计算机研究与发展 ISSN 1000-1239/CN 11-1777/TP
Journal of Computer Research and Development 43(10):1719～1723,2006

基于覆盖的粗糙模糊集模型研究

魏 莱 苗夺谦 徐菲菲 夏富春
(同济大学计算机科学与技术系 上海 201804)
(weily105@hotmail.com)

Research on a Covering Rough Fuzzy Set Model

Wei Lai, Miao Duoqian, Xu Feifei, and Xia Fuchun
(Department of Computer Science and Technology, Tongji University, Shanghai 201804)

Abstract Based on the research of a covering rough set model, the definitions about the upper approximation of the covering rough set are found to be inconsistent. Through the analysis of the difference among the several models, a more reasonable model is adopted, and, combined with the theory about covering reduction, the covering rough set model is redefined. Its several properties are given. Meanwhile, the model is extended, and the covering rough fuzzy set model is defined. Then some of its good properties are also proved.

Key words covering; covering reduction; rough set; rough fuzzy set

摘 要 在研究覆盖粗糙集模型中,发现对覆盖粗糙集上近似的定义并不一致. 简述了各个模型的区别,并在一个较合理的覆盖粗糙集上近似定义上,结合覆盖约简理论,重新定义了基于覆盖的粗糙集模型,讨论了它的一些性质. 另外,将模型进行推广,定义了基于覆盖的粗糙模糊集模型,证明了它具有一些较好的性质.

关键词 覆盖;覆盖约简;粗糙集;粗糙模糊集

中图法分类号 TP18

1982 年波兰数学家 Pawlak 提出的粗糙集理论[1]是对经典集合论的推广,是用来处理不完全、不精确、不相容数据的强有力工具. 其在数据挖掘、模式识别、人工智能等领域发挥了重要的作用. 粗糙集理论通过等价关系对论域进行划分,用论域上的一对可定义的(crisp)子集——称之为上、下近似——来逼近讨论的对象集合,从而可以发现信息系统中一些隐藏的知识[2]. 可以说等价关系是经典粗糙集理论中最基本也是最重要的概念. 但是对象间满足等价关系这样的要求过于严格,限制了粗糙集的应用范围. 为此学者们对等价关系进行推广,提出了一些泛化的粗糙集模型,如基于相似关系下的粗糙集模型[3]、基于覆盖的粗糙集模型[4-6]、粗糙模糊集和模糊粗糙集模型[7-8]等,这些模型的提出丰富了粗糙集理论,为粗糙集理论的应用研究拓宽了道路.

本文研究了基于覆盖的粗糙集模型,分析了几个模型关于上近似的不同定义,阐述了它们之间的联系,并在覆盖约简[3]的基础上重新定义了基于覆盖的粗糙集模型,给出了它的一些性质. 同时,结合粗糙模糊集理论,定义了基于覆盖的粗糙模糊集模型,证明了其主要的性质.

1 基于覆盖的粗糙集的上近似

定义 1[1]. 设 $U=\{x_1, x_2, x_3, \cdots, x_n\}$是一个

收稿日期:2006-04-29;**修回日期**:2006-05-29
基金项目:国家自然科学基金项目(60175016,60475019)

非空有限论域, R 是一等价关系(满足自反性、对称性、传递性), 我们称 (U, R) 为一个近似空间. 由 R 可以导出 U 上的一个等价划分 $U/R=\{X_1, X_2, \cdots, X_m\}$, 满足 $X_i \cap X_j = \varnothing$, 且 $\bigcup_{i=1}^{m} X_i = U$. 对于任意 $A \subset U$, 定义 A 关于 R 的上、下近似为

$$\overline{R}(A) = \cup\{X_i \in U/R \mid X_i \cap A \neq \varnothing\},$$

$$\underline{R}(A) = \cup\{X_i \in U/R \mid X_i \subseteq A\}.$$

这就是 Pawlak 关于粗糙集最原始的定义, 它们的性质在这里不再复述, 具体可以参见文献[1].

为了讨论基于覆盖的粗糙模糊集模型, 为此先介绍基于覆盖的粗糙集的定义. 关于覆盖粗糙集模型我们可以发现有几种不同定义, 其不一致处是关于上近似定义的不同.

先引入覆盖和最小描述的概念.

定义 2[5]. 设 U 是一个非空有限论域, C 是 U 的一个子集族, 如果 $\cup C = U$ 且 $\forall C_i \neq \varnothing$, 称 C 是 U 上的一个覆盖, (U, C) 为一个覆盖近似空间. 对 $\forall x \in U$, 关于 x 的最小描述定义为

$$md(x) = \{K \in C \mid x \in K \wedge (\forall S \in C \wedge x \in S \wedge S \subseteq K \Rightarrow K = S)\}.$$

可见覆盖 C 是对等价关系 R 的扩展. 基于自反关系和相似关系的粗糙集模型都可以纳入基于覆盖的粗糙集模型的定义中[9].

定义 3[5]. 对 $\forall X \subseteq U$, 集族 $C_*(X) = \{K \in C \mid K \subseteq X\}$ 称为 X 的覆盖下近似集族.

1) 集合 $\underline{C}(X) = \cup C_*(X)$ 称为 X 的覆盖下近似.

2) 集合 $C_*^*(X) = X - \underline{C}$ 称为 X 的覆盖边界.

3) 集族 $Bn(X) = \{md(x) \mid x \in C_*^*(X)\}$ 称为 X 的覆盖边界近似集族.

4) 集族 $C^*(X) = C_*(X) \cup Bn(X)$ 称为 X 的覆盖上近似集族.

5) 集合 $\overline{C}(X) = \cup C^*(X)$ 称为 X 的覆盖上近似.

如果 $\overline{C}(X) = \underline{C}(X)$ 称 X 是可定义的, 否则称其为不可定义的.

事实上, 这个关于覆盖粗糙集的定义是有其局限性的, 主要是对上近似的定义. 我们用文献[9]中的例子说明.

例 1. 设 $U=\{a, b, c, d\}$, $C=\{\{a, b\}, \{b, c\}, \{d\}\}$, $X=\{a, b\}$, 那么有上述定义可得 $\underline{C}(X) = \{a, b\}$, $\overline{C}(X) = \{a, b\}$, 但事实上 $K=\{b, c\} \in C$ 且 $K \cap X \neq \varnothing$, 按照 Pawlak 对粗糙集的理解来说, 应该有 $K \subseteq \overline{C}(X)$, 故此定义有其不合理处.

我们再来看另一种定义.

定义 4[6]. 对 $\forall X \subseteq U$, 集合 $\overline{C}(X) = \{x \in \cup \mid \exists K \in C, x \in K, K \cap X \neq \varnothing\}$ 称为 X 的覆盖上近似.

这个定义的上近似是对 Pawlak 经典粗糙集上近似定义的简单推广, 但对于基于覆盖的粗糙集不一定合适. 我们也举个例子.

例 2. 设 $U=\{a, b, c, d\}$, $C=\{\{a, b\}, \{b, c\}, \{b, c, d\}, \{d\}\}$, $X=\{a, b\}$, 那么根据定义 4 可知, $\overline{C}(X) = \{a, b, c, d\}$, 上近似中含有元素 d 是因为 $K=\{b, c, d\} \cap X \neq \varnothing$, 但是由于 $X=\{a, b\}$, 覆盖中存在 $\{b, c\} \subset K$, 所以 $K=\{b, c, d\}$ 看来是不需要的.

以上两个例子的具体分析可以参见文献[9].

显然对 $\forall X \subseteq U$, 定义 3 中的上近似 $\overline{C}_1(X)$ 和定义 4 中的上近似 $\overline{C}_2(X)$, 有这样的关系: $\overline{C}_1(X) \subseteq \overline{C}_2(X)$. 从上面看来, 上述两种覆盖粗糙集模型关于上近似的定义都不够完善. 为此文献[9]给出了比较合理的定义.

定义 5[9]. 设 U 是一个非空有限论域, C 是 U 上的一个覆盖, 对 $X \subseteq U$, X 的上近似定义成:

$$\overline{C}(X) = \{y \in md(x) \mid x \in X\}.$$

这样有 $\overline{C}_1(X) \subseteq \overline{C}(X) \subseteq \overline{C}_2(X)$. 根据这个定义, 例 1、例 2 中基于覆盖的粗糙集的上近似都为

$$\overline{C}(X) = \{a, b, c\}.$$

2 基于覆盖约简的粗糙集模型

定义 6[4]. 设 U 是一个非空有限论域, C 是 U 上的一个覆盖, $K \in C$, 若 K 是 $C - \{K\}$ 中一些集合的并, 那么称 K 是 C 的一个可约简元素, 否则为一个不可约简的元素.

覆盖 C 中约去所有可约简元素后, 仍然是一个覆盖, 称 C 的覆盖约简, 记为 $reduct(C)$. 且 $reduct(C)$ 与 C 对 $\forall x \in U$ 具有相同的最小描述, 即 $md(x) = md_{reduct(X)}(x)$. 因此, 对于任一个覆盖我们都可以先进行约简, 得到其覆盖约简, 再在其覆盖约简上讨论相关的问题, 这样可以去除掉一些冗余信息, 使计算简化.

有了上述的准备, 我们完整地给出基于覆盖约简的粗糙集的定义.

定义 7. 设 C 是有限非空论域 U 上的一个覆盖, $reduct(C)$ 是 C 的覆盖约简, $\forall x \in U$ 的最小描述为

$$md_r(x) = \{K \in reduct(C) \mid x \in K \wedge$$

$(\forall S \in reduct(C) \wedge x \in S \wedge S \subseteq K \Rightarrow K = S)\} \forall X \subseteq U.$

基于覆盖的粗糙集的上近似定义为

$$\overline{C}(X) = \{y \in md_r(x) \mid x \in X\},$$

可以等价地描述为

$$\overline{C}(X) = \cup\{K \in md_r(x) \mid x \in X\}. \quad (1)$$

下近似:

$$\underline{C}(X) = \{y \in K \mid K \subseteq X, K \in md_r(x), x \in X,$$

可以等价地描述为

$$\underline{C}(X) = \cup\{K \in md_r(x) \mid K \subseteq X, x \in X\}. \quad (2)$$

我们知道,覆盖约简是去除冗余的有效手段,这样定义下的覆盖粗糙集模型能够更快速有效地求出覆盖粗糙集上、下近似.

我们看一下此模型具有的性质:

性质 1.

1) $\underline{C}(\varnothing) = \overline{C}(\varnothing) = \varnothing$;

2) $\underline{C}(U) = \overline{C}(U) = U$;

3) $\underline{C}(X) \subseteq X \subseteq \overline{C}(X)$;

4) $X \subseteq Y \Rightarrow \underline{C}(X) \subseteq \underline{C}(Y), \overline{C}(X) \subseteq \overline{C}(Y)$;

5) $\underline{C}(\underline{C}(X)) = \underline{C}(X)$;

6) $\underline{C}(\overline{C}(X)) = \overline{C}(X)$.

证明. 我们只证明 5) 和 6),其他的根据定义即可得.

因为$\underline{C}(\underline{C}(X)) \subseteq \underline{C}(X)$,我们只要证明$\underline{C}(X) \subseteq \underline{C}(\underline{C}(X))$. 任意 $K \subseteq \underline{C}(X)$,有$\underline{C}(K) \subseteq \underline{C}(\underline{C}(X))$. 又 K 在 X 中元素的最小描述里,故$\underline{C}(K) = K$,所以 $K \subseteq \underline{C}(\underline{C}(X))$,故$\underline{C}(X) \subseteq \underline{C}(\underline{C}(X))$,所以$\underline{C}(\underline{C}(X)) = \underline{C}(X)$.

同理(6)也可得.

但是我们发现有些性质一般情况下是不成立的,只有当对象集的覆盖变成划分时才成立. 这些性质是:

$$\underline{C}(X) \cap \underline{C}(Y) = \underline{C}(X \cap Y),$$

$$\overline{C}(X) \cup \overline{C}(Y) = \overline{C}(X \cup Y),$$

$$\underline{C}(X) = (\overline{C}(X^c))^c,$$

$$\overline{C}(X) = (\underline{C}(X^c))^c,$$

其中 X^c 表示 X 的补集.

3 基于覆盖约简的粗糙模糊集模型

粗糙模糊集是粗糙集理论的扩展,它扩大了粗糙集理论的应用范围,有着广泛的理论和应用价值. 为此我们想到将覆盖约简理论在上述基础上引入粗糙模糊集.

经典的粗糙模糊集的定义如下:

定义 8[7]. 设 U 是一个非空有限论域, R 是 U 上的一个等价关系. $A \in F(U)$ 表示 U 上的一个模糊集合. U 上的一对模糊集 $R(A) = (\underline{R}(A), \overline{R}(A))$ 为粗糙模糊集,隶属度函数分别为

$$\mu_{\underline{R}(A)}(x) = \inf\{\mu_A(y) \mid y \in [x]_R\},$$

$$\mu_{\overline{R}(A)}(x) = \sup\{\mu_A(y) \mid y \in [x]_R\},$$

其中 $x \in U$.

基于覆盖约简的粗糙模糊集的定义,就有了如下的形式:

定义 9. 设 U 是一个非空有限论域, C 是 U 上的一个覆盖,它的约简记为 $reduct(C)$, $A \in F(U)$, U 上的基于覆盖的粗糙模糊集为一对模糊集合 $CA = (\underline{C}A, \overline{C}A)$,其隶属度函数如下:

$$\mu_{\underline{C}(A)}(x) = \inf\{\mu_A(y) \mid y \in \cup K, K \in md_r(x)\},$$

$$\mu_{\overline{C}(A)}(x) = \sup\{\mu_A(y) \mid y \in \cup K, K \in md_r(x)\},$$

其中 $supp(A)$ 表示 A 的支集, $x \in U$.

如果 $\overline{C}(X)\underline{C}(A)$,称 A 是可定义的,否则称为粗糙模糊的.

我们来看基于覆盖的粗糙模糊集具有的性质.

性质 2. 设 U 是一个非空有限论域, C 是 U 上的一个覆盖, $A, B \in F(U)$, $x \in U$.

1) $\underline{C}(A) \subseteq A \subseteq \overline{C}(A)$.

2) $\underline{C}(U) = \overline{C}(U) = U$.

3) $\underline{C}(\varnothing) = \overline{C}(\varnothing) = \varnothing$.

4) 单调性: $A \subseteq B \Rightarrow \overline{C}(A) \subseteq \overline{C}(B), \underline{C}(A) \subseteq \underline{C}(B)$.

5) 可加性:

$\overline{C}(A \cup B) = \overline{C}(A) \cup \overline{C}(B)$.

$\underline{C}(A \cup B) = \underline{C}(A) \cup \underline{C}(B)$.

6) 幂等性:

$\overline{C}(A) = \overline{C}(\overline{C}(A)), \underline{C}(A) = \underline{C}(\underline{C}(A))$.

7) $\overline{C}(A^c) = (\underline{C}(A))^c, \underline{C}(A^c) = (\overline{C}(A))^c$,

其中, A^c 表示 A 的补集.

证明.

1) 根据定义显然.

2) 对于 $\forall x \in U$,

$$\mu_{\underline{C}(U)}(x) = \inf\{\mu_U(y) \mid y \in \cup K, K \in md_r$$

$(x)\} = 1$,同样 $\mu_{\overline{C}(U)}(x) = \sup\{\mu_U(y) | y \in \cup K, K \in md_r(x)\} = 1$,而 $\mu_U(x) = 1$,所以有 $\mu_{\overline{C}(U)}(x) = \mu_{\underline{C}(U)}(x) = \mu_U(x) = 1$,故 $\underline{C}(U) = \overline{C}(U) = U$.

3) 证明同(2).

4) 因为 $A \subseteq B$,所以 $\forall x \in U$,有 $\mu_A(x) \le \mu_B(x)$,所以根据定义 $\mu_{\underline{C}(A)}(x) \le \mu_{\underline{C}(B)}(x)$, $\mu_{\overline{C}(A)}(x) \le \mu_{\overline{C}(B)}(x)$,即有 $\overline{C}(A) \subseteq \overline{C}(B)$, $\underline{C}(A) \subseteq \underline{C}(B)$.

5) $\mu_{\overline{C}(A \cup B)}(x) = \sup\{\mu_{A \cup B}(y) | y \in \cup K, K \in md_r(x)\} = \sup\{\mu_A(y) \vee \mu_B(y) | y \in \cup K, K \in md_r(x)\} = \sup\{\mu_A(y) | y \in \cup K, K \in md_r(x)\} \vee \sup\{\mu_B(y) | y \in \cup K, K \in md_r(x)\} = \mu_{\overline{C}(A)}(x) \vee \mu_{\overline{C}(B)}(x)$,所以 $\overline{C}(A \cup B) = \overline{C}(A) \cup \overline{C}(B)$,同理有 $\underline{C}(A \cap B) = \underline{C}(A) \cap \underline{C}(B)$.

6) $\mu_{\overline{C}(\overline{C}(A))}(x) = \sup\{\mu_{\overline{C}(A)}(y) | y \in \cup K, K \in md_r(x)\} = \sup\{\sup\{\mu_A(y) | y \in \cup K, K \in md_r(x)\} | y \in \cup K, K \in md_r(x)\} = \sup\{\sup\{\mu_A(y) | y \in \cup K, K \in md_r(x)\} | y \in \cup K, K \in md_r(x)\} = \sup\{\mu_A(y) | y \in \cup K, K \in md_r(x)\} = \mu_{\overline{C}(A)}(x)$,所以 $\overline{C}(A) = \overline{C}(\overline{C}(A))$,同理有 $\underline{C}(A) = \underline{C}(\underline{C}(A))$.

7) $\mu_{\overline{C}(A^c)} = \sup\{\mu_{A^c}(x) | y \in \cup K, K \in md_r(x)\} = \sup\{1 - \mu_A(x) | y \in \cup K, K \in md_r(x)\} = 1 - \inf\{\mu_A(x) | y \in \cup K, K \in md_r(x)\} = 1 - \mu_{\overline{C}(A)}(x)$, 所以 $\overline{C}(A^c) = (\underline{C}(A))^c$,同理 $\underline{C}(A^c) = (\overline{C}(A))^c$.

我们比较可以看出,性质 5),7) 是覆盖粗糙集所不具有的.

4 结束语

粗糙集有着广泛的应用领域,但实际的应用需要完善的理论作为支撑. 覆盖粗糙集是对粗糙集的扩展,在研究覆盖粗糙集的过程中,我们发现对覆盖粗糙集上近似的定义不一致,这影响了粗糙集的应用. 为此本文简述了各个关于覆盖粗糙集的上近似的不同定义,并研究了覆盖约简,在此基础上重新定义了基于覆盖约简的粗糙集,给出了一些性质.

同时,模糊集理论是用来处理模糊现象和模糊概念的强有力工具,它与粗糙集理论互相联系,互为补充. Dubois 和 Prade[8]将两种理论联系起来,定义了粗糙模糊集和模糊粗糙集模型,引起了许多学者的关注[10-12]. 本文在上文中提出的覆盖粗糙集模型的基础上又进一步结合粗糙模糊集理论定义了基于覆盖的粗糙模糊集,对它的主要性质做了证明. 对进一步探索粗糙集和模糊集理论具有一定意义.

参 考 文 献

[1] Z Pawlak. Rough sets [J]. International Journal of Information and Computer Science, 1982, 11(5): 341-356

[2] M kryszkiewicz. Rough set approach to incomplete information system [J]. Information Science, 1998, 11(2): 39-49

[3] R Slowinski, D Vanderpooten. A generalized definition of rough approximation based on similarity [J]. IEEE Trans on Knowledge and Data Engineering, 2000, 12(2): 331-336

[4] Willian Zhu, Feiyue Wang. Reduction and axiomization of covering generalized rough sets [J]. Information Science, 2003, 152(1): 217-230

[5] Z Bonikowski, E Bryniarski, U Wybraniec. Extension and intentions in the rough set theory [J]. Information Science, 1998, 107(1): 149-167

[6] J N Mordeson. Rough set theory applied to (fuzzy) ideal theory [J]. Fuzzy Sets and System, 2001, 121(2): 315-324

[7] M Banerjee, Sankar K Pal. Roughness of a fuzzy set [J]. Information and Computer Science, 1996, 93(3): 235-245

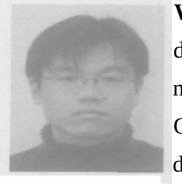
[8] D Dubois, H Prade. Rough fuzzy sets and fuzzy rough sets [J]. Information and Computer Science, 1990, 17(2): 191-209

[9] Eric C C Tsang, De Gang Chen, John W T Lee, *et al*. On the upper approximation of covering generalized rough sets [C]. The 3rd International Conference on Machine Learning and Cybernetics, Shanghai, 2004

[10] Xi Zhao Wang, Yan Ha, De Gang Chen. On the reduction of fuzzy rough sets [C]. The 3rd Int'l Conf on Machine Learning and Cybernetics, Guangzhou, 2005

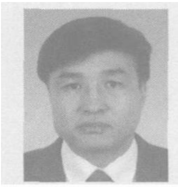
[11] Daniel S Yeung, De Gang Chen, Eric C C Tsang, *et al*. On the generalization of fuzzy rough sets [J]. IEEE Trans on Fuzzy System, 2005, 13(3): 343-361

[12] Wu Weizhi, Mi Jusheng, Zhang Wenxiu. Generalized fuzzy rough sets [J]. Information Science, 2003, 151(5): 263-282



Wei Lai, born in 1980. Received his B. A's degree in the Department of Applied Mathematics from Tongji University, Shanghai, China, in 2004. Now he is Ph D degree candidate in the Department of Computer Science of Tongji University. His current research interests include rough set theory, data mining, artificial intelligence, *etc*.

魏莱,1980 年生,博士研究生,主要研究方向为粗糙集理论、数据挖掘、人工智能等.



Miao Duoqian, born in 1961. Professor and Ph D supervisor of Tongji university. His main research interests include rough set theory, pattern recognition, artificial intelligence, *etc*.

苗夺谦,1961 年生,教授,博士生导师,主要研究方向为粗糙集理论、模式识别、人工智能等.

Xu Feifei, born in 1983. Received his B A's degree in the Department of Computer Science from Shanghai Electric College, China, in 2004. Now she is Ph D candidate in the Department of Computer Science of Tongji University. Her current research interests include rough set theory, data mining, artificial intelligence, *etc*.

徐菲菲,1983 年生,博士研究生,主要研究方向为粗糙集理论、数据挖掘、人工智能等.

Xia Fuchun, born in 1982. Received his B A's degree in the Department of Computer Science from Tongji University. China, in 2004. Now he is master candidate in the Department of Computer Science, Tongji University, Shanghai. His current research interests include rough set theory, data mining, *etc*.

夏富春,1982 年生,硕士研究生,主要研究方向为粗糙集理论、数据挖掘.

Research Background

Rough set theory is a powerful mathematical tool to deal with uncertainty and vagueness data. It, as well as fuzzy set theory, is an extension of standard set theory. The combination of the two different theories has created great influence on the development of data mining, artificial intelligence, pattern recognition and so on. The wide application requests the consistent theory. But the upper approximation of covering rough set has not been well defined. In this paper, we adopt the theory of covering reduction, redefined the covering rough set model. Then we combine the fuzzy set theory and extend the model, define the covering fuzzy rough set model, and prove its several properties. Our work is supported by the National Natural Science Foundation of China (60175016, 60475019).

《计算机研究与发展》被 EI Compendex 数据库收录

据 EI **中国** 2006 **年** 8 **月** 8 **日新闻报道**,**《计算机研究与发展》被** EI Compendex **数据库收录**,**成为** EI **核心期刊**(http://www.ei.org.cn). **我们向长期给予本刊关心与厚爱的广大专家和学者表示诚挚的感谢**!

《计算机研究与发展》编辑部