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

刘伟斌[1] 李天瑞[1] 邹维丽[2] 胡成祥[1]
(西南交通大学信息科学与技术学院 成都610031)[1] (西南交通大学数学学院 成都610031)[2]

摘 要 不完备信息系统是一般信息系统的推广,在现实中具有广泛的应用。信息系统动态变化时,对象的近似集会产生相应的变化。研究如何利用原有近似集信息来进行近似集的更新具有重要意义。信息系统动态变化主要可以从属性值粗化细化、属性集粗化细化、对象集粗化细化3个方面考虑。现仅讨论属性值粗化细化时近似集的增量更新方法,给出了不完备信息系统中属性值粗化细化的定义,讨论了在不完备信息系统下的特性关系粗糙集模型中属性值粗化细化时近似集的增量更新方法,并通过实例验证了方法的有效性。

关键词 粗糙集,粒计算,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

Approaches for Incrementally Updating Approximations under Characteristic Relation-based Rough Sets While Attribute Values Coarsening and Refining

LIU Wei-bin[1] LI Tian-rui[1] ZOU Wei-li[2] HU Cheng-xiang[1]
(School of Information Science and Technology, Southwest Jiaotong University, Chengdu 610031, China)[1]
(School of Mathematics, Southwest Jiaotong University, Chengdu 610031, China)[2]

Abstract An incomplete information system is a generalization of an information system, which has wide applications in many domains. The approximations of concepts may vary when the information system is dynamically changing. It is meaningful to study how to update the approximations by using the original information. The dynamic changes of information systems can be considered in the three aspects: the coarsening and refining of attribute values, the attribute set and the object set. This paper only focused on the approaches for updating approximations when coarsening and refining attribute values. The definitions of attribute values coarsening and refining in the incomplete information system were given. Then approaches for updating the approximations were discussed when coarsening and refining attribute values under the characteristic relation. Finally, several examples were given to illustrate the validity of the proposed approaches.

Keywords Rough set, Granular computing, Incomplete information system, Characteristic relation, Incremental updating

1 引言

粗糙集理论[1,2](Rough Set Theory, RST)是由波兰科学家Pawlak于20世纪80年代初提出的一种处理含糊和不精确性问题的新型数学工具。近年来,为了适应实际工程应用的需求,已有许多的学者提出了各种各样的粗糙集模型,如可变精度粗糙集[3]、概率粗糙集[4]、模糊粗糙集[5]等,其中不完备信息系统[6-9](Incomplete Information System, IIS)中各种拓展粗糙集模型的应用已成为粗糙集理论发展的一个重要方面。对于IIS而言,数据的缺损可能是由两种原因造成的:一是在某种特殊的环节丢失的数据,比如在数据录入中遗漏的数据、因误操作而删除的数据,称这种缺损数据为丢失数据,用"?"表示。Stefanowski[7]等人构建了其中的相似关系(满足自反、传递性)来处理这种丢失数据。二是缺损数据对信息系统而言是不相关或不重要的数据,将其删除不会影响最终结果,称这种缺损数据为不关心数据,用"*"表示。Kryszkiewicz[6]提出了IIS中的容差关系(满足自反、对称性)来处理这种数据。但是,当IIS同时存在丢失数据和不关心数据的时候,相似关系和容差关系就不适用。因此,Grzymala-Busse提出了特性关系[10,11](Characteristic relation)。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,它可处理这种同时具有丢失型和不关心未知属性值的IIS。

经典粗糙集理论处理信息系统中的数据是精确的和不缺损的,即它不包含对原始不完全数据的处理。如今面对日益复杂和动态变化的数据,特别是当属性值动态变化(属性值粗化细化)时,经典粗糙集理论方法已经不能满足实际需求。然

到稿日期:2009-07-14 返修日期:2009-09-28 本文受国家自然科学基金《基于粒计算的动态知识发现中若干关键问题研究》(编号:60873108)资助。

刘伟斌(1983-),男,硕士生,主要研究方向为数据挖掘与知识发现等,E-mail:lwb-07@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。

而特性关系粗糙集凭借其同时处理丢失数据和不关心数据的优势，可为改进不完备信息系统中的数据挖掘技术提供强有力的理论支撑，有望实现从 IIS 中高效动态地获取规则。然而在对象集和属性集均不变、属性值变化的情形下，相当于对某一属性的取值范围进一步细化或粗化。此情形下基于特性关系粗糙集的粒计算模型的知识更新与维护问题至今还无人涉及。本文旨在讨论如何利用特性关系粗糙集理论来探讨属性值粗化细化情况下的近似集增量更新理论分析，以充分体现特性关系粗糙集在处理不确定性问题的优势，并完善粗糙集的理论与方法，这对拓展粗糙集的应用领域有十分重要的意义。

2 基本概念

定义 1[11] 在信息系统 $S=(U,A)$ 中，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 被称为 a 的域。如果存在空缺值 $a(x)$，$a\in A,x\in U$，则称 S 为不完备信息系统。系统中的空缺值分别用“?”或“ * ”表示，“?”表示丢失值，“ * ”表示不关心的值。

在文献[11]中，Grzymala-Busse 提出了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[11] 在不完备信息系统 $S=(U,A)$ 中，令 b 是属性，v 是 b 的一个属性取值，$t=(b,v)$ 是属性值对。如果 $v\neq?$ 和 $*$，那么可以按照 $[t]=\{x\in U\mid b(x)=v\}$ 形式定义 $[t]$。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 ($v\neq?$ 和 $*$)，都有 $x\notin[(b,v)]$；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($v\neq?$ 和 $*$)，都有 $x\in[(b,v)]$。

表 1　不完备信息系统

	Outlook	Temperature	Humidity	Windy
1	overcast	hot	?	not
2	*	hot	high	medium
3	sunny	*	normal	not
4	sunny	mild	high	medium
5	?	cool	*	?
6	rain	?	normal	not

例 1　表 1 是一个不完备信息系统。

定义 3[11] 在不完备信息系统 $S=(U,A)$ 中，令 $B\subseteq A$ 是一个属性子集，$x\in U$，那么特性集被定义为如下形式：

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

在表 1 中，设 $B=A$，由定义 3 可计算表 1 的特性集为：

$I_B^C(1)=\{1,2\}\cap\{1,2,3\}\cap\{1,3,6\}=\{1\}$

$I_B^C(2)=\{1,2,3\}\cap\{2,4,5\}\cap\{2,4\}=\{2\}$

$I_B^C(3)=\{2,3,4\}\cap\{3,5,6\}\cap\{1,3,6\}=\{3\}$

$I_B^C(4)=\{2,3,4\}\cap\{3,4\}\cap\{2,4,5\}\cap\{2,4\}=\{4\}$

$I_B^C(5)=\{3,5\}$

$I_B^C(6)=\{2,6\}\cap\{3,5,6\}\cap\{1,3,6\}=\{6\}$

在有了特性集的概念以后，下面给出特性关系的定义。

定义 4[12] 在不完备信息系统 $S=(U,A)$ 中，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：

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

例 2　在表 1 中，由定义 4 得各属性在 U 下的特性关系为：

$C_B=\{(1,1),(2,2),(3,3),(4,4),(5,5),(5,3),(6,6)\}$

显然，特性关系只具有自反性，并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下近似和上近似来描述。

定义 5[12] 在特性关系的意义下，相对于属性子集 B 而言，X 的下近似和上近似分别是：

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

$X_C^B=\bigcup\{I_B^C(x)\mid x\in X,I_B^C(x)\cap X\neq\emptyset\}=\bigcup\{I_B^C(x)\mid x\in X\}$

X 的边界域定义为 $B(X)=X_C^B-X_B^C$。

例 3　在表 1 中，令 $B=\{Temperature\}$，则可以计算在属性 B 下的特性集为：

$I_B^C(1)=I_B^C(2)=\{1,2,3\}$，$I_B^C(4)=\{3,4\}$，$I_B^C(5)=\{3,5\}$

令 $X=\{1,3,4,6\}$，则可计算出 X 关于属性 B 的上下近似集为：

$X_C^B=\{1,2,3,4,5\}$，$X_B^C=\{3,4\}$

再由上下近似集可以计算出 X 边界域为：

$B(X)=X_C^B-X_B^C=\{1,2,5\}$

定义 6[12] 特性关系的意义下，相对于属性子集 B 而言，X 的下边界集和上边界集分别是：

$\Delta X_B^C=X-X_B^C$，$\Delta X_C^B=X_C^B-X$

例 4　由例 3 中的计算结果，再根据定义 6 可以得到 X 的下边界集和上边界集分别为：

$\Delta X_B^C=X-X_B^C=\{1,6\}$，$\Delta X_C^B=X_C^B-X=\{2,5\}$

定义 7(属性值粗化)　设 $S=(U,A)$ 为不完备信息系统，$B\subseteq A,a_l\in B$，v_{il} 为对象 i 在属性 a_l 下的属性值，即 $v_{il}=f(i,a_l)$。v_{kl} 为对象 k 在属性 a_l 下的属性值，且 $v_{il}\neq v_{kl}$。记 $U_{a_l}=\{i'\in U\mid v_{il}=f(i',a_l)\}$，令 $f(i',a_l)=v_{kl}$，$\forall i'\in U_{a_l}$，则称属性值 v_{il} 被粗化为 v_{kl}。

为便于区分，记 $a_l^\wedge$ 为经粗化后的属性 a_l，$B^\wedge$ 为经粗化后的属性集 B，$V_{a_l}^\wedge$ 为属性 $a_l^\wedge$ 的值域。

定义 8(属性值细化)　设 $S=(U,A)$ 为不完备信息系统，$B\subseteq A,a_l\in B$，v_{il} 为对象 i 在属性 a_l 下的属性值，$U_{a_l}=\{i'\in U\mid v_{il}=f(i',a_l)\}$，若 $\exists v\notin V_l$，$\exists i'\in U_{a_l}$，令 $f(i',a_l)=v$，则称对象 i' 下的属性值 v_{il} 被细化为 v。

记 $a_l^\vee$ 为经细化后的属性 a_l，$B^\vee$ 为经细化后的属性集 B，$V_{a_l}^\vee$ 为属性 $a_l^\vee$ 的值域。

3 属性值粗化细化时近似集增量更新方法

3.1 单个属性的属性值粗化细化时近似集增量更新方法

引理 1　设 (U,A) 为不完备信息系统，$X\subseteq U,B\subseteq A,a_l\in B$，属性 a_l 的属性值粗化后，近似集的变化情况为：

$X_B^C\supseteq X_{B^\wedge}^C$，$X_C^B\subseteq X_C^{B^\vee}$

证明：根据属性值粗化的定义可得 $\forall(x,y)\in C_B$，有 $(x,y)\in C_{B^\wedge}$，即有 $I_B^C(x)\subseteq I_{B^\wedge}^C(x)$。

(1) $\forall x\in X_{B^\wedge}^C$，$I_{B^\wedge}^C(x)\subseteq X$，所以 $I_B^C(x)\subseteq X$，则 $x\in X_B^C$，即有 $X_B^C\supseteq X_{B^\wedge}^C$。

(2) $\forall x\in X_C^B$，$I_C^B(x)\cap X\neq\emptyset$。因为 $I_C^B(x)\subseteq I_C^{B^\wedge}(x)$，所以 $I_C^{B^\wedge}(x)\cap X\neq\emptyset$，则 $x\in X_C^B$，即有 $X_C^B\subseteq X_C^{B^\wedge}$。可得 $I_C^{B^\wedge}(x)\cap$

$X\neq\emptyset$,所以 $x\in X_C^B$,则 $X_C^B\subseteq X_C^{B^\wedge}$。

例 5 令 $B=\{Temperature\}$,$X=\{1,3,4,6\}$,则:

$I_B^C(1)=I_B^C(2)=\{1,2,3\}$,$I_B^C(4)=\{3,4\}$,$I_B^C(5)=\{3,5\}$,$X_B^C=\{3,4\}$,$X_C^B=\{1,2,3,4,5\}$

将属性 *Temperature* 下的属性值 *mild* 粗化成 *hot*,则:

$I_{B^\wedge}^C(1)=I_{B^\wedge}^C(2)=I_{B^\wedge}^C(4)=\{1,2,3,4\}$,$I_{B^\wedge}^C(5)=\{3,5\}$,$X_{B^\wedge}^C=\emptyset$,$X_C^{B^\wedge}=\{1,2,3,4,5\}$

推论 1 设 (U,A) 为不完备信息系统,$X\subseteq U$,$B\subseteq A$,$a_l\in B$,属性 a_l 的属性值粗化后,X 的边界域的变化情况为:

$B(X)\subseteq B^\wedge(X)$

推论 2 设 (U,A) 为不完备信息系统,$X\subseteq U$,$B\subseteq A$,$a_l\in B$,属性 a_l 的属性值粗化后,X 的边界集的变化情况为:

$\Delta X_B^C\subseteq\Delta X_{B^\wedge}^C$,$\Delta X_C^{B^\wedge}\supseteq\Delta X_C^B$

引理 2 设 (U,A) 为不完备信息系统,$X\subseteq U$,$B\subseteq A$,$a_l\in B$,属性 a_l 的属性值细化后,近似集的变化情况为:

$X_B^C\subseteq X_{B^\wedge}^C$,$X_C^B\supseteq X_C^{B^\vee}$

证明:根据属性值细化的定义可得 $\forall(x,y)\in C_{B^\vee}$,有 $(x,y)\in C_B$,即有 $I_{B^\vee}^C(x)\subseteq I_B^C(x)$。

(1) $\forall x\in X_B^C$,$I_B^C(x)\subseteq X$,所以 $I_{B^\vee}^C(x)\subseteq X$,则 $x\in X_{B^\vee}^C$,即有 $X_B^C\subseteq X_{B^\vee}^C$。

(2) $\forall x\in X_C^{B^\vee}$,$I_C^{B^\vee}(x)\cap X\neq\emptyset$。因为 $I_C^{B^\vee}(x)\subseteq I_C^B(x)$,所以 $I_C^B(x)\cap X\neq\emptyset$,则 $x\in X_C^B$。

例 6 令 $B=\{Temperature\}$,$X=\{1,3,4,6\}$,将对象 1 在属性 *Temperature* 下的属性值 *hot* 细化为 *very hot*,则:

$I_{B^\vee}^C(1)=\{1,3\}$,$I_{B^\vee}^C(2)=\{2,3\}$,$I_{B^\vee}^C(4)=\{3,4\}$

$I_{B^\vee}^C(5)=\{3,5\}$,$X_{B^\vee}^C=\{1,3,4\}\supseteq X_B^C$

$X_C^{B^\vee}=\{1,2,3,4,5\}\subseteq X_C^B$

推论 3 设 (U,A) 为不完备信息系统,$X\subseteq U$,$B\subseteq A$,属性 a_l 的属性值细化后,X 的边界域的变化情况为:

$B(X)\supseteq B^\vee(X)$

推论 4 设 (U,A) 为不完备信息系统,$X\subseteq U$,$B\subseteq A$,$a_l\in B$,属性 a_l 的属性值细化后,边界集的变化情况为:

$\Delta X_B^C\supseteq\Delta X_{B^\vee}^C$,$\Delta X_C^{B^\vee}\subseteq\Delta X_C^B$

定理 1 设 $B\subseteq A$,$a\in B$,则 $X_{B^\wedge}^C=X_B^C-\Delta X_{B^\wedge}^C$,其中

$\Delta X_{B^\wedge}^C=\{x\in\bigcap_{a\in B^\wedge}\Delta X_{\{a\}}^C\mid\bigcap_{a\in B^\wedge}X_{\{a\}}^C\not\subset X\}$

证明:因为 $X_{B^\wedge}^C\subseteq X_B^C$,所以 $\Delta X_B^C\subseteq\Delta X_{B^\wedge}^C$。又因为 $X_{B^\wedge}^C\cup\Delta X_{B^\wedge}^C=X=X_B^C\cup\Delta X_B^C$,所以

$X_{B^\wedge}^C=X_B^C\cup\Delta X_B^C-\Delta X_{B^\wedge}^C=X_B^C-\Delta X_{B^\wedge}^C$

例 7 令 $B=\{Outlook,Temperature\}$,$X=\{1,3,4,6\}$,其中 $P=\{Temperature\}$,$Q=\{Outlook\}$,将属性 *Temperature* 下的属性值 *mild* 粗化为 *hot*,则:

$X_B^C==\{3,4\}$,$X_{B^\wedge}^C=\emptyset$,$\Delta X_Q^C=\{1,3,4,6\}$

因为 $\bigcap_{a\in B^\wedge}I_B^C(1)=\{1,2\}\not\subset X$,$\bigcap_{a\in B^\wedge}I_B^C(3)=\{2,3,4\}\not\subset X$,$\bigcap_{a\in B^\wedge}I_B^C(4)=\{2,3,4\}\not\subset X$,$\bigcap_{a\in B^\wedge}I_B^C(6)=\{2,6\}\not\subset X$,所以 $\Delta X_{B^\wedge}^C=\{1,3,4,6\}$,则有:

$X_{B^\wedge}^C=X_B^C-\Delta X_{B^\wedge}^C=\{3,4\}-\{1,3,4,6\}=\emptyset$

定理 2 设 $B\subseteq A$,$a\in B$,则 $X_C^{B^\wedge}=X\cup\Delta X_C^B\cup Y$,其中 $Y=\{x\in\bigcap_{a\in B^\wedge}\Delta X_C^{\{a\}}\mid\bigcap_{a\in B^\wedge}I_{\{a\}}^C(x)\not\subset\bigcap_{a\in B^\wedge}\Delta X_C^{\{a\}}\}$。

证明:"⇒" $\forall x\in X_C^{B^\wedge}(X)$,$x\notin X$,则 $x\in\Delta X_C^{B^\wedge}$。因为 $\Delta X_C^B\subseteq\Delta X_C^{B^\wedge}$,若 $x\notin\Delta X_C^{B^\wedge}X$,下面证明 $x\in Y$。若 $\bigcap_{a\in B^\wedge}I_{\{a\}}^C(x)\subseteq\bigcap_{a\in B^\wedge}\Delta X_C^{\{a\}}$,因为 $\bigcap_{a\in B^\wedge}\Delta X_C^{\{a\}}\cap X=\emptyset$,所以 $\bigcap_{a\in B^\wedge}I_{\{a\}}^C(x)\cap X=\emptyset$,则有 $x\notin X_C^{a\in B^\wedge}(X)$,这与 $x\in X_C^{B^\wedge}$ 矛盾,则 $\bigcap_{a\in B^\wedge}I_{\{a\}}^C(x)\not\subset\bigcap_{a\in B^\wedge}\Delta X_C^{\{a\}}$,即 $x\in Y$。所以 $x\in X\cup\Delta X_C^{B^\wedge}\cup Y$,即有 $X_C^{B^\wedge}\subseteq X\cup\Delta X_C^{B^\wedge}\cup Y$。

"⇐"因为 $X\cup\Delta X_C^B\subseteq X\cup\Delta X_C^{B^\wedge}=X_C^{B^\wedge}$,所以只需证明 $Y\subseteq X_C^{B^\wedge}$。$\forall x\in Y$,若 $\bigcap_{a\in B^\wedge}I_{\{a\}}^C(x)\cap X=\emptyset$,则 $\bigcap_{a\in B^\wedge}I_{\{a\}}^C(x)\subseteq\Delta X_C^B$ $(\forall a\in B^\wedge)$,所以 $x\notin Y$,与 $x\in Y$ 矛盾,则有 $\bigcap_{a\in B^\wedge}I_{\{a\}}^C(x)\cap X\neq\emptyset$,即 $x\in X_C^{B^\wedge}$。所以 $Y\subseteq X_C^{B^\wedge}$,即 $X\cup\Delta X_C^B\cup Y\subseteq X_C^{B^\wedge}$。

例 8 令 $B=\{Outlook,Temperature\}$,$X=\{1,3,4,6\}$,其中 $P=\{Temperature\}$,$Q=\{Outlook\}$,将属性 *Temperature* 下的属性值 *mild* 粗化 *hot*,则 $X\cup\Delta X_C^B=X_C^B=\{1,2,3,4,5,6\}$,$X_C^{P^\wedge}=\{1,2,3,4,5\}$,$X_C^Q=\{1,2,3,4,6\}$,所以 $\Delta X_C^{P^\wedge}=\{2,6\}$,$\Delta X_C^Q=\{2\}$,则有 $\bigcap_{a\in B^\wedge}\Delta X_C^{\{a\}}=\Delta X_C^{P^\wedge}\cap\Delta X_C^Q=\{2\}$。

因为 $\bigcap_{a\in B^\wedge}I_{\{a\}}^C(2)=\{1,2,3,4\}\not\subset\bigcap_{a\in B^\wedge}\Delta X_C^{\{a\}}$,所以 $Y=\{2\}$,则 $X_C^{B^\wedge}=X\cup\Delta X_C^B\cup Y=\{1,2,3,4,5,6\}\cup\{2\}=\{1,2,3,4,5,6\}$。

定理 3 设 $B\subseteq A$,$a\in B$,则 $X_{B^\vee}^C=X_B^C\cup X_{\{a^\vee\}}^C\cup Z$,这里

$Z=\{x\in\Delta X_B^C\cap\Delta X_{\{a^\vee\}}^C\mid\bigcap_{a\in B^\vee}I_{\{a\}}^C(x)\subseteq X\}$

证明:"⇒" $\forall x\in X_{B^\vee}^C$,$\bigcap_{a\in B^\vee}I_{\{a\}}^C(x)\subseteq X$。若 $x\notin X_B^C\cup X_{\{a^\vee\}}^C$,则 $x\in\Delta X_B^C\cap\Delta X_{\{a^\vee\}}^C$。所以 $x\in Z$,即有 $X_{B^\vee}^C\subseteq X_B^C\cup X_{\{a^\vee\}}^C\cup Z$。

"⇐"因为 $X_B^C\subseteq X_{B^\vee}^C$,$X_{\{a^\vee\}}^C\subseteq X_{B^\vee}^C$,$Z\subseteq X_{B^\vee}^C$,所以 $X_B^C\cup X_{\{a^\vee\}}^C\cup Z\subseteq X_{B^\vee}^C$。

例 9 令 $B=\{Outlook,Temperature\}$,$X=\{1,3,4,6\}$,其中 $P=\{Temperature\}$,$Q=\{Outlook\}$,将对象 1 在属性 *Temperature* 下的属性值 *hot* 细化为 *very hot*,则 $X_B^C=\{3,4\}$,$X_{P^\vee}^C=\{1,3,4\}$,$\Delta X_B^C=X-X_B^C=\{1,3,4,6\}-\{3,4\}=\{1,6\}$,$\Delta X_{P^\vee}^C=X-X^C_{P^\vee}=\{1,3,4,6\}-\{1,3,4\}=\{6\}$,所以 $\Delta X_B^C\cap\Delta X_{\{a^\vee\}}^C=\Delta X_B^C\cap\Delta X_{P^\vee}^C=\{1,6\}\cap\{6\}=\{6\}$。因为对象 6 在属性 P 下的属性值为?,所以 $Z=\emptyset$,则有 $X_{B^\vee}^C=X_B^C\cup X_{P^\vee}^C\cup Z=\{3,4\}\cup\{1,3,4\}\cup\emptyset=\{1,3,4\}$。

定理 4 设 $B\subseteq A$,$a\in B$,则 $X_C^{B^\vee}=X_C^B\cap(X_C^{\{a^\vee\}}\cup X)-Z'$,这里 $Z'=\{x\in\Delta X_C^B\cap\Delta X_C^{\{a^\vee\}}\mid\bigcap_{a\in B^\vee}I_{\{a\}}^C(x)\subseteq\bigcap_{a\in B^\vee}\Delta X_C^{\{a\}}\}$。

证明:"⇒" $\forall x\in X_C^{B^\vee}$,有

$$\bigcap_{a\in B^\vee}I_{\{a\}}^C(x)\cap X\neq\emptyset \tag{1}$$

若 $x\notin X$,则 $x\in\Delta X_C^{B^\vee}X\subseteq\Delta X_C^B$。同理可得 $x\in\Delta X_C^{\{a^\vee\}}$,所以 $x\in\Delta X_C^{B^\vee}\cap\Delta X_C^{\{a^\vee\}}$。

下面证明 $x\notin Z'$。由于

$$\left(\bigcap_{a\in B^\vee}\Delta X_C^{\{a\}}\right)\cap X=\emptyset \tag{2}$$

由式(1),式(2)得 $\bigcap_{a\in B^\vee}I_{\{a\}}^C(x)\not\subset\bigcap_{a\in B^\vee}\Delta X_C^{\{a\}}$,因此 $x\notin Z'$。则 $x\notin X\cup(\Delta X_C^{B^\vee}\cap\Delta X_C^{\{a^\vee\}}-Z')=X_C^B\cap(X_C^{\{a^\vee\}}\cup X)-Z'$,所以 $X_C^{B^\vee}\subseteq X_C^B\cap(X_C^{\{a^\vee\}}\cup X)-Z'$。

"⇐" $\forall x\in X_C^B\cap(X_C^{\{a^\vee\}}\cup X)-Z'$,$x\notin X$。若 $x\notin X_C^{B^\vee}$,

则 $\bigcap_{a\in B^\vee} I^C_{\{a\}}(x)\cap X=\emptyset$。所以 $\bigcap_{a\in B^\vee} I^C_{\{a\}}(x)\subseteq \Delta X_C^{\{a\}}X(\forall a\in B^\vee)$。即有 $\bigcap_{a\in B^\vee} I^C_{\{a\}}(x)\subseteq \bigcap_{a\in B^\vee}\Delta X_C^{\{a\}}$，所以 $x\in Z'$。这与 $x\in X_C^B\cap(X_C^{\{a^\vee\}}\cup X)-Z'$ 矛盾。则 $x\in X_C^{B^\vee}$，即 $X_C^{B^\vee}\supseteq X_C^B\cap(X_C^{\{a^\vee\}}\cup X)-Z'$。

综上可得，$X_C^{B^\vee}=X_C^B\cap(X_C^{\{a^\vee\}}\cup X)-Z'$。

例 10　令 $B=\{Outlook, Temperature\}$，$X=\{1,3,4,6\}$，其中 $P=\{Temperature\}$，$Q=\{Outlook\}$，将对象 1 在属性 *Temperature* 下的属性值 *hot* 细化为 *very hot*，则 $X_C^B=\{1,2,3,4,5,6\}$，$X_C^{P^\vee}=\{1,2,3,4,5\}$，所以 $\Delta X_C^B=X_C^B-X=\{2,5\}$，$\Delta X_C^{P^\vee}=X_C^{P^\vee}-X=\{2,5\}$，$\Delta X_C^Q=X_C^Q-X=\{1,2,3,4,6\}-\{1,3,4,6\}=\{2\}$，则 $\Delta X_C^B\cap\Delta X_C^{P^\vee}=\{2,5\}$，$\bigcap_{a\in B^\vee}\Delta X_C^{\{a\}}=\Delta X_C^{P^\vee}\cap\Delta X_C^Q=\{2\}$。

另外，由于 $\bigcap_{a\in B^\vee} I^C_{\{a\}}(2)=\{2,3\}\not\subset\bigcap_{a\in B^\vee}\Delta X_C^{\{a\}}$，$\bigcap_{a\in B^\vee} I^C_{\{a\}}(5)=\{3,5\}\not\subset\bigcap_{a\in B^\vee}\Delta X_C^{\{a\}}$，因此 $Z'=\emptyset$，则有 $X_C^{B^\vee}=X_C^B\cap(X_C^{\{P^\vee\}}\cup X)-Z'=\{1,2,3,4,5,6\}$。

以上定理和实例都说明了在特性关系粗糙集模型下单个属性的属性值粗化细化时近似集的变化情况。在现实生活中，多个属性的属性值同时粗化细化是非常常见的，因此以下研究多个属性的属性值粗化细化时近似集的变化情况。

3.2　多个属性的属性值粗化细化时近似集增量更新方法

引理 3　设(U,A)为集值信息系统，$X\subseteq U$，$B\subseteq A$，属性集 B 的属性值粗化后，近似集的变化情况为 $X^C_{B^\wedge}\subseteq X^C_B$，$X_C^{B^\wedge}\supseteq X_C^B$。

例 11　令 $B=\{Outlook, Temperature\}$，$X=\{1,3,4,6\}$，将属性 *Outlook* 下的属性 *overcast* 粗化成 *sunny*，*Temperature* 下的属性值 *mild* 粗化成 *hot*，则：

$I^C_{B^\wedge}(1)=I^C_{B^\wedge}(2)=I^C_{B^\wedge}(3)=I^C_{B^\wedge}(4)=\{1,2,3,4\}$，$I^C_{B^\wedge}(5)=\{3,5\}$，$I^C_{B^\wedge}(6)=\{2,6\}$，$X^C_{B^\wedge}=\emptyset\subseteq X^C_B$，$X_C^{B^\wedge}=\{1,2,3,4,5,6\}\supseteq X_C^B$。

推论 5　设(U,A)为不完备信息系统，$X\subseteq U$，$B\subseteq A$，属性集 B 的属性值粗化后，X 的边界域的变化情况为 $B(X)\subseteq B^\wedge(X)$。

推论 6　设(U,A)为不完备信息系统，$X\subseteq U$，$B\subseteq A$，属性集 B 的属性值粗化后，X 的边界集的变化情况为 $\Delta X^C_B\subseteq\Delta X^C_{B^\wedge}$，$\Delta X_C^{B^\wedge}\supseteq\Delta X_C^B$。

引理 4　设(U,A)为不完备信息系统，$X\subseteq U$，$B\subseteq A$，属性集 B 的属性值细化后，近似集的变化情况为 $X^C_B\subseteq X^C_{B^\wedge}$，$X_C^B\supseteq X_C^{B^\vee}$。

例 12　令 $B=\{Outlook, Temperature\}$，$X=\{1,3,4,6\}$，将对象 1 在属性 *Outlook* 下的属性值 *overcast* 细化成 *fairness*，*Temperature* 下的属性值 *hot* 细化为 *very hot*，则：

$I^C_{B^\vee}(1)=\{1\}$，$I^C_{B^\vee}(2)=I^C_{B^\vee}(3)=\{2,3\}$，$I^C_{B^\vee}(3)=I^C_{B^\vee}(4)=\{3,4\}$，$I^C_{B^\vee}(5)=\{3,5\}$，$I^C_{B^\vee}(6)=\{2,6\}$，$X^C_{B^\vee}=\{1,3,4\}\supseteq X^C_B$，$X_C^{B^\vee}=\{1,2,3,4,5,6\}\subseteq X_C^B$。

推论 7　设(U,A)为不完备信息系统，$X\subseteq U$，$B\subseteq A$，属性集 B 的属性值细化后，X 的边界域的变化情况为 $B(X)\supseteq B^\vee(X)$。

推论 8　设(U,A)为不完备信息系统，$X\subseteq U$，$B\subseteq A$，属性集 B 的属性值细化后，X 边界集的变化情况为 $\Delta X^C_B\supseteq\Delta X^C_{B^\vee}$，$\Delta X_C^{B^\vee}\subseteq\Delta X_C^B$。

定理 5　设 $B\subseteq A$，$C\subseteq B$，则 $X^C_{B^\wedge}=X^C_B-\Delta X^C_{B^\wedge}$，$X_C^{B^\wedge}=X\cup\Delta X_C^B\cup Y$，其中

$$\Delta X^C_{B^\wedge}=\{x\in\bigcap_{a\in B^\wedge}\Delta X^C_{\{a\}}\mid\bigcap_{a\in B^\wedge}X^C_{\{a\}}\not\subset X\}$$

$$Y=\{x\in\bigcap_{a\in B^\wedge}\Delta X_C^{\{a\}}\mid\bigcap_{a\in B^\wedge}I^C_{\{a\}}(x)\not\subset\bigcap_{a\in B^\wedge}\Delta X_C^{\{a\}}\}$$

定理 6　设 $B\subseteq A$，$C\subseteq B$，则 $X^C_{B^\vee}=X^C_B\cup X^C_{\{a_l^\vee\}}\cup Z$，$X_C^{B^\vee}=X_C^B\cap(X_C^{\{a_l^\vee\}}\cup X)-Z'$，这里

$$Z=\{x\in\Delta X^C_B\cap\Delta X^C_{\{a_l^\vee\}}\mid\bigcap_{a\in B^\vee}I^C_{\{a\}}(x)\subseteq X\}$$

$$Z'=\{x\in\Delta X_C^B\cap\Delta X_C^{\{a_l^\vee\}}\mid\bigcap_{a\in B^\vee}I^C_{\{a\}}(x)\subseteq\bigcap_{a\in B^\vee}\Delta X_C^{\{a\}}\}$$

结束语　本文仅讨论了特性关系下粗糙集模型中的属性值粗化细化时近似集的变化情况。当属性值粗化后，上近似集变大，下近似集变小，因此边界变大；当属性值细化后，上近似集变小，下近似集变大，因此边界变小。今后的工作将讨论在不完备决策信息系统中，特性关系下粗糙集模型里属性值粗化细化、属性集粗化细化和对象集粗化细化时增量式规则的获取方法。

参 考 文 献

[1] Pawlak Z. Rough set theory and its applications to data analysis [J]. Cybernetics and Systems, 1998, 29: 661-688

[2] Pawlak Z. Rough sets and intelligent data analysis[J]. Information Sciences, 2002, 147: 1-12

[3] Wojciech Z. Variable precision rough set model[J]. Journal of Computer and System Sciences, 1993, 46: 39-59

[4] Wang J, Xu L. Probabilistic rough set models[J]. Computer Science, 2002, 29(8): 76-78

[5] Nanda S, Majumdar S. Fuzzy rough sets[J]. Fuzzy Sets and Systems, 1992, 45(2): 157-160

[6] Kryszkiewicz M. Rough set approach to incomplete information systems[J]. Information Sciences, 1998, 112: 39-49

[7] Stefanowski J, Tsoukias A. Incomplete information tables and rough classification[J]. Computational Intelligence, 2001, 17: 54-56

[8] Wu W, Zhang W, Li H. Knowledge acquisition in incomplete fuzzy information systems via the rough set approach[J]. Expert Systems, 2003, 20(5): 280-286

[9] Wang G Y. Extension of rough set under incomplete information systems[J]. Journal of Computer Research and Development, 2002, 39(10): 1238-1243

[10] Grzymala-Busse J W. Data with missing attribute values: generalization of indiscernibility relation and rule induction[C] // Transactions on rough sets I, Lecture Notes in Computer Science. vol. 3100, Berlin: Springer-Verlag, 2004: 78-95

[11] Grzymala-Busse J W. Characteristic relations for incomplete data: a generalization of the indiscernibility relation[C] // Proceeding of the Third International Conference on Rough Sets and Current Trends in Computing (RSCTC 2004). Lecture Notes in Artificial Intelligence. vol. 3066, Berlin: Springer-Verlag, 2004: 244-253

[12] Li T, Ruan D, Greet W, et al. A rough set based characteristic relation approach for dynamic attribute generalization in data mining[J]. Knowledge-Based Systems, 2007, 20(2): 485-494