

## 粗糙集信度一致属性约简

高 灿<sup>1,2\*</sup>, 苗夺谦<sup>1,2</sup>, 张志飞<sup>1,2</sup>, 张红云<sup>1,2</sup>

(1. 同济大学 计算机科学与技术系, 上海 201804; 2. 同济大学 嵌入式系统与服务计算教育部重点实验室, 上海 201804)

(\* 通信作者电子邮箱 2005gaocan@163.com)

**摘 要:**为了消除现有概率粗糙集模型约简过程中出现的诸多约简异常问题,通过引入对象最大信度概念,提出了非参与带参最大决策熵属性约简模型,阐明了带参最大决策熵测度的单调性,给出了带参最大决策熵核和相对不必要属性的定义,并分析了其约简与已有概率粗糙集模型约简的关系。其次将对象置信度引入差别矩阵,构建了带参与非参与信度差别矩阵,讨论了其定义与经典差别矩阵对不确定对象刻画的差异性。最后运用实验验证了方法的有效性。

**关键词:**概率粗糙集;属性约简;约简异常;最大决策熵;信度差别矩阵

**中图分类号:** TP18 **文献标志码:** A

### Rough set based attribute reduction with consistent confidence

GAO Can<sup>1,2\*</sup>, MIAO Duo-qian<sup>1,2</sup>, ZHANG Zhi-fei<sup>1,2</sup>, ZHANG Hong-yun<sup>1,2</sup>

(1. Department of Computer Science and Technology, Tongji University, Shanghai 201804, China;

2. Key Laboratory of Embedded System and Service Computing, Ministry of Education, Tongji University, Shanghai 201804, China)

**Abstract:** In order to solve the problem of reduction anomaly in the existing probabilistic rough set models, non-parameterized and parameterized maximum decision entropy measures for attribute reduction were proposed by using the concept of maximum confidence of uncertain object. The monotonicity of the parameterized maximum decision entropy was explained and the relationship between its attribute reduction and other ones was analyzed. The definitions for core and relatively dispensable attributes in the proposed model were also given. Moreover, non-parameterized and parameterized confidence discernibility matrixes were put forward and the difference of classical discernibility matrix and the proposed ones in characterizing the uncertain object were discussed. Finally, a case study was given to show the validity of the proposed model.

**Key words:** probabilistic rough set; attribute reduction; reduction anomaly; maximum decision entropy; confidence discernibility matrix

## 0 引言

粗糙集理论作为处理不精确、不一致和不完备信息的有效方法<sup>[1]</sup>,在机器学习、数据挖掘和人工智能等领域中得到广泛的应用<sup>[2]</sup>。粗糙集理论通过不可分辨关系对对象空间进行划分,运用上下近似算子来刻画不确定知识。但其本身对数据的要求过于严格,缺乏柔性或鲁棒性。针对经典粗糙集的不足,研究学者通过引入概率方法提出了变精度粗糙集、决策论粗糙集、贝叶斯粗糙集等概率模型<sup>[3]</sup>。

文献[4]在经典粗糙集模型中引入了误分率的概念,将集合的完全包含关系扩展至部分包含,提出了能有效处理噪声数据的变精度粗糙集模型,但其约简仍采用经典粗糙集的定义,约简过程中存在对象置信度不一致、区间动态性、约简跳跃和分类异常等问题。文献[5]分析了文献[4]约简定义的缺陷,通过限制约简前后的分类精度区间解决了约简分类异常问题。文献[6]进一步指出约简过程中应保持区间交集非空或相等,消除了约简过程中的区间动态异常。文献[7]给出了最大分布约简的概念,但其约简仅保持对象的最大分布决策不变,各类约简异常都可能发生。文献[8]提出了变精度 $\beta$ 上、下近似分布约简, $\beta$ 下近似分布约简仅能消除约简

过程中的分类异常,而 $\beta$ 上近似分布约简比 $\beta$ 下近似条件更弱,约简前后正域都会发生变化。文献[9]提出了快速求解最大分布约简的新方法,但亦未消除各类约简异常。以上概率粗糙集模型存在约简异常的根本原因是约简度量的非单调性。本文基于对象的分类及信度信息,提出了带参最大决策熵单调约简模型及信度差别矩阵,有效地消除了约简过程中各类约简异常,达到了属性约简的目的。

## 1 基本概念

**定义 1<sup>[1]</sup>** 信息系统可表示为  $S = (U, A, V, f)$ , 其中:  $U$  是对象集合;  $A$  是属性非空集合;  $V = \cup V_a, V_a$  表示属性  $a$  的值域;  $f: U \times A \rightarrow V$  是一个映射, 指定  $U$  中每一对象  $x$  的属性值, 即对  $x \in U, a \in A$  有  $a(x) \in V_a$ 。

如果属性集  $A$  可分为条件属性集  $C$  和决策属性集  $D$ , 即  $A = C \cup D, C \cap D = \emptyset$ , 则该信息系统称为决策信息系统或决策表。

**定义 2<sup>[1]</sup>** 给定信息系统  $S = (U, A, V, f)$ , 对于任意属性子集  $B \subseteq A$ , 可定义不可分辨关系  $IND(B) = \{(x, y) \in U \times U \mid \forall a \in B, a(x) = a(y)\}$ 。

$IND(B)$  是一个等价关系, 构成  $U$  的一个划分, 用

收稿日期: 2011-09-08; 修回日期: 2011-11-28。

基金项目: 国家自然科学基金资助项目(60970061, 61075056, 61103067); 上海市重点学科建设项目(B004)。

作者简介: 高灿(1983-), 男, 湖南南县人, 博士研究生, CCF 会员, 主要研究方向: 粗糙集理论、机器学习; 苗夺谦(1964-), 男, 山西祁县人, 教授, 博士生导师, CCF 高级会员, 主要研究方向: 粗糙集理论、Web 智能、机器学习; 张志飞(1986-), 男, 江苏南通人, 博士研究生, CCF 会员, 主要研究方向: 粒计算、文本挖掘; 张红云(1972-), 女, 江苏丹阳人, 讲师, 博士, CCF 会员, 主要研究方向: 主曲线、粒计算。

$U/IND(B)$  表示, 简记为  $U/B$ 。

**定义 3<sup>[1]</sup>** 给定信息系统  $S = (U, A, V, f)$ , 对于任意属性子集  $B \subseteq A$ , 不可分辨关系  $IND(B)$  的等价类可表示为  $[x]_B = \{y \in U \mid \langle x, y \rangle \in IND(B)\}$ 。

**定义 4<sup>[4]</sup>** 给定决策表  $S = (U, A = C \cup D, V, f)$  及分类精度  $\beta \in (0.5, 1]$ , 设  $X \subseteq U$ , 对任一属性子集  $B \subseteq C$ ,  $X$  关于  $B$  的  $\beta$  上、下近似集分别表示如下:

$$\underline{B}_\beta(X) = \cup \{[x]_B \mid P(X \mid [x]_B) \geq \beta\} \quad (1)$$

$$\overline{B}_\beta(X) = \cup \{[x]_B \mid P(X \mid [x]_B) \geq 1 - \beta\} \quad (2)$$

**定义 5<sup>[4]</sup>** 给定决策表  $S = (U, A = C \cup D, V, f)$  及分类精度  $\beta \in (0.5, 1]$ , 则  $D$  关于条件属性  $C$  的  $\beta$  正域、边界域和负域可分别表示为:

$$POS(C, D, \beta) = \cup \{\underline{C}_\beta(D_j) \mid D_j \in U/D\} \quad (3)$$

$$BND(C, D, \beta) = \cup \{\overline{C}_\beta(D_j) - \underline{C}_\beta(D_j) \mid D_j \in U/D\} \quad (4)$$

$$NEG(C, D, \beta) = U - POS(C, D, \beta) - BND(C, D, \beta) \quad (5)$$

## 2 信度一致属性约简

### 2.1 最大决策熵属性约简

为了阐述方便, 先给出条件类最大包含度、最大包含度决策及分布等相关概念。

**定义 6** 给定决策表  $S = (U, A = C \cup D, V, f)$ , 条件类  $C_i \in U/C$  相对于决策的最大包含度表示为  $MP(C_i) = \max\{|C_i \cap D_1|/|C_i|, |C_i \cap D_2|/|C_i|, \dots, |C_i \cap D_{|U/D|}|/|C_i|\}$ , 则条件属性集  $C$  下条件类最大包含度分布为  $MS(C) = (MP(C_1), MP(C_2), \dots, MP(C_{|U/C|}))$ 。

**定义 7** 给定决策表  $S = (U, A = C \cup D, V, f)$ , 条件属性集  $C$  下任意对象  $x \in U$  的决策值表示为  $VD(x) = \cup\{D(z) \mid z \in [x]_C\}$ , 则最大包含度决策值为  $MD(x) = \{D(z) \mid z \in D_j \wedge P(D_j \mid [x]_C) = MP([x]_C)\}$ 。

**定义 8** 给定决策表  $S = (U, A = C \cup D, V, f)$ , 属性子集  $B (B \subseteq C)$  的最大包含度分布为  $MS(B) = (MP(B_1), MP(B_2), \dots, MP(B_{|U/B|}))$ , 则条件属性集  $B$  的非参最大决策熵表示为  $ME(B) = - \sum_{i=1}^{|U/B|} \frac{|B_i|}{|U|} MP(B_i) \lg MP(B_i)$ 。

**定义 9** 给定决策表  $S = (U, A = C \cup D, V, f)$  及分类精度  $\beta \in (0.5, 1]$ , 属性子集  $B (B \subseteq C)$  的最大包含度分布为  $MS(B) = (MP(B_1), MP(B_2), \dots, MP(B_{|U/B|}))$ , 则条件属性集  $B$  的带参最大决策熵表示为:

$$ME_\beta(B) = - \sum_{B_i \in U/B \wedge MP(B_i) > \beta} \frac{|B_i|}{|U|} MP(B_i) \lg MP(B_i)$$

根据贝叶斯原则, 规则的误分率主要取决于较小决策样本, 因此属性约简过程中各等价类的最大决策信息应保持不变。非参最大决策熵考虑了各等价类的最大决策类信息, 因此约简后能保证各对象的最大决策类信息不发生变化。而带参最大决策熵引入了变精度分类参数, 仅反映最大包含度大于分类精度阈值的等价类信息, 也即保持约简前后变精度正域不发生变化, 达到了约简的目的。最大决策熵函数可表示为图 1。

**命题 1** 给定任意决策表, 带参最大决策熵是单调的。

**命题 2** 给定决策表, 非参最大决策熵不一定满足单调性质。

从图 1 可看出, 函数在区间  $(0.5, 1]$  上严格单调下降, 因此带参最大决策熵测度是单调的。而在整个论域上, 最大决

策熵测度为凸函数, 因此非参最大决策熵不具单调性。

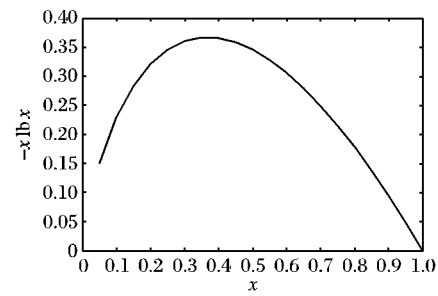


图 1 函数  $y = -x \lg x$

根据以上概念可知, 非参与带参最大决策熵的区别主要在于是否加入最大包含度小于参数  $\beta$  的不确定条件等价类信息。而这种差异导致完全不同的结论, 以至只能采用带参最大决策熵作为约简度量标准。在以上结论基础上, 可定义最大决策熵核及约简算法。

**定义 10** 给定决策表  $S = (U, A = C \cup D, V, f)$ , 分类精度  $\beta \in (0.5, 1]$ , 且  $B \subseteq C$ , 任意属性  $a \in C - B$  的重要性  $Sig(a, B, D)$  定义为  $Sig(a, B, D) = ME_\beta(B \cup \{a\}) - ME_\beta(B)$ 。

**定义 11** 给定决策表  $S = (U, A = C \cup D, V, f)$  及分类精度  $\beta \in (0.5, 1]$ , 则任意属性  $a \in C$  为决策表核属性的充分必要条件为  $Sig(a, C, D) > 0$  或  $ME_\beta(C) \neq ME_\beta(C - \{a\})$ 。

**定义 12** 给定决策表  $S = (U, A = C \cup D, V, f)$  及分类精度  $\beta \in (0.5, 1]$ , 则任意属性  $a \in C$  为决策表的相对冗余属性的充分必要条件为  $Sig(a, C, D) = 0$  或  $ME_\beta(C) = ME_\beta(C - \{a\})$ 。

**定义 13** 给定决策表  $S = (U, A = C \cup D, V, f)$  及分类精度  $\beta \in (0.5, 1]$ , 任意属性子集  $P \subseteq C$  为决策表  $S$  的约简当且仅当以下条件成立:

- 1)  $ME_\beta(P) = ME_\beta(C)$
- 2)  $\forall a \in P, ME_\beta(P - \{a\}) \neq ME_\beta(P)$

基于上述概念, 可构造带参最大决策熵约简算法, 具体描述如下:

**算法 1** 带参最大决策熵约简 (Parameterized Maximum Decision Entropy-based Attribute Reduction, PMDEAR)

输入: 决策表  $S = (U, A = C \cup D, V, f)$  及分类精度  $\beta$

输出: 约简  $P$

步骤 1 求条件等价类及最大分布包含;

步骤 2 计算核属性  $Core, P = Core$ ;

步骤 3 如果  $ME_\beta(P) \neq ME_\beta(C)$ 。

计算条件属性集合  $C - P$  中各属性的重要性;

选择属性重要性最大的属性  $a, P = P \cup \{a\}$ ;

步骤 4 返回约简  $P$ , 结束。

算法 1 先在条件属性  $C$  和决策属性  $D$  下进行等价类划分, 计算等价类的最大包含度分布。在核属性基础上, 依次加入属性重要性最大的属性, 直至当前属性集合的最大决策熵与原决策表的最大决策熵相等。假设决策表对象  $|U| = n$ , 且条件属性  $|C| = m$ , 等价类划分的时间复杂度为  $O(mn)$ , 而算法 1 循环  $|C| = m$  次必定结束, 所以总体时间复杂度为  $O(m^2n)$ , 空间复杂度为  $O(n)$ 。

### 2.2 信度矩阵属性约简

**定义 14** 非参信度差别矩阵。给定决策表  $S = (U, A =$

$C \cup D, V, f)$ , 条件属性集  $C$  下非参信度差别矩阵的元素  $m(i, j)$  定义为:

$$m(i, j) = \begin{cases} \emptyset, MP([x_i]_C) = MP([x_j]_C) \wedge MD([x_i]_C) = MD([x_j]_C) \\ \{a_k | a_k \in C \wedge a_k(x_i) \neq a_k(x_j)\}, \text{其他} \end{cases}$$

定义 15 带参信度差别矩阵。给定决策表  $S = (U, A = C \cup D, V, f)$  及分类精度参数  $\beta$ , 条件属性集  $C$  下带参信度差别矩阵的元素  $m_\beta(i, j)$  定义为:

$$m_\beta(i, j) = \begin{cases} \emptyset, \max\{MP([x_i]_C), MP([x_j]_C)\} < \beta \\ \emptyset, MP([x_i]_C) = MP([x_j]_C) \wedge MD([x_i]_C) = MD([x_j]_C) \wedge \min\{MP([x_i]_C), MP([x_j]_C)\} \geq \beta \\ \{a_k | a_k \in C \wedge a_k(x_i) \neq a_k(x_j)\}, \text{其他} \end{cases}$$

原有差别矩阵仅涉及对象的决策信息,即决策不同的对象需生成区分信息以保证决策表的分类能力。由于概率粗糙集模型下决策表的分类能力体现在对象决策及置信度两方面,因此应在差别矩阵中引入对象置信度信息。非参信度差别矩阵无需区分最大包含度决策相同且置信度相等的对象,而带参信度差别矩阵进一步忽略了最大包含度都小于分类精度阈值的对象分类信息(即边界域和负域),仅保持相对正域对象不发生变化。相对于原有差别矩阵,新的信度差别矩阵定义更严格,反映的分类信息亦更为丰富。

### 2.3 与其他方法比较

文献[7]分析了分布约简、最大分布约简、分配约简和下近似约简之间的关系,指出分布约简必为分配约简或最大分布约简,而最大分布约简与分配约简无强弱关系。同时,分配约简与近似约简具有等价性。本文引入的最大决策熵约简与以上各约简都不同,约简条件弱于分布约简,但强于最大分布约简。各约简间的关系如下:

分布约简(条件熵)  $\Rightarrow$  最大决策熵约简  $\Rightarrow$  最大分布约简  
分布约简(条件熵)  $\Rightarrow$  分配约简  $\Leftrightarrow$  近似约简

表 2 约简模型及其约简信息(分类质量  $\gamma = 1$ )

条件属性集	条件类及最大包含度	约简区间	备注
$C$	$1.0/C_1 = \{o_1\}, 1.0/C_2 = \{o_2\}, 0.75/C_3 = \{o_3, o_4, o_5, o_6\},$ $0.67/C_4 = \{o_7, o_8, o_9\}, 0.67/C_5 = \{o_{10}, o_{11}, o_{12}\}$	$\beta((0.5, 0.67))$	分布约简
$\{a_1, a_3, a_4\}$	$1.0/C_1 = \{o_1\}, 1.0/C_2 = \{o_2\}, 0.75/C_3 = \{o_3, o_4, o_5, o_6\},$ $0.67/C_4 = \{o_7, o_8, o_9, o_{10}, o_{11}, o_{12}\}$	$\beta((0.5, 0.67))$	最大决策熵约简
$\{a_1\}$	$0.67/C_1 = \{o_1, o_2, o_3, o_4, o_5, o_6\},$ $0.67/C_2 = \{o_7, o_8, o_9, o_{10}, o_{11}, o_{12}\}$	$\beta((0.5, 0.67))$	文献[4]约简 文献[5]约简
$\{a_4\}$	$1.0/C_1 = \{o_1\},$ $0.73/C_2 = \{o_2, o_3, o_4, o_5, o_6, o_7, o_8, o_9, o_{10}, o_{11}, o_{12}\}$	$\beta((0.5, 0.73))$	下近似约简 最大分布约简

表 3 示例决策表  $S$  的信度差别矩阵

行	列				
	$C_1$	$C_2$	$C_3$	$C_4$	$C_5$
$C_1$	$\emptyset$				
$C_2$	$\{a_4\}$	$\emptyset$			
$C_3$	$\{a_3, a_4\}$	$\{a_3\}$	$\emptyset$		
$C_4$	$\{a_1, a_3, a_4\}$	$\{a_1, a_3\}$	$\{a_1\}$	$\emptyset$	
$C_5$	$\{a_1, a_2, a_3, a_4\}$	$\{a_1, a_2, a_3\}$	$\{a_1, a_2\}$	$\emptyset$	$\emptyset$

运用吸收律后,信度差别矩阵的分辨函数为  $f_C(D) = a_4 \wedge a_3 \wedge a_1$ , 因此条件属性集  $\{a_1, a_3, a_4\}$  为示例决策表的

## 3 实例分析

给定决策表表 1, 已有概率粗糙集模型<sup>[4-9]</sup>将出现对象置信度不一致、区间动态性、分类异常和约简跳跃等诸多问题,本文的方法能消除以上异常,获得较优质量的约简。

表 1 示例决策表  $S$

$U$	$a_1$	$a_2$	$a_3$	$a_4$	$d$
$o_1$	0	0	0	0	1
$o_2$	0	0	0	1	2
$o_3$	0	0	1	1	2
$o_4$	0	0	1	1	2
$o_5$	0	0	1	1	2
$o_6$	0	0	1	1	3
$o_7$	1	0	1	1	1
$o_8$	1	0	1	1	2
$o_9$	1	0	1	1	2
$o_{10}$	1	1	1	1	2
$o_{11}$	1	1	1	1	2
$o_{12}$	1	1	1	1	3

在条件属性集  $C$  下,表 1 对象形成划分  $U/C = \{\{o_1\}, \{o_2\}, \{o_3, o_4, o_5, o_6\}, \{o_7, o_8, o_9\}, \{o_{10}, o_{11}, o_{12}\}\}$ 。给定分类质量  $\gamma = 1$ , 各约简模型的约简信息如表 2。

从表 2 可见,最大分布约简及下近似约简  $\{a_4\}$  能保证决策表的正域不发生变化,但对象  $o_2$  的置信度由 1 降至 0.73,且出现区间扩张现象。文献[4]约简  $\{a_1\}$  的约简区间和分类质量都不发生变化,但对象  $o_1$  出现分类异常现象。分布约简  $C$  能保持对象各决策类的隶属程度不变,但实际分类器仅选用对象最大决策类规则,因此约简条件过于严格。最大决策熵约简  $\{a_1, a_3, a_4\}$  相对于分布约简更为简洁,能有效地去除冗余属性,同时保持各对象的最大决策及置信度都不发生变化,达到属性约简的目的。

表 1 对象决策属性划分为  $U/D = \{\{o_1, o_7\}, \{o_2, o_3, o_4, o_5, o_8, o_9, o_{10}, o_{11}\}, \{o_6, o_{12}\}\}$ 。依据信度差别矩阵的定义,等价类划分后的信度差别矩阵如表 3。

属性约简。

## 4 结语

在粗糙集理论中,属性约简是保持决策表分类能力不变的条件属性子集。经典粗糙集模型的分类质量、正域及下近似三者等价且约简后的规则具有确定性,因此可定义单调的度量获取决策表的约简。但在概率粗糙集模型下,三类概念不再等价且对象生成的往往是概率规则,此时决策表的分类能力不能仅用分类质量、正域或下近似表示,需要考虑各对象

(下转第 1073 页)



进行比较。每组实验做3次,取平均值。结果如表1所示。

从表1可以看出,用WIFSCM模型进行图像分割所用的时间远远小于IFCM算法的平均运算时间。当图像尺寸不断变大时,两种算法所耗费时间的比例却不断降低,更加显示了WIFSCM模型对大数据样本集进行聚类的优势。



图2 待聚类图像

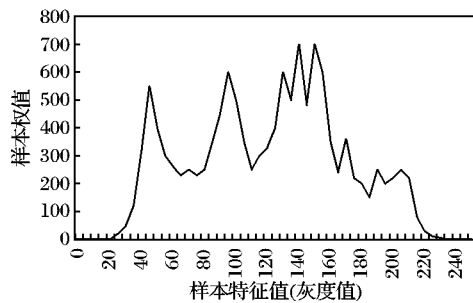


图3 样本权值分布



图4 本文模型聚类结果

表1 WIFSCM模型和IFCM对图像分割的性能比较

图像尺寸/ 像素	C	待聚类样本数		平均运算时间/ms		时间花费 比例/%
		WIFSCM	IFCM	WIFSCM	IFCM	
256 × 256	3	229	65 536	165	15 995	1.03
	4	229	65 536	216	51 312	0.42
	5	229	65 536	334	63 796	0.52
512 × 512	3	234	262 144	187	64 877	0.30
	4	234	262 144	569	223 568	0.26
	5	234	262 144	413	259 997	0.16

### 3 结语

本文受到已有研究的启发,提出了基于加权直觉模糊集合的聚类模型,该模型主要针对大样本集的聚类。该模型利用直觉模糊集描述样本,从而使模型具有更强的表达能力,更具灵活性;用等价样本策略缩减直觉模糊集合中的样本数,得到加权直觉模糊集合,从而提高了聚类速度;利用加权直觉模糊集合密度函数确定初始聚类中心,减少了迭代的次数,提高了收敛的速度。

#### 参考文献:

- [1] ZADEH L A. Fuzzy sets[J]. Information and Control, 1965(8): 338-356.
- [2] ATANASSOV K. Intuitionistic fuzzy sets[J]. Fuzzy Sets and Systems, 1986, 20(1): 87-96.
- [3] PASHA E A, FATEMI A. Intuitionistic fuzzy sets clustering (IFSC) with an application in psychology[J]. Journal of Mathematics and Applications, 2006(28): 113-120.
- [4] 雷英杰, 王宝树, 胡军红. 直觉模糊等价矩阵构造方法[J]. 系统工程理论与实践, 2007, 16(7): 127-131.
- [5] 张洪美, 徐泽水, 陈琦. 直觉模糊集的聚类方法研究[J]. 控制与决策, 2007, 22(8): 882-888.
- [6] 吴成茂. 模糊C-均值算法在直觉模糊数聚类中的应用[J]. 计算机工程与应用, 2009, 45(16): 141-145.
- [7] 徐小米, 雷英杰, 赵学军. 基于直觉模糊熵的直觉模糊聚类[J]. 空军工程大学学报, 2008, 9(2): 80-83.
- [8] 申晓勇, 雷英杰, 蔡茹, 等. 一种基于密度函数的直觉模糊聚类初始化方法[J]. 计算机科学, 2009, 36(5): 197-199.
- [9] 申晓勇, 雷英杰, 蔡茹, 等. 基于目标函数的直觉模糊集合数据的聚类方法[J]. 系统工程与电子技术, 2009, 31(11): 2732-2735.

(上接第1069页)

约简后的置信度,以保证约简前后对象决策及置信度的一致性。本文提出的带参最大决策熵及信度矩阵能保证约简后对象正域和置信度都不发生变化,因此能有效地消除已有模型各类约简异常,达到属性约简的目的。进一步工作将研究邻域粗糙集下最大决策熵约简模型,并通过实验来分析和验证模型的效率。

#### 参考文献:

- [1] PAWLAK Z. Rough sets[J]. International Journal of Computer and Information Science, 1982, 11(5): 341-356.
- [2] PAWLAK Z. Rough sets: Theoretical aspects of reasoning about data[M]. Dordrecht: Kluwer Academic Publishers, 1991.
- [3] YAO Y Y. Probabilistic rough set approximations[J]. International Journal of Approximate Reasoning, 2008, 49(2): 255-271.
- [4] ZIARKO W. Variable precision rough set model[J]. Journal of

Computer and System Science, 1993, 46(1): 39-59.

- [5] BEYNON M. Reducts within the variable precision rough sets model: An further investigation[J]. European Journal of Operational Research, 2001, 134(3): 592-605.
- [6] WANG J Y, ZHOU J. Research of reduct features in the variable precision rough set model[J]. Neurocomputing, 2009, 72(10/11/12): 2643-2648.
- [7] 张文修, 米据生, 吴伟志. 不协调目标信息系统的知识约简[J]. 计算机学报, 2003, 26(1): 12-18.
- [8] MI J S, WU W Z, ZHANG W X. Approaches to knowledge reduction based on variable precision rough set model[J]. Information Sciences, 2004, 159(3/4): 255-272.
- [9] 余承依, 李进金. 不协调决策信息系统最大分布约简新方法[J]. 计算机应用, 2011, 31(6): 1645-1659.