

控制与决策

Control and Decision

区间粗糙数信息系统的覆盖分类冗余度与属性约简

吕跃进, 程林海, 张玉, 何莹莹

引用本文:

吕跃进, 程林海, 张玉, 等. 区间粗糙数信息系统的覆盖分类冗余度与属性约简[J]. *控制与决策*, 2021, 36(3): 677–685.

在线阅读 View online: <https://doi.org/10.13195/j.kzyjc.2019.0744>

您可能感兴趣的其他文章

Articles you may be interested in

[基于矩阵的双论域模糊概率粗糙集增量更新算法](#)

Incremental updating of fuzzy probability rough sets over two universes based on matrix method

控制与决策. 2021, 36(3): 553–564 <https://doi.org/10.13195/j.kzyjc.2019.0692>

[基于知识粒度特征的多目标粗糙集属性约简算法](#)

Multi objective rough set attribute reduction algorithm based on characteristics of knowledge granularity

控制与决策. 2021, 36(1): 196–205 <https://doi.org/10.13195/j.kzyjc.2019.0490>

[考虑时间序列的动态大群体应急决策方法](#)

Dynamic large group emergency decision-making method considering time series

控制与决策. 2020, 35(11): 2609–2618 <https://doi.org/10.13195/j.kzyjc.2019.0088>

[基于社交网络的双知识表达分类方法](#)

Double knowledge representations based classification method from perspective of social networks

控制与决策. 2020, 35(11): 2653–2664 <https://doi.org/10.13195/j.kzyjc.2019.0141>

[基于无标签、不均衡、初值不确定数据的设备健康评估方法](#)

Equipment health risk assessment based on unlabeled, unbalanced data under uncertain initial condition

控制与决策. 2020, 35(11): 2687–2695 <https://doi.org/10.13195/j.kzyjc.2018.1493>

区间粗糙数信息系统的覆盖分类冗余度与属性约简

吕跃进^{1,2†}, 程林海³, 张玉³, 何莹莹³

(1. 广西大学 数学与信息科学学院, 南宁 530004; 2. 广西科技大学 鹿山学院, 广西 柳州 545616; 3. 广西大学 电气工程学院, 南宁 530004)

摘要: 属性约简是知识获取的重要内容, 目前针对区间粗糙数信息系统属性约简的研究十分匮乏. 考虑到现有研究中论域的分类结果存在冗余, 且尚未有研究对分类冗余进行度量, 在区间粗糙数信息系统中提出覆盖分类冗余度的概念, 并且在提出相似类和 β -极大相容类的概念之后, 提出 β -等价类的概念, 使得以 β -等价类对论域进行分类所得的结果为论域的一个划分, 将覆盖分类冗余度降低为 0. 在此基础上, 对区间粗糙数信息系统提出一种保持 β -等价类不变的属性约简定义, 同时根据 β -等价类的特点, 给出仅以 β -等价类中某个元素作为区分对象建立区分矩阵的属性约简方法. 实例分析验证了所提出方法的有效性.

关键词: 区间粗糙数信息系统; 属性约简; β -极大相容类; β -等价类; 覆盖分类冗余度; 区分矩阵

中图分类号: TP18

文献标志码: A

DOI: 10.13195/j.kzyjc.2019.0744

开放科学(资源服务)标识码(OSID):



引用格式: 吕跃进, 程林海, 张玉, 等. 区间粗糙数信息系统的覆盖分类冗余度与属性约简[J]. 控制与决策, 2021, 36(3): 677-685.

Coverage classification redundancy and attribute reduction of interval rough number information system

LYU Yue-jin^{1,2†}, CHENG Lin-hai³, ZHANG Yu³, HE Ying-ying³

(1. College of Mathematics and Information Science, Guangxi University, Nanning 530004, China; 2. Lushan College, Guangxi University of Science and Technology, Liuzhou 545616, China; 3. College of Electrical Engineering, Guangxi University, Nanning 530004, China)

Abstract: Attribute reduction is an important part of knowledge acquisition. At present, the research on attribute reduction of the interval rough number information system is very scarce. Firstly, considering the existence of redundancy in the classification results of existing researches and the fact that the classification redundancy is not yet measured, the concept of coverage classification redundancy is proposed in the interval rough number information system. After similar class and β -maximal consistent class are proposed, the concept of β -equivalent class is proposed, so that the result of classifying the domain by β -equivalent class is a division of the domain, and the coverage classification redundancy decrease to 0. On this basis, an attribute reduction definition which keeps the β -equivalent class invariant is proposed for the interval rough number information system. At the same time, according to the characteristics of the β -equivalent class, an attribute reduction method is given, which only uses an element in the β -equivalent class as a distinguishing object to establish the distinguishing matrices. Finally, examples are given to illustrate the effectiveness of the proposed method.

Keywords: interval rough number information system; attribute reduction; β -maximal consistent class; β -equivalent class; coverage classification redundancy; discernibility matrix

0 引言

区间粗糙数是由 Liu^[1] 于 2002 年结合粗糙集理论与不确定性变量提出的粗糙变量的概念演变而来. 区间粗糙数利用一个下近似区间和一个上近似区间来描述数据的不确定性, 是不确定数据的一种表

现形式. 与区间数、集值、模糊数等不确定数据形式相比, 区间粗糙数在刻画数据的不确定性时能够体现数据不确定性中一定的确定性, 在处理某些问题时更为合适. 目前, 针对区间粗糙数的研究主要集中于区间粗糙数的性质、排序方法和多属性决策方法等方

收稿日期: 2019-05-29; 修回日期: 2019-08-22.

基金项目: 国家自然科学基金项目(71361002).

责任编辑: 王凌.

†通讯作者. E-mail: lvyjin@126.com.

面^[2-5].

属性约简是知识获取中的重要内容. 属性约简的方法主要分为两类: 1) 仅求得一个或多个约简的属性约简方法, 这类方法主要是基于启发式信息进行属性约简^[6-10], 不一定能求得所有约简; 2) 求得所有约简的属性约简方法, 这类方法主要是基于区分矩阵进行属性约简^[11-17]. 对于第1类属性约简方法, 如文献[7]在符号型和数值型属性共存的邻域信息系统中, 定义了邻域组合熵的概念, 进而提出了以邻域组合熵作为启发式信息的属性约简算法; 对于第2类属性约简方法, 如文献[17]在区间值信息系统中, 在相似类的基础上提出极大相容类的概念, 之后提出了一种保持极大相容类不变的属性约简定义, 并给出了基于区分矩阵的属性约简方法. 针对区间粗糙数信息系统属性约简的研究十分匮乏, 只查阅到文献[16]在参考了文献[17]的基础上, 在区间粗糙数信息系统中提出了一种保持极大相容类不变的属性约简定义, 并给出了两种属性约简方法: 1) 以论域中的元素作为区分对象建立区分矩阵进行属性约简; 2) 以基于极大相容类的信息熵作为启发式信息进行属性约简. 在文献[16]中, 虽然以论域中的元素作为区分对象建立区分矩阵进行约简的方法能够求得所有约简, 但是当论域数据量较大时, 区分矩阵的规模也会较大, 导致计算复杂度和空间复杂度较高.

在存在不确定数据的信息系统中, 由于等价关系的成立条件(同时满足自反性、对称性和连续性)较为严格, 难以合理表达对象之间的关系, 众多学者对等价关系进行了拓展并取得了丰富的成果^[18-27]. 这些拓展所得的非等价关系只满足等价关系的部分性质, 如文献[18]在区间值信息系统中定义了一种新的区间相似率, 进而将等价关系拓展为只满足自反性和对称性的相似关系; 文献[19]在集值信息系统中将等价关系拓展为只满足自反性和对称性的相似关系; 文献[20]在直觉模糊信息系统中利用三角模和三角余模将等价关系拓展为只满足自反性和连续性的3种优势关系(强优势关系、弱优势关系和平均优势关系). 在这些非等价关系下对论域进行分类所得的结果通常是论域的一个覆盖, 存在冗余, 而现有研究成果中尚未见到对分类冗余进行度量.

针对上述问题, 本文首先在区间相似度的基础上给出一种新的区间粗糙数的相似度, 进而在区间粗糙数信息系统中提出相似类和 β -极大相容类的概念, 并且在 β -极大相容类的基础上提出 β -等价类的概念,

使得以 β -等价类对论域进行分类所得的结果为论域的一个划分; 同时, 根据对象在论域的分类结果中重复出现的次数, 提出覆盖分类冗余度的概念以度量分类结果的分类冗余, 并论证论域分别以相似类、 β -极大相容类和 β -等价类进行分类所得结果的覆盖分类冗余度之间的关系. 在此基础上, 本文对区间粗糙数信息系统提出一种保持 β -等价类不变的属性约简定义, 并根据 β -等价类中对象之间不可区分的特点, 给出仅以 β -等价类中某个元素作为区分对象建立区分矩阵的属性约简方法. 与以论域中的所有元素作为区分对象建立区分矩阵的属性约简方法相比, 该方法可以有效地减小区分矩阵的规模, 避免不必要的计算. 最后通过例子验证该方法的有效性.

1 基本概念

定义1^[2](区间粗糙数) 一个区间粗糙数是下近似和上近似均为区间的粗糙变量, 记为 $([c, d], [e, g])$. 其中: $e \leq c \leq d \leq g$; $[c, d]$ 为下近似区间, 表示粗糙变量的极有可能取值范围; $[e, g]$ 为上近似区间, 表示粗糙变量的确定取值范围; $[e, c]$ 和 $[d, g]$ 为边界区间, 表示粗糙变量的较小可能取值范围; $[e, c]$ 为下边界区间; $[d, g]$ 为上边界区间. 如决策者在判断某项目的产出时, 用区间粗糙数 $([4, 6], [3, 7])$ 来表示该项目的产出, 则可以解释为该决策者认为这个项目的产出一定介于3到7之间, 并且极有可能在4到6之间, 还有较小可能在3到4或者6到7之间.

定义2^[21](区间的交运算) 给定 E, F 为实数轴上任意的两个闭区间, $E = [l_E, u_E], F = [l_F, u_F]$, 则 E 与 F 的交为

$$E \cap F = \begin{cases} [l_E, u_E], & l_F \leq l_E \leq u_E \leq u_F; \\ [l_F, u_F], & l_E \leq l_F \leq u_F \leq u_E; \\ [l_E, u_F], & l_F \leq l_E \leq u_F \leq u_E; \\ [l_F, u_E], & l_E \leq l_F \leq u_E \leq u_F; \\ \emptyset, & \text{otherwise.} \end{cases} \quad (1)$$

定义3(区间相似度) 给定实数轴上任意的两个闭区间 E, F , 则 E 与 F 的相似度为

$$\mu_{EF} = \frac{|E \cap F|}{\max(|E|, |F|)}. \quad (2)$$

其中: μ_{EF} 表示区间 E 与 F 的相似度, $|\cdot|$ 表示闭区间的区间长度(若“ \cdot ”表示集合, 则 $|\cdot|$ 表示该集合的元素个数), 若 $E = [l_E, u_E]$, 则 $|E| = u_E - l_E$, 并规定空区间的区间长度为0. 显然, $\mu_{EF} \in [0, 1]$; 当且仅当区间 E 与 F 相同时, $\mu_{EF} = 1$; 当区间 E 与 F 相交部分

为单值或空区间时, $\mu_{EF} = 0$.

根据定义1, 区间粗糙数可以理解为是由上近似区间和下近似区间组成的粗糙变量, 也可以理解为由下近似区间和边界区间组成的粗糙变量, 那么两个区间粗糙数的相似程度可以由它们的下近似区间的相似程度和边界区间的相似程度来决定.

定义4(区间粗糙数的相似度) 给定任意两个区间粗糙数 $Q_i = ([c_i, d_i], [e_i, g_i])$ 和 $Q_j = ([c_j, d_j], [e_j, g_j])$, 参数 $\alpha \in (0, 1)$, $\mu_{Q_i Q_j}^l$ 、 $\mu_{Q_i Q_j}^{Bn}$ 分别为 Q_i 与 Q_j 的下近似区间和边界区间的相似度, 则 Q_i 与 Q_j 的相似度 $\gamma_{Q_i Q_j}^\alpha$ 为

$$\begin{cases} \gamma_{Q_i Q_j}^\alpha = \alpha \mu_{Q_i Q_j}^l + (1 - \alpha) \mu_{Q_i Q_j}^{Bn}, \\ \mu_{Q_i Q_j}^l = \frac{|[c_i, d_i] \cap [c_j, d_j]|}{\max(|[c_i, d_i]|, |[c_j, d_j]|)}, \\ \mu_{Q_i Q_j}^{Bn} = \frac{(|[e_i, c_i] \cap [e_j, c_j]| + |[e_i, c_i] \cap [d_j, g_j]| + |[d_i, g_i] \cap [e_j, c_j]| + |[d_i, g_i] \cap [d_j, g_j]|)}{\max(|[e_i, c_i]| + |[d_i, g_i]|, |[e_j, c_j]| + |[d_j, g_j]|)}. \end{cases} \quad (3)$$

显然, $\gamma_{Q_i Q_j}^\alpha \in [0, 1]$: 当且仅当 $\mu_{Q_i Q_j}^l = 1$ 和 $\mu_{Q_i Q_j}^{Bn} = 1$ 同时成立时, $\gamma_{Q_i Q_j}^\alpha = 1$; 当且仅当 $\mu_{Q_i Q_j}^l = 0$ 且 $\mu_{Q_i Q_j}^{Bn} = 0$ 时, $\gamma_{Q_i Q_j}^\alpha = 0$. 参数 α 的引入提高了处理数据的灵活性, α 的具体取值可以按照具体情况来决定, 根据决策者对下近似区间和边界区间的重视程度, α 的取值可分为如下3种情况.

情况1 若下近似区间比边界区间重要, 则 $1/2 < \alpha < 1$, 表示决策者在计算区间粗糙数的相似度时, 更加倚重区间粗糙数在极有可能取值范围上的相似程度.

情况2 若边界区间比下近似区间重要, 则 $0 < \alpha < 1/2$, 表示决策者在计算区间粗糙数的相似度时, 更加倚重区间粗糙数在较小可能取值范围上的相似程度.

情况3 若下近似区间与边界区间重要性相同, 则 $\alpha = 1/2$, 表示决策者在计算区间粗糙数的相似度时, 对区间粗糙数在极有可能取值范围上的相似程度和较小可能取值范围上的相似程度是同等重要的.

2 区间粗糙数信息系统

2.1 区间粗糙数信息系统

定义5^[3](区间粗糙数信息系统) 称 $S = (U, A, V, f)$ 为一个区间粗糙数信息系统. 其中: U 为非空

有限论域; A 为非空有限属性集合; V 为属性值的集合, $V_{ij} = ([c_{ij}, d_{ij}], [e_{ij}, g_{ij}])$ 表示第 i 个对象在第 j 个属性下的取值; $f : U \times A \rightarrow V$ 为信息函数, 表示 $x_i \in U, a_j \in A, f(x_i, a_j) = ([c_{ij}, d_{ij}], [e_{ij}, g_{ij}]) \in V$.

2.2 β -相似关系和相似类

定义6(β -相似关系和相似类) 给定区间粗糙数信息系统 $S = (U, A, V, f)$, 属性子集 $B (B \subseteq A)$, 参数 α, β , 则 β -相似关系 R_B^β 和相似类 $[x_i]_B^\beta$ 为

$$R_B^\beta = \{(x_i, x_j) \in U \times U | \forall b \in B, \gamma_{x_i^b x_j^b}^\alpha \geq \beta\}, \quad (4)$$

$$[x_i]_B^\beta = \{x_j \in U | (x_i, x_j) \in R_B^\beta\}. \quad (5)$$

其中: $\alpha \in (0, 1), \beta \in (0, 1], \gamma_{x_i^b x_j^b}^\alpha$ 表示 x_i 与 x_j 在属性 b 下的取值的相似度. 参数 β 的引入增强了信息处理的灵活性, β 的取值大小体现了决策者对对象之间满足相似关系的严格程度. β 越大, 则对象之间满足相似关系要求越严格; 反之, 则对象之间满足相似关系要求越宽松. 显然, R_B^β 满足自反性和对称性, 不满足传递性. 令 $G_B^\beta = \{[x]_B^\beta | x \in U\}$, 则 G_B^β 表示论域 U 基于相似类的分类结果, 且 G_B^β 是论域 U 上的一个覆盖.

2.3 β -极大相容类

在区间值信息系统中, 因为以相似类作为分类依据所得分类结果的分类精度不高, 存在冗余^[17]. 因此, 文献[17]在区间值信息系统中提出了 α -极大相容类的概念, 并将其作为论域的分类依据. 同样地, 在区间粗糙数信息系统中, 以相似类作为分类依据所得的分类结果存在同样的问题^[16]. 因此, 文献[16]将文献[17]中的 α -极大相容类的概念推广到了区间粗糙数信息系统中. 本文在文献[16]的基础上, 结合定义6给出 β -极大相容类的概念如下.

定义7(β -极大相容类) 给定区间粗糙数信息系统 $S = (U, A, V, f)$, 属性子集 $B (B \subseteq A)$, 参数 $\alpha, \beta, \alpha \in (0, 1), \beta \in (0, 1]$. R_B^β 为 β -相似关系, 若 $M_B^\beta \subseteq U$, 对于任意 $x, y \in M_B^\beta$ 均满足 $x, y \in R_B^\beta$, 且对 $\forall z \in U - M_B^\beta$, 一定存在 $z' \in M_B^\beta$ 满足 $(z, z') \notin R_B^\beta$, 则称 M_B^β 为区间粗糙数信息系统 S 关于 R_B^β 的一个 β -极大相容类.

定义7说明 M_B^β 是 U 中对象满足两两相似的最大集合. 因为包含 x 的 β -极大相容类可能存在多个, 所以区间粗糙数信息系统 S 在 R_B^β 下的 β -极大相容类的集合 $\xi_B^\beta(U) = \bigcup \xi_B^\beta(x_i)$ 是 U 的一个覆盖, 其中 $\xi_B^\beta(x_i) = \{M_B^\beta | x_i \in M_B^\beta\}$ 表示包含对象 x_i 的 β -极大相容类的集合, 因此 $\xi_B^\beta(U)$ 也表示论域 U 基于 β -极大相容类的分类结果.

定理1 (β -极大相容类和相似类的关系) 给定区间粗糙数信息系统 $S = (U, A, V, f)$, 属性子集 $B (B \subseteq A)$, 在 β -相似关系 R_B^β 下, 对于 $\forall x_i \in U$ 有

$$[x_i]_B^\beta = \bigcup_{M_B^\beta \in \xi_B^\beta(x_i)} M_B^\beta. \quad (6)$$

证明 1) 由相似类和 β -极大相容类的定义易知, 对于 $\forall x_j \in [x_i]_B^\beta, x_j \in \bigcup_{M_B^\beta \in \xi_B^\beta(x_i)} M_B^\beta$, 即 $[x_i]_B^\beta \subseteq$

$$\bigcup_{M_B^\beta \in \xi_B^\beta(x_i)} M_B^\beta;$$

2) 对于 $\forall x_j \in M_B^\beta$, 因为 $M_B^\beta \in \xi_B^\beta(x_i)$, 所以 $x_i \in M_B^\beta$, 根据 β -极大相容类的定义可得 $(x_i, x_j) \in R_B^\beta$, 所以 $x_j \in [x_i]_B^\beta$, 因此 $\bigcup_{M_B^\beta \in \xi_B^\beta(x_i)} M_B^\beta \subseteq [x_i]_B^\beta$.

由1)和2)可得, $[x_i]_B^\beta = \bigcup_{M_B^\beta \in \xi_B^\beta(x_i)} M_B^\beta$. \square

定理1说明 β -极大相容类是相似类的细化, 任意一个 β -极大相容类都可以扩展为相似类, 对于 $\forall [x]_B^\beta \in G_B^\beta$, 总有一个 $\xi_B^\beta(x) \subseteq \xi_B^\beta(U)$ 与其对应.

2.4 β -等价类

由于基于 β -极大相容类对论域进行分类所得的分类结果仍然是论域的一个覆盖, 存在冗余, 所以本文在 β -极大相容类的基础上提出了 β -等价类的概念如下.

定义8 (β -等价关系和 β -等价类) 给定区间粗糙数信息系统 $S = (U, A, V, f)$, 属性子集 $B (B \subseteq A)$, 参数 $\alpha, \beta, \alpha \in (0, 1), \beta \in (0, 1]$, 在 β -相似关系 R_B^β 下, 对于 $\forall x_i \in U, \xi_B^\beta(x_i)$ 为包含对象 x_i 的所有 β -极大相容类的集合, 那么 β -等价关系 N_B^β 和 x_i 的 β -等价类 $N_B^\beta(x_i)$ 为

$$N_B^\beta = \{(x_i, x_j) \in U \times U | \forall b \in B, \xi_b^\beta(x_j) = \xi_b^\beta(x_i)\}, \quad (7)$$

$$N_B^\beta(x_i) = \{x_j \in U | \forall b \in B, (x_i, x_j) \in N_B^\beta\}. \quad (8)$$

显然, β -等价关系 N_B^β 满足自反性、对称性和传递性.

定理2 所有 β -等价类 $N_B^\beta(x_i)$ 的集合构成了论域 U 上的一个划分. 给定区间粗糙数信息系统 $S = (U, A, V, f)$, 属性子集 $B (B \subseteq A)$, 令论域 U 基于 β -等价类的分类结果为 $N_B^\beta(U) = \{N_B^\beta(x_i) | x_i \in U\}$, 则 $N_B^\beta(U)$ 构成了论域 U 上的一个划分.

证明 1) 根据 β -等价类的定义, 必定有 $x_i \in N_B^\beta(x_i)$, 所以 $\bigcup_{x_i \in U} N_B^\beta(x_i) = U$;

2) 假设存在两个不同的 β -等价类 $N_B^\beta(x_i)$ 和 $N_B^\beta(x_j)$ 满足 $N_B^\beta(x_i) \cap N_B^\beta(x_j) \neq \emptyset$, 令 $N_B^\beta(x_i) \cap$

$N_B^\beta(x_j) \neq \{x_m\}$, 根据 β -等价类的定义可知, 对于 $\forall b \in B, \forall x \in N_B^\beta(x_i), \forall y \in N_B^\beta(x_j)$, 都有 $\xi_b^\beta(x) = \xi_b^\beta(x_m) = \xi_b^\beta(y)$, 所以 $N_B^\beta(x_i) = N_B^\beta(x_j)$, 这与假设矛盾. 因此, 对于 $N_B^\beta(U)$ 中的任意两个 β -等价类 $N_B^\beta(x_i)$ 和 $N_B^\beta(x_j)$, 均满足 $N_B^\beta(x_i) \cap N_B^\beta(x_j) = \emptyset$.

由1)和2)可得, $N_B^\beta(U) = \{N_B^\beta(x_i) | x_i \in U\}$ 构成了论域 U 上的一个划分. \square

定理2说明通过定义 β -等价类, 实现了对论域 U 的分类由覆盖到划分的转变.

定理3 (β -等价类与 β -极大相容类的关系) 给定区间粗糙数信息系统 $S = (U, A, V, f)$, 属性子集 $B (B \subseteq A)$, 则对于 $\forall M_B^\beta \in \xi_B^\beta(U)$, 有

$$M_B^\beta = \bigcup_{x \in M_B^\beta} N_B^\beta(x). \quad (9)$$

证明 1) 对于 $\forall x \in M_B^\beta$, 均满足 $N_B^\beta(x) \subseteq \bigcup_{x \in M_B^\beta} N_B^\beta(x)$, 因为 $x \in N_B^\beta(x)$, 所以 $x \in \bigcup_{x \in M_B^\beta} N_B^\beta(x)$, 即 $M_B^\beta \subseteq \bigcup_{x \in M_B^\beta} N_B^\beta(x)$;

2) 对于 $\forall x' \in N_B^\beta(x)$, 均满足对 $\forall b \in B, \xi_b^\beta(x') = \xi_b^\beta(x)$, 所以 $\xi_b^\beta(x') = \xi_b^\beta(x)$, 若 $x \in M_B^\beta$, 则 $M_B^\beta \in \xi_B^\beta(x)$, 所以 $M_B^\beta \in \xi_B^\beta(x')$, 因此 $x' \in M_B^\beta$, 即 $\bigcup_{x \in M_B^\beta} N_B^\beta(x) \subseteq M_B^\beta$.

由1)和2)可得, $M_B^\beta = \bigcup_{x \in M_B^\beta} N_B^\beta(x)$ 成立. \square

定理3说明 β -等价类是 β -极大相容类的细化, 任意一个 β -等价类都可以扩展为 β -极大相容类, 任意一个 β -极大相容类都可以分解为若干个 β -等价类.

3 覆盖分类冗余度

基于相似关系对信息系统中的对象进行分类所得的分类结果通常是论域的一个覆盖, 存在冗余. 为了度量分类结果的冗余程度, 本文提出覆盖分类冗余度的概念.

定义9 (覆盖分类冗余度) 给定非空有限论域 U, T 为论域 U 的分类结果 (覆盖), $t \in T$, 则论域 U 的覆盖分类冗余度 $L(T)$ 为

$$L(T) = \frac{\sum_{t \in T} |t|}{|U|} - 1, \quad (10)$$

其中 $|\cdot|$ 表示集合的元素个数. 覆盖分类冗余度的概念表示论域 U 中的对象在基于 t 所得的分类结果 T 中重复出现的总次数与 U 中的对象个数的比例. 显然, $L(T) \geq 0$, 当且仅当 T 为 U 的一个划分时, $L(T) = 0$. 在给定 U 的情况下, 覆盖分类冗余度越大, 说明 U

中的对象在分类结果 T 中重复出现的总次数越大;反之,说明 U 中的对象在分类结果 T 中重复出现的总次数越小.

根据定义9,给定区间粗糙数信息系统 $S = (U, A, V, f)$,属性子集 $B(B \subseteq A)$,参数 $\alpha, \beta, \alpha \in (0, 1), \beta \in (0, 1]$,在 β -相似关系 R_B^β 下,论域 U 基于相似类、 β -极大相容类和 β -等价类的覆盖分类冗余度分别为

$$L(G_B^\beta) = 0 = \frac{\sum_{[x]_B^\beta \in G_B^\beta} |[x]_B^\beta|}{|U|} - 1, \quad (11)$$

$$L(\xi_B^\beta(U)) = 0 = \frac{\sum_{M_B^\beta \in \xi_B^\beta(U)} |M_B^\beta|}{|U|} - 1, \quad (12)$$

$$L(N_B^\beta(U)) = 0 = \frac{\sum_{N_B^\beta(x) \in N_B^\beta(U)} |N_B^\beta(x)|}{|U|} - 1. \quad (13)$$

定理4(3种分类方法的覆盖分类冗余度对比) 给定区间粗糙数信息系统 $S = (U, A, V, f)$,属性子集 $B(B \subseteq A)$,论域 U 基于相似类、 β -极大相容类和 β -等价类的覆盖分类冗余度满足

$$L(G_B^\beta) \geq L(\xi_B^\beta(U)) \geq L(N_B^\beta(U)) = 0. \quad (14)$$

证明 1) 根据定理1可知,对于 $\forall [x]_B^\beta \in G_B^\beta$,有 $[x]_B^\beta = \bigcup_{M_B^\beta \in \xi_B^\beta(U) \wedge x \in M_B^\beta} M_B^\beta$,即 G_B^β 中任意一个相似类都可以由一个或者多个不同的 β -极大相容类合并而得. 假设 $\xi_B^\beta(U)$ 中包含 x 的 M_B^β 有 $p(x)$ 个, G_B^β 中包含 x 的相似类有 $q(x)$ 个,下面分两种情况讨论:

① 若 $\forall x \in U$ 满足 $p(x) = 1$,则 $[x]_B^\beta = M_B^\beta(x)$,所以 $p(x) = q(x) = 1$,因此 $L(G_B^\beta) = L(\xi_B^\beta(U)) = 0$.

② 若 $\exists x \in U$ 满足 $p(x) \geq 2$,根据 β -极大相容类的定义可知,对于 $\forall M, M' \in \xi_B^\beta(U)$,若 $M \neq M'$,则 M 与 M' 互不包含,即 $\xi_B^\beta(U)$ 中任意两个 M_B^β 不存在包含关系,所以 $\xi_B^\beta(U)$ 中任意两个包含 x 的 M_B^β (假设为 M_1 和 M_2) 满足: $\exists x_1 \in M_1 \wedge x_1 \notin M_2, \exists x_2 \in M_2 \wedge x_2 \notin M_1$. 因此 $[x]_B^\beta, [x_1]_B^\beta$ 和 $[x_2]_B^\beta$ 两两互不相同,但它们都属于 G_B^β 且都包含 x ,由此可知 $p(x) < q(x)$,所以 $L(G_B^\beta) > L(\xi_B^\beta(U))$.

由①和②得 $L(G_B^\beta) \geq L(\xi_B^\beta(U))$.

2) 因为 $\xi_B^\beta(U)$ 是论域 U 的覆盖,而 $N_B^\beta(U)$ 是论域 U 的划分,所以 $\sum_{M_B^\beta \in \xi_B^\beta(U)} |M_B^\beta| \geq \sum_{N_B^\beta(x) \in N_B^\beta(U)} |N_B^\beta(x)| = |U|$,因此 $L(\xi_B^\beta(U)) \geq L(N_B^\beta(U)) = 0$.

由1)和2)可得, $L(G_B^\beta) \geq L(\xi_B^\beta(U)) \geq L(N_B^\beta(U)) = 0$. \square

定理4说明相比于根据相似类对论域进行分类,根据 β -极大相容类进行分类可以获得更低的覆盖分类冗余度,并且根据 β -等价类进行分类又使覆盖分类冗余度进一步降低为0,使得分类不存在冗余.

4 属性约简

文献[16]在区间粗糙数信息系统中提出了一种保持极大相容类不变的属性约简定义,极大相容类是基于相似类给出的概念,而本文在 β -极大相容类的基础上,提出了 β -等价类的概念,并在第3节说明了根据 β -等价类进行分类可以使覆盖分类冗余度降低为0,使得分类不存在冗余. 因此,本文在区间粗糙数信息系统中,新提出一种基于保持 β -等价类不变的属性约简定义如下.

定义10(属性约简) 给定区间粗糙数信息系统 $S = (U, A, V, f)$,属性子集 $B(B \subseteq A)$,参数 $\alpha, \beta, \alpha \in (0, 1), \beta \in (0, 1]$. 在 β -相似关系 R_B^β 下,若 $\forall x \in U$ 满足 $N_B^\beta(x) = N_A^\beta(x)$,并且对于 $\forall C \subset B, \forall x \in U$ 满足 $N_C^\beta(x) \neq N_A^\beta(x)$,则称属性子集 B 为区间粗糙数信息系统 S 基于保持 β -等价类不变的一个属性约简,令 red^β 为 S 的属性约简集合,则称 $\text{cor}^\beta = \bigcap_{B \in \text{red}^\beta} B$ 为 A 的核.

根据 β -等价类的定义可知,对于 $\forall x, y \in U$,若 $x, y \in N_B^\beta(x)$,则 x 和 y 在属性集 B 下是不可区分的,即对于 $\forall b \in B$,有 $\xi_b^\beta(x) = \xi_b^\beta(y)$;对于 $\forall z \in U - N_B^\beta(x), \forall b \in B$,若属性 b 能区分 x 和 z ,即 $\xi_b^\beta(x) \neq \xi_b^\beta(z)$,则属性 b 也一定能区分 y 和 z . 因此,在区分 U 中的对象时,只需对 β -等价类中的某个元素进行区分即可. 由此,保持 β -等价类不变的属性约简对应的区分矩阵 Dis^β 及其区分函数 Δ^β 可以定义如下.

定义11(区分矩阵及其区分函数) 给定区间粗糙数信息系统 $S = (U, A, V, f)$,参数 $\alpha, \beta, \alpha \in (0, 1), \beta \in (0, 1]$,在 β -相似关系 R_B^β 下,对于 $\forall (N_A^\beta(x_i), N_A^\beta(x_j)) \in N_A^\beta(U) \times N_A^\beta(U)$,保持 β -等价类不变的属性约简对应的区分矩阵 Dis^β 及其区分函数 Δ^β 分别为

$$\text{Dis}^\beta(x_i, x_j) = \{a \in A \mid \xi_a^\beta(x_i) \neq \xi_a^\beta(x_j)\}, \quad (15)$$

$$\Delta^\beta = \bigwedge_{\forall (N_A^\beta(x_i), N_A^\beta(x_j)) \in N_A^\beta(U) \times N_A^\beta(U)} \bigvee \text{Dis}^\beta(x_i, x_j). \quad (16)$$

定理5(区分函数的性质) 根据文献[29]对区分

函数的论述可知,区分函数有如下性质:给定区间粗糙数信息系统 $S = (U, A, V, f)$, 区分函数 Δ^β 满足 Δ^β 的极小析取范式中的所有合取式是属性集 A 的所有约简.

定理5说明约简是能区别由整个属性集区别的所有对象的属性极小子集. 易知,对于 $\forall(x_i, x_j) \in U \times U$, 若 $B \subseteq A$ 是满足条件 $B \cap \text{Dis}(x_i, x_j) \neq \emptyset$ ($\text{Dis}(x_i, x_j)$ 表示能区分对象 x_i 和 x_j 的所有属性的集合,且 $\forall \text{Dis}(x_i, x_j) \neq \emptyset$) 的极小子集,则 B 是 A 的一个约简.

在区间粗糙数信息系统 $S = (U, A, V, f)$ 中,给定

参数 α, β , 则保持 β -等价类不变的约简步骤如下:

step 1: 求出 $G_{a_i}^\beta, \xi_{a_i}^\beta(U)$ 和 $N_A^\beta(U)$;

step 2: 根据定义11建立区分矩阵 Dis^β , 并列区分函数 Δ^β ;

step 3: 计算区分函数 Δ^β 的极小析取范式, 根据定理5得出属性集 A 的所有约简.

5 实验分析

一个区间粗糙数信息系统如表1所示. 其中: 论域 $U = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7\}$; 属性集 $A = \{a_1, a_2, a_3, a_4\}$; 属性值 $f(x_i, a_j) = ([c_{ij}, d_{ij}], [e_{ij}, g_{ij}])$ 是区间粗糙数, 如 $f(x_1, a_1) = ([4, 8], [2, 10])$.

表1 区间粗糙数信息系统

	a_1	a_2	a_3	a_4
x_1	([4,8], [2,10])	([8,12], [5,15])	([12,17], [10,20])	([17,22], [15,25])
x_2	([5,8], [1,9])	([7,11], [7,14])	([12,16], [12,21])	([17,21], [16,24])
x_3	([7,10], [5,12])	([8,11], [5,12])	([13,17], [9,17])	([18,22], [14,22])
x_4	([7,11], [6,13])	([7,11], [6,12])	([13,16], [11,18])	([18,21], [16,23])
x_5	([10,14], [7,15])	([11,15], [9,17])	([16,20], [13,23])	([21,25], [18,28])
x_6	([11,15], [9,16])	([11,14], [8,18])	([16,19], [16,22])	([22,26], [20,27])
x_7	([11,14], [8,17])	([13,16], [8,17])	([18,21], [14,23])	([22,25], [20,26])

本文通过3个例子来分析说明覆盖分类冗余度的计算和论域基于相似类、 β -极大相容类和 β -等价类的覆盖分类冗余度三者的大小关系, 并且验证本文提出的属性约简方法的有效性.

例1 在表1所示的区间粗糙数信息系统中, 假设 $\alpha = 0.8, \beta = 0.5$, 由 β -相似关系的定义, U 中各元素的相似类 $[x_i]_A^{0.5}$ 分别为

$$\begin{aligned}
 [x_1]_A^{0.5} &= [x_2]_A^{0.5} = \{x_1, x_2\}, \\
 [x_3]_A^{0.5} &= [x_4]_A^{0.5} = \{x_3, x_4\}, \\
 [x_5]_A^{0.5} &= \{x_5, x_6, x_7\}, \\
 [x_6]_A^{0.5} &= \{x_5, x_6\}, \\
 [x_7]_A^{0.5} &= \{x_5, x_7\}.
 \end{aligned}$$

因为 $G_A^{0.5} = \{[x]_A^{0.5} | x \in U\}$, 所以 $G_A^{0.5} = \{\{x_1, x_2\}, \{x_3, x_4\}, \{x_5, x_6, x_7\}, \{x_5, x_6\}, \{x_5, x_7\}\}$.

根据 β -极大相容类的定义, 包含对象 x_i 的 β -极大相容类 $M_A^{0.5}(x_i)$ 分别为

$$\begin{aligned}
 M_A^{0.5}(x_1) &= M_A^{0.5}(x_2) = \{x_1, x_2\}, \\
 M_A^{0.5}(x_3) &= M_A^{0.5}(x_4) = \{x_3, x_4\}, \\
 M_A^{0.5}(x_5)_1 &= M_A^{0.5}(x_6) = \{x_5, x_6\}, \\
 M_A^{0.5}(x_5)_2 &= M_A^{0.5}(x_7) = \{x_5, x_7\}.
 \end{aligned}$$

因此, 包含对象 x_i 的 β -极大相容类的集合 $\xi_A^{0.5}(x_i)$ 分别为

$$\begin{aligned}
 \xi_A^{0.5}(x_1) &= \{M_A^{0.5}(x_1) = \{x_1, x_2\}\}, \\
 \xi_A^{0.5}(x_2) &= \{M_A^{0.5}(x_2) = \{x_1, x_2\}\}, \\
 \xi_A^{0.5}(x_3) &= \{M_A^{0.5}(x_3) = \{x_3, x_4\}\}, \\
 \xi_A^{0.5}(x_4) &= \{M_A^{0.5}(x_4) = \{x_3, x_4\}\}, \\
 \xi_A^{0.5}(x_5) &= \\
 &\{M_A^{0.5}(x_5)_1 = \{x_5, x_6\}, M_A^{0.5}(x_5)_2 = \{x_5, x_7\}\}, \\
 \xi_A^{0.5}(x_6) &= \{M_A^{0.5}(x_6) = \{x_5, x_6\}\}, \\
 \xi_A^{0.5}(x_7) &= \{M_A^{0.5}(x_7) = \{x_5, x_7\}\}.
 \end{aligned}$$

因为 $\xi_B^\beta(U) = \bigcup \xi_B^\beta(x_i)$, 所以 $\xi_A^{0.5}(U)$ 为

$$\xi_A^{0.5}(U) = \{\{x_1, x_2\}, \{x_3, x_4\}, \{x_5, x_6\}, \{x_5, x_7\}\}.$$

根据 β -等价类的定义, 各对象的 β -等价类 $N_A^{0.5}(x_i)$ 分别为

$$\begin{aligned}
 N_A^{0.5}(x_1) &= N_A^{0.5}(x_2) = \{x_1, x_2\}, \\
 N_A^{0.5}(x_3) &= N_A^{0.5}(x_4) = \{x_3, x_4\}, \\
 N_A^{0.5}(x_5) &= \{x_5\}, \\
 N_A^{0.5}(x_6) &= \{x_6\},
 \end{aligned}$$

$$N_A^{0.5}(x_7) = \{x_7\}.$$

因此

$$N_A^{0.5}(U) = \{\{x_1, x_2\}, \{x_3, x_4\}, \{x_5\}, \{x_6\}, \{x_7\}\}.$$

根据上述计算结果和覆盖分类冗余度的定义,有

$$\begin{aligned} L(G_A^{0.5}) &= \frac{11}{7} - 1 = \frac{4}{7}, \\ L(\xi_A^{0.5}(U)) &= \frac{8}{7} - 1 = \frac{1}{7}, \\ L(N_A^{0.5}(U)) &= \frac{7}{7} - 1 = 0. \end{aligned}$$

因此

$$L(G_A^{0.5}) \geq L(\xi_A^{0.5}(U)) \geq L(N_A^{0.5}(U)) = 0.$$

下面用本文所提的属性约简方法对表1所示的区间粗糙数信息系统进行属性约简.

首先,计算 $G_{a_i}^{0.5}$,得

$$G_{a_1}^{0.5} = \{\{x_1, x_2\}, \{x_3, x_4\}, \{x_5, x_6, x_7\}\},$$

$$G_{a_2}^{0.5} = G_{a_3}^{0.5} =$$

$$\{\{x_1, x_2, x_3, x_4\}, \{x_5, x_6, x_7\}, \{x_5, x_6\}, \{x_5, x_7\}\},$$

$$G_{a_4}^{0.5} = \{\{x_1, x_2, x_3, x_4\}, \{x_5, x_6, x_7\}\}.$$

根据 $G_{a_i}^{0.5}$, 计算 $\xi_{a_i}^{0.5}(U)$,得

$$\xi_{a_1}^{0.5}(U) = \{\{x_1, x_2\}, \{x_3, x_4\}, \{x_5, x_6, x_7\}\},$$

$$\xi_{a_2}^{0.5}(U) = \xi_{a_3}^{0.5}(U) =$$

$$\{\{x_1, x_2, x_3, x_4\}, \{x_5, x_6\}, \{x_5, x_7\}\},$$

$$\xi_{a_4}^{0.5}(U) = \{\{x_1, x_2, x_3, x_4\}, \{x_5, x_6, x_7\}\}.$$

根据定义11建立区分矩阵 $\text{Dis}^{0.5}$ 如表2所示. 因为区分矩阵 $\text{Dis}^{0.5}$ 是对称的,所以表中只给出区分矩阵的一半. 表2中的数字表示由对应的属性组成的属性集合,如“2,3”表示 $\{a_2, a_3\}$.

表2 区分矩阵 $\text{Dis}^{0.5} (\alpha = 0.8)$

	x_1	x_3	x_5	x_6	x_7
x_1	\emptyset				
x_3	1	\emptyset			
x_5	1,2,3,4	1,2,3,4	\emptyset		
x_6	1,2,3,4	1,2,3,4	2,3	\emptyset	
x_7	1,2,3,4	1,2,3,4	2,3	2,3	\emptyset

由表2得区分函数 $\Delta^{0.5}$ 为

$$\begin{aligned} \Delta^{0.5} &= a_1 \wedge (a_1 \vee a_2 \vee a_3 \vee a_4) \wedge (a_2 \vee a_3) = \\ &(a_1 \wedge a_2) \vee (a_1 \wedge a_3). \end{aligned}$$

因此 $\text{red}^{0.5} = \{\{a_1, a_2\}, \{a_1, a_3\}\}, \text{cor}^{0.5} = \{a_1\}$.

例2 在表1所示的区间粗糙数信息系统中,假设 $\alpha = 0.8, \beta = 0.6$,与例1类似,通过计算可得以下

结果:

$$G_A^{0.6} = \{\{x_1, x_2\}, \{x_3, x_4\}, \{x_5, x_6\}, \{x_7\}\},$$

$$\xi_A^{0.6}(U) = \{\{x_1, x_2\}, \{x_3, x_4\}, \{x_5, x_6\}, \{x_7\}\},$$

$$N_A^{0.6}(U) = \{\{x_1, x_2\}, \{x_3, x_4\}, \{x_5, x_6\}, \{x_7\}\}.$$

所以 $L(G_A^{0.6}) = L(\xi_A^{0.6}(U)) = L(N_A^{0.6}(U)) = \frac{7}{7} - 1 = 0$,满足 $L(G_A^{0.6}) \geq L(\xi_A^{0.6}(U)) \geq L(N_A^{0.6}(U)) = 0$.

下面用本文所提的属性约简方法对表1所示的区间粗糙数信息系统进行属性约简.

与例1类似,根据定义11建立区分矩阵 $\text{Dis}^{0.6}$ 如表3所示.

表3 区分矩阵 $\text{Dis}^{0.6}$

	x_1	x_3	x_5	x_7
x_1	\emptyset			
x_3	1	\emptyset		
x_5	1,2,3,4	1,2,3,4	\emptyset	
x_7	1,2,3,4	1,2,3,4	2,3	\emptyset

由表3得区分函数 $\Delta^{0.6}$ 为

$$\begin{aligned} \Delta^{0.6} &= a_1 \wedge (a_1 \vee a_2 \vee a_3 \vee a_4) \wedge (a_2 \vee a_3) = \\ &(a_1 \wedge a_2) \vee (a_1 \wedge a_3). \end{aligned}$$

所以 $\text{red}^{0.6} = \{\{a_1, a_2\}, \{a_1, a_3\}\}, \text{cor}^{0.6} = \{a_1\}$.

例3 在表1所示的区间粗糙数信息系统中,假设 $\alpha = 0.2, \beta = 0.5$,与例1类似,通过计算可得以下结果:

$$G_A^{0.5} = \{\{x_1\}, \{x_2\}, \{x_3, x_4\}, \{x_5\}, \{x_6\}, \{x_7\}\},$$

$$\xi_A^{0.5}(U) = \{\{x_1\}, \{x_2\}, \{x_3, x_4\}, \{x_5\}, \{x_6\}, \{x_7\}\},$$

$$N_A^{0.5}(U) = \{\{x_1\}, \{x_2\}, \{x_3, x_4\}, \{x_5\}, \{x_6\}, \{x_7\}\}.$$

所以 $L(G_A^{0.5}) = L(\xi_A^{0.5}(U)) = L(N_A^{0.5}(U)) = \frac{7}{7} - 1 = 0$,满足 $L(G_A^{0.5}) \geq L(\xi_A^{0.5}(U)) \geq L(N_A^{0.5}(U)) = 0$.

下面用本文所提的属性约简方法对表1所示的区间粗糙数信息系统进行属性约简.

与例1类似,根据定义11建立区分矩阵 $\text{Dis}^{0.5}$ 如表4所示.

表4 区分矩阵 $\text{Dis}^{0.5} (\alpha = 0.2)$

	x_1	x_2	x_3	x_5	x_6	x_7
x_1	\emptyset					
x_2	2	\emptyset				
x_3	1,3,4	1,2,3,4	\emptyset			
x_5	1,2,3,4	1,2,3,4	1,2,3,4	\emptyset		
x_6	1,2,3,4	1,2,3,4	1,2,3,4	1,3,4	\emptyset	
x_7	1,2,3,4	1,2,3,4	1,2,3,4	1,4	1,3	\emptyset

由表4得区分函数 $\Delta^{0.5}$ 为

$$\begin{aligned} \Delta^{0.5} = & a_2 \wedge (a_1 \vee a_3 \vee a_4) \wedge (a_1 \vee a_2 \vee \\ & a_3 \vee a_4) \wedge (a_1 \vee a_4) \wedge (a_1 \vee a_3) = \\ & (a_1 \wedge a_2) \vee (a_2 \wedge a_3 \wedge a_4). \end{aligned}$$

所以 $\text{red}^{0.5} = \{\{a_1, a_2\}, \{a_2, a_3, a_4\}\}$, $\text{cor}^{0.5} = \{a_2\}$.

例1~例3不仅从实例角度说明了覆盖分类冗余度的计算方法和定理4的正确性,而且经检验,例1~例3所求得属性约简结果都是表1所示的区间粗糙数信息系统在相应条件下的属性约简,从而验证了本文提出的属性约简方法的有效性.同时,通过表2~表4可以看出,与以论域中所有元素作为区分对象建立区分矩阵的属性约简方法相比,本文所提的属性约简方法可以有效减小区分矩阵的规模,避免不必要的计算.

6 结论

本文在区间粗糙数信息系统中提出了覆盖分类冗余度和 β -等价类的概念,将论域以 β -等价类进行分类所得结果的覆盖分类冗余度降低为0.本文对区间粗糙数信息系统提出了一种保持 β -等价类不变的属性约简定义,并给出了基于区分矩阵的属性约简方法.通过例子验证了该方法的有效性,为区间粗糙数信息系统的属性约简提供了一条新的思路.下一步的工作将探讨在区间粗糙数信息系统中快速求取相似类、 β -极大相容类和 β -等价类的方法,以期能对本文提出的属性约简方法做出改进.

参考文献(References)

- [1] Liu B D. Theory and practice of uncertain programming[M]. Heidelberg: Physica-Verlag, 2002: 111-128.
- [2] 曾玲, 曾祥艳. 一类区间粗糙数型多属性决策方法研究[J]. 控制与决策, 2010, 25(11): 1757-1760.
(Zeng L, Zeng Y X. Research on a class of multiple attribute decision making method with interval rough numbers[J]. Control and Decision, 2010, 25(11): 1757-1760.)
- [3] 翁世洲, 吕跃进. 区间粗糙数的排序方法及其应用[J]. 南京大学学报: 自然科学, 2015, 51(4): 818-825.
(Weng S Z, Lv Y J. Sorting method with interval rough number and its application[J]. Journal of Nanjing University: Natural Sciences, 2015, 51(4): 818-825.)
- [4] 田泽金, 黄锐露, 吕跃进. 区间粗糙数互补判断矩阵的两种一致性[J]. 计算机工程与应用, 2019(24): 235-240.
(Tian Z J, Huang R L, Lv Y J. Two consistencies of complementary judgment matrix based on interval rough numbers[J]. Computer Engineering and Applications, 2019(24): 235-240.)
- [5] 吕跃进, 杨燕华. 区间粗糙数层次分析法[J]. 系统工程理论与实践, 2018, 38(3): 786-793.
(Lv Y J, Yang Y H. Analytic hierarchy process based on interval rough number[J]. System Engineering — Theory and Practice, 2018, 38(3): 786-793.)
- [6] 姚晟, 徐风, 吴照玉, 等. 基于邻域粗糙互信息熵的非单调性属性约简[J]. 控制与决策, 2019, 34(2): 353-361.
(Yao S, Xu F, Wu Z Y, et al. Non-monotonic attribute reduction based on neighborhood rough mutual information entropy[J]. Control and Decision, 2019, 34(2): 353-361.)
- [7] 王光琼. 基于邻域组合熵的属性约简算法[J]. 计算机应用与软件, 2018, 35(12): 269-273.
(Wang G Q. Attribute reduction algorithm based on neighborhood combinatorial entropy[J]. Computer Applications and Software, 2018, 35(12): 269-273.)
- [8] Chen M, Yuan J, Li L, et al. Heuristic attribute reduction and resource-saving algorithm for energy data of data centers[J]. Knowledge and Information Systems, 2019, 61(1): 277-299.
- [9] Kang Y, Wu S, Li Y, et al. A variable precision grey-based multi-granulation rough set model and attribute reduction[J]. Knowledge-Based Systems, 2018, 148: 131-145.
- [10] Fan X, Zhao W, Wang C, et al. Attribute reduction based on max-decision neighborhood rough set model[J]. Knowledge-Based Systems, 2018, 151: 16-23.
- [11] 葛浩, 李龙澍, 杨传健. 差别矩阵约简表示及其快速算法实现[J]. 控制与决策, 2016, 31(1): 12-20.
(Ge H, Li L S, Yang C J. Discernibility matrix-based reduct representation and quick algorithms[J]. Control and Decision, 2016, 31(1): 12-20.)
- [12] 荣梓景. 关系决策系统的分布约简[J]. 计算机工程与应用, 2018, 54(17): 62-66.
(Rong Z J. Distribution reduction algorithms for relational decision systems[J]. Computer Engineering and Applications, 2018, 54(17): 62-66.)
- [13] 吴晓瑛, 魏巍, 崔军彪. 维度增量决策表的区分矩阵属性约简方法[J]. 小型微型计算机系统, 2017, 38(2): 411-416.
(Wu X Y, Wei W, Cui J B. Dimension incremental attribute reduction methods based on discernibility matrix[J]. Journal of Chinese Computer Systems, 2017, 38(2): 411-416.)
- [14] Elsayed M K. Attribute reduction based on covering granules and boolean reasoning[J]. Journal of

- Computational and Theoretical Nanoscience, 2018, 15(3): 1072-1075.
- [15] Dai J, Hu H, Wu W Z, et al. Maximal discernibility pairs based approach to attribute reduction in fuzzy rough sets[J]. IEEE Transactions on Fuzzy Systems, 2018, 26(4): 2174-2187.
- [16] 谢凤平. 区间粗糙数多属性决策方法及属性约简问题的研究[D]. 南宁: 广西大学数学与信息科学学院, 2016: 24-38.
(Xie F P. Research on multiple attribute decision making method and attributes reduction with interval rough numbers[D]. Nanning: College of Mathematics and Information Science, Guangxi University, 2016: 24-38.)
- [17] 张楠, 苗夺谦, 岳晓冬. 区间值信息系统的知识约简[J]. 计算机研究与发展, 2010, 47(8): 1362-1371.
(Zhang N, Miao D Q, Yue X D. Approaches to knowledge reduction in interval-valued information systems[J]. Journal of Computer Research and Development, 2010, 47(8): 1362-1371.)
- [18] Miao D Q, Zhang N, Yue X D. Knowledge reduction in interval-valued information systems[C]. Proceedings of the 8th IEEE International Conference on Cognitive Informatics. Hong Kong: IEEE, 2009: 320-327.
- [19] Zhang J B, Li T R, Ruan D, et al. Rough sets based matrix approaches with dynamic attribute variation in set-valued information systems[J]. International Journal of Approximate Reasoning, 2012, 53(4): 620-635.
- [20] 梁美社, 米据生, 赵天娜. 广义优势多粒度直觉模糊粗糙集及规则获取[J]. 智能系统学报, 2017, 12(6): 883-888.
(Liang M S, Mi J S, Zhao T N. Generalized dominance-based multi-granularity intuitionistic fuzzy rough set and acquisition of decision rules[J]. CAAI Transactions on Intelligent Systems, 2017, 12(6): 883-888.)
- [21] Suo Z Y, Cheng S Y, Ren J S. Probability rough set model based on the semantic in set-valued information system[C]. Proceedings of the IEEE International Conference on Computer and Communications. Chengdu: IEEE, 2017: 1244-1249.
- [22] Lin B Y, Xu W H. Multi-granulation rough set for incomplete interval-valued decision information systems based on multi-threshold tolerance relation[J]. Symmetry, 2018, 10(6): 208.
- [23] Dai J H, Wei B J, Shi H, et al. Uncertainty measurement for incomplete interval-valued information systems by θ -rough set model[C]. Proceedings of the International Conference on Information Management. Chengdu: IEEE, 2017: 212-217.
- [24] Du W S, Hu B Q. Dominance-based rough set approach to incomplete ordered information systems[J]. Information Sciences, 2016, 346/347: 106-129.
- [25] Zhang X, Chen D, Tsang E C C. Generalized dominance rough set models for the dominance intuitionistic fuzzy information systems[J]. Information Sciences, 2017, 378: 1-25.
- [26] Liu Y, Yin X, Cao B, et al. Variable precision rough set model and application based on grey similarity incidence relationship[J]. Journal of Grey System, 2017, 29(3): 45-57.
- [27] Dai J H, Gao S C, Zheng G J. Generalized rough set models determined by multiple neighborhoods generated from a similarity relation[J]. Soft Comput, 2018(22): 2081-2094.
- [28] Yamaguchi D, Li G D, Nagai M. A grey-based rough approximation model for interval data processing[J]. Information Sciences, 2007, 177(21): 4727-4744.
- [29] 张文修, 吴伟志, 梁吉业, 等. 粗糙集理论与方法[M]. 北京: 科学出版社, 2001: 22-25.
(Zhang W X, Wu W Z, Liang J Y, et al. Rough sets theory and method[M]. Beijing: Science Press, 2001: 22-25.)

作者简介

吕跃进(1958—), 男, 教授, 博士, 从事粗糙集、预测与决策、数据挖掘等研究, E-mail: lvyjin@126.com;

程林海(1995—), 男, 硕士生, 从事粗糙集、预测与决策、数据挖掘的研究, E-mail: chenglh@ sina.com;

张玉(1996—), 女, 硕士生, 从事不确定数学、数据挖掘的研究, E-mail: zyyw17@foxmail.com;

何莹莹(1994—), 女, 硕士生, 从事数据挖掘的研究, E-mail: hyy3720@126.com.

(责任编辑: 齐 霖)