合成信息系统与子信息系统*)

邵明文　张文修　米据生

（西安交通大学理学院信息与系统科学研究所　西安710049）

摘　要　本文给出了对象合成信息系统、属性合成信息系统、对象子信息系统及属性子信息系统的定义，分别讨论了它们的上下近似算子与原信息系统的上下近似算子之间的关系，并给出了它们的一些实际应用。

关键词　粗糙集，信息系统，下近似，上近似

Combined Information Systems and Subinformation Systems

SHAO Ming-Wen　ZHANG Wen-Xiu　MI Ju-Sheng

(Institute of Information and System Science, Xi'an Jiaotong University, Xi'an 710049)

Abstract　In this paper, by giving definition of object combined information systems, attribute combined information systems, sub-object information systems and sub-attibute information systems, discuss the relationship between them and the original information systems respectively ,at last giving some applications.

Keywords　Rough set, Information systems, Lower approximation, Upper approximation

1　引言

粗糙集[1]理论是一种新的处理不确定知识的数学工具，其主要思想是，在保持信息系统的分类能力不变的前提下，通过知识约简，导出问题的决策或分类规则。近20年来，粗糙集理论已经在理论和应用上取得了长足的发展，特别是由于20世纪90年代在知识发现等领域得到了成功的应用而受到国际学术界广泛关注，基于 Rough 集方法的知识发现与知识约简被许多专家所研究。目前，它正在被广泛应用于机器学习、决策分析、过程控制、模式识别和数据挖掘等领域[2~4]。

对一个目标信息系统(U,A,F,D,G)来说，其知识库是确定的，可以按照不同的要求来进行知识发现与规则提取[5~9]。有时在实际应用中，我们在原目标信息系统的基础上添加部分对象，增加或删除部分对决策有重要影响或很小影响的条件属性，分别得到新的目标信息系统，这时新目标信息系统与原目标信息系统在知识库、知识发现 、规则提取等方面的关系是一个有意义的问题。C-C. Chan 在文[10]中给出了 Rough 集近似的属性递归算法，本文的属性合成信息系统与子属性信息系统两部分中利用和推广了其结果。本文讨论了合成信息系统和子信息系统与原信息系统的上、下近似算子之间的关系。

2　对象合成信息系统

定义2.1　称(U,A,F)为一个信息系统，或者数据库系统。

其中U为有限对象集，即：$U=\{x_1,x_2,\cdots,x_n\}$，U中的每个x_i，$(i\leqslant n)$，称为一个对象。而A为属性集，即$A=\{a_1,a_2,\cdots,a_m\}$，A中的每个a_j，$(j\leqslant m)$，称为一个属性。F为U和A的关系集。即$F=\{f_j:U\rightarrow V_j;j\leqslant m\}$，其中$V_j$为属性$a_j$的有限值域。

定义2.2　设(U_1,A,F_1)和(U_2,A,F_2)为两个信息系统，称$(U_1\cup U_2,A,F_1\cup F_2)$为对象合成信息系统，其中：

$$F_1=\{f_l^1:U_1\rightarrow V_l^1;l\leqslant m\},$$

$$F_2=\{f_l^2:U_2\rightarrow V_l^2;l\leqslant m\},$$

$$F_1\cup F_2=\{f_l^1\cup f_l^2:U_1\cup U_2\rightarrow V_l^1\cup V_l^2;l\leqslant m\},$$

$$(f_l^1\cup f_l^2)(x)=\begin{cases}f_l^1(x) & x\in U_1\\ f_l^2(x) & x\in U_2\end{cases}$$

易证：$[x]^{U_1\cup U_2}=[x]^{U_1}\cup[x]^{U_2}$。

在对象合成信息系统中，我们总是把对象集U_1与U_2中的对象看作是不同的，即：$U_1\cup U_2=\emptyset$。

定理2.1　设$(U_1\cup U_2,A,F_1\cup F_2)$为信息系统$(U_1,A,F_1)$和$(U_2,A,F_2)$的对象合成信息系统，$X=X_1\cup X_2\subseteq U_1\cup U_2$，其中$X_1\subseteq U_1$，$X_2\subseteq U_2$，则：

$$\underline{R}^{U_1\cup U_2}(X)=\underline{R}^{U_1\cup U_2}(X_1\cup X_2)=(\underline{R}^{U_1}(X_1)-Z_1)\cup(\underline{R}^{U_2}(X_2)-Z_2),$$

其中：$Z_1=\{x\in U_1\mid[x]_A^{U_1}\subseteq X_1,[x]_A^{U_2}\not\subseteq X_2\}$，

$Z_2=\{x\in U_2\mid[x]_A^{U_2}\subseteq X_2,[x]_A^{U_1}\not\subseteq X_2\}$。

证明：(⇒)：设$x\in\underline{R}^{U_1\cup U_2}(X_1\cup X_2)$，则$[x]_A^{U_1\cup U_2}\subseteq X_1\cup X_2\Rightarrow[x]_A^{U_1}\subseteq X_1,[x]_A^{U_2}\subseteq X_2\Rightarrow x\in(\underline{R}^{U_1}(X_1)-Z_1)$或$x\in(\underline{R}^{U_2}(X_2)-Z_2)\Rightarrow x\in(\underline{R}^{U_1}(X_1)-Z_1)\cup(\underline{R}^{U_2}(X_2)-Z_2)$，因此，$\underline{R}^{U_1\cup U_2}(X_1\cup X_2)\subseteq(\underline{R}^{U_1}(X_1)-Z_1)\cup(\underline{R}^{U_2}(X_2)-Z_2)$

(⇐)：设$x\in(\underline{R}^{U_1}(X_1)-Z_1)\cup(\underline{R}^{U_2}(X_2)-Z_2)$，则：$x\in(\underline{R}^{U_1}(X_1)-Z_1)$或$x\in(\underline{R}^{U_2}(X_2)-Z_2)\Rightarrow[x]_A^{U_1}\subseteq X_1,[x]_A^{U_2}\subseteq X_2\Rightarrow[x]_A^{U_1\cup U_2}=[x]_A^{U_1}\cup[x]_A^{U_2}\subseteq X_1\cup X_2\Rightarrow x\in\underline{R}^{U_1\cup U_2}(X_1\cup X_2)\Rightarrow(\underline{R}^{U_1}(X_1)-Z_1)\cup(\underline{R}^{U_2}(X_2)-Z_2)\subseteq\underline{R}^{U_1\cup U_2}(X_1\cup X_2)$。 □

定理2.2　设$(U_1\cup U_2,A,F_1\cup F_2)$为信息系统$(U_1,A,F_1)$和$(U_2,A,F_2)$的对象合成信息系统，$X=X_1\cup X_2\subseteq U_1\cup U_2$，其中$X_1\subseteq U_1$，$X_2\subseteq U_2$，则：

$$\overline{R}^{U_1\cup U_2}(X)=\overline{R}^{U_1\cup U_2}(X_1\cup X_2)=\overline{R}^{U_1}(X_1)\cup\overline{R}^{U_2}(X_2)\cup Z_1\cup Z_2。$$

其中：$Z_1=\{x\in U_1\mid[x]_A^{U_1}\cap X_1=\emptyset,[x]_A^{U_2}\cap X_2\neq\emptyset\}$

$Z_2=\{x\in U_2\mid[x]_A^{U_2}\cap X_2=\emptyset,[x]_A^{U_1}\cap X_1\neq\emptyset\}$。

证明：(⇐)设$x\in\overline{R}^{U_1}(X_1)\cup\overline{R}^{U_2}(X_2)\cup Z_1\cup Z_2$，则$x\in\overline{R}^{U_1}(X_1)$或$x\in\overline{R}^{U_2}(X_2)$或$x\in Z_1$或$x\in Z_2\Rightarrow[x]_A^{U_1}\cap X_1\neq\emptyset$或$[x]_A^{U_2}\cap X_2\neq\emptyset$或$[x]_A^{U_1\cup U_2}\cap X_2\neq\emptyset$或$[x]_A^{U_1\cup U_2}\cap X_1\neq\emptyset\Rightarrow$

*）本文得到国家973项目基金资助(No. 2002CB312200)。**邵明文**　博士生，主要研究方向为粗糙集、神经网络、代数系统。**张文修**　教授，博导，主要研究方向为粗糙集、遗传算法、模糊集、人工智能等。**米据生**　教授，博士生，主要研究方向为粗糙集、随机集、近似推理。

$[x]_A^{U_1\cup U_2}\cap X_1\neq\emptyset$ 或 $[x]_A^{U_1\cup U_2}\cap X_2\neq\emptyset$ 或 $[x]_A^{U_1\cup U_2}\cap X_2\neq\emptyset$ 或 $[x]_A^{U_1\cup U_2}\cap X_1\neq\emptyset\Rightarrow[x]_A^{U_1\cup U_2}\cap(X_1\cup X_2)\neq\emptyset\Rightarrow x\in\overline{R}^{U_1\cup U_2}(X)\Rightarrow\overline{R}^{U_1}(X_1)\cup\overline{R}^{U_2}(X_2)\cup Z_1\cup Z_2\subseteq\overline{R}^{U_1\cup U_2}(X)$。

$(\Rightarrow)$设 $x\in\overline{R}^{U_1\cup U_2}(X_1\cup X_2)$，则$[x]_A^{U_1\cup U_2}\cap(X_1\cup X_2)\neq\emptyset\Rightarrow([x]_A^{U_1}\cup[x]_A^{U_2})\cap(X_1\cup X_2)\neq\emptyset\Rightarrow[x]_A^{U_1}\cap X_1\neq\emptyset$ 或$[x]_A^{U_2}\cap X_2\neq\emptyset$ 或$[x]_A^{U_1}\cap X_2\neq\emptyset$ 或$[x]_A^{U_2}\cap X_1\neq\emptyset\Rightarrow x\in\overline{R}^{U_1}(X_1)$或 $x\in\overline{R}^{U_2}(X_2)$或 $x\in Z_1$或 $x\in Z_2\Rightarrow x\in\overline{R}^{U_1}(X_1)\cup\overline{R}^{U_2}(X_2)\cup Z_1\cup Z_2\Rightarrow\overline{R}^{U_1\cup U_2}(X_1\cup X_2)\subseteq\overline{R}^{U_1}(X_1)\cup\overline{R}^{U_2}(X_2)\cup Z_1\cup Z_2$。 □

例2.1 表1,2,3给出了两个信息系统和它们的对象合成信息系统。

表1 信息系统(U_1,A,F_1)

U	a_1	a_2	a_3
x_1	2	1	3
x_2	3	2	1
x_3	2	1	3
x_4	1	1	4
x_5	1	1	2

表2 信息系统(U_2,A,F_2)

U	a_1	a_2	a_3
x_1^*	1	1	4
x_2^*	1	2	3
x_3^*	1	2	3
x_4^*	3	2	1

表3 对象合成信息系统$(U_1\cup U_2,A,F_1\cup F_2)$

U	a_1	a_2	a_3
x_1	2	1	3
x_2	3	2	1
x_3	2	1	3
x_4	1	1	4
x_5	1	1	2
x_6	1	1	4
x_7	1	2	3
x_8	1	2	3
x_9	3	2	1

令 $X=\{x_2,x_4,x_7,x_9\}$，则 $X_1=\{x_2,x_4\}\subseteq U_1$，$X_2=\{x_7,x_9\}\subseteq U_2$。

$\underline{R}^{U_1}(X_1)=\{x\in U_1\mid[x]_A^{U_1}\subseteq X_1\}=\{x_2,x_4\}$

$\underline{R}^{U_2}(X_2)=\{x\in U_2\mid[x]_A^{U_2}\subseteq X_2\}=\{x_9\}$

$Z_1=\{x\in U_1\mid[x]_A^{U_1}\subseteq X_1,[x]_A^{U_2}\not\subseteq X_2\}=\{x_4\}$

$Z_2=\{x\in U_2\mid[x]_A^{U_2}\subseteq X_2,[x]_A^{U_1}\not\subseteq X_1\}=\emptyset$

则：$\underline{R}^{U_1\cup U_2}(X)=(\underline{R}^{U_1}(X_1)-Z_1)\cup(\underline{R}^{U_2}(X_2)-Z_2)=\{x_2,x_9\}$

$\overline{R}^{U_1}(X_1)=\{x_2,x_4\}$，$\overline{R}^{U_2}(X_2)=\{x_7,x_8,x_9\}$

$Z_1=\{x\in U_1\mid[x]_A^{U_1}\cap X_1=\emptyset,[x]_A^{U_2}\cap X_2\neq\emptyset\}=\emptyset$，

$Z_2=\{x\in U_2\mid[x]_A^{U_2}\cap X_2=\emptyset,[x]_A^{U_1}\cap X_1\neq\emptyset\}=\{x_6\}$

$\overline{R}^{U_1\cup U_2}(X)=\overline{R}^{U_1}(X_1)\cup\overline{R}^{U_2}(X_2)\cup Z_1\cup Z_2=\{x_2,x_4,x_6,x_7,x_8,x_9\}$

3 属性合成信息系统

定义3.1 设(U,A,F_1)和(U,B,F_2)为信息系统，其中 U 为有限对象集，A,B 为互不相同属性集，称$(U,A\cup B,F_1\cup F_2)$为属性合成信息系统。

其中：U 为对象集合，$U=\{x_1,x_2,\cdots,x_n\}$；A,B 为属性集，$A=\{a_1,a_2,\cdots,a_m\}$，$B=\{a_1^*,a_2^*,\cdots,a_h^*\}$；$A\cup B=\{a_1,a_2,\cdots,a_m,a_1^*,a_2^*,\cdots,a_h^*\}$；$F_1$是 U 和 A 的关系集，$F_1=\{f_k^1:k\leqslant m\}$，$V_k$ 是 a_k 的有限值域；F_2是 U 和 B 的关系集，$F_2=\{f_k^2:k\leqslant h\}$，$V_k^*$ 是 a_k^* 的有限值域；$F_1\cup F_2=\{f_l:l\leqslant m+h\}$，$f_l=$

$$\begin{cases} f_l^1 & l\leqslant m, \\ f_l^2 & l>m。\end{cases}$$

设(U,A,F)为信息系统，记$\underline{A}(X)$为 X 关于属性集 A 的等价关系 R_A 下的下近似，$\overline{A}(X)$为 X 关于属性集 A 的等价关系 R_A 下的上近似，记：$\underline{\Delta}_A(X)=X-\underline{A}(X)$，$\overline{\Delta}_A(X)=\overline{A}(X)-X$。

性质3.1 设(U,A,F)为信息系统，$B\subseteq A$，$X\subseteq U$，则：(1)$\underline{\Delta}_A(X)\subseteq\underline{\Delta}_B(X)$；(2)$\overline{\Delta}_A(X)\subseteq\overline{\Delta}_B(X)$。

证明(略)

定理3.1 设$(U,A\cup B,F_1\cup F_2)$为信息系统(U,A,F_1)和(U,B,F_2)的属性合成信息系统，对于任意 $X\subseteq U$，有：$\underline{A\cup B}(X)=\underline{A}(X)\cup\underline{B}(X)\cup Y$。其中：$Y=\{x\in\underline{\Delta}_A(X)\cap\underline{\Delta}_B(X)\mid[x]_A\cap[x]_B\subseteq X\}$。

证明：首先，显然有 $\underline{A}(X)\cup\underline{B}(X)\subseteq\underline{A\cup B}(X)$。从而 $x\notin\underline{A}(X)\cup\underline{B}(X)$且 $x\in X$ 当且仅当 $x\in(X-\underline{A}(X))\cap X-\underline{B}(X)=\underline{\Delta}_A(X)\cap\underline{\Delta}_B(X)$。则证。

其次，$x\in\underline{A\cup B}(X)$当且仅当$[x]_{A\cup B}\subseteq X$，从而$[x]_{A\cup B}=[x]_A\cap[x]_B\subseteq X$。则证。 □

定理3.2 设$(U,A\cup B,F_1\cup F_2)$为信息系统(U,A,F_1)和(U,B,F_2)的属性合成信息系统，则对于任意 $X\subseteq U$，有：

$\overline{A\cup B}(X)=X\cup(\overline{\Delta}_A(X)\cap\overline{\Delta}_B(X)-Z)$。

其中：$Z=\{x\in\overline{\Delta}_A(X)\cap\overline{\Delta}_B(X)\mid[x]_{A\cup B}\subseteq\overline{\Delta}_A(X)\cap\overline{\Delta}_B(X)\}$。

证明：因为$\overline{A\cup B}(X)=X\cup\overline{\Delta}_{A\cup B}(X)$，所以只须证：$\overline{\Delta}_{A\cup B}(X)=\overline{\Delta}_A(X)\cap\overline{\Delta}_B(X)-Z$。由性质3.1(2)：$\overline{\Delta}_{A\cup B}(X)\subseteq\overline{\Delta}_A(X)\cap\overline{\Delta}_B(X)$，对任意 $x\in\overline{\Delta}_{A\cup B}(X)\Rightarrow x\in\overline{\Delta}_A(X)\cap\overline{\Delta}_B(X)$且$[x]_{A\cup B}\cap X\neq\emptyset\Rightarrow x\in\overline{\Delta}_A(X)\cap\overline{\Delta}_B(X)$且$[x]_{A\cup B}\not\subseteq\overline{\Delta}_A(X)\cap\overline{\Delta}_B(X)\Rightarrow x\in\overline{\Delta}_A(X)\cap\overline{\Delta}_B(X)$且 $x\notin Z\Rightarrow x\in(\overline{\Delta}_A(X)\cap\overline{\Delta}_B(X)-Z)\Rightarrow\overline{\Delta}_{A\cup B}(X)\subseteq\overline{\Delta}_A(X)\cap\overline{\Delta}_B(X)-Z$。反之，对任意 $x\in\overline{\Delta}_A(X)\cap\overline{\Delta}_B(X)-Z\Rightarrow x\notin X$ 且$[x]_{A\cup B}\not\subseteq\overline{\Delta}_A(X)\cap\overline{\Delta}_B(X)\Rightarrow x\notin X$ 且$[x]_{A\cup B}\not\subseteq\overline{\Delta}_{A\cup B}(X)\Rightarrow x\in\overline{A}(X)-X\Rightarrow x\in\overline{\Delta}_{A\cup B}(X)\Rightarrow\overline{\Delta}_A(X)\cap\overline{\Delta}_B(X)-Z\subseteq\overline{\Delta}_{A\cup B}(X)$。 □

例3.1 设$(U,A\cup B,F_1\cup F_2)$为信息系统(U,A,F_1)和(U,B,F_2)的属性合成信息系统，其中：

$U=\{x_1,x_2,\cdots,x_8\}$，$A=\{a,b\}$，$B=\{c,d\}$，

$\mathscr{A}(A)=\{\{x_1,x_3\},\{x_2,x_4\},\{x_5,x_6\},\{x_7,x_8\}\}$，

$\mathscr{B}(B)=\{\{x_1,x_3,x_5,x_6\},\{x_2,x_4\},\{x_7,x_8\}\}$.

令 $X=\{\{x_2,x_5,x_6\}\}$，则：

(1)：$\underline{A}(X)=\{x_5,x_6\}$，$\underline{B}(X)=\emptyset$。

$\underline{\Delta}_A(X)=X-\underline{A}(X)=\{x_2\}$，$\underline{\Delta}_B(X)=X-\underline{B}(X)=\{x_2,x_5,x_6\}$。

$Y=\{x\in\underline{\Delta}_A(X)\cap\underline{\Delta}_B(X)\mid[x]_A\cap[x]_B\subseteq X\}=\emptyset$。

因此：$\underline{A\cup B}(X)=\underline{A}(X)\cup\underline{B}(X)\cup Y=\{x_5,x_6\}$。

(2)：$\overline{A}(X)=\{x_2,x_4,x_5,x_6\}$，$\overline{B}(X)=\{x_1,x_3,x_5,x_6,x_2,x_4\}$。

$\overline{\Delta}_A(X)=\overline{A}(X)-X=\{x_4\}$，$\overline{\Delta}_B(X)=\overline{B}(X)-X=\{x_1,x_3,x_4\}$。

$\overline{\Delta}_A(X)\cap\overline{\Delta}_B(X)=\{x_4\}$，$\{x_4\}_A\cap[x_4]_B=\{x_2,x_4\}\not\subseteq\overline{\Delta}_A(X)\cap\overline{\Delta}_B(X)$，得 $Z=\emptyset$。

所以：$\overline{A\cup B}(X)=X\cup(\overline{\Delta}_A(X)\cap\overline{\Delta}_B(X)-\emptyset)=\{x_2,x_4,x_5,x_6\}$。

4 子对象信息系统

定义4.1 设(U,A,F)为信息系统，称信息系统(U_1,A,F_1)为(U,A,F)的子对象信息系统。其中：$U_1\subseteq U$，$F_1=\{f_l^1:U_1\rightarrow V_l:l\leqslant m\}$且当 $x\in U_1$时，任意 $f_l\in F$，有 $f_l(x)=f_l^1(x)(l\leqslant m)$。

定理4.1 设(U_1,A,F_1)为(U,A,F)的子对象信息系统，

则对于任意 $X\subseteq U_1$ 有:$\underline{A}^{U_1}(X)=\underline{A}^{U}(X)\cup Z_1$。其中:$Z_1=\{x\in U_1\mid[x]_A^{U_1}\subseteq X,[x]_A^{U}\not\subseteq X\}$。

证明:$x\in\underline{A}^{U_1}(X)\Leftrightarrow[x]_A^{U_1}\subseteq X\Leftrightarrow[x]_A^{U_1}\subseteq X,[x]_A^{U}\subseteq X$ 或 $[x]_A^{U_1}\subseteq X,[x]_A^{U}\not\subseteq X\Leftrightarrow x\in\underline{A}^{U}(X)$ 或 $x\in Z_1$。 □

定理4.2 设(U_1,A,F_1)为(U,A,F)的子对象信息系统,则对于任意 $X\subseteq U_1$,有:$\overline{A}^{U_1}(X)=\overline{A}^{U}(X)-U_2$。其中:$U_2=U-U_1$。

证明:(⇒)设 $X\in\overline{A}^{U_1}(X)$,则$[x]_A^{U_1}\cap X\neq\emptyset\Rightarrow[x]_A^{U}\cap X\neq\emptyset$ 且 $x\notin U_2\Rightarrow x\in\overline{A}^{U}(X)-U_2\Rightarrow\overline{A}^{U_1}(X)\subseteq\overline{A}^{U}(X)-U_2$。

(⇐)设 $x\in\overline{A}^{U}(X)-U_2$,则 $x\in U_1$ 且$[x]_A^{U}\cap X\neq\emptyset\Rightarrow x\in U_1,([x]_A^{U_1}\cup[x]_A^{U_2})\cap X\neq\emptyset\Rightarrow x\in U_1,([x]_A^{U_1}\cap X)\cup([x]_A^{U_2}\cap X)\neq\emptyset\Rightarrow x\in U_1,[x]_A^{U_1}\cap X\neq\emptyset\Rightarrow x\in\overline{A}^{U_1}(X)\Rightarrow\overline{A}^{U}(X)-U_2\subseteq\overline{A}^{U_1}(X)$。 □

5 子属性信息系统

定义5.1 设(U,A,F)为信息系统,称信息系统(U,B,F_1)为(U,A,F)的子属性信息系统,若 $B\subseteq A$,$F_1\subseteq F$。其中:U 是对象集合,$U=\{x_1,x_2,\cdots,x_n\}$;A,B 为属性集,$B=\{a_1,a_2,\cdots,a_k\}$,$A=\{a_1,a_2,\cdots,a_k,\cdots,a_m\}$;$F$ 是 U 和 A 的关系集,$F=\{f_l;l\leqslant m\}$,V_l 是 a_l 的有限值域;F_1是 U 和 B 的关系集,$F_1=\{f_l:l\leqslant k\}$,V_l 是 a_l 的有限值域。

定理5.1 设(U,B,F_1)为(U,A,F)的子属性信息系统,则对于任意 $X\subseteq U$ 有:

$$\underline{B}(X)=\underline{A}(X)-\underline{\Delta}_B(X)=\underline{A}(X)-\{x\in X\mid[x]_B\not\subseteq X\}。$$

证明:由于 $\underline{B}(X)\subseteq\underline{A}(X)\subseteq X$,因此:$\underline{\Delta}_B(X)\subseteq\underline{\Delta}_A(X)$;而 $\underline{A}(X)\cap\underline{\Delta}_A(X)=\emptyset$,于是 $\underline{B}(X)=\underline{A}(X)-(\underline{\Delta}_B(X)-\underline{\Delta}_A(X))=\underline{A}(X)-\underline{\Delta}_B(X)$。

又因:$x\in\underline{\Delta}_B(X)\Leftrightarrow x\in X$ 且$[x]_B\not\subseteq X$,
所以:$\underline{B}(X)=\underline{A}(X)-\underline{\Delta}_B(X)=\underline{A}(X)-\{x\in X\mid[x]_B\not\subseteq X\}$。□

定理5.2 设(U,B,F_1)为(U,A,F)的子属性信息系统,则对于任意 $X\subseteq U$ 有:$\overline{B}(X)=\overline{A}(X)\cup Z$。其中:$Z=\{x\in(\bigcap_{b\in B}\overline{\Delta}_{\{b\}}(X)-\bigcap_{a\in A}\overline{\Delta}_{\{a\}}(X))\mid[x]_B\not\subseteq\bigcap_{b\in B}\overline{\Delta}_{\{b\}}(X)\}$。

证明:由于 $\overline{B}(X)=X\cup\overline{\Delta}_B(X)$,因此,当 $x\in\overline{B}(X)$,且 $x\notin X$ 时,必有 $x\in\overline{\Delta}_B(X)$。又由于 $x\in\overline{\Delta}_B(X)$当且仅当 $x\in\bigcap_{b\in B}\overline{\Delta}_{\{b\}}(X)$且$[x]_B\not\subseteq\bigcap_{b\in B}\overline{\Delta}_{\{b\}}(X)$。

容易证明:$\overline{\Delta}_B(X)-\overline{\Delta}_A(X)\subseteq\bigcap_{b\in B}\overline{\Delta}_{\{b\}}(X)-\bigcap_{a\in A}\overline{\Delta}_{\{a\}}(X)$。

由于 $\overline{\Delta}_B(X)\subseteq\overline{\Delta}_A(X)$,所以 $x\in\overline{\Delta}_B(X)$,$x\notin\overline{\Delta}_A(X)$当且仅当 $x\in\bigcap_{b\in B}\overline{\Delta}_{\{b\}}(X)-\bigcap_{a\in A}\overline{\Delta}_{\{a\}}(X)$且$[x]_B\not\subseteq\bigcap_{b\in B}\overline{\Delta}_{\{b\}}(X)$,即 $x\in X$,所以 $\overline{\Delta}_B(X)=\overline{\Delta}_A(X)\cup Z$。因此 $\overline{B}(X)=X\cup\overline{\Delta}_A(X)\cup Z=\overline{A}(X)\cup Z$。 □

例5.1 设(U,B,F_1)为(U,A,F)的子属性信息系统,$U=\{x_1,x_2,\cdots,x_8\}$,$A=\{a,b\}$,$B=\{b\}$,且 $\mathscr{A}(a)=\{\{x_1,x_3\},\{x_2,x_4,x_5,x_6\},\{x_7,x_8\}\}$,$\mathscr{A}(b)=\{\{x_1,x_2,x_3,x_4,x_5,x_6\},\{x_7,x_8\}\}$。

令:$X=\{x_1,x_3,x_5,x_6\}$,

(1):由于 $\underline{A}(X)=\{x_1,x_3\}$,$\underline{\Delta}_A(X)=\{x_5,x_6\}$,$Y=\underline{\Delta}_B(X)=\{x_1,x_3,x_5,x_6\}$,

于是:$\underline{B}(X)=\underline{A}(X)-Y=\emptyset$。

(2):$\overline{A}(X)=\{x_1,x_2,x_3,x_4,x_5,x_6\}$,$\overline{\Delta}_A(X)=\{x_2,x_4\}$,$Z=\{x\in(\bigcap_{b\in B}\overline{\Delta}_{\{b\}}(X)-\bigcap_{a\in A}\overline{\Delta}_{\{a\}}(X))\mid[x]_B\not\subseteq\bigcap_{b\in B}\overline{\Delta}_{\{b\}}(X)\}=\{x_2,x_4\}$,从而:$\overline{B}(X)=X\cup\overline{\Delta}_A(X)\cup Z=\{x_1,x_2,x_3,x_4,x_5,x_6\}$。

6 应用

在不协调目标信息系统中,我们增加条件属性直至信息系统是协调的,这时我们得到确定的命题知识。

例6.1 一个关于某些病人的决策如表4,其中 $U=\{x_1,x_2,\cdots,x_8\}$,$A=\{$头痛,肌肉痛$\}$,$D=\{$流感$\}$。

表4 目标信息系统

U	头痛	肌肉痛	流感
x_1	是	是	否
x_2	是	否	是
x_3	是	是	是
x_4	否	是	否
x_5	否	否	否
x_6	否	是	是
x_7	否	否	是
x_8	否	是	否

表5 属性合成信息系统

U	头痛	肌肉痛	体温	流感
x_1	是	是	正常	否
x_2	是	是	高	是
x_3	是	是	很高	是
x_4	否	是	正常	否
x_5	否	否	高	否
x_6	否	是	很高	是
x_7	否	否	高	否
x_8	否	是	很高	是

令 $a_1=$ 头痛,$a_2=$ 肌肉痛,则:

$U/\{a_1\}=\{\{x_1,x_2,x_3\},\{x_4,x_5,x_6,x_7,x_8\}\}$,

$U/\{a_2\}=\{\{x_1,x_2,x_3,x_4,x_6,x_8\},\{x_5,x_8\}\}$。

$U/D=\{\{x_1,x_4,x_5,x_7\},\{x_2,x_3,x_6,x_8\}\}$,

$U/\{a_1,a_2\}=\{\{x_1,x_2,x_3\},\{x_4,x_6,x_8\},\{x_5,x_7\}\}$。

显然,上表是一个不协调目标信息系统,如果在上表中增加条件属性 $a_3=\{$体温$\}$,则得到决策表5。这时,$U/\{a_3\}=\{\{x_1,x_4\},\{x_2,x_5,x_7\},\{x_3,x_6,x_8\}\}$,令:$A=\{a_1,a_2\}$,$B=\{a_3\}$。

对任意 $X_i\subseteq U/D$,计算 $\underline{B}(X_i)$,$\overline{B}(X_i)$,$\overline{\Delta}_B(X_i)$,$\underline{\Delta}_B(X_i)$,($i\leqslant 2$)。计算两个近似算子$\underline{A\cup B}(X_i)$,$\overline{A\cup B}(X_i)$,因为$\underline{A\cup B}(X_i)=\overline{A\cup B}(X_i)$,($i\leqslant 2$)。所以该属性合成目标信息系统是协调的,我们可以得到命题知识。

结语 本文由对象合成信息系统、属性合成信息系统、对象子信息系统及属性子信息系统的定义,得到了合成信息系统和子信息系统的上、下近似算子与原信息系统的上、下近似算子之间的关系,利用它们的关系可以由原信息系统计算合成信息系统和属性子信息系统的上、下近似算子。

参考文献

1 Pawlak Z. Rough sets. International Journal of Computer and Information Science, 1982, 11: 341～356
2 Pawlak Z. Rough Sets Thoery and It's Application to Data Analysis [J]. Cybernetics Systems, An International Journal, 1998, 29: 661～688
3 An A, et al. Applying knowledge discovery to predict water-supply consumption [J]. IEEE Expert, 1997, 72～78
4 Pawlak Z. Slowinski R. Rough set approach to multiattribute decision analysis, invited review [J]. European Journal of Operational Research, 1994, 72: 443～459
5 Wong S K M, Ziarko W. On optional decision rules in decision tables [J]. Bulletin of Polish Academy of Science, 1985, 33: 693～696
6 Hu X H, Cercone N. Learning in relational database: a rough set approach [J]. Computational Intelligence, 1995, 11(2): 323～338
7 Marzena K. Comparative study of alternative types of knowledge reduction in inconsistent systems [J]. International J of Intelligent Systems, 2001, 16: 105～120
8 常梨云,王国胤,吴渝. 一种基于 Rough Set 理论的属性约简及规则提取方法. 软件学报, 1999, 10(11): 1206～1211
9 张文修,米据生,吴伟志. 不协调目标信息系统的知识约简. 计算机学报, 2003, 26(1): 12～18
10 Chan C C. A rough set approach to attribution generalization in data mining [J]. Information science, 1998, 107: 169～176