文章编号: 1001-7402(2011)02-0146-10

基于最小描述交的覆盖广义粗糙集*

张晓燕[1], 徐伟华[1], 张文修[2]

(1. 重庆理工大学 数学与统计学院, 重庆　400054;
2. 西安交通大学 理学院, 陕西 西安　710049)

摘　要: 针对 Bonikowski 覆盖广义粗糙集模型的不足, 给出了基于最小描述交的覆盖上下近似算子。通过和 Pawlak 经典粗糙集以及 Bonikowski 的覆盖广义粗糙集比较, 发现给出的覆盖上、下近似算子具有了对偶关系, 并得到了相关重要性质; 进一步讨论了在新定义下覆盖广义粗糙集的约简和公理化问题, 丰富了覆盖广义粗糙集理论, 并为覆盖广义粗糙集的应用提供了更确切的理论根据。

关键词: Pawlak 粗糙集; Bonikowski 覆盖近似; 覆盖约简; 公理化方法
中图分类号: TP18　　　文献标识码: A

1　引言

粗糙集理论是波兰数学家 Z. Pawlak 于 20 世纪 80 年代初提出的用于数据分析的理论[13]。由于粗糙集理论能够分析处理不精确、不协调和不完备信息, 因此作为一种具有极大潜力和非常有效的知识获取工具受到了人工智能工作者的广泛关注[1-3,7-12]。目前, 粗糙集理论已被成功地应用在人工智能、数据挖掘、模式识别、智能信息处理等领域, 该理论已发展成为计算机科学以及信息科学的研究热点之一[14-22,25]。

经典粗糙集理论中基本的概念和运算都是通过代数学的等价关系和集合运算来定义的。然而现实生活中遇到的很多论域并不是基于等价关系的。为此, Zakowski 将等价关系对论域的划分推广为覆盖, 从而建立了覆盖广义粗集理论[23], 而且该模型被认为是在数据挖掘中具有广泛应用前景的模型, 并有许多学者对其进行了较深入的讨论[4-6], 给出了覆盖广义粗集的一些基本结果。遗憾的是 Zakowski 提出的覆盖广义粗集的覆盖上近似过于粗糙; 而且其给出的上近似与下近似不具有对偶性。另外, 文献[26]中虽然说明了上下近似算子是相互确定的, 也给出了下近似运算的公理, 但没能给出上近似的公理刻画, 也没给出上下近似算子之间具体的数量关系。

针对上述不足, 我们在 Z. Bonikowski 覆盖广义粗糙集模型的基础上, 对覆盖近似算子进行了修改, 给出了基于最小描述交的覆盖广义近似算子, 使得覆盖上、下近似算子具有了对偶关系, 并讨论了在新定义的覆盖广义粗糙集的约简和公理化问题, 给出了相关重要结论, 进一步丰富了覆盖广义粗糙集模型理论。

* 收稿日期: 2009-09-20; 修订日期: 2010-06-01
基金项目: 重庆理工大学青年科研项目
作者简介: 张晓燕(1979-), 女, 山西怀仁人, 重庆理工大学数学与统计学院讲师, 研究方向: 粗糙集, 模糊集; 徐伟华(1979-), 男, 山西浑源人, 重庆理工大学数学与统计学院副教授, 博士, 研究方向: 不确定性推理及应用, 人工智能的数学基础; 张文修(1940-), 男, 山西翼城人, 西安交通大学理学院教授, 博士生导师, 研究方向: 模糊集, 粗糙集, 人工智能的数学基础。

2 Z. Bonikowski 覆盖广义粗糙

本节简要介绍Z. Bonikowski 覆盖广义粗糙近似及其相关性质。

定义 **2. 1**[4-6]　设 U 是有限论域, $\{\mathscr{C}_i\}(i=1,2,\cdots,n)$ 是 U 上的子集族。若 $\mathscr{C}_i\neq\varnothing$ 且 $\bigcup_i\mathscr{C}_i=U$, 则称 $\{\mathscr{C}_i\}$ 是论域 U 的一个覆盖。称二元组$(U,\mathscr{C})$为覆盖广义近似空间显然, 论域 U 上由等价关系形成的划分也是一个覆盖, 覆盖是划分的推广。

设 $S=(U,\mathscr{C})$ 为一覆盖广义近似空间。对 $u\in U$, 称 $Md(u)=\{K\in\mathscr{C}\mid u\in K\wedge(\forall H\in\mathscr{C}\wedge u\in H\wedge H\subseteq K\Rightarrow K=H)\}$ 为 u 的最小描述。

类似于 Pawlak 近似空间, 为了用已知的知识去解释未知的知识, 须引入上、下近似的概念。

Z. Bonikowski 覆盖广义粗糙近似如下:

定义 **2. 2**[4-6]　对集合 $X\subseteq U$,

集族$\underline{S\mathscr{C}}(X)=\{K\in\mathscr{C}\mid K\subseteq X\}$称为 X 的覆盖下近似。

集族集合$\underline{\mathscr{C}}(X)=\bigcup\underline{S\mathscr{C}}(X)$称为 X 的覆盖下近似。

集族$Bn(X)=\bigcup\{Md(x)\mid x\in X-\underline{\mathscr{C}}(X)\}$称为 X 的覆盖边界集族。

集族$\overline{S\mathscr{C}}(X)=\underline{S\mathscr{C}}(X)\cup Bn(X)$称为 X 的覆盖上近似集族。

集合$\overline{\mathscr{C}}(X)=\bigcup\overline{S\mathscr{C}}(X)$称为 X 的覆盖上近似。

如果$\underline{\mathscr{C}}(X)=\overline{\mathscr{C}}(X)$, 则称 X 为精确的, 否则为粗糙的。

注 **2. 1**　对照定义 2. 3 的下近似可以自然的想到覆盖的另外一种上近似的定义, 即

$$\bigcap\{K\in\mathscr{C}\mid X\subseteq K\}$$

需要说明的是这样定义的上近似和 Bonikowski 的不同。

例如, 取论域为 $U=\{a,b,c,d\}$ 及其覆盖 $\mathscr{C}=\{\{a,b\},\{a,c,d\}\}$。

若记: $\mathscr{C}^{\#}(X)=\bigcap\{K\in\mathscr{C}\mid X\subseteq K\}$, 并取 $X_1=\{a\}$, 则有 $\mathscr{C}^{\#}(X_1)=\{a,b\}\cap\{a,c,d\}=\{a\}$。

Bonikowski 定义的覆盖下近似为$\underline{\mathscr{C}}(X_1)=\varnothing$.

于是 $X_1-\underline{\mathscr{C}}(X_1)=\{a\}$, 及 $Md(a)=\{\{a,b\},\{a,c,d\}\}$, 故有 $Bn(X_1)=U$.

因此 Bonikowski 定义的覆盖上近似为$\overline{\mathscr{C}}(X_1)=U$.

显然, $\overline{\mathscr{C}}(X_1)\neq\mathscr{C}^{\#}(X)$。

命题 **2. 1**[4-6]　Bonikowski 覆盖近似集族$\underline{S\mathscr{C}}(X)$和$\overline{S\mathscr{C}}(X)$有以下性质:

(1) $\underline{S\mathscr{C}}(\varnothing)=\overline{S\mathscr{C}}(\varnothing)=\varnothing$, $\underline{S\mathscr{C}}(U)=\overline{S\mathscr{C}}(U)=U$

(2) $\underline{S\mathscr{C}}(X)\subseteq\overline{S\mathscr{C}}(X)$

(3) $\underline{S\mathscr{C}}(\underline{\mathscr{C}}(X))=\underline{S\mathscr{C}}(X)=\overline{S\mathscr{C}}(\underline{\mathscr{C}}(X))$

(4) $X\subseteq Y\Rightarrow\underline{S\mathscr{C}}(X)\subseteq\underline{S\mathscr{C}}(Y)$

命题 **2. 2**[4-6]　如果 $\mathscr{C}$ 是一个划分, 则$\underline{\mathscr{C}}(X)$和$\overline{\mathscr{C}}(X)$是 Pawlak 上、下近似。

命题 **2. 3**[4-6]　设 $\mathscr{C}$ 是 U 的覆盖, 其上、下近似有下面性质:

(1L) $\underline{\mathscr{C}}(X)\subseteq X$　(下近似的收缩性)

(1U) $X\subseteq\overline{\mathscr{C}}(X)$　(上近似的扩张性)

(3L) $\underline{\mathscr{C}}(\varnothing)=\varnothing$　(下近似的正规性)

(3U) $\overline{\mathscr{C}}(\varnothing)=\varnothing$　(上近似的正规性)

(4L) $\underline{\mathscr{C}}(U)=U$　(下近似的余正规性)

(4U) $\overline{\mathscr{C}}(U)=U$　(上近似的余正规性)

(6L) $X\subseteq Y\Rightarrow\underline{\mathscr{C}}(X)\subseteq\underline{\mathscr{C}}(Y)$　(下近似的单调性)

(7L) $\underline{\mathscr{C}}(\underline{\mathscr{C}}(X)) = \underline{\mathscr{C}}(X)$ (下近似的幂等性)

(7U) $\overline{\mathscr{C}}(\overline{\mathscr{C}}(X)) = \overline{\mathscr{C}}(X)$ (上近似的幂等性)

命题 **2.4**[4-6]　设 $\mathscr{C}$ 是 U 的覆盖, 其上、下近似的下面性质不再成立。

(2) $\underline{\mathscr{C}}(\sim X) = \sim\overline{\mathscr{C}}(X)$ (上下近似的对偶性)

$\overline{\mathscr{C}}(\sim X) = \sim\underline{\mathscr{C}}(X)$ (上下近似的对偶性)

(5L) $\underline{\mathscr{C}}(X\cap Y) = \underline{\mathscr{C}}(X)\cap\underline{\mathscr{C}}(Y)$ (下近似的可乘性)

(5U) $\overline{\mathscr{C}}(X\cup Y) = \overline{\mathscr{C}}(X)\cup\overline{\mathscr{C}}(Y)$ (上近似的可加性)

(6U) $X\subseteq Y\Rightarrow\overline{\mathscr{C}}(X)\subseteq\overline{\mathscr{C}}(Y)$ (上近似的单调性)

由上面命题 2.4 可以看出 Bonikowski 定义的覆盖上下近似算子不是对偶的一对算子, 这就为很多问题的研究带来不便, 故有必要寻找更好的更方便的覆盖上下近似算子。

3　基于最小描述交的覆盖广义粗糙近似算子

本节给出覆盖广义粗糙集的新近似算子, 并讨论其性质, 进而和 Pawlak 经典粗糙集的上下近似以及 Bonikowski 定义的覆盖上近似做了对比。

定义 3.1　设 U 为一非空论域, $\mathscr{C}$ 为 U 上的覆盖。对任意的 $X\subseteq U$, X 关于覆盖近似空间 $(U,\mathscr{C})$ 的上下、近似分别定义为:

$$\mathscr{C}_*(X) = \{x\in U \mid (\cap Md(x))\subseteq X\}$$
$$\mathscr{C}^*(X) = \{x\in U \mid (\cap Md(x))\cap X\neq\varnothing\}$$

集合 $\mathscr{C}_*(X)$ 称为 X 关于覆盖近似空间 $(U,\mathscr{C})$ 的正域, 记作: $pos_{\mathscr{C}}(X)$。

集合 $\sim\mathscr{C}^*(X)$ 称为 X 关于覆盖近似空间 $(U,\mathscr{C})$ 的负域, 记作: $neg_{\mathscr{C}}(X)$。

集合 $\mathscr{C}^*(X)-\mathscr{C}_*(X)$ 称为 X 关于覆盖近似空间 $(U,\mathscr{C})$ 的边界, 记作: $bn_{\mathscr{C}}(X)$。

当 $\mathscr{C}_*(X)=\mathscr{C}^*(X)$ 时, 称 X 为精确的, 否则为粗糙的。

定理 3.1　设 $\mathscr{C}$ 是 U 的一个划分, 则 $\mathscr{C}_*(X)$ 和 $\mathscr{C}^*(X)$ 分别退化为 Pawlak 上、下近似进一步, 上述覆盖近似算子有如下结论。

证明　由定义 3.1 直接可得。

定理 3.2　设 U 为一非空论域, $\mathscr{C}$ 为 U 上的覆盖。覆盖下近似 $\mathscr{C}_*$ 和上近似 $\mathscr{C}^*$ 有下面性质:

(1L) $\mathscr{C}_*(X)\subseteq X$ (下近似的收缩性)

(1U) $X\subseteq\mathscr{C}^*(X)$ (上近似的扩张性)

(2) $\mathscr{C}_*(\sim X) = \sim\mathscr{C}^*(X)$ (上下近似的对偶性)

$\mathscr{C}^*(\sim X) = \sim\mathscr{C}_*(X)$ (上下近似的对偶性)

(3L) $\mathscr{C}_*(\varnothing) = \varnothing$ (下近似的正规性)

(3U) $\mathscr{C}^*(\varnothing) = \varnothing$ (上近的正规性)

(4L) $\mathscr{C}_*(U) = U$ (下近似的余正规性)

(4U) $\mathscr{C}^*(U) = U$ (上近似的余正规性)

(5L) $\mathscr{C}_*(X\cap Y) = \mathscr{C}_*(X)\cap\mathscr{C}_*(Y)$ (下近似的可乘性)

(5U) $\mathscr{C}^*(X\cup Y) = \mathscr{C}^*(X)\cup\mathscr{C}^*(Y)$ (上近似的可加性)

(6L) $X\subseteq Y\Rightarrow\mathscr{C}_*(X)\subseteq\mathscr{C}_*(Y)$ (下近似的单调性)

(6U) $X\subseteq Y\Rightarrow\mathscr{C}^*(X)\subseteq\mathscr{C}^*(Y)$ (上近似的单调性)

(7L) $\mathscr{C}_*(\mathscr{C}_*(X)) = \mathscr{C}_*(X)$ (下近似的幂等性)

(7U) $\mathscr{C}^*(\mathscr{C}^*(X)) = \mathscr{C}^*(X)$ (上近似的幂等性)

(9L) $\forall K \in \mathscr{C}$, $\mathscr{C}_*(K) = K$ 　　(下近似的颗粒性)

证明　(1L) 对于任意的 $x \in \mathscr{C}_*(X)$, 有 $\cap Md(x) \subseteq X$. 又 $x \in \cap Md(x)$, 于是得: $x \in X$. 因此有 $\mathscr{C}_*(X) \subseteq X$.

(1U) 对于任意的 $x \in X$, 有 $(\cap Md(x)) \cap X \neq \varnothing$. 故有 $x \in \mathscr{C}^*(X)$, 即 $X \subseteq \mathscr{C}^*(X)$。

(2) 设 $x \in \mathscr{C}_*(X)$, 则 $x \in \mathscr{C}_*(X) \Leftrightarrow \cap Md(x) \subseteq X \Leftrightarrow (\cap Md(x)) \cap \sim X = \varnothing \Leftrightarrow x \neq \mathscr{C}^*(\sim X)$, 于是有 $\mathscr{C}_*(X) = \sim \mathscr{C}^*(\sim X)$。因此, $\mathscr{C}^*(\sim X) = \sim \mathscr{C}_*(X)$。

另一个同上可得。

(3L) 由(1L)和(1U)知 $\mathscr{C}_*(\varnothing) \subseteq \varnothing$. 故有 $\varnothing \subseteq \mathscr{C}_*(\varnothing)$。因此可得 $\mathscr{C}_*(\varnothing) = \varnothing$.

(3U) 若 $\mathscr{C}^*(\varnothing) \neq \varnothing$, 则必存在一个 $x \in \mathscr{C}^*(\varnothing)$。因此, $(\cap Md(x)) \cap \varnothing \neq \varnothing$. 但显然 $(\cap Md(x)) \cap \varnothing = \varnothing$, 矛盾。故有: $\mathscr{C}^*(\varnothing) = \varnothing$.

(4L) 由(2)和(3L)直接可得 $\mathscr{C}_*(U) = \mathscr{C}_*(\sim \varnothing) = \sim \mathscr{C}^*(\varnothing) = \sim \varnothing = U$.

(4U) 由(4L)直接获证。

(5L) 设 $x \in \mathscr{C}_*(X \cap Y)$, 则

$$
\begin{aligned}
& x \in \mathscr{C}_*(X \cap Y) \\
\Leftrightarrow & (\cap Md(x)) \subseteq X \cap Y \\
\Leftrightarrow & (\cap Md(x)) \subseteq X \text{ 且 } (\cap Md(x)) \subseteq Y \\
\Leftrightarrow & x \in \mathscr{C}_*(X) \text{ 且 } x \in \mathscr{C}_*(Y) \\
\Leftrightarrow & x \in \mathscr{C}_*(X) \cap x \in \mathscr{C}_*(Y)
\end{aligned}
$$

故有: $\mathscr{C}_*(X \cap Y) = \mathscr{C}_*(X) \cap \mathscr{C}_*(Y)$。

(5U) 由(2)和(5L)可得。

(6L) 设 $x \in \mathscr{C}_*(X)$, 则 $\cap Md(x) \subseteq X$. 又 $X \subseteq Y$, 于是 $\cap Mdx \subseteq X \subseteq Y$, 即 $x \in \mathscr{C}_*(Y)$。故可有: $\mathscr{C}_*(X) \subseteq \mathscr{C}_*(Y)$。因此, $X \subseteq Y \Rightarrow \mathscr{C}_*(X) \subseteq \mathscr{C}_*(Y)$ 成立。

(6U) 由(6L)可得。

(7L) 对于任意的 $x \in \mathscr{C}_*(X)$, 有 $\cap Md(x) \subseteq X$. 又由(6L)可有 $\mathscr{C}_*(\cap Md(x)) \subseteq \mathscr{C}_*(X)$。取 $y \in \cap Md(x)$, 则对任意的 $K \in Md(x)$ 都有 $y \in K$. 那么 $\cap Md(y) \subseteq \cap Md(x)$。故 $y \in \mathscr{C}_*(\cap Md(x))$。因此, $\cap Md(x) \subseteq \mathscr{C}_*(\cap Md(x)) \subseteq \mathscr{C}_*(X)$。于是有: $x \in \mathscr{C}_*(\mathscr{C}_*(X))$。所以 $\mathscr{C}_*(X) \subseteq \mathscr{C}_*(\mathscr{C}_*(X))$。另外, 显然有 $\mathscr{C}_*(\mathscr{C}_*(X)) \subseteq \mathscr{C}_*(X)$。因此可得: $\mathscr{C}_*(\mathscr{C}_*(X)) = \mathscr{C}_*(X)$。

(7U) 设 $X \subseteq U$, 由(L)可得 $\mathscr{C}_*(\mathscr{C}_*(\sim X)) = \mathscr{C}_*(\sim X)$。又由(2)有

$$
\begin{aligned}
& \mathscr{C}_*(\mathscr{C}_*(\sim X)) = \mathscr{C}_*(\sim X) \\
\Rightarrow\ & \mathscr{C}_*(\mathscr{C}_*(\sim X)) \supseteq \mathscr{C}_*(\sim X) \\
\Rightarrow\ & \mathscr{C}_*(\sim \mathscr{C}^*(X)) \supseteq \sim \mathscr{C}^*(X) \\
\Rightarrow\ & \sim \mathscr{C}^*(\mathscr{C}^*(X)) \supseteq \sim \mathscr{C}^*(X) \\
\Rightarrow\ & \mathscr{C}^*(\mathscr{C}^*(X)) \subseteq \mathscr{C}^*(X)
\end{aligned}
$$

这样便有 $\mathscr{C}^*(\mathscr{C}^*(X)) = \mathscr{C}^*(X)$ 成立。

(9L) 由(1)得: $\mathscr{C}_*(K) \subseteq K$. 若设 $x \in K$, 则有 $\cap Md(x) \subseteq K$, 所以有 $x \in \mathscr{C}_*(K)$, 即 $K \subseteq \mathscr{C}_*(K)$。因此: $\mathscr{C}_*(K) = K$.

注 3.1　设 U 为一非空论域, $\mathscr{C}$ 为 U 上的覆盖, 覆盖下近似 $\mathscr{C}_*$ 和上近似 $\mathscr{C}^*$ 下面性质不再成立:

(8L) $\mathscr{C}_*(\sim \mathscr{C}_*(X)) = \sim \mathscr{C}_*(X)$ 　　(下近似的补性)

(8U) $\mathscr{C}^*(\sim \mathscr{C}^*(X)) = \sim \mathscr{C}^*(X)$ 　　(上近似的补性)

(9U) $\forall K \in \mathscr{C}$, $\mathscr{C}^*(K) = K$ 　　(上近似的颗粒性)

例 **3.1** 覆盖下近似 $\mathscr{C}_*$ 和上近似 $\mathscr{C}^*$ 上述性质不在成立的反例如下:

设 $U=\{a,b,c\}$, $K_1=\{a,b\}$, $K_2=\{b,c\}$, 即 $\mathscr{C}=\{K_1,K_2\}$。

若取 $X=\{a,b\}$, $Y=\{a\}$, 有

$$\mathscr{C}_*(X)=\{a,b\}\sim\mathscr{C}_*(X)=\{c\}$$

$$\mathscr{C}_*(X)(\sim\mathscr{C}_*(X))=\mathscr{C}_*(X)(\{c\})=\varnothing\neq\{c\}$$

$$\mathscr{C}^*(Y)=\{a\},\ \sim\mathscr{C}^*(Y)=\{b,c\}$$

$$\mathscr{C}^*(\sim\mathscr{C}^*(Y))=\mathscr{C}^*(\{b,c\})=\{a,b,c\}\neq\{b,c\}$$

以及 $\mathscr{C}^*(K_1)=U\neq K_1$.

由以上定理可以看出, 我们给出的覆盖的上、下近似算子是对偶的一对算子, 这是 Z. Bonikowski 覆盖近似算子不满足的, 而且 Z. Bonikowski 覆盖近似不成立的性质在我们给出的近似算子中有许多都成立。我们给出的覆盖广义粗糙近似算子与经典的 Pawlak 粗糙近似算子对比一下会发现只是性质(L)、($8U$)、($9U$) 不在成立, 其余性质都成立, 这样为研究覆盖广义粗糙集带来很大的方便。

为描述方便, 下文没有特别说明覆盖广义粗糙集均指基于最小描述的覆盖广义粗糙集。

4 覆盖广义粗糙集的约简理论

Pawlak 经典粗糙集理论一个非常重要的问题就是属性约简, 通过对信息表或决策表的属性约简可以删除知识库中不相关或者不重要的知识, 从而对信息表或决策表的知识进行获取。本节主要讨论基于覆盖的粗糙集中覆盖的约简。

定义 4.1 设 U 是有限非空论域, $\mathscr{C}_1$, $\mathscr{C}_2$ 是 U 的两个覆盖, 若对任意的 $u\in U$ 都有: $\cap Md_{\mathscr{C}_1}(u)=\cap Md_{\mathscr{C}_2}(u)$, 则称覆盖 $\mathscr{C}_1$ 与覆盖 $\mathscr{C}_2$ 等价, 其中 $Md_{\mathscr{C}_1}(u)$, $Md_{\mathscr{C}_2}(u)$ 分别指 u 在覆盖 $\mathscr{C}_1$, $\mathscr{C}_2$ 下的最小描述。记做: $\mathscr{C}_1\sim\mathscr{C}_2$.

显然, 对于 U 上的覆盖 $\mathscr{C}$, $\mathscr{C}_1$, $\mathscr{C}_2$, $\mathscr{C}_3$, 满足下面性质:

(1) (自反性) $\mathscr{C}\sim\mathscr{C}$

(2) (对称性) $\mathscr{C}_1\sim\mathscr{C}_2\Rightarrow\mathscr{C}_2\sim\mathscr{C}_1$

(3) (传递性) $\mathscr{C}_1\sim\mathscr{C}_2$ 且 $\mathscr{C}_2\sim\mathscr{C}_3\Rightarrow\mathscr{C}_1\sim\mathscr{C}_3$ 由以上可知 U 上的覆盖的等价是一个等价关系。

定义 4.2[26] 设 U 是有限非空论域, $\mathscr{C}$ 是 U 的一个覆盖, $K\in\mathscr{C}$, 如果 K 是 $\mathscr{C}-\{K\}$ 中某些集合的并集, 则称 K 是 $\mathscr{C}$ 的可约简元素, 否则称 K 是 $\mathscr{C}$ 的不可约简元素。

定义 4.3[26] 设 U 是有限非空论域, $\mathscr{C}$ 是 U 的一个覆盖, 如果 $\mathscr{C}$ 的每一个元素都是不可约简元素, 则称 $\mathscr{C}$ 是不可约简的, 否则称为是可约简的。

显然, 在覆盖 $\mathscr{C}$ 中去掉一个可约简的元素 K, 集族 $\mathscr{C}-K$ 仍然是一个覆盖, 更一般的在去掉覆盖中所有可约简元素之后, 剩下的集族仍然是一个覆盖, 这样覆盖的约简被定义为:

定义 4.4[26] 设 U 是有限非空论域, $\mathscr{C}$ 是 U 的一个覆盖, 在 $\mathscr{C}$ 中去掉所有可约简元素后得到的覆盖称为 $\mathscr{C}$ 的约简, 记作: $reduct(\mathscr{C})$。

命题 4.1[26] 设 U 是有限非空论域, $\mathscr{C}$ 是 U 的一个覆盖, K 是 $\mathscr{C}$ 的可约简元素, 若记: $K_1\in\mathscr{C}-\{K\}$, 则 K_1 是 $\mathscr{C}$ 的可约简元素的充分必要条件为 K_1 是 $\mathscr{C}-\{K\}$ 的可约简元素。

命题 4.1 表明: 在一个覆盖中去掉一个可约简元素后, 不会产生新的可约简元素, 也不会是原来的覆盖中的可约简元素在新的覆盖中变为不可约简元素, 因此, 在计算一个覆盖的约简时, 可以同时去掉所有的可约简元素, 也可以分步分步去掉可约简元素。

定理 4.1 设 U 是有限非空论域, 若 $\mathscr{C}$ 是 U 的一个覆盖, K 是 $\mathscr{C}$ 的可约简元素。则对于任意的 $x\in U$, 有 $Md_{\mathscr{C}}(x)=Md_{\mathscr{C}-\{K\}}(x)$。

证明 因为 K 是 $\mathscr{C}$ 的可约简元素, 由 $Md_{\mathscr{C}}(x)$ 的定义可知: $K\notin Md_{\mathscr{C}}(x)$。又对于 $\mathscr{C}$ 中的任意元素

K_1, 若 $K\subseteq K_1$, 则必然存在 $\mathscr{C}$ 中的元素 K' 使得有 $K\subseteq K\subseteq K_1$ 成立。因此对于任意的 $x\in U$, 有 $Md_{\mathscr{C}}(x)=Md_{\mathscr{C}-\{K\}}(x)$。

定理 4.2　设 U 是有限非空论域, $\mathscr{C}$ 是 U 的一个覆盖, 则 $\mathscr{C}\sim reduct(\mathscr{C})$。

证明　由定义 4.2 和定理 4.1 直接获证。

于是, 我们可以直接得出下面结论。

定理 4.3　设 U 是有限非空论域, $\mathscr{C}_1$、$\mathscr{C}_2$ 是 U 的一个覆盖, 则 $\mathscr{C}_1\sim\mathscr{C}_2\Leftrightarrow reduct(\mathscr{C}_1)\sim reduct(\mathscr{C}_2)$。

注 4.1　在上述定理中, 若 $\mathscr{C}_1\sim\mathscr{C}_2$, 未必有 $reduct(\mathscr{C}_1)=reduct(\mathscr{C}_2)$。

例如, 取论域为 $U=\{a,b,c\}$ 及其覆盖 $\mathscr{C}_1=\{\{a,b\},\{b,c\}\}$, $\mathscr{C}_2=\{\{a,b\},\{b\},\{b,c\}\}$, 显然, $\cap Md_{\mathscr{C}_1}(a)=\cap Md_{\mathscr{C}_2}(a)$; $\cap Md_{\mathscr{C}_1}(b)=\cap Md_{\mathscr{C}_2}(b)$; 即 $\mathscr{C}_1\sim\mathscr{C}_2$, 但 $\mathscr{C}_1=reduct(\mathscr{C}_1)\neq reduct(\mathscr{C}_2)=\mathscr{C}_2$.

定理 4.4　设 U 是有限非空论域, $\mathscr{C}_1$, $\mathscr{C}_2$ 是 U 的一个覆盖, 若对任意的 $X\subseteq U$ 关于覆盖 $\mathscr{C}_1$ 与 $\mathscr{C}_2$ 产生相同的覆盖下近似, 则 $\mathscr{C}_1$ 与 $\mathscr{C}_2$ 等价。

证明　设 $\forall u\in U$, 取 $X=\cap Md_{\mathscr{C}_2}(u)$, 显然有 $\cap Md_{\mathscr{C}_2}(u)\subseteq\cap Md_{\mathscr{C}_2}(u)$。因此, $u\in\mathscr{C}_{2*}(X)$。

又对任意的 $X\subseteq U$ 关于覆盖 $\mathscr{C}_1$ 与 $\mathscr{C}_2$ 产生相同的覆盖下近似, 即 $\mathscr{C}_{1*}(X)=\mathscr{C}_{2*}(X)$。于是, $u\in\mathscr{C}_{1*}(X)$。故 $\cap Md_{\mathscr{C}_1}(u)\subseteq X$, 即 $\cap Md_{\mathscr{C}_1}(u)\subseteq\cap Md_{\mathscr{C}_2}(u)$。

同理可证 $\cap Md_{\mathscr{C}_2}(u)\subseteq\cap Md_{\mathscr{C}_1}(u)$。因此有 $\cap Md_{\mathscr{C}_1}(u)=\cap Md_{\mathscr{C}_2}(u)$, 即 $\mathscr{C}_1$ 与 $\mathscr{C}_2$ 等价。

定理 4.5　设 U 是有限非空论域, $\mathscr{C}_1$, $\mathscr{C}_2$ 是 U 的一个覆盖, 若对任意的 $X\subseteq U$ 关于覆盖 $\mathscr{C}_1$ 与 $\mathscr{C}_2$ 产生相同的覆盖上近似, 则 $\mathscr{C}_1$ 与 $\mathscr{C}_2$ 等价。

证明　由于对任意的 $X\subseteq U$ 关于覆盖 $\mathscr{C}_1$ 与 $\mathscr{C}_2$ 产生相同的覆盖上近似, 即 $\mathscr{C}_1^*(X)=\mathscr{C}_2^*(X)$。故有: $\mathscr{C}_1^*(\sim X)=\mathscr{C}_2^*(\sim X)$。由覆盖上下近似的对偶性知, 对任意的 $X\subseteq U$, 此时有 $\sim\mathscr{C}_{1*}(X)=\sim\mathscr{C}_{2*}(X)$。因此, $\forall X\subseteq U$, 有 $\mathscr{C}_{1*}(X)=\mathscr{C}_{2*}(X)$。由定理 4.4 便知, 对任意的 $u\in U$, 有 $\cap Md_{\mathscr{C}_1}(u)=\cap Md_{\mathscr{C}_2}(u)$, 即 $\mathscr{C}_1$ 与 $\mathscr{C}_2$ 等价。定理获证。

定理 4.6　设 U 是有限非空论域, $\mathscr{C}_1$, $\mathscr{C}_2$ 是 U 的一个覆盖, 对任意的 $X\subseteq U$ 关于覆盖 $\mathscr{C}_1$ 与 $\mathscr{C}_2$ 产生相同的覆盖下近似的充分必要条件是 $\mathscr{C}_1$ 与 $\mathscr{C}_2$ 等价。

证明　由定义 5 和定理 4.5 直接可得。

定理 4.7　设 U 是有限非空论域, $\mathscr{C}_1$, $\mathscr{C}_2$ 是 的一个覆盖, 对任意的 $X\subseteq U$ 关于覆盖 $\mathscr{C}_1$ 与 $\mathscr{C}_2$ 产生相同的覆盖上近似的充分必要条件是 $\mathscr{C}_1$ 与 $\mathscr{C}_2$ 等价。

证明　由定义 5 和定理 4.5 直接可得。

定理 4.8　设 U 是有限非空论域, $\mathscr{C}_1$, $\mathscr{C}_2$ 是 U 的一个覆盖, 对任意的 $X\subseteq U$ 关于覆盖 $\mathscr{C}_1$ 与 $\mathscr{C}_2$ 产生相同的覆盖上近似的充分必要条件是产生相同的覆盖下近似。

证明　由定理 4.6 和定理 4.7 直接可得。

由定理 4.8 可以看出这样定义的覆盖上下近似不仅是一对对偶算子, 而且还是相互决定的。我们在考虑覆盖的约简时, 只需考虑一个算子即可。因此结合上面定理我们可以有下面推论。

推论 4.1　设 U 是有限非空论域, 对任意的 $X\subseteq U$ 关于 U 的若干覆盖产生相同的覆盖近似的充分必要条件是这些覆盖相互等价。

推论 4.2　设 U 是有限非空论域, 对任意的 $X\subseteq U$ 关于 U 的若干覆盖产生相同的覆盖近似的充分必要条件是这些覆盖有相互等价的约简。

5　覆盖广义粗糙集的公理化方法

粗糙集理论的公理化方法是把下近似和上近似视为最基本的概念, 并利用一个公理集来刻画它们, 以揭示公理集与此条件下所产生的代数系统之间的联系。不同粗糙集模型中的近似算子要用不同的公

理集来刻画, 一般二元关系下的粗糙集, 基于模糊关系的模糊粗糙集模型等等的公理化方法已经获得, 详见文献[24]。而且 W. Zhu 也在文[26]中对提出的近似算子进行了公理化讨论。但是由于文中讨论的上下近似算子并不是一对对偶的算子, 故只给出了下近似的公理化刻画, 不是很完善。本节我们就我们给出的覆盖近似算子进行公理化方法的讨论。

定义 5.1[24] 设 U 是有限非空论域, 对 $P(U)$ 上的算子 L, H, 若 $\forall X \in P(U)$, 它们满足下面公理:

(L_0) $L(X)=\sim H(\sim X)$;

(L_0) $H(X)=\sim L(\sim X)$,

则称算子 L 和 H 是对偶的。

由于 L, H 的对偶性, 我们可以用其中的一个来定义另一个, 但为了方便, 通常仍用两个。

定理 5.1 设 U 是有限非空论域, 若 $L: P(U) \to P(U)$ 是论域幂集到自身的映射, 则一定在 U 上存在一个覆盖 $\mathscr{C}$ 使得对任意的 $X \in P(U)$ 都有 $L(X)=\mathscr{C}_*(X)$ 成立的充分必要条件是 L 满足下列公理:

(L_1) $L(U)=U$;

(L_2) $L(X\cap Y)=L(X)\cap L(Y)$, 其中 $X, Y \in P(U)$;

(L_3) $L(X)\subseteq X$;

(L_4) $L(L(X))=L(X)$。

证明 "$\Leftarrow$": 由定理 2 直接可得。

"$\Rightarrow$": 取 $\mathscr{C}=\{K\subseteq U \mid L(K)=K\}$, 显然它是 U 上的覆盖。下证对任意的 $X\subseteq U$ 有 $\mathscr{C}_*(X)=L(X)$。设 $x\in\mathscr{C}_*(X)$, 则又下近似的定义有 $\cap Md(x)\subseteq X$. 又由公理 L_2 可知: 对任意的 $X\subseteq Y$ 有 $L(X)=L(X\cap Y)=L(X)\cap L(Y)$。所以有 $X\subseteq Y\Rightarrow L(X)\subseteq L(Y)$ 成立。故 $L(\cap Md(x))\subseteq L(X)$。由 $Md(x)$ 的定义知 $\cap Md(x)=\cap K_i$, $K_i\in\mathscr{C}$. 于是结合公理 L_1 和 L_2 可得: $L(\cap Md(x))=L(\cap K_i)=\cap K_i$. 又由定理 2 有 $\mathscr{C}_*(K_i)=K_i$, 故有: $\cap K_i=\cap\mathscr{C}_*(K_i)=\mathscr{C}_*(\cap K_i)=\mathscr{C}_*(\cap Md(x))$。所以有: $L(\cap Md(x))=\mathscr{C}_*(\cap Md(x))$。于是可得 $\mathscr{C}_*(\cap Md(x))\subseteq L(X)$。又 $x\in\mathscr{C}_*(\cap Md(x))$, 故有 $x\in L(X)$。即 $\mathscr{C}_*(X)\subseteq L(X)$ 成立。

另一方面, 如果 $L(X)=\varnothing$, 显然有 $L(X)\subseteq\mathscr{C}_*(X)$。如果 $L(X)\neq\varnothing$, 由公理 L_4 可知 $L(X)\in\mathscr{C}$. 所以对于任意的 $x\in L(X)$, 有 $\cap Md(x)\subseteq L(X)$。再由公理 L_3 可得: $\cap Md(x)\subseteq L(X)\subseteq X$, 故 $x\in\mathscr{C}_*(X)$。即 $L(X)\subseteq\mathscr{C}_*(X)$ 成立。

因此, 对于任意的 $X\in P(U)$, 定理成立。

上述定理用公理化的方法对覆盖下近似进行了刻画, 由于上下近似算子的对偶性, 直接可得下面公理化的方法刻画上近似的结论。

定理 5.2 设 U 是有限非空论域, 若 $H: P(U)\to P(U)$ 是论域幂集到自身的映射, 则一定在 U 上存在一个覆盖 $\mathscr{C}$ 使得对任意的 $X\in P(U)$ 都有 $H(X)=\mathscr{C}^*(X)$ 成立的充分必要条件是 H 满足下列公理:

(H_1) $H(\varnothing)=\varnothing$;

(H_2) $H(X\cup Y)=H(X)\cup H(Y)$, 其中 $X, Y\in P(U)$;

(H_3) $H(X)\supseteq X$;

(H_4) $H(H(X))=H(X)$。

定义 5.2 设 U 是有限非空论域, $L, H: P(U)\to P(U)$ 是论域幂集到自身的一对对偶算子, 如果 L 满足公理 $L_1\sim L_4$, 或者等价的 H 满足公理 $H_1\sim H_4$, 则称系统 $(P(U), \cap, \cup, \sim, L, H)$ 为一个覆盖广义粗糙集代数, L, H 称为近似算子。

在给定 (L_1) 的条件下, L_2 等价于下面的公理 K:

(K) $L(\sim X\cup Y)\subseteq\sim L(X)\cup L(Y)$。

利用对偶性易知 (K) 等价于 (K'):

(K')　$\sim H(X)\cap H(Y)\subseteq H(\sim X\cap Y)$。

我们所讨论的满足公理$(L_1)\sim(L_4)$和$(H_1)\sim(H_4)$的一对对偶算子L,H中L是可乘的, 即对交运算满可分配; H是可加的, 即对并运算满足可分配, 但显然L不是可加的, H也不是可乘的。

尽管L不是可加的, H也不是可乘的, 但它们却满足较弱的条件:

(L_5)　$L(X\cup Y)\supseteq L(X)\cup L(Y)$;

(H_5)　$H(X\cap Y)\subseteq H(X)\cap H(Y)$。

显然, (L_5)和(H_5)可以等价的表述为L和H对集合包含关系的单调性, 即

(L_6)　$X\subseteq Y\Rightarrow L(X)\subseteq L(Y)$;

(H_6)　$X\subseteq Y\Rightarrow H(X)\subseteq H(Y)$。

容易看出, (L_2)蕴涵(L_5)和L_6, (H_2)蕴涵(H_5)和H_6. 但反之不对, 实际上它们等价于

(L_7)　$L(X\cap Y)\subseteq L(X)\cap L(Y)$;

(H_7)　$H(X\cup Y)\supseteq H(X)\cup H(Y)$。

由以上分析各公理之间的关系, 可以总结为下面的结论。

定理 **5.3**　对于近似算子L和H, 有下面的性质成立:

(a)　$(L_1),(L_2),(L_3),(L_4)\Leftrightarrow(L_1),(K),(L_3),(L_4)$;

　　$(H_1),(H_2),(H_3),(H_4)\Leftrightarrow(H_1),(K'),(H_3),(H_4)$。

(b)　$(L_2)\Rightarrow(L_5),(L_6),(L_7)$;

　　$(H_2)\Rightarrow(H_5),(H_6),(H_7)$。

(c)　$(L_5)\Leftrightarrow(L_6)\Leftrightarrow(L_7)$;

　　$(H_5)\Leftrightarrow(H_6)\Leftrightarrow(H_7)$。

在一个覆盖广义粗糙集代数$(P(U),\cap,\cup,\sim,L,H)$中, 两个一元算子加上"～"复合可以产生八个一元算子, 即$LL,HH,HL,LH,\sim L,\sim H,L\sim,H\sim$, 其中$\sim L=H\sim$, $\sim H=L\sim$, 因此可以看作是六个新算子$LL,HH,HL,LH,\sim L,\sim H$. 一般来讲, 它们不同于原来的算子$L,H,\sim$。这就意味着重复使用一元算子一般是不可以简化的。但上面六个复合后的六个一元算子恰好构成三对对偶算子, 即: LL和HH, LH和HL, $\sim L$和$\sim H$.

由于L和H满足$(L_1)\sim(L_4)$和$(H_1)\sim(H_4)$, 我们有

$$\begin{aligned}
&LL(U)=U\\
&LL(X\cap Y)=LL(X)\cap LL(Y)\\
&LL(X)\subseteq X\\
&LL(L(X))=L(X)\\
&HH(\varnothing)=\varnothing\\
&HH(X\cup Y)=HH(X)\cup HH(Y)\\
&HH(X)\supseteq X\\
&HH(H(X))=H(X)
\end{aligned}$$

所以, 系统$(P(U),\cap,\cup,\sim,LL,HH)$也是一个覆盖广义粗糙集代数。易知, 对任意的自然数n, $(P(U),\cap,\cup,\sim,\underbrace{LL\cdots L}_{n个},\underbrace{HH\cdots H}_{n个})$也是覆盖广义粗糙集代数。

对于对偶算子HL和LH, 由于

$$HL(X\cap Y)=H(L(X)\cap L(Y))\subseteq HL(X)\cap HL(Y)$$
$$LH(X\cup Y)=L(H(X)\cup H(Y))\supseteq LH(X)\cup LH(Y)$$

故算子HL和LH满足(L_5)和(H_5), 而不满足(L_2)和(H_2), 所以系统$(P(U),\cap,\cup,\sim,LH,HL)$是比

覆盖广义粗糙集代数弱的系统。

另外, 由 DeMorgan 定律以及公理$(L_1)\sim(L_4)$和$(H_1)\sim(H_4)$, 对偶算子$\sim L$ 和$\sim H$ 满足下面性质:

$(\sim L_1)$ $\sim L(U)=\varnothing$;

$(\sim L_2)$ $\sim L(X\cap Y)=\sim L(X)\cup\sim L(Y)$;

$(\sim L_3)$ $\sim L(X)\supseteq\sim X$;

$(\sim L_4)$ $\sim L(\sim H(X))=H(X)$;

$(\sim H_1)$ $\sim H(\varnothing)=U$;

$(\sim H_2)$ $\sim H(X\cup Y)=\sim H(X)\cap\sim H(Y)$;

$(\sim H_3)$ $\sim H(X)\subseteq\sim X$;

$(\sim H_4)$ $\sim H(\sim L(X))=L(X)$。

显然, $\sim L$ 和$\sim H$ 满足$(\sim L_1)\sim(\sim L_4)$和$(\sim H_1)\sim(\sim H_4)$当且仅当L 和H 满足$(L_1)\sim(L_4)$和$(H_1)\sim(H_4)$。所以系统$(P(U),\cap,\cup,\sim,\sim L,\sim H)$可以看作是覆盖广义粗糙集代数的另一种表现形式。

6 结论

本文对基于覆盖的广义粗糙集做了详细的研究, 在 Z. Bonikowski 的覆盖广义粗糙集模型的基础上, 对覆盖近似算子做了适当的修改, 使得覆盖上、下近似算子具有对偶关系, 而且详细的研究了它们的性质, 并和 Pawlak 经典粗糙集以及 Z. Bonikowski 的覆盖广义粗糙集做了比较。详细讨论了在新定义的覆盖广义粗糙集的约简问题, 研究了覆盖约简与覆盖近似的关系, 给出了覆盖公理化相关重要的结论。本文是在覆盖广义粗糙集在现有的情况下做了深入的探讨和推广, 使广义粗糙集理论得到了很好的完善, 丰富了覆盖广义粗糙集模型理论。本文所采用的方法以及得到的相关结论, 对覆盖广义信息系统中知识描述的进一步研究, 奠定了一定的理论基础。

参考文献:

[1] Banerjee M, Pal S K. Roughness of a fuzzy set[J]. Inform. Sci., 1995, 93: 235～246.

[2] Biswas R. On rough sets and fuzzy rough sets[J]. Bull. Pol. Acad. Sci. Math., 1992, 42: 343～349.

[3] Biswas R. On rough fuzzy sets[J]. Bull. Pol. Acad. Sci. Math., 1994, 42: 352～355.

[4] Bonikowski Z. Algebraic structures of rough sets[C]//Ziarko W. Rough Sets, Fuzzy Sets and Knowledge Discovery. London: Springer-Verlag, 1994: 243～247.

[5] Bonikowski Z, Bryniarski E, Wybraniec U. Extensions and intentions in the rough set theory[J]. Inform. Sci., 1998, 107: 149～167.

[6] Bryniaski E. A calculus of rough sets of the first order[J]. Bull. Pol. Acad. Sci., 1998, 106: 71～77.

[7] Cattaneo G. Abstract approximation spaces for rough theories[C]//Polkowski L, Skowron A. Rough Sets in Knowledge Discovery 1: Methodology and Applications. Heidelberg: Physica-Verlag, 1998: 59～98.

[8] Chakrabarty K, Biswas R, Nanda S. Fuzziness in rough sets[J]. Fuzzy Sets and Systems, 2000, 110: 247～251.

[9] Dubois D, Prade H. Rough fuzzy sets and fuzzy rough sets[J]. Internat. J. Gen. Systems, 1990, 17: 191～208.

[10] Kryszkiewicz M. Rough set approach to incomplete information systems[J]. Inform. Sci., 1998, 112: 39～49.

[11] Nakamura A. Fuzzy rough sets[J]. Note on Multiple-valued Logic in Japan, 1988, 9: 1～8.

[12] Nanda S, Majumda S. Fuzzy rough sets[J]. Fuzzy Sets and Systems, 1992, 45: 157～160.

[13] Pawlak Z. Rough sets[J]. Internat. J. Comput. Inform. Sci., 1982, 11: 341～356.

[14] Pomykala J A. Approximation operations in approximation space[J]. Bull. Pol. Acad. Sci., 1987, 9～10: 653～662.

[15] Skowron A, Stepaniuk J. Tolerance approximation spaces[J]. Fundam. Inform., 1996, 27: 245～253.

[16] Slowinski R, Vanderpooten D. A generalized definition of rough approximations based on similarity[J]. IEEE

Trans. Knowledge and Data Eng., 2000, 12: 331～336.

[17] Wasilewska A. Topological rough algebras[C]//Lin T Y, Cercone N. Rough Sets & Data Mining. Boston: Kluwer Academic Publishers, 1997: 441～425.

[18] Xu W H, Zhang W X. Measuring roughness of generalized rough sets induced by a covering[J]. Fuzzy Sets and Systems, 2007, 158: 2443～2455.

[19] 徐伟华. 基于广义粗糙集和粒计算的知识发现方法研究[D]. 西安: 西安交通大学, 2007.

[20] Yao Y Y. Constructive and algebraic methods of theory of rough sets[J]. Inform. Sci., 1998, 109: 21～47.

[21] Yao Y Y. Relational interpretations of neighborhood operators and rough set approximation operators[J]. Inform. Sci., 1998, 111: 239～259.

[22] Zadeh L A. Fuzzy sets[J]. Inform. Contr. 1965, 8: 338～353.

[23] Zakowski W. Approximations in the space (U, Π)[J]. Demonstratio Mathematica, 1983, 16: 761～769.

[24] 张文修, 吴伟志, 梁吉业, 李德玉. 粗糙集理论与方法[J]. 北京: 科学出版社, 2001.

[25] 张文修, 仇国芳, 吴伟志. 粗糙集属性约简的一般理论[J]. 中国科学 E 辑, 2005, 35(12): 1304～1313.

[26] Zhu W, Wang F Y. Reduction and axiomization of coveringgereralized rough sets[J]. Inform. Sci., 2003, 152: 217～230.

Covering Generalized Rough Sets Based on the Intersection of Minimal Description

ZHANG Xiao-yan[1], XU Wei-hua[1], ZHANG Wen-xiu[2]

(1. School of Mathematics and Statistics, Chongqing University of Technology, Chongqing 400054, China;
2. School of Science, Xi'an Jiaotong University, Xi'an 710049, China)

Abstract: Bonikowski has got the model of covering generalized rough sets in [4], but there exist some shortcoming. In this paper, we propose new lower and upper approximations based on the intersection of minimal description. And some important properties have been acquired in the generalized rough set induced by a covering. Especially, it shows that the two approximations are dual approximation by comparing with ones of Pawlak's rough sets and Bonikowski's covering generalized rough sets respectively. Moreover, we investigate the covering reduction based on generalized rough sets with the new approximations and discuss some significant properties of axiomatization.

Key words: Pawlak Rough Sets; Bonikowski Approximation; Reductions of Covering; Axiomatization