

## 基于 $\alpha$ 信息熵的模糊粗糙属性约简方法

潘瑞林<sup>†</sup>, 李园沁, 张洪亮, 伊长生, 樊杨龙, 杨庭圣

(安徽工业大学 管理科学与工程学院, 安徽 马鞍山 243032)

**摘要:** 基于邻域粗糙集以及模糊粗糙集等价关系下的属性约简方法, 引入 $\alpha$ 信息熵, 建立模糊相似关系下的 $\alpha$ 信息熵不确定性度量, 提出基于 $\alpha$ 信息熵的属性重要度度量, 并以此构建混合属性约简算法. 利用UCI数据集与几种相关的约简方法进行比较, 验证了该方法可以选择较少属性的同时保证较高的分类精确性. 实际应用中, 对参数 $\alpha$ 的有效调节, 可获得多个约简结果, 进而可根据需要选择最佳约简.

**关键词:** 信息熵; 模糊相似关系; 模糊粗糙集; 约简

中图分类号: TP18

文献标志码: A

## Fuzzy-rough attribute reduction algorithm based on $\alpha$ information entropy

PAN Rui-lin<sup>†</sup>, LI Yuan-qin, ZHANG Hong-liang, YI Chang-sheng, FAN Yang-long, YANG Ting-sheng

(School of Management Science and Engineering, Anhui University of Technology, Maanshan 243032, China)

**Abstract:** Based on the attribute reduction algorithms computed with the equivalence relation in the neighborhood-rough set or fuzzy-rough set, the paper introduces the  $\alpha$  information entropy for measuring fuzzy similarity relation, and proposes an attribute significance measure model based on it. Using the measure as heuristic information, a reduction algorithm for hybrid attribute is proposed. Comparing to other methods, experimental results show that the proposed method can select a few attributes but keep, even improve classification ability in several UCI datasets. In practice, by tuning  $\alpha$  we can obtain multiple reduction results, and choose the optimum attribute reduction set according to the actual needs.

**Keywords:**  $\alpha$  information entropy; fuzzy similarity relation; fuzzy rough set; reduction

## 0 引言

现实世界中, 高维数据普遍存在, 例如在微阵列数据的基因选择以及文本分类中, 原始数据中的属性个数从几百到几个不等. 然而, 不相关或冗余属性的存在, 增加了数据存储的代价和知识获取的难度<sup>[1]</sup>. 因此, 有必要从众多属性中约简掉不相关、冗余的属性, 即属性约简. 属性约简的结果一般要求约简后集合中的属性不仅能维持原有数据的知识, 而且具有最小的冗余. 求解所有约简和最小属性约简是 NP-hard 问题, 一些求解次优或满意约简的相关算法被陆续提出, 例如基于代数观<sup>[2-4]</sup>、信息观<sup>[5-7]</sup>以及图表示<sup>[8]</sup>等的约简方法. 粗糙集理论已经被证明是处理不确定、不精确问题的强有力方法, 被广泛应用于属性约简、知识获取、数据挖掘以及粒计算中.

经典的粗糙集模型只能有效地处理名义型或离散型的数据, 对连续型数据进行属性约简前需要离散化处理. 这种方式由于离散化后的属性值没有保留其在实数值上存在的差异, 可能会导致原始数据的信息损失<sup>[1]</sup>. 而大量连续型以及连续、离散混合型数据的存在, 极大限制了经典粗糙集的应用范围. 文献[9]利用拓扑空间中球形邻域的概念, 提出了邻域粗糙集模型, 构造了数值特征(属性)的选择算法. 该方法在选择特征的数量和分类精度方面都有较大的优势, 拓展了经典粗糙集理论. 文献[10]将粗糙集与模糊集相结合, 提出了模糊粗糙集的概念, 并描述了其相关特性. 模糊粗糙集已被应用于混合型数据的属性约简中<sup>[11]</sup>. 文献[12]基于模糊等价关系(自反性、对称性、传递性), 分别从代数观与信息观出发, 定义了模糊等

收稿日期: 2015-12-07; 修回日期: 2016-03-21.

基金项目: 国家自然科学基金项目(71172219, 71302056); 安徽省科技厅软科学重大项目(1502052006).

作者简介: 潘瑞林(1973—), 男, 副教授, 博士, 从事生产运作管理、智能优化算法等研究; 李园沁(1989—), 女, 硕士生, 从事生产运作管理、机器学习的研究.

<sup>†</sup>通讯作者. E-mail: rlpn@163.com

价关系下的属性依赖度与重要度量方式,并以此作为启发式信息进行属性约简.实验分析表明,基于模糊粗糙集的属性约简方法优于先离散化再约简的方法.针对等价关系中的传递性过于严格的问题,文献[13]基于信息熵的概念,从属性区分能力角度提出了一种新的知识不确定性度量方法—— $\alpha$ 信息熵,并给出了其在一般二元关系(自反性)下的重要性质,证明了提出的 $\alpha$ 信息熵是现有多种不确定性度量的扩展.在文献[13]的基础上,滕书华等在熵的物理意义的基础上构造了两种新的不确定性度量函数—— $\alpha$ 熵和 $\alpha$ 粒度<sup>[14]</sup>.此外,滕书华等在一般二元关系下构建了一种带有可调参数的不确定性度量方式——加权 $\alpha$ 熵,并利用 $\alpha$ 精度设计了一种加权属性约简算法<sup>[15]</sup>.虽然该文献分别从完备(等价关系、优势关系)以及不完备(容差关系、非对称相似关系、限制容差关系)信息系统角度对加权 $\alpha$ 熵度量性质进行了讨论,但未将 $\alpha$ 熵引入模糊粗糙集中.

区别于文献[13]和文献[15],本文将 $\alpha$ 信息熵引入模糊粗糙集中,通过定义一种模糊相似关系(自反性,对称性)下的 $\alpha$ 信息熵,给出属性重要度的度量方式,并以此作为启发式信息,提出一种基于 $\alpha$ 信息熵的模糊粗糙属性约简方法.通过对参数 $\alpha$ 的调节,在多个UCI数据集上进行实验比较,结果表明新方法能够在获得较小约简集合的同时保持较高的分类精确性.

### 1 基于 $\alpha$ 信息熵的模糊相似关系不确定性度量模型

**定义1** 四元组 $S = (U, A = C \cup D, V, f)$ 称为一个信息系统, $U$ 为非空有限集合,称为论域,其中元素称为对象; $A = C \cup D$ 为非空有限集合,其中元素称为属性, $C$ 为条件属性集, $D = \{d\}$ 为决策属性集; $V = \bigcup_{a \in A} V_a$ 为属性的取值构成的集合,即值域; $f$ 为信息函数,它为每个属性赋予一个信息值,且对于任意 $x \in U, a \in A$ ,有 $f(x, a) \in V_a$ .若 $D = \emptyset$ ,则信息系统为信息表,否则为决策表.

给定一个非空有限对象集合 $X, R$ 为定义在 $X$ 上的一般二元关系,可用关系矩阵表示为

$$M(R) = \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ r_{21} & r_{22} & \cdots & r_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ r_{n1} & r_{n2} & \cdots & r_{nn} \end{bmatrix}, \quad (1)$$

其中 $r_{ij} \in [0, 1]$ 表示对象 $x_i$ 与 $x_j$ 的相关性大小.

如果 $R$ 满足:1)自反性, $\forall x \in X, R(x, x) = 1$ ;2)对称性, $\forall x, y \in X, R(x, y) = R(y, x)$ .则 $R$ 为模糊相似关系.如果进一步还满足:3)传递性, $\forall x, y, z \in X, R(x, z) \geq \min(R(x, y), R(y, z))$ ,则 $R$ 为模糊等价关系.模糊等价关系形成了对论域的模糊划分,模糊相似关系形成了对论域的覆盖.

模糊等价关系中的传递性过于严格,因此,本文基于模糊相似关系建立了模糊相似关系矩阵.通常相似关系的度量有夹角余弦法、最大最小法等方法.根据Yu在文献[16]提出的基于三角隶属函数给出的模糊相似关系度量,本文给出如下相关定义.

**定义2** 信息系统 $(U, A, V, f), \forall a \in A, \forall x_j \in U$ ,在属性 $a$ 上,对象 $x_j$ 与其他对象间的模糊相似关系定义为:

1) 属性 $a$ 为离散型,即

$$R_a(x_i, x_j) = \begin{cases} 1, & f(x_i, a) = f(x_j, a); \\ 0, & \text{others.} \end{cases} \quad (2)$$

2) 属性 $a$ 为连续型,即

$$R_a(x_i, x_j) = 1 - \frac{|f(x_i, a) - f(x_j, a)|}{\varepsilon}, \quad (3)$$

其中 $\varepsilon$ 为邻域半径<sup>[17]</sup>.

给定非空有限对象集合 $X, \forall x, y \in X, R$ 为定义在 $X$ 上的模糊相似关系,在关系矩阵上的一些运算定义如下:

1)  $R_1 = R_2 \Leftrightarrow R_1(x, y) = R_2(x, y)$ ;

2)  $R = R_1 \cup R_2 \Leftrightarrow R(x, y) = \max(R_1(x, y), R_2(x, y))$ ;

3)  $R = R_1 \cap R_2 \Leftrightarrow R(x, y) = \min(R_1(x, y), R_2(x, y))$ ;

4)  $R_1 \subseteq R_2 \Leftrightarrow R_1(x, y) \leq R_2(x, y)$ .

**定义3**  $U$ 为论域, $R$ 为定义在 $U$ 上的模糊相似关系,则由 $R$ 形成对 $U$ 的覆盖,定义为

$$\frac{U}{R} = \bigcup_{i=1}^n [x_i]_R, \quad (4)$$

其中 $[x_i]_R = \frac{r_{i1}}{x_1} + \frac{r_{i2}}{x_2} + \cdots + \frac{r_{in}}{x_n}$ 为由模糊相似关系 $R$ 产生的模糊相似关系类.

针对粗糙集代数观点的上、下近似集不易被人理解的问题,苗夺谦等提出了粗糙集中概念和运算的信息表示<sup>[18]</sup>;王国胤等将信息熵引入到属性约简中,建立了知识的粗糙性与信息之间的关系,提出了一种基于信息熵( $H$ 信息熵)的知识不确定性度量<sup>[19]</sup>;Liang等提出了用信息量( $E$ 信息熵)来度量信息系统的确定性,并使用该方法度量了粗糙集理

论中的不确定性和模糊性<sup>[20]</sup>.  $H$  信息熵和  $E$  信息熵是信息系统中两种典型的不确定性度量. 区别于文献 [15] 中的相关定义及性质, 本文给出了模糊相似关系下  $\alpha$  信息熵的相关定义及其性质.

**定义 4** 信息系统  $(U, A, V, f)$  中,  $U = \{x_1, x_2, \dots, x_{|U|}\}$ ,  $P \subseteq A, R$  为模糊相似关系,  $U/R(P) = \{[x_1]_R, [x_2]_R, \dots, [x_{|U|}]_R\}$  为论域的覆盖, 而  $[x_i]_R = \{x_j \in U | (x_i, x_j) \in R(P)\}$  为模糊相似关系下与  $x_i$  不可区分对象的最大集合, 则对象  $x_i$  的模糊相似  $\alpha$  信息熵定义为

$$H^\alpha(P, x_i) = (1 - 2^{\alpha-1})^{-1} \left[ \left( \frac{|[x_i]_R|}{|U|} \right)^{\alpha-1} - 1 \right], \alpha \neq 1. \quad (5)$$

**定理 1** 若  $P \subseteq A, R$  为模糊相似关系, 则  $H^\alpha(P, x_i)$  为  $\alpha$  的连续函数. 对于  $H^\alpha(P, x_i)$ , 调整  $\alpha$  值可得到不同的信息熵度量. 令  $H^\alpha(P, x) = (1 - 2^{\alpha-1})^{-1}(x^{\alpha-1} - 1), 0 < x \leq 1$ , 则当  $x = \frac{|[x_i]_R|}{|U|}$  时,  $H^\alpha(P, x) = H^\alpha(P, x_i)$ .

对  $H^\alpha(P, x)$  关于  $\alpha$  求导, 可知  $H^\alpha(P, x)$  在定义域  $\alpha \in R$  且  $\alpha \neq 1$  内单调递减, 即  $H^\alpha(P, x_i)$  随着  $\alpha$  的减小而增大.

**定理 2**  $x_i \in U, P \subseteq A, \alpha, \alpha' \in R$ . 如果  $\alpha > \alpha'$ , 则有  $H^\alpha(P, x_i) \leq H^{\alpha'}(P, x_i)$ , 当且仅当  $[x_i]_R = U$  时等号成立, 此时  $H^\alpha(P, x_i) = H^{\alpha'}(P, x_i) = 0$ .

具体证明过程可参考文献 [15] 中的定理 4.

**定义 5** 集合  $U$  的模糊相似关系  $R$  下的  $\alpha$  信息熵定义为

$$H^\alpha(P) = (1 - 2^{\alpha-1})^{-1} \left[ \frac{1}{|U|^\alpha} \sum_{i=1}^{|U|} |[x_i]_R|^{\alpha-1} - 1 \right], \quad (6)$$

其中:  $\alpha$  为实数, 且  $\alpha \neq 1$ ;  $[x_i]_R = \sum_{j=1}^n r_{ij}$  为  $[x_i]_R$  的基数. 不同的  $\alpha$  值可以得到不同的信息熵度量. 由于  $H^{\alpha'}(P, x_i)$  是单调的, 根据函数单调性的运算法则可知  $H^\alpha(P)$  关于  $\alpha$  是单调递减的.

**定理 3** 若  $P \subseteq A, \alpha, \alpha' \in R, \alpha' < \alpha$ , 则  $H^\alpha(P) \leq H^{\alpha'}(P)$ .

模糊相似关系下  $\alpha$  信息熵是基于香农信息熵对信息系统不确定性度量的扩展,  $\alpha$  信息熵越大, 信息系统所含的信息量越少, 该系统的无序程度越高. 对于论域中任意对象由模糊相似关系所构成的分类, 若其所包含的对象越多, 则认为与该对象相似的对象越

多, 所构成的知识粒度越大,  $\alpha$  信息熵越小, 分辨能力越弱.

**定理 4** (单调性) 信息系统  $(U, A, V, f)$  中,  $P, Q \subseteq A, R$  为模糊相似关系, 则  $P \prec Q$  表示  $\forall x_i \in U, [x_i]_{R(P)} \subset [x_i]_{R(Q)}$ , 意味着知识  $P$  比知识  $Q$  较细, 即  $H^\alpha(Q) < H^\alpha(P)$ ; 当  $P \approx Q$  时,  $H^\alpha(Q) = H^\alpha(P)$ .

**证明** 1) 当  $\alpha \leq 0$  时, 式 (6) 可变换为

$$H^\alpha(P) = \left(1 - \frac{1}{2^{1-\alpha}}\right)^{-1} \left[ \sum_{i=1}^{|U|} \frac{|U|^{-\alpha}}{|[x_i]_{R(P)}|^{1-\alpha}} - 1 \right],$$

此时  $\left(1 - \frac{1}{2^{1-\alpha}}\right)^{-1} > 0$ . 若  $P \prec Q$ , 则  $\forall x_i \in U$ , 有且  $\exists x_j \in U$  满足  $[x_i]_{R(P)} \subseteq [x_i]_{R(Q)}$ , 即  $\forall x_i \in U$ , 有且  $\exists x_j \in U$  满足  $|[x_j]_{R(P)}| < |[x_j]_{R(Q)}|$ . 由此可得  $H^\alpha(Q) < H^\alpha(P)$ , 即知识粒度变细,  $H^\alpha(P)$  单调增加.

2) 当  $0 < \alpha < 1$  时, 式 (6) 可变换为

$$H^\alpha(P) = \left(1 - \frac{1}{2^{1-\alpha}}\right)^{-1} \left[ \frac{1}{|U|^\alpha} \sum_{i=1}^{|U|} \frac{1}{|[x_i]_{R(P)}|^{1-\alpha}} - 1 \right],$$

此时  $\left(1 - \frac{1}{2^{1-\alpha}}\right)^{-1} > 0$ . 由式 (1) 中的分析可知, 当  $P \prec Q$  时有  $H^\alpha(Q) < H^\alpha(P)$ .

3) 当  $\alpha > 1$  时, 有  $(1 - 2^{\alpha-1})^{-1} < 0$ , 则由式 (5) 可知  $H^\alpha(Q) < H^\alpha(P)$ .

如果  $P \approx Q$ , 即  $\forall x_i \in U, [x_i]_{R(P)} = [x_i]_{R(Q)}$ , 则有  $H^\alpha(Q) = H^\alpha(P)$ .  $\square$

**定理 5** 信息系统  $(U, A, V, f)$  中,  $P \subseteq A, R$  为模糊相似关系, 有如下结论:

1) 当  $\alpha \rightarrow 1$  时,  $\lim_{\alpha \rightarrow 1} H^\alpha(P) = H(P)$ , 模糊相似关系下  $\alpha$  信息熵退化为  $H$  信息熵.

2) 当  $\alpha = 2$  时,  $H^\alpha(P) = E(P)$ , 模糊相似关系下  $\alpha$  信息熵退化为  $E$  信息熵.

**证明** 1) 当  $\alpha \rightarrow 1$  时, 由式 (6) 可得

$$\begin{aligned} \lim_{\alpha \rightarrow 1} H^\alpha(P) &= \lim_{\alpha \rightarrow 1} (1 - 2^{\alpha-1})^{-1} \left[ \frac{1}{|U|^\alpha} \sum_{i=1}^{|U|} |[x_i]_R|^{\alpha-1} - 1 \right] = \\ &= \frac{\lim_{\alpha \rightarrow 1} \left[ \sum_{i=1}^{|U|} \left( \frac{|[x_i]_R|}{|U|} \right)^\alpha \times \frac{1}{|[x_i]_R|} - 1 \right]}{(1 - 2^{\alpha-1})}. \end{aligned}$$

由洛必塔法则可得

$$\lim_{\alpha \rightarrow 1} H^\alpha(P) =$$

$$\lim_{\alpha \rightarrow 1} \frac{\sum_{i=1}^{|U|} \frac{1}{|[x_i]_R|} \times \left(\frac{|[x_i]_R|}{|U|}\right)^\alpha \times \ln \frac{|[x_i]_R|}{|U|}}{-2^{\alpha-1} \times \ln 2} = -\sum_{i=1}^{|U|} \frac{1}{|U|} \log_2 \frac{|[x_i]_R|}{|U|} = H(P).$$

2) 将 $\alpha = 2$ 代入式(6),得

$$H^\alpha(P) = -\left[\sum_{i=1}^{|U|} \frac{|[x_i]_R|^2}{|U|^2 - 1}\right] = \sum_{i=1}^{|U|} \frac{|[x_i]_R|}{|U|} \left(1 - \frac{|[x_i]_R|}{|U|}\right) = E(P). \quad \square$$

由定理5可知,模糊相似关系下的 $H$ 信息熵和 $E$ 信息熵是 $\alpha$ 信息熵的特殊形式.因此,模糊相似关系下的 $\alpha$ 信息熵是知识不确定性度量的扩展.

上述定义中的 $R$ 若为等价关系,则提出的信息度量即为等价关系下的 $\alpha$ 信息熵.因此,模糊相似关系下的 $\alpha$ 条件熵和联合熵与等价关系下的条件熵和联合熵有相似的特征.

**定义6** 信息系统 $(U, A, V, f)$ ,  $A$ 为离散或连续属性集.  $B$ 和 $E$ 为 $A$ 的两个子集,  $[x_i]_B$ 和 $[x_i]_E$ 分别表示由 $B$ 和 $E$ 产生的模糊相似类.则 $B$ 和 $E$ 的联合信息熵定义为

$$H^\alpha(BE) = H^\alpha(R_E R_B) = (1 - 2^{\alpha-1})^{-1} \left[ \frac{1}{|U|^\alpha} \sum_{i=1}^{|U|} \left| [x_i]_B \cap [x_i]_E \right|^{\alpha-1} - 1 \right]. \quad (7)$$

**定义7** 信息系统 $(U, A, V, f)$ ,  $A$ 为离散或连续属性集.  $B$ 和 $E$ 为 $A$ 的两个子集,  $[x_i]_B$ 和 $[x_i]_E$ 分别是由 $B$ 和 $E$ 产生的模糊相似类,则 $B$ 和 $E$ 的条件信息熵定义为

$$H^\alpha(E|B) = (1 - 2^{\alpha-1})^{-1} \left[ \frac{1}{|U|^\alpha} \sum_{i=1}^{|U|} \left( \frac{|[x_i]_E \cap [x_i]_B|}{|[x_i]_B|} \right)^{\alpha-1} - 1 \right]. \quad (8)$$

**性质1**  $H^\alpha(E|B) = H^\alpha(BE) - H^\alpha(B)$ .

- 性质2**
- 1)  $\forall B \subseteq A, H^\alpha(B) \geq 0$ ;
  - 2)  $H^\alpha(BE) \geq \max\{H^\alpha(B), H^\alpha(E)\}$ ;
  - 3) 若 $B \subseteq E$ 或 $R_B \subseteq R_E$ ,则 $H^\alpha(BE) = H^\alpha(E)$ ;
  - 4) 若 $B \subseteq E$ 或 $R_B \subseteq R_E$ ,则 $H^\alpha(E|B) = 0$ .

性质2中,1)说明 $\alpha$ 信息熵为非负数;2)表示两个属性的联合熵大于等于其中任何一个属性的熵值,即当加入一个属性时,属性总体的分辨能力保持不变或者增加;3)说明若 $B$ 包含于 $E$ 中,则加入属性 $B$ 不会引起 $\alpha$ 信息熵的变化;4)说明若 $B$ 包含于 $E$ 中,则属性 $B$ 和 $E$ 的条件熵为0.

**定义8** 信息系统 $(U, A, V, f)$ ,  $B \subseteq A, \forall a \in B$ 是冗余属性,则有 $H^\alpha(B) = H^\alpha(B - a)$ ;如果 $B$ 是独立的,则 $\forall a \in B$ 有 $H^\alpha(B) > H^\alpha(B - a)$ .因此,如果 $B$ 和 $A$ 满足下列条件:

- 1)  $H^\alpha(B) = H^\alpha(A)$ ;
- 2)  $\forall a \in B, H^\alpha(B) > H^\alpha(B - a)$ ;

则称 $B$ 为 $A$ 的一个约简.

**定义9** 决策信息系统 $(U, A, V, f)$ ,  $A = C \cup d$ .  $B$ 为属性集 $C$ 的子集.若 $\forall a \in B$ 相对于决策属性 $d$ 是冗余的,则 $H^\alpha(d|B - a) = H^\alpha(d|B)$ .因此,如果满足下列条件:

- 1)  $H^\alpha(d|B) = H^\alpha(d|C)$ ;
- 2)  $\forall a \in B, H^\alpha(d|B - a) > H^\alpha(d|B)$ ;

则称 $B$ 为 $A$ 的一个相对约简.

## 2 基于 $\alpha$ 信息熵的模糊粗糙属性约简算法

由于寻找最小约简是个NP-hard问题,基于属性重要度的启发式算法已经被应用到属性约简中,这些算法能够得到一些次优约简结果的同时只需较少的代价.经典的机器学习算法(如ID3、KNN)将香农信息熵作为属性重要性的度量.前文提出的建立在模糊相似关系下的 $\alpha$ 信息熵不确定性度量,能够有效度量属性的分辨能力.信息表中的熵值越大,属性的分辨能力越强,属性的重要度越高.根据提出的 $\alpha$ 信息熵的性质2中的2),当增加一个属性,熵值是增加的,表明属性个数的增加使得属性的分辨能力增强.而对于决策表,条件熵的值越大,属性对于决策的参考重要度越小.因此,下面给出模糊相似关系下基于 $\alpha$ 信息熵的属性重要度的定义.

**定义10** 信息系统 $(U, A, V, f)$ ,  $B \subseteq A, \forall a \in B$ 的重要度 $SIG(a, B)$ 定义为

$$SIG(a, B) = H^\alpha(B) - H^\alpha(B - a). \quad (9)$$

**定义11** 决策信息系统 $(U, A, V, f)$ ,  $A = C \cup d$ .  $C$ 为条件属性集,  $d$ 为决策属性.  $B \subseteq A, \forall a \in B$ 对于决策属性 $d$ 的重要度 $SIG(a, B, d)$ 定义为

$$SIG(a, B, d) = H^\alpha(d|B - a) - H^\alpha(d|B). \quad (10)$$

针对混合属性,将以上两种属性重要度作为启发式信息,本文采用前向搜索构建了基于 $\alpha$ 信息熵的模糊粗糙属性约简方法( $\alpha$ -FRS),包括约简以及相对约简算法两部分,具体描述如下.

**算法1** 基于 $\alpha$ 信息熵的模糊粗糙混合属性约简算法.

输入: 一个完备信息表 $(U, C, V, f)$ ;

输出: 该信息表的约简.

Step 1:  $\forall a \in C$ , 计算模糊相似关系  $M(R) = (r_{ij})_{n \times n}$ .

Step 2: 初始化约简集,  $\text{reduct} = \emptyset$ .

Step 3:  $\forall a \in C - \text{reduct}$ , 计算每一个属性  $a$  的重要度  $S_i = \text{SIG}(a, C - \text{reduct})$ .

Step 4: 选择属性重要度最大的属性, 即

$$\text{SIG}(a, C - \text{reduct}) = \max_i(S_i).$$

Step 5: 如果  $\text{SIG}(a, C - \text{reduct}) > \delta$  ( $\delta$  为很小的值), 则将该属性加入到约简集中, 即  $a \rightarrow \text{reduct}$ , 返回 Step 3.

**算法 2** 基于  $\alpha$  信息熵的模糊粗糙混合属性相对约简算法.

输入: 一个完备决策表  $(U, C \cup d, V, f)$ ;

输出: 该决策表的一个相对约简.

Step 1:  $\forall a \in C$ , 计算模糊相似关系  $M(R) = (r_{ij})_{n \times n}$ .

Step 2: 初始化约简集,  $\text{reduct} = \emptyset$ .

Step 3:  $\forall a \in C - \text{reduct}$ , 计算每一个属性  $a$  相对于决策属性的重要度

$$S_i = \text{SIG}(a, C - \text{reduct}, d).$$

Step 4: 选择属性重要度最大的属性, 即

$$\text{SIG}(a, C - \text{reduct}, d) = \max_i(S_i).$$

Step 5: 如果  $\text{SIG}(a, C - \text{reduct}, d) > \delta$  ( $\delta$  为很小的值), 则将该属性加入到约简集中, 即  $a \rightarrow \text{reduct}$ , 返回 Step 3.

**算法终止条件的设定** 在经典粗糙集理论中, 约简属性集与原始属性集的信息熵(或条件信息熵)需要保持一致, 然而在模糊相似关系下, 该条件不一定成立<sup>[16]</sup>. 所以, 对终止条件设置一个阈值, 当满足条件  $\text{SIG}(a, C - \text{reduct}) \leq \delta$  或  $\text{SIG}(a, C - \text{reduct}, d) \leq \delta$  时, 算法终止. 在实际应用中, 有效的  $\delta$  值可以根据原有属性个数和总体的条件信息熵进行设定. 为了便于比较, 本文基于文献[17]中的实验结果, 在实验部分中设定  $\delta = 0.05$ .

**算法的时间复杂度分析** 在算法 1 中, Step 1 需要计算每个属性所对应的相似关系矩阵, 因此其时间复杂度为  $O(|C||U|^2)$ ; Step 3 的核心在于求属性信息熵的变化, 每次选择一个属性进行判断, 在最坏的情况下, 其时间复杂度为  $O(|C|^2|U|)$ . 因此, 算法 1 的时间复杂度为  $O(|C||U|^2) + O(|C|^2|U|)$ , 与文献[12]中算法的时间复杂度相当, 要小于文献[19]中相应属性约简算法的时间复杂度. 同理可知, 算法 2 与算法 1

的时间复杂度相同.

### 3 实验设计与结果分析

实验环境为: CPU 为 Intel(R) Core(TM) i7-4790 3.60 GHz, 内存为 4 G, 操作系统为 Windows 7, 程序在 Matlab 2010b 软件中实现.

#### 3.1 实验环境和数据

实验选择了 12 组 UCI 数据集来评估算法的性能, 如表 1 所示. 其中: Ecoli、Heart Disease、Sick 和 Thyroid 为包含离散型和连续型数据的混合型数据集, 其他为连续型数据集. 实验对比分析了本文提出的  $\alpha$ -FRS 算法与其他相关算法, 其中包括:

1) 基于依赖度的可变精度邻域粗糙集约简算法 ( $\gamma$ -VPNRS)<sup>[21]</sup>;

2) 在文献[21]的基础上, 引入信息熵, 基于信息熵的邻域粗糙集约简算法 (ENRS);

3) 基于信息熵的模糊粗糙混合属性约简算法 (H-FR)<sup>[12]</sup>;

4) 在文献[12]的基础上, 引入 Liang 提出的信息量<sup>[20]</sup>, 基于信息量的模糊粗糙混合属性约简算法 (E-FR).

表 1 UCI 实验数据集

数据集	样本数	数值属性个数	符号属性个数	分类个数
Wine	178	13	0	3
Ionosphere	351	34	0	2
Ecoli	336	5	2	2
Heart Disease	270	7	6	2
WPBC	198	33	0	2
Seed	210	7	0	3
Australian	690	14	0	2
Sick	1947	7	22	2
Thyroid	5787	7	22	2
WDBC	569	30	0	2
Sonar	208	60	0	2
Derm	366	34	0	6

在算法对比过程中, 基于 Hu 在文献[9]中实验所得出的邻域半径以及精度的理想范围, 本文设置邻域半径参数  $\epsilon = 0.25$ , 精度参数  $k = 0.8$ . 前文已经证明, 在不同的  $\alpha$  值下求得的模糊相似关系  $\alpha$  信息熵不同, 因此, 不同的  $\alpha$  值求得的属性重要度也将不同, 对应的约简结果也不同. 对于  $\alpha$ -FRS 算法, 实验给出了信息熵因子  $\alpha$  在  $0 \sim 3$  变化, 以 0.05 为步长, 记录最好的分类精度所对应的约简结果. 为比较约简后属性的分类能力, 实验采用经典的 LibSVM 和 CART 分类算法, 以 10 折交叉验证的方式得出分类精确性.

### 3.2 算法对比结果与分析

表2给出了表1中前10个数据集基于不同算法约简前后的属性个数. 结果显示每种算法所约简的属性个数不同, 但是有些算法在一些数据集中并不能很好实现对冗余属性的删除. 例如, H-FR算法没有实现对Ecoli、Heart Disease以及Seed三个数据集的约简,  $\gamma$ -VPNRS和ENRS算法没有实现对Ecoli与Heart Disease两个数据集的约简. 而 $\alpha$ -FRS算法和E-FR算法对属性的约简相对更加有效. 表3给出了 $\alpha$ -FRS算法与H-FR算法约简后属性的比较结果, 可见两种算法对相同的数据集所选择的属性是不同的, 而且所选择的第1个属性对后续属性的选取有重要影响. 以数据集Wine为例, H-FR算法约简结果为: 7, 1, 10, 13, 5, 2, 8, 11, 3, 9, 12;  $\alpha$ -FRS算法约简结果为: 12, 13, 1, 2, 8, 10, 6, 5, 4, 3.

表4和表5分别给出了不同算法约简后属性的分类质量. 从中可以看出: 绝大多数情况下,  $\alpha$ -FRS算法约简后的属性都能维持甚至增加分类精确性. 例如数据集WPBC共有33个属性, 利用 $\alpha$ -FRS算法约简后仅剩2个属性, 其基于LibSVM分类算法的分类精确性变化仅为1.52(100  $\rightarrow$  98.48)个百分点, 在基于CART分类算法中分类精确性保持不变; 对于数据集

Heart Disease, 基于CART分类算法,  $\alpha$ -FRS算法约简前后的分类精确性分别为0.693 6和0.774 4, 约简后的属性分类精确性却增加了8.08个百分点; 而对于数据集WDBC, 基于两种分类算法, 其约简后的属性分类精确性都增加了0.17个百分点. 由此可知, 并不是属性越多分类精确性越高, 冗余或者不相关属性的删除有时反而能够提高分类算法的分类精确性.

虽然E-FR算法在约简后属性个数上较少, 但从表4和表5中可以看到: 除Thyroid数据集外, 根据LibSVM分类算法, 经 $\alpha$ -FRS算法约简后属性的分类精确性明显要优于E-FR算法; 对于数据集Wine、

表2 基于不同属性约简算法所选择属性个数的比较

数据集	原始属性个数	$\gamma$ -VPNRS	ENRS	H-FR	E-FR	$\alpha$ -FRS
Wine	13	10	13	11	6	10
Ionosphere	34	12	11	9	4	10
Ecoli	7	7	7	7	6	4
Heart disease	13	13	13	13	7	7
WPBC	33	2	2	2	1	2
Seed	7	5	6	7	5	3
Australian	14	5	5	5	2	5
Sick	29	1	3	3	2	3
Thyroid	29	5	15	16	11	11
1 WDBC	30	12	11	11	6	11

表3  $\alpha$ -FRS算法与H-FR算法选择属性的比较

数据集	H-FR	$\alpha$ -FRS
Wine	7, 1, 10, 13, 5, 2, 8, 11, 3, 9, 12	12, 13, 1, 2, 8, 10, 6, 5, 4, 3
Ionosphere	5, 1, 7, 21, 23, 15, 17, 8, 6, 29, 4, 33	5, 6, 8, 25, 28, 34, 29, 3, 23, 18
Ecoli	6, 1, 5, 2, 3, 7, 4	6, 1, 5, 2
Heart Disease	13, 12, 8, 3, 10, 7, 9, 11, 1, 4, 2, 5, 6	12, 3, 11, 13, 7, 4, 9
WPBC	5, 1	5, 1
Seed	1, 6, 7, 5, 4, 3, 2	1, 6, 7
Australian	14, 2, 13, 7, 1	14, 13, 2, 5, 7
Sick	20, 26, 22	22, 20, 26
Thyroid	22, 19, 9, 2, 13, 10, 7, 8, 5, 18, 11, 3, 15, 4, 12, 16	22, 19, 18, 2, 10, 9, 5, 15, 13, 17, 8
WDBC	28, 21, 22, 11, 8, 29, 16, 7, 1, 20, 12	28, 21, 22, 25, 7, 29, 10, 12, 2, 8, 16

表4 基于不同属性约简算法的分类精确性的比较(LibSVM分类算法)

数据集	原始数据	$\gamma$ -VPNRS	ENRS	H-FR	E-FR	$\alpha$ -FRS	$\alpha$ 值
Wine	0.983 1	0.983 1	0.988 7	0.983 1	0.938 2	<b>0.988 7</b>	1.1
Ionosphere	0.963 0	0.934 5	0.937 3	0.920 2	0.880 3	0.954 4	1.05
Ecoli	0.687 5	0.687 5	0.687 5	0.687 5	0.687 5	<b>0.699 4</b>	1.4
Heart Disease	0.603 0	0.603 0	0.603 0	0.603 0	0.562 3	<b>0.606 1</b>	1.9
WPBC	1	0.984 8	0.984 8	0.984 8	0.848 5	0.984 8	0.9
Seed	0.923 8	0.923 8	0.914 3	0.923 8	0.923 8	<b>0.933 3</b>	1.05
Australian	0.997 1	0.991 3	0.991 3	0.966 7	0.814 5	0.991 3	0.95
Sick	0.974 8	0.918 9	0.933 7	0.936 3	0.933 7	0.936 3	0.9
Thyroid	0.669 7	0.666 7	0.669 7	0.669 7	0.669 7	<b>0.669 7</b>	1.8
WDBC	0.959 6	0.942 0	0.957 8	0.956 1	0.947 3	<b>0.961 3</b>	1.3

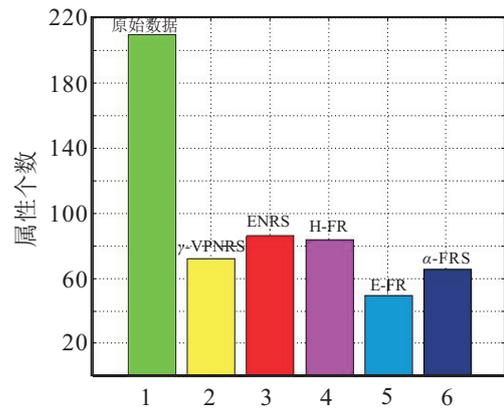
表5 基于不同属性约简算法的分类精确性的比较(CART分类算法)

数据集	原始数据	$\gamma$ -VPNRS	ENRS	H-FR	E-FR	$\alpha$ -FRS	$\alpha$ 值
Wine	0.9568	0.9568	0.9438	0.9568	0.9623	<b>0.9623</b>	1.1
Ionosphere	0.9117	0.9430	0.9658	0.9800	0.9060	<b>0.9858</b>	1.05
Ecoli	0.9137	0.9137	0.9137	0.9137	0.9137	0.9137	1.4
Heart Disease	0.6936	0.6936	0.6936	0.6936	0.6263	<b>0.7744</b>	1.9
WPBC	0.7626	0.7626	0.7626	0.7626	0.7626	0.7626	0.9
Seed	0.9714	0.9190	0.9714	0.9714	0.9714	0.9714	1.05
Australian	0.8826	0.7841	0.7841	0.7971	0.6246	0.7841	0.95
Sick	0.9949	0.9872	0.9188	0.9872	0.9779	0.9872	0.9
Thyroid	0.7079	0.6958	0.6667	0.6958	0.6958	<b>0.7172</b>	1.8
WDBC	0.9842	0.9754	0.9508	0.9859	0.9596	<b>0.9859</b>	1.3

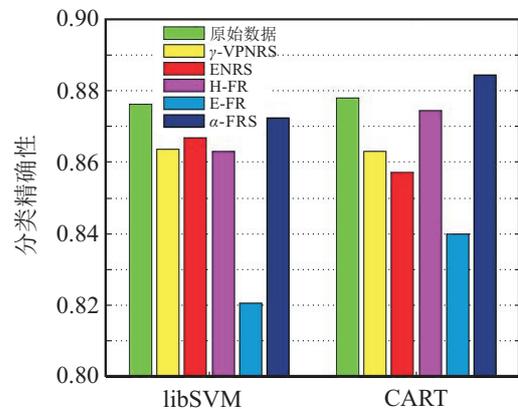
Ionosphere、Heart Disease、Australian、Sick、Thyroid 和 WDBC, 根据 CART 分类算法, 经  $\alpha$ -FRS 算法约简后属性的分类精确性同样也要优于 E-FR 算法, 而对于数据集 Ecoli、WPBC 和 Seed, 经  $\alpha$ -FRS 算法约简后属性的分类精确性能保持不低于 E-FR 算法. 因此, 评价一个属性约简算法的有效性, 不能以约简后属性的个数多少为唯一标准, 还要考虑约简结果的分类精确性.

图1从两个方面即约简集中属性个数的总和以及分类精确性的平均值, 给出了不同约简算法综合对比的结果. 结果表明, 虽然经 E-FR 算法约简后的属性个数较少, 但从分类精确性的角度来看, 并不能取得较为满意的结果. 而  $\alpha$ -FRS 算法在约简属性的个数以及约简后属性的分类精确性两方面进行综合对比, 整体优于其他4种算法. 之所以能够获得较好的约简结果, 是因为该算法可以通过调节信息熵因子  $\alpha$  的取值, 得到多种不确定性度量, 进而从中选择恰当的  $\alpha$ , 因此所计算的属性重要度更加精确.

为了明确邻域半径  $\epsilon$  和信息熵因子  $\alpha$  对约简算法的影响, 图2和图3分别给出在两个连续型数据集 (WPBC 和 WDBC) 上约简后属性个数和 CART 分类精确性随两个参数变化的情况; 而图4~图6分别给出了在3个混合型数据集 (Thyroid、Heart Disease 和 Sick) 上约简后属性个数和 CART 分类精确性随两个参数变化的情况, 其中参数  $\epsilon$  设定取值范围在 0~0.5 之间, 以 0.05 为步长变化; 信息熵因子  $\alpha$  设定取值范围在 0~3 (不等于 1) 之间, 并以 0.05 为步长变化. 从图 2(a)、3(a)、4(a)、5(a) 及 6(a) 中可以发现, 当信息熵因子  $\alpha$  的取值在一定范围内逐渐增加时, 属性个数也基本呈递减趋势 (图 2(a) 中属性选择个数虽然在一定范围内有波动, 但图形总体显示, 属性个数呈递减趋势), 但是图 2(b)、3(b)、4(b)、5(b) 和 6(b) 给出的约简后属性的分类精确性却基本保持不变 (图 5(b) 中, 分类精

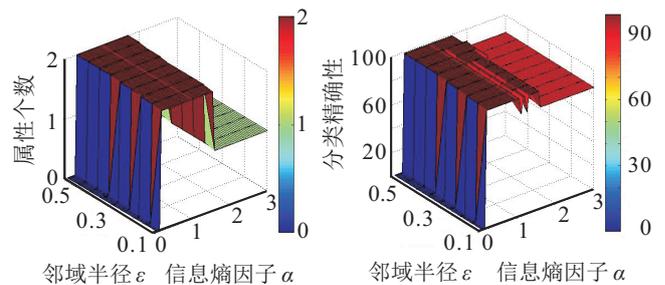


(a) 约简后属性个数总和



(b) 约简后分类精确性的平均值

图1 不同约简算法在所有数据集上的约简后属性个数总和以及分类精确性的平均值



(a) 约简后属性个数 (b) 约简后分类精确性

图2 约简后属性个数和分类精确性随邻域半径  $\epsilon$  和信息熵因子  $\alpha$  的变化 (WPBC)

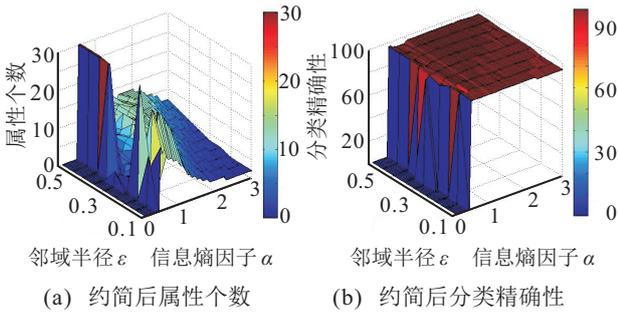


图3 约简后属性个数和分类精确性随邻域半径 $\epsilon$ 和信息熵因子 $\alpha$ 的变化(WDBC)

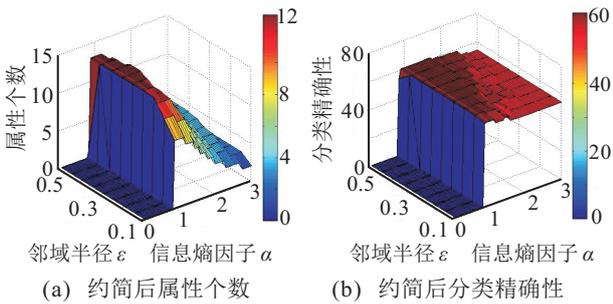


图4 约简后属性个数和分类精确性随邻域半径 $\epsilon$ 和信息熵因子 $\alpha$ 的变化(Heart Disease)

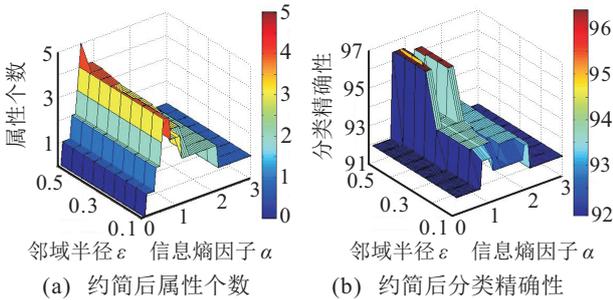


图5 约简后属性个数和分类精确性随邻域半径 $\epsilon$ 和信息熵因子 $\alpha$ 的变化(Sick)

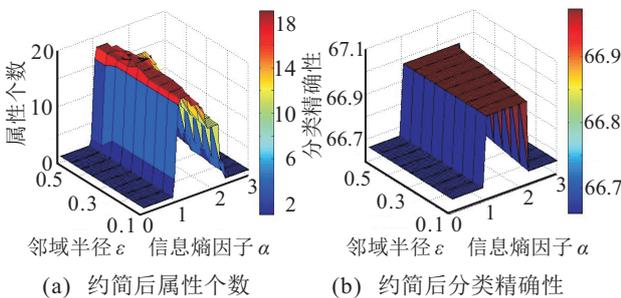


图6 约简后属性个数和分类精确性随邻域半径 $\epsilon$ 和信息熵因子 $\alpha$ 的变化(Thyroid)

确性变化幅度较大,因为属性选择个数虽然一致,但属性选择结果不一致),从而进一步验证了属性个数与属性的分类精确性之间没有必然的线性关系.还可以发现,当邻域半径 $\epsilon$ 在 $0.2 \sim 0.4$ 之间且信息熵因子 $\alpha$ 在 $0.9 \sim 1.8$ 之间时,约简的属性个数相对较少,并且能够保持较好的分类精确性.

此外,为了与文献[15]中提出的 $ARA_{\alpha}A$ 算法进行比较,本文参照该文献中 $\alpha$ 参数的相关设定分别在5个UCI数据集上进行实验,结果如表6所示.由表6可以看出,对于数据集WDBC, $\alpha$ -FRS算法在属性约简个数和分类精确性上都好于 $ARA_{\alpha}A$ 算法;对于数据集Ionosphere和Derm, $\alpha$ -FRS算法虽然在分类精确性上略低于 $ARA_{\alpha}A$ 算法,但在属性约简个数结果上优于 $ARA_{\alpha}A$ 算法;对于Wine和Sonar数据集, $\alpha$ -FRS算法在分类精确性上占优,在属性约简个数上多于 $ARA_{\alpha}A$ 算法.总体而言, $\alpha$ -FRS算法在属性约简个数的平均值上占优,而 $ARA_{\alpha}A$ 算法在分类精确性的平均值上占优.

表6 两种基于 $\alpha$ 熵属性约简算法属性约简个数与分类精确性的比较

数据集	平均属性个数		CART 分类精度	
	$ARA_{\alpha}A$	$\alpha$ -FRS	$ARA_{\alpha}A$	$\alpha$ -FRS
Wine	5.75	6.96	<b>0.9438</b>	<b>0.9586</b>
WDBC	10.6	<b>10.11</b>	0.9473	<b>0.9650</b>
Sonar	6.20	11.63	0.8077	<b>0.8873</b>
Ionosphere	8.0	<b>2.37</b>	0.9145	0.8542
Derm	9.75	<b>7.30</b>	0.9645	0.8190
<b>Average</b>	8.06	<b>7.67</b>	0.9156	0.8968

### 4 结论

模糊粗糙集已被成功应用于包含离散型和连续型的混合型数据属性约简中.本文基于邻域粗糙集和模糊粗糙集等价关系下的约简算法,通过引入 $\alpha$ 信息熵,建立了模糊相似关系下的 $\alpha$ 信息熵不确定性度量,给出了基于 $\alpha$ 信息熵的属性重要度度量方式,并以此作为启发式信息,构建了一种新的混合型数据的模糊粗糙属性约简算法.利用多个UCI数据集与其他相关算法( $\gamma$ -VPNRS、ENRS、H-FR和E-FR)进行比较分析,实验结果表明,该方法在约简后属性的个数和分类精确性两方面整体优于其他算法;而 $ARA_{\alpha}A$ 算法与本文算法各有优劣.在实际应用中,通过对参数 $\alpha$ 的调节,可以得到信息系统的多种不确定性度量,然后得出不同的属性约简结果,进而可以根据实际需要选择相应的约简结果.本文中邻域半径的优化值采用步长的实验方法获得,如何从理论分析的方法进行选择以及其选择的值与 $\alpha$ 的关系,需要进一步研究.

### 参考文献(References)

[1] Jensen R, Shen Q. New approaches to fuzzy-rough feature selection[J]. IEEE Trans on Fuzzy Systems, 2009, 17(4): 824-838.  
 [2] 杨明. 一种基于一致性准则的属性约简算法[J]. 计算机学报, 2010, 33(2): 231-239.

- (Yang M. A novel algorithm for attribute reduction based on consistency criterion[J]. Chinese J of Computers, 2010, 33(2): 231-239.)
- [3] 尹林子, 阳春华, 王晓丽, 等. 基于标记可辨识矩阵的增量式属性约简算法[J]. 自动化学报, 2014, 40(3): 397-404.  
(Yin L Z, Yang C H, Wang X L, et al. An incremental algorithm for attribute reduction based on labeled discernibility matrix[J]. Acta Automatica Sinica 2014, 40(3): 397-404.)
- [4] 胡峰, 代劲, 王国胤. 一种决策表增量属性约简算法[J]. 控制与决策, 2007, 22(3): 268-277.  
(Hu F, Dai J, Wang G Y. Incremental algorithms for attribute reduction in decision table[J]. Control and Decision, 2007, 22(3): 268-277.)
- [5] 苗夺谦, 王国胤, 刘清. 粒计算: 过去, 现在与展望[M]. 北京: 科学出版社, 2007: 142-180.  
(Miao D Q, Wang G Y, Liu Q. Granular computing: Past, now and hope[M]. Beijing: Science Press, 2007: 142-180.)
- [6] 马建敏, 张文修, 朱朝晖. 基于信息量的序信息系统的属性约简[J]. 系统工程理论与实践, 2010, 30(8): 1679-1683.  
(Ma J M, Zhang W X, Zhu Z H. Information quantity-based attribute reduction in ordered information systems[J]. System Engineering—Theory & Practice, 2010, 30(8): 1679-1683.)
- [7] 江峰, 王莎莎, 杜军威, 等. 基于近似决策熵的属性约简[J]. 控制与决策, 2015, 30(1): 65-70.  
(Jiang F, Wang S S, Du J W, et al. Attribute reduction based on approximation decision entropy[J]. Control and Decision, 2015, 30(1): 65-70.)
- [8] 苗夺谦, 陈玉明, 王睿智, 等. 图表示下的知识约简[J]. 电子学报, 2010, 38(7): 1952-1957.  
(Miao D Q, Chen Y M, Wang R Z, et al. Knowledge reduction algorithm under graph view[J]. Acta Electronica Sinica, 2010, 38(7): 1952-1957.)
- [9] 胡清华, 于达仁, 谢宗霞. 基于邻域粒化和粗糙逼近的数值属性约简[J]. 软件学报, 2008, 19(3): 640-649.  
(Hu Q H, Yu D R, Xie Z X. Numerical attribute reduction based on neighborhood granulation and rough approximation[J]. J of Software, 2008, 19(3): 640-649.)
- [10] Dubois D, Prade H. Putting rough sets and fuzzy sets together[M]. Springer Netherlands: Intelligent Decision Support, 1992: 203-232.
- [11] Chen D G, Zhang L, Zhao S Y, et al. A novel algorithm for finding reducts with fuzzy rough sets[J]. IEEE Trans on Fuzzy Systems, 2012, 20(2): 385-389.
- [12] Hu Q H, Yu D R, Xie Z X. Information-preserving hybrid data reduction based on fuzzy-rough techniques[J]. Pattern Recognition Letters, 2006, 27(5): 414-423.
- [13] 陶午沙, 滕书华, 孙即祥, 等. 基于一般二元关系的不确定性度量方法研究[J]. 国防科技大学学报, 2011, 33(2): 63-67.  
(Tao W S, Teng S H, Sun J X, et al. Study on uncertainty measure based on general binary relation[J]. J of National University of Defense Technology, 2011, 33(2): 63-67.)
- [14] 滕书华, 鲁敏, 张军, 等. 信息系统中的熵理论和信息粒度[J]. 计算机工程与科学, 2012, 34(4): 94-101.  
(Teng S H, Lu M, Zhang J, et al. Entropy theory and information granularity in information systems[J]. Computer Engineering & Science, 2012, 34(4): 94-101.)
- [15] 滕书华, 鲁敏, 杨阿锋, 等. 基于一般二元关系的粗糙集加权不确定性度量[J]. 计算机学报, 2014, 37(3): 649-665.  
(Teng S H, Lu M, Yang A F, et al. A weighted uncertainty measure of rough sets based on general binary relation[J]. Chinese J of Computers, 2014, 37(3): 649-665.)
- [16] Yu D R, Hu Q H, Wu C X. Uncertainty measures for fuzzy relations and their applications[J]. Applied Soft Computing, 2007, 7(3): 1135-1143.
- [17] 杨明. 决策表中基于条件信息熵的近似约简[J]. 电子学报, 2007, 35(11): 2156-2160.  
(Yang M. Approximate reduction based on conditional information entropy in decision tables[J]. Acta Electronica Sinica, 2007, 35(11): 2156-2160.)
- [18] 苗夺谦, 王珏. 粗糙集理论中概念与运算的信息表示[J]. 软件学报, 1999, 10(2): 113-116.  
(Miao D Q, Wang J. An information representation of the concepts and operations in rough set theory[J]. J of Software, 1999, 10(2): 113-116.)
- [19] 王国胤, 于洪, 杨大春. 基于条件信息熵的决策表约简[J]. 计算机学报, 2002, 25(6): 759-766.  
(Wang G Y, Yu H, Yang D C. Decision table reduction based on conditional information entropy[J]. J of Computers, 2002, 25(6): 759-766.)
- [20] Liang J Y, Qian Y H. Information granules and entropy theory in information systems[J]. Science in China Series F: Information Sciences, 2008, 51(9): 1427-1444.
- [21] Hu Q H, Liu J F, Yu D R. Mixed feature selection based on granulation and approximation[J]. Knowledge-Based Systems, 2008, 21(4): 294-304.

(责任编辑: 孙艺红)