量化粗糙集的单调性属性约简方法

鞠恒荣[1,2] 杨习贝[1,2,3] 戚 湧[3] 杨静宇[2]
(江苏科技大学计算机科学与工程学院 镇江 212003)[1]
(高维信息智能感知与系统教育部重点实验室 南京 210094)[2] (南京理工大学经济管理学院 南京 210094)[3]

摘 要 单调性在经典粗糙集属性约简过程中发挥着重要的作用。然而,在一些泛化模型(如量化粗糙集模型)中该性质并不存在。针对该问题,提出了量化粗糙集模型中下近似单调约简的定义,并给出了求得该约简的启发式方法。实验结果表明,相较于下近似保持约简算法,下近似单调约简算法不仅耗时短,而且增加了由正域和边界域表示的确定性,同时降低了由边界域带来的不确定性。

关键词 单调性,下近似保持,下近似单调,粗糙集
中图法分类号 TP18 **文献标识码** A **DOI** 10.11896/j.issn.1002-137X.2015.8.007

Approach to Monotonicity Attribute Reduction in Quantitative Rough Set

JU Heng-rong[1,2] YANG Xi-bei[1,2,3] QI Yong[3] YANG Jing-yu[2]
(School of Computer Science and Engineering, Jiangsu University of Science and Technology, Zhenjiang 212003, China)[1]
(Key Laboratory of Intelligent Perception and Systems for High-Dimensional Information
(Nanjing University of Science and Technology), Ministry of Education, Nanjing 210094, China)[2]
(School of Economics and Management, Nanjing University of Science and Technology, Nanjing 210094, China)[3]

Abstract It is well-known that the monotonicity plays an important role in attribute reduction of classical rough set. However, such property does not always hold in some generalization models, and quantitative rough set is a typical example. From this point of view, the definition of lower approximate monotonicity attribute reduction was presented in quantitative rough set model, and the heuristic approach was also given to compute the reduct. The experiment results show that compared with lower approximate preservation reduct, the lower approximate monotonicity can not only save the time consuming, but also increase the certainties which are expressed by positive and negative regions, and decrease the uncertainty coming from boundary region.

Keywords Monotonicity, Lower approximate preservation, Lower approximate monotonicity, Rough set

1 引言

粗糙集理论[1](Rough Set)是Pawlak于20世纪80年代提出的一种能够处理不精确、不确定性问题的数学工具。目前,粗糙集理论已在数据挖掘、模式识别、人工智能和分类等领域吸引了众多学者的研究兴趣[2-5]。

经典粗糙集模型的基石是不可分辨关系,该不可分辨关系是一族等价关系的交集,因此依然是一个等价关系。在不可分辨关系的基础上,Pawlak给出了下近似和上近似的定义,使用已有的知识近似逼近目标。在Pawlak工作的基础上,随着实际工程应用的需要,出现了很多扩展粗糙集模型,如利用多粒化粗糙集处理多源数据[6],利用测试代价敏感粗糙集模型处理包含测试代价的数据[7]等。值得一提的是,在文献[8]中,Zhao等人从不可分辨关系的定义入手,认为不可分辨关系的要求过于严格,即要求两个对象在所有的属性上都具有相同的取值,他们将此情形称为强不可分辨关系。为解决这一问题,他们提出了弱不可分辨关系和α量化不可分辨关系。弱不可分辨关系要求两个对象在至少一个属性上具有相同的属性值即可,而α量化不可分辨关系从定量的角度对满足要求的属性个数进行限定。

属性约简是粗糙集理论研究中的核心内容之一。经典粗糙集模型下的属性约简是在保持信息系统中某些度量不变的条件下,删除其中不相关或者不重要的属性,从而达到提取有用信息、简化知识处理的目的。在约简过程中,单调性发挥着重要作用。例如在启发式约简算法中,随着属性的不断增多,下近似不断增大,上近似不断减少,直到度量指标不再发生变

到稿日期:2014-05-25 返修日期:2014-07-28 本文受国家自然科学基金(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),江苏省高校自然科学基金(13KJB520003),江苏省普通高校研究生科研创新计划项目(CXLX13_707)资助。

鞠恒荣(1989—),男,硕士生,主要研究方向为粗糙集、粒计算、代价敏感,E-mail:justjuhngrong@126.com;杨习贝(1980—),男,博士后,副教授,主要研究方向为粗糙集、粒计算、知识发现,E-mail:zhenjiangyangxibei@163.com(通信作者);戚 湧(1970—),男,博士,教授,博士生导师,主要研究方向为人工智能与信息处理,E-mail:qyong@njust.edu.cn;杨静宇(1941—),博士,教授,博士生导师,主要研究方向为模式识别、计算机视觉。

化。然而，在量化粗糙集模型中，随着属性的不断增多，下近似有可能增大也有可能减少，上近似的情形类似。因此，量化粗糙集模型并不具有经典粗糙集模型所具备的单调性准则。针对单调性问题，Yao 和 Zhao 在研究决策粗糙集理论过程中也发现了类似的问题，他们提出了决策单调约简的概念[9]，该约简的目标是使得决策者可以获得更大的下近似集与上近似集。然而遗憾的是，Yao 和 Zhao 所讨论的内容都基于决策粗糙集理论，对量化粗糙集模型中的单调约简问题并未涉及，同时也未给出求解决策单调约简的方法，这些将是本文的主要研究内容。

2 不可分辨关系和粗糙集

2.1 强不可分辨关系

形式化地，一个决策信息系统可被定义为二元组 $S=\langle U, AT\cup D\rangle$，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：$U\to V_a$，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，可得到一个强不可分辨关系形如：

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

强不可分辨关系 $IND(AT)$ 描述的是两个对象不可分辨当且仅当这两个对象在所有属性上的属性值都相等。显然强不可分辨关系满足自反性、对称性和传递性，因而是一个等价关系。

定义 1 令 S 为决策信息系统，$\forall X\subseteq U$，X 基于强不可分辨关系的下近似集合 $\underline{AT}_S(X)$ 与上近似集合 $\overline{AT}_S(X)$ 分别定义为：

$$\underline{AT}_S(X)=\{x\in U:[x]^S_{AT}\subseteq X\} \quad (2)$$

$$\overline{AT}_S(X)=\{x\in U:[x]^S_{AT}\cap X\neq\varnothing\} \quad (3)$$

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

2.2 弱不可分辨关系

在决策信息系统 S 中，根据条件属性集合 AT，可得到一个弱不可分辨关系：

$$WIND(AT)=\{(x,y)\in U^2:\exists a\in AT, a(x)=a(y)\} \quad (4)$$

弱不可分辨关系 $WIND(AT)$ 描述的是两个对象不可分辨当且仅当这两个对象在一个或一个以上的属性上取值相等。显然，弱不可分辨关系满足自反性对称性，因而是一个相容关系。

定义 2 令 S 为决策信息系统，$\forall X\subseteq U$，X 基于弱不可分辨关系的下近似集合 $\underline{AT}_W(X)$ 和上近似集合 $\overline{AT}_W(X)$ 分别定义为：

$$\underline{AT}_W(X)=\{x\in U:[x]^W_{AT}\subseteq X\} \quad (5)$$

$$\overline{AT}_W(X)=\{x\in U:[x]^W_{AT}\cap X\neq\varnothing\} \quad (6)$$

其中，$[x]^W_A=\{y\in U:(x,y)\in WIND(A)\}$ 表示 U 中所有与 x 具有弱不可分辨关系对象的集合。

2.3 量化不可分辨关系

根据上述定义，从二元关系构建的角度来看，强不可分辨关系要求两对象在所有属性上的取值都相等，而弱不可分辨关系则只要求两对象在一个属性上取值相同即可。这两种二元关系都并未体现满足条件的属性个数在关系构建中发挥的作用，为此，Zhao 等人提出了 α 量化不可分辨关系，具体形式如定义 3 所示。

定义 3 令 S 为决策信息系统，α 量化不可分辨关系定义如下：

$$ind_\alpha(AT)=\{(x,y)\in U^2:\frac{|a\in AT:a(x)=a(y)|}{|AT|}\geqslant\alpha\} \quad (7)$$

其中，$\alpha\in(0,1]$，$|Y|$ 表示集合 Y 的基数。

根据量化不可分辨关系的定义，可得到相对应的粗糙集下近似集和上近似集定义。

定义 4 令 S 为决策信息系统，$\forall X\subseteq U$，X 基于量化不可分辨关系的下近似集合 $\underline{AT}_\alpha(X)$ 和上近似集合 $\overline{AT}_\alpha(X)$ 分别定义为：

$$\underline{AT}_\alpha(X)=\{x\in U:[x]^\alpha_A\subseteq X\} \quad (8)$$

$$\overline{AT}_\alpha(X)=\{x\in U:[x]^\alpha_A\cap X\neq\varnothing\} \quad (9)$$

其中，$[x]^\alpha_A=\{y\in U:(x,y)\in ind_\alpha(AT)\}$ 表示 U 中所有与 x 具有量化不可分辨关系对象的集合。

基于量化不可分辨关系的下、上近似集，可将论域划分为 3 个互不相交的区域，分别为正域、边界和负域，即：

$$POS(AT,X)=\underline{AT}_\alpha(X) \quad (10)$$

$$BND(AT,X)=\overline{AT}_\alpha(X)-\underline{AT}_\alpha(X) \quad (11)$$

$$NEG(AT,X)=U-\overline{AT}_\alpha(X)=(\overline{AT}_\alpha(X))^c \quad (12)$$

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 $x\in U$，根据强不可分辨关系的定义有 $[x]^S_B\subseteq[x]^S_A$，假设 $[x]^S_A\subseteq X$，必定有 $[x]^S_B\subseteq X$。因此由 $A\subseteq B$ 可以得到 $\underline{A}_S(X)\subseteq\underline{B}_S(X)$。然而，在量化不可分辨关系中，量化不可分辨关系的确定受到阈值 α 的影响，因此由 $A\subseteq B$，并不能得到 $ind_\alpha(A)\subseteq ind_\alpha(B)$ 或者 $ind_\alpha(B)\subseteq ind_\alpha(A)$，也就无法得到下近似集的包含关系。上近似集的讨论类似。由此可见，基于量化不可分辨关系的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出下、上近似集包含关系的单调性准则。

3 属性约简

由量化不可分辨关系的定义可以发现，量化粗糙集模型并不具有传统 Pawlak 粗糙集的单调性质。即随着属性集中属性的增加（或减少），量化粗糙集的下近似集并不单调增大（减少）或者减少（增大）。因此，采用传统意义下的近似质量来构建属性重要度并不合理。为解决这一问题，本文将下近似分布约简引入量化粗糙集模型的属性约简问题中并采用对称差构建重要度，具体讨论如下。

3.1 下近似保持约简定义

定义 5 令 $S=\langle U, AT\cup D\rangle$ 为决策信息系统，$U/IND(D)=\{X_1,X_2,\cdots,X_n\}$ 是由决策属性 D 诱导出的划分，基于条件属性集 AT 的下近似分布集可表示为 $L_\alpha(AT)=\{\underline{AT}_\alpha(X_1),\underline{AT}_\alpha(X_2),\cdots,\underline{AT}_\alpha(X_n)\}$。对于任意的属性子集 $A\subseteq AT$，A 为量化粗糙集的下近似分布保持约简当且仅当：

1）A 为量化粗糙集的下近似分布协调集，即对于任意的 $X_j\in U/IND(D)$，$\underline{A}_\alpha(X_j)=\underline{AT}_\alpha(X_j)$；

2）对于 A 中的任意真子集 A'，A' 不为量化粗糙集下的下近似分布协调集。

令 S 为决策信息系统，$\alpha\in(0,1]$，$\forall A\subseteq AT$，$\forall a\in A$，可定义：

$$Sig_{in}(a,A)=\frac{\sum_{j=1}^{n}\{|\underline{A}_{\alpha}(X_j)\oplus\underline{A-\{a\}}_{\alpha}(X_j)|\}}{|U|}$$

其中，$X\oplus Y$ 表示集合 X 与集合 Y 之间的对称差。$Sig_{in}(a,A)$ 用来描述属性 a 从当前属性集 A 中删除后下近似的变化程度，相应地也可定义：

$$Sig_{out}(a,A)=\frac{\sum_{j=1}^{n}\{|\underline{A\cup\{a\}}_{\alpha}(X_j)\oplus\underline{A}_{\alpha}(X_j)|\}}{|U|}$$

其中，$a\in AT-A$，$Sig_{out}(a,A)$ 即表示属性集 A 增加属性 a 后下近似的变化程度。

3.2 下近似单调约简

由定义 4 可以发现，量化粗糙集的下近似与阈值 α 的选取息息相关，下近似随着 α 的单调减少而单调减少。据此，可推断若 α 值很小时，会出现下近似集合为空的情形。在这种情况下，保持下近似为空集没有任何实际意义。因此，很自然的想法是，在属性约简过程中可讨论寻求下近似集合尽可能地增大的最小属性子集。

同时，上节讨论了量化粗糙集模型的属性单调性问题，在文献[9]中，Yao 等人基于决策粗糙集模型的非单调性问题提出了决策单调约简概念。该约简的目标是使得决策者可以获得更大的下近似集与上近似集。然而上近似集由正域和边界域构成，在粗糙集的语义解释中，正域表示正确分类而边界域代表不确定性分类。在实际工程中，决策者往往希望得到更多的正确分类。因此本文仅考虑正域尽可能地增大，其定义如下所示。

定义 6 令 $S=\langle U,AT\cup D\rangle$ 为决策信息系统，对于任意的属性子集 $A\subseteq AT$，A 为量化粗糙集的下近似分布单调约简当且仅当：

1）A 为量化粗糙集下的下近似分布单调协调集，即对于任意的 $X_j\in U/IND(D)$，$\underline{A}_{\alpha}(X_j)\supseteq\underline{AT}_{\alpha}(X_j)$；

2）对于 A 中的任意真子集 A'，A' 不为量化粗糙集下的下近似分布单调协调集。

与定义 5 不同的是，定义 6 得到的属性子集是使得下近似集合尽可能增大的最小属性子集。同时也可推断：若 A 为量化粗糙集的一个下近似分布协调集，那么 A 也为量化粗糙集的一个下近似分布单调协调集，反之则不成立。

令 S 为决策信息系统，$\alpha\in(0,1]$，$\forall A\subseteq AT$，$\forall a\in A$，为实现下近似单调约简，重要度可定义为：

$$Sig_{in}(a,A)=\frac{\sum_{j=1}^{n}\{\underline{A}_{\alpha}(X_j)\odot\underline{A-\{a\}}_{\alpha}(X_j)\}}{|U|}$$

其中：

$$A\odot B=\begin{cases}|B-A|, & A\subseteq B\\ -|B-A|, & \text{其他}\end{cases}$$

相应地，对于任意的 $a\in AT-A$，属性集 A 增加属性 a 后下近似的变化程度可定义为：

$$Sig_{out}(a,A)=\frac{\sum_{j=1}^{n}\{\underline{A}_{\alpha}(X_j)\odot\underline{A\cup\{a\}}_{\alpha}(X_j)\}}{|U|}$$

3.3 属性约简算法

针对属性约简算法，众多学者根据不同的需求设计了很多约简算法，就目前的研究进展而言，大致可分为两类，即穷举法和启发式算法[7,10,11]。穷举算法最大的优点是可以求得数据的所有约简，但是在大规模数据处理中其时间复杂度过高，给研究者带来了很大的困扰。相比较而言，启发式算法可以快速求得数据集中的一个约简，因其高效率的特性越来越受到研究者的青睐。本文亦采取启发式算法开展下近似保持约简和下近似单调约简，其算法流程如下。

算法 启发式约简算法

输入：决策信息系统 S=〈U，AT∪D〉，阈值 α；

输出：约简 red

步骤 1 计算下近似分布集 $L_{\alpha}(AT)$；

步骤 2 red←∅；

步骤 3 ∀a∈AT，计算属性 a 的重要度 $Sig_{in}(a,A)$；

步骤 4 若 a_i 满足 $Sig_{in}(a_i,A)=\max\{\forall a\in AT:Sig_{in}(a,A)\}$，则 red←$a_i$，计算 $L_{\alpha}(red)$；

步骤 5 若 $L_{\alpha}(red)\neq L_{\alpha}(AT)$，重复以下循环，否则转步骤 6；

(1) ∀a∈AT−red，计算 $Sig_{out}(a,red)$；

(2)若 $Sig_{out}(a_j,red)=\max\{Sig_{out}(a,red):a\in AT-red\}$，则 red=red∪$\{a_j\}$；

(3)计算 $L_{\alpha}(red)$；

步骤 6 ∀a∈red，若 $L_{\alpha}(red-a)=L_{\alpha}(red)$，则 red=red−a；

步骤 7 输出 red

上述启发式算法可用来分别求解下近似保持约简和下近似单调约简。在求解下近似保持约简时，算法中的属性重要度对应于下近似保持约简的属性重要度；在求解下近似单调约简时，算法中所用到的属性重要度则对应于下近似单调约简中定义的重要度。

4 实验分析

本节将通过实验对比分析下近似保持约简算法与下近似单调约简算法。本次实验选取了 UCI 数据中的 4 组数据进行分析，这 4 组数据的具体描述如表 1 所列。实验运行环境为 Windows 7&Matlab R2012b。

表 1 实验数据基本信息

序号	数据集	样本个数	属性个数	决策类个数
1	Adult	1605	14	2
2	Annealing	798	38	5
3	Chess	3196	36	2
4	Wdbc	569	30	4

如上所述，下近似保持约简寻求保持下近似不发生变化的最小属性子集，而下近似单调约简则寻求下近似尽可能增大的最小属性子集。因此有必要考察两种约简算法后正域、边界域和负域的变化程度。假设属性集 $A\subseteq AT$ 为算法得到的约简，则定义

$$PRC=\frac{\sum_{j=1}^{n}(|POS(A,X_j)|-|POS(AT,X_j)|)}{|U|}$$

$$BRC=\frac{\sum_{j=1}^{n}(|BND(A,X_j)|-|BND(AT,X_j)|)}{|U|}$$

$$NRC=\frac{\sum_{j=1}^{n}(|NEG(A,X_j)|-|NEG(AT,X_j)|)}{|U|}$$

表示约简后正域、边界域和负域的变化程度。实验结果如表 2−表 5 所列。

表 2　两种约简算法的比较(Adult 数据)

α	下近似保持约简			下近似单调约简		
	PRC	BRC	NRC	PRC	BRC	NRC
0.1	0	0	0	0	0	0
0.2	0	0	0	0	0	0
0.3	0	0	0	0	0	0
0.4	0	0	0	0	0	0
0.5	0	0	0	0	0	0
0.6	0	0	0	0.1827	−0.3655	0.1827
0.7	0	0	0	0.4151	−0.8304	0.4151
0.8	0	0	0	0.4287	−0.8573	0.4287
0.9	0	0	0	0.3820	−0.7641	0.3820
1.0	0	0	0	0	0	0

表 3　两种约简算法的比较(Annealing 数据)

α	下近似保持约简			下近似单调约简		
	PRC	BRC	NRC	PRC	BRC	NRC
0.1	0	0	0	0.0752	−1.3008	1.2256
0.2	0	0	0	0.0752	−1.3008	1.2256
0.3	0	0	0	0.0752	−1.3008	1.2256
0.4	0	0	0	0.0752	−1.3008	1.2256
0.5	0	0	0	0.0752	−1.3008	1.2256
0.6	0	−0.7858	0.7857	0.4975	−3.6917	3.1942
0.7	0	−1.1140	1.1140	0.7945	−4.5301	3.7356
0.8	0	0.0113	−0.0113	0.9236	−3.3922	2.4687
0.9	0	0	0	0.3559	−0.7155	0.3596
1.0	0	0	0	0	0	0

表 4　两种约简算法的比较(Chess 数据)

α	下近似保持约简			下近似单调约简		
	PRC	BRC	NRC	PRC	BRC	NRC
0.1	0	0	0	0	0	0
0.2	0	0	0	0	0	0
0.3	0	0	0	0	0	0
0.4	0	0	0	0	0	0
0.5	0	0	0	0	0	0
0.6	0	0	0	0.1827	−0.3655	0.1827
0.7	0	0	0	0.4152	−0.8304	0.4152
0.8	0	0	0	0.4287	−0.8573	0.4287
0.9	0	0	0	0.3820	−0.7641	0.3820
1.0	0	0	0	0	0	0

表 5　两种约简算法的比较(Wdbc 数据)

α	下近似保持约简			下近似单调约简		
	PRC	BRC	NRC	PRC	BRC	NRC
0.1	0	0	0	0.2443	−0.4886	0.2443
0.2	0	0	0	0.2443	−0.4886	0.2443
0.3	0	0	0	0.2443	−0.4886	0.2443
0.4	0	0	0	0.2443	−0.4886	0.2443
0.5	0	0	0	0	0	0
0.6	0	0	0	0.9543	−1.9156	0.9613
0.7	0	0	0	0.9543	−1.9156	0.9613
0.8	0	0	0	0.1248	−0.2478	0.1230
0.9	0	0	0	0	0	0
1.0	0	0	0	0	0	0

从上述表中的实验结果可得到以下结论：

1)对于 *PRC*，在下近似保持约简算法中，*PRC* 值均为 0，这表明该算法实现了保持下近似不发生变化的目标；在下近似单调约简算法中，*PRC* 值均大于或者等于 0，表明本文提出的下近似单调约简算法使数据的下近似不变或得到了增大。

2)在下近似单调约简算法中，*PRC* 和 *NRC* 值均大于或者等于 0，*BRC* 值小于或者等于 0。同时仔细观察可发现 $PRC+NRC+BRC=0$，此现象表明正域和负域中增加的对象来自于边界域中减少的对象。因此，下近似单调约简降低了由边界域带来的不确定性，增加了对象分类的确定性。而在下近似保持约简算法中此效果并不理想。

表 6 列出了下近似保持约简算法和下近似单调约简算法在 4 组数据集上的平均运行时间。从平均运行时间来看，下近似单调约简算法所需要的时间也是远小于下近似保持约简算法所需要的时间。

表 6　两种约简算法的平均时间对比

数据集		Adult	Annealing	Chess	Wdbc
平均时间 (s)	下近似保持	13.1976	17.9025	174.23	5.1658
	下近似单调	6.5630	8.1702	141.81	2.3841

综上所述，相较于下近似保持约简算法，下近似单调约简算法不仅运行时间较短，而且增加了正域和负域所表示的确定性，降低了由边界域带来的不确定性。

结束语　不同形式的粗糙集泛化模型对于粗糙集理论适应不同的实际需求起到了很大的促进作用。在泛化过程中，必然会导致原有粗糙集理论的一些性质发生变化，如传统意义上的单调性。本文从下近似保持与下近似单调两个角度研究了量化粗糙集模型下的属性约简问题。实验结果表明，下近似单调约简算法不仅运行时间短，而且增加了正域和负域所表示的确定性，降低了由边界域带来的不确定性。

参 考 文 献

[1] Zdzisław P. Rough sets-theoretical aspects of reasoning about data [M]. Dordrecht: Kluwer Academic, 1991

[2] Luo Gong-zhi, Yang Xi-bei. Limited dominance-based rough set model and knowledge reductions in incomplete decision system [J]. Journal of Information Science and Engineering, 2010, 26(6): 2199-2211

[3] Xu Wei-hua, Wang Qiao-rong, Zhang Xian-tao. Multi-granulation fuzzy rough sets in a fuzzy tolerance approximation space [J]. International Journal of Fuzzy Systems, 2011, 14: 246-259

[4] Yang Xi-bei, Song Xiao-ning, Dou Hui-li, et al. Multi-granulation rough set: from crisp to fuzzy case [J]. Annals Fuzzy Mathematics Information, 2011, 1(1): 55-70

[5] Hu Qing-hua, Che Xun-jian, Zhang Lei, et al. Rank entropy based decision trees for monotonic classification [J]. IEEE Transactions on Knowledge and Data Engineering, 2012, 24(11): 2052-2064

[6] Qian Yu-hua, Liang Ji-ye, Yao Yi-yu, et al. MGRS: A multi-granulation rough set [J]. Information Sciences, 2010, 180: 949-970

[7] Yang Xi-bei, Qi Yun-song, Song Xiao-ning, et al. Test cost sensitive multigranulation rough set: model and minimal cost selection [J]. Information Sciences, 2013, 250: 184-199

[8] Zhao Yan, Yao Yi-yu, Luo Feng. Data analysis based on discernibility and indiscernibility [J]. Information Sciences, 2007, 177: 4959-4976

[9] Yao Yi-yu, Zhao Yan. Attribute reduction in decision-theoretic rough set model [J]. Information Sciences, 2008, 178(17): 3356-3373

[10] Yang Xi-bei, Song Xiao-ning, She Yan-hong, et al. Hierarchy on multigranulation structures: a knowledge distance approach [J]. International Journal of General Systems, 2013, 42(7): 754-773

[11] Min Fan, Zhu William. Attribute reduction of data with error ranges and test costs [J]. Information Sciences, 2012, 211: 48-67