonos, "Close form solutions to the structure from motion problem from line correspondence," in *Proc. AAAI*, Seattle, WA, 1987, pp. 738–743.

[6] J. K. Aggarwal and Y. F. Wang, "Analysis of a sequence of images using point and line correspondence," in *Int. Conf. Robotics Automation*, Raleigh, NC, 1987, pp. 1275–1280.

[7] J. K. Aggarwal and N. Nandhakumar, "On the computation of motion from sequence of images: A review," *Proc. IEEE*, vol. 76, pp. 971–935, Aug. 1988.

[8] R. Safaee-Rad, K. C. Smith, B. Benhabib, and I. Tchoukanov, "3-D location estimation of circular features for machine vision," *IEEE Trans. Robot. Automat.*, vol. 8, pp. 624–640, Oct. 1992.

[9] G. L. Foresti and V. Murino, "Moving-object recognition from an image sequence for autonomous vehicle driving," in *Time-Varying Image Processing and Moving Object Recognition*, V. Cappelini, Ed. Amsterdam, The Netherlands: Elsevier, 1994, vol. 3, pp. 383–390.

[10] M. Yamamoto, "Pose from motion: A direct method," in *Time-Varying Image Processing and Moving Object Recognition*, V. Cappelini, Ed. Amsterdam, The Netherlands: Elsevier, 1994, vol. 3, pp. 287–294.

[11] J. Hwang, Y. Ooi, and S. Ozawa, "An adaptive sensing system with tracking and zooming a moving object," *IEICE Trans. Inform. Syst.*, vol. E76-D, pp. 926–934, Aug. 1993.

[12] D. Murray and A. Basu, "Active tracking," in *IEEE Int. Conf. Intelligent Robots Systems*, Yokohama, Japan, 1993, pp. 1021–1028.

[13] S. Shmuel and M. Werman, "Active vision: 3-D from an image sequence," in *IEEE Int. Conf. Pattern Recognition*, Atlantic City, NJ, 1990, vol. 1, pp. 8–54.

[14] J. T. Feddema and C. S. G. Lee, "Adaptive image feature prediction and control for visual tracking with a hand-eye coordinated camera," *IEEE Trans. Syst., Man, Cybern.*, vol. 20, pp. 1172–1183, Sept./Oct. 1990.

[15] M. J. Magee and J. K. Aggarwal, "Determine the position of a robot using a single calibration object," *IEEE Int. Conf. Robotics Automation*, Atlanta, GA, 1984, pp. 140–149.

[16] B. Ussain and M. R. Kabuka, "Real-time system for accurate three-dimensional position determination and verification," *IEEE Trans. Robot. Automat.*, vol. 6, pp. 31–43, Feb. 1990.

[17] Y. Fainman, L. Feng, and Y. Koren, "Estimation of absolute spatial position of mobile systems by hybrid opto-electronic processor," *IEEE, Int. Conf. Systems, Man, Cybernetics*, Cambridge, MA, 1989, pp. 651–657.

[18] Y. C. Shin and S. Ahmad, "3D location of circular spherical features by monocular model-based vision," *IEEE Int. Conf. Syst., Man, Cybern.*, Boston, MA, 1989, pp. 576–581.

[19] R. Safaee-Rad, K. C. Smith, B. Benhabib, and I. Tchoukanov, "Accurate estimate of quadratic curves from grey-level images," *J. Comput. Vision, Graphics, Image Processing: Image Understanding*, vol. 54, pp. 259–274, Sept. 1991.

[20] N. Ayache and O. D. Faugeras, "HYPER: A new approach for the recognition and positioning of two-dimensional objects," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-8, no. 1, pp. 44–54, 1986.

[21] R. Y. Tsai, "A versatile camera calibration technique for high-accuracy 3-D machine vision metrology using off-the-shelf TV cameras and lenses," *IEEE J. Robot. Automat.*, vol. RA-3, pp. 323–344, Aug. 1987.

[22] X. He, B. Benhabib, and K. C. Smith, "CAD-based off-line planning for active vision," *J. Robot. Syst.*, vol. 12, pp. 677–691, Oct. 1995.

[23] R. Safaee-Rad, "An active vision system for 3-D object recognition in robotic assembly workcells," Ph.D. dissertation, Dept. Mech. Eng., Univ. Toronto, Toronto, Ont., Canada, 1991.

[24] D. He, "An active-vision system for the recognition of moving objects, M.A.Sc. thesis, Dept. Mech. Eng., Univ. Toronto, Toronto, Ont., Canada, 1996.

[25] D. He, D. Hujic, J. K. Mills, and B. Benhabib, "Moving object recognition using pre-marking and active vision system," in *IEEE Int. Conf. Robotics Automation*, Minneapolis, MN, 1996, pp. 1980–1985.

Sufficient Conditions on Uniform Approximation of Multivariate Functions by General Takagi–Sugeno Fuzzy Systems with Linear Rule Consequent

Hao Ying

***Abstract*—We have constructively proved a general class of multi-input single-output Takagi–Sugeno fuzzy systems to be universal approximators. The systems use any types of continuous fuzzy sets, any types of fuzzy logic AND, fuzzy rules with linear rule consequent and the generalized defuzzifier. We first prove that the TS fuzzy systems can uniformly approximate any multivariate polynomial arbitrarily well, and then prove they can uniformly approximate any multivariate continuous function arbitrarily well. We have derived a formula for computing the minimal upper bounds on the number of fuzzy sets and fuzzy rules necessary to achieve prespecified approximation accuracy for any given bivariate function. A numerical example is furnished. Our results provide a solid theoretical basis for fuzzy system applications, particularly as fuzzy controllers and models.**

I. Introduction

With respect to fuzzy control and modeling applications, the existing fuzzy systems can be classified into two major types, namely Mamdani fuzzy systems and Takagi–Sugeno (TS) fuzzy systems. The primary difference between them lies in the fuzzy rule consequent. Mamdani fuzzy systems use fuzzy sets as rule consequent whereas

Manuscript received April 20, 1997; revised December 7, 1997. This work was supported in part by a Biomedical Engineering Research Grant from The Whitaker Foundation and by the Texas Higher Education Coordinating Board under Grant 004952-054.

The author is with the Department of Physiology and Biophysics, Biomedical Engineering Center, The University of Texas Medical Branch, Galveston, TX 77555-0456 USA (e-mail: hying@utmb.edu).

Publisher Item Identifier S 1083-4427(98)04348-3.

TS fuzzy systems employ linear functions of input variables as rule consequent [14]. Both types of fuzzy systems have been used widely as effective tools in various practical applications, especially in the areas of control and modeling. From mathematics 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 are 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 to which fuzzy systems 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. All the results in the literature, nevertheless, are only on Mamdani fuzzy systems (e.g., [1], [4], [5], [7], [10]–[13], [15], [16], [20]). Many of these results are of rather limited usefulness because were derived by using the Stone-Weierstrass theorem. As a result, they are just existence results on some particular configurations of Mamdani fuzzy systems. Existence results are interesting but are far from enough. In classical function approximation theory covering such approximators as polynomials and spline functions, quantitative results are the norm. Fuzzy approximation theory should be established to the same level. Keep this in mind, we have studied sufficient conditions for general and typical multiple-input single-output (MISO) Mamdani fuzzy systems as universal approximators [17]. We have also investigated necessary conditions for MISO as well as SISO fuzzy systems as universal approximators [18]. In [21], detailed approximation accuracy analysis for some particular fuzzy systems is carried out.

At present, there do not exist any approximation results in the literature for the common TS fuzzy systems. By "common," we mean those that use linear rule consequent and the centroid defuzzifier, as originally proposed by Takagi and Sugeno [14]. The interesting existence result in [3] is on an uncommon two-input one-output TS fuzzy system because it uses a linear defuzzifier, (i.e., the defuzzifier does not have denominator) and requires rule consequent be polynomials of the input variables. In reality, TS fuzzy systems use only linear functions of input variables as rule consequent. Linear rule consequent is critical to the practicality and usefulness of TS fuzzy systems. This is because when nonlinear rule consequent are used, proper determination of the rule consequent structure and parameters is extremely difficult, if not impossible. Furthermore, compared with traditional polynomial approximators that have been well established, the fuzzy system with polynomial rule consequent is greatly disadvantageous in terms of complexity and practical usefulness.

In present paper, we will use a two-step constructive proof approach to theoretically demonstrate that general MISO TS fuzzy systems with linear rule consequent are universal approximators. The TS fuzzy systems in this study are general because they use any types of continuous fuzzy sets, any types of fuzzy logic AND, fuzzy rules with linear consequent and the generalized defuzzifier containing the centroid defuzzifier as a special case. Furthermore, we have achieved some quantitative approximation results; we have derived a formula that relates the number of fuzzy sets and rules needed to the bivariate function to be approximated as well as the prespecified approximation error bound. The formula can compute the minimal upper bounds on the number of fuzzy sets and fuzzy rules. We provide a numerical example to demonstrate how to use the formula and its usefulness.

II. Configuration of General MISO TS Fuzzy Systems

The general TS fuzzy systems use r input variables, continuous-time or discrete-time or both. The variables are represented by a vector

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

where t is time. For notational simplicity, in the rest of the paper, we will use $\boldsymbol{x}$ and x_i, respectively, instead of $\boldsymbol{x}(\boldsymbol{t})$ and $x_i(t)$. Because one can always scale the variables so that they all fall in $[-1, 1]$, without loss of generality, we suppose that

$$-1 \le x_i \le 1, \qquad i = 1, 2, \cdots, r.$$

We partition $[-1, 1]$ into $2n$ $(n \ge 1)$ equal intervals, each of which is $[k/n, (k+1)/n]$ where $k = -n, \cdots, n-1$. Over the $2n$ intervals, we define $2n+1$ fuzzy sets for fuzzifying the variables. Each of the fuzzy sets is denoted by $A_{i,j}$ $(j = 0, \pm 1, \cdots, \pm n)$. The membership function of $A_{i,j}$, designated by $\mu_{i,j}$, can be any continuous functions whose values are bounded between 0 and 1. We impose little restriction on the membership functions because we want to include all the commonly-used fuzzy sets (e.g., triangular type, trapezoidal types, and bell-shape type) as well as any reasonable types. We are able to impose so little restriction because the approximation results we have established are independent of the shapes of the membership functions, as we will show later. 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 $2n-1$ ones over $[(k-1)/n, (k+1)/n]$, where $-(n-1) \le k \le n-1$. The fuzzy sets are not required to be identical for different input variables or for the same input variable, as the shapes of the membership functions will be irrelevant in the theoretical development in the next section.

$(2n+1)^r$ fuzzy rules with linear consequent are used to cover all the possible combinations of $A_{i,j}$. Rule $\#m, 1 \le m \le (2n+1)^r$, is

IF x_1 is $A_{1,p_{1,m}}$ AND x_2 is $A_{2,p_{2,m}}$ AND $\cdots$ AND x_r is $A_{r,p_{r,m}}$
THEN $F_n(\boldsymbol{x}) = a_{0,m} + a_{1,m}x_1 + a_{2,m}x_2 + \cdots + a_{r,m}x_r$

where $F_n(\boldsymbol{x})$ is output of the fuzzy systems. Here, we use the subscript n to signify that system output is parameterized by n. In other words, different number of fuzzy sets and rules will result in different structures of the fuzzy systems. The subscript $p_{i,m}$, where $-n \le p_{i,m} \le n$, is an integer. In each rule, there are $r+1$ parameters, namely $a_{0,m}, a_{1,m}, \cdots, a_{r,m}$. For $(2n+1)^r$ rules, there are $(r+1)(2n+1)^r$ parameters. The values of these parameters are chosen by the system developer.

Fuzzy logic AND in the rules can be any types of T-norm (e.g., Zadeh AND operation and/or product AND operation [8], [9]). A mixture of different types may be used in the same rule or in different rules. We can make such a broad assumption on fuzzy AND operators because their choices will not affect the final approximation results established in this paper. Using symbol $\otimes$ to represents any types of fuzzy logic AND operations ($\otimes$ represents as many types of AND operations as used), we obtain the combined membership for the consequent of Rule $\#m$ as follows:

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

We use the generalized defuzzifier [6] to calculate output of the fuzzy systems, which is

$$F_n(\boldsymbol{x}) = \frac{\sum_{m=1}^{(2n+1)^r} \mu_m^\alpha (a_{0,m} + a_{1,m}x_1 + \cdots + a_{r,m}x_r)}{\sum_{m=1}^{(2n+1)^r} \mu_m^\alpha}.$$

Different defuzzifiers are obtained by using different values, where $0 \leq \alpha < +\infty$. When $\alpha = 1$, the popular centroid defuzzifier is realized. The mean of maximum defuzzifier is realized when $\alpha = \infty$. One sees that this generalized defuzzifier is indeed general.

Mathematically, $F_n(\boldsymbol{x})$ is a function sequence in terms of n. It is a mapping F_n: $C^r[-1,1] \to (-\infty, \infty)$, where $C^r[-1,1]$ represents r-dimensional product space.

III. The General MISO TS Fuzzy Systems as Universal Approximators

To prove that the general MISO TS fuzzy systems are universal approximators, we will use the same two-step constructive proof approach as we did in [17]. The key of this approach is to use polynomials as a "bridge" to connect the two proof steps. In the first step, we will prove that the general MISO TS fuzzy systems can uniformly approximate any multivariate polynomial to any degree of accuracy. In the second step, we will utilize the fact that any multivariate continuous function can always be approximated by a 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 suppose that $P_M(\boldsymbol{x})$ is a multivariate polynomial of degree M defined in $C^r[-1,1]$

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

The proof of the following theorem completes the first of the two steps.

Theorem 1: The general MISO TS fuzzy systems with linear rule consequent can uniformly approximate $P_M(\boldsymbol{x})$, defined on a compact domain, with an arbitrarily small approximation error bound. That is $\forall\varepsilon > 0$ there exists a positive integer n^* such that $n > n^*$,

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

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

We assume a polynomial of two variables is expressed by

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

where $M_1 + M_2 = M$ is the degree of the polynomial. We suppose $P_M(x_1,x_2)$ is explicitly given. Using $P_M(x_1,x_2)$, we construct $(2n+1)^2$ TS fuzzy rules with linear rule consequent as follows:

Rule #1: IF x_1 is $A_{1,p_{1,1}}$ AND x_2 is $A_{2,p_{2,1}}$ THEN

$$F_n(x_1,x_2) = P_M\left(\frac{p_{1,1}}{n}, \frac{p_{2,1}}{n}\right) - \beta_{1,0}\frac{p_{1,1}}{n} - \beta_{0,1}\frac{p_{2,1}}{n} + \beta_{1,0}x_1 + \beta_{0,1}x_2$$

$\vdots$

Rule #m: IF x_1 is $A_{1,p_{1,m}}$ AND x_2 is $A_{2,p_{2,m}}$ THEN

$$F_n(x_1,x_2) = P_M\left(\frac{p_{1,m}}{n}, \frac{p_{2,m}}{n}\right) - \beta_{1,0}\frac{p_{1,m}}{n} - \beta_{0,1}\frac{p_{2,m}}{n} + \beta_{1,0}x_1 + \beta_{0,1}x_2$$

$\vdots$

One notices that in constructing the fuzzy rules, we let the parameters in Rule $\#m$ be

$$a_{0,m} = P_M\left(\frac{p_{1,m}}{n}, \frac{p_{2,m}}{n}\right) - \beta_{1,0}\frac{p_{1,m}}{n} - \beta_{0,1}\frac{p_{2,m}}{n}$$
$$a_{1,m} = \beta_{1,0}$$
$$a_{2,m} = \beta_{0,1}$$

where $m = 1, 2, \cdots, (2n+1)^2$. We point out that $a_{0,m}, a_{1,m}$, and $a_{2,m}$ are still constants. Output of the two-input fuzzy systems is shown in (1) at the bottom of the page. Because at any time

$$\frac{p_{i,m}}{n} \leq x_i \leq \frac{p_{i,m}+1}{n}, \qquad i = 1,2 \tag{2}$$

we have

$$\lim_{n\to\infty}\frac{p_{i,m}}{n} \leq x_i \leq \lim_{n\to\infty}\frac{p_{i,m}+1}{n}$$

which leads to

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

Thus,

$$\lim_{n\to\infty}\beta_{1,0}\frac{p_{1,m}}{n} = \beta_{1,0}x_1 \quad \text{and} \quad \lim_{n\to\infty}\beta_{0,1}\frac{p_{2,m}}{n} = \beta_{0,1}x_2.$$

As a result, we have from (1)

$$\lim_{n\to\infty} F_n(x_1,x_2) = \lim_{n\to\infty}\frac{\sum_{m=1}^{(2n+1)^2}\mu_m^\alpha P_M\left(\frac{p_{1,m}}{n}, \frac{p_{2,m}}{n}\right)}{\sum_{m=1}^{(2n+1)^2}\mu_m^\alpha} = P_M(x_1,x_2).$$

This means when the number of fuzzy sets and rules is very large, the fuzzy systems will approach the polynomial and, to the limit, will become the polynomial. Despite of this, we still need to prove the approximation to be uniform. To accomplish this, we will derive

$$F_n(x_1,x_2) = \frac{\sum_{m=1}^{(2n+1)^2}\mu_m^\alpha\left[\left(P_M\left(\frac{p_{1,m}}{n}, \frac{p_{2,m}}{n}\right) - \beta_{1,0}\frac{p_{1,m}}{n} - \beta_{0,1}\frac{p_{2,m}}{n} + \beta_{1,0}x_1 + \beta_{0,1}x_2\right)\right]}{\sum_{m=1}^{(2n+1)^2}\mu_m^\alpha}. \tag{1}$$

a formula that can calculate a positive integer n^*, based on given approximation error bound $\varepsilon > 0$, such that $\forall n > n^*$

$$||F_n - P_M||_{C^2[-1,1]} = \max_{x_1,x_2\in[-1,1]} |F_n(x_1,x_2) - P_M(x_1,x_2)| < \varepsilon.$$

According to (1), this inequality is achieved if the following inequality can hold:

$$\left|P_M\left(\frac{p_{1,m}}{n}, \frac{p_{2,m}}{n}\right) - \beta_{1,0}\frac{p_{1,m}}{n} - \beta_{0,1}\frac{p_{2,m}}{n} + \beta_{1,0}x_1 + \beta_{0,1}x_2 - P_M(x_1,x_2)\right| < \varepsilon.$$

We are going to determine n^* from this inequality. We have

$$\begin{aligned}&\left|P_M\left(\frac{p_{1,m}}{n}, \frac{p_{2,m}}{n}\right) - \beta_{1,0}\frac{p_{1,m}}{n} - \beta_{0,1}\frac{p_{2,m}}{n} + \beta_{1,0}x_1 + \beta_{0,1}x_2 - P_M(x_1,x_2)\right| \\ &\le \left|P_M\left(\frac{p_{1,m}}{n}, \frac{p_{2,m}}{n}\right) - P_M(x_1,x_2)\right| + \left|\beta_{1,0}\left(x_1 - \frac{p_{1,m}}{n}\right)\right| + \left|\beta_{0,1}\left(x_2 - \frac{p_{2,m}}{n}\right)\right|. \end{aligned} \quad (3)$$

Due to (2), we have

$$\frac{p_{i,m}}{n} = x_i - \frac{\theta_{i,m}}{n}, \quad \text{for } i = 1, 2,$$

where $0 \le \theta_{i,m} \le 1$, and

$$\left|x_i - \frac{p_{i,m}}{n}\right| = \frac{|\theta_{i,m}|}{n} < \frac{1}{n}.$$

Hence, for the second and third terms in the last part of (3), the following inequalities hold:

$$\left|\beta_{1,0}\left(x_1 - \frac{p_{1,m}}{n}\right)\right| \le \frac{|\beta_{1,0}|}{n} \quad \text{and} \quad \left|\beta_{0,1}\left(x_2 - \frac{p_{2,m}}{n}\right)\right| \le \frac{|\beta_{0,1}|}{n}.$$

For the first part of (3), the following is true:

$$\begin{aligned}&\left|P_M\left(\frac{p_{1,m}}{n}, \frac{p_{2,m}}{n}\right) - P_M(x_1,x_2)\right| \\ &= \left|\sum_{d_1=0}^{M_1}\sum_{d_2=0}^{M_2} \beta_{d_1,d_2}\left[\left(\frac{p_{1,m}}{n}\right)^{d_1}\left(\frac{p_{2,m}}{n}\right)^{d_2} - x_1^{d_1}x_2^{d_2}\right]\right| \\ &\le \sum_{d_1=0}^{M_1}\sum_{d_2=0}^{M_2} |\beta_{d_1,d_2}| \left|\left(x_1 - \frac{\theta_{1,m}}{n}\right)^{d_1}\left(x_2 - \frac{\theta_{2,m}}{n}\right)^{d_2} - x_1^{d_1}x_2^{d_2}\right|. \end{aligned} \quad (4)$$

Note that

$$\begin{aligned}&\left|\left(x_1 - \frac{\theta_{1,m}}{n}\right)^{d_1}\left(x_2 - \frac{\theta_{2,m}}{n}\right)^{d_2} - x_1^{d_1}x_2^{d_2}\right| \\ &= \left|x_2^{d_2}\left[-C_{d_1}^1 x_1^{d_1-1}\frac{\theta_{1,m}}{n} + \cdots + (-1)^{d_1}C_{d_1}^{d_1}\left(\frac{\theta_{1,m}}{n}\right)^{d_1}\right]\right. \\ &\quad + x_1^{d_1}\left[-C_{d_2}^1 x_2^{d_2-1}\frac{\theta_{2,m}}{n} + \cdots + (-1)^{d_2}C_{d_2}^{d_2}\left(\frac{\theta_{2,m}}{n}\right)^{d_2}\right] \\ &\quad + \left[-c_{d_1}^1 x_1^{d_1-1}\frac{\theta_{1,m}}{n} + \cdots + (-1)^{d_1}C_{d_1}^{d_1}\left(\frac{\theta_{1,m}}{n}\right)^{d_1}\right] \\ &\quad \left.\cdot \left[-c_{d_2}^1 x_2^{d_2-1}\frac{\theta_{2,m}}{n} + \cdots + (-1)^{d_2}C_{d_2}^{d_2}\left(\frac{\theta_{2,m}}{n}\right)^{d_2}\right]\right| \\ &\le \frac{C_{d_1}^1}{n} + \frac{C_{d_1}^2}{n} + \cdots + \frac{C_{d_1}^{d_1}}{n} + \frac{C_{d_2}^1}{n} + \frac{C_{d_2}^2}{n} + \cdots \\ &\quad + \frac{C_{d_2}^{d_2}}{n} + \left(\frac{C_{d_1}^1}{n} + \cdots + \frac{C_{d_1}^{d_1}}{n}\right)\left(\frac{C_{d_2}^1}{n} + \cdots + \frac{C_{d_2}^{d_2}}{n}\right) \\ &= \frac{2^{d_1} - 1}{n} + \frac{2^{d_2} - 1}{n} + \frac{(2^{d_1} - 1)(2^{d_2} - 1)}{n^2} \\ &\le \frac{2^{d_1+d_2} - 1}{n}. \end{aligned}$$

In the above derivation, we utilized the following relations:

$$|x_i^\eta| \le 1 \quad \text{where} \quad \eta \ge 1,$$

$$\frac{|\theta_{i,m}|}{n} \le \frac{1}{n},$$

$$\frac{1}{n} \ge \frac{1}{n^\eta} \quad \text{where} \quad \eta \ge 1,$$

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

Therefore, (4) becomes

$$\sum_{d_1=0}^{M_1}\sum_{d_2=0}^{M_2} |\beta_{d_1,d_2}| \left|\left(x_1 - \frac{\theta_{1,m}}{n}\right)^{d_1}\left(x_2 - \frac{\theta_{2,m}}{n}\right)^{d_2} - x_1^{d_1}x_2^{d_2}\right| \le \frac{\sum_{d_1=0}^{M_1}\sum_{d_2=0}^{M_2} |\beta_{d_1,d_2}|(2^{d_1+d_2} - 1)}{n}.$$

Combining all the inequalities derived, we have

$$\max_{x_1,x_2\in[-1,1]} |F_n(x_1,x_2) - P_M(x_1,x_2)| \le \frac{|\beta_{1,0}| + |\beta_{0,1}| + \sum_{d_1=0}^{M_1}\sum_{d_2=0}^{M_2} |\beta_{d_1,d_2}|(2^{d_1+d_2} - 1)}{n} < \varepsilon.$$

Hence, we have derived what we wanted as follows:

$$n^* \ge \frac{|\beta_{1,0}| + |\beta_{0,1}| + \sum_{d_1=0}^{M_1}\sum_{d_2=0}^{M_2} |\beta_{d_1,d_2}|(2^{d_1+d_2} - 1)}{\varepsilon}. \quad (5)$$

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

We now need to complete the second step of our two-step proof of the general TS fuzzy systems being universal approximators. In the second step, we will use the Weierstrass approximation theorem, which basically states that any multivariate continuous function can always be uniformly approximated by a multivariate polynomial no matter how small the desired approximation error bound is. The following theorem is proved for the general MISO fuzzy systems with r input variables.

Theorem 2 (Universal Approximation Theorem): The general MISO TS fuzzy systems with linear rule consequent can uniformly approximate any multivariate continuous function on a compact domain to any degree of accuracy.

Proof: Designate any multivariate continuous function to be approximated as $G(\boldsymbol{x})$ and the desired uniform approximation error bound as $\varepsilon > 0$. According to the Weierstrass approximation theorem, we can always find a multivariate polynomial $P_M(\boldsymbol{x})$ that can uniformly approximate $G(\boldsymbol{x})$ with accuracy ε. This is to say, $\forall \varepsilon_1 > 0, ||P_M - G|| < \varepsilon_1$. Furthermore, on the basis of Theorem 1, $\forall \varepsilon_2 > 0, ||F_n - P_M|| < \varepsilon_2$. Hence, if we choose ε_1 and ε_2 to make $\varepsilon_1 + \varepsilon_2 < \varepsilon$, then

$$||F_n - G|| \le ||P_M - G|| + ||F_n - P_M|| = \varepsilon_1 + \varepsilon_2 < \varepsilon.$$

That is,

$$|F_n(\boldsymbol{x}) - G(\boldsymbol{x})| < \varepsilon,$$

which means $F_n(\boldsymbol{x})$ can uniformly approximate $G(\boldsymbol{x})$ arbitrarily well. ■

To this point, we have proved the general MISO TS fuzzy systems to be universal approximators in that they can uniformly approximate any continuous multivariate functions arbitrarily well. Although Theorems 1 and 2 are qualitative results, the proof of Theorem 1 actually contains a quantitative result, that is, inequality (5). Combining these two theorems and using (5), we obtain the following quantitative result concerning sufficient requirement on the number of fuzzy sets and rules for the general TS fuzzy systems as universal approximators. Though the result is stated for bivariate functions, similar results can be established for functions with more variables. For simplicity and better presentation, however, we do not provide the more general results in the paper.

Theorem 3 (Sufficient Conditions): Given a continuous function $G(x_1, x_2)$, defined on a compact domain, to be approximated with desired uniform approximation error bound $\varepsilon > 0, |F_n(x_1, x_2) - G(x_1, x_2)| < \varepsilon$ when $n > n^*$ where

$$n^* \geq \frac{|\beta_{1,0}| + |\beta_{0,1}| + \sum_{d_1=0}^{M_1}\sum_{d_2=0}^{M_2} |\beta_{d_1,d_2}|(2^{d_1+d_2} - 1)}{\varepsilon - \varepsilon_1}. \quad (6)$$

Here, we assume that ε_1 to be such chosen that $|P_M(x_1, x_2) - G(x_1, x_2)| < \varepsilon_1$ and $\varepsilon_1 < \varepsilon$.

Proof: As we just stated above, the formula was already derived in the proof of Theorem 1. By simply replacing ε in (5) with $\varepsilon - \varepsilon_1$, we get this formula. ■

Once n^* is determined, the number of fuzzy sets needed to fuzzify x_1 and x_2 is obviously $2n^* + 1$ and the number of fuzzy rules is $(2n^* + 1)^2$.

IV. Discussion

Using any integer larger than the number calculated by (6) to compute the number of fuzzy sets and fuzzy rules guarantees to achieve uniform approximation by the general MISO fuzzy systems. As any integer larger than n^* is an upper bound, there exists an infinite number of upper bounds. But this is purely from mathematics point of view. Practically, though, one should always use as few fuzzy sets and rules as possible. Thus, only the computed n^* is the sensible upper bound because it is the smallest. We call $2n^*+1$ and $(2n^*+1)^2$ minimal upper bounds on the number of fuzzy sets and fuzzy rules, respectively. One should always use the computed n^* (if it is an integer) or the integer just larger than the calculated n^* (if n^* is not an integer) to calculate the minimal upper bounds.

We point out that n^* calculated by (6) could be conservative. In other words, depending on the function to be approximated, the minimal upper bounds determined could be somewhat too large. This is because (6) represents sufficient conditions on the number of fuzzy sets and rules; it does not represent necessary conditions, nor necessary and sufficient ones. Consequently, there may exist some smaller numbers of fuzzy sets and rules that will approximate the function with the desired approximation accuracy. Overestimation is natural and inevitable as (6) is independent of shape of the membership functions, fuzzy logic AND operators and the defuzzifier type (i.e., value of α). In practice, the common TS fuzzy systems use triangular, trapezoidal or Gaussian membership functions, Zadeh or product AND operators and the centroid defuzzifier. Under these specific constraints, how to numerically determine the minimal upper bounds or, ideally the exact number of fuzzy sets and rules needed, is an interesting, important but technically challenging research topic. The minimal upper bounds found under these constraints are expected to be less conservative than those computed by using (6).

A question related to the minimal upper bounds naturally arises: can lower bounds on the number of fuzzy sets and rules be numerically estimated? The answer is affirmative but a full presentation of our study in this aspect is beyond the scope of the present paper. The interested reader is referred to our recent paper [19] in which we have established necessary conditions for the general MISO TS fuzzy systems with linear rule consequent as universal approximators. We have found that for those continuous functions that are complex in mathematical representation but only have a small number of extrema, a few fuzzy sets and rules *may* suffice for uniform approximation with arbitrarily high accuracy. On the other hand, for those functions that are simple in mathematical formulation but have many extrema, such as periodic functions or highly-oscillatory functions, a large number of fuzzy sets and rules *must* be used. Once a function is explicitly given, the number of fuzzy sets and rules necessary for the approximation can easily be determined, which represent lower bounds. The numbers that are necessary as well as sufficient are between the lower bounds and the minimal upper bounds.

We now put the new approximation results in the context of fuzzy control and modeling, two major applications of fuzzy systems. For fuzzy control, the task is to develop a nonlinear controller that can control any given nonlinear system with desired control performance. As the system model is mathematically unavailable in most cases, the desired controller is mathematically unknown, too. Under this kind of circumstances, analytical design of the controller using conventional control theory is virtually impossible. One can construct the controller via fuzzy control methodology in a trial-and-error manner, but one may wonder whether such a construction approach will lead to the desired controller or at least to an approximate version of it. The results established in present paper guarantee its success provided that the controller is continuous and, as many fuzzy sets and rules as necessary are permissible. The essence of fuzzy modeling is the same as fuzzy control. One wants to construct a nonlinear model to accurately represent any given continuous and nonlinear physical dynamical system that is explicitly unknown. The approximation results are useful because they ensure such a nonlinear model is achievable.

V. Numerical Example

Example: What are the minimal upper bounds on the number of fuzzy sets and fuzzy rules for the general TS fuzzy systems to uniformly approximate $f(x_1, x_2) = e^{x_1+x_2}$, where $x_1 \in [-0.5, 0.5]$ and $x_2 \in [-0.5, 0.5]$, with approximation error less than 0.2 or 0.1?

Solution: The function $f(x) = e^x$ on the interval $[-1, 1]$ can be approximated by the following polynomial

$$P_3(x) = \frac{191}{192} + x + \frac{13x^2}{24} + \frac{x^3}{6}$$

with truncation error slightly less than 0.071. Hence, $f(x_1, x_2) = e^{x_1+x_2}$ can be approximated uniformly by the following third-order polynomials

$$P_3(x_1, x_2) = \tfrac{191}{192} + x_1 + x_2 + \tfrac{13}{24}x_1^2 + \tfrac{13}{24}x_2^2 + \tfrac{13}{12}x_1x_2 + \tfrac{1}{6}x_1^3 + \tfrac{1}{6}x_2^3 + \tfrac{1}{2}x_1^2x_2 + \tfrac{1}{2}x_1x_2^2$$

with truncation error slightly less than 0.071. We now use this polynomial to compute the minimal upper bounds on the number of fuzzy sets and fuzzy rules. Obviously, $\varepsilon_1 = 0.071$. For approximation accuracy $\varepsilon = 0.2$, according to (6), we obtain

$$n^* \geq \frac{1}{0.2-0.071}[1+1+(1+1)(2^1-1)+(\tfrac{13}{24}+\tfrac{13}{24}+\tfrac{13}{12}) \cdot (2^2-1)+(\tfrac{1}{6}+\tfrac{1}{6}+\tfrac{1}{2}+\tfrac{1}{2})(2^3-1)] = 153.7.$$

We choose $n^* = 154$ and the minimal upper bounds on the number of fuzzy sets and fuzzy rules are 309 and 95 481, respectively. For approximation accuracy $\varepsilon = 0.1$, it is straightforward to compute $n^* \geq 683.9$. We use $n^* = 684$, and consequently the minimal upper bounds on the number of fuzzy sets and rules are 1369 and 1 874 161, respectively. As expected, the higher the approximation accuracy, the larger the number of fuzzy sets and fuzzy rules.

One may wonder whether the minimal upper bounds computed will be different if a higher order or a lower order polynomial is used to approximate the function. Indeed, they will be different and we now continue the example to demonstrate this.

Note that $f(x_1, x_2) = e^{x_1+x_2}$ can also be approximated by the following fourth-order polynomial

$$\begin{aligned} P_4(x_1, x_2) = &1 + x_1 + x_2 + \tfrac{1}{2}x_1^2 + \tfrac{1}{2}x_2^2 + x_1x_2 \\ &+ \tfrac{1}{6}x_1^3 + \tfrac{1}{6}x_2^3 + \tfrac{1}{2}x_1^2x_2 + \tfrac{1}{2}x_1x_2^2 + \tfrac{1}{24}x_1^4 \\ &+ \tfrac{1}{24}x_2^4 + \tfrac{1}{6}x_1^3x_2 + \tfrac{1}{6}x_1x_2^3 + \tfrac{1}{4}x_1^2x_2^2 \end{aligned}$$

with truncation error slightly less than 0.024. That is, $\varepsilon_1 = 0.024$. Using (6), it is easy to compute $n^* \geq 174.3$ for $\varepsilon = 0.2$ and $n^* \geq 403.6$ for $\varepsilon = 0.1$. The minimal upper bounds on the number of fuzzy sets are 351 and 809, respectively. The corresponding minimal upper bounds on the number of fuzzy rules are 123 201 and 654 481, respectively.

These upper bounds appear to be excessively large; but one should remember that they represent sufficient conditions only. The numbers are big because they hold regardless of the shapes of the fuzzy sets, fuzzy logic AND operators and type of the defuzzifier. According to the necessary conditions in [21], a few fuzzy sets for x_1 and x_2 and a dozen of fuzzy rules *may* suffice for the fuzzy approximation required. Unfortunately, the exact numbers cannot be determined theoretically at present. Nevertheless, we believe that when typical fuzzy sets, fuzzy logic AND operators and defuzzifier are used, the number of fuzzy sets and fuzzy rules actually needed for the approximation should be far smaller than the numbers computed above.

VI. Conclusions

The MISO TS fuzzy systems covered in present paper are very general. They use any types of continuous fuzzy sets, any types of fuzzy logic AND, fuzzy rules with linear rule consequent and the generalized defuzzifier containing the popular centroid defuzzifier as a special case. Our contribution is two-fold:

1) we have constructively proved, in a two-step approach using polynomials as the bridge, that these fuzzy systems are universal approximators;
2) we have derived a formula for computing the minimal upper bounds on the number of fuzzy sets and rules needed for approximating any bivariate function with prespecified accuracy. The formula can be extended to cover functions with more variables.

The usefulness of these results for fuzzy control and modeling is discussed.

Prior to this paper, there did not even exist any qualitative results in the literature regarding TS fuzzy systems with linear rule consequent as universal approximators, letting alone quantitative ones. Thus, our qualitative and quantitative results are important and valuable; they provide a solid theoretical basis for various applications of TS fuzzy systems, particularly as fuzzy controllers and models.

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 Systems*, Yokohama, Japan, Mar. 1995, pp. 919–926.

[8] M. M. Gupta and J. Qi, "Theory of T-norms and fuzzy inference methods," *Fuzzy Sets Syst.*, vol. 40, pp. 431–450, 1991.

[9] ——, "Design of fuzzy logic controllers based on generalized T-operators," *Fuzzy Sets Syst.*, vol. 40, pp. 473–490, 1991.

[10] B. Kosko, "Fuzzy systems as universal approximators," in *Proc. IEEE Int. Conf. Fuzzy Systems*, San Diego, CA., Mar. 8–12, 1992.

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

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

[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 Systems*, San Diego, CA., Mar. 8–12, 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, no. 2, pp. 157–167, 1998.

[20] ——, "Constructing nonlinear variable gain controllers via the Takagi-Sugeno fuzzy control," *IEEE Trans. Fuzzy Syst.*, vol. 6, no. 2, pp. 226–234, 1998.

[21] 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," submitted for publication.

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

[23] ——, "Approximation accuracy analysis of fuzzy systems as function approximators," *IEEE Trans. Fuzzy Syst.*, vol. 4, pp. 44–63, 1996.