

特性关系粗糙集中对象变化时近似集增量更新方法研究

周秋山 李天瑞 陈红梅 张钧波

(西南交通大学信息科学与技术学院 成都 610031)

摘要 特性关系信息系统是一般信息系统的推广模型,仅满足自反性,具有同时处理不完备信息系统中的丢失数据和不关心数据的功能。在实际应用中,信息系统随着对象的变化而不断地动态变化。研究了在特性关系粗糙集模型下单个对象增加和删除时近似集变化的规律,提出了相应的增量更新算法并给予了实验验证。

关键词 粗糙集,不完备信息系统,特性关系,增量更新

中图法分类号 TP18 文献标识码 A

Research on Properties of Approximations in Rough Sets Based on Characteristic Relation When Object Varies with Time

ZHOU Qiu-shan LI Tian-rui CHEN Hong-mei ZHANG Jun-bo

(School of Information Science and Technology, Southwest Jiaotong University, Chengdu 610031, China)

Abstract The information system based on characteristic relation is an extension of the general model of the information system. It only satisfies the reflexive and can deal with “lost” and “do not care” data in incomplete information systems simultaneously. In real-life applications, the information system may change dynamically according to the variation of objects. This paper discussed properties of approximations in rough sets based on the characteristic relation when one object is added to or deleted from the information system. An incremental algorithm for updating approximations was proposed and experimental evaluation was employed to validate the proposed method.

Keywords Rough sets, Incomplete information systems, Characteristic relation, Incremental updating

1 引言

粗糙集理论是由波兰科学家 Pawlak 于 20 世纪 80 年代初提出的一种处理不一致问题的数学工具^[1]。为了适应实际应用的需求,提出了各种扩展的粗糙集模型,如可变精度粗糙集^[2]、概率粗糙集^[3]、模糊粗糙集^[4]等,其中对不完备信息系统进行处理的各种扩展粗糙集模型已成为粗糙集理论发展的一个重要方面。对于不完备信息系统而言,人们也提出了很多处理方法,Grzymala-Busse 将不完备的信息系统转换为完备信息系统进行处理,然后提取规则^[10];Kryszkiewicz 将粗糙集理论的相关概念在不完备信息系统中进行扩展,提出了容差关系直接从不完备信息系统产生规则的方法^[6];Slowinski 等人提出了相似关系的粗糙集扩展模型^[11];Stefanowski 等人提出了非对称相似关系及量化容差关系处理和分析不完备信息系统^[5]。对于不完备信息系统而言,数据的缺失存在如下两种情况:丢失数据和不关心数据。相似关系处理的是丢失数据,容差关系处理的是不关心数据。而对于两种情况同时存在的情况,相似关系和容差关系就不适用,因此 Grzymala-Busse 提出了特性关系^[7,8]。特性关系是结合相似和容差关系的一种推广形式,仅满足自反性,它可处理丢失数据和不关心数据同时存在的不完备信息系统。

对于特性关系模型粗糙集的研究, Li 等在属性集变化的情况下,提出了基于特性关系粗糙集模型的增量更新近似集的方法,并使其与直接计算近似集的方法进行了比较。实验结果验证了所提方法的优越性,同时实现了特性关系粗糙集模型框架下多个属性同时增删时的规则提取方法^[12]。Song 和 Li 等提出了基于云变换和特性关系粗糙集模型下的决策树构造方法^[13]。刘伟斌等对特性关系粗糙集下属性值粗化细化时近似集增量更新方法进行了研究^[9]。在特性关系粗糙集模型中,对象增加和删除时上、下近似集的变化和增量更新方法的研究还无人涉及。本文主要对单个对象增加和删除时上、下近似集的变化规律进行了讨论,给出了相应的增量更新算法,并给予了实验验证。

2 基本概念

定义 1^[8] 设 $S=(U,A)$ 为信息系统, U 是由对象组成的非空有限集合; $A=C \cup D$ 为属性集合,这里的 C 为条件属性, D 为决策属性,并且 $C \cap D = \emptyset$ 。对于 $\forall a \in A$ 都有一个 a 的取值集合 V_a 。与之关联, V_a 称为属性 a 的值域, $a(x)$ 为对象 x ($x \in U$) 在属性 a 的值。如果 $a(x)$ 存在空缺值,则称 S 为不完备信息系统,系统中的空缺值分别用“?”或“*”表示。其中“?”表示丢失值,“*”表示不关心的值。

到稿日期:2011-07-01 返修日期:2011-09-02 本文受国家自然科学基金(60873108)资助。

周秋山(1985—),男,硕士生,主要研究方向为数据挖掘与知识发现等,E-mail:zhouqiushan570821@163.com;李天瑞(1969—),男,教授,博士生导师,主要研究方向为数据挖掘与知识发现、粗糙集与粒计算等,E-mail:trli@swjtu.edu.cn(通信作者)。

在文献[8]中,Grzymala-Busse 提出了属性值对、特性集和特性关系的概念。特性集和特性关系可以用下面的属性值对的思想来定义。

定义 2^[8] 设 $S=(U,A)$ 为不完备信息系统,令 b 是一个属性, $b \in A$, v 是对应 b 的一个属性取值, 设 $t=(b,v)$ 是属性值对, 即一组 t 可以定义为 $[t]=\{x \in U | b(x)=v\}$, 通过 $[t]$ 分类。如果存在一个 $x \in U$ 使得 $b(x)=?$, 那么对于 b 的任意属性值 $v(v \neq ? \text{ 和 } *)$, 都有 $x \notin [t]$; 如果存在一个 $x \in U$ 使得 $b(x)=*$, 那么对于 b 的任意属性值 $v(v \neq ? \text{ 和 } *)$, 都有 $x \in [t]$ 。

定义 3^[8] 设 $S=(U,A)$ 为不完备信息系统, 令 $B \subseteq A$ 是一个属性子集, $x \in U$, 那么特性集 $I_B^x(x)$ 被定义为如下形式:

$$I_B^x(x) = \bigcap_{b \in B} [(b, b(x))]$$

令 $B=\{a,b,c,d\}$, 由定义 3 可得表 1 的特性集:

$$\begin{aligned} I_B^x(1) &= \{1,3,4\} \cap \{1,2,5,6\} \cap \{1,2,4,5,6\} \cap \{1,2,6\} \\ &= \{1\} \\ I_B^x(2) &= \{2,3,6\} \cap \{1,2,4,5,6\} \cap \{1,2,6\} = \{2,6\} \\ I_B^x(3) &= \{3\} \\ I_B^x(4) &= \{1,3,4\} \cap \{2,4,5\} \cap \{1,2,4,5,6\} \cap \{3,4,5,6\} \\ &= \{4\} \\ I_B^x(5) &= \{1,2,4,5,6\} \cap \{3,4,5,6\} = \{4,5,6\} \\ I_B^x(6) &= \{2,3,6\} \cap \{1,2,5,6\} \cap \{1,2,4,5,6\} = \{2,6\} \end{aligned}$$

表 1 不完备信息系统

	a	b	c	d
1	1	1	1	1
2	2	*	1	1
3	*	?	2	2
4	1	2	1	2
5	?	*	1	2
6	2	1	1	*

定义 4^[8] 设 $S=(U,A)$ 为不完备信息系统, 令 $B \subseteq A$ 是一个属性子集, $x, y \in U$, 特性关系 C_B 被定义为如下形式:

$$(x, y) \in C_B \Leftrightarrow y \in I_B^x(x)$$

定义 5^[8] 设 $S=(U,A)$ 为不完备信息系统, 令 $B \subseteq A$ 是一个属性子集, $x, y \in U$, 特性关系 C_B 同时可被定义为如下形式:

$$C_B = \{(x, y) \in U^2 | \forall a \in B \wedge a(x) \neq ?, a(x) = a(y) \vee a(x) = * \vee a(y) = *\}$$

特性关系只具有自反性, 不满足对称性和传递性。对于这种不完备信息系统中的粗糙集, 仍然可以用两个精确集, 即特性关系意义下的粗糙集的上、下近似集来描述。

定义 6^[8] 在特性关系的意义下, 相对于属性子集 B 而言, 集合 $X(X \subseteq U)$ 的下近似集 X_B^L 定义为:

$$X_B^L = \bigcup \{I_B^x(x) | x \in X, I_B^x(x) \subseteq X\}$$

X 的上近似集 X_B^R 定义为:

$$\begin{aligned} X_B^R &= \bigcup \{I_B^x(x) | x \in X, I_B^x(x) \cap X \neq \emptyset\} \\ &= \bigcup \{I_B^x(x) | x \in X\} \end{aligned}$$

3 新增单个对象对特性关系粗糙集中上、下近似集的影响

记 x_{n+1} 为新增的对象, X_B^L 为新增对象后的下近似集, X_B^R 为新增对象后的上近似集, $I_B^x(x)$ 为新增对象后的特性集, C_B' 为新增对象后的特性关系, X 为对象集中的一个子集。

对于信息系统而言, 这里只针对 X 集合不变的情况进行讨论。

定理 1 当增加对象 x_{n+1} 时, 有 $x_{n+1} \notin X$, 则 $X_B^L \subseteq X_B^L$, $X_B^R \subseteq X_B^R$ 。

证明: 由 $I_B^x(x)$ 的定义式可知, 增加对象 x_{n+1} 时,

$$I_B^x(x) \subseteq I_B^x(x)$$

又 $\because X_B^L = \bigcup \{I_B^x(x) | x \in X, I_B^x(x) \subseteq X\}$, $\therefore \forall I_B^x(x) \subseteq X$, 可能 $\exists I_B^x(x) \not\subseteq X$, $\therefore X_B^L \subseteq X_B^L$ 。

由于 $I_B^x(x) \subseteq I_B^x(x)$, $X_B^R = \bigcup \{I_B^x(x) | x \in X, I_B^x(x) \cap X \neq \emptyset\}$, 则 $X_B^R \subseteq X_B^R$ 。

例 1 表 1 中, 设 $B=\{1,2,3,4\}$, $B \subseteq U$ 。通过定义 6 可得 $X_B^L = \{1,3,4\}$, $X_B^R = \{1,2,3,4,6\}$ 。当增加对象 7 对应的属性值为 $\{2,?,1,*\}$ 时, 可得 $X_B^L = \{1,3,4\} \subseteq X_B^L$, $X_B^R \subseteq X_B^R = \{1,2,3,4,6,7\}$ 。

4 删除单个对象对特性关系粗糙集中集合 X 的上、下近似集的影响

记 x_n 为删除的对象, X_B^L 为删除对象后的下近似集, X_B^R 为删除对象后的上近似集, $I_B^x(x)$ 为删除对象后的特性集, C_B' 为删除对象后的特性关系。对于删除对象 x_n 是否属于集合 X , 这里分两类进行讨论(参见以下定理 2 和定理 3)。

定理 2 当 $x_n \notin X$ 时, 则 $X_B^L \subseteq X_B^L$, $X_B^R \subseteq X_B^R$ 。

证明: 由 $I_B^x(x)$ 的定义式可知, 当删除对象 x_n 时,

$$I_B^x(x) \subseteq I_B^x(x)$$

又 $\because X_B^L = \bigcup \{I_B^x(x) | x \in X, I_B^x(x) \subseteq X\}$, $\therefore X_B^L \subseteq X_B^L$ 。

由于 $I_B^x(x) \subseteq I_B^x(x)$, $X_B^R = \bigcup \{I_B^x(x) | x \in X, I_B^x(x) \cap X \neq \emptyset\}$, 则 $X_B^R \subseteq X_B^R$ 。

例 2 就表 1 中, 设 $B=\{1,2,3,4\}$, $B \subseteq U$ 。通过定义 6 可得 $X_B^L = \{1,3,4\}$, $X_B^R = \{1,2,3,4,6\}$ 。当删除对象 6 时, 且 $6 \notin B$, $X_B^L = \{1,2,3,4\}$, $X_B^R = \{1,2,3,4\}$, 所以 $X_B^L \subseteq X_B^L$, $X_B^R \subseteq X_B^R$ 。

定理 3 当 $x_n \in X$ 时, 则 $X_B^L \subseteq X_B^L$, $X_B^R \subset X_B^R$ 。

证明: 由 $I_B^x(x)$ 的定义式可知, 当删除对象 x_n 时,

$$I_B^x(x) \subseteq I_B^x(x)$$

又 $\because X_B^L = \bigcup \{I_B^x(x) | x \in X, I_B^x(x) \subseteq X\}$, $\therefore X_B^L \subseteq X_B^L$ 。

由于 $I_B^x(x) \subseteq I_B^x(x)$, $X_B^R = \bigcup \{I_B^x(x) | x \in X, I_B^x(x) \cap X \neq \emptyset\}$, 则 $X_B^R \subset X_B^R$ 。

例 3 表 1 中, 设 $B=\{4,5,6\}$, $B \subseteq U$, 通过定义 6 求得 $X_B^L = \{4,5,6\}$, $X_B^R = \{2,4,5,6\}$ 。当删除对象 6 时, 且 $6 \in B$, $X_B^L = \{4,5\}$, $X_B^R = \{4,5\}$, 所以 $X_B^L \subseteq X_B^L$, $X_B^R \subset X_B^R$ 。

5 单个对象增加、删除时近似集增量更新算法

当单个对象增加或删除时, 静态的方法(传统的求上、下近似集的方法)是根据对象的变化按定义重新求取特性关系粗糙集下集合的上、下近似集。

增量更新方法是指当对象变化时根据变化前所求得的特性和上、下近似集, 来进一步设计算法, 以求取对象更新后的上、下近似集。

5.1 增加对象时增量更新上、下近似集的算法描述

$(x_{n+1} \notin X)$

输入: $S=(U,A,V,F), X, x_{n+1}$

输出: X_B^L, X_B^R

步骤1 $\forall x_i \in U$, 根据 $I_B^f(x_i)$ 的定义计算 $I_B^f(x_i)$ 。

步骤2 若 $x_i \in X$, 且 $I_B^f(x_i) \subseteq X$, 则根据 $X_B^f = \bigcup \{I_B^f(x) | x \in X, I_B^f(x) \subseteq X\}$, 从而求得 X_B^f , 并求得一组 x_i 的数组 array[0]。

步骤3 若 $x_i \in X$, 且 $I_B^f(x_i) \cap X \neq \emptyset$, 则根据 $X_B^f = \bigcup \{I_B^f(x) | x \in X, I_B^f(x) \cap X \neq \emptyset\}$, 从而求得 X_B^f , 并求得一组对应 x_i 的数组 array[1]。

步骤4 增加对象 x_{n+1} 时, $I_B^f(x)$ 根据步骤1求得的结果 $I_B^f(x_i)$ 以及增加对象时对各 $I_B^f(x_i)$ 的变化来求解。如果 $I_B^f(x_i)$ 发生变化, 则把 x_i 加入到一个数组 array[2]。

步骤5 求数组 array[0] 与数组 array[2] 的差, 找出其中所有元素的 $I_B^f(x_i)$, 并求其并集, 以得到下近似集 X_B^f 。

步骤6 如果数组 array[1] 与 array[2] 交集为空, 则上近似集不变。如果交集不为空, 则将 x_{n+1} 加入到 X_B^f 中, 即得到上近似集 X_B^f 。

5.2 删除对象时(在 $x_n \notin X$ 的情况下)增量更新上、下近似集的算法描述

输入: $S=(U, A, V, F), X$

输出: X_B^f, X_B^g

步骤1 $\forall x_i \in U$, 根据 $I_B^f(x_i)$ 的定义计算 $I_B^f(x_i)$ 。

步骤2 若 $x_i \in X$, 且 $I_B^f(x_i) \subseteq X$, 则根据 $X_B^f = \bigcup \{I_B^f(x) | x \in X, I_B^f(x) \subseteq X\}$, 从而求得 X_B^f , 并求得一组 x_i 的数组 array[0]。

步骤3 若 $x_i \in X$, 且 $I_B^f(x_i) \cap X \neq \emptyset$, 则根据 $X_B^f = \bigcup \{I_B^f(x) | x \in X, I_B^f(x) \cap X \neq \emptyset\}$, 从而求得 X_B^f , 并求得一组对应 x_i 的数组 array[1]。

步骤4 删除对象 x_n 时, $I_B^f(x)$ 根据步骤1求得的结果 $I_B^f(x_i)$ 以及删除对象时对各 $I_B^f(x_i)$ 的变化来求解。如果 $I_B^f(x_i)$ 发生变化, 则把 x_i 加入到一个数组 array[2]。

步骤5 如果数组 array[0] 与 array[2] 的交集为空, 则下近似集不变, 否则根据公式 $X_B^f = \bigcup \{I_B^f(x) | x \in X, I_B^f(x) \subseteq X\}$ 求下近似集。

步骤6 如果数组 array[1] 与 array[2] 交集为空, 则上近似集不变; 如果交集不为空, 则将 x_n 从 X_B^f 中删除, 即得到上近似集 X_B^f 。

5.3 删除对象时(在 $x_n \in X$ 的情况下)增量更新上、下近似集的算法描述

输入: $S=(U, A, V, F), X$

输出: X_B^f, X_B^g

步骤1 $\forall x_i \in U$, 根据 $I_B^f(x_i)$ 的定义计算 $I_B^f(x_i)$ 。

步骤2 若 $x_i \in X$, 且 $I_B^f(x_i) \subseteq X$, 则根据 $X_B^f = \bigcup \{I_B^f(x) | x \in X, I_B^f(x) \subseteq X\}$, 从而求得 X_B^f , 并求得一组 x_i 的数组 array[0]。

步骤3 若 $x_i \in X$, 且 $I_B^f(x_i) \cap X \neq \emptyset$, 则根据 $X_B^f = \bigcup \{I_B^f(x) | x \in X, I_B^f(x) \cap X \neq \emptyset\}$, 从而求得 X_B^f , 并求得一组对应 x_i 的数组 array[1]。

步骤4 删除对象 x_n 时, $I_B^f(x)$ 根据步骤1求得的结果 $I_B^f(x_i)$ 以及删除对象时对各 $I_B^f(x_i)$ 的变化来求解。如果 $I_B^f(x_i)$ 发生变化, 则把 x_i 加入到一个数组 array[2]。

步骤5 查找集合 $I_B^f(x_n) - x_n$ 中的元素是否在 array[0] 中元素所对应的 $I_B^f(x)$ 中。如果存在则退出, 如果不存在则从 X_B^f 集合中删除该元素(在遍历过程中, 如果存在 x_n , 则删

除 x_n), 得到更新后的 X_B^f 即为下近似集 X_B^f 。

步骤6 求数组 array[1] 与数组 array[2] 的差以及 array[1] 与 array[2] 的交集中满足条件 $I_B^f(x_i) \cap X \neq \emptyset$ 的对象, 找出其中所有元素的 $I_B^f(x_i)$, 并求其并集, 得到的集合即为上近似集 X_B^f 。

6 实验验证

为了验证增量更新方法的有效性和适用性, 将其与增加对象和删除对象时直接计算粗糙集近似集的方法进行时间性能比较。实验环境 CPU 为 2.53GHz、操作系统 Window 7, 所有程序使用 C# 开发。实验数据选择 UCI 中 5 个数据集。实验数据集的基本特性如表 2 所列。表 2 中, E 表示原数据集中含有“?”或者“*”, R 表示在原数据集中用“?”或者“*”任意替换了其中部分数据(其中替换的数据约为十分之一)。

表 2 实验数据特性描述表

编号	数据集名称	记录数	属性数	?	*
1	lung-cancer	32	56	E	R
2	zoo	101	17	R	R
3	soybean	307	35	E	R
4	breast-cancer-wisconsin	699	10	R	R
5	contraceptive method choice	1473	9	R	R

当增加或删除单个对象时, 我们对直接计算方法和增量更新方法的性能利用上述 5 个数据集进行了比较, 其中集合 X 是在 5 个数据集中随机抽取的。图 1 是当增加对象($x_{n+1} \in X$)时, 两种方法计算 X 的上、下近似集的性能比较结果; 图 2 和图 3 是当删除对象时, 分别在 $x_n \in X$ 和 $x_n \notin X$ 两种情况下计算 X 的上、下近似集的性能比较结果。

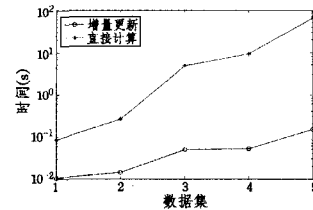


图 1 单个对象增加时, 直接计算与增量更新方法的性能比较

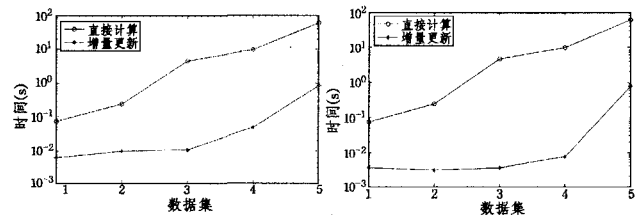


图 2 单个对象删除时且 $x_n \in X$, 直接计算与增量更新方法的性能比较

图 3 单个对象删除时且 $x_n \notin X$, 直接计算与增量更新方法的性能比较

从图 1—图 3 中的实验结果可以得出, 在不完备信息系统中, 增量更新方法比直接计算方法的性能要好, 前者所耗费的计算时间比后者少得多, 并验证了特性关系下粗糙集模型中增加或删除单个对象时增量更新方法在不同数据集应用中的有效性和适用性。

结束语 本文讨论了在特性关系粗糙集模型中单个对象增加和删除时近似集的变化情况。当增加单个对象时, 集合 X 不包含增加的对象, 所以集合 X 的上近似集变大, 下近似

(下转第 197 页)

希大大减少了消歧时语义密度的计算次数,从而降低了消歧的计算复杂度。最后在 SemCor 语料库上对本文提出的算法进行了测试。结果显示,本文提出的算法与同类算法相比有较好的性能。然而本文的算法仅针对名词消歧,如何处理名词以外的词以及 WordNet 中没有登记的词,则是本文所述算法没有解决的问题,有待进一步研究和改进。

参考文献

- [1] Turdakov D Y. Word Sense Disambiguation Methods [J]. Programming and Computer Software, 2010, 36(6): 309-326
- [2] Lesk M. Automatic sense disambiguation using machine readable dictionaries; how to tell a pine cone from an ice cream cone [C]// Proceedings of the 5th Annual International Conference on Systems Documentation (SIGDOC '86). New York, NY, USA: ACM, 1986: 24-26
- [3] Miller G A. Wordnet: A Lexical Database for English [J]. Communications of the ACM, 1995, 38(11): 39-41
- [4] Cao D D, Basili R, Luciani M, et al. Robust and Efficient Page Rank for Word Sense Disambiguation [C]// Proceedings of the 2010 Workshop on Graph-based Methods for Natural Language Processing. Uppsala, Sweden: ACL, 2010: 24-32
- [5] Liu Hong-fang, Teller V, Friedman C. A Multi-aspect Comparison Study of Supervised Word Sense Disambiguation [J]. Journal of the American Medical Informatics Association, 2004, 11(4): 320-331
- [6] Berkhin P. Survey of Clustering Data Mining Techniques [R]. CA, USA: Accrue Software, Inc., 2002
- [7] Miller G A, Leacock C, Teng R, et al. A Semantic Concordance [C]// Proceedings of the Workshop on Human Language Technology (HLT'93). Morristown, NJ, USA: Association for Computational Linguistics, 1993: 303-308
- [8] Roventini, Alonge, Bertagna, et al. ItalWordNet: Building a Large Semantic Database for the Automatic Treatment of Italian [J]. Linguistica Computazionale, 2003, 18: 745-791
- [9] Vossen P. EuroWordNet: building a multilingual database with wordnets for European languages [J]. The ELRA Newsletter, 1998, 3(1): 7-10
- [10] 于江生, 俞士汶. 中文概念词典的结构 [J]. 中文信息学报, 2002, 16(4): 12-20
- [11] 黄居仁, 谢舒凯, 洪嘉麒, 等. 中文词汇网络: 跨语言知识处理基础架构的设计理念与实践 [J]. 中文信息学报, 2010, 24(2): 14-23
- [12] Ravichandran D, Pantel P, Hovy E. Randomized algorithms and NLP: using locality sensitive hash function for high speed noun clustering [C]// Proceedings of the 43rd Annual Meeting of the ACL. Ann Arbor, USA: ACL, 2005: 622-629
- [13] Schmid H. Probabilistic Part-of-Speech Tagging Using Decision Trees [C]// Proceedings of the International Conference on New Methods in Language Processing. Manchester, UK, 1994
- [14] Agirre E, Rigau G. Word sense disambiguation using conceptual density [C]// Proceedings of the 16th Conference on Computational linguistics. Morristown, NJ, USA, 1996: 16-22

(上接第 193 页)

集变小。当删除单个对象时,如果删除的对象不属于集合 X , 则集合 X 的上近似集变小,下近似集变大;如果删除的对象属于集合 X ,则集合 X 的上、下近似集都变小。最后提出了动态增量更新算法,并给予了实验验证。对于对象集和属性集同时变化的研究将另文介绍。

参考文献

- [1] Pawlak Z. Rough sets [J]. International Journal of Computer and Information Sciences, 1982, 11: 341-356
- [2] Ziarko W. Variable precision rough set model [J]. Journal of Computer and System Sciences, 1993, 46: 39-59
- [3] Wang J, Xu L. Probabilistic rough set models [J]. Computer Science, 2002, 29(8): 76-78
- [4] Nanda S, Majumdar S. Fuzzy rough sets [J]. Fuzzy Sets and Systems, 1992, 45(2): 157-160
- [5] Stefanowski J, Tsoukias A. Incomplete information tables and rough classification [J]. Computational Intelligence, 2001, 17: 54-566
- [6] Kryszkiewicz M. Rough set approach to incomplete information systems [J]. Information Sciences, 1998, 112: 39-49
- [7] Grzymala-Busse J W. Data with missing attribute values; generalization of indiscernibility relation and rule induction [J]. Transactions on rough sets I, Lecture Notes in Computer Science, Springer-Verlag, 2004, 3100: 78-95
- [8] Grzymala-Busse J W. Characteristic relations for incomplete data; a generalization of the indiscernibility relation [C]// Proceedings of the Third International Conference on Rough Sets and Current Trends in Computing (RSCTC 2004). Lecture Notes in Artificial Intelligence, Springer-Verlag, 2004, 3066: 244-253
- [9] 刘伟斌, 李天瑞, 邹维丽, 等. 特性关系粗糙集下属性值粗化细化时近似集增量更新方法研究 [J]. 计算机科学, 2010, 37(6): 248-251
- [10] Grzymala-Busse J W. A new version of the rule induction system LERS [J]. Fundamenta Informaticae, 1997, 31: 27-39
- [11] Slowinski R, Vanderpooten D. A generalized definition of rough approximations based on similarity [J]. IEEE Transactions on Knowledge and Data Engineering, 2000, 12: 331-336
- [12] Li T, Ruan D, Geert W, et al. A rough sets based characteristic relation approach for dynamic attribute generalization in data mining [J]. Knowledge-based Systems, 2007, 20(5): 485-494
- [13] Song J, Li T, Ruan D. An integration of cloud transform and rough set theory to induction of decision trees [J]. Fundamenta Informaticae, 2009, 94(2): 261-273