

基于粗-模糊神经网络的决策控制

张倩生

(广东外语外贸大学 信息科学技术学院, 广东 广州 510420; 中山大学 数学与计算科学学院, 广东 广州 510275)

摘要: 通过将模糊集和粗集, 神经网络结合, 提出了一种基于模糊规则的新的粗模糊神经网络, 它通过利用误差反向传播算法实时修正该新型网络中的权值参数, 从而能被有效地应用于不确定系统的决策分类与模式识别问题. 最后通过对一个不确定决策系统的模式识别的仿真结果表明该粗模糊神经网络能大大提高模式识别决策的准确率.

关键词: 粗集; 模糊神经网络; 决策表; 模式识别

中图分类号: TP18 **文献标识码:** A

Decision control based on rough-fuzzy neural network

ZHAN Qian-sheng

(School of information science and technology, Guangdong University of Foreign Studies, Guangzhou Guangdong 510420, China;

School of Mathematics and computing science, Zhongshan University, Guangzhou Guangdong 510275, China)

Abstract: A new rough-fuzzy neural network based on extracting of fuzzy rules is proposed by integrating fuzzy set and rough set with neural network. It can be effectively used in decision controlling for any uncertain decision system by timely correcting the parameter value of network with error back propagation algorithm. The simulation results of a model recognition indicates that the construction approach of the rough-fuzzy neural network can improve model recognition probabilities.

Key words: rough set; rough-fuzzy neural network; decision table; pattern recognition

1 引言 (Introduction)

对具有不确定性因素对象的这个极富挑战性的控制问题, 目前主要有自适应控制和鲁棒控制法, 但我们注意到在应用上述方法时, 往往要求将控制对象描述为某些标准的数学模型, 而客观实际中要处理的控制对象往往具有严重的不确定性, 无法精确建模, 近年来兴起的人工神经网络方法因其具有模拟人脑的形象直觉思维的能力, 且具有较强的鲁棒容错性和并行处理能力所以在解决高度非线性, 具有大量不确定性的系统控制方面显示了极大潜力. 而针对不确定, 不完整的数据对象控制方面, 波兰学者 Z. Pawlak 早在 1982 年就提出了粗集理论这种有效的智能数据分析工具, 因其具有极大的模拟人类抽象逻辑思维功能, 故其在对不完整, 不确定数据对象进行分析推理简化, 知识发现提取方面已受到模式识别, 决策系统控制, 机器学习等许多领域的重视. 本文正基于此, 用模糊规则提取的方法构建了一

种具有广泛应用前景的新的粗-模糊神经网络模型, 用来对任意复杂的不确定对象实现自学习或自优化的控制决策识别. 其基本思想是先利用粗-模糊神经网络的映射功能对不确定对象过程实时建模, 然后根据误差反向传播梯度算法实时修正网络模型的权值参数和调整因子, 从而实现对复杂不确定对象的自学习控制识别. 最后用一个模式识别的仿真试验结果表明, 该粗-模糊神经网络模型是可行的且具有很高的正确识别率.

2 智能决策系统简介 (Intelligent decision system description)

设智能决策系统表 $T = (U, A, C, D)$, 其中 A 为属性集, C 为其中的条件属性集, D 为决策属性 (见文献[1]); 而作者知道智能决策系统中的规则挖掘是基于由一定实验数据所得出的决策表, 然而决策表是典型的不完备信息系统, 针对于此, Z. Pawlak 提出了粗集理论这种高效的智能数据处理工

具,但智能系统学习能力提高的标志在于其不介入新的程序设计器前提下获取和发现知识规则,并用其所获规则来提高本身技能。

下面简单介绍用粗集理论提取模糊规则.任取条件属集 C 的属性子集 B ,由不可分辨关系 $R = IND(B)$ 可得论域 U 关于 B 的一个模糊划分 $B^* = \{\tilde{B}_1, \tilde{B}_2, \tilde{B}_3, \dots, \tilde{B}_k\}$,而由决策属性 D 也可得到论域 U 关于决策 D 的一个模糊划分 $D^* = \{\tilde{D}_1, \tilde{D}_2, \tilde{D}_3, \dots, \tilde{D}_r\}$,模糊划分的定义参见文献[2].则可由 FRS 理论^[3]得:对 $\forall i = 1, 2, \dots, k; \forall t = 1, 2, \dots, r$ 有

$$\mu_{R(\tilde{D}_t)}(\tilde{B}_i) = \inf_{x \in U} \max\{1 - \tilde{B}_i(x), \tilde{D}_t(x)\}$$

表示模糊类 \tilde{B}_i 在模糊决策类 \tilde{D}_t 中的必然隶属度;

$$\mu_{R(\tilde{D}_t)}(\tilde{B}_i) = \sup_{x \in U} \min\{\tilde{B}_i(x), \tilde{D}_t(x)\}$$

表示模糊类 \tilde{B}_i 在 \tilde{D}_t 类中的可能隶属度。

故而上面两个隶属度分别表示了模糊规则“ $\text{If } \tilde{B}_i \text{ Then } \tilde{D}_t$ ”成立的确信度和可能度。

显然有 $\mu_{R(\tilde{D}_t)}(\tilde{B}_i) \leq \mu_{R(\tilde{D}_t)}(\tilde{B}_i)$ (见文献[1]);若 $\mu_{R(\tilde{D}_t)}(\tilde{B}_i) = 1$,则模糊规则必然真,若 $\mu_{R(\tilde{D}_t)}(\tilde{B}_i) = 0$ 则规则必然假.若 $\mu_{R(\tilde{D}_t)}(\tilde{B}_i) > 0$,则可能真.又由于粗集理论只能处理离散型不确定数据,所以当信息决策表中的属性值是连续的时,必须在用粗集处理样本之前,将属性值量化,即离散化,一般我们可采用全局聚类法^[4]或动态层次聚类法^[5]来进行属性值的离散归一化处理,就将得到离散化的决策信息表,然后我们才再用粗糙集理论将其进行条件属性的简化和条件属性值的简化,获取最小决策算法^[1],得到尽可能最少条数的决策规则。

3 基于粗集理论的粗-模糊神经网络的构造和学习控制方法 (Creating rough-fuzzy neural network and its learning algorithm based on rough set theory)

由于粗集方法对噪音较敏感,而神经网络又具有强大的抗干扰和非线性并行处理能力.所以可将它们结合起来,并用上面所获取的最少规则来构造神经网络。

3.1 网络结构 (Rough-fuzzy neural network structure)

根据以上第 2 节中所述思路,假设 $T = (U, A, C, D)$ 经最小简化后条件属性只剩 m 个,决策属性值离散化分为 N 类.本文可给出新构造的粗-模糊神经网络,它描述了一个多输入多输出系统,网络共 7 层:

第 1 层 输入层.表示输入样本 $X_p = (x_{p1}, x_{p2},$

$\dots, x_{pm})^T$,其中 x_{pi} 为样本 P 的第 i 个条件属性值。

第 2 层 模糊化层,即分别将这个 m 个输入量 $(x_{p1}, x_{p2}, \dots, x_{pm})$ 依某种不可分辨关系进行模糊划分后得到其模糊化值,也即是将 x_{pi} 转化为 r_i 个不同的模糊化值, r_i 是决策表离散化后第 i 个条件属性离散化后的个数,可定义该层神经元的作用函数为

$$\tilde{A}_i^j(x_{pi}) = e^{-\frac{(x_{pi} - b_{ij})^2}{\sigma_{ij}}}, i = 1, 2, \dots, m, j = 1, 2, \dots, r_i.$$

\tilde{A}_i^j 即为第 i 个条件属性所对应的第 j 个模糊类;其中 b_{ij}, σ_{ij} 是模糊隶属函数的均值与方差,可由专家经验,竞争学习,遗传算法等方法进行预确定。

第 3 层 求与层,即该层每个节点代表一个模糊规则的前件部分,节点实现最小化,即模糊与操作,用来匹配模糊规则前件,与之连接的第 2 层节点为该条规则前件中所含属性值的模糊化节点.实际上该层每一个节点代表一个由第二节介绍的粗集方法生成的最小决策算法中的某一条模糊规则的 IF 部分.总的节点个数恰为最优规则的条数,每个节点对应一个模糊类 \tilde{Z}_k , (其中 $\tilde{Z}_k = \bigcap \tilde{A}_i^j$); s 为第 k 条规则前件中所含的条件属性, j 为条件属性 s 的第 j 个离散化。

第 4 层 信息综合层.该层总的节点个数为 N 个,且该层每一个节点代表对具有相同决策类属(后件)的模糊规则 IF 部分的综合信息运算,其作用函数为 $\tilde{U}_t = T(\tilde{Z}_{tk})$, \tilde{Z}_{tk} 是与该层第 t 个节点相联结的上层节点,它可能有多个(因为对于同一个决策类可能是由不同的规则前件导致的),若只有一个时则 $\tilde{U}_t = \tilde{Z}_{tk}$;若有多个时则 \tilde{U}_t 为各个 \tilde{Z}_{tk} 的信息综合运算; T 为信息综合运算符号^[6]。

第 5 层 由 N 个粗糙神经元组成(N 为按决策属性对论域所划分的类的个数),该层每个粗神经元节点对应一个模糊决策类 \tilde{D}_t ,其输出分两部分:粗神经元的上输出为 $\bar{R}_{\tilde{D}_t}(\tilde{U}_t)$,粗神经元的下输出为 $R_{\tilde{D}_t}(\tilde{U}_t)$

第 6 层 与第 5 层相对应的粗神经元层,该层的每个粗神经元与其相对应的第 5 层粗神经元全连接,有 4 条连接权;例如 \tilde{V}_t 与 \tilde{D}_t 间的 4 条连接权分别为: ω_t^{11} 表示 \tilde{D}_t^1 与 \tilde{V}_t^1 的连接权, ω_t^{12} 表示 \tilde{D}_t^1 与 \tilde{V}_t^2 的连接权, ω_t^{21} 表示 \tilde{D}_t^2 与 \tilde{V}_t^1 的连接权, ω_t^{22} 表示 \tilde{D}_t^2 与 \tilde{V}_t^2 的连接权.该层第 t 个节点 \tilde{V}_t 的上粗神经元 \tilde{V}_t^1 输入为 $\text{input}(\tilde{V}_t^1) = \omega_t^{11}\bar{R}_{\tilde{D}_t}(\tilde{U}_t) + \omega_t^{21}R_{\tilde{D}_t}(\tilde{U}_t)$,

下粗神经元 \tilde{V}_t^2 的输入为 $\text{input}(\tilde{V}_t^2) = \omega_t^{12}\bar{R}_{\tilde{D}_t}(\tilde{U}_t) + \omega_t^{22}R_{\tilde{D}_t}(\tilde{U}_t)$.其相对应的输出为

$$\begin{aligned} \text{output}(\tilde{V}_i^1) &= \max(f_1(\text{input}(\tilde{V}_i^1)), f_2(\text{input}(\tilde{V}_i^2))), \\ \text{output}(\tilde{V}_i^2) &= \min(f_1(\text{input}(\tilde{V}_i^1)), f_2(\text{input}(\tilde{V}_i^2))). \end{aligned}$$

本文取 $f_1(x) = f_2(x) = \frac{1}{1 + e^{-x}} \in [0, 1]$.

第7层 普通神经元层,共 N 个.该层第 t 个节点输入为 $\text{output}(\tilde{V}_i^1), \text{output}(\tilde{V}_i^2)$, 输出即为 $y_{pt} = \frac{\text{output}(\tilde{V}_i^1) + \text{output}(\tilde{V}_i^2)}{2}$.

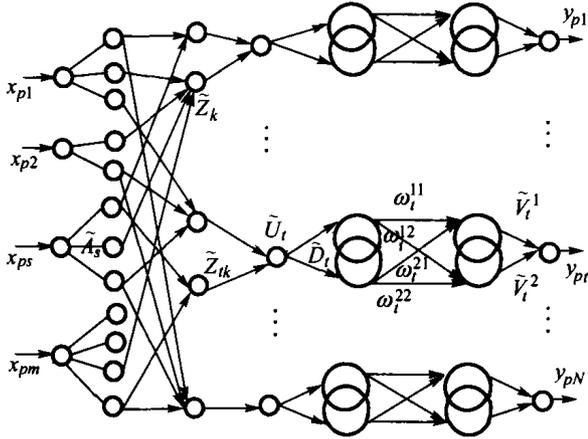


图1 粗模糊神经网络模型结构

Fig. 1 Structure of rough-fuzzy neural network

3.2 自适应控制学习算法 (Self-adaptive learning algorithm)

在本网络模型中,只需通过对6,7层间连接权的学习,即可获得对给定问题的描述,下面是网络误差反向传播梯度学习算法.

假设用于训练粗模糊神经网络的训练样本是选取了 l 个样本.则可建立该网络平均误差代价函数为

$$\begin{aligned} E &= \frac{1}{2l} \sum_{p=1}^l \sum_{t=1}^N (\hat{y}_{pt}(x_p) - y_{pt}(x_p))^2, \\ t &= 1, 2, \dots, N. \end{aligned}$$

其中 $y_{pt}(x_p)$ 为样本的实际计算输出, $\hat{y}_{pt}(x_p)$ 为样本的理想期望输出;

当 x_p 属于决策类 t 类时, $\hat{y}_{pt}(x_p) = 1$; 当 x_p 属于其他类情况时, $\hat{y}_{pt}(x_p) = 0$, 则第6,7层间的网络权值修正如下:

$$\begin{aligned} \Delta \omega_i^{11} &= -\eta_1 \frac{\partial E}{\partial \omega_i^{11}} = \\ &= -\eta_1 \frac{\partial E}{\partial y_{pt}} \cdot \frac{\partial y_{pt}}{\partial f(\text{input} \tilde{V}_i^1)} \cdot \frac{\partial f(\text{input} \tilde{V}_i^1)}{\partial (\text{input} \tilde{V}_i^1)} \cdot \frac{\partial (\text{input} \tilde{V}_i^1)}{\partial \omega_i^{11}} = \\ &= -\eta_1 \sum_{p=1}^l \sum_{t=1}^N -\frac{2}{2l} (\hat{y}_{pt}(x_p) - y_p(x_p)) \cdot \\ &= \frac{1}{2} \frac{e^{-\text{input} \tilde{V}_i^1}}{(1 + e^{-\text{input} \tilde{V}_i^1})^2} \cdot \bar{R}_{\tilde{D}_i}(\tilde{U}_i) = \end{aligned}$$

$$\begin{aligned} &= \frac{\eta_1}{2l} \sum_{p=1}^l \sum_{t=1}^N (\hat{y}_{pt}(x_p) - y_p(x_p)) \cdot \\ &= \frac{e^{-\text{input} \tilde{V}_i^1}}{(1 + e^{-\text{input} \tilde{V}_i^1})^2} \cdot \bar{R}_{\tilde{D}_i}(\tilde{U}_i), \\ \Delta \omega_i^{12} &= -\eta_2 \frac{\partial E}{\partial \omega_i^{12}} = \\ &= -\eta_2 \frac{\partial E}{\partial y_{pt}} \cdot \frac{\partial y_{pt}}{\partial f(\text{input} \tilde{V}_i^2)} \cdot \frac{\partial f(\text{input} \tilde{V}_i^2)}{\partial (\text{input} \tilde{V}_i^2)} \cdot \frac{\partial (\text{input} \tilde{V}_i^2)}{\partial \omega_i^{12}} = \\ &= \frac{\eta_2}{2l} \sum_{p=1}^l \sum_{t=1}^N (\hat{y}_{pt}(x_p) - y_p(x_p)) \cdot \\ &= \frac{e^{-\text{input} \tilde{V}_i^2}}{(1 + e^{-\text{input} \tilde{V}_i^2})^2} \cdot \bar{R}_{\tilde{D}_i}(\tilde{U}_i), \\ \Delta \omega_i^{21} &= -\eta_1 \frac{\partial E}{\partial \omega_i^{21}} = \\ &= \frac{\eta_1}{2l} \sum_{p=1}^l \sum_{t=1}^N (\hat{y}_{pt}(x_p) - y_p(x_p)) \cdot \\ &= \frac{e^{-\text{input} \tilde{V}_i^1}}{(1 + e^{-\text{input} \tilde{V}_i^1})^2} \cdot \bar{R}_{\tilde{D}_i}(\tilde{U}_i), \\ \Delta \omega_i^{22} &= -\eta_2 \frac{\partial E}{\partial \omega_i^{22}} = \\ &= \frac{\eta_2}{2l} \sum_{p=1}^l \sum_{t=1}^N (\hat{y}_{pt}(x_p) - y_p(x_p)) \cdot \\ &= \frac{e^{-\text{input} \tilde{V}_i^2}}{(1 + e^{-\text{input} \tilde{V}_i^2})^2} \cdot \bar{R}_{\tilde{D}_i}(\tilde{U}_i). \end{aligned}$$

其中 $\eta_1, \eta_2 \in (0, 1)$ 为选取的两个学习效率参数.

用该误差反向传播梯度学习算法^[7],即可修正网络权值,从而使粗-模糊神经网络收敛.本文规定当误差 $E < 0.01$ 时,即认为已达目标网络,可停止迭代学习修正,此时权值即为网络训练好的最终网络权值.

下面采用上述已经训练好的粗-模糊神经网络来识别预测未知类型样本的类别属性.

假设待识别样本的输入为 $[d_1, d_2, \dots, d_m]$, 经过该训练好的网络,计算得到的其实际输出为 $[y_{p1}, y_{p2}, \dots, y_{pN}]$

若当 $y_{pk} = \max\{y_{p1}, \dots, y_{pN}\} \geq \theta_p$ (θ_p 为预设的门槛一般大于 0.5), 则可认为待识别样本应属决策类中的第 k 类.

若 $y_{pk} = \max\{y_{p1}, \dots, y_{pN}\} < \theta_p$, 则不能判断其所属决策类类别,即网络拒识该样本.

4 模式识别仿真试验 (Model recognition simulation illustration)

用文献[4]中 16 个样本数据作为构造粗-模糊

神经网络的训练样本,按照第 3 节中介绍的粗-模糊神经网络构建步骤,需先对决策系统化简,求得其最小算法规则,由于该决策系统表中条件属性 a_1, a_2, a_3, a_4 和 a_5 是刻划凝聚的 5 个物理性质的检验结果,最后一个属性 class 是决策属性,1 表示抗冻,0 表示不抗冻,由于条件属性值均为连续值,所以必须先离散化处理;文献[5]中是通过用动态层次聚类的连续属性离散化算法,其离散化后的最终决策表(消去重复的行后)如表 1,其核值表如表 2.

表 1 决策表

Table 1 Decision chart

U	A				
	a1	a2	a4	a5	class
1	0	0	0	0	1
2	0	0	1	1	1
3	0	1	0	0	1
4	0	1	0	1	0
5	1	1	0	0	0
6	1	1	1	1	0
7	1	1	1	0	0
8	0	0	1	0	0
9	0	0	0	1	0

表 2 核值表

Table 2 Core value chart

U	A				
	a1	a2	a4	a5	class
1	-	-	0	0	1
2	-	-	1	1	1
3	0	-	-	0	1
4	-	-	-	1	0
5	1	-	-	-	0
6	-	-	-	-	0
7	-	-	-	-	0
8	-	-	1	0	0
9	-	-	0	1	0

由表 2,即可求得该决策系统有 2 个最小决策简化表(见表 3,4).

可选表 3 或表 4 作为一个最小决策算法^[6],不妨取最小决策表(1),则可得其决策系统的最优规则共 5 条,如下:

- If $a_1 = 0 \ \& \ a_4 = 0 \ \& \ a_5 = 0$ Then class = 1;
- If $a_2 = 0 \ \& \ a_4 = 1 \ \& \ a_5 = 1$ Then class = 1;
- If $a_4 = 0 \ \& \ a_5 = 1$ Then class = 0;
- If $a_1 = 1$ Then class = 0;
- If $a_4 = 1 \ \& \ a_5 = 0$ Then class = 0.

表 3 简化决策表(1)

Table 3 Reduction decision chart (1)

U	A				
	a1	a2	a4	a5	class
1	0	x	0	0	1
2	x	0	1	1	1
4	x	x	0	1	0
5	1	x	x	x	0
8	x	x	1	0	0

表 4 简化决策表(2)

Table 4 Reduction decision chart (2)

U	A				
	a1	a2	a4	a5	class
1	0	x	0	0	1
2	0	x	1	1	1
4	x	x	0	1	0
5	1	x	x	x	0
8	x	x	1	0	0

用上面获取的最少规则去构建粗-模糊神经网络(见图 2)^[8,9].该新建网络:第 1 层是输入层,有 4 个属性值输入(4 个节点);第 2 层是 4 个属性值的模糊化节点,共 8 个节点;第 3 层有 5 个节点(即 5 条规则的前件);第 4 层有 2 个节点;第 5~7 层都是 2 个节点(2 个决策类).

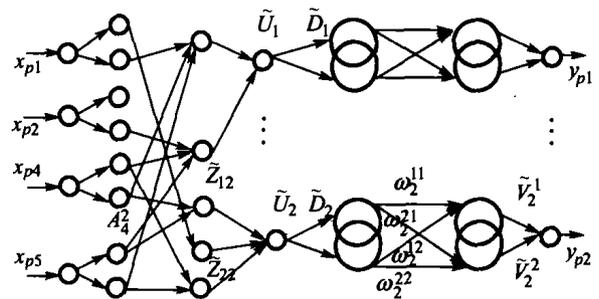


图 2 模式识别的仿真粗-模糊神经网络结构
Fig. 2 Rough-fuzzy neural network structure of model recognition simulation

然后再用误差反向传播梯度算法^[7,10]来训练网络,最后用另外 1000 个样本(其中抗冻的有 753 个,不抗冻的有 247 个)作为测试样本,输入到该网络其实际检验识别结果如表 5 所示.

通过表 5 可看出:

对抗冻性物体来说网络正确识别率为 746/753, 误识率为 4/753,拒识率为 3/753;

对不抗冻性物体网络的正确识别率为 239/247, 误识率为 3/247,拒识率为 5/247;

网络对物体总的正确分类识别率为 $(746 + 239)/1000 = 98.5\%$ 。

表5 模式识别仿真结果

Table 5 Results of model recognition simulation

	识别为 抗冻性物	识别为 不抗冻性物	不能识别的
抗冻性物	746	4	3
不抗冻物	3	239	5

5 结论(Conclusion)

本文介绍了一种利用决策系统表简化进行最优规则提取的新方法,并在此基础上构建了一种新的粗-模糊神经网络,网络利用FRS理论中模糊类对给定范畴的上、下隶属度,将模糊规则的获取和简化融入网络自身中,使网络更加精练,不含冗余规则。因此它具有很强的模式识别和抗干扰功能,其正确分类识别率也非常高,通过上面的仿真试验结果表明本文所采用的方法在模式识别和决策系统控制方面是很成功的。

参考文献(References):

- [1] 曾黄麟. 粗集理论及其应用——关于数据推理的新方法[M]. 重庆:重庆大学出版社,1998.
(ZENG Huangling. *Rough Set Theory and Its Application-New Approach for Data Reasoning* [M]. Chongqing: Chongqing University Press, 1998)
- [2] 于剑,程乾生. 模糊聚类方法中的最佳聚类数的搜索范围[J]. 中国科学(E), 2002, 32(2): 274 - 280.
(YU Jian, CHENG Qiansheng. The search scope of premium clustering number in fuzzy clustering method [J]. *China Science (E)*, 2002, 32(2): 274 - 280.)
- [3] SARKAR M, YEGNARAYANA B. Fuzzy-rough membership functions [C] // *Proc of IEEE Int Conf on Fuzzy Systems*. Anchorage, Alaska, USA: [s. n.], 1998: 2028 - 2033.

- [4] LENARCIK A, PIASTA Z. Discretization of condition attributes space [M] // *Intelligent Decision Support*. Kluwer: Roman Slowinski, 1992: 373 - 389.
- [5] 苗夺谦. Rough Set 理论中连续属性的离散化方法[J]. 自动化学报, 2001, 27(3): 296 - 302.
(MIAO Duoqian. A new method of discretization of continuous attributes in rough sets [J]. *Acta Automatica Sinica*, 27(3): 296 - 302.)
- [6] 胡丹,莫智文. 基于粗集的规则提取 LBR 和 LEM3 [J]. 模式识别与人工智能, 2002, 15(2): 129 - 133.
(HU Dan, MO Zhiwen. Method of extracting rules based on rough set-LBR and LEM3 [J]. *Model Recognition and Artificial Intelligence*, 2002, 15(2): 129 - 133.)
- [7] 王士同. 神经模糊系统及其应用[M]. 北京:北京航空航天大学出版社, 1998, 6: 37 - 43.
(WANG Shitong. *Neural Fuzzy System and Its Application* [M]. Beijing: Beijing University of Aeronautics and Astronautics Press, 1998, 6: 37 - 43.)
- [8] 刘璨,陈统坚. 基于粗集理论的模糊神经网络建模方法研究[J]. 中国机械工程, 2001, 12(11): 1256 - 1259.
(LIU Can, CHEN Tongjian. Research for creating fuzzy neural network model based on rough set theory [J]. *China Mechanics Engineering*, 2001, 12(11): 1256 - 1259.)
- [9] 胡丹,莫智文. 基于规则提取的粗-模糊神经网络及其应用[J]. 模式识别与人工智能, 2001, 14(3): 327 - 330.
(HU Dan, MO Zhiwen. Rough-fuzzy neural network based on rule-extracting and its application [J]. *Pattern Recognition & Artificial Intelligence*, 2001, 14(3): 327 - 330.)
- [10] 漆进,莫智文. 基于粗糙集与神经网络结合的室性 QRS 波分类方法的研究[J]. 模式识别与人工智能, 2002, 15(1): 121 - 123.
(QI Jin, MO Zhiwen. Research for class method of chamber QRS wave based on rough set and neural network [J]. *Pattern Recognition & Artificial Intelligence*, 2002, 15(1), 121 - 123.)

作者简介:

张倩生 (1975—),男,博士,在广东外语外贸大学信息科学技术学院工作,主要感兴趣的研究领域为动力系统与分形和模糊系统控制等, E-mail: zhqiansh01@sina.com.cn.