

差别矩阵约简表示及其快速算法实现

葛浩^{1a,2}, 李龙澍², 杨传健^{1b}

(1. 滁州学院 a. 电子与电气工程学院, b. 计算机与信息工程学院, 安徽 滁州 239000;
2. 安徽大学 计算智能与信号处理教育部重点实验室, 合肥 230601)

摘要: 差别矩阵可以拥有不同的信息, 根据差别矩阵描述的区分信息量不同, 给出4种差别矩阵定义, 并提出相应H-约简、S-约简、B-约简和P-约简的概念; 研究4种约简之间的关系, 构建通用约简算法模型. 为了提高约简算法的效率, 给出相对分辨能力约简定义(RD-约简), 揭示相对分辨能力约简与4种差别矩阵约简之间的等价性, 进而设计相对分辨能力快速约简算法. 最后, 通过实例和UCI数据集验证了所提出约简算法的有效性和时空性能.

关键词: 粗糙集; 差别矩阵; 分辨能力; 核属性; 约简

中图分类号: TP181

文献标志码: A

Discernibility matrix-based reduct representation and quick algorithms

GE Hao^{1a,2}, LI Long-shu², YANG Chuan-jian^{1b}

(1a. School of Electronic and Electrical Engineering, 1b. School of Computer and Information Engineering, Chuzhou University, Chuzhou 239000, China; 2. Key Laboratory of Computation Intelligence and Signal Processing of Ministry of Education, Anhui University, Hefei 230601, China. Correspondent: GE Hao, E-mail: togehao@126.com)

Abstract: The discernibility matrix can possess different information. In this paper, according to different discernible information quantity, four descriptions of the discernibility matrix are proposed, and the four reduct definitions, i.e., H-reduct, S-reduct, B-reduct and P-reduct are presented. The relationship and equivalence for four reducts are researched, and two general reduction algorithms are designed. In order to improve the efficiency of the reduction algorithm, the related concepts of the relative discernibility-based reduct(denoted as RD-reduct) are proposed. The equivalence between the relative discernibility-based reduct and four discernibility matrix-based reducts are revealed. Two effective reduction algorithms based on the relative discernibility are designed. Finally, the example and experiments are used to explain the feasibility and effectiveness of the proposed algorithms.

Keywords: rough set; discernibility matrix; discernibility; core attributes; attribute reduct

0 引言

粗糙集理论^[1]作为一种强有力的软计算工具, 被广泛应用于决策分析、数据挖掘、机器学习和故障诊断等方面. 属性约简是粗糙集理论的核心内容之一, 是在保持知识库中某种性质不变的情况下, 删除不相关和冗余属性, 只保留能够保持该特性的最小属性集. 许多学者对粗糙集约简方法进行研究, 并取得了较好的结果^[2-4].

约简定义与维持目标特性息息相关, 对于不同的目的, 约简的概念和性质则不同, 获得的约简结果未必一样. 因此, Miao等^[5]指出约简无所谓对错, 约简的

标准不同, 约简结果不一定相同. Pawlak^[1]给出的是正域约简概念, 它保持了知识库正区域的不变性. 杨明^[6]提出了正域差别矩阵构建方法, 并设计了正域约简求解算法. 根据 Skowron 差别矩阵^[7], Hu等^[8]设计了差别矩阵约简方法, 它保持了知识库分辨能力的不变性. Kryzkiewicz^[9]提出了分布约简和分配约简的概念, 前者保持了每个对象在决策类的隶属度不变, 后者保持了每个对象可能的决策类不变. 张文修等^[10]在 Kryzkiewicz 研究的基础上提出了最大分布约简的概念, 该约简保持了对象最大决策类的不变性. Miao等^[5]针对不协调决策表提出了正域约简、决策约简和

收稿日期: 2014-11-29; **修回日期:** 2015-02-07.

基金项目: 国家自然科学基金项目(51307011, 61402005); 安徽省自然科学基金项目(1308085QF114, 1508085MF126, 1508085MF127); 安徽省高等学校省级自然科学基金项目(KJ2013A015, KJ2012A212); 滁州学院科技优秀人才基金重点项目(2013RC003); 计算智能与信号处理教育部重点实验室开放课题基金项目.

作者简介: 葛浩(1976-), 男, 副教授, 博士生, 从事数据挖掘、粒计算和粗糙集的研究; 李龙澍(1956-), 男, 教授, 博士生导师, 从事不精确信息处理、智能软件等研究.

关系约简, 并针对这些约简给出一个统一的约简框架. Zhou等^[11]研究了13种约简以及它们之间的关系, 指出13种约简中对不协调决策表存在6种约简, 协调决策表上仅存在2种约简. 针对不完备决策表, Meng等^[12]研究了8种约简及其之间的关系, 指出本质上只有5种约简, 并构建分辨函数, 设计了约简求解算法. 为了提高不同约简算法的效率, 蒋云良等^[13]从条件信息熵着手, 通过哈希分类实现高效的分布约简算法. 针对分布约简、分配约简和最大分布约简, 为了避免直接采用定义求解约简的不便, Li等^[14]分别针对分布约简、最大分布约简和分配约简给出新的属性重要性度量标准, 并设计了3种快速约简算法.

上述方法从不同侧面研究在保持决策表某种特性不变的情况下, 不同约简概念、性质、之间关系, 以及基于差别矩阵或差别函数的约简实现. 然而, 对于差别矩阵中可以拥有怎样信息、信息多少与约简表示形式之间的关系, 并没有给予深入讨论. 本文则从差别矩阵可含有描述信息不同着手, 定义4种差别矩阵, 研究它们之间的关系, 并给出相应约简定义; 同时提出添加和删除两种策略的通用差别矩阵约简模型. 为了解决差别矩阵约简算法效率低下的不足, 提出相对分辨能力约简概念, 研究差别矩阵约简和相对分辨能力约简之间的关系, 设计快速启发式属性约简算法. 最后通过实例表明了本文方法的有效性, 并利用UCI中的数据对比分析了差别矩阵约简与相对分辨能力约简的时空性能.

1 差别矩阵及其属性约简

决策表可以表示为 $DT = (U, A = C \cup D, V, f)$. 其中: A 为属性集, 由条件属性集 C 和决策属性集 D 构成, 且 $C \cap D = \emptyset$; $V = \bigcup_{a \in A} V_a$, V_a 是属性 a 的值域; $f: U \times A \rightarrow V$ 为信息函数, $\forall a \in A, x \in U$, 有 $f(x, a) \in V_a$. 设 $P \subseteq C$, $IND(P) = \{(x, y) \in U \times U \mid f(x, a) = f(y, a), \forall a \in P\}$ 称为 U 的不可区分关系. 不可区分关系实为等价关系, 含 x 的等价类记为 $[x]_P = \{y \mid y \in U, (x, y) \in IND(P)\}$.

1.1 4种差别矩阵

定义1 决策表为DT, 设 $R \subseteq C$, 称 $M_H^R = \{m_{ij}^R \mid 1 \leq i, j \leq |U|\}$ 为属性集 R 的Hu差别矩阵, m_{ij}^R 表示矩阵中 i 行 j 列的元素, 其满足

$$m_{ij}^R = \begin{cases} \{a \mid a \in R, \\ f(x_i, a) \neq f(x_j, a) \wedge f(x_i, D) \neq f(x_j, D); \\ \emptyset, \text{ otherwise.} \end{cases} \quad (1)$$

定义2 决策表为DT, 设 $a \in C$, 称 $M_B^{\{a\}} = \{\lambda_{ij}^a \mid 1 \leq i, j \leq |U|\}$ 为属性集 a 的布尔差别矩阵, λ_{ij}^a 表示矩阵 i 行 j 列的元素, 其满足

$$\lambda_{ij}^a = \begin{cases} 1, f(x_i, a) \neq f(x_j, a) \wedge f(x_i, D) \neq f(x_j, D); \\ 0, \text{ otherwise.} \end{cases} \quad (2)$$

定义3 决策表为DT, 设 $a \in C$, 属性 a 的布尔差别矩阵为 $M_B^{\{a\}}$, 则 $M_B^{\{a\}}$ 的势记为 $p(a)$, 其满足

$$p(a) = \sum_{i=1}^{|U|} \sum_{j=1}^{|U|} \lambda_{ij}^a. \quad (3)$$

$p(a)$ 实际为 $M_B^{\{a\}}$ 中1的个数, 也称为属性 a 的频率.

定义4 决策表为DT, 设 $R \subseteq C$, 称 $M_S^R = \{\delta_{ij}^R \mid 1 \leq i, j \leq |U|\}$ 为属性集 R 的结构差别矩阵, δ_{ij}^R 表示矩阵 i 行 j 列的元素, 其满足

$$\delta_{ij}^R = \sum_{k=1}^{|R|} \lambda_{ij}^{a_k}, a_k \in R. \quad (4)$$

定义5 决策表为DT, 设 $R \subseteq C$, 称 $M_B^R = \{\lambda_{ij}^R \mid 1 \leq i, j \leq |U|\}$ 为属性集 R 的布尔差别矩阵, λ_{ij}^R 表示矩阵 i 行 j 列的元素, 其满足

$$\lambda_{ij}^R = \begin{cases} 1, \exists a \in R, f(x_i, a) \neq f(x_j, a) \wedge \\ f(x_i, D) \neq f(x_j, D); \\ 0, \text{ otherwise.} \end{cases} \quad (5)$$

定义6 决策表为DT, 设 $R \subseteq C$, 属性 R 的布尔差别矩阵为 M_B^R , 用 $|M_B^R|$ 表示矩阵的势, 其满足

$$|M_B^R| = \sum_{i=1}^{|U|} \sum_{j=1}^{|U|} \lambda_{ij}^R. \quad (6)$$

$|M_B^R|$ 实际为 M_B^R 中非0元素的个数(即 M_B^R 中元素的累加和).

差别矩阵直观地反映了对象之间的区分关系. 在上面的描述形式中, M_H^R 不仅描述了矩阵中某行某列是否含有元素, 而且还反映了该元素是由哪些属性组成; M_S^R 不仅描述了某行某列是否非空, 而且还能描述该元素由几个属性构成, 但具体由哪些属性组成, 则不明确; M_B^R 仅能描述某行某列是否非空, 具体由几个属性构成, 则不明确; $|M_B^R|$ 能描述矩阵中非空元素的数量.

例1 表1为一个决策表, 其中 $\{a, b, c\}$ 为条件属性 C , D 为决策属性, 有4个样本对象 $U = \{x_1, x_2, x_3, x_4\}$.

设 $R = \{ac\}$, 则有

$$M_H^R = \begin{bmatrix} \emptyset & ac & c & \emptyset \\ ac & \emptyset & a & ac \\ c & a & \emptyset & \emptyset \\ \emptyset & ac & \emptyset & \emptyset \end{bmatrix}, M_S^R = \begin{bmatrix} 0 & 2 & 1 & 0 \\ 2 & 0 & 1 & 2 \\ 1 & 1 & 0 & 0 \\ 0 & 2 & 0 & 0 \end{bmatrix},$$

$$M_B^R = \begin{bmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}, |M_B^R| = 8.$$

表1 决策表S

U	a	b	c	D
x_1	1	0	1	0
x_2	0	0	0	1
x_3	1	1	0	2
x_4	1	1	1	2

在属性集 R 上, 以 x_1 和 x_3 为例, 由 M_H^R 可知 x_1 和 x_3 可被属性 c 区分; 从 M_S^R 可以发现 x_1 和 x_3 可被 1 个属性区分, 具体是哪个属性则不明了; 从 M_B^R 仅能发现 x_1 和 x_3 可被区分, 具体由哪些和几个属性区分则不明确; $|M_H^R|$ 仅反映由属性集 R 区分的对象对的个数是 8, 但 x_1 和 x_3 互相是否可以被区分, 则不清楚.

1.2 核属性和属性约简

定义 7 决策表为 DT, 属性集 C 的 Hu 差别矩阵为 M_H^C , 核属性集 $\text{CORE}_H(C)$ 定义为^[8]

$$\text{CORE}_H(C) = \{a \in C \mid m_{ij}^C = \{a\}\}. \quad (7)$$

定义 8 决策表为 DT, 属性集 C 的结构差别矩阵为 M_S^C , 核属性集 $\text{CORE}_S(C)$ 定义为

$$\text{CORE}_S(C) = \{a \in C \mid \delta_{ij}^C = 1 \text{ 且 } \delta_{ij}^a = 1\}. \quad (8)$$

定义 9 决策表为 DT, 属性集 C 的布尔差别矩阵为 M_B^C , 核属性集 $\text{CORE}_B(C)$ 定义为

$$\text{CORE}_B(C) = \{a \in C \mid M_B^C \neq M_B^{C-\{a\}}\}. \quad (9)$$

定义 10 决策表为 DT, 属性集 C 的布尔差别矩阵为 M_B^C 且势为 $|M_B^C|$, 核属性集 $\text{CORE}_P(C)$ 定义为

$$\text{CORE}_P(C) = \{a \in C \mid |M_B^C| \neq |M_B^{C-\{a\}}|\}. \quad (10)$$

性质 1 决策表为 DT, 有

$$\begin{aligned} \text{CORE}_H(C) &= \text{CORE}_S(C) = \\ \text{CORE}_B(C) &= \text{CORE}_P(C). \end{aligned} \quad (11)$$

定义 11 设 RedS_H 为 DT 基于 Hu 差别矩阵的 H-约简集, 对于 $\forall R \subseteq C, R \in \text{RedS}_H$ 满足以下条件^[8]:

- 1) $\forall \emptyset \neq m_{ij}^C \in M_H^C$, 有 $m_{ij}^R \neq \emptyset$;
- 2) $\forall a \in R, \exists \emptyset \neq m_{ij}^R \in M_H^C$, 有 $m_{ij}^{R-\{a\}} = \emptyset$.

定义 12 设 RedS_S 为 DT 基于结构差别矩阵的 S-约简集, 对于 $\forall R \subseteq C, R \in \text{RedS}_S$ 满足以下条件:

- 1) $\forall 0 \neq \delta_{ij}^C \in M_S^C$, 有 $\delta_{ij}^R \neq 0$;
- 2) $\forall a \in R, \exists 0 \neq \delta_{ij}^R \in M_S^C$, 有 $\delta_{ij}^{R-\{a\}} = 0$.

定义 13 设 RedS_B 为 DT 基于布尔差别矩阵的 B-约简集, 对于 $\forall R \subseteq C, R \in \text{RedS}_B$ 满足以下条件:

- 1) $M_B^C = M_B^R$;
- 2) $\forall a \in R, M_B^C \neq M_B^{R-\{a\}}$.

定义 14 设 RedS_P 为 DT 基于结构差别矩阵势的 P-约简集, 对于 $\forall R \subseteq C, R \in \text{RedS}_P$ 满足以下条件:

- 1) $|M_B^C| = |M_B^R|$;
- 2) $\forall a \in R, |M_B^C| \neq |M_B^{R-\{a\}}|$.

上述 4 个定义依据所提供的区分信息量不同, 从不同层面给出 4 种差别矩阵约简描述. 定义 11 从矩阵元素中区分属性具体描述形式来表示, 定义 12 从矩阵元素中区分属性的量化标准来表示, 定义 13 从矩阵中某一位置是否存在区分属性的标准来表示, 定义 14 从矩阵中区分元素的量化标准来表示. 可以发现, 定义 11~定义 13 均需要创建相应的差别矩阵, 通过相应差别矩阵可以直观地观察到某行某列的两个对象是否可以被区分(或分辨), 以及不同的区分信息量; 定义 14 仅需执行差别矩阵中元素的生成运算, 而无需存储. 从分辨能力上看, 4 个约简定义均可以反映知识的分辨能力, 从这个层次上来说 4 种约简定义的结果应是一致的. 下面给出相关定理.

定理 1 对于决策表 DT, M_H^R 为 Hu 差别矩阵, M_S^R 为结构差别矩阵, 设 $R \subseteq C, \forall \emptyset \neq m_{ij}^C$, 有 $m_{ij}^R \neq \emptyset \Leftrightarrow \forall 0 \neq \delta_{ij}^C$, 有 $\delta_{ij}^R \neq 0$.

定理 2 对于决策表 DT, M_H^R 为 Hu 差别矩阵, M_S^R 为结构差别矩阵, 设 $R \subseteq C, \forall a \in R, \exists \emptyset \neq m_{ij}^R$, 有 $m_{ij}^{R-\{a\}} = \emptyset \Leftrightarrow \forall a \in R, \exists 0 \neq \delta_{ij}^R$, 有 $\delta_{ij}^{R-\{a\}} = 0$.

定理 1 和定理 2 均可由反证法得证, 此略.

由定理 1 和定理 2 可知, Hu 差别矩阵约简与结构差别矩阵约简的结果一致. 同理, Hu 差别矩阵约简与差别矩阵势的约简以及布尔差别矩阵约简结果均相同. 因此, 上述 4 个约简定义获得的约简结果相同.

性质 2 设 Hu 差别矩阵约简的空间开销为 SPACE_H , 结构差别矩阵约简的空间开销为 SPACE_S , 布尔差别矩阵约简的空间开销为 SPACE_B , 基于差别矩阵势的约简空间开销为 SPACE_P , 则

$$\text{SPACE}_H > \text{SPACE}_S > \text{SPACE}_B > \text{SPACE}_P. \quad (12)$$

为了便于一般化描述, 采用 M_Δ 表示某种差别矩阵(其中 $\Delta \in \{H, S, B, P\}$), $\Delta(C)$ 表示属性集 C 对应于差别矩阵 M_Δ 所拥有的特性. $\Delta(C) = \Delta(R)$ 表示属性集 C 与属性集 R 关于差别矩阵 M_Δ 具有相同的特性; 否则, 称属性集 C 与属性集 R 关于差别矩阵 M_Δ 具有不同的特性.

给出如下核属性和约简一般表示形式.

定义 15 决策表为 DT, 属性集 C 的 Δ -差别矩阵为 M_Δ , 其核属性集 $\text{CORE}_\Delta(C)$ 可表示为

$$\text{CORE}_\Delta(C) = \{a \in C \mid \Delta(C) \neq \Delta(C - \{a\})\}. \quad (13)$$

定义 16 决策表为 DT, 一般差别矩阵为 M_Δ , 设 RedS_Δ 为 DT 关于 M_Δ 的 Δ -约简集, 对于 $\forall R \subseteq C$, $R \in \text{RedS}_\Delta$ 满足以下两个条件:

- 1) $\Delta(C) = \Delta(R)$;
- 2) $\forall a \in R, \Delta(R) \neq \Delta(R - \{a\})$.

2 差别矩阵约简实现

通过差别矩阵或差别函数可以获得所有约简和最小约简, 但是均为 NP-hard 问题^[15], 采用启发式方法可以获得一个约简. 添加策略和删除策略是两种常用的启发式搜索策略^[16]. 添加策略通常从空集出发或以核属性为初始约简集 R , 每次选择具有最大重要性的属性加入 R 中, 直到 R 与原始条件属性集 C 具有相同的分类能力或分辨能力. 删除策略一般以属性集 C 的一个序关系为出发点, 依次从当前属性序中选择一个属性, 判断其是否可以删除, 直到所有属性均被检查.

2.1 添加策略属性约简

在差别矩阵中, 属性初始频率 $p(a)$ 反映了属性 a 的区分能力: 频率越大, 说明区分的对象对越多, 重要性越强; 反之, 重要性越弱. 添加策略约简以核属性为初始约简集, 条件属性的初始频率为启发式信息, 不断将频率大的属性加入约简集中, 直到找到约简. 下面给出 4 种差别矩阵添加策略的约简框架.

算法 1 添加策略的差别矩阵约简算法(简称为 Δ -约简-Add, 其中 $\Delta \in \{H, S, B, P\}$).

输入: 决策表 $\text{DT} = (U, C \cup D, V, f)$;

输出: 属性约简 R .

Step 1: $R = \emptyset$, 创建相应的差别矩阵 M_Δ .

Step 2: 计算条件属性频率 $p(a)$ 以及核属性 $\text{CORE}_\Delta(C)$, $R = \text{CORE}_\Delta(C)$.

Step 3: while (1)

if (R 满足 Δ -约简定义条件 1)) break;

从 $C - R$ 中选择具有最大初始值的属性, 记为 a , $R \leftarrow R \cup \{a\}$.

Step 4: Output 约简 R .

对于算法 1, H-约简算法的时间复杂度为 $O(|C||U|^2)$, 空间复杂度为 $O(|C||U|^2)$; S-约简算法的时间复杂度为 $O(|C||U|^2)$, 空间开销为 $O((|C|+1)|U|^2)$, 空间复杂度为 $O(|C||U|^2)$; B-约简算法时间和空间复杂度均为 $O(|C||U|^2)$; P-约简算法时间复杂度为 $O(|C||U|^2)$, 空间复杂度为 $O(|C|)$.

2.2 删除策略属性约简

在删除策略约简中, 依据属性初始频率递增顺序构建属性序关系; 然后, 根据该属性序依次判断当前属性必要性, 直到所有属性均检测一遍, 可以获得一个约简. 下面给出 4 种差别矩阵删除策略的约简描述.

算法 2 删除策略的差别矩阵约简算法(简称为 Δ -约简-Del, 其中 $\Delta \in \{H, S, B, P\}$).

输入: 决策表 $\text{DT} = (U, C \cup D, V, f)$;

输出: 属性约简 R .

Step 1: 创建相应差别矩阵 M_Δ , 获得条件属性频率并递增排序, 使得 $R = \{a_1, a_2, \dots, a_{|C|}\}$;

Step 2: for $i = 1$ to $|C|$ Do

if (a_i 不满足 Δ -约简定义条件 2)), 则

$R \leftarrow R - \{a_i\}$;

Step 3: Output 约简 R .

对于算法 2, H-约简、S-约简和 B-约简算法的时间和空间复杂度均分别为 $O(|C||U|^2)$ 和 $O(|C||U|^2)$; P-约简算法时间和空间复杂度分别为 $O(|C||U|^2)$ 和 $O(|C|)$.

2.3 实例分析

例 2 以表 1 所示决策表 DT 为例来说明 4 种差别矩阵约简过程.

$$M_H^C = \begin{bmatrix} \emptyset & ac & bc & b \\ ac & \emptyset & ab & abc \\ bc & ab & \emptyset & \emptyset \\ b & abc & \emptyset & \emptyset \end{bmatrix}, M_H^{\{ab\}} = \begin{bmatrix} \emptyset & a & b & b \\ a & \emptyset & ab & ab \\ b & ab & \emptyset & \emptyset \\ b & ab & \emptyset & \emptyset \end{bmatrix},$$

$$M_H^{\{bc\}} = \begin{bmatrix} \emptyset & c & bc & b \\ c & \emptyset & b & bc \\ bc & b & \emptyset & \emptyset \\ b & bc & \emptyset & \emptyset \end{bmatrix}, M_B^{\{a\}} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix},$$

$$M_B^{\{b\}} = \begin{bmatrix} 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{bmatrix}, M_B^{\{c\}} = \begin{bmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix},$$

$$M_S^C = \begin{bmatrix} 0 & 2 & 2 & 1 \\ 2 & 0 & 2 & 3 \\ 2 & 2 & 0 & 0 \\ 1 & 3 & 0 & 0 \end{bmatrix}, M_B^C = \begin{bmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{bmatrix},$$

$$M_B^{\{ab\}} = \begin{bmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{bmatrix}, M_B^{\{bc\}} = \begin{bmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{bmatrix},$$

$$M_S^{\{ab\}} = \begin{bmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 2 & 2 \\ 1 & 2 & 0 & 0 \\ 1 & 2 & 0 & 0 \end{bmatrix}, M_S^{\{bc\}} = \begin{bmatrix} 0 & 1 & 2 & 1 \\ 1 & 0 & 1 & 2 \\ 2 & 1 & 0 & 0 \\ 1 & 2 & 0 & 0 \end{bmatrix}.$$

对于决策表DT, 创建Hu差别矩阵 M_H^C , 结构差别矩阵 $M_S^{\{ab\}}$ 、 $M_S^{\{bc\}}$ 、 M_S^C 以及属性布尔差别矩阵 $M_B^{\{a\}}$ 、 $M_B^{\{b\}}$ 、 $M_B^{\{c\}}$ 、 $M_B^{\{ab\}}$ 、 $M_B^{\{bc\}}$ 、 M_B^C ; 有 $|M_B^C| = 10$ 、 $|M_B^{\{a\}}| = 6$ 、 $|M_B^{\{b\}}| = 8$ 、 $|M_B^{\{c\}}| = 8$, 以及 $p(a) = 6$ 、 $p(b) = 8$ 、 $p(c) = 6$. 进而得到 $\text{CORE}_H(C) = \text{CORE}_S(C) = \text{CORE}_B(C) = \text{CORE}_P(C) = \{b\}$. 最终, 约简集为 $\{ab, bc\}$.

对于H-约简算法, 采用添加策略, 先将核属性加入约简 $R = \{b\}$; 然后, 选择 $C - R = \{a, c\}$ 中最大初始频率的属性 a , 添加到约简集中, 得 $R = \{ab\}$; 最终 $R = \{ab\}$. 采用删除策略, 依据 $p(a) = 6$ 、 $p(b) = 8$ 、 $p(c) = 6$, 属性的序关系为 $a = c \prec b$, 对当前最小初始频率的属性 a 检测其必要性, a 可以删除; 再分别检查 b 和 c , 均不可删除; 最终 $R = \{bc\}$.

同样, 对于S-约简、B-约简和P-约简算法, 采用添加策略, 获得约简为 $R = \{ab\}$; 采用删除策略, 获得约简为 $R = \{bc\}$. 由例2可以发现, 添加策略与删除策略的约简结果不一样. 这是由于两种策略侧重点不同: 添加策略侧重于将重要的属性先添加到约简属性集中; 删除策略侧重于将不重要的属性先删除掉.

3 相对分辨能力约简实现

4种差别矩阵约简均有较高的时间复杂度. 空间复杂度方面, H-约简、S-约简和B-约简的空间复杂度也高达 $O(|C||U|^2)$. 虽然P-约简的空间复杂度仅为 $O(|C|)$, 但在约简过程中, 需要重新执行差别矩阵中元素的生成运算, 其实际的时间开销通常要高于前3种约简算法. 总之, 4种差别矩阵约简方法的时空效率较低, 不利于大规模数据集处理.

Pawlak的粗糙集理论认为知识代表了分辨事物的能力. 差别矩阵直观地反映了知识的分辨能力, 前

面4种差别矩阵均能反映出知识的分辨能力, 只不过拥有的区分信息量不同而已. 在信息系统中属性的分辨能力体现为可分辨对象对的数目, 数目越多分辨能力越强, 也越重要.

下面将从分辨能力的角度来研究约简求解, 揭示分辨能力约简与前面4种约简之间的关系; 另外, 为了提高算法求解效率, 设计两种策略的快速约简算法.

3.1 分辨能力相关概念和性质

定义 17 设信息系统 $\text{IS} = (U, A, V, f)$, $R \subseteq A$, 给出如下分辨关系的相关定义:

$$\begin{aligned} \text{DIS}(R) &= \{(x, y) \in U \times U \mid \exists a \in R, f(x, a) \neq f(y, a)\}, \\ \text{NDIS}(R) &= \{(x, y) \in U \times U \mid \forall a \in R, f(x, a) = f(y, a)\}, \\ W(R) &= |\text{DIS}(R)| = |\{(x, y) \in U \times U \mid \exists a \in R, f(x, a) \neq f(y, a)\}|, \\ \overline{W}(R) &= |\text{NDIS}(R)| = |\{(x, y) \in U \times U \mid \forall a \in R, f(x, a) = f(y, a)\}|. \end{aligned} \quad (14)$$

其中: $\text{DIS}(R)$ 表示属性集 R 的分辨关系, $\text{NDIS}(R)$ 表示属性集 R 的不可分辨关系; $W(R)$ 表示属性集 R 的分辨能力, $\overline{W}(R)$ 表示属性集 R 的不可分辨能力. 显然, 不可分辨关系为一个等价关系. 由 R 在 U 上导出等价类的划分记为 U/R , 同一等价类中任意两个对象是不可区分的, 不同等价类中任意两个对象是可以区分的.

定义 18 设决策表 $\text{DT} = (U, C \cup D, V, f)$, $R \subseteq A$, 给出如下相对于决策属性分辨关系的相关定义:

$$\begin{aligned} \text{DIS}(R|D) &= \{(x, y) \in U \times U \mid \exists a \in R, f(x, a) \neq f(y, a) \wedge f(x, D) \neq f(y, D)\}, \\ \text{NDIS}(R|D) &= \{(x, y) \in U \times U \mid \forall a \in R, f(x, a) = f(y, a) \vee f(x, D) = f(y, D)\}, \\ W(R) &= |\text{DIS}(R|D)| = |\{(x, y) \in U \times U \mid \exists a \in R, f(x, a) \neq f(y, a) \wedge f(x, D) \neq f(y, D)\}|, \\ \overline{W}(R|D) &= |\text{NDIS}(R|D)| = |\{(x, y) \in U \times U \mid \forall a \in R, f(x, a) = f(y, a) \vee f(x, D) = f(y, D)\}|. \end{aligned} \quad (15)$$

其中: $\text{DIS}(R|D)$ 表示属性集 R 的相对分辨关系, $\text{NDIS}(R|D)$ 表示属性集 R 的相对不可分辨关系; $W(R|D)$ 表示属性集 R 的相对分辨能力, $\overline{W}(R|D)$ 表示属性集 R 的相对不可分辨能力.

根据以上定义, 给出如图1所示的关系, 并不难

得到以下性质.

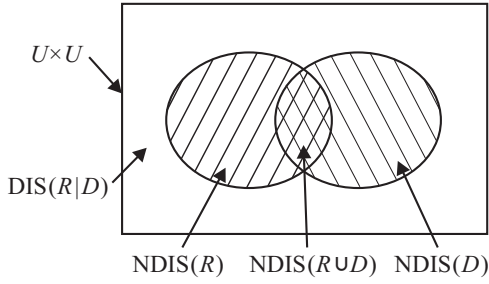


图1 关系图

性质3 决策表为DT, $R \subseteq C$, $W(R|D)$ 为 R 的相对分辨能力, 则有

$$W(R|D) = |U|^2 - \overline{W}(D) - \overline{W}(R) + \overline{W}(R \cup D). \quad (16)$$

由容斥原理, 性质3可以得证, 此略.

性质3表明, 相对分辨能力可由不可分辨能力获得, 采用这种方法计算相对分辨能力可以大大减少计算量.

性质4 决策表为DT, 设 $R \subseteq C$, 令 $U/R = \{X_1, X_2, \dots, X_n\}$, $U/D = \{Y_1, Y_2, \dots, Y_m\}$, $U/(D \cup R) = \{Z_1, Z_2, \dots, Z_t\}$, R 的相对分辨能力 $W(R|D)$ 为

$$W(R|D) = |U|^2 - \overline{W}(D) - \overline{W}(R) + \overline{W}(R \cup D) = |U|^2 - \sum_{i=1}^m |Y_i|^2 - \sum_{j=1}^n |X_j|^2 + \sum_{k=1}^t |Z_k|^2. \quad (17)$$

例3 针对表1中的决策表DT, 设 $R = \{ac\}$, 有 $U/D = \{\{x_1\}, \{x_2\}, \{x_3, x_4\}\}$, $U/R = \{\{x_1, x_4\}, \{x_3\}, \{x_2\}\}$ 和 $U/D \cup R = \{\{x_1\}, \{x_2\}, \{x_3\}, \{x_4\}\}$. 经计算可得 $\overline{W}(D) = 2 \times 1^2 + 2^2 = 6$, $\overline{W}(R) = 2^2 + 2 \times 1^2 = 6$, $\overline{W}(D \cup R) = 4 \times 1^2 = 4$. 最终有 $W(R|D) = |U|^2 - \overline{W}(D) - \overline{W}(R) + \overline{W}(R \cup D) = 4^2 - 6 - 6 + 4 = 8$.

定义19 决策表为DT, $R \subseteq C$, R 为相对分辨能力的属性约简, 其满足:

- 1) $W(R|D) = W(C|D)$;
- 2) $\forall a \in R, W(R - \{a\}|D) < W(R|D)$.

Hu 差别矩阵通过描述某个位置是否存在元素, 以及元素是由哪些属性构成, 直观地反映了知识的分辨关系和分辨能力. 从分辨能力角度上, 相对分辨能力约简(简称RD-约简)与H-约简、S-约简、B-约简和P-约简是一致的. 下面仅给出RD-约简与H-约简等价性的证明.

定理3 设 M_H^C 为Hu差别矩阵, $R \subseteq C, \forall \emptyset \neq m_{ij}^C \in M_H^C$, 则有

$$R \cap m_{ij}^C \neq \emptyset \Leftrightarrow W(R|D) = W(C|D). \quad (18)$$

证明 1) \Rightarrow . 反证法. 假设 $W(R|D) < W(C|D)$, 由定义18可知 $\exists x_i, x_j \in U (i \neq j)$, 有 $f(x_i, R) = f(x_j, R) \wedge f(x_i, D) \neq f(x_j, D)$, 但 $f(x_i, C) \neq f(x_j, C)$, 则 $\exists a \in C - R$, 使得 $f(x_i, a) \neq f(x_j, a) \wedge f(x_i, D) \neq f(x_j, D)$. 由定义1知 $a \in m_{ij}^C$, 使得 $m_{ij}^C \neq \emptyset$. 可得 $\exists \emptyset \neq m_{ij}^C \in M_H^C$, 有 $R \cap m_{ij}^C = \emptyset$. 此与定理3前件矛盾. 故 \Rightarrow 得证.

2) \Leftarrow . 反证法. 假设 $\exists \emptyset \neq m_{ij}^C \in M_H^C$, 有 $R \cap m_{ij}^C = \emptyset$, 则 $f(x_i, R) = f(x_j, R) \wedge f(x_i, D) \neq f(x_j, D)$. 由定义18知 $(x_i, x_j) \in \text{NDIS}(R|D)$, 即 $(x_i, x_j) \notin \text{DIS}(R|D)$. 由于 $\exists \emptyset \neq m_{ij}^C \in M_H^C$, 则 $\exists a \in C - R$, 使得 $f(x_i, a) \neq f(x_j, a) \wedge f(x_i, D) \neq f(x_j, D)$, 即 $f(x_i, C) \neq f(x_j, C)$, 故 $W(R|D) \neq W(C|D)$. 此与定理3后件矛盾, 故 \Leftarrow 得证. \square

定理4 设 M_H^C 为Hu差别矩阵, 若 $R \subseteq C, \forall a \in R$, 则 $\exists m_{ij}^C \in M_H^C$,

$$m_{ij}^C \cap R - \{a\} = \emptyset \Leftrightarrow W(R - \{a\}|D) < W(R|D). \quad (19)$$

证明过程同定理3, 采用反证法即可得证, 此略.

由定理3和定理4可知, RD-约简与H-约简是等价的. 同理, 可以得证RD-约简与S-约简、B-约简、P-约简均是等价的. 因此, 可以通过求RD-约简来获得差别矩阵约简, 避免相应差别矩阵的创建和存储, 从而提高约简求解效率.

性质5 决策表为DT, 设 $R \subseteq C$, 则 $W(R|D) = |M_B^R|$; 若 $W(R|D) = W(C|D)$, 则 $|M_B^R| = |M_B^C|$.

定义20 决策表为DT, 设相对分辨能力的核属性集为 $\text{CORE}_W(C), \forall a \in C, a \in \text{CORE}_W(C)$, 有 $W(C - \{a\}|D) < W(C|D)$.

定理5 决策表为DT, Hu差别矩阵核属性为 $\text{CORE}_H(C)$, 相对分辨能力核属性为 $\text{CORE}_W(C)$, 则 $\text{CORE}_H(C) = \text{CORE}_W(C)$.

定义21 决策表为DT, 设 $a \in R, R \subseteq C$, 则 a 的重要性 $\text{Sig}^-(a, R, D)$ 定义为

$$\text{Sig}^-(a, R, D) = W(R|D) - W(R - \{a\}|D). \quad (20)$$

定义22 决策表为DT, 设 $a \in C, R \subseteq C$ 且 $a \in C - R$, 则 a 的重要性 $\text{Sig}^+(a, R, D)$ 定义为

$$\text{Sig}^+(a, R, D) = W(R \cup \{a\}|D) - W(R|D). \quad (21)$$

性质6 决策表为DT, $\forall a \in C$ 且 $a \in \text{CORE}_W(C)$, 有 $\text{Sig}^-(a, C, D) > 0$.

3.2 基于相对分辨能力的启发式约简算法

添加策略的启发式约简算法是以核属性为初始约简集, 每次将重要性 $\text{Sig}^+(a, R, D)$ 最大的属性加入约简集 R 中, 直到满足条件 $W(R|D) = W(C|D)$, 便

可获得一个约简.

算法 3 添加策略的相对分辨能力约简算法 (RD-约简-Add).

输入: 决策表 $DT = (U, C \cup D, V, f)$;

输出: 属性约简 R .

Step 1: $R = \emptyset$, 求解 $\text{CORE}_W(C)$.

Step 2: $R = \text{CORE}_W(C)$, 计算 $W(R|D)$.

Step 3: while ($W(R|D) \neq W(C|D)$)

$\forall a_i \in C - R$, 求解 $\text{Sig}^+(a_i, R, D)$, 选择具有最大值的属性, 记为 a ;

$R \leftarrow R \cup \{a\}$, 并更新 $W(R|D)$.

Step 4: Output R .

本文中等价类划分均采用基数排序方法实现. 因此, Step 1 的时间复杂度为 $O(|C|^2|U|)$; Step 2 的时间复杂度为 $O(|\text{CORE}_W(C)||U|)$; Step 3 的时间复杂度为 $O(|C|^2|U|)$. 算法 3 的总时间复杂度为 $O(|C|^2|U|)$.

删除策略的启发式约简算法按照 $W(a|D)$ 对条件属性排序, 获得一个非递减的属性序列, 依次对该序列中每个属性检测其必要性, 直到所有属性均检测一遍后, 即可获得一个约简.

算法 4 删除策略的相对分辨能力约简算法 (RD-约简-Del).

输入: 决策表 $DT = (U, C \cup D, V, f)$;

输出: 属性约简 R .

Step 1: 对于 $\forall a_i \in C$, 求解 $W(a_i|D)$, 按值递增排序, 获得 $R = \{a_1, a_2, \dots, a_{|C|}\}$ 和 $W(R|D)$.

Step 2: for $i = 1$ to $|C|$ do

计算 $W(R - \{a_i\}|D)$;

if ($\text{Sig}^-(a_i, R, D) = 0$) $R = R - \{a_i\}$.

Step 3: Output R .

算法 4 中, Step 1 的时间复杂度为 $O(|C||U|)$; Step 2 的时间复杂度为 $O(|C|^2|U|)$. 因此, 算法 4 的总

时间复杂度为 $O(|C|^2|U|)$.

对表 1 采用算法 3, 经计算求得 $W(C|D) = 10$, $W(C - \{a\}|D) = W(C - \{c\}|D) = 10$, $W(C - \{b\}|D) = 8$, 故 $b \in \text{CORE}_W(C)$. 令 $R = \{b\}$, 有 $W(R|D) = 8$, 并有 $\text{Sig}^+(a, R, D) = \text{Sig}^+(c, R, D) = 2$; 选择 a 加入 R , 有 $W(R|D) = W(C|D) = 10$. 获得约简 $R = \{ab\}$.

对表 1 采用算法 4, 首先求得 $W(a|D) = W(c|D) = 6$, $W(b|D) = 8$, $W(C|D) = 10$, 有 $a = c < b$, 令 $R = \{a, c, b\}$. 因 $W(R - \{a\}|D) = W(C|D) = 10$, 故 a 可以被删除, 有 $R = \{c, b\}$; 此时, $W(R - \{b\}|D) = 6 < W(C|D) = 10$, b 不可删除; 同时 $W(R - \{c\}|D) = 8 < W(C|D) = 10$, c 也不可删除. 最终 $R = \{bc\}$.

4 实验比较

为了说明 4 种差别矩阵约简算法和相对分辨能力约简算法的时空性能, 在 Intel Core i3-3110 2.4 GHz, RAM 3.7 G, Windows 7, Visual Studio.NET 2010 平台的 VC++ 实验环境下, 采用 UCI 数据库中 6 个数据集为测试数据 (基本信息见表 2) 进行对比实验. 表 2 中, $|\text{COR}|$ 表示核属性的数目, Class 表示类别数. 实验结果中, T_Δ 表示算法执行时间 (单位: s), S_Δ 表示算法的空间开销 (单位: 兆), Add 表示添加策略, Del 表示删除策略.

表 2 数据集描述

Data sets	U	C	$ \text{COR} $	Class
Cancer	683	9	1	8
Car	1 728	6	6	4
Gene	3 190	60	0	3
Mushroom	8 124	22	0	2
Poker	20 510	10	4	10
Shuttler	58 000	10	2	7

1) 比较文中 5 种共 10 个约简算法的时间和空间性能, 结果如表 3 和表 4 所示.

表 3 约简算法时间开销比较

Data sets	H-约简		S-约简		B-约简		P-约简		RD-约简	
	Add	Del	Add	Del	Add	Del	Add	Del	Add	Del
Cancer	0.084	0.093	0.053	0.065	0.049	0.057	0.082	0.130	0.010	0.005
Car	0.134	0.099	0.125	0.136	0.098	0.102	0.111	0.264	0.008	0.005
Gene	15.098	64.624	4.331	6.065	3.906	5.381	9.763	57.240	2.087	1.428
Mushroom	Memory Overflow		10.756	14.436	8.943	12.546	15.931	46.830	0.389	0.287
Poker	Memory Overflow		Memory Overflow		23.17	42.782	55.292	81.289	0.184	0.15
Shuttle	Memory Overflow		Memory Overflow		Memory Overflow		362.411	681.401	0.721	0.379

表4 约简算法空间开销比较

Data sets	H-约简	S-约简	B-约简	P-约简	RD-约简
Cancer	10.660	4.524	2.782	0.636	0.624
Car	47.556	21.196	11.935	0.828	0.904
Gene	1239.036	343.328	206.518	3.860	5.516
Mushroom	Memory Overflow	843.764	463.630	4.304	6.152
Poker	Memory Overflow	Memory Overflow	641.569	7.616	9.188
Shuttle	Memory Overflow	Memory Overflow	Memory Overflow	17.000	29.056

由表3可见, 同策略约简算法(添加或删除策略), RD-约简算法的效率明显高于4种差别矩阵约简. 在4种差别矩阵约简算法中, 同策略约简算法的时间开销关系大致是 $T_H \approx T_P > T_S > T_B$. $T_H > T_S$ 是因为H-约简算法中矩阵元素是属性集合, 采用元素比较运算; S-约简算法中矩阵元素是属性数目, 通过加和减运算实现, 而数值存储和运算的代价小于集合存储和运算的代价, 故 $T_H > T_S$. 在B-约简算法中, 虽然布尔差别矩阵元素是0或者1, 但 $T_B < T_S$, 是因为B-约简算法在求解元素运算中, 只要有一个条件属性比较结果不等, 便可获得元素为1, 运算可以提前结束; 而S-约简算法中需要将R(或者C)中所有属性比较运算均执行完, 方可获得元素值. $T_P \approx T_H$ 并且 T_P 要高于 T_S 和 T_B , 因为P-约简算法没有存储差别矩阵, 在约简过程需要多次重新执行布尔差别矩阵元素的生成运算, 造成时间开销增加, 以至于与 T_H 相当, 甚至在处理某些数据集时 T_P 开销更多.

从表4可以发现, 同一数据集空间开销关系基本满足 $S_H > S_S > S_B > S_{RD} > S_P$, 这个结果与性质2基本一致. 具体情况是: H-约简、S-约简和B-约简在约简过程中, 均需要创建相应的差别矩阵, 因此在处理Mushroom、Poker和Shuttle数据集时, H-约简算法由于内存开销过大而造成溢出错误; 在处理Poker和Shuttle数据集时, S-约简算法因内存开销过大而导致溢出错误; 在处理Shuttle数据集时, B-约简算法也因内存开销过大而导致内存溢出. 而P-约简和RD-约简算法无需存储差别矩阵, 因此, 均可以顺利处理这8个数据集, 且空间开销较低. 虽然P-约简算法与RD-约简算法的空间开销相差不大, 并且可以处理较大的数据集Poker和Shuttle, 但从表3可以发现P-约简算法的时间开销较大, RD-约简算法较为高效.

2) 对比相对分辨能力约简算法与其他典型启发式约简算法的时间性能.

采用表2中6个数据集作为实验数据, 将RD-约简的两个算法(RD-约简-Add和RD-约简-Del)与文献[17]约简算法PR和FAPA-PR, 以及文献[18]约简

算法(REPR)进行比较, 结果见表5. 其中Average表示各算法关于6个数据集的平均时间开销.

表5 RD-约简与典型启发式约简算法时间比较

Data Sets	RD-约简		文献[17]		文献[18]
	Add	Del	PR	FAPA-PR	REPR
Cancer	0.010	0.005	0.011	0.008	0.113
Car	0.008	0.005	0.008	0.008	0.104
Gene	2.087	1.428	2.076	1.745	18.087
Mushroom	0.389	0.287	0.391	0.326	2.839
Poker	0.184	0.151	0.185	0.170	1.184
Shuttle	0.721	0.379	0.720	0.692	4.721
Average	0.567	0.376	0.565	0.492	4.508

由表5可见: RD-约简-Add算法的性能与PR算法相当, 略小于FAPA-PR, 但远优于REPR算法, 这是因为在RD-约简和文献[17]的约简算法中等价类划分均采用基数排序方法, 而REPR采用快速排序方法; RD-约简-Del算法效率明显高于RD-约简-Add算法和PR算法, 同时也优于FAPA-PR算法, 这是因为RD-约简-Del算法采用删除策略, 避免求核操作, 节省了大量的时间. 另外, 从平均时间开销Average上可以明显看出, 5个约简算法中, REPR约简算法性能较差, RD-约简-Add算法与PR算法性能相当, FAPA-PR算法性能优于RD-约简-Add算法和PR算法, RD-约简-Del算法效率最高.

5 结论

属性约简是粗糙集理论研究的核心内容之一, 而差别矩阵约简方法是属性约简的一种常用方法, 是从保持知识库分辨能力不变的角度来实现属性约简. 针对差别矩阵含有信息量不同, 本文给出4种差别矩阵描述, 提出H-约简、S-约简、B-约简和P-约简的定义, 揭示了4种约简定义的等价性; 基于上述4种约简定义, 构建了添加和删除两种策略的差别矩阵通用约简模型. 但是, 差别矩阵需要占用较大的内存空间, 而且常因数据集过大引发内存紧张, 甚至导致内存溢出.

为解决这个问题,提高约简求解效率,本文提出了分辨能力相关概念,研究分辨能力相关性质和相对分辨能力计算方法,给出相对分辨能力约简定义,证明了该约简与上述4种差别矩阵约简是一致的,并且设计了快速相对分辨能力约简算法.利用实例分析了文中5类10个约简算法的有效性,并选用UCI数据库中的数据集验证了算法的时空性能.

参考文献(References)

- [1] Pawlak Z. Rough sets[J]. *Int J of Computer and Information Science*, 1982, 11(5): 341-356.
- [2] 苏跃斌, 郭进, 郭瑞. 基于幂图的属性约简[J]. *控制与决策*, 2014, 29(4):743-747.
(Su Y B, Guo J, Guo R. Attribute reduceion based on power graph[J]. *Control and Decision*, 2014, 29(4): 743-747.)
- [3] Liang J Y, Mi J R, Wei W, et al. An accelerator for attribute reduceion based on perspective of objects and attributes[J]. *Knowledge-Based Systems*, 2013, 44: 90-100.
- [4] Shen Q, Jensen R. Selecting informative features with fuzzy-rough sets and its application for complex systems monitoring[J]. *Pattern Recognition*, 2014, 37(7): 1351-1363.
- [5] Miao D Q, Zhao Y, Yao Y Y, et al. Relative reduces in consistent and inconsistent decision tables of the Pawlak rough set model[J]. *Information Sciences*, 2009, 179(24): 4140-4150.
- [6] 杨明. 一种基于改进差别矩阵的属性约简增量式更新算法[J]. *计算机学报*, 2007, 30(5): 815-822.
(Yang M. An incremental updating algorithm for attribute reduction based on improved discernibility matrix[J]. *Chinese J of Computers*, 2007, 30(5): 815-822.)
- [7] Skowron A, Rauszer C. The discernibility matrices and functions in information systems[C]. *Intelligent Decision Support—Handbook of Applications and Advances of the Rough Sets Theory*. Dordrecht: Kluwer Academic Publisher, 1991: 331-362.
- [8] Hu X H, Cercone N. Learning in relational databases: A rough set approach[J]. *Computational Intelligence*, 1995, 11(2): 323-337.
- [9] Kryzkiewicz M. Comparative study of alternative types of knowledge reduction in inconsistent systems[J]. *Int J of Intelligent Systems*, 2001, 16(1): 105-120.
- [10] 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): 1-7.
(Zhang W X, Mi J S, Wu W Z. Knowledge reductions in inconsistent information systems[J]. *Chinese J of Computers*, 2003, 26(1): 1-7.)
- [11] Zhou J, Miao D Q, Pedrycz W, et al. Analysis of alternative objective functions for attribute reduction in complete decision tables[J]. *Soft Computing*, 2011, 15(8): 1601-1616.
- [12] Meng Z Q, Shi Z Z. Extended rough set-based attribute reduction in inconsistent incomplete decision systems[J]. *Information Sciences*, 2012, 204(10): 44-69.
- [13] 蒋云良, 杨章显, 刘勇. 不协调信息系统快速属性分布约简方法[J]. *自动化学报*, 2012, 38(3): 382-388.
(Jiang Y L, Yang Z X, Liu Y. Quick distribution reduction algorithm in inconsistent information system[J]. *Acta Automatica Sinica*, 2012, 38(3): 382-388.)
- [14] Li M, Shang C X, Feng S Z, et al. Quick attribute reduction in inconsistent decision tables[J]. *Information Sciences*, 2014, 254: 155-180.
- [15] Wong S, Ziarko W. On optimal decision rules in decision tables[J]. *The Bulletin of the Polish Academy of Sciences*, 1985, 33(11/12): 693-696.
- [16] Yao Y Y, Zhao Y, Wang J. On reduct construction algorithms[C]. *The 1st Int Conf on Rough Sets and Knowledge Technology*. Chongqing, 2006: 297-304.
- [17] Qian Y H, Liang J Y, Pedrycz W, et al. Positive approximation: An accelerator for attribute reduction in rough set theory[J]. *Artificial Intelligence*, 2010, 174(9/10): 597-618.
- [18] 刘少辉, 盛秋戩, 吴斌, 等. Rough集高效算法的研究[J]. *计算机学报*, 2003, 26(5): 524-529.
(Liu S H, Sheng Q J, Wu B, et al. Research on efficient algorithms for rough set methods[J]. *Chinese J of Computers*, 2003, 26(5): 524-529.)

(责任编辑: 李君玲)