

文章编号: 1001-0920(2002)05-0614-03

# 基于粗糙集理论的规则修正方法

郝丽娜<sup>1</sup>, 徐心和<sup>2</sup>

(1. 东北大学 机械工程与自动化学院, 辽宁 沈阳 110004; 2. 东北大学 控制与仿真研究中心, 辽宁 沈阳 110004)

**摘要:** 基于粗糙集理论分别给出了确定性决策规则和可能性决策规则的获取与修正的理论和方法, 给出一种利用粗糙集理论解决在增加样本数量情况下的动态规则获取方法。滚动轴承的故障诊断实例证明了该方法的有效性。

**关键词:** 粗糙集; 规则获取; 规则修正; 故障诊断

中图分类号: TP 11 文献标识码: A

## Rule modification based on rough set theory

HAO Li-na<sup>1</sup>, XU Xin-he<sup>2</sup>

(1. College of Mechanical Engineering and Automation, Northeastern University, Shenyang 110004, China;  
2. Research Center of Control and Simulation, Northeastern University, Shenyang 110004, China)

**Abstract:** The theory and method of acquiring and modifying determinant rules and possible rules based on rough set are presented to give a method of dynamical rule acquisition under the condition of increasing sample numbers. An example of bearing-ball fault diagnosis shows the validity of this method.

**Key words:** rough set theory; rule acquisition; rule modification; fault diagnosis

## 1 引言

粗糙集(RS)理论的主要思想是在保持分类能力不变的前提下,通过知识约简导出概念的分类规则。RS方法是用上、下近似构成粗糙集,构成不确定区间。在近似的基础上可产生两类对应的规则:确定性规则和可能性规则<sup>[1]</sup>。文献[2,3]给出了几种规则获取和修正的方法。本文在上述研究的基础上,从粗糙集理论的基本概念出发,进一步给出适用于确定性规则和可能性规则的规则修正方法。

## 2 粗糙集理论的有关概念

知识表达系统<sup>[2]</sup>:一般表示为

$$S = (U, A, V, f) \quad (1)$$

其中, $S$ 为知识表达系统; $U$ 为论域; $A$ 为属性集, $A = C \cup D, C \cap D = \emptyset, C$ 为条件属性集合, $D$ 为决策属性集合; $V$ 为属性值域, $V = \bigcup_{a \in A} V_a, a \in A, V_a$ 是属性 $D$ 的值域; $f: U \times A \rightarrow V$ 为一信息函数,它指定 $U$ 中每一对象 $x$ 在属性 $a$ 下的值。

确定性规则和可能性规则<sup>[2]</sup>:信息系统

$S = (U, C, D, V, f), \forall x \in U \quad (2)$   
 $x$ 唯一确定了一个 $C$ 基本公式和 $D$ 基本公式,分别为

收稿日期: 2001-07-05; 修回日期: 2001-08-13

基金项目: 西安交通大学机械制造系统工程国家重点实验室开放基金项目

作者简介: 郝丽娜(1968—),女,辽宁庄河人,副教授,博士,从事智能混合系统、CIMS工程应用等研究;徐心和(1940—),

© 1994-2011, China Academic Electronic Journal Publishing House. All rights reserved. http://www.cnki.net

$$\text{desc}_c(x) = \underset{c}{c} (c, f(x, c)) \quad (3)$$

$$\text{desc}_d(x) = \underset{d}{d} (d, f(x, d)) \quad (4)$$

决策规则为

$$r_x^{C,D}: \text{desc}_c(x) \quad \text{desc}_d(x) \quad (5)$$

其中,  $\text{desc}_c(x)$  和  $\text{desc}_d(x)$  分别称为规则  $r_x^{C,D}$  的条件部分和决策部分。

$$|\text{desc}_d(x)|_s \quad U | \text{IND}(D) \quad (6)$$

表示某个决策类, 若  $x \in \underline{C}(|\text{desc}_d(x)|_s)$ , 则  $r_x^{C,D}$  是确定性决策规则; 若  $x \in \overline{C}(|\text{desc}_d(x)|_s)$ , 则  $r_x^{C,D}$  是可能性决策规则。

每条规则的置信度可用粗糙隶属函数的值表示, 它表示匹配规则的对象数与匹配规则条件部分的对象数之比。即规则  $r_x^{C,D}$  的置信度为

$$\alpha(r_x^{C,D}) = \frac{|\text{desc}_c(x)|_s \quad |\text{desc}_d(x)|_s}{|\text{desc}_c(x)|_s} \quad (7)$$

### 3 规则获取与规则修正方法

#### 3.1 规则获取

决策矩阵<sup>[2]</sup>的概念来于分辨矩阵, 其定义为: 假定属于某一类和不属于某一类的对象分别用下标  $i(i = 1, 2, \dots, n)$  和  $j(j = 1, 2, \dots, m)$  表示, 则系统  $S$  的决策矩阵  $M(S)$  为

$$M(S) = (M_{ij})_{n \times m} \quad (8)$$

其中,  $M_{ij} = \{(a, a(i)) | a(i) \quad a(j)\}$ ,  $a(i) \in V_a$ ,  $a(j) \in V_a$ 。

决策函数<sup>[2]</sup>定义为

$$B_i = \bigwedge_j M_{ij} \quad (9)$$

其中,  $\bigwedge$  表示析取运算,  $\bigvee$  表示合取运算,  $B_i$  是一个布尔函数, 它是利用决策矩阵  $M$  的第  $i$  行构造的, 并把  $M_{ij}$  中每个属性看作一个布尔变量。对决策函数  $B_i$  运用布尔代数中的分配律和吸收律化简成析取范式, 则每个合取项对应一个决策规则, 于是所有的决策规则为

$$\text{RULE} = |B_i|, \quad i = 1, 2, \dots, n \quad (10)$$

利用决策矩阵和决策函数进行规则获取的方法<sup>[3]</sup>, 可求出概括性最强(含有最少条件属性)的分类规则, 即长度最小的规则。求确定性决策规则和可能性决策规则的方法是: 先求出每一类别的确定性决策矩阵和可能性决策矩阵, 再对每一矩阵求出决策规则, 它们就是确定性决策矩阵和可能性决策矩阵。规则获取流程如图 1 所示。

#### 3.2 规则修正

我们的目的是利用粗糙集理论给出一种在增加样本数量情况下的动态规则获取方法。当增加新的数据(如学习样本)时, 只需修正现有的决策矩阵和决策函数, 实现对原有的规则进行修正, 便可得到新数据集的规则, 而不需重新计算决策矩阵和决策函数。本文方法首先分别获取确定性决策规则和可能性决策规则, 再在文献[3]基于决策矩阵和决策函数进行规则修正的增量式算法的基础上, 分别进行确定性决策规则和可能性决策规则的修正。其修正算法如下:

Step1: 根据决策属性值确定新样本属于哪个目标概念(类别); 如果不属于任何一类, 则定义一个新概念(类别), 进行规则求取。设  $M_{ij}$  为当前决策矩阵, 新类为  $k$ 。

Step2: 如果新类  $k$  包含在当前概念的正域中, 则:

- 1) 生成决策矩阵的新行  $M_{kj}$ ;
- 2) 计算新类的决策函数  $B_k = \bigvee_j M_{kj}$ ;
- 3) 修正规则  $\text{RULE} = \text{RULE} \quad |B_k|$ 。

Step3: 如果新类  $k$  不包含在当前概念的正域中, 则:

- 1) 生成决策矩阵的新列  $M_{jk}$ ;
- 2) 计算新类的决策函数  $B_i = B_i \quad M_{jk}$ ;
- 3) 修正规则  $\text{RULE} = |B_i|$ 。

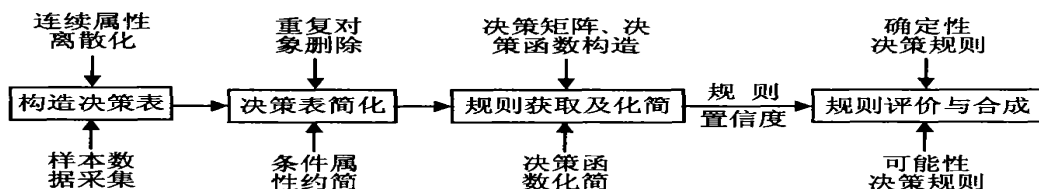


图 1 RS 理论自动获取规则流程

表 1 D = 0 类的确定性决策矩阵(a = s4, b = s9, c = s10)

	5	11	27	28	33	37	38	39	40	41	12
(1)	(a,2)(b,1)	(a,2)(b,1)	(b,1)	(a,2)(b,1)	(a,2)(c,1)	(a,2)(b,1)	(a,2)(b,1)	(a,2)(b,1)	(c,1)	(a,2)(b,1)	(a,2)(b,1)
1	(c,1)			(c,1)		(c,1)	(c,1)	(c,1)			
(2)	(a,1)(b,1)	(a,1)(b,1)	(a,1)(b,1)	(a,1)(b,1)	(a,1)(c,1)	(a,1)(b,1)	(a,1)(b,1)	(a,1)(b,1)	(a,1)(c,1)	(a,1)(b,1)	(a,1)(b,1)
2	(c,1)			(c,1)		(c,1)	(c,1)	(c,1)			
(3)	(b,4)	(a,5)(b,4)	(a,5)(b,4)	(a,5)(b,4)	(b,4)	(a,5)(b,4)	(a,5)(b,4)	(a,5)(b,4)	(a,5)(b,4)	(a,5)(b,4)	(b,4)(c,2)
3		(c,2)	(c,2)							(c,2)	
(4)	(b,1)(c,1)	(a,5)(b,1)	(a,5)(b,1)	(a,5)(b,1)	(c,1)	(a,5)(b,1)	(a,5)(b,1)	(a,5)(b,1)	(a,5)(c,1)	(a,5)(b,1)	(b,1)
7				(c,1)		(c,1)	(c,1)	(c,1)			
(5)	(a,1)(b,2)	(a,1)	(a,1)	(a,1)(c,1)	(a,1)(b,2)	(a,1)(b,2)	(a,1)(b,2)	(a,1)(c,1)	(a,1)(b,2)	(a,1)	(a,1)(b,2)
8	(c,1)				(c,1)	(c,1)	(c,1)	(a,1)(c,1)	(c,1)		
(6)	(b,4)	(a,5)(b,4)	(a,5)(b,4)	(a,5)(b,4)	(b,4)(c,3)	(a,5)(b,4)	(a,5)(b,4)	(a,5)(b,4)	(a,5)(b,4)	(a,5)(b,4)	(b,4)(c,3)
17	(c,3)	(c,3)	(c,3)	(c,3)		(c,3)	(c,3)	(c,3)	(c,3)	(c,3)	
(7)	(a,3)(b,5)	(a,3)(b,5)	(a,3)(b,5)	(a,3)(b,5)	(a,3)(b,5)	(b,5)	(a,3)(b,5)	(b,5)	(a,3)(b,5)	(b,5)(c,2)	(a,3)(b,5)
36		(c,2)	(c,2)							(c,2)	
(8)	(b,5)(c,4)	(a,5)(b,5)	(a,5)(b,5)	(a,5)(b,5)	(b,5)(c,4)	(a,5)(b,5)	(a,5)(b,5)	(a,5)(b,5)	(a,5)(b,5)	(a,5)(b,5)	(b,5)(c,4)
44		(c,4)	(c,4)	(c,4)		(c,4)	(c,4)	(c,4)	(c,4)	(c,4)	
(9)	(b,2)(c,3)	(a,5)(c,3)	(a,5)(c,3)	(a,5)(c,3)	(b,2)(c,3)	(a,5)(b,2)	(a,5)(b,2)	(a,5)(c,3)	(a,5)(b,2)	(a,5)(c,3)	(b,2)(c,3)
47						(c,3)	(c,3)		(c,3)		
(10)	(a,4)(c,1)	(b,3)	(a,4)(b,3)	(b,3)(c,1)	(a,4)(b,3)	(a,4)(c,1)	(c,1)	(a,4)(b,3)	(a,4)(b,3)	(a,4)(b,3)	(a,4)
49					(c,1)			(c,1)	(c,1)		

If s4 = 5 且 s10 = 1 Then D = 0

同理,可求出 D = 1 类的确定性决策规则和可能性决策规则,在此不再赘述。

### 4 应用实例

#### 4.1 规则获取

滚动轴承故障诊断数据集见文献[4]的表1。采用图1的步骤,得到 D = 0 类确定性决策规则如下

If s9 = 4 Then D = 0

If s4 = 1 Then D = 0

If s10 = 3 Then D = 0

If s9 = 5 Then D = 0

If s10 = 4 Then D = 0

If s4 = 4 且 s9 = 3 且 s10 = 1 Then D = 0

If s9 = 1 且 s10 = 1 Then D = 0

可以证明,上述规则的置信度为1,为确定性决策规则,且包含了所有的对象。而文献[1]中 D = 0 类的规则2是错误的,且对象3,4,35,45和52无法判定。D = 0 类的可能性决策规则如下

If s9 = 1 且 s10 = 1 Then D = 0

If s9 = 4 Then D = 0

If s4 = 1 Then D = 0

If s10 = 3 Then D = 0

If s9 = 5 Then D = 0

If s10 = 4 Then D = 0

If s9 = 3 且 s10 = 1 Then D = 0

#### 4.2 规则修正

规则修正过程如下。以样本41为例,该样本的属性值如表2所示。

表2 样本41的属性值

No.	s4	s9	s10	D
41	3	2	1	1

D = 0 类的确定性决策规则:对表1去掉样本41有关的元素后,得到 D = 0 类的确定性决策矩阵生成一新列,实质上就是表1的第10列。再计算各行构成的决策函数,其结果必然与4.1节的结果相同。

### 5 结论

本文给出了基于决策矩阵和决策函数的规则获取和修正的理论和方法。由于该规则修正方法是在原有的决策规则基础上进行,不需重新计算整个知识表达系统的决策矩阵和决策函数,因此能有效地解决动态环境下的规则修正问题。该方法同样适用于专家系统知识库的实时获取与修改。

(下转第620页)

control[J]. *Information and Control*, 1995, 24(4): 59-62.)

- [2] 贾磊, 朱摩西. 用论域缩小逼近法消除模糊控制器的余差[J]. 信息与控制, 1995, 24(4): 251-256.

(Jia Lei, Zhu Mo-xi. Eliminating steady-state error of fuzzy controller with the successive approximation of universe of discourse [J]. *Information and Control*, 1995, 24(4): 251-256.)

- [3] 路兆梅, 于跃海. 多项式重置增量型模糊控制器[J]. 东南大学学报, 1999, 29(1): 49-53.

(Lu Zhao-mei, Yu Yue-hai. Polynomial reset incremental fuzzy controller[J]. *J of Southeast University*, 1999, 29(1): 49-53.)

- [4] 范晓英, 陆培新, 陈文楷. 一种新型的模糊控制器[J]. 控制理论与应用, 1995, 12(5): 597-601.

(Fan Xiao-ying, Lu Pei-xin, Chen Wen-kai. A new type of fuzzy controller[J]. *Control Theory and Application*, 1995, 12(5): 597-601.)

- [5] 李华. 一种自适应模糊控制器的设计方法[J]. 电气自动化, 1998, 20(1): 35-37.

(Li Hua. A design method of self-adaptive fuzzy controller[J]. *Electrical Drive Automation*, 1998, 20(1): 35-37.)

- [6] 李娟, 陈佳娟. 模糊控制静差产生原因的理论分析[J]. 甘肃科技, 2000, 16(3): 28.

(Li Juan, Chen Jia-juan. The theory analysis of causing the static error for fuzzy control[J]. *Gansu Science and Technology*, 2000, 16(3): 28.)

- [7] 张乃尧. 典型模糊控制器的结构分析[J]. 模糊系统与数学, 1997, 21(2): 10-21.

(Zhang Nai-yao. Structure analysis of typical fuzzy controllers[J]. *Fuzzy System and Mathematic*, 1997, 21(2): 10-21.)

- [8] Change C C. Improved DNC design for nonlinear process control[J]. *J of AIChE*, 1992, 38(4): 602-608.

(上接第 616 页)

#### 参考文献(References):

- [1] 韩祯祥, 张琦, 文福拴. 粗糙集理论及其应用综述[J]. 控制理论与应用, 1999, 16(2): 153-157.

(Han Zhenxiang, Zhang Qi, Wen Fushuan. A survey on rough set theory and its application[J]. *Control and Applications*, 1999, 16(2): 153-157.)

- [2] 曾黄麟. 粗糙集理论及其应用[M]. 重庆: 重庆大学出版社, 1998. 71, 94-117.

- [3] Ning Shan, Wojciech Ziarko. Data-based acquisition and incremental modification of classification rules[J]. *Computational Intelligence*, 1995, 11(2): 357-370.

- [4] Ryszard Nowicki. Evaluation of vibroacoustic diagnostic symptoms by means of the rough sets theory[J]. *Computers in Industry*, 1992, 20(2): 141-152.

## 2003 年 IFAC 冶金自动化新技术国际会议征文通知

IFAC 冶金自动化新技术国际会议将于 2003 年 10 月 11 日~13 日在上海浦东召开。会议由 IFAC 采矿、矿物和金属加工自动化技术委员会以及 IFAC 低成本自动化技术委员会、IFAC 发展中国家技术委员会联合发起。受中国自动化学会委托, 经 IFAC 批准, 会议由大连理工大学和上海宝信软件股份有限公司联合承办。

征文范围: 与冶金工业生产过程有关的检测与自动化装置、建模与控制、故障诊断、计划与调度、管理与优化等, 以及与上述内容有关的方法与技术的实际应用。

征文要求: 论文作者应向程序委员会秘书处提交英文论文全文, 且 Email 投稿。论文应按 IFAC 论文版式(双列)排版, 不超过 6 页, 详见 IFAC 网站 [www.ifac-control.org](http://www.ifac-control.org), 也可直接访问 [www.elsevier.com/locate/ifac](http://www.elsevier.com/locate/ifac)。论文经国际程序委员会审稿录用后, 要求至少有一位作者注册并参加会议。

论文出版: 按 IFAC 惯例, 会议向与会代表提供会议论文预印集(Preprints); 参加会议并在会上宣读的论文将

收入 IFAC 正式论文集(Proceedings); 论文集由 Elsevier Science Ltd, Oxford UK 出版。优秀论文可推荐到 *Control Engineering Practice* 等国际期刊。

论文提交日期: 2003 年 2 月 28 日

录用通知: 2003 年 5 月 31 日

终稿(按 IFAC 版式)提交日期: 2003 年 7 月 30 日

国际程序委员会秘书处地址: 116024 大连理工大学信息与控制研究中心

联系人: 潘学军 副教授

电话: 0411-4707576 传真: 0411-4707579

Email: [panxj@dlut.edu.cn](mailto:panxj@dlut.edu.cn)

组织委员会秘书处地址: 201203 上海宝信软件有限责任公司

联系人: 丛力群 博士

电话: 021-50803342 传真: 021-50800701

Email: [congliqun@baosight.com](mailto:congliqun@baosight.com)

英文征文通知见 <http://www.baosight.com/ifac03>。