Fifth IFSA World Congress (1993), 1414-1417

ON APPROXIMATION OF CONTROLS BY FUZZY SYSTEMS

Hung T. Nguyen[1], **Vladik Kreinovich**[2]
[1]*Department of Systems Science, Tokyo Institute of Technology*
4259 Nagatsuta-cho, Midori-ku, Yokohama 227 Japan
[2]*Computer Science Department*
University of Texas at El Paso, El Paso, TX 79968-0518

Abstract. Wang and Mendel proved (1991) that fuzzy systems with product inference, centroid defuzzification, and everywhere positive membership functions (in particular, Gaussians, Wang, 1992) are capable of approximating any real continuous control function on a compact set to arbitrary accuracy. Kosko (1992) proved that fuzzy systems, in which membership functions have compact support, and combination operation ($\vee$-operation) for rules is the sum, are also universal approximators.

In this paper, we generalize this result of Kosko and prove that for any &- and $\vee$-operations, any defuzzification procedure, and any basic membership function with a compact support, the resulting fuzzy controls are universal approximators. Also, Wang's result is transfered to min-inference.

1. INTRODUCTION

It has been recognized that fuzzy systems and neural networks share similar structures. Approximation capabilities of neural networks are justified by Hornik et al [HSW89] who used Stone-Weierstrass approximation theorem to prove that neural networks are universal approximators. Using similar technique, Mendel and Wang [MW91, W92] and Kosko [K92] showed that some classes of fuzzy systems are also universal approximators.

In the present paper we prove that for any membership function with a compact support, &- and $\vee$-operations, and a defuzzification rule, an arbitrary continuous control can be approximated by fuzzy controls. In other words, the resulting fuzzy systems are universal approximators.

Specifically, with the rule base consisting of the rules of the following type:

if x_1 is A_1^j and x_2 is A_2^j and... and x_n is A_n^j, then u is B^j

where A_i^j, B^j are the terms of natural language that are used in describing j-th rule (e.g., "small", "medium", etc).

The property "u is a proper control" (which we will denote by $C(u)$) can be described as follows:

$$\begin{aligned} C(u) \equiv\ & (A_1^1(x_1)\,\&\,A_2^1(x_2)\,\&\ldots\&\,A_n^1(x_n)\,\&\,B^1(u))\vee \\ & (A_1^2(x_1)\,\&\,A_2^2(x_2)\,\&\ldots\&\,A_n^2(x_n)\,\&\,B^2(u))\vee \\ & \ldots \\ & (A_1^j(x_1)\,\&\,A_2^j(x_2)\,\&\ldots\&\,A_n^j(x_n)\,\&\,B^j(u))\vee \\ & \ldots \\ & (A_1^K(x_1)\,\&\,A_2^K(x_2)\,\&\ldots\&\,A_n^K(x_n)\,\&\,B^K(u)) \end{aligned}$$

After we translate $A_i^j(x)$ as $\mu_i^j(x)$, $B^j(u)$ as $\mu_j(u)$, and & and $\vee$ as $f_\&$ and $f_\vee$, we form the membership function: $\mu_C(u) = f_\vee(p_1, \ldots, p_K)$, where $p_j = f_\&(\mu_{j,1}(x_1), \mu_{j,2}(x_2), \ldots, \mu_{j,n}(x_n), \mu_j(u)), j = 1, \ldots, K$. We can now apply a defuzzification operator D to this membership function $\mu_C(u)$ and thus obtain the desired value $f(\vec{x})$ of the control that corresponds to $\vec{x} = (x_1, \ldots, x_n)$.

In the next Section, we will specify the terminology and notations used in this paper, and formulate the results. The proofs will be in Section 3.

2. DEFINITIONS AND MAIN RESULTS

First, Wang's result [W92] remains valid when the product inference ($f_{\&}(a,b) = ab$) is replaced by min inference ($\min(a,b)$). Specifically, let us denote by $\mathcal{F}$ the class of functions $f : R^n \to R$ of the form $f(\vec{x}) = [\sum_{j=1}^{K} \bar{u}_j \min(\mu_{j,i}(x_i), i = 1,...,n)]/[\sum_{j=1}^{K} \min(\mu_{j,i}(x_i), i = 1,...,n)]$, where $\mu_{j,i}(x)$ are Gaussian functions, and $\bar{u}_j$ are any real numbers. By $\mathcal{F}_{|\mathcal{U}}$ we mean the restrictions of elements of $\mathcal{F}$ to a domain U. The set of all continuous functions on U is denoted by $C(U)$. Then:

THEOREM 1. *For any compact $U \subset R^n$, $\mathcal{F}_{|\mathcal{U}}$ is dense in the sup-norm in $C(U)$.*

We now state a similar result, but for membership functions of a different type. All membership functions are assumed to be continuous on the real line R. By a *basic membership function* we mean a continuous function $\mu_0(x)$ that is positive for all x from some interval (a, b), and is equal to 0 outside that interval. A membership function $\mu(x)$ is *of the type* μ_0 if $\mu(x) = \mu_0(kx + l)$, for some real numbers $k \neq 0$ and l. Next, by an *&−operation* we mean a continuous binary operation $f : [0,1] \times [0,1] \to [0,1]$ such that $f(0,0) = f(0,1) = f(1,0) = 0, f(1,1) = 1$, $f(p,q) = f(q,p)$ for all p,q, $f(p,q) \leq p$ for all p and q, and if $p > 0$ and $q > 0$, then $f(p,q) > 0$. For an *$\vee$−operation*, we require the following: $f(0,0) = 0, f(0,1) = f(1,0) = f(1,1) = 1$, $f(p,q) = f(q,p)$ for all p,q, and $f(p,q) \geq p$ for all p and q.

A *defuzzification procedure* D transforms a membership function $\mu(x)$ into a number in such a way that if $\mu(x) = 0$ outside an interval (a,b), then $D(\mu) \in [a,b]$ (both centroid and center-of-maximum are defuzzification procedures in this sense). By a *fuzzy methodology* $\mathcal{M}$ we mean a triple consisting of a basic membership function μ_0, a pair of &− and $\vee$−operations $f_{\&}(p,q)$ and $f_{\vee}(p,q)$, and a defuzzification procedure D.

Assume that a fuzzy methodology $\mathcal{M}$ is fixed. By $\mathcal{F}(\mathcal{M})$ we denote the class of all functions $f : R^n \to R$ that are equal to $f(\vec{x}) = D(\mu_C)$, where $\mu_C(u) = f_{\vee}(p_1, p_2, ..., p_j, ..., p_K)$, p_j is defined previously, and all membership functions $\mu_{j,i}(x)$, $\mu_j(u)$ are of type μ_0. Then:

THEOREM 2. *For any given fuzzy methodology $\mathcal{M}$, and compact $U \subset R^n$, $\mathcal{F}(\mathcal{M})_{|\mathcal{U}}$ is dense in the sup-norm in $C(U)$.*

3. PROOFS

Theorem 1: As in [W92], it suffices to verify the hypotheses of Stone-Weierstrass theorem.

The fact that $\mathcal{F}_{|\mathcal{U}}$ is an algebra of functions, follows simply from the fact that if a_i, b_J are all positive, then $(\min(a_i, i = 1,...,p)) \times (\min(b_j, j = 1,...,q)) = \min_i \min_j a_i b_j$. Indeed, if $f(\vec{x}) = [\sum_{j=1}^{K} \bar{u}_j \min(\mu_{j,i}(x_i), i = 1,...,n)]/[\sum_{j=1}^{K} \min(\mu_{j,i}(x_i), i = 1,...,n)]$ and $g(\vec{x}) = [\sum_{j=1}^{L} \bar{v}_j \min(\nu_{j,i}(x_i), i = 1,...,n)]/[\sum_{j=1}^{L} \min(\nu_{j,i}(x_i), i = 1,...,n)]$, and α is a real number, then it is obvious that $\alpha f(vecx)$ is in $\mathcal{F}$. The sum $f(\vec{x}) + g(\vec{x})$ and product $f(\vec{x})g(\vec{x})$ of the expressions f and g, involve terms of the form $\min(\mu_{j,i}(x_i), i = 1,...,n) \min(\nu_{k,i}(x_i), i = 1,...,n)$. The last term is equal to $\min(\mu_{j,i}(x_i)\nu_{k,i}(x_i), i = 1,...,n)$, because the $\mu_{j,i}(x_i)$'s and the $\nu_{k,i}(x_i)$'s are positive. The result follows from the obvious fact that products of Gaussians are Gaussians.

$\mathcal{F}_{|\mathcal{U}}$ vanishes at no point: $\min(a_i, i = 1,...,p)$ is always positive.

$\mathcal{F}_{|\mathcal{U}}$ separates points: for $\vec{y}, \vec{z} \in R^n$, consider the element of $\mathcal{F}_{|\mathcal{U}}$: $f(\vec{x}) = [\min_i(\exp(-1/2(x_i - y_i)^2)]/[\min_i(\exp(-1/2(x_i - y_i)^2) + \min_i(\exp(-1/2(x_i - z_i)^2)]$.
Then $f(\vec{z})/f(\vec{y}) = \min_i(\exp(-1/2(z_i - y_i)^2) = 1$ if and only if $z_i = y_i$ for all i, which is impossible if $\vec{y} \neq \vec{z}$.

Theorem 2: Since U is compact, there exist a finite $(\delta/2)-$net, i.e., a finite set of points $\vec{x}^1, ..., \vec{x}^K \in U$ such that for every $\vec{x} \in U$ there exists a j, for which $\rho(\vec{x}, \vec{x}^j) \leq \delta/2$. Let us fix such a net.

Let g belong to $C(U)$, and $\varepsilon > 0$. Let us form a rule base with K rules, one rule for each point $\vec{x}^j = (x_1^j, x_2^j, ..., x_n^j)$ from the chosen $\delta/2-$net. This rule will take the form

$$A_1^j(x_1) \,\&\, A_2^j(x_2) \,\&\, ... A_n^j(x_n) \rightarrow B^j(u),$$

where the corresponding membership functions are $\mu_{j,i}(x) = \tilde{\mu}_0((x - x_i^j)/(\delta))$, and $\mu_j(u) = \tilde{\mu}_0((u - u^j)/(\varepsilon/2)$, and by u^j we denoted $u^j = g(\vec{x}^j)$, and where $\tilde{\mu}_0(x) = \mu_0(x((b-a)/2 + (a+b)/2$ is a function of type μ_0 that is > 0 for $x \in [-1, 1]$.

It suffices to show that $\mu_C(u)$ is not identically 0, and is zero outside $[g(\vec{x}) - \varepsilon, g(\vec{x}) + \varepsilon]$. Then the resulting inequality will follow from the definition of a defuzzification procedure D.

First, let us prove that it is not identically 0. Since we chose the set $\{\vec{x}^j\}$ as a $\delta/2-$net, there exists a j such that $|x_i^j - x_i| \leq \delta/2$ for all i. Therefore, $|x_i^j - x_i| < \delta$, hence $x_i \in (x_i^j - \delta, x_i^j + \delta)$, and, $\mu_{j,i}(x_i) > 0$ for all i. If we take $u = g(\vec{x}^j)$, then we conclude that $\mu_j(u) > 0$. Since we demanded that $f_\&(p,q) > 0$ if $p > 0$ and $q > 0$, we can conclude that

$$p_j = f_\&(\mu_{j,1}(x_1), \mu_{j,2}(x_2), ..., \mu_{j,n}(x_n), \mu_j(u))) > 0$$

Since $f_\vee(p,q) \geq \max(p,q)$, we conclude that for this u, $\mu_C(u) = f_\vee(p_1, ..., p_j, ..., p_K) \geq p_j > 0$, so $\mu_C(u)$ is not identically 0.

Let us now prove that if $|u - g(\vec{x})| > \varepsilon$, then $\mu_C(u) = 0$. Namely, we will prove that in this case, $p_1 = p_2 = ... = p_j = ... = p_K = 0$; then $\mu_C(u) = f_\vee(p_1, p_2, ..., p_K) = f_\vee(0, 0, ..., 0) = 0$.

Let us take an arbitrary j from 1 to K and prove that $p_j = 0$. Indeed, since

$$p_j = f_\&(\mu_{j,1}(x_1), \mu_{j,2}(x_2), ..., \mu_{j,n}(x_n), \mu_j(u))) > 0$$

and $f_\&(0,p) = 0$, the only possibility for p_j to be positive is for all the terms $\mu_{j,1}(x_1)$, $\mu_{j,2}(x_2)$,...,$\mu_{j,n}(x_n)$,and $\mu_j(u)$ to be positive. Because of our choice of $\mu_{j,i}$, the first term is positive only for $|x_1 - x_1^j| \leq \delta$. In a similar manner, the first n terms are positive if $|x_i - x_i^j| \leq \delta$ for all i. But in this case, by the choice of δ, $|g(\vec{x}) - g(\vec{x}^j)| = |g(\vec{x}) - u^j| \leq \varepsilon/2$. We assumed that $|u - g(\vec{x})| > \varepsilon$; therefore, $|u - u^j| \geq |u - g(\vec{x})| - |g(\vec{x}) - u^j| > \varepsilon/2$, and $\mu_j(u) = 0$.

So, for every $\vec{x}$, either one of terms $\mu_{j,i}(x_i)$ is equal to 0, or they are all positive, and then $\mu_j(u) = 0$. In both cases, $p_j = 0$. So, $\mu_C(u) = f_\vee(p_1, ..., p_K) = f_\vee(0, 0, ..., 0) = 0$ for all u outside an interval $[g(\vec{x}) - \varepsilon, g(\vec{x}) + \varepsilon]$, and thence $f(\vec{x}) = D(\mu_C)$ belongs to this same interval. Q.E.D.

4. CONCLUSION

In view of the results of Wang and Mendel, Wang, Kosko, and ours, many classes of fuzzy systems are universal approximators. The choice of a fuzzy methodology should be based upon additional considerations such as smoothness, stability, or robustness of the resulting control (see, e.g., [KQL91], [KQLFLKBR92], [NKT92]).

Acknowledgements. This work was supported by a NSF Grant No. CDA-9015006, NASA Research Grant No. 9-482 and the Institute for Manufacturing and Materials Management grant. The work of the first named author was carried out at the Laboratory for International Fuzzy Engineering Research, Chair of Fuzzy Theory, Japan.

REFERENCES

[HSW89] K. Hornik, M. Stinchcombe, and H. White. *Multilayer feedforward networks are universal approximators*, Neural Networks, 1989, Vol. 2, pp. 359–366.

[K92] B. Kosko. *Fuzzy systems as universal approximators.* Proceedings of the IEEE International Conference on Fuzzy Systems 1992, San Diego, CA, 1992, pp. 1153–1162.

[KQL91] V. Kreinovich, C. Quintana, and R. Lea. *What procedure to choose while designing a fuzzy control? Towards mathematical foundations of fuzzy control*, Working Notes of the 1st International Workshop on Industrial Applications of Fuzzy Control and Intelligent Systems, College Station, TX, 1991, pp. 123–130.

[KQLFLKBR92] V. Kreinovich, C. Quintana, R. Lea, O. Fuentes, A. Lokshin, S. Kumar, I. Boricheva, and L. Reznik. *What non- linearity to choose? Mathematical foundations of fuzzy control.* Proceedings of the 1992 International Fuzzy Systems and Intelligent Control Conference, Louisville, KY, 1992, pp. 349–412.

[NKT92] H. T. Nguyen, V. Kreinovich, and D. Tolbert. *On robustness of fuzzy logics*, University of Texas at El Paso, Computer Science Department, Technical Report, May 1992.

[W92] L.-X. Wang. *Fuzzy systems are universal approximators*, Proceedings of the IEEE International Conference on Fuzzy Systems 1992, San Diego, CA, 1992, pp. 1163–1169.

[WM91] L.-X. Wang and J. Mendel. *Generating fuzzy rules from numerical data, with applications*, University of Souther California, Signal and Image Processing Institute, Technical Report USC-SIPI # 169, 1991.