

决策粗糙集的属性约简准则研究

鞠恒荣^{1,2}, 杨习贝^{1,2,3}, 于化龙¹, 戚湧³, 杨静宇²

(1. 江苏科技大学计算机科学与工程学院, 江苏 镇江 212003)

(2. 高维信息智能感知与系统教育部重点实验室, 江苏 南京 210094)

(3. 南京理工大学经济管理学院, 江苏 南京 210094)

[摘要] 属性约简是粗糙集理论研究的重要内容之一。在传统 Pawlak 粗糙集模型中, 随着属性数量的单调变化, 下、上近似集也单调变化。然而, 在决策粗糙集模型中, 随着属性的单调增加, 下、上近似集有可能增加也有可能减少。针对这一问题, 从优化角度给出了决策单调准则、一般性准则和代价准则的适应性函数并通过遗传算法求得三种准则下的约简。实验结果表明: 决策单调准则约简获得了更多的正域规则; 一般性准则约简获取了最多的正域规则; 代价准则约简获得了最小的决策代价。

[关键词] 属性约简, 代价, 决策粗糙集, 决策单调, 一般性准则

[中图分类号] TP18 **[文献标志码]** A **[文章编号]** 1001-4616(2015)01-0041-07

Research on Attribute Reduction Criteria in Decision-Theoretic Rough Set

Ju Hengrong^{1,2}, Yang Xibei^{1,2,3}, Yu Hualong¹, Qi Yong³, Yang Jingyu²

(1. School of Computer Science and Engineering, Jiangsu University of Science and Technology, Zhenjiang 212003, China)

(2. Key Laboratory of Intelligent Perception and Systems for High-Dimensional Information, Nanjing University of Science and Technology, Ministry of Education, Nanjing 210094, China)

(3. School of Economics and Management, Nanjing University of Science and Technology, Nanjing 210094, China)

Abstract: Attribute reduction is one of the important research issues in rough set theory. In classical Pawlak rough set, the lower and upper approximations are monotonic with respect to the set inclusion of attributes. However, in decision-theoretic rough set model, the lower and upper approximations may increase or decrease with respect to the increasing of attributes. To address this issue, from the viewpoint of optimization, fitness functions of the decision-monotonicity criterion, generality criterion and cost criterion have been proposed respectively. Genetic algorithm is also applied to compute reducts. The experimental results show that: the reducts based on decision-monotonicity criterion can generate more positive rules; the reducts based on generality criterion can generate most positive rules; the reducts based on cost criterion can obtain lowest decision costs.

Key words: attribute reduction, cost criterion, decision-theoretic rough set, decision-monotonicity, generality criterion

在粗糙集理论研究发展过程中, 为了使粗糙集模型在处理分类决策问题时具有容错性和风险代价敏感性, 加拿大华人学者 Yiyu Yao 教授在 20 世纪 90 年代初提出了一种新的粗糙集理论与方法: 决策粗糙集 (Decision-theoretic Rough Set)^[1]。决策粗糙集理论将 Bayes 风险决策方法引入粗糙集理论模型中, 通过分析各种决策的风险代价, 找出最小风险代价的决策, 以此作为把对象划分到正域、负域和边界区域的依据, 形成了接受决策、拒绝决策和延迟决策的三支决策语义^[2]。决策粗糙集理论及三支决策理论的最新研究进展可参见文献[3-13]。

属性约简是粗糙集理论研究的一个重要内容。与传统粗糙集属性约简研究不同, 决策粗糙集模型下的属性约简问题面临一些新的挑战。Pawlak 粗糙集模型下近似集的定义建立在集合的代数包含关系基础

收稿日期: 2014-08-16。

基金项目: 国家自然科学基金(61100116、61272419、61305058)、江苏省自然科学基金(BK2011492、BK2012700、BK20130471)、高维信息智能感知与系统教育部重点实验室(南京理工大学)开放基金(30920130122005)、中国博士后科学基金(2014M550293)。

通讯联系人: 杨习贝, 博士后, 副教授, 研究方向: 粗糙集、粒计算、知识发现。E-mail: zhenjiangyangxibei@163.com

上,因此由属性集的包含关系可以诱导出下近似集的包含关系. 这一单调性准则在 Pawlak 粗糙集属性约简计算中发挥着巨大作用. 而在决策粗糙集模型中,下、上近似集由条件概率和阈值共同决定. 因此,随着属性的不断增多,下近似有可能增大也有可能减少,上近似亦类似. 针对这一问题,众多研究者做出了巨大的努力,例如:Ma 等人详细分析了现阶段约简定义和约简算法的局限性并基于信息熵提出了一种新的属性重要度来实现决策区域保持这一目标^[14];Li 等人提出了使得正域中元素个数至少不会减少的约简^[15];Jia 等人设计了求得较小代价的约简算法^[16]等等.

在决策粗糙集非单调性背景下,Yao 和 Zhao 从三支决策规则以及决策代价的角度出发,分别提出了 3 种不同类型的约简准则,即决策单调准则、一般性准则和代价准则^[17,18]. 决策单调准则要求通过属性约简,决策者能够获得比原始信息系统更大的下近似集和更大的上近似集;一般性准则追求获得更大的下近似集;代价准则要求通过属性约简,降低信息系统的决策代价. 然而遗憾的是,Yao 与 Zhao 不仅并未给出求解基于这三种准则约简的方法,而且也没有就约简前后以及三种准则相互之间的关系进行对比分析,这正是笔者所要讨论的内容.

本文的主要内容安排如下:第二节简要介绍 Pawlak 粗糙集和决策粗糙集;第三节分析了决策粗糙集的单调性,并给出了三种属性约简准则最优化下的适应性函数;第四节通过 8 组 UCI 上的数据对三种约简准则进行了对比分析;第五节总结全文.

1 Pawlak 粗糙集与决策粗糙集

一般而言,决策信息系统^[19-21]可被定义为 1 个二元组 $S = \langle U, AT \cup D \rangle$,其中 U 表示所有对象的集合,称为论域; AT 表示所有条件属性的集合, D 为决策属性集合.

对于 $\forall a \in AT \cup D$,定义映射: $U \rightarrow V_a, V_a$ 表示属性 a 的值域,即 $a(x) \in V_a (\forall x \in U)$.

在决策信息系统 S 中,根据属性集合 AT ,可得到 1 个不可分辨关系形如

$$IND(AT) = \{ (x, y) \in U^2 : \forall a \in AT, a(x) = a(y) \}. \quad (1)$$

显然不可分辨关系满足自反性、对称性和传递性,因而是 1 个等价关系.

定义 1 令 S 为一决策信息系统,对于 $\forall X \in U, X$ 的下近似集合 $\underline{AT}(X)$ 与上近似集合 $\overline{AT}(X)$ 分别定义为:

$$\underline{AT}(X) = \{ x \in U : [x]_{AT} \subseteq X \}, \overline{AT}(X) = \{ x \in U : [x]_{AT} \cap X \neq \emptyset \}, \quad (2)$$

其中 $[x]_{AT} = \{ y \in U : (x, y) \in IND(AT) \}$ 表示 U 中所有与 x 具有不可分辨关系 $IND(AT)$ 的对象的集合,即 x 的等价类.

决策粗糙集模型是 Bayes 决策理论的 1 个简单应用. 相较于其他泛化模型,决策粗糙集的主要贡献之一就是将误分类代价与延迟决策代价引入粗糙集模型中,并基于误分类代价和延迟决策代价得到了构建概率粗糙集所需要的阈值. 其描述如下:对于任意的子集 $X \subseteq U$,可构建 1 个含有两个状态的集合 $\Omega = \{ X, \sim X \}$,分别表示对象属于 X 还是不属于 X ;令 $P(X|[x]_{AT})$ 表示任意对象在属于 $[x]_{AT}$ 的条件下属于 X 的条件概率. 对应于粗糙集中的 3 个域,可以构造 1 个决策动作集 $\tau = \{ e_P, e_B, e_N \}$,其中 e_P, e_B, e_N 分别代表对 1 个对象分类的动作,即选择对象属于正域($POS(X)$)、边界域($BND(X)$)或者负域($NEG(X)$). 不同的决策会引导不同的分类错误,也将产生不同的代价. 这可表示为 1 个由 3×2 的矩阵定义的损失函数:

其中, $\lambda_{PP}, \lambda_{BP}, \lambda_{NP}$ 分别表示当 1 个对象属于集合 X 时,采取动作 e_P, e_B, e_N 的损失; $\lambda_{PN}, \lambda_{BN}, \lambda_{NN}$ 分别表示当对象不属于集合 X 时,采取这些动作的损失.

对于 $[x]_{AT}$ 中的对象,采取不同的动作所产生的损失表示如下:

$$\begin{aligned} R(e_P|[x]_{AT}) &= \lambda_{PP} \cdot P(X|[x]_{AT}) + \lambda_{PN} \cdot P(\sim X|[x]_{AT}), \\ R(e_B|[x]_{AT}) &= \lambda_{BP} \cdot P(X|[x]_{AT}) + \lambda_{BN} \cdot P(\sim X|[x]_{AT}), \\ R(e_N|[x]_{AT}) &= \lambda_{NP} \cdot P(X|[x]_{AT}) + \lambda_{NN} \cdot P(\sim X|[x]_{AT}). \end{aligned}$$

根据 Bayes 决策理论,可得到以下最小风险决策规则:

表 1 决策粗糙集中的代价矩阵

Table 1 Cost matrix in decision-theoretic rough set

	X	$\sim X$
e_P	λ_{PP}	λ_{PN}
e_B	λ_{BP}	λ_{BN}
e_N	λ_{NP}	λ_{NN}

(P)规则:如果 $R(e_P|[x]_{AT}) \leq R(e_B|[x]_{AT})$ 且 $R(e_P|[x]_{AT}) \leq R(e_N|[x]_{AT})$, 则选择 $x \in \text{POS}(X)$,
 (B)规则:如果 $R(e_B|[x]_{AT}) \leq R(e_P|[x]_{AT})$ 且 $R(e_B|[x]_{AT}) \leq R(e_N|[x]_{AT})$, 则选择 $x \in \text{BND}(X)$,
 (N)规则:如果 $R(e_N|[x]_{AT}) \leq R(e_P|[x]_{AT})$ 且 $R(e_N|[x]_{AT}) \leq R(e_B|[x]_{AT})$, 则选择 $x \in \text{NEG}(X)$.
 为了简化决策规则, Yao 进一步假设如下^[2]:

$$\lambda_{PP} \leq \lambda_{BP} \leq \lambda_{NP}, \lambda_{NN} \leq \lambda_{BN} \leq \lambda_{PN};$$

$$(\lambda_{NP} - \lambda_{BP}) \cdot (\lambda_{PN} - \lambda_{BN}) > (\lambda_{BP} - \lambda_{PP}) \cdot (\lambda_{BN} - \lambda_{NN}).$$

基于以上的假设,可得到如下的简化规则:

(P)规则:如果 $P(X|[x]_{AT}) \geq \alpha$, 则选择 $x \in \text{POS}(X)$,
 (B)规则:如果 $\beta < P(X|[x]_{AT}) < \alpha$, 则选择 $x \in \text{BND}(X)$,
 (N)规则:如果 $P(X|[x]_{AT}) \leq \beta$, 则选择 $x \in \text{NEG}(X)$.

$$\text{其中: } \alpha = \frac{(\lambda_{PN} - \lambda_{BN})}{(\lambda_{PN} - \lambda_{BN}) + (\lambda_{BP} - \lambda_{PP})}, \beta = \frac{(\lambda_{BN} - \lambda_{NN})}{(\lambda_{BN} - \lambda_{NN}) + (\lambda_{NP} - \lambda_{BP})}.$$

根据如上所述的 3 种决策规则,可以构建如下所示的决策下近似集与决策上近似集:

$$\underline{AT}_{DT}(X) = \{x \in U : P(X|[x]_{AT}) \geq \alpha\}, \quad (3)$$

$$\overline{AT}_{DT}(X) = \{x \in U : P(X|[x]_{AT}) > \beta\}. \quad (4)$$

所以,在决策粗糙集模型下, X 的正域 $\text{POS}(X) = \underline{AT}_{DT}(X)$, X 的边界域 $\text{BND}(X) = \overline{AT}_{DT}(X) - \underline{AT}_{DT}(X)$, X 的负域 $\text{NEG}(X) = U - \overline{AT}_{DT}(X)$.

令 S 为决策信息系统, $U/\text{IND}(D) = \{X_1, X_2, \dots, X_n\}$ 是由决策属性 D 诱导的划分,根据损失函数与决策规则,可得到决策系统下决策规则的代价如下:

$$\text{(P)规则代价: } \text{COST}_{AT}^{\text{POS}} = \sum_{j=1}^n \sum_{x \in \underline{AT}_{DT}(X_j)} (P(X_j|[x]_{AT}) \cdot \lambda_{PP} + (1 - P(X_j|[x]_{AT})) \cdot \lambda_{PN}).$$

$$\text{(B)规则代价: } \text{COST}_{AT}^{\text{BND}} = \sum_{j=1}^n \sum_{x \in \overline{AT}_{DT}(X_j) - \underline{AT}_{DT}(X_j)} (P(X_j|[x]_{AT}) \cdot \lambda_{BP} + (1 - P(X_j|[x]_{AT})) \cdot \lambda_{BN}).$$

$$\text{(N)规则代价: } \text{COST}_{AT}^{\text{NEG}} = \sum_{j=1}^n \sum_{x \in U - \overline{AT}_{DT}(X_j)} (P(X_j|[x]_{AT}) \cdot \lambda_{NP} + (1 - P(X_j|[x]_{AT})) \cdot \lambda_{NN}).$$

根据决策规则的代价,决策系统的总代价表示为:

$$\text{COST}_{AT} = \text{COST}_{AT}^{\text{POS}} + \text{COST}_{AT}^{\text{BND}} + \text{COST}_{AT}^{\text{NEG}}.$$

2 属性约简

在 Pawlak 粗糙集模型中,任意的属性子集 $A, B \subseteq AT$, 并且 $A \subseteq B$. 那么对于任意的 $x \in U$, 根据不可分辨关系的定义有 $[x]_B \subseteq [x]_A$, 假设 $[x]_A \subseteq X$, 必定有 $[x]_B \subseteq X$. 因此由 $A \subseteq B$ 可以得到 $\underline{A}(X) \subseteq \underline{B}(X)$. 然而,在决策粗糙集模型中,由于受到条件概率和阈值的影响,有可能存在 $X_j \in U/\text{IND}(D)$, 使得 $\underline{A}_{DT}(X_j) \supseteq \underline{B}_{DT}(X_j)$. 上近似集的讨论可类似. 由此可见,决策粗糙集模型并不具有 Pawlak 经典粗糙集中由属性集的包含关系诱导下、上近似集包含关系的单调性准则. 为解决这一问题, Yao 等人提出了决策单调准则、一般性准则和代价准则.

定义 2 令 $S = \langle U, AT \cup D \rangle$ 为一决策信息系统, 其中 $A \subseteq AT$, $U/\text{IND}(D) = \{X_1, X_2, \dots, X_n\}$ 是由决策属性 D 诱导的划分,

(1) A 被称为 S 的 1 个决策单调准则约简当且仅当对任意的 $X_j \in U/\text{IND}(D)$, A 是保证 $\underline{A}_{DT}(X_j) \supseteq \underline{AT}_{DT}(X_j)$ 与 $\overline{A}_{DT}(X_j) \supseteq \overline{AT}_{DT}(X_j)$ 都成立的最小属性子集;

(2) A 被称为 S 的 1 个一般性准则约简当且仅当对任意的 $X_j \in U/\text{IND}(D)$, A 是保证 $\underline{A}_{DT}(X_j) \supseteq \underline{AT}_{DT}(X_j)$ 成立的最小属性子集;

(3) A 被称为 S 的 1 个代价准则约简当且仅当 A 是保证 $\text{COST}_A \leq \text{COST}_{AT}$ 成立的最小属性子集.

决策单调准则致力于通过约简冗余的属性,寻找使得决策粗糙集的下近似集与上近似集都尽可能增大的最小属性子集;一般性准则脱胎于传统 Pawlak 属性约简准则,传统 Pawlak 属性约简的 1 个重要准则就是保持下近似集不发生变化,由于决策粗糙集丢失了传统的单调性准则,因此,一般性准则致力于通过属性约

简,寻找使得决策粗糙集的下近似集尽可能增大的最小属性子集;引入决策代价是决策粗糙集的一个重要贡献,所以,代价准则致力于通过属性约简,寻找使得决策粗糙集总代价尽可能降低的最小属性子集.

然而,值得注意的是,上述决策单调准则、一般性准则和代价准则的定义都是一种定性表述,而非定量定义.为解决这一问题,本文将属性约简视为属性优化问题.从优化角度考虑,决策单调准则约简寻找使得决策粗糙集下近似集和上近似集最大的最小属性子集;一般准则约简寻找使得决策粗糙集下近似集最大的最小属性子集;代价准则约简是寻找决策粗糙集决策代价最小的最小属性子集.在优化问题中,适应性函数发挥着关键作用,适应性函数设计的好与坏将会直接影响优化的效果.针对上述三种约简准则, $A \subseteq AT$,可分别定义适应性函数如下:

$$f_{DM} = \sum_{j=1}^n (\underline{AT}_{DT}(X_j) \odot \underline{A}_{DT}(X_j) + \overline{AT}_{DT}(X_j) \odot \overline{A}_{DT}(X_j)),$$
$$f_G = \sum_{j=1}^n (\underline{AT}_{DT}(X_j) \odot \underline{A}_{DT}(X_j)),$$
$$f_C = COST_A,$$

其中:

$$A \odot B = \begin{cases} |B-A| : A \subseteq B \\ -\infty : A \not\subseteq B \end{cases}$$

在上述定义中, f_{DM} 表示决策单调准则的适应性函数,通过运算符 \odot 选取使得 f_{DM} 值最大的属性子集作为决策单调准则约简; f_G 表示一般性准则的适应性函数,通过运算符 \odot 选取使得 f_G 值最大的属性子集作为一般性准则约简; f_C 表示代价准则适应性函数,使 f_C 值最小的属性子集即是代价准则约简.

本文选取遗传算法实现优化问题,在遗传算法中,每条染色体由长度为 $|AT|$ 的二进制序列表示,其中“1”表示相应的属性存在于该染色体中,相反“0”表示该属性不在该染色体中.基于遗传算法的属性约简算法如下所示:

基于遗传优化的属性约简算法(GAAR)				
输入:决策信息系统 S;				
输出:约简.				
步骤 1 初始化:产生最初种群数(数量为 70),最大演化代数设为 100;				
步骤 2 计算适应性函数,对适应性函数进行评估,若满足停止条件,转第 6 步,否则执行第 3 步;				
步骤 3 选择:采用转盘赌选择方式选择对象产生下一代;				
步骤 4 交叉与变异:对生成的新一代进行交叉操作和变异操作;				
步骤 5 评估新一代的适应性函数;				
步骤 6 从当前种群中选择最优的约简.				

3 实验分析

本节将通过实验对比分析三种准则在属性约简过程中发挥的作用.本文选择了 UCI 上的 8 组数据进行实验.这 8 组数据的描述如表 2 所示.实验运行环境为 Windows 7 & Matlab R2012b.

表 2 数据描述
Table 2 Data sets description

序号	数据集	对象个数	属性个数	决策类
1	Credit Approval	690	15	2
2	Dermatology	366	34	6
3	Ionosphere	351	34	2
4	Soybean	307	35	4
5	Spect Heart	267	22	2
6	Tic-Tac-Toe Endgame	958	8	2
7	Wdbc	569	30	4
8	Zoo	101	16	7

在本节的实验中,所有数据均采取了离散化预处理方法.针对每一个数据集,利用随机函数分别生成了 10 组不同的代价矩阵,并假设正确分类不产生代价,即 $\lambda_{pp} = \lambda_{nn} = 0$.为了方便实验,适应性函数中涉及

到的 $-\infty$,在实验中以1个非常小的实数(-10^8)代替.实验中所得到的决策规则数目、决策代价、属性个数都是在10组不同代价矩阵下所得结果的平均数,分类精度则以平均数 \pm 均分差形式表示.在以下实验结果中,最大值用粗体表示,最小值以斜体表示.

表3列出了根据原始数据、经过三种不同准则约简后数据所得到的(P)规则数目和(B)规则数目,经观察不难得到如下结论:

(1)相比于原始数据而言,在分别经过决策单调准则、一般性准则、代价准则约简后所得到的数据,可以获得同样数目的(P)规则或者更多的(P)规则.在三种约简准则对比中,经过一般性准则约简得到的(P)规则数目最多,决策单调准则约简得到的(P)规则数目其次,代价准则约简得到的(P)规则数量在多数数据中与原始数据保持一致.

(2)在(B)规则数目的对比中,决策单调准则约简在大多数数据集上得到最多的(B)规则,这主要是由于在决策单调准则定义中要求上近似集尽可能增大,因此在一定程度上增加了(B)规则的数目.在多数数据集中,一般性准则约简得到的(B)规则数目最少.

表3 三种约简准则下决策规则的对比

Table 3 Comparisons of decision rule numbers among the three reduction criteria

数据序号	(P)规则数目				(B)规则数目			
	原始数据	决策单调准则	一般性准则	代价准则	原始数据	决策单调准则	一般性准则	代价准则
1	137.2	168.2	180.7	<i>135.6</i>	866.9	1 051	<i>829.6</i>	866.9
2	329.3	335.6	341.2	<i>329.3</i>	58.8	81	<i>49.7</i>	58.8
3	299.8	303.9	313.8	<i>299.8</i>	60.4	68.1	<i>59</i>	60.4
4	302.3	305.9	307	<i>302.3</i>	6.5	34.4	<i>0</i>	6.5
5	237.3	246.4	249.3	<i>237.3</i>	50.1	<i>37.7</i>	<i>31.7</i>	50.1
6	514.2	527.9	527.9	<i>514.2</i>	855.5	890.4	<i>886.2</i>	<i>855.5</i>
7	531.3	548	567.6	<i>531.3</i>	63.8	119.4	<i>16.3</i>	64.1
8	91.5	92.3	94.8	<i>91.5</i>	15.8	22.4	<i>14.3</i>	15.8

表4列出了根据原始数据、经过三种不同准则约简后数据所得到的决策代价和属性个数,经观察不难得到如下结论:

(1)在决策代价对比中,通过属性约简,原有数据集的决策代价都得到大幅度的降低.在三种约简准则对比中,代价准则约简得到最小的决策代价,决策单调准则约简得到最多的决策代价.

(2)在属性个数对比中,通过属性约简,剔除了原有数据集中的冗余属性.其中,决策单调准则约简的属性个数最少而代价准则约简的属性个数最多.由此可见,决策单调准则的属性约简率最高,代价准则的属性约简率最低.

(3)在运行时间的对比中,由实验结果不难发现决策单调约简相比较于其他两种约简准则而言需要花费更多的时间.

为了进一步对比分析三种属性约简准则的分类效果,本文采用SVM分类器和Bayes分类器对经过三种准则约简后的数据集进行分类,实验采取十折交叉验证机制.

表4 三种约简准则下决策代价,属性个数以及运行时间的对比

Table 4 Comparisons of decision cost, attribute numbers and run time among the three reduction criteria

数据序号	决策代价				属性个数				运行时间/s		
	原始数据	决策单调准则	原始数据	决策单调准则	一般性准则	代价准则	一般性准则	代价准则	决策单调准则	一般性准则	代价准则
1	721.31	289.49	15	9.3	10	11.6	278.19	<i>273.83</i>	240.4	213.2	<i>204.6</i>
2	310.88	33.63	34	23.9	24.6	27.8	30.78	<i>22.74</i>	159.29	<i>67.84</i>	97.88
3	265.15	50.49	34	<i>19.4</i>	21.1	26.1	44.23	<i>42.66</i>	94.5	65.62	<i>61.61</i>
4	248.43	21.56	35	<i>14.6</i>	17.2	17.5	12.34	<i>10.74</i>	62.95	54.40	<i>51.43</i>
5	230.76	60.17	22	<i>10.1</i>	<i>11.7</i>	18.3	54.07	<i>23.86</i>	30.46	<i>27.64</i>	28.47
6	988.32	329.26	8	7.3	7.4	8	329.26	<i>307.27</i>	294.2	<i>254.1</i>	260.6
7	475.13	119.91	30	<i>11.5</i>	12.9	23.4	86.55	<i>51.39</i>	242.0	<i>210.5</i>	<i>237.6</i>
8	80.38	14.24	16	<i>10</i>	11	12	11.62	<i>6.01</i>	9.74	7.83	8.47

表 5 列出了分别经过三种约简准则约简后数据集的分类精度. 由表中实验结果可以发现:
(1)在大多数数据集上, Bayes 分类器的分类精度低于 SVM 分类器的分类精度;
(2)在大多数数据集上, 经过代价准则约简后的数据集的分类精度高于经过决策单调准则和一般性准则约简后的数据集的分类精度.

综上所述, 可得到如下结论:
(1)基于决策单调准则的属性约简可以得到更多的(P)规则和最多的(B)规则, 并且拥有最少的属性数, 属性约简率最高. 但是却在一定程度上损失了分类精度并付出了更多的决策代价;
(2)基于一般性准则的属性约简在付出较小决策代价的基础上获得了最多的(P)规则, 并且降低了不确定性规则, 即(B)规则的数目, 但是并未获得最佳的分类精度;
(3)基于代价准则的属性约简在获取和原始数据相同(P)规则的情况下使得决策代价最小, 并且获得了最佳的分类精度, 但相较于其他两个准则, 其属性数目最多, 属性约简率最低.

表 5 三种约简准则下约简分类精度的对比
Table 5 Comparisons of classification accuracies among the three reduction criteria

数据序号	SVM 分类器精度			Bayes 分类器精度		
	决策单调准则	一般性准则	代价准则	决策单调准则	一般性准则	代价准则
1	0.669±0.041	0.688±0.009	0.691±0.001	0.503±0.061	0.537±0.012	0.535±0.008
2	0.869±0.020	0.877±0.009	0.879±0.007	0.468±0.024	0.477±0.031	0.469±0.040
3	0.715±0.009	0.718±0.006	0.721±0.006	0.606±0.039	0.643±0.012	0.645±0.014
4	0.951±0.042	0.969±0.010	0.975±0.006	0.864±0.176	0.930±0.139	0.947±0.087
5	0.7940±0.00	0.9744±0.001	0.9751±0.003	0.7407±0.00	0.7407±0.00	0.7444±0.00
6	0.653±0.000	0.653±0.000	0.653±0.000	1.000±0.000	1.000±0.000	1.000±0.000
7	0.853±0.081	0.889±0.012	0.895±0.003	0.733±0.087	0.755±0.067	0.822±0.011
8	0.825±0.055	0.847±0.045	0.860±0.022	0.805±0.122	0.824±0.121	0.881±0.046

4 结束语

决策粗糙集的非单调性现象带来了更多值得探讨的科学问题. 本文紧紧围绕 Yao 和 Zhao 提出的三种属性约简准则对决策粗糙集的属性约简进行了讨论. 考虑到此三种约简准则的初始定义是一种定性刻画, 本文将属性约简问题转化为属性优化问题, 给出了三种约简准则的适应性函数并采用遗传优化算法分别求得三种准则下的约简, 实验结果表明, 三种约简准则各有优点和不足, 例如基于代价准则的属性约简以最小的决策代价获取了和原始数据集相同的正域规则, 但是相较于其他两种准则, 其属性约简率最低.

在本文工作的基础上, 笔者下一步将对以下几个问题进行深入的探讨:
(1)如何利用决策粗糙集的属性约简处理实际工程中的复杂数据是一项艰巨的任务;
(2)在实际工程中, 测试代价也是一个非常重要的概念, 在决策粗糙集属性约简过程中同时考虑测试代价和决策代价以及两种之间的联系是值得探讨的科学问题.

[参考文献]

[1] Yao Y Y, Wong S K M. A decision theoretic framework for approximating concepts[J]. International Journal of Man-Machine Studies, 1992, 37: 793–809.
[2] Yao Y Y. The superiority of three-way decisions in probabilistic rough set models[J]. Information Sciences, 2011, 181: 1 080–1 096.
[3] Jia X Y, Tang Z M, Liao W H, et al. On an optimization representation of decision-theoretic rough set model[J]. International Journal of Approximate Reasoning, 2014, 55: 156–166.
[4] 贾修一, 李伟漳, 商琳, 等. 一种自适应求三枝决策中决策阈值的算法[J]. 电子学报, 2011, 39(11): 2 520–2 525.
[5] 贾修一, 商琳. 一种求三支决策阈值的模拟退火算法[J]. 小型微型计算机系统, 2013, 34(11): 2 603–2 606.
[6] Li H X, Zhou X Z, Huang B, et al. Cost-sensitive three-way decision: a sequential strategy[C]//Lingras P, Wolski M, Cornelis C, et al. RSKT 2013. LNCS, Heidelberg: Springer, 2013, 8 171: 325–337.

- [7] Liang D C, Liu D, Pedrycz W, et al. Triangular fuzzy decision-theoretic rough sets[J]. International Journal of Approximate Reasoning, 2013, 54(8) : 1 087–1 106.
- [8] Liang D C, Liu D. Systematic studies on three-way decisions with interval-valued decision-theoretic rough sets[J]. Information Sciences, 2014, 276: 186–203.
- [9] Liu D, Li T R, Li H X. A multiple-category classification approach with decision-theoretic rough sets[J]. Fundamenta Informaticae, 2012, 115: 173–188.
- [10] Liu D, Li T R, Liang D C. Incorporating logistic regression to decision-theoretic rough sets for classification[J]. International Journal of Approximate Reasoning, 2014, 55(1) : 197–210.
- [11] Qian Y H, Zhang H, Sang Y L, et al. Multigranulation decision-theoretic rough sets[J]. International Journal of Approximate Reasoning, 2013, 55(1) : 225–237.
- [12] Zhou B. Multi-class decision-theoretic rough sets[J]. International Journal of Approximate Reasoning, 2013, 55(1) : 211–224.
- [13] Deng X F, Yao Y Y. Decision-theoretic three-way approximations of fuzzy sets[J]. Information Sciences, 2014, 279: 702–715.
- [14] Ma X A, Wang G Y, Yu H, et al. Decision region distribution preservation reduction in decision-theoretic rough set model[J]. Information Sciences, 2014, 278: 614–640.
- [15] Li H X, Zhou X Z, Zhao J B, et al. Non-monotonic attribute reduction in decision-theoretic rough sets[J]. Fundamenta Informaticae, 2013, 126(4) : 415–432.
- [16] Jia X Y, Liao W H, Tang Z M, et al. Minimum cost attribute reduction in decision-theoretic rough set models[J]. Information Sciences, 2013, 219: 151–167.
- [17] Yao Y Y, Zhao Y. Attribute reduction in decision-theoretic rough set models[J]. Information Sciences, 2008, 178: 3 356–3 373.
- [18] Zhao Y, Wong S K, Yao Y Y. A note on attribute reduction in the decision-theoretic rough set model[C]//Peters J F, Skowron A, Chan C C, et al. TRS XIII. LNCS, Heidelberg: Springer, 2011, 6 499: 260–275.
- [19] 鞠恒荣, 马兴斌, 杨习贝, 等. 不完备信息系统中测试代价敏感的可变精度分类粗糙集[J]. 智能系统学报, 2014, 9(2) : 219–223.
- [20] Yang X B, Qi Y S, Song X N, et al. Test cost sensitive multigranulation rough set: model and minimal cost selection[J]. Information Sciences, 2013, 250: 184–199.
- [21] Yang X B, Yang J Y. Incomplete Information System and Rough Set Theory: Model and Attribute Reductions[M]. Beijing: Science Press, and Berlin Heidelberg: Springer, 2012.

[责任编辑: 顾晓天]