

# 不完备系统中基于多粒度的变精度粗糙集方法

周君仪<sup>1</sup> 窦慧莉<sup>2</sup> 杨习贝<sup>2,3</sup>

(江苏科技大学经济管理学院 镇江 212003)<sup>1</sup> (江苏科技大学计算机科学与工程学院 镇江 212003)<sup>2</sup>  
(南京理工大学计算机科学与技术学院 南京 210094)<sup>3</sup>

**摘 要** 多粒度方法是近年来新兴起的一种数据处理模型。为了使多粒度方法适用于不完备信息系统,借鉴非对称相似关系和可变精度的基本思想,提出了基于多粒度的变精度粗糙集模型,其分别包括乐观和悲观两种不同的形式。对这些模型的基本性质进行了讨论,并在多粒度框架下,将变精度方法与严格包含的方法进行了对比分析,最后通过实例分析,说明了如何在不完备信息系统中根据所提出的多粒度变精度粗糙集模型来获取“或”决策规则。

**关键词** 多粒度粗糙集,变精度粗糙集,不完备信息系统,相似关系

中图法分类号 TP18 文献标识码 A

## Multigranulation Based Variable Precision Rough Set Approach to Incomplete Systems

ZHOU Jun-yi<sup>1</sup> DOU Hui-li<sup>2</sup> YANG Xi-bei<sup>2,3</sup>

(School of Economics and Management, Jiangsu University of Science and Technology, Zhenjiang 212003, China)<sup>1</sup>

(School of Computer Science and Engineering, Jiangsu University of Science and Technology, Zhenjiang 212003, China)<sup>2</sup>

(School of Computer Science and Technology, Nanjing University of Science and Technology, Nanjing 210094, China)<sup>3</sup>

**Abstract** In recent years, multigranulation approach is a new arisen data analyzing model. To deal with the incomplete information system through the multigranulation approach, by employing the basic ideas of non-symmetric similarity relation and variable precision, the multigranulation based variable precision rough sets were proposed, which include optimistic and pessimistic cases, respectively. Not only the basic properties about these models were proved, but also the relationships between variable precision approach and strict inclusion approach were discussed under the multigranulation frame. Finally, an example was used to show the “OR” decision rules can be generated from the incomplete information systems by the proposed multigranulation based variable precision rough set models.

**Keywords** Multigranulation rough sets, Variable precision rough set, Incomplete information system, Similarity relation

## 1 引言

粗糙集理论是由 Pawlak 提出的一种处理含糊和不确定问题的数学工具。Pawlak 的经典粗糙集模型是通过构建下、上近似集,从而达到对目标概念进行近似逼近的目的。Pawlak 的粗糙集模型是建立在知识基的基础上的,知识基中一些等价关系的交集仍然是一个等价关系, Pawlak 将其称为不可分辨关系。以粒计算的观点来看,不可分辨关系仍然是一个等价关系,这个等价关系对应于论域上的一个划分,这个划分就是一种最为简单的信息粒化。

就目前的研究成果来看,很多学者所讨论的粗糙集及其相关的扩展模型都是建立在仅仅一个粒空间的基础上,这个粒空间可以是划分、覆盖<sup>[4]</sup>、邻域系统<sup>[5]</sup>等多种表现形式。而我国著名的粒计算与粗糙集学者钱宇华和梁吉业教授<sup>[6]</sup>认为,在单个粒空间中进行目标概念的近似逼近存在一定的不足之处,如不同的决策者之间的关系是独立的,所以对所有决策者的知识求交集是不符合实际问题需要的;再如对分布式

信息的处理,为了降低约简及知识获取的时空复杂度,无需对不同来源的信息做交集处理。为了解决上述问题,钱宇华和梁吉业教授提出了多粒度粗糙集的基本思想。所谓多粒度粗糙集,就是在一族而非单个粒空间中进行目标概念的近似逼近。在钱宇华和梁吉业提出的多粒度粗糙集理论中,有两种不同的基本模型:一种称为乐观多粒度粗糙集模型,另一种称为悲观多粒度粗糙集模型<sup>[7]</sup>。值得一提的是,经典的多粒度粗糙集依然是建立在等价关系的基础上的,因而只能用于处理具有不同信息来源的完备信息系统。然而,在现实世界中由于各种原因,面临的信息系统不一定是完备的,因而将多粒度方法引入不完备信息系统就显得尤其重要。

尽管钱宇华和梁吉业在文献[8]中已将多粒度方法引入不完备信息系统,但是他们仅考虑了乐观多粒度粗糙集模型,并且假设不完备信息系统中的所有未知属性值与其他值都是可以比较的,因而采用了一族容差关系来对目标概念进行乐观近似逼近。值得一提的是,在粗糙集方法中,未知属性值还有另一种解释,即对象被不完全描述的原因不仅可能是由于

本文受国家自然科学基金(61100116),江苏省自然科学基金(BK2011492),江苏省高校自然科学基金(11KJB520004),中国博士后科学基金(20100481149),江苏省博士后科学基金(1101137C)资助。

周君仪(1983—),女,硕士生,主要研究方向为粗糙集与粒计算, E-mail: zhoujy@just.edu.cn。

知识不精确,还可能是由于不可能用所有的属性来描述它们<sup>[9]</sup>,从这个角度来看,所有的未知属性值会被认为是丢失的,是不允许被比较的,鉴于此,Stefanowski 等人采用了相似关系<sup>[10,11]</sup>在不完备信息系统中对粗糙集模型进行了扩展。

本文主要考虑不完备信息系统中多粒度粗糙集方法,并将可变精度的基本思想融入其中,采用非对称相似关系提出了基于多粒度的变精度粗糙集模型。这主要是因为可变精度是建立在集合的多数包含的基础上的,其允许上、下近似存在一定的分类错误,能处理某种程度上的包含和属于,放宽了边界的严格定义,柔化了边界,更加适应于不完备和噪声数据的处理。

本文第2节介绍不完备信息系统中相似关系和基于相似关系的多粒度粗糙集模型;第3节将可变精度的思想引入相似关系多粒度粗糙集模型,并对基于多粒度的可变精度粗糙集模型的基本性质进行了讨论;第4节进行了实例分析,并说明了如何根据所提出的模型从不完备信息系统中获取“或”决策规则;最后总结全文。

## 2 基本概念

一个信息系统可被定义为二重组  $I = \langle U, AT \rangle$ , 其中  $U$  表示所有对象的集合,称为论域; $AT$  表示所有属性的集合。对于  $\forall a \in AT$ , 定义映射  $a: U \rightarrow V_a$ ,  $V_a$  表示属性  $a$  的值域,即  $a(x) \in V_a (x \in U)$ 。一个决策系统是一个信息系统  $I = \langle U, AT \cup \{d\} \rangle$ , 其中  $AT$  为条件属性集合,  $d$  是决策属性。

在决策系统中,当部分对象的条件属性值未知时,该决策系统就被称为不完备决策系统,若假设所有的属性值是丢失的,是不允许被比较的,未知的条件属性值用“?”表示。针对这种解释,Stefanowski 等人构建了非对称相似关系形如:

$$S_A = \{(x, y) \in U^2 : \forall a \in A, a(x) = ? \vee a(x) = a(y)\} \quad (1)$$

根据非对称相似关系,Stefanowski 给出了如下两种信息粒:  $S_A(x) = \{y \in U : (y, x) \in S_A\}$ , 表示所有非对称相似于  $x$  的集合;  $S_A^{-1}(x) = \{y \in U : (x, y) \in S_A\}$  表示所有  $x$  与之非对称相似的集合。

**定义 1** 设  $I$  为一不完备信息系统,  $\forall A \subseteq AT$ , 对于  $\forall X \subseteq U$ ,  $X$  根据相似关系  $S_A$  所得到的下近似集合  $\underline{A}_S(X)$  与上近似集合  $\overline{A}_S(X)$  分别定义为:

$$\underline{A}_S(X) = \{x \in U : S_A^{-1}(x) \subseteq X\}$$

$$\overline{A}_S(X) = \bigcup \{S_A(x) : x \in X\} = \{x \in U : S_A^{-1}(x) \cap X \neq \emptyset\}$$

根据定义 1, Stefanowski 的粗糙集模型显然是建立在一个相似关系的基础上的。然而根据钱宇华和梁吉业教授的多粒度粗糙集方法,可以在多个具有独立分布的粒空间上构建粗糙集模型,即可以采用多个二元关系进行目标概念的近似逼近。借鉴这个思想,很容易将多粒度粗糙集模型推广到不完备信息系统中,采用多个不同的相似关系来逼近目标概念。

**定义 2** 令  $I$  为一不完备决策系统,  $A_1, \dots, A_m \subseteq AT$ , 对于  $\forall X \subseteq U$ ,  $X$  根据相似关系族  $\{S_{A_1}, \dots, S_{A_m}\}$  所得到的乐观多粒度下近似集合  $\sum_{i=1}^m A_i^O(X)$  与乐观多粒度上近似集合  $\sum_{i=1}^m A_i(X)$  分别定义为:

$$\sum_{i=1}^m A_i^O(X) = \{x \in U : S_{A_1}^{-1}(x) \subseteq X \vee \dots \vee S_{A_m}^{-1}(x) \subseteq X\} \quad (2)$$

$$\sum_{i=1}^m A_i(X) = \sim \sum_{i=1}^m A_i^O(\sim X) \quad (3)$$

在定义 2 所示的乐观多粒度粗糙集中,某个对象属于乐观多粒度下近似当且仅当这个对象在某一个条件属性集合中所得到的非对称相似类包含于目标概念,乐观多粒度上近似采用乐观多粒度下近似补集的方式定义。

**定义 3** 令  $I$  为一不完备决策系统,  $A_1, \dots, A_m \subseteq AT$ , 对于  $\forall X \subseteq U$ ,  $X$  根据相似关系族  $\{S_{A_1}, \dots, S_{A_m}\}$  所得到的悲观多粒度下近似集合  $\sum_{i=1}^m A_i^P(X)$  与悲观多粒度上近似集合  $\sum_{i=1}^m A_i(X)$  分别定义为:

$$\sum_{i=1}^m A_i^P(X) = \{x \in U : S_{A_1}^{-1}(x) \subseteq X \wedge \dots \wedge S_{A_m}^{-1}(x) \subseteq X\} \quad (4)$$

$$\sum_{i=1}^m A_i(X) = \sim \sum_{i=1}^m A_i^P(\sim X) \quad (5)$$

在定义 3 所示的悲观多粒度粗糙集中,某个对象属于悲观多粒度下近似当且仅当这个对象在所有条件属性集合中所得到的非对称相似类都包含于目标概念,悲观多粒度上近似依然采用悲观多粒度下近似补集的方式定义。

## 3 基于多粒度的变精度粗糙集模型

变精度粗糙集模型首先是由 Ziarko 提出的。相比于传统的粗糙集模型来说,变精度方法能够处理某种程度上的包含和属于,更加适用于噪声及不完整数据的分析。为了在多粒度环境下沿袭变精度思想的优点,有必要将其引入到乐观和悲观多粒度粗糙集模型中。

**定义 4** 令  $I$  为一不完备决策系统,  $A_1, \dots, A_m \subseteq AT$ , 对于  $\forall X \subseteq U$ ,  $X$  根据相似关系族  $\{S_{A_1}, \dots, S_{A_m}\}$  所得到的基于多粒度的变精度乐观下近似集合和上近似集合分别定义为:

$$\sum_{i=1}^m A_{i\beta}^O(X) = \{x \in U : P(S_{A_1}^{-1}(x), X) \geq \beta \vee \dots \vee P(S_{A_m}^{-1}(x), X) \geq \beta\} \quad (6)$$

$$\sum_{i=1}^m A_{i\beta}(X) = \sim \sum_{i=1}^m A_{i\beta}^O(\sim X) \quad (7)$$

式中,  $\beta \in (0.5, 1]$ ,  $P(S_{A_i}^{-1}(x), X) = |S_{A_i}^{-1}(x) \cap X| / |S_{A_i}^{-1}(x)|$ 。

**定理 1** 令  $I$  为一不完备决策系统,  $A_1, \dots, A_m \subseteq AT$ , 对于  $\forall X \subseteq U$ , 有

$$\sum_{i=1}^m A_{i\beta}(X) = \{x \in U : P(S_{A_1}^{-1}(x), X) > 1 - \beta \wedge \dots \wedge P(S_{A_m}^{-1}(x), X) > 1 - \beta\} \quad (8)$$

证明: 根据定义 4, 有

$$\begin{aligned} \sum_{i=1}^m A_{i\beta}(X) &= \sim \sum_{i=1}^m A_{i\beta}^O(\sim X) \\ &= \{x \in U : P(S_{A_1}^{-1}(x), \sim X) < \beta \wedge \dots \wedge P(S_{A_m}^{-1}(x), \sim X) < \beta\} \\ &= \{x \in U : \frac{|S_{A_1}^{-1}(x) \cap (\sim X)|}{|S_{A_1}^{-1}(x)|} < \beta \wedge \dots \wedge \frac{|S_{A_m}^{-1}(x) \cap (\sim X)|}{|S_{A_m}^{-1}(x)|} < \beta\} \\ &= \{x \in U : 1 - \frac{|S_{A_1}^{-1}(x) \cap X|}{|S_{A_1}^{-1}(x)|} < \beta \wedge \dots \wedge 1 - \frac{|S_{A_m}^{-1}(x) \cap X|}{|S_{A_m}^{-1}(x)|} < \beta\} \\ &= \{x \in U : P(S_{A_1}^{-1}(x), X) > 1 - \beta \wedge \dots \wedge P(S_{A_m}^{-1}(x), X) > 1 - \beta\} \end{aligned}$$

**定义 5** 令  $I$  为一不完备决策系统,  $A_1, \dots, A_m \subseteq AT$ , 对于  $\forall X \subseteq U$ ,  $X$  根据相似关系族  $\{S_{A_1}, \dots, S_{A_m}\}$  所得到的基于多

粒度的变精度悲观下近似集合和上近似集合分别定义为:

$$\sum_{i=1}^m A_i^P(X) = \{x \in U; P(S_{A_1}^{-1}(x), X) \geq \beta \wedge \dots \wedge P(S_{A_m}^{-1}(x), X) \geq \beta\} \quad (9)$$

$$\sum_{i=1}^m A_i^P(X) = \sim \sum_{i=1}^m A_i^P(\sim X) \quad (10)$$

**定理 2** 令  $I$  为一不完备决策系统,  $A_1, \dots, A_m \subseteq AT$ , 对于  $\forall X \subseteq U$ , 有

$$\sum_{i=1}^m A_i^P(X) = \{x \in U; P(S_{A_1}^{-1}(x), X) > 1 - \beta \vee \dots \vee P(S_{A_m}^{-1}(x), X) > 1 - \beta\} \quad (11)$$

证明: 根据定义 5, 与定理 1 的证明类似。

定理 1 和定理 2 说明了虽然乐观和悲观多粒度上近似是采用下近似补集的方式定义的, 但它们仍然可以表示为一族信息粒和目标概念之间的变精度包含关系。

**定理 3** 令  $I$  为一不完备决策系统,  $A_1, \dots, A_m \subseteq AT$ , 对于  $\forall X \subseteq U$ , 基于多粒度的变精度乐观粗糙集有如下性质:

$$\sum_{i=1}^m A_i^O(U) = \sum_{i=1}^m A_i^O(U) = U \quad (12)$$

$$\sum_{i=1}^m A_i^O(\emptyset) = \sum_{i=1}^m A_i^O(\emptyset) = \emptyset$$

$$\beta_1 \leq \beta_2 \Rightarrow \sum_{i=1}^m A_i^O(X) \supseteq \sum_{i=1}^m A_i^O(X) \quad (13)$$

$$\sum_{i=1}^m A_i^O(X) \subseteq \sum_{i=1}^m A_i^O(X)$$

$$\sum_{i=1}^m A_i^O(\bigcap_{j=1}^n X_j) \subseteq \bigcap_{j=1}^n \sum_{i=1}^m A_i^O(X_j) \quad (14)$$

$$\sum_{i=1}^m A_i^O(\bigcup_{j=1}^n X_j) \supseteq \bigcup_{j=1}^n \sum_{i=1}^m A_i^O(X_j) \quad (15)$$

$$\sum_{i=1}^m A_i^O(\bigcap_{j=1}^n X_j) \subseteq \bigcap_{j=1}^n \sum_{i=1}^m A_i^O(X_j) \quad (16)$$

$$\sum_{i=1}^m A_i^O(\bigcup_{j=1}^n X_j) \supseteq \bigcup_{j=1}^n \sum_{i=1}^m A_i^O(X_j) \quad (17)$$

证明:

1) 根据定义 4, 式(12)显然成立。

2) 对于  $\forall x \in \sum_{i=1}^m A_i^O(X)$ , 根据定义 4, 必定  $\exists A_i \in \{A_1, \dots, A_m\}$ , 使得  $P(S_{A_i}^{-1}(x), X) \geq \beta_2$ 。又因为  $\beta_1 \leq \beta_2$ , 所以就有  $P(S_{A_i}^{-1}(x), X) \geq \beta_1$ , 即  $x \in \sum_{i=1}^m A_i^O(X)$ 。类似地, 易证  $\sum_{i=1}^m A_i^O(X) \subseteq \sum_{i=1}^m A_i^O(X)$ 。

3) 对于  $\forall x \in \sum_{i=1}^m A_i^O(\bigcap_{j=1}^n X_j)$ , 根据定义 4, 必定  $\exists A_i \in \{A_1, \dots, A_m\}$  使得  $P(S_{A_i}^{-1}(x), \bigcap_{j=1}^n X_j) \geq \beta$ , 从而对于  $\forall X_j (j=1, \dots, n)$ , 就有  $P(S_{A_i}^{-1}(x), X_j) \geq \beta$ , 于是根据定义 4 可知  $x \in \sum_{i=1}^m A_i^O(X_j)$ , 即  $x \in \bigcap_{j=1}^n \sum_{i=1}^m A_i^O(X_j)$ , 从而证得  $\sum_{i=1}^m A_i^O(\bigcap_{j=1}^n X_j) \subseteq \bigcap_{j=1}^n \sum_{i=1}^m A_i^O(X_j)$ 。

根据上述证明过程, 易证式(15), 式(16)和式(17)。

**定理 4** 令  $I$  为一不完备决策系统,  $A_1, \dots, A_m \subseteq AT$ , 对于  $\forall X \subseteq U$ , 基于多粒度的变精度悲观粗糙集有如下性质:

$$\sum_{i=1}^m A_i^P(U) = \sum_{i=1}^m A_i^P(U) = U$$

$$\sum_{i=1}^m A_i^P(\emptyset) = \sum_{i=1}^m A_i^P(\emptyset) = \emptyset \quad (18)$$

$$\beta_1 \leq \beta_2 \Rightarrow \sum_{i=1}^m A_i^P(X) \supseteq \sum_{i=1}^m A_i^P(X) \quad (19)$$

$$\sum_{i=1}^m A_i^P(X) \subseteq \sum_{i=1}^m A_i^P(X)$$

$$\sum_{i=1}^m A_i^P(\bigcap_{j=1}^n X_j) \subseteq \bigcap_{j=1}^n \sum_{i=1}^m A_i^P(X_j) \quad (20)$$

$$\sum_{i=1}^m A_i^P(\bigcup_{j=1}^n X_j) \supseteq \bigcup_{j=1}^n \sum_{i=1}^m A_i^P(X_j) \quad (21)$$

$$\sum_{i=1}^m A_i^P(\bigcap_{j=1}^n X_j) \subseteq \bigcap_{j=1}^n \sum_{i=1}^m A_i^P(X_j) \quad (22)$$

$$\sum_{i=1}^m A_i^P(\bigcup_{j=1}^n X_j) \supseteq \bigcup_{j=1}^n \sum_{i=1}^m A_i^P(X_j) \quad (23)$$

证明: 与定理 3 的证明过程类似。

钱宇华等人的多粒度粗糙集模型虽然建立在多个相互独立的二元关系上, 但是他仍然使用的是严格包含关系。而本文提出的基于多粒度的变精度粗糙集模型则采用了多数包含的概念, 因而有必要对这两种不同的多粒度粗糙集模型进行对比分析。

**定理 5** 令  $I$  为一不完备决策系统,  $A_1, \dots, A_m \subseteq AT$ , 对于  $\forall X \subseteq U$ , 有

$$\sum_{i=1}^m A_i^O(X) \subseteq \sum_{i=1}^m A_i^O(X) \quad (24)$$

$$\sum_{i=1}^m A_i^O(X) \supseteq \sum_{i=1}^m A_i^O(X) \quad (25)$$

$$\sum_{i=1}^m A_i^P(X) \subseteq \sum_{i=1}^m A_i^P(X) \quad (26)$$

$$\sum_{i=1}^m A_i^P(X) \supseteq \sum_{i=1}^m A_i^P(X) \quad (27)$$

证明: 对于  $\forall x \in \sum_{i=1}^m A_i^O(X)$ , 根据定义 2, 必定  $\exists A_i \in \{A_1, \dots, A_m\}$ , 使得  $S_{A_i}^{-1}(x) \subseteq X$ 。根据多数包含的定义可知  $P(S_{A_i}^{-1}(x), X) = 1 \geq \beta$ , 再由定义 4 可以看出  $x \in \sum_{i=1}^m A_i^O(X)$  显然成立, 即  $\sum_{i=1}^m A_i^O(X) \subseteq \sum_{i=1}^m A_i^O(X)$ 。类似地, 易证其他包含式。

定理 5 说明了基于多粒度的变精度粗糙集与钱宇华等人的多粒度粗糙集之间的关系。在多粒度环境下, 无论是乐观还是悲观的情形, 采用错误分类率的方法都可以得到更大的下近似集和更小的上近似集, 从而使得多粒度的粗糙近似精度得以提高, 这与单粒度环境下所得到的结论是一致的。

## 4 实例分析

根据第 3 节所示的基于多粒度的可变精度粗糙集模型, 本节以一个具体实例来进行讨论。表 1 是文献[10]中用来进行分析的一个不完备决策系统, 其中  $U = \{x_1, x_2, \dots, x_{12}\}$  为论域,  $AT = \{a_1, a_2, a_3, a_4, a_5\}$  为条件属性集合,  $d$  为决策属性。

表 1 不完备决策系统实例

| U               | a <sub>1</sub> | a <sub>2</sub> | a <sub>3</sub> | a <sub>4</sub> | a <sub>5</sub> | d |
|-----------------|----------------|----------------|----------------|----------------|----------------|---|
| x <sub>1</sub>  | 1              | 0              | ?              | 2              | 0              | 1 |
| x <sub>2</sub>  | 2              | 2              | 0              | 0              | ?              | 0 |
| x <sub>3</sub>  | 0              | 1              | 2              | ?              | 1              | 0 |
| x <sub>4</sub>  | 2              | 2              | 0              | 1              | 1              | 0 |
| x <sub>5</sub>  | 1              | 0              | 1              | 2              | 0              | 1 |
| x <sub>6</sub>  | 2              | 2              | 0              | ?              | 1              | 0 |
| x <sub>7</sub>  | ?              | ?              | 1              | 0              | 1              | 2 |
| x <sub>8</sub>  | 0              | 1              | 2              | 1              | ?              | 2 |
| x <sub>9</sub>  | 1              | 0              | 0              | 2              | 0              | 2 |
| x <sub>10</sub> | ?              | 0              | ?              | 2              | 0              | 1 |
| x <sub>11</sub> | 0              | 1              | 1              | 1              | 0              | 0 |
| x <sub>12</sub> | 0              | 1              | 1              | 1              | 0              | 1 |

根据决策属性  $d$ , 所得到的论域上的划分为  $U/\{d\} = \{X_0, X_1, X_2\} = \{\{x_2, x_3, x_4, x_6, x_{11}\}, \{x_1, x_5, x_{10}, x_{12}\}, \{x_7, x_8, x_9\}\}$ 。令  $A_1 = \{a_1, a_2\}, A_2 = \{a_3\}, A_3 = \{a_4, a_5\}$  为 3 个不同的条件属性集合,  $\beta = 0.9$ 。因而, 根据定义 4, 可得到如下所示的基于多粒度的变精度乐观近似集:

$$\begin{aligned} \sum_{i=1}^3 A_i^O(X_0) &= \{x_{10}\} \\ \sum_{i=1}^3 A_i^O(X_1) &= \{x_2, x_3, x_4, x_6\} \\ \sum_{i=1}^3 A_i^O(X_2) &= \{x_7, x_8\} \\ \sum_{i=1}^3 A_{i\beta}(X_0) &= \{x_1, x_5, x_{10}, x_{11}, x_{12}\} \\ \sum_{i=1}^3 A_{i\beta}(X_1) &= \{x_2, x_3, x_4, x_6, x_{11}, x_{12}\} \\ \sum_{i=1}^3 A_{i\beta}(X_2) &= \{x_7, x_8, x_9\} \end{aligned}$$

根据定义 5, 可得到如下所示的基于多粒度的变精度悲观近似集:

$$\begin{aligned} \sum_{i=1}^3 A_i^P(X_0) &= \emptyset \\ \sum_{i=1}^3 A_i^P(X_1) &= \emptyset \\ \sum_{i=1}^3 A_i^P(X_2) &= \emptyset \\ \sum_{i=1}^3 A_{i\beta}(X_0) &= \{x_1, x_3, x_5, x_7, x_8, x_9, x_{10}, x_{11}, x_{12}\} \\ \sum_{i=1}^3 A_{i\beta}(X_1) &= U \\ \sum_{i=1}^3 A_{i\beta}(X_2) &= U \end{aligned}$$

根据多粒度粗糙集方法, 可以从决策系统中生成决策规则, 但是这种决策规则与从 Pawlak 粗糙集模型所生成的决策规则是不同的, 因为决策规则的前件的逻辑连接词是“ $\vee$ ”而非“ $\wedge$ ”。在本文所讨论的不完备决策系统中, 根据提出的多粒度变精度粗糙集模型, 可以得到如下所示的决策规则形式:  $r_x: y \in S_{A_1}^{-1}(x) \vee \dots \vee y \in S_{A_m}^{-1}(x) \rightarrow y \in X, r_x$  的可信度记为  $C(r_x)$  且  $C(r_x) = \max\{P(S_{A_i}^{-1}(x), X); i = 1, \dots, m\}$ 。

在阈值  $\beta$  的控制下, 若决策规则的可信度大于等于  $\beta$ , 则此规则可被视作确定决策规则; 若此规则的可信度大于  $1 - \beta$ , 则此规则可被视作可能决策规则。

**定理 6** 令  $I$  为一不完备决策系统,  $A_1, \dots, A_m \subseteq AT$ , 对于  $\forall x \in U$ , 有

$$\begin{aligned} 1) x \in \sum_{i=1}^m A_i^O(X) &\Leftrightarrow C(r_x) \geq \beta; \\ 2) x \in \sum_{i=1}^m A_{i\beta}(X) &\Leftrightarrow C(r_x) > 1 - \beta. \end{aligned}$$

证明: 仅证明 1), 2) 的证明类似。

$$\begin{aligned} x \in \sum_{i=1}^m A_i^O(X) &\Leftrightarrow \exists j \in \{1, \dots, m\}, P(S_{A_j}^{-1}(x), X) \geq \beta \\ &\Leftrightarrow \max\{P(S_{A_i}^{-1}(x), X); i = 1, \dots, m\} \geq \beta \\ &\Leftrightarrow C(r_x) \geq \beta \end{aligned}$$

根据定理 6 可知, 由乐观多粒度可变精度下近似中元素所支持的决策规则是确定决策规则, 而由悲观多粒度可变精度上近似中元素所支持的决策规则是可能决策规则。

根据表 1 中计算所得出的乐观和悲观多粒度可变精度近似集, 可以生成相应的决策规则, 鉴于篇幅起见, 此处仅列出

所有的确定决策规则如下所示:

$$\begin{aligned} r_1: a_2(y) = 0 \vee (a_4(y) = 2 \wedge a_5(y) = 0) &\rightarrow y \in X_0 // \text{由对象 } x_{10} \text{ 支持} \\ r_2: (a_1(y) = 2 \wedge a_2(y) = 2) \vee a_3(y) = 0 \vee a_4(y) = 0 &\rightarrow y \in X_1 // \text{由对象 } x_2 \text{ 支持} \\ r_3: (a_1(y) = 0 \wedge a_2(y) = 1) \vee a_3(y) = 2 \vee a_5(y) = 1 &\rightarrow y \in X_1 // \text{由对象 } x_3 \text{ 支持} \\ r_4: (a_1(y) = 2 \wedge a_2(y) = 2) \vee a_3(y) = 0 \vee (a_4(y) = 1 \wedge a_5(y) = 1) &\rightarrow y \in X_1 // \text{由对象 } x_4 \text{ 支持} \\ r_5: (a_1(y) = 2 \wedge a_2(y) = 2) \vee a_3(y) = 0 \vee a_5(y) = 1 &\rightarrow y \in X_1 // \text{由对象 } x_6 \text{ 支持} \\ r_6: a_3(y) = 1 \vee (a_4(y) = 0 \wedge a_5(y) = 1) &\rightarrow y \in X_2 // \text{由对象 } x_7 \text{ 支持} \\ r_7: (a_1(y) = 0 \wedge a_2(y) = 1) \vee a_3(y) = 2 \vee a_4(y) = 1 &\rightarrow y \in X_2 // \text{由对象 } x_8 \text{ 支持} \end{aligned}$$

**结束语** 近年来, 关于多粒度粗糙集模型的研究已经成为粗糙集理论中的一个热点问题。笔者在不完备信息系统中, 考虑一族相似关系, 引入了变精度的思想, 进而提出了基于多粒度的变精度乐观和悲观粗糙集模型。相比于基于严格包含的多粒度粗糙集方法来说, 采用变精度的方法, 依然可以获得更大的下近似和更小的上近似, 使近似精度得以提高, 这与单粒度环境下的结论是一致的。

在本文工作的基础上, 笔者将分别采用分辨矩阵和启发式算法对新提出的基于多粒度的变精度粗糙集约简问题进行研究。

## 参 考 文 献

- [1] Pawlak Z. Rough sets—theoretical aspects of reasoning about data[M]. London: Kluwer Academic Publishers, 1991
- [2] Lin T Y. Granular computing: practices, theories, and future directions[M]. Encyclopedia on Complexity of Systems Science, 2009; 4339-4355
- [3] Lin T Y. Granular computing I: the concept of granulation and its formal model[J]. International Journal of Granular Computing, Rough Sets and Intelligent Systems, 2009, 1(1): 21-42
- [4] Zhu W. Relationship between generalized rough sets based on binary relation and covering[J]. Information Sciences, 2009, 179(3): 210-225
- [5] Yang X B, Zhang M, Dou H L, et al. Neighborhood systems-based rough sets in incomplete information system[J]. Knowledge-Based Systems, 2011, 24(6): 858-867
- [6] Qian Y H, Liang J Y, Yao Y Y, et al. MGRS: A multi-granulation rough set[J]. Information Sciences, 2010, 180(6): 949-970
- [7] Qian Y H, Liang J Y, Wei W. Pessimistic rough decision[C]// Second International Workshop on Rough Sets Theory. Zhoushan, China, October 2010: 440-449
- [8] Qian Y H, Liang J Y, Dang C Y. Incomplete multigranulation rough set[J]. IEEE Transactions on Systems, Man and Cybernetics, Part A, 2010, 40(2): 420-431
- [9] 王国胤. 粗糙集理论在不完备信息系统中的扩充[J]. 计算机研究与发展, 2002, 39(10): 1238-1243
- [10] Stefanowski J, Tsoukias A. Incomplete information tables and rough classification[J]. Computational Intelligence, 2001, 17(3): 545-566
- [11] 杨习贝, 於东军, 吴陈, 等. 不完备信息系统中基于相似关系的知识约简[J]. 计算机科学, 2008, 35(2): 163-165, 177