# New Mass-Assignment-Based Fuzzy CMAC and Its Learning Convergence*

WANG Shi-tong[1,2], J.F. Baldwin[2], T.P. Martin[2]

[1](*Department of Computer Science, Eastchina Shipbuilding Institute, Zhenjiang* 212003, *China*)

[2](*Advanced Computing Research Center, Bristol University, UK*)

E-mail: zjstwang@public.zj.js.cn

http://www.zj.js.cn

Received June 1, 1999; accepted November 24, 1999

**Abstract**: In this paper, based on the mass-assignment theory proposed by J.F. Baldwin *et al.*, the new mass-assignment-based fuzzy CMAC is presented. Accordingly, its learning rules are also investigated. The theoretical research results reveal that this new mass-assignment-based fuzzy CMAC is a universal approximator, and has its learning convergence. Therefore, this new fuzzy CMAC has very important potentials of applications.

**Key words**: fuzzy CMAC; learning rule; mass-assignment theory; learning convergence

Fuzzy CMAC, as the extension of CMAC presented by Dr. Albus, has been abstracting more and more scholars and obtaining more and more applications[1~3]. In Ref.[3], based on $f_g$, $f_V$, we presented a generalized fuzzy CMAC, and then proved its learning convergence. Generalized fuzzy CMAC is so far the most general extension of fuzzy CMAC.

Recently, J.F. Baldwin *et al.* developed mass-assignment theory[4], and obtained many excellent applications. This theory is a very effective tool for processing uncertainties. When we attempt to combine it with multilayer fuzzy neural networks, some difficulties occur due to the existence of the reordering procedures in least prejudiced distributions of this theory. However, can we combine this theory with fuzzy CMAC? The answer is certain. In this paper, we will extend the theory to fuzzy CMAC naturally and further present the new mass-assignment-based fuzzy CMAC. In this new fuzzy CMAC, it adopts a new overall mapping function which is completely different from that in generalized fuzzy CMAC. Thus, this new fuzzy CMAC represents another family of fuzzy CMACs, that is to say, it cannot be included in generalized fuzzy CMACs. Because of its universal approximation capability and its learning convergence, it will have very good application potentials.

## 1 Mass-Assignment-Based Fuzzy CMAC

Mass-assignment theory[4] consists of many concepts including the definition of mass assignment, voting model, support pairs, point/interval value semantic unifications, and least prejudiced distribution. Here, we give the following concepts which will be used in this paper.

* Supported by the National Natural Science Foundation of China under Grant No.69843002 (国家自然科学基金)

**WANG Shi-tong** was born in 1964. He is a professor and doctoral supervisor at the Department of Computer Science, Eastchina Shipbuilding Institute. His research areas include artificial intelligence, neural networks and fuzzy systems. **J.F. Baldwin** is a professor and head of advanced computing research center, University of Bristol, U.K. His research areas include artificial intelligence, neural networks and fuzzy systems. **T.P. Martin** is a Reader of Advanced Computing Research Center, University of Bristol, U.K. His research areas include artificial intelligence, neural networks and fuzzy systems.

Suppose $A$ is a fuzzy set defined on the discrete domain $\{a_1, a_2, \ldots, a_n\}$:

$$A=\lambda_1/a_1+\lambda_2/a_2+\ldots+\lambda_n/a_n$$

where $\lambda_i$ corresponds to the membership value of the element $a_i$ and $\lambda_i \geq \lambda_{i+1}$, then the corresponding mass assignment is:

$$\{\{a_1, \ldots, a_i\}: \lambda_i-\lambda_{i+1}; \varnothing: 1-\lambda_1\} \quad with\ \lambda_{n+1}=0$$

By redistributing the mass assignment with a group of elements equally among these elements, we get its least prejudiced distribution which is a probability distribution:

$$a_1: \lambda_1-\lambda_2+(\lambda_2-\lambda_3)/2+(\lambda_3-\lambda_4)/3+\ldots+\lambda_n/n+(1-\lambda_1)/n$$

$$a_2: (\lambda_2-\lambda_3)/2+(\lambda_3-\lambda_4)/3+\ldots+\lambda_n/n+(1-\lambda_1)/n$$

$$a_3: (\lambda_3-\lambda_4)/3+\ldots+\lambda_n/n+(1-\lambda_1)/n$$

$$\vdots$$

$$a_n: \lambda_n/n+(1-\lambda_1)/n$$

For example, suppose fuzzy set $A=0.8/a+0.7/b+0.3/c$, then its mass assignment is: $\{a\}:0.1$, $\{a,b\}:0.4$, $\{a,b,c\}:0.3$, $\varnothing:0.2$, thus, we may get its least prejudiced distribution:

$$a: 0.1+0.2+0.1+0.2/3=0.4+0.2/3$$

$$b: 0.2+0.1+0.2/3=0.3+0.2/3$$

$$c: 0.1+0.2/3$$

Now, let us describe the new mass-assignment-based fuzzy CMAC as follows. The architecture of this new fuzzy CMAC is shown in Fig. 1. In Fig. 1, all the membership functions take the form of Gaussian functions $\exp(-(x-m)^2/\sigma)$. As shown in Fig. 1, this new fuzzy CMAC inherits the preferred features of learning and parallel processing from conventional CMAC, and the capability of acquiring and incorporating human knowledge into a system and the capability of processing information based on fuzzy inferences from fuzzy logic. The combination of CMAC, fuzzy logic and mass-assignment theory yields an advanced intelligent system architecture.

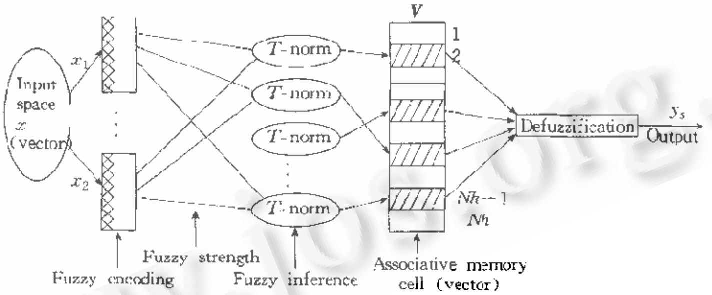


Fig. 1

At the input stage, the fuzzy CMAC uses the same fuzzification method as that in fuzzy logic controller to get its input encoding scheme. Fuzzy rules can be assigned to each associative memory cell. These rules may not necessarily have a crisp consequent part. The output generation uses the overall mapping function which is based on mass-assignment theory. Suppose $m_i$ is the number of knot points (Fig. 1) on the $i$th input dimension, $j_i=1, 2, \ldots, m_i$, then there are $M$ weights $v_p (p=1,2,\ldots,M)$ which will be chosen to get the overall output. $M=j_1$, if $N=1$; $M=j_i$ for $N>1$, $i=1,2,\ldots,N$. From Fig. 1, we know that suppose $\mu_{i,ji}(x_i)$ is a Gaussian membership function, we can define $c_p(x_s)=\mu_{1,j1}(x_{s1}) * \mu_{2,j2}(x_{s2}) * \ldots * \mu_{N,jN}(x_{sN})$ as a fuzzy basis function, where $ji$ (including the following similar representations) denotes $j_i$ and $x_s$ is the $s$th input vector. Since there are $N * \sum_{j=1}^{N} m_j \prod_{i=1}^{N} m_i$ fuzzy membership functions in fuzzy encoding, as shown in Fig. 1, we can get $Nh=\prod_{i=1}^{N} m_i$ fuzzy basis functions, i.

e: $p=1,2,\ldots,Nh$.

Let $a_s^T=[a_{s,1},a_{s,2},\ldots,a_{s,Nh}]$ denote a selection vector of fuzzy basis functions which has $M$ 1's, and we define the possibility of taking $v_i$ as $a_{s,i}c_i(x_s)$, i.e.,

$$v_1: a_{s,1}c_1(x_s)$$
$$v_2: a_{s,2}c_2(x_s)$$
$$\vdots$$
$$v_{Nh}: a_{s,Nh}c_{Nh}(x_s)$$

Thus, we can construct a fuzzy set:

$$a_{s,1}c_1(x_s)/v_1+a_{s,2}c_2(x_s)/v_2+\ldots+a_{s,Nh}c_{Nh}(x_s)/v_{Nh}$$

For simplicity, we assume $a_{s,1}c_1(x_s)>=a_{s,2}c_2(x_s)>=\ldots>=a_{s,Nh}c_{Nh}(x_s)$. As stated previously, we may get the corresponding least prejudiced distribution:

$$v_1: p_1=a_{s,1}c_1(x_s)-a_{s,2}c_2(x_s)+(a_{s,2}c_2(x_s)-a_{s,3}c_3(x_s))/2+(a_{s,3}c_3(x_s)-a_{s,4}c_4(x_s))/3+\ldots-a_{s,Nh}c_{Nh}(x_s)/N_h+(1-a_{s,1}c_1(x_s))/N_h$$

$$v_2: p_2=(a_{s,2}c_2(x_s)-a_{s,3}c_3(x_s))/2+(a_{s,3}c_3(x_s)-a_{s,4}c_4(x_s))/3+\ldots+a_{s,Nh}c_{Nh}(x_s)/N_h+(1-a_{s,1}c_1(x_s))/Nh$$

$$v_3: p_3=(a_{s,3}c_3(x_s)-a_{s,4}c_4(x_s))/3+\ldots+a_{s,Nh}c_{Nh}(x_s)/Nh+(1-a_{s,1}c_1(x_s))/Nh.$$

$$\ldots$$

$$v_{Nh}: p_{Nh}=a_{s,Nh}c_{Nh}(x_s)/Nh+(1-a_{s,1}c_1(x_s))/Nh$$

Note: $p_1+p_2+\ldots+p_{Nh}=1$. So, the overall mapping function of this new fuzzy CMAC can be defined as

$$y_s=\sum_{Nh}^{i=1} p_i v_i \tag{1}$$

We should note that in generalized fuzzy CMAC, the overall mapping function is

$$y_s=\sum_{i=1}^{Nh} a_{s,i}c_i(x_s)v_i/a_{s,i}c_i(x_s) \tag{2}$$

Hence, Eq. (1) is different from Eq. (2). Thus, mass-assignment-based fuzzy CMAC is a new model of CMAC. When implementing this new fuzzy CMAC, we should note, since the value $a_{s,i}c_i(x_s)$ is indeterministic, we must reorder them every time to get $p_i$. It can easily be done.

## 2 Mass-Assignment-Based Fuzzy CMAC as a Universal Approximator

In this section, we will prove a very good characteristic of mass-assignment-based CMAC, that is, it is a universal approximator, namely, it can approximate any given real continuous function on a compact domain to arbitrary accuracy.

**Theorem 1**[5]. If the basis functions $f_i(x)$ are differentiable in arbitrary order, then their linear combination $\sum w_i f_i(x)$ is a universal approximator, where $w_i$ is a real coefficient.

Now, let us apply the above theorem to prove Theorem 2.

**Theorem 2.** Mass-assignment-based fuzzy CMAC is a universal approximator.

*Proof.* In terms of Eq. (1), the overall output of mass-assignment-based fuzzy CMAC is

$$y_s=\sum_{i=1}^{NH} p_i v_i$$

where $p_i$ is the probability in the least prejudiced distribution, and $p_i$ is the linear combination of $a_{s,i}c_i(x_s)$ ($j=1,2,\ldots,Nh$). Because $c_i(x_s)$ is the multiplication of Gaussian membership functions, then

$$a_{s,i}c_i(x_s)=\begin{cases}0 & \text{if } a_{s,i}=0\\ c_i(x_s) & \text{if } a_{s,i}=1\end{cases}$$

thus, $y_s=\sum_{i=1}^{Nh} p_i v_i$ is also the linear combination of the multiplications of Gaussian membership functions. Because the multiplication function of Gaussian membership functions can easily be proved to be differentiable in arbitrary order, in terms of Theorem 1, this theorem obviously holds, i.e., this fuzzy CMAC is a universal approximator.

## 3 The Learning Rule and Its Convergence of Mass-Assignment-Based Fuzzy CMAC

In this section, we will first derive the learning rule of mass-assignment-based fuzzy CMAC, then prove that it has the learning convergence.

Suppose given input $x_s$, $Y_s$ is the actual output, we choose the least square error function

$$e=(Y_s-y_s)^2/2$$

then the learning rule of this new fuzzy CMAC is as follows:

$$\Delta v_k=\gamma\partial e/\partial v_k=\gamma(Y_s-y_s)\partial y_s/\partial v_k=\gamma(Y_s-y_s)p_k=\gamma\Big(Y_s-\sum_{i=1}^{Nh}p_i v_i\Big)p_k \tag{3}$$

where $\gamma$ is the learning rate.

In the following, we investigate the convergence problem of the above learning rule. In terms of Eq. (1), we can re-express it into $y=Bv$, where $B=(b_{ij})_{1\times Nh}$ which is bounded, i.e., all elements are bounded in terms of the definitions of $p_j$; $v=(v_1,v_2,\ldots,v_{Nh})^T$. Given the real output $y^*$ for input vector $x$, the learning rule is

$$\Delta v_k=\gamma\Big(y^*(x)-\sum_{i=1}^{Nh}b_{1i}v_i\Big)b_{1k} \tag{4}$$

where $\gamma$ is the learning rate.

Comparing (4) with (3), we have

$$Y_s=y^*(x_s)$$
$$b_{1k}=p_k$$
$$y(x)=\sum_{i=1}^{Nh}b_{1i}v_i$$

From the above analysis, we derive that in order to prove that the learning rule (3) has the learning convergence, we only need to prove that (4) has the learning convergence.

We define $v_s^{(i)}$ as the vector of weights before the $s$th sample is presented in the $i$th iteration of learning; $N_s$ denotes the number of samples. We consider the case that a set of $N_s$ training data is repeatedly presented to the learning rule.

With (4), we have

$$\begin{aligned} v_s^{(i)}&=v_{s-1}^{(i)}+\Delta v_{s-1}^{(i)}=v_{s-1}^{(i)}+\gamma(y_{s-1}^*-B_{s-1}v_{s-1}^{(i)})(b_{11}^{s-1},\ldots,b_{1Nh}^{s-1})^T\\ &=v_{s-1}^{(i)}+\gamma(y_{s-1}^*-B_{s-1}v_{s-1}^{(i)})R_{s-1} \end{aligned} \tag{5}$$

where $R_{s-1}=(b_{11}^{s-1},\ldots,b_{1Nh}^{s-1})^T$. With (5), the difference in the vector $v_s^{(i)}$ between two consecutive iterations $i$ and $i+1$ is calculated as

$$\begin{aligned} Dv_s^{(i)}&=v_s^{(i+1)}-v_s^{(i)}=v_{s-1}^{(i+1)}+\Delta v_{s-1}^{(i+1)}-(v_{s-1}^{(i)}+\Delta v_{s-1}^{(i)})\\ &=Dv_{s-1}^{(i)}+\gamma(y_{s-1}^*-B_{s-1}v_{s-1}^{(i+1)})R_{s-1}-\gamma(y_{s-1}^*-B_{s-1}v_{s-1}^{(i)})R_{s-1}=(E-\gamma B_{s-1}R_{s-1})Dv_{s-1}^{(i)} \end{aligned} \tag{6}$$

where $E$ is the identity matrix. We define $Dv_0^{(i)}=Dm_{N_s}^{(i-1)}$, $I(x_0)=I(x_{N_s})$. For simplicity, we define $E_{s-1}=(E-\gamma B_{s-1}R_{s-1})$, which is a bounded function about $x_s$. Thus, we have

$$\begin{aligned} Dv_s^{(i)}&=E_{s-1}E_{s-2}\ldots E_1 Dv_0^{(i)}=(E_{s-1}E_{s-2}\ldots E_1E_{Ns}\ldots E_s)Dv_s^{(i-1)}\\ &=(E_{s-1}E_{s-2}\ldots E_1E_{Ns}\ldots E_s)^i Dv_s^{(0)} \end{aligned} \tag{7}$$

Define $F_s=(E_{s-1}E_{s-2}\ldots E_1E_{Ns}\ldots E_s)$, then $F_s^i=(E_{s-1}E_{s-2}\ldots E_1E_{Ns}\ldots E_s)^i$. Please note that $Dv_s^{(0)}$ is the accumulation of $\Delta m$, namely

$$Dv_s^{(0)}=\Delta v_s^{(0)}+\Delta v_{s+1}^{(0)}+\ldots+\Delta v_{N_s}^{(0)}+\Delta v_1^{(0)}+\ldots-\Delta v_{s-1}^{(0)} \tag{8}$$

**Theorem 3.** The learning rule (4) (i.e., (3)) has the learning convergence if the number of iterations approaches infinity and the learning rate approaches 0.

*Proof.* In terms of (6), we have

$$v^{(i+1)}s=Dv_s^{(i)}+v_s^{(i)}=Dv_s^{(i)}+Dv_s^{(i-1)}+v_s^{(i-1)}=Dv_s^{(i)}+Dv_s^{(i-1)}+\ldots+Dv_s^{(1)}+Dv_s^{(0)}=\sum_{k=0}^{i}Dv_s^{(k)}F_s^k+v_s^{(0)}$$

Because $v_s^{(0)}-\Delta v_1^{(0)}+\ldots+\Delta v_{s-1}^{(0)}+v_l^{(0)}$, in terms of (8), we further have

$$\begin{aligned}
v_s^{(i+1)} &= \sum_{k=0}^{i}(\Delta v_s^{(0)}+\Delta v_{s+1}^{(0)}+\ldots+\Delta v_{N_s}^{(0)}+\Delta v_1^{(0)}+\ldots+\Delta v_{s-1}^{(0)}F_s^k+v_s^{(0)} \\
&= \sum_{k=0}^{i}\gamma F_s^k[(y_s^*-B_sv_s^{(0)})R_s+(y_{s+1}^*-B_{s+1}v_{s+1}^{(0)})R_{s+1}+\ldots+(y_{N_s}^*-B_{N_s}v_{N_s}^{(0)})R_{N_s}+ \\
&\quad (y_{s-1}^*-B_1v_1^{(1)})R_1+\ldots+(y_{s-1}^*-B_{s-1}v_{s-1}^{(1)}R_{s-1}]+\Delta v_1^{(0)}+\ldots+\Delta v_{s-1}^{(0)}+v_1^{(0)} \\
&= \sum_{k=0}^{i}\gamma F_s^k\Big[\sum_{l=1}^{N_s}y_l^*R_l-B_sv_s^{(0)}R_s-B_{s+1}v_{s+1}^{(0)}R_{s-1}-B_{N_s}v_{N_s}^{(0)}R_{N_s}-B_1v_1^{(0)}R_1-\ldots-B_{s-1}v_{s-1}^{(0)}R_{s-1}\Big]- \\
&\quad v_1^{(0)}+((y_{s-1}^*-B_1v_1^{(1)})R_1+\ldots+(y_{s-1}^*-B_{s-1}v_{s-1}^{(1)})R_{s-1})\gamma \\
&= \sum_{k=0}^{i}\gamma F_s^k\Big[\sum_{l=1}^{N_s}y_l^*R_l-B_sv_s^{(0)}R_s-B_{s+1}v_{s+1}^{(0)}R_{s-1}-B_{N_s}v_{N_s}^{(0)}R_{N_s}-B_1v_1^{(1)}R_1-\ldots-B_{s-1}v_{s-1}^{(1)}R_{s-1}\Big]-v_1^{(0)}+O(\gamma)
\end{aligned} \tag{9}$$

Let us observe:

$$\begin{aligned}
B_jv_j^{(1)}R_j &= B_j(v_{j-1}^{(1)}+\Delta v_{j-1}^{(1)})R_j=B_j(v_1^{(0)}+\Delta v_1^{(0)}+\ldots+\Delta v_{N_s}^{(0)}+\Delta v_1^{(0)}-\ldots+\Delta v_{j-1}^{(0)})R_j \\
&= B_jv_1^{(0)}R_j+O(\gamma u_j)
\end{aligned}$$

where $u_j$ is bounded. Similarly, we have

$$\begin{aligned}
B_jv_j^{(0)}R_j &= B_j(v_{j-1}^{(0)}+\Delta v_{j-1}^{(0)})R_j=B_j(v_1^{(0)}+\Delta v_1^{(0)}+\ldots+\Delta v_{j-1}^{(0)})R_j \\
&= B_jv_1^{(0)}R_j+O(\gamma v_j)
\end{aligned}$$

where $v_j$ is also bounded. Thus, from (9), we have

$$\begin{aligned}
\lim_{\gamma\to0}\lim_{i\to\infty}v_s^{(i+1)} &= \lim_{\gamma\to0}\lim_{i\to\infty}\Big\{\sum_{k=0}^{i}\gamma F_s^k\Big[\sum_{l=1}^{N_s}(y_l^*-B_lv_1^{(0)})R_l+O(\gamma u_j)+O(\gamma v_j)\Big]+v_1^{(0)}+O(\gamma)\Big\} \\
&= \lim_{\gamma\to0}\lim_{i\to\infty}\left\{\gamma(E-F_s)^{-1}\begin{bmatrix}\sum_{l=1}^{N_s}(y_l^*-B_lv_1^{(0)})R_l+O(\gamma u_j)+ \\ O(\gamma v_j)\end{bmatrix}+v_1^{(0)}+O(\gamma)\right\}
\end{aligned} \tag{10}$$

Because

$$F_s=E_{s-1}E_{s-2}\ldots E_1E_{N_s}\ldots E_s=E-\gamma\sum_{l=1}^{N_s}B_lR_l+O(\gamma^2) \tag{11}$$

Therefore, substituting (11) into (10), we obtain

$$\begin{aligned}
\lim_{\gamma\to0}\lim_{i\to\infty}v_s^{(i+1)} &= \lim_{\gamma\to0}\lim_{i\to\infty}\Big\{\Big(\sum_{l=1}^{N_s}B_lR_l\Big)^{-1}\Big[\sum_{l=1}^{N_s}(y_l^*-B_lv_1^{(0)})R_l+O(\gamma u_j)+O(\gamma v_j)\Big]+v_1^{(0)}+O(\gamma)\Big\} \\
&= \Big(\sum_{l=1}^{N_s}B_lR_l\Big)^{-1}\sum_{l=1}^{N_s}(y_l^*R_l)
\end{aligned}$$

This means that the learning rule has the learning convergence, and converges to $\Big(\sum_{l=1}^{N_s}B_lR_l\Big)^{-1}y_l^*R_l$, therefore, this theorem holds.

## 4 Conclusion

In this paper, we present a novel mass-assignment based fuzzy CMAC scheme. Mass-assignment theory can be naturally combined with fuzzy CMAC. We proved that mass-assignment-based fuzzy CMAC is a universal approximator and has the learning convergence. These important properties provide a solid foundation for its wide application potential. Further research work is how to compare this new fuzzy CMAC with other fuzzy CMAC or other CMAC, etc.

**Acknowledgements** This work here is supported by British Royal Society and the National Natural Science Foundation of China.

**References:**

[1] Lin, C. S, *et al.* Learning convergence of CMAC technique. Journal of IEEE Neural Networks, 1997,8(6):1000~1012.
[2] Xu, H, *et al.* Real-time adaptive on-line traffic incident detection. Journal of Fuzzy Sets and Systems, 1997,93(5):173~183.
[3] Wang, Shitong. Neural-Fuzzy Systems and Their Applications. The Publishing House of Beijing Aeronautics and Astronautics University, 1998.
[4] Baldwin, J. F, *et al.* FRIL-Fuzzy and Evidential Reasoning in AI. Longdon: Research Studies Press, John Wiley, 1995.
[5] Wei, Guang, *et al.* Research on global approximation of feedforward neural networks. Journal of Information and Control, 1997,14(2):1~12.
[6] Albus, J. A new approach to manipulator control: the cerebellar model articulation controller (CMAC). Journal of Dynamic Systems, Measurement and Control Transcations on ASME, 1975,97:200~227.
[7] Glanz, F, *et al.* An overview of CMAC neural network. In: Proceedings of the Conference Neural Networks for Ocean Engineering. Washington, 1991. 301~308.
[8] Lane, S, *et al.* Development of high-order CMAC. Journal of IEEE Control Systems Magazine, 1992,12(2):23~30.
[9] Wong, Y, *et al.* Learning convergence of CMAC. Journal of IEEE Neural Networks, 1992,(1):115~121.
[10] Lin, C. T, *et al.* Fuzzy Systems and Fuzzy Neural Networks. New York: Prentice-Hill Press, 1997.
[11] Wang, Shitong, Baldwin, J. F, Martin, T. P. Research on the convergence of fuzzy CMAC. Journal of Advanced Software Research, 1999,11(11):1440~1450.

# 新的基于 mass-assignment 的模糊 CMAC 神经网络及其学习收敛性

王士同[1,2], J.F. Baldwin[2], T.P. Martin[2]

[1](华东船舶工程学院 计算机系,江苏 镇江 212003)
[2](英国 Bristol 大学 高级计算研究中心,英国)

**摘要**: 基于 J. F. Baldwin 等人提出的 mass-assignment 理论,提出了新的基于 mass-assignment 的模糊 CMAC 神经网络,接着研究了其学习规则.理论研究结果揭示出,此新模糊 CMAC 是一个全局逼近器,并且具有学习收敛性.故此新模糊 CMAC 有非常重要的应用潜力.

**关键词**: 模糊 CMAC;学习规则;mass-assignment 理论;学习收敛性

**中图法分类号**: TP18 **文献标识码**: A