

# 用 HCM 聚类 and 遗传算法实现多级模糊神经网络

赵海军<sup>1</sup> 崔梦天<sup>2</sup>

(西华师范大学计算机学院 南充637002)<sup>1</sup> (西华师范大学数学与信息学院 南充637002)<sup>2</sup>

**摘要** 模糊集理论适用于一些实验数据中不确定性和模糊性的建模问题,而模糊推理系统拥有模糊 IF-THEN 格式的结构化知识表示,但缺少适应性。神经网络本身具有对外部很强的适应性和从过去数据中学习的机制,但基于线性推理的模糊神经网络(FNN)模型作为模糊推理方法不能得到存在于参数间的最终关系,也不能影响接着发生的模糊集合。因此,我们提出了一个多级模糊神经网络(Multi-FNN),使用硬C均值聚类和进化模糊颗粒,利用处理为近似推理的一个线性推理,获得信息微粒和模糊集之间的关系。

**关键词** 多级模糊神经网络,模糊规则,HCM 聚类,遗传算法

## Development of Multi-fuzzy Neural Networks by the Means of HCM Clustering and GAS

ZHAO Hai-Jun<sup>1</sup> CUI Meng-Tian<sup>2</sup>

(School of Computer,China-West Normal University ,Nanchong 637002)<sup>1</sup>

(School of Mathematics and Information ,China-West Normal University ,Nanchong 637002)<sup>2</sup>

**Abstract** Fuzzy sets theory has been introduced to model uncertain and/or ambiguous characteristics in any experimental data. Fuzzy inference system is expressed by the form "if-then", but it lacks of fitness. While the essential advantage of neural networks lies in their adaptive nature and mechanisms of learning from historical data. The drawback of fuzzy neural networks(FNNs) model based on linear inference treated as fuzzy inference method is that eventual relationships existing between the variables cannot be captured in this manner and reflected in the form of the ensuing fuzzy sets. To deal with shortcoming, we propose an idea of Multi-FNNs. They use a Hard C-Means (HCM) clustering algorithm and evolutionary fuzzy granulation and obtain relationship between information granulation and fuzzy sets by the linear inference method treated as approximation inference. The results demonstrated the effectiveness of the proposed model

**Keywords** Multi-Fuzzy-Neural Networks, Fuzzy rule, Hard C-Means clustering, Genetic algorithm

## 1 引言

模糊集合理论提供了系统的、以语言表示不确定和模糊信息的计算工具,通过使用隶属函数表示语言变量,还可以进行数值计算。合理选择模糊规则是模糊推理系统(FIS)的关键因素,它可以有效地对特定应用领域中的人类专门知识进行建模。尽管模糊推理系统拥有模糊 IF-THEN 格式的结构化知识表示,但仍缺少对变化的外部环境进行适应的能力。而神经网络的根本优点是对外部很强的适应性和从过去数据中学习的机制<sup>[1,2]</sup>。规则方面,学习强调的是隶属函数的参数。

我们用基于线性推理的 FNN 模型作为模糊推理方法。基本的 FNN 兼有神经网络的"IF-THEN"规则,可以通过反传算法(BP)获得最优化<sup>[2]</sup>。网络的结构是通过为每一个个体的输入变量划分模糊输入-输出空间来构建的。但是,该方法不能得到存在于参数间的最终关系,也不能影响接着发生的模糊集合。因此,我们提出了一个多级模糊神经网络(Multi-FNN),使用硬C均值(HCM)聚类和进化模糊颗粒,利用被处理为近似推理的一个线性推理,获得信息微粒和模糊集之间的关系。我们讨论了一个目标函数来考虑训练数据和检验数据<sup>[2,3]</sup>,旨在我们所提出的模型的近似和预测能力之间实现一个很好的平衡。最后通过例子验证了所提出的模型的有效性。

## 2 基于数据颗粒的多级 FNN

多级 FNN 的体系是基于一个用 HCM 形成的信息颗粒的一个有效的结构。它是伴随着通过每一个单独 FNN 的输

入参数的空间划分形成的模糊颗粒而产生的。

### 2.1 基于多级 FNN 的线性模糊推理

线性模糊推理的"IF-THEN"类型,其规则的结论部分以它的输入和输出变量之间的一些线性关系来表达。多级 FNN 的基本模型如图1所示。在个体空间(变量)形成的模糊集合形成了 FNN 的预处理块。图1中,当"N"确定为输入变量  $x_i$  应用于隶属度标准化过程时,"圆""方形"指 FNN 单元。"Π"和"Σ"作为乘积和求和神经元。"Σ"神经元的输出  $f_i(x_i)$  通过一些非线性函数  $f_i$  表示。FNN  $\hat{y}$  的输出由下列表达式决定:

$$\hat{y} = f_1(x_1) + f_2(x_2) + \dots + f_m(x_m) = \sum_{i=1}^m f_{i1}(x_i) \quad (1)$$

其中  $m$  为输入变量的数量。FNN 由每个输入变量的模糊划分表示。从这个意义,可以把式(1)所给出的  $f_i$  看成下列的规则。

$$R^1: \text{IF } x_i \text{ is } A_{i1} \text{ then } Cy_{i1} = w_{0i1} + x_i w_{ai1}$$

⋮

$$R^j: \text{IF } x_i \text{ is } A_{ij} \text{ then } Cy_{ij} = w_{0ij} + x_i w_{aij}$$

⋮

$$R^n: \text{IF } x_i \text{ is } A_{in} \text{ then } Cy_{in} = w_{0in} + x_i w_{ain}$$

更具体地说,当  $a_{ij}$  表示一个相应的模糊规则的一个前提的模糊变量时,  $R_j$  是第  $j$  个模糊规则。它代表图2所示的隶属函数  $\mu_j$ 。  $w_{0ij}$  和  $w_{aij}$  包含相应的模糊规则的结果,表示存在于神经元之间的联系(图1所示)。此外,我们仅限于研究三角隶属函数并使隶属度累加到1(因此,这可导致变量的模糊划分)。在模糊规则前提部分的每个隶属函数用邻近的一个数赋

值成互补。形式如图2所示<sup>[4]</sup>。

从  $x_i$  到  $f_i(x_i)$  的数值对应关系由模糊推理来决定：

$$f_i(x_i) = \frac{\sum_{j=1}^n \mu_j(x_i)(w_{0,j} + x_i w_{a,j})}{\sum_{j=1}^n \mu_j(x_i)} \quad (3)$$

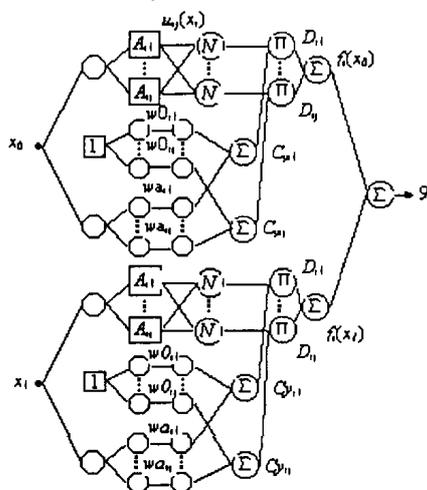


图1 基于线性模糊推理的FNN结构

考虑到模糊规则划分的特性,输入信号  $x$  只活动于通过  $k$  和  $k+1$  标记的两个邻近的隶属函数之间(图2)。然后式(3)表达成如下：

$$f_i(x_i) = \mu_k(x_i)(w_{0,k} + x_i w_{a,k}) + \mu_{k+1}(x_i)(w_{0,k+1} + x_i w_{a,k+1}) \quad (4)$$

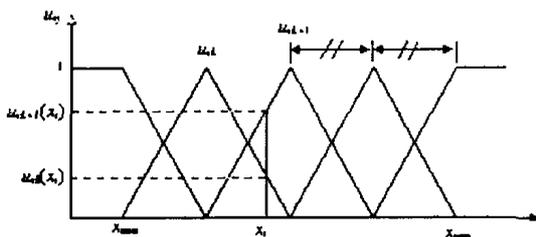


图2 进行基因优化前的MF

FNN 的学习是通过调整神经元的联接实现的。其值的修改,  $w_{0,j}$  和  $w_{a,j}$  通过一个标准的反传算法(BP)实现。在这个研究中,使用了2个量(行为指数)。

i) 使用欧几里得几何空间误差作为一个行为量：

$$E_p = (y_p - \hat{y}_p)^2 \quad (5)$$

这里,  $E_p$  是第  $p$  个数据的一个误差,  $\hat{y}_p$  是第  $p$  个目标输出数据,  $y_p$  代表针对这个具体的数据点,第  $p$  个模型的实际的输出。对于  $n$  个输入-输出数据组,全部的行为指数由误差和给出。

$$E = \frac{1}{n} \sum_{p=1}^n (y_p - \hat{y}_p)^2 \quad (6)$$

ii) 基于百分比的平均数误差采用的形式：

$$E = \frac{1}{n} \sum_{p=1}^n \frac{|y_p - \hat{y}_p|}{y_p} \times 100\% \quad (7)$$

就学习而言,联接发生变化为：

$$wa(new) = wa(old) + \Delta wa \quad (8)$$

这里,用梯度下降法表示,即

$$\Delta wa_{ij} = \eta_e \left( -\frac{\partial E_p}{\partial wa_{ij}} \right) \quad (9)$$

其中  $\eta_e$  作为一个正学习率,自身的激励函数定义为：

$$m(t) = wa_{ij}(t) - wa_{ij}(t-1) \quad (10)$$

最终更新后的公式如下：

$$\Delta wa_{ij} = 2\eta_e (y_p - \hat{y}_p) \mu_j(x_i) + \alpha_e (wa_{ij}(t) - wa_{ij}(t-1)) \quad (11)$$

其中  $\alpha_e$  限制在一个单元区间。

## 2.2 多级 FNN 结构

通常的 FNN 用 HCM 聚类方法通过系统输入空间的划分决定一些隶属函数的初始参数值。我们研究了特别的 FNN,通过 HCM 聚类构建数据群。群的数量相当于模型数量(FNN 结构),如图3所示。

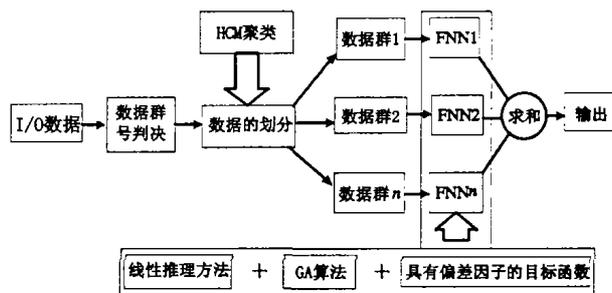


图3 多级模糊神经结构

## 2.3 HCM 聚类方法

HCM 聚类方法不仅广泛用于数据的组织和分类,也用于数据压缩和模型构建。用该方法生成的多级 FNN 结构的信息数据微粒按下列次序产生<sup>[2]</sup>。

i) 按照下列 HCM 聚类算法将训练数据划分为  $c$  个组。

步骤1: 固定组的数量  $c(2 \leq c < n)$ , 初始化划分矩阵  $U^{(0)} \in M_c$ 。

$$M_c = \{U | u_{ik} \in \{0, 1\}, \sum_{i=1}^c u_{ik} > 0, \sum_{k=1}^n u_{ik} < n\} \quad (12)$$

步骤2: 计算每个组的中心。

步骤3: 更新划分矩阵  $U^{(r)}$ , 这些修改按照欧几里得几何空间函数,在数据点和聚类中心之间进行。

步骤4: 检测终止条件。IF  $(\|U^{(r+1)} - U^{(r)}\| < \epsilon)$  停止,否则令  $r=r+1$ , 返回步骤2。

ii) 计算训练数据的中心矢量与检验数据的元素间的距离。

iii) 划分基于预先构成的组中心的检验数据。

## 3 基于进化模糊颗粒的 FNN 的最优化

优化任何一个复杂模型有两个目的:一是必须选择一些优化算法以适应于着手要解决的问题的需求。二是优化算法的各种参数需要调谐以便能更好地执行算法。

GAs 是不严格地建立在自然选择和自然进化概念基础上的一种非导数的随机优化技术。从本质上说,它们是并行搜索算法,是在自然界基因中发现的操作,通过参数空间来指导搜索。无论理论上还是实践上, GAs 都已被证明在复杂空间中,能提供强大的搜索能力,提供了有效的方法解决高效搜索的问题。

### 3.1 遗传算法(GA)

GAs 算法是基于种群数量最优化技术。在 GAs 中,有3个基本操作:选择,交换,变异<sup>[2,6]</sup>。选择是从当前一代生成一组新的种群。选择操作决定哪些个体参与生成下一代的个体。个体基因链按照它们的适应度函数值被复制到交配池中。交换处理分两步:1)来自交配池的成员随机交配。2)每对基因链按照下列方式交换:位置  $L$  沿着基因链在区间  $[1, l-1]$  随机被

选择,这里,  $l$  是链的长度。两个链通过在位置  $K$  和  $L$  之间交换产生。变异,是基因链位置的值随机交替的过程。变异能防止整个种群在任何局部的任一位置收敛到一个值,并且更为重要的是,它能防止种群收敛并停滞在任何局部点。因此,由交换操作中得到的好染色体不会丢失。遗传算法如下:

```

Begin
  t=0;
  Initialize P(0) := {a1(0), ..., an(0)} ∈ In
  Evaluate Fitness:
  {
    ζ, if  $\sum_{k=1}^N P_k \geq 0$ 
     $\sum_{k=1}^N P_k \geq 0$ , otherwise
  }
  Where Pk = Amin(Θk(x)) (k=1, ..., N)
  While (Fitness ≠ termination condition) Do
  Begin
    recombine ak(t) := r(P(t))
    mutate ak(t) := m(ak(t))
    evaluate P'(t) := Ft(x(t)) ∪ Ft(x'(t))
    select P(t+1) from P(t):
      IF (Ft(x(t)) < Ft(x'(t)))
        x(t+1) = x'(t)
      Else
        x(t+1) = x(t)
    t=t+1
  End
End
  
```

在种群  $P(t)$  中,被重组的种群  $a_k(t)$  将通过重组操作  $r$  获得。变异操作数  $m$  将生成  $a_k(t)$ ,  $P'(t)$  是这两个操作之后,由  $x'(t)$  个体组成的新的种群。下一个种群中的个体,如  $x(t+1)$ ,将通过估计适应度函数  $F(t)$  来选择。

为优化 FNN 模型,本文在选择操作数时,GA 算法使用二进制数,旋转赌轮法;在交叉操作数中进行一点交叉;在变异操作数中使用倒置法。这里,我们用 60 代,种群规模为 40,每个基因链有 8 位,交叉概率为 0.7,变异概率为 0.1,图 4 为用于 FNN 的 GA 递阶结构。这里,变量  $x_1, x_2, \dots, x_k$  表示输入变量 FNN 模型,变量  $\eta, \alpha$  分别表示学习率和激励系数。

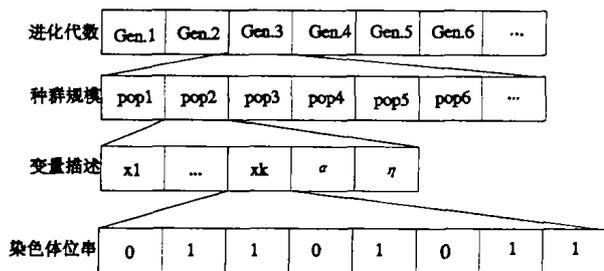


图4 用 GA 实现的 FNN 模型的最优化结构

### 3.2 带有偏差因子的目标函数

系统建模的通常方法是使用一些训练数据建构模型,然后使用测验数据进行评价<sup>[4]</sup>。但这样就存在一个问题:这不能保证所要求的行为是令人满意的,因为新建的模型是针对训练数据定制的,结果会产生近似误差,并且应进一步减少把模型作为预测器使用的情况。因此,下列的目标函数被用来减少误差,并增加模型的预测能力。因此目标函数应包含:训练数据的行为指数,检验数据的行为指数和偏差因子<sup>[5]</sup>。

目标函数包含训练数据的行为指数 ( $PI$ )、检验数据的行为指数 ( $E-PI$ ) 和作为两个指数 ( $(PI)$ 、 $(E-PI)$ ) 的凸的和  $\theta$ 。

$$f(PI, E-PI) = \theta PI + (1-\theta)E-PI \quad (13)$$

$PI$  和  $E-PI$  表示划分训练数据和检验数据的行为指数。此外,  $\theta$  是偏差因子,它容许我们打破实现模型的训练数据和检验数据行为之间的平衡。依据偏差因子,几个目标函数的具体

情况是可以区分的。

$$\text{Case 1. } 0=1: f(PI, E-PI) = PI$$

这个情况中,目标函数变成  $f(PI, E-PI) = PI$ ,模型优化是基于训练数据,而检验数据不考虑进去。这个情况表明,突出近似能力而预测能力相对近似能力变得低些。

$$\text{Case 2. } 0=1: f(PI, E-PI) = E-PI$$

这个情况中,目标函数变成  $f(PI, E-PI) = E-PI$ ,从检验数据得到的模型优化通过训练数据建构,然后,从  $E-PI$  那里得到优化。该效果表明较低的近似能力,而预测能力相对 Case 1.0 增加了。

$$\text{Case 3. } 0=1: f(PI, E-PI) = 0.5 PI + 0.5 E-PI$$

$PI$  和  $E-PI$  都被考虑。效果表明有较低的近似能力(与 Case 1.0 比),有较低的预测能力(与 Case 1.0 和 Case 2.0 比)。

$$\text{Case 4. } 0 = \alpha (\alpha \in [0, 1]): f(PI, E-PI) = \theta PI + (1-\theta) E-PI$$

$PI$  和  $E-PI$  都被考虑。恰当地选择  $\theta$  可实现近似和预测能力之间的平衡。这种情况下,  $PI$  从模型的训练数据中得到,  $E-PI$  从该模型的检验数据中得到。模型选择是通过优化与 FNN 相关的参数,从目标函数最小化中完成。

当考虑带有很强的非线性复杂系统时,如果偏差因子  $\theta$  选择很小的值(与 Case 3.0 相比),那么目标函数将几乎等于 Case 2.0。而且在这个情况下,  $PI$  的效果(近似能力)将减小。其次,这个方法既考虑了近似能力,又考虑了预测能力,选择恰当的模型满足需要须通过选择合适的偏差因子。

## 4 模拟和结果的讨论

我们用著名的 Box-Jenikin 实验<sup>[7]</sup>这个例子来评价所提出的方法的优点和有效性。这是一个煤气炉过程的时间序列数据集。煤气炉的输入  $u(t)$  为煤气流量,  $CO_2$  浓度  $y(t)$  为炉子的输出。用到的行为指数 ( $PI$ ) 为:式(6)为煤气炉过程数据和  $CO_2$  气体过程数据,式(7)为 3 个有效的非线性函数数据。

我们用 296 组  $[u(t), y(t)]$  数据。甲烷流速  $u_m(t)$  范围从  $-2.5 \sim 2.5$ , 煤气流量  $u(t)$  范围为  $0.5 \sim 0.7$ , 要遵循下列公式:

$$u(t) = 0.6 - 0.048 \times u_m(t) \quad (14)$$

建模时,我们用 2 个输入  $u(t-1)$  和  $y(t-1)$  及一个输出  $y(t)$ 。这里,使每个输入变量形成颗粒的模糊集的数量为 3。总共的数据集分为 2 部分:即训练数据集(145 个)和检验数据集(145 个)。模型的结构和参数是用 HCM 聚类方法和 GA 算法实现的。遗传算法通过自调谐 FNN 模型参数如隶属函数的峰值、学习率、激励系数提高模型的执行。图 5 为  $c=4$  时的聚类结果。

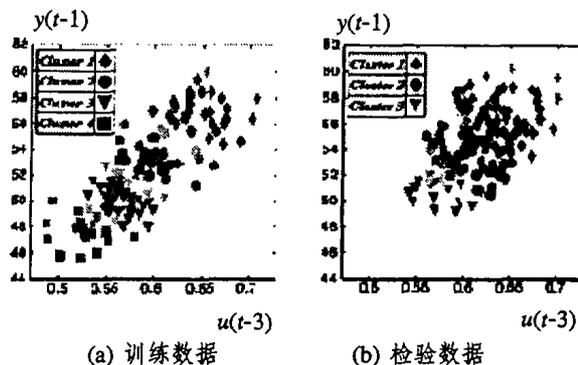


图5 用 HCM 聚类方法划分数据

图6结果表明,使用 HCM 聚类算法划分数据要比使用任

意数据划分方法会得到更好的多级模糊神经网络执行。

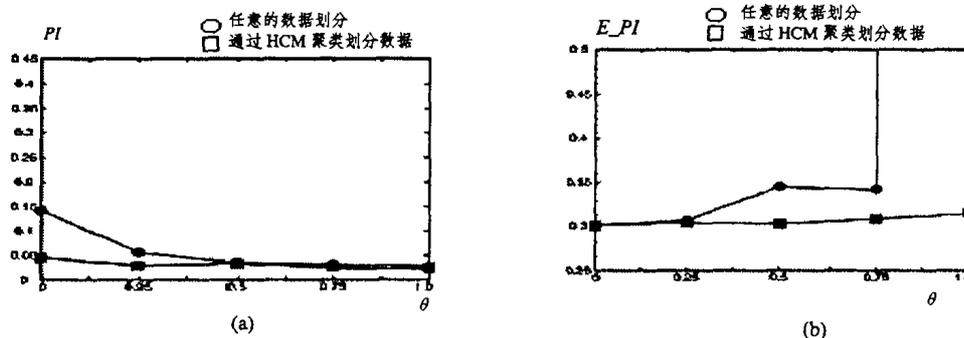


图6 多级模糊神经网络的行为指数(a)(b)

**结论** 本文提出了一个作为并行结构的模糊模型神经网络的多级 FNN。多级 FNN 包含多种技术:聚类 HCM,基于 FNN 和 GAs 算法的线性模糊推理。HCM 聚类可帮助构建结构优良有效的模型。处理信息颗粒的优化问题最重要的思想是通过利用聚类和进化计算。所提出的行为指数可帮助完成近似与预测能力之间的平衡。实验结果表明了该模型的有效性。

参考文献

1 刘保路. 随机规划与模糊规划[M]. 北京:清华大学出版社,1998  
 2 张智星. 神经-模糊与软计算[M]. 西安:西安交通大学出版社,

2000  
 3 Dubois D, Prade H. Fuzzy sets and systems: theory and applications. New York: Academic press, 1990  
 4 Bezdek J C. Weighted fuzzy pattern matching. Fuzzy Sets and Systems, 1988, 28: 313~331  
 5 Oh Sung-Kwun, et al. Rule-based multi-FNN identification with the aid of evolutionary fuzzy granulation. Knowledge-Based Systems 17, 2002  
 6 Goldberg D E. Genetic algorithms in search, optimization, and machine learning. Addison-Wesley, Reading, MA, 1989  
 7 Box, Jenkins. Time Series Analysis, Forecasting and Control, Holden Day, San Francisco, CA, 1976. 532~533

(上接第225页)

文[9]中提出对于 ANFIS 进行建模,方法是通过训练样本来拟合非线性函数,通过 ANFIS 优化隶属函数,本文也对这种情况进行仿真,本文对函数  $y = 1.8\sin(\pi x) + 0.2\sin(5\pi x)$  进行拟合,采用五个隶属函数变量。初始的隶属函数参数如表1,采用本文的隶属函数优化方法,优化后的隶属函数保持了完整性和语义性。拟合的结果如图10,优化前后得到的隶属函数参数如表2。

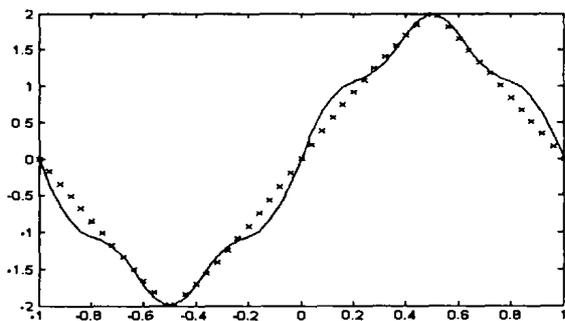


图10

表2

隶属函数	NB	NS	ZO	PS	PB
优化前	$[-1 \ -1 \ 0]$	$[-1 \ -0.5 \ 0]$	$[-0.5 \ 0 \ 0.5]$	$[0 \ 0.5 \ 1]$	$[0.5 \ 1 \ 1]$
优化后	$[-1 \ -1 \ -0.499]$	$[-1 \ -0.4824 \ 0]$	$[-0.5033 \ 0 \ 0.5118]$	$[-0.0177 \ 0.5012 \ 0.9745]$	$[0.4949 \ 1 \ 1]$

**结论** 本文在对用梯度下降法、GA 等模糊系统学习方法优化隶属函数时存在的问题进行深入分析的基础上,针对一类简单、有效、并广泛使用的三角形隶属度函数,提出了一种保证隶属度函数 e-完备性和语义性的参数调整的方法。通过实例进行了模拟,验证了该方法的可行性,以及保证模糊系统隶属度函数完备性和模糊集合语义一致性的优点。

参考文献

1 刘曙光,等. 模糊控制技术. 北京:中国纺织出版社,2001  
 2 冯冬青,谢宋和. 智能控制技术. 北京:化学工业出版社,1998  
 3 Lotfi A, Tsoi A C. Learning fuzzy inference system using an adaptive membership function scheme. IEEE transaction on system, man and cybernetics-part B, 1996, 26: 326~331  
 4 Tarnq Y S, Yeh Z M, Nian C Y. Genetic synthesis of fuzzy logic controllers in tuning. Fuzzy Sets and Systems, 1996, 83: 301~310  
 5 Shimojima O, Fukuda T, Hasegawa Y. Self-tuning fuzzy modeling with adaptive membership function, rules and hierarchical structure based on genetical algorithm. Fuzzy Sets and Systems, 1995, 71: 295~309  
 6 Farag W A, Quintana V H, Lambert-Torres G. A genetic-based neuro-fuzzy approach for modeling and control of dynamical systems. IEEE transaction on neural networks, 1998, 9: 756~767  
 7 张智星,孙春在,水谷英二. 神经-模糊和软计算. 西安交通大学出版社,2000  
 8 武研,等. 一种保证隶属度函数完备性和模糊集合语义一致性的学习方法. 计算机研究与发展, 1999, 9: 1080~1085  
 9 侯北平,卢佩,等. MATLAB 下模糊控制器的设计与应用. 测控技术, 2001, 20: 40~42