第 15 卷第 3 期
1998 年 6 月

控制理论与应用
CONTROL THEORY AND APPLICATIONS

Vol. 15, No. 3
Jun., 1998

A Kind of Adaptive Fuzzy System *

YAO Min

(Department of Computer Science, Hangzhou University • Hangzhou, 310028, PRC)

(State Key Laboratory for Novel Software Technology, Nanjing University • Nanjing, 210093, PRC)

Abstract: By means of the method of combining fuzzy computing with neurocomputing, this paper proposes a kind of fuzzy system model: AFS——Adaptive Fuzzy System. AFS adopts forward neural networks to realize fuzzy reasoning rules, presents fuzzy consistent matrix method to accomplish dynamic adaptation, and raises max-relational membership principle to carry out fuzzy decision making. The performance of AFS is compared with those of six currently representative reasoning methods by application examples.

Key words: fuzzy systems; fuzzy reasoning; dynamic adaptation; fuzzy decision making

1 Introduction

fuzziness is one of the universal attributes of human thinking and objective things. Fuzzy computing, based on fuzzy set theory, is one of the effective methods for simulating partial human brain thought activities, especially the fuzziness of human brain thinking. Zadeh's CRI (Compositional Rule of Inference), Sugeno's P (Position type method) and PG (Position-Gradient type method), Turksen's IVCRI (Interval-valued Version of CRI) and IVAAR (Interval-valued Version of the Approximate Analogical Reasoning method) are current representative reasoning methods for fuzzy system models[1]. All of them consider fuzzy inference as an interpolation from the set of linguistic rules and thus introduce some reasonable properties to the reasoning. However, it is difficult to acquire linguistic rules for the reasoning process. Whereas neurocomputing has strong ability to approximate any function. Therefore, we construct a new fuzzy system model: AFS—— Adaptive Fuzzy System by means of fusing fuzzy computing and neurocomputing together. This paper explains the structure and the working principle of AFS, introduces new methods and new techniques presented in realizing AFS, and gives the application examples of AFS.

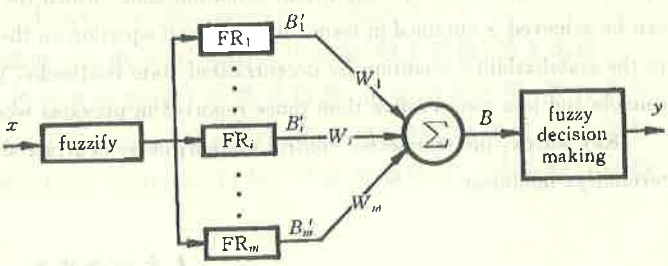

Fig. 1 AFS architecture

2 AFS——Adaptive Fuzzy System

2.1 The Architecture of AFS

The architecture of AFS is shown in Fig. 1, where

* The project is supported by Zhejiang Provincial Natural Science Foudation of China (9609315).
Manuscript received Jun. 4, 1996, revised May13, 1997.

1) $FR_i (i=1,\cdots,m)$ are fuzzy reasoning rules,

$$FR_i = (A_{i1}, A_{i2}, \cdots, A_{it}; B_i)$$

indicates "If x_1 is A_{i1}, x_2 is A_{i2}, $\cdots$, x_t is A_{it}, then y is B_i", where A_{ij} and $B_i (i = 1,\cdots,m; j = 1, \cdots, t)$ are all fuzzy subsets. Because FR_i affects reasoning results directly, and the acquirement of fuzzy reasoning rules lacks solid theory foundation and effective means, we adopt forward neural networks to realize fuzzy reasoning rules $FR_i (i=1,\cdots,m)$. In other words, m forward neural networks (FNNs) are needed to take the place of m fuzzy rules in Fig. 1, while the structural parameters of FNNs are determined by combined learning algorithm with training samples as described in [2], the detail is omitted here.

2) $W = (w_1, w_2 \cdots, w_m)^T$ is a weight vector. In AFS, weights vector W is not fixed, but adjustable according to input situations in the running process of AFS, so that AFS can adapt varied reasoning situations automatically. We will introduce the fuzzy consistent matrix method for determining weights vector W dynamically in Section 2. 2.

3) The task of fuzzy decision making is to accomplish transformation from fuzzy determination to accurate determination. In AFS, fuzzy decision making employs max-relational membership principle, introduced in Section 2. 3, to obtain a precise countermeasure or conclusion from output fuzzy subset B.

The working process of AFS is as follows:

Firstly, by fuzzilizing processing, combined fuzzy set $A = (A_1, \cdots, A_t)$ is produced from input information $X = (X_1, \cdots, X_t)^T$. Secondly, the combined fuzzy set A activates m forward neural networks at the same time, and thus m fuzzy subsets are obtained from the output terminations of FNNs,

$$B_i' = FR_i(A) = (b'_{i1}, \cdots, b'_{iq})^T.$$

Then, the output fuzzy subset B is produced by weighted combination of m fuzzy subsets B_i'

$$B = \sum_{i=1}^{m} W_i' B_i'$$

where

$$W_i' = \begin{cases} W_i, & \text{if } W_i > \theta, \\ 0, & \text{else}. \end{cases}$$

θ is called dynamic adaptive threshold and ranged from 0 to 1. From above equation, we know that if some $w_j < \theta$, then let $W_j' = 0$, i. e. close the jth passageway to the accumulator $\sum$, so as to prevent most incorrelative signals from disturbing output results and increase reasoning precision further.

It is necessary to note that in order to construct AFS, some time is needed to train m forward neural networks, while the working time of AFS is fundamentally real-time.

2. 2 Dynamic Adaptive Principle

As mentioned earlier, in AFS, the weights vector W changes continuously with input information, i. e.

$$W = W(X).$$

This varied weight vector may be determined by fuzzy consistent matrix method.

Definition 2. 1 Let fuzzy matrix $F = (f_{ij})_{m\times m}$, if $f_{ij} + f_{ji} = 1$, then matrix F is called fuzzy mutual-complementing matrix.

Definition 2.2 Fuzzy mutual-complementing matrix $R = (r_{ij})_{m\times m}$ is called fuzzy consistent matrix if satisfying $r_{ij} = r_{ik} - r_{jk} + 0.5$ for any k.

In AFS, the element r_{ij} in fuzzy consistent matrix R indicates the degree of relative importance between B_i' and B_j' for output B, i.e.

1) $r_{ij} = 0.5$, indicates B_i' is as important as B_j';

2) $0 < r_{ij} < 0.5$, indicates B_j' is more important than B_i' and the less r_{ij}, the more important B_j' than B_i';

3) $0.5 < r_{ij} < 1$, indicates B_i' is more important than B_j' and the larger r_{ij}, the more important B_i' than B_j'.

Theorem 2.1 If $R = (r_{ij})_{m\times m}$ is a fuzzy consistent matrix, then

1) $r_{ii} = 0.5$;

2) R^{T} and R^{C} are also fuzzy consistent matrix, where R^{T} and R^{C} are transposition matrix and complementary matrix of R respectively;

3) The sub-matrix, obtained by deleting any row and its correspondent column from R, is also fuzzy consistent matrix;

4) R satisfies center-division transitivity, i.e.

a) When $\lambda > 0.5$, if $r_{ij} \geqslant \lambda$, $r_{jk} \geqslant \lambda$, then $r_{ik} \geqslant \lambda$;

b) When $\lambda < 0.5$, if $r_{ij} \leqslant \lambda$, $r_{jk} \leqslant \lambda$, then $r_{ik} \leqslant \lambda$.

Proof It is not difficult to prove those properties in above theorem based on Definition 2.1 and Definition 2.2. Therefore, only the proof of property 4) is given as follows:

According to definition 2.2, $r_{ij} = r_{ik} - r_{jk} + 0.5$.

a) When $\lambda > 0.5$, by given condition, we have

$$r_{ik} = r_{ij} + r_{jk} - 0.5 \geqslant \lambda + \lambda - 0.5 \geqslant \lambda.$$

b) When $\lambda < 0.5$, by given condition, we have

$$r_{ik} = r_{ij} + r_{jk} - 0.5 \leqslant \lambda + \lambda - 0.5 \leqslant \lambda.$$

It is worthy to note that the significance of property 3) is that when we construct AFS in Fig. 1 by hardware, if some FR_i has not response because of breakdown, reduced rank matrix $R' = (r_{ij}')_{(m-1)\times(m-1)}$ is still fuzzy consistent, thus the consistency of the other weight values is not affected; while the property 4) conforms to the consistency of human decision making thinking, namely,

a) Let $\lambda > 0.5$, if B_i' is more important than B_j' $(r_{ij} \geqslant \lambda)$, B_j' is more important than B_k' $(r_{jk} \geqslant \lambda)$, then B_i' is more important than B_k' certainly $(r_{jk} \geqslant \lambda)$;

b) Let $\lambda < 0.5$, if B_i' is less important B_j' $(r_{ij} \leqslant \lambda)$, B_j' is less important than B_k' $(r_{jk} \leqslant \lambda)$, the B_i' is less important than B_k' certainly $(r_{ik} \leqslant \lambda)$.

Theorem 2.2 Assume $F = (f_{ij})_{m\times m}$ is a fuzzy mutual-complementing matrix, write

$$r_i = \sum_{k=1}^{m} f_{ik}, \quad i = 1, \cdots, m$$

and carry out following transformation

$$r_{ij} = \frac{r_i - r_j}{2m} + 0.5$$

then the matrix $R = (r_{ij})_{m\times m}$ built from above is fuzzy consistent.

Proof

1) $$r_{ij} + r_{ji} = \frac{r_i - r_j}{2m} + 0.5 + \frac{r_j - r_i}{2m} + 0.5 = 1$$

then R is fuzzy mutual-complementary;

2) $$r_{ij} = \frac{r_i - r_j}{2m} + 0.5 = \frac{r_i - r_k - (r_j - r_k)}{2m} + 0.5$$

$$= \frac{r_i - r_k}{2m} + 0.5 - (\frac{r_j - r_k}{2m} + 0.5) + 0.5 = r_{ik} - r_{jk} + 0.5.$$

Thereby R is fuzzy consistent.

Thereupon, the process of determining weight vector by fuzzy consistent method is as follows:

1) For $i = 1,\cdots,m$, find $s(i) = \sigma(B_i, B_i')$;

Where $\sigma(B_i, B_i')$ is the propinquity between fuzzy subsets B_i and B_i'[3].

2) Build priority relation matrix $F = (f_{ij})_{m\times m}$

$$f_{ij} = \begin{cases} 0.5, & \text{if } s(i) = s(j), \\ 1, & \text{if } s(i) > s(j), \\ 0, & \text{if } s(i) < s(j). \end{cases}$$

Clearly, F is a fuzzy mutual-complementing matrix;

3) Reform F into fuzzy consistent matrix R by Theorem 2.2;

4) Obtain weight vector by means of root-squaring method[4].

It is necessary to note that:

1° The weight vector W in AFS may also be obtained by the method of neural network learning[3]. However, there are following disadvantages for this method: a) it is not easy to determine training samples; b) training processing converges hardly; c) for fixed weights, the ability to generalize is weak.

2° It is the special properties of fuzzy consistent matrix, especially center-division transitivity that make fuzzy consistent matrix conform to psychological characteristic of human decision making. Therefore, it is practical to make use of fuzzy consistent matrix method to define dynamic weights.

2.3 Fuzzy Decision Making Method

Fuzzy decision making realizes transformation from fuzzy determination to accurate determination. For the present, maximal membership principle and fuzzy centroid approach are the methods used frequently for fuzzy decision making. Maximal membership principle only considers the maximal component in membership vector and pays no attention to other components so that some information is not taken into account in the decision making. Therefore, it is possible to produce large error, even obtain wrong conclusions. Fuzzy centroid approach makes use of the information provided by all components in membership vector, especially fit to continuous universe situation. However, for a discrete universe, a clear and unambiguous answer is not gained by fuzzy centroid approach directly. Consider the above situation, we propose max-relational membership principle for decision making in our AFS.

Definition 2. 3 Assume F_I is a discrete function set, γ is a mapping, $\gamma: F_I \times F_I \to [0, 1]$, if γ satisfies

1) Normality for any $f \in F_I, \gamma(f,f) = 1$.

2) Symmetry for any $f,g \in F_I, \gamma(f,g) = \gamma(g,f)$.

3) Transitivity for any $f,g,h,s \in F_I$, if $\gamma(f,s) \geqslant \gamma(g,s), \gamma(g,s) \geqslant \gamma(h,s)$ then $\gamma(f,s) \geqslant \gamma(h,s)$,

then the value of γ is called relational grade[5].

Relational grade is the propinquity of discrete functions (or vectors, sequences). Assume $B = (b_1, \cdots, b_q)^T$ in Fig. 1 is a vector to be compared, $B_{i0}(i = 1, \cdots, q)$ are reference vectors

$$B_{i0} = (b_{i10}, b_{i20}, \cdots, b_{iq0})^T$$

where
$$b_{ik0} = \begin{cases} 1, & \text{if} \quad i = k, \\ 0, & \text{if} \quad i \neq k. \end{cases}$$

Clearly B stands for a fuzzy subset over discrete domain $v = \{v_1, \cdots, v_q\}$, but $B_{i0}(i = 1, \cdots, q)$ stand for distinct sets including only one element v_i.

Definition 2. 4 According to definition 2. 3, we define

$$\gamma(B, B_{i0}) = \frac{1}{q}\sum_{k=1}^{q}\xi_i(k)$$

as relational grade between fuzzy vector B and reference vector B_{i0}. Where

$$\xi_i(k) = \frac{\min_i \min_k \Delta_{ik} + \rho \max_i \max_k \Delta_{ik}}{\Delta_{ik} + \rho \max_i \max_k \Delta_{ik}}$$

is relational coefficient in kth component for B and B_{i0}. Where $\Delta_{ik} = |b_k - b_{ik0}|$ is absolute difference in kth component for B and B_{i0}; ρ is resolution coefficient and is a positive real number between 0 and 1.

It is necessary to note that the above calculating formula for relational coefficient may be simplified as

$$\xi_i(k) = \frac{\Delta S + \rho \Delta B}{\Delta_{ik} + \rho \Delta B}.$$

Where

$$\Delta S = \min(\min b_k, 1 - \max b_k), \quad \Delta B = \max(1 - \min b_k, \max b_k).$$

Max-relational membership principle indicates: If there is $k \in \{1, \cdots, q\}$ such that

$$\gamma(B, B_{k0}) = \max\{\gamma(B, B_{10}), \cdots, \gamma(B, B_{q0})\}$$

then the accurate decision may be considered as $y = v_k$.

If we consider different importance for each element in decision process, we may weight relational coefficients and thus obtain weighted relational grade $\gamma_w(B, B_{i0})$.

$$\gamma_w(B, B_{i0}) = \sum_{k=1}^{q} w_k \xi_i(k) = W^T \Xi.$$

Where $\Xi = [\xi_i(1), \xi_i(2), \cdots, \xi_i(q)]^T$ is relation vector, $W = (w_1, \cdots, w_q)^T$ is weight vector and satisfies

$$\sum_{k=1}^{q} w_k = 1.$$

In this case, if there is $k \in \{1, \cdots, q\}$ such that

$$\gamma_w(B, B_{k0}) = \max\{\gamma_w(B, B_{10}), \cdots, \gamma_w(B, B_{q0})\}$$

then distinct decision is $y = v_k$.

3 Application Examples

H. Nakanishi et al. compare six fuzzy reasoning methods with respect to three real-life examples[1]. Now, we apply AFS to these examples and compare its experiment results with those obtained by above six fuzzy reasoning methods in reasoning precision, the number of valid input samples and the degree of difficulty for realizing.

In reference [1], the samples of each example are divided into two groups, one is training set and the other is test set. Training sets are used to determine fuzzy reasoning rules and test sets are used to compare reasoning precision and application scope for varied methods. We follow the division in reference [1] and use training set to train artificial neural networks in AFS. The training process is as follows:

1) Classify all output data in training set by fuzzy clustering approach (such as k- mean algorithm) and thus form m classes;

2) Train m forward neural networks with m class samples respectively and the m trained FNNs are considered m fuzzy reasoning rules.

When training is over, we input the input variable data of test sample to AFS, and obtain actual output $\hat{y}$ at output terminal of AFS.

Let RMSE be root-mean-square error between actual outputs and expected outputs, i. e.

$$\mathrm{RMSE} = [\frac{1}{N}\sum_{i=1}^{N}(\hat{y}_i - y_i)^2]^{1/2}.$$

Where y's are expected outputs and $\hat{y}$'s are actual output.

The reasoning precision Pr is computed as a relative value with respect to Sugeno's P&PG method, $\mathrm{Pr} = \mathrm{RMSE}/\mathrm{RMSE}_0$.

Where RMSE_0 is the RMSE of Sugeno's P & PG.

3.1 Example A——Nonlinear System

There are 25 training samples and 25 test samples in example A. Each sample datum is $(x_1, x_2, x_3, x_4; y)$, where $x_i (i = 1, \cdots, 4)$ is input variables and y is output variable. We classify training samples into $m = 6$ classes by means of fuzzy clustering, and train six neural networks to realize m fuzzy reasoning rules. Then test samples are inputted to AFS. When dynamic adaptive threshold $\theta = 0.13$, the reasoning precision (Pr) together with the number of valid input samples (n) of seven methods are shown in Table 1.

Table 1 Reasoning results of example A

	Sugeno's P&PG	Mamdani's P	Sugeno's PG	Sugeno's IVCRI	Turksen's IVAAR	Turksen's	AFS
Pr	1	1.16	1.58	1.17	1.45	1.45	0.78
n	25	24	24	25	25	20	25

3.2 Example B——Human Operation of a Chemical Plant

There are 35 samples in training set and test set respectively. Each sample datum is $(x_1,$

$x_2, x_3, x_4; y$). We classify training samples into $m = 5$ classes. When $\theta = 0.20$, the reasoning results of seven methods are shown in Table 2.

Table 2 Reasoning results of example B

	Sugeno's P&PG	Mamdani's	Sugeno's P	Sugeno's PG	Turksen's IVCRI	Turksen's IVAAR	AFS
Pr	1	0.90	1.53	2.59	0.94	1.42	0.83
n	35	34	34	35	34	33	35

3.3 Example C——Daily Data of a Stock in Stock Market

There are 50 training samples and 50 test samples in example C. Each sample is ($x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9, x_{10}; y$). By means of fuzzy clustering method, the training samples are divided into $m = 4$ groups. When $\theta = 0.20$, the reasoning results of seven reasoning methods are shown in Table 3.

Table 3 Reasoning results of example C

	Sugeno's P&PG	Mamdani's	Sugeno's P	Sugeno's PG	Turksen's IVCRI	Turksen's IVAAR	AFS
Pr	1	0.66	0.84	1.34	0.99	0.669	0.55
n	50	50	50	50	50	48	50

3.4 Discussion

The reasoning results of various reasoning methods have been shown in Table 1, Table 2 and Table 3, we compare their reasoning precision, number of valid input samples and the degree of difficult for realizing as following.

1) Reasoning precision.

For three real-life examples, our AFS is superior to all other methods to a different extent. If the same number of valid input samples are considered, the precision of example A raised by from 22% (as compared with Sugeno's P&PG) to 46% (compared with Turksen's IVCRI); the precision of example B raised by from 17% to 60%; the precision of example C raised by from 16.5% (compared with Mamdani's) to more than fifty percent (compared with Sugeno's PG).

2) Number of valid input samples.

From above tables, the number of valid input samples of AFS equals to the number of test samples, i.e. in AFS, the reasoning can be carried out for all sample data points.

3) The degree of difficulty for realizing.

The six methods introduced in reference[1] need determining fuzzy reasoning rules from training sets. The approach to obtain fuzzy rules is to make clustering for all output data in training set first and then determine the membership functions of input variables x_i ($i = 1, \cdots, t$) and output variable y from clustering results. This process is very tedious. While in AFS, it is enough to train the neural networks only.

4 Conclusion

According to previous discussion and analysis, we have following conclusions:

1) In constructing AFS, we not only combine neurocomputing with fuzzy computing

close, but also propose many new techniques such as fuzzy consistent matrix technique for realizing dynamic adaptation and max-relational membership principle for fuzzy decision making.

2) Application examples of AFS indicate that our AFS has many advantages such as appropriate precision, better reasoning ability and flexibility to cope with all sample data points, and thus is more perspective.

References

1 Nakanishi, H., Turksen, I. B. and Sugeno, M. A.. Review and comparison of six reasoning methods. Fuzzy Sets and Systems, 1993, 57(6): 257－294

2 Yao, M.. Fuzzy entropy based combined learning algorithm for neural networks. Journal of Systems Engineering and Electronics, 1996, 7(1): 15－22

3 Kosko, B.. Neural Networks and Fuzzy Systems. Englewood Cliffs, NJ: Prentice Hall, 1992

4 Wang, Z. T.. Introduction to System Engineering. Beijing: Electronic Industry Press, 1984 (in Chinese)

5 Deng, J. L.. Relational space of grey systems. Fuzzy Mathematics, 1985, 5(2): 1－10 (in Chinese)

6 Serra, R. and Zanarini, G.. Complex Systems and Cognitive Process. Berlin: Springer-Verlag, 1990

7 Simpson, P. K.. Artificial Neural Systems: Foundations, Paradigms, Applications, and Implementations. New York: Pergamon Press, 1990

一种自适应模糊系统

姚　敏

(杭州大学计算机系·杭州，310028)

(南京大学计算机软件新技术国家重点实验室·南京，210093)

摘要：采用模糊计算与神经计算相结合的方法，本文提出一种自适应模糊系统模型——AFS. AFS 采用前向神经网络来实现模糊推理规则，运用模糊一致矩阵方法实现动态自适应以及最大关联隶属原则执行模糊决策. 最后通过若干实例以说明 AFS 的性能.

关键词：模糊系统；模糊推理；动态自适应；模糊决策

本文作者简介

姚　敏　1954 年生. 分别于 1982 年和 1986 年在合肥工业大学获工学学士学位和工学硕士学位，1995 年在浙江大学获工学博士学位. 现为杭州大学计算机系软件教研室主任、副教授. 发表论文六十余篇. 主要研究方向为计算智能，模糊系统，模式识别和神经网络.