

区间值模糊决策序信息系统的分布约简*

史德容, 徐伟华⁺

重庆理工大学 理学院, 重庆 400054

Distribution Reduction in Interval-Valued Fuzzy Decision Ordered Information Systems*

SHI Derong, XU Weihua⁺

School of Sciences, Chongqing University of Technology, Chongqing 400054, China

+ Corresponding author: E-mail: chcuwh@gmail.com

SHI Derong, XU Weihua. Distribution reduction in interval-valued fuzzy decision ordered information systems. Journal of Frontiers of Computer Science and Technology, 2017, 11(4): 652-658.

Abstract: Because of the complexity and uncertainty of information systems, it's hard to use accurate value to represent the object's attribute value. The interval-value which blurred is used to deal with the issue. This paper introduces dominance relations to establish inconsistent interval-valued fuzzy ordered decision information system. Then, this paper considers distribution reduction and maximum distribution reduction to simplify the expression of knowledge, at the same time, finds out the relationship between them. Moreover, this paper obtains judgment theorem of distribution reduction and maximum distribution reduction and discernibility matrix in the system. It also provides specific method of distribution reduction and maximum distribution reduction in interval-valued fuzzy ordered decision information system. Furthermore, this paper analyzes a specific case about the venture investment and discusses the significance of study on distribution reduction. Finally, this experiment enriches rough set method for interval-valued fuzzy ordered decision information system.

Key words: rough set; interval-value; ordered information system; distribution reduction

* The National Natural Science Foundation of China under Grant Nos. 61105041, 61472463, 61402064 (国家自然科学基金); the Natural Science Foundation of Chongqing under Grant No. cstc2015jcyjA1390 (重庆市自然科学基金); the Graduate Innovation Foundation of Chongqing under Grant No. CYS16217 (重庆市研究生创新基金); the Graduate Innovation Foundation of Chongqing University of Technology under Grant Nos. YCX2015227, YCX2016227 (重庆理工大学研究生创新基金).

Received 2016-02, Accepted 2016-06.

CNKI网络优先出版: 2016-06-02, <http://www.cnki.net/kcms/detail/11.5602.TP.20160602.1144.002.html>

摘要:因信息系统的复杂性和不确定性,对象的属性值难以用精确的数值来表达,而是采用区间形式表示。针对这一问题,对区间值进一步模糊化,并引进优势关系,建立了不协调区间值模糊序决策信息系统。通过分布约简和最大分布约简来简化知识的表达,找出二者之间的关系,得到了分布约简和最大分布约简的判定定理以及可辨识属性集和可辨识矩阵;提供了不协调的区间值模糊序信息系统的分布约简和最大分布约简的具体方法;结合投资风险这一具体案例的求解分析,进一步阐述了对分布约简研究的意义,丰富了区间值模糊序决策信息系统中的粗糙集方法。

关键词:粗糙集;区间值;序信息系统;分布约简

文献标志码:A **中图分类号:**TP18

1 引言

粗糙集理论^[1-2]最早由数学家Pawlak在1982年提出,是数据分析的一种数学工具,是经典集合论的一种推广形式,其主要思想是在保持分类不变的情况下,经过属性约简推出问题的决策准则。目前,国内对粗糙集的理论基础及应用^[3-4]研究取得了很大的进步,许多学者已在该领域出版了相应的专著,同时也发表了数百篇论文。当然粗糙集^[5-6]的应用不仅仅是限制在知识理论方面,它已经在人工智能、故障检测、数据挖掘、医疗诊断、股票数据分析、模式识别、智能信息处理等领域得到了普遍的应用。

粗糙集理论的核心问题之一就是知识约简^[7-9]。约简就是实际的知识库中所描述的知识的属性并不都是同等重要的,甚至有些属性是多余的。所谓知识约简,就是从知识库中去掉一些不重要的属性,使得知识得以简化,又不丢失其主要信息。

实际问题中,事物的一些属性值介于某个范围之间,常被用来刻画信息系统^[10]中的不确定信息。为了表达这种情况,属性值通常用区间值^[11-12]形式来表示,进一步将其模糊化,这种信息系统就是区间值模糊信息系统。本文通过在带有决策的区间值模糊信息系统中引入优势关系^[13-16],建立区间值模糊序决策系统;在不协调的区间值模糊序信息系统^[15-17]中引进了分布约简以及最大分布约简^[18],探讨二者之间的关系,得到了分布约简和最大分布约简的判定定理以及辨识属性集和可辨识矩阵;提供了不协调的区间值模糊序信息系统的分布约简和最大分布约简的具体方法;同时经过例题验证了此方法的有效性,丰富了区间值模糊序决策信息系统中的粗糙集方法。

2 基于区间值模糊的序决策信息系统

信息决策系统是既有条件属性又有决策属性的一种特殊信息系统。信息决策系统主要是研究条件属性和决策属性之间的关系问题。为了便于理解,下面给出一些基本概念。

定义1^[15] 称一个四元组 $I=(U, AT \cup DT, F, G)$ 为一个决策信息系统。其中 $I=(U, AT, F)$ 是信息系统, AT 称为条件属性集, DT 称为目标属性集。

U 是有限对象集, $U=\{x_1, x_2, \dots, x_n\}$;

AT 是有限条件属性集, $AT=\{a_1, a_2, \dots, a_p\}$;

DT 是有限决策属性集, $DT=\{d_1, d_2, \dots, d_q\}$;

F 是 U 与 AT 的关系集, $F=\{f:U \rightarrow V_a, a \in AT\}$, V_a 为 a 的有限值域;

G 是 U 与 DT 的关系集, $G=\{g:U \rightarrow V_d, d \in DT\}$, V_d 为 d 的有限值域。

设 $I=(U, AT \cup DT, F, G)$ 为一个决策信息系统,若对任意 $f \in F$, $a \in AT$ 和 $x_i \in U$ 都有:

$$f(x_i, a)=[a^l(x_i), a^u(x_i)]$$

则称 $I=(U, AT, F)$ 为区间值模糊信息系统, $I=(U, AT \cup DT, F, G)$ 为区间值模糊决策信息系统。其中 $a^l(x_i)$, $a^u(x_i) \in [0, 1]$, 且 $a^l(x_i) \leq a^u(x_i)$, $f(x_i, a)$ 是 x_i 在属性 a 下的属性值范围(区间数)。特别地,当 $a^l(x_i) \equiv a^u(x_i)$ 时, $f(x_i, a)$ 就退化成了一个模糊数。因此区间值模糊信息系统是一般形式,单值模糊信息系统是其特殊形式。

设 $I=(U, AT, F)$ 为区间值模糊信息系统。对任意的 $a \in AT$, 在区间值模糊信息系统中可对属性值进行比较,定义:

$$f(x_i, a) \leq f(x_j, a) \Leftrightarrow (\forall a \in AT)[a^l(x_i) \leq a^l(x_j), a^u(x_i) \leq a^u(x_j)]$$

$$f(x_i, a) \geq f(x_j, a) \Leftrightarrow (\forall a \in AT)[a^l(x_i) \geq a^l(x_j), a^u(x_i) \geq a^u(x_j)]$$

其中“≤”和“≥”可在区间值模糊信息系统中分别构建一个递增的偏序和一个递减的偏序。如果区间值模糊信息系统中属性的值域为递增的或者递减的偏序,那么称该属性是区间值模糊信息系统中的—个准则。本文只考虑由递增偏序构成的优势关系的情景,递减的偏序情形可以类似地得到相同的结论。

定义 2 设 $I=(U, AT \cup DT, F, G)$ 为区间值模糊决策信息系统,若 I 中所有条件属性都是准则,则称 I 是区间值模糊序决策信息系统,记作 I^\geq 。称 $I^\geq=(U, AT, F)$ 为区间值模糊序信息系统。

在区间值模糊序信息系统中,设 $a \in AT$ 为准则,存在优势关系“ \geq_a ”,“ $x_j \geq_a x_i$ ”表示 x_j 关于准则 a 优于 x_i 。设 $A \subseteq AT$ 是准则集,那么 $x_j \geq_a x_i \Leftrightarrow (\forall a \in A)[x_j \geq_a x_i]$,优势关系 R_A^\geq 可定义为:

$$R_A^\geq = \{(x_i, x_j) \in U \times U \mid x_j \geq_a x_i, \forall a \in A\} = \{(x_i, x_j) \in U \times U \mid (\forall a \in A)[a^l(x_i) \leq a^l(x_j), a^u(x_i) \leq a^u(x_j)]\}$$

由区间值模糊优势关系 R_A^\geq 诱导的 $[x_i]_A^\geq$ 为:

$$[x_i]_A^\geq = \{x_j \in U \mid (x_i, x_j) \in R_A^\geq\} = \{x_j \in U \mid (\forall a \in A)[a^l(x_i) \leq a^l(x_j), a^u(x_i) \leq a^u(x_j)]\}$$

称为区间值模糊优势类,简称为优势类。用 $U/R_A^\geq = \{[x_i]_A^\geq \mid x \in U\}$ 表示论域 U 上由区间值模糊优势关系 R_A^\geq 诱导的区间值模糊优势类全体。一般地, U/R_A^\geq 中的优势类不一定构成 U 上的—个划分,而是仅仅构成 U 上的—个覆盖。

定义 3 设 $I^\geq=(U, AT \cup \{d\}, F, G)$ 为区间值模糊序决策信息系统,如果 $R_{AT}^\geq \subseteq R_{d^r}^\geq$,则称该区间值模糊序决策信息系统是协调的,否则为不协调的。

本文仅仅考虑不协调区间值模糊序决策信息系统。

3 区间值模糊序信息决策系统的分布约简

已经知道了序信息系统中属性约简理论定义的

分布函数和最大分布函数,下面将给出区间值模糊序信息系统中的分布函数和最大分布函数的定义。

定义 4 设 $I^\geq=(U, AT \cup \{d\}, F, G)$ 为区间值模糊序决策信息系统。对于任意的 $A \subseteq AT$, $x \in U$,记:

$$U/R_A^\geq = \{[x_i]_A^\geq \mid x \in U\}$$

$$U/R_d^\geq = \{D_1, D_2, \dots, D_r\}$$

$$\mu_A(x) = \left(\frac{|D_1 \cap [x]_A^\geq|}{|U|}, \frac{|D_2 \cap [x]_A^\geq|}{|U|}, \dots, \frac{|D_r \cap [x]_A^\geq|}{|U|} \right)$$

$$\gamma_A(x) = \max \left\{ \frac{|D_1 \cap [x]_A^\geq|}{|U|}, \frac{|D_2 \cap [x]_A^\geq|}{|U|}, \dots, \frac{|D_r \cap [x]_A^\geq|}{|U|} \right\}$$

称 $\mu_A(x)$ 为论域 U 上的关于准则集 A 和决策 d 的分布函数, $\gamma_A(x)$ 为论域 U 上的关于准则集 A 和决策 d 的最大分布函数。

定义 5^[15] 设 $\alpha=(a_1, a_2, \dots, a_n)$ 和 $\beta=(b_1, b_2, \dots, b_n)$ 为两个 n 维向量,若 $a_i = b_i (i=1, 2, \dots, n)$,称向量 α 等于向量 β ,记作 $\alpha = \beta$;若 $a_i \leq b_i (i=1, 2, \dots, n)$,称向量 α 小于等于向量 β ,记作 $\alpha \leq \beta$;否则如果存在某个 $i_0 (i_0 \in \{1, 2, \dots, n\})$,使得 $a_{i_0} > b_{i_0}$,称向量 α 不小于等于向量 β ,记作 $\alpha \not\leq \beta$ 。

显然由以上可立即得到下面命题。

定理 1 设 $I^\geq=(U, AT \cup \{d\}, F, G)$ 为区间值模糊序决策信息系统。对于任意的 $A \subseteq AT$:

- (1) 对 $\forall x \in U$, 当 $B \subseteq A$ 时,有 $\mu_A(x) \leq \mu_B(x)$;
- (2) 对 $\forall x \in U$, 当 $B \subseteq A$ 时,有 $\gamma_A(x) \leq \gamma_B(x)$;
- (3) 对 $\forall x, y \in U$, 当 $[y]_A^\geq \subseteq [x]_A^\geq$ 时,有 $\mu_A(y) \leq \mu_A(x)$;
- (4) 对 $\forall x, y \in U$, 当 $[y]_A^\leq \subseteq [x]_A^\leq$ 时,有 $\gamma_A(y) \leq \gamma_A(x)$ 。

定义 6 设 $I^\geq=(U, AT \cup \{d\}, F, G)$ 为区间值模糊序决策信息系统。 $A \subseteq AT$,对于任意的 $x \in U$,如果有 $\mu_A(x) = \mu_{AT}(x)$,则称 A 是 I^\geq 中关于区间值模糊优势关系 R_{AT}^\geq 的分布协调集;如果 A 的任何真子集均不是分布协调集,则称 A 是 I^\geq 中关于区间值模糊优势关系 R_{AT}^\geq 的分布约简。

定义 7 设 $I^\geq=(U, AT \cup \{d\}, F, G)$ 为区间值模糊序决策信息系统。 $A \subseteq AT$,对于任意的 $x \in U$,如果有 $\gamma_A(x) = \gamma_{AT}(x)$,则称 A 是最大分布协调集;若 A 的任何真子集均不是最大分布协调集,则称 A 是 I^\geq 中关于

区间值模糊优势关系 R_{AT}^{\geq} 的最大分布约简。

下面具体给出区间值模糊序决策信息系统的分布约简与最大分布约简的关系及其判定定理。

定理 2 设 $I^{\geq}=(U,AT\cup\{d\},F,G)$ 为区间值模糊序决策信息系统,则分布协调集一定是最大分布协调集。

证明 由定义直接可得。□

推论 1 设 $I^{\geq}=(U,AT\cup\{d\},F,G)$ 为区间值模糊序决策信息系统,则分布协调集一定是最大分布约简。

定理 3 设 $I^{\geq}=(U,AT\cup\{d\},F,G)$ 为区间值模糊序决策信息系统, $A\subseteq AT$, 则 A 是分布协调集当且仅当对 $\forall x,y\in U$, 当 $\mu_{AT}(y)\prec\mu_{AT}(x)$ 时有 $[y]_A^{\geq}\not\subseteq[x]_A^{\geq}$ 。

证明 反证法。

假设当 $\mu_{AT}(y)\prec\mu_{AT}(x)$ 时有 $[y]_A^{\geq}\subseteq[x]_A^{\geq}$ 不成立, 则有 $[y]_A^{\geq}\not\subseteq[x]_A^{\geq}$, 由定理 1 可知 $\mu_A(y)\leq\mu_A(x)$ 。由于是分布协调集, 可得 $\mu_{AT}(x)=\mu_A(x)$ 且 $\mu_{AT}(y)=\mu_A(y)$, 故有 $\mu_{AT}(y)=\mu_{AT}(x)$, 与 $\mu_{AT}(y)\prec\mu_{AT}(x)$ 矛盾。□

用同样的方法可以得到最大分布协调集的充要条件。

定理 4 设 $I^{\geq}=(U,AT\cup\{d\},F,G)$ 为区间值模糊序决策信息系统, $A\subseteq AT$ 是最大分布协调集, 则对 $\forall x,y\in U$, 当 $\gamma_{AT}(y)\prec\gamma_{AT}(x)$ 时有 $[y]_A^{\geq}\not\subseteq[x]_A^{\geq}$ 。

4 区间值模糊序决策信息系统的分布约简方法

第 3 章给出了不协调的区间值模糊序决策信息系统的分布协调集和最大分布协调集的等价刻画, 这是判断准则集是否协调的理论所在, 因此可以得出分布约简和最大分布约简的方法。下面给出辨识属性集以及辨识属性矩阵的相关概念。

定义 8 设 $I^{\geq}=(U,AT,F)$ 为区间值模糊序信息系统, 记:

$$Dis_{\geq AT}^{\mu}(x_i,x_j)=\{a\in AT \mid (x_i,x_j)\notin R_a^{\geq}=\{a\in AT \mid a^l(x_i)>a^l(x_j)\text{或}a^u(x_i)>a^u(x_j)\}\}$$

称 $Dis_{\geq AT}^{\mu}(x_i,x_j)$ 为 I^{\geq} 中 x_i, x_j 关于区间值模糊优势关系 R_{AT}^{\geq} 的分布可辨识属性集。记:

$$Dis_{\geq AT}^{\mu}=(Dis_{\geq AT}^{\mu}(x_i,x_j))_{|U|\times|U|}$$

称 $Dis_{\geq AT}^{\mu}$ 为 I^{\geq} 中 x_i, x_j 关于区间值模糊优势关系 R_{AT}^{\geq} 的分布可辨识矩阵。特别地, 对任意 $x_i,x_j\in U$, 有:

$$Dis_{\geq AT}^{\mu}(x_i,x_i)=\emptyset$$

$$Dis_{\geq AT}^{\mu}(x_i,x_j)\cap Dis_{\geq AT}^{\mu}(x_j,x_i)=\emptyset$$

定理 5 设 $I^{\geq}=(U,AT\cup\{d\},F,G)$ 为区间值模糊序决策信息系统, $A\subseteq AT$, 对象关于 R_{AT}^{\geq} 的分布可辨识属性集为 $Dis_{\geq AT}^{\mu}(x_i,x_j)$, 则:

$$R_A^{\geq}=R_{AT}^{\geq}\Leftrightarrow(Dis_{\geq AT}^{\mu}(x_i,x_j)\neq\emptyset)[A\cap Dis_{\geq AT}^{\mu}(x_i,x_j)\neq\emptyset]$$

证明 必要性。由 $R_A^{\geq}=\bigcap_{a\in A}R_a^{\geq}$, 对于任意的 $a\in A$, 如果 $(x_i,x_j)\notin R_a^{\geq}$, 则有 $(x_i,x_j)\notin R_A^{\geq}$ 。设 $R_A^{\geq}=R_{AT}^{\geq}$, 由区间值模糊优势关系的定义, 对任意的 $x_i\in U$, 有 $[x_i]_A^{\geq}=[x_i]_{AT}^{\geq}$ 。又 $(x_i,x_j)\notin R_{AT}^{\geq}$, $(x_i,x_j)\notin R_A^{\geq}$, 当 $Dis_{\geq AT}^{\mu}(x_i,x_j)\neq\emptyset$ 时, 存在 $a\in AT$ 使得 $(x_i,x_j)\notin R_a^{\geq}$ 。因而存在 $a'\in A$ 使得 $(x_i,x_j)\notin R_{a'}^{\geq}$, 于是 $a'\in Dis_{\geq AT}^{\mu}(x_i,x_j)$, 由 $A\subseteq AT$ 及分布可辨识公式定义得 $Dis_{\geq A}^{\mu}(x_i,x_j)\subseteq Dis_{\geq AT}^{\mu}(x_i,x_j)$, 因此 $a'\in Dis_{\geq AT}^{\mu}(x_i,x_j)$ 。即证得:

$$R_A^{\geq}=R_{AT}^{\geq}\Leftrightarrow(Dis_{\geq AT}^{\mu}(x_i,x_j)\neq\emptyset)[A\cap Dis_{\geq AT}^{\mu}(x_i,x_j)\neq\emptyset]$$

充分性。对任意 $(x_i,x_j)\notin R_{AT}^{\geq}$, 存在 $a\in AT$ 满足 $(x_i,x_j)\notin R_a^{\geq}$, 那么 $Dis_{\geq AT}^{\mu}(x_i,x_j)\neq\emptyset$ 。从而 $A\cap Dis_{\geq AT}^{\mu}(x_i,x_j)\neq\emptyset$ 。于是存在 $a\in A$ 使得 $a\in Dis_{\geq AT}^{\mu}(x_i,x_j)$, 即 $a\in A$ 且 $(x_i,x_j)\notin R_a^{\geq}$, 得到 $(x_i,x_j)\notin R_A^{\geq}$, 即 $(x_i,x_j)\notin R_{AT}^{\geq}\Rightarrow(x_i,x_j)\notin R_A^{\geq}$ 。其逆否命题 $(x_i,x_j)\in R_A^{\geq}\Rightarrow(x_i,x_j)\in R_{AT}^{\geq}$ 。由此得到 $R_A^{\geq}\subseteq R_{AT}^{\geq}$ 。另一方面由 $A\subseteq AT$ 显然可得 $R_A^{\geq}\supseteq R_{AT}^{\geq}$, 即证得 $R_A^{\geq}=R_{AT}^{\geq}$ 。□

定义 9 设 $I^{\geq}=(U,AT\cup\{d\},F,G)$ 为区间值模糊序决策信息系统, 辨识矩阵为 $Dis_{\geq AT}^{\mu}$ 。称

$$M_{\geq AT}^{\mu}=\bigwedge\{\bigvee\{a \mid a\in Dis_{\geq AT}^{\mu}(x_i,x_j)\} \mid \forall x_i,x_j\in U\}$$

为该区间值模糊序决策信息系统的分布可辨识公式。

定理 6 设 $I^{\geq}=(U,AT\cup\{d\},F,G)$ 为区间值模糊序决策信息系统, 分布辨识公式 $M_{\geq AT}^{\mu}$ 的极小析取范式为 $M_{\geq \min}^{\mu}=\bigvee_{k=1}^p(\bigwedge_{s=1}^{q_k}a_s)$, 若记 $B_{\mu}^k=\{a_s, s=1,2,\dots,q_k\}$, 则 $\{B_{\mu}^k, k=1,2,\dots,p\}$ 是所有分布约简形式的集合。

5 实例分析

设 $I^{\geq}=(U,AT\cup\{d\},F,G)$ 为区间值模糊序决策信息系统; $U=\{x_1,x_2,\dots,x_6\}$ 为论域, 代表 6 个投资对象;

$A = \{a_1, a_2, a_3\}$, 分别代表着市场风险、技术风险、管理风险; $\{d\}$ 为决策属性, 表示风险, 其中 3 表示“高”, 2 表示“中”, 1 表示“低”。统计数据如表 1。

Table 1 Interval-valued fuzzy ordered decision information system in venture investment

表 1 风险投资的区间值模糊序决策信息系统

U	a_1	a_2	a_3	d
x_1	[0.1,0.3]	[0.2,0.3]	[0.1,0.4]	3
x_2	[0.3,0.5]	[0.2,0.6]	[0.2,0.8]	2
x_3	[0.1,0.5]	[0.1,0.4]	[0.2,0.7]	1
x_4	[0.2,0.7]	[0.1,0.5]	[0.3,0.7]	2
x_5	[0.3,0.6]	[0.3,0.7]	[0.2,0.9]	3
x_6	[0.3,0.9]	[0.2,0.7]	[0.3,0.8]	1

由表 1 可得到:

$$[x_1]_A^{\geq} = \{x_1, x_2, x_5, x_6\}$$

$$[x_2]_A^{\geq} = \{x_2, x_5, x_6\}$$

$$[x_3]_A^{\geq} = \{x_2, x_3, x_4, x_5, x_6\}$$

$$[x_4]_A^{\geq} = \{x_4, x_6\}$$

$$[x_5]_A^{\geq} = \{x_5\}$$

$$[x_6]_A^{\geq} = \{x_6\}$$

$$[x_1]_d^{\geq} = [x_5]_d^{\geq} = \{x_1, x_5\}$$

$$[x_2]_d^{\geq} = [x_4]_d^{\geq} = \{x_1, x_2, x_4, x_5\}$$

$$[x_3]_d^{\geq} = [x_6]_d^{\geq} = \{x_1, x_2, x_3, x_4, x_5, x_6\}$$

显然, $R_A^{\geq} \not\subseteq R_d^{\geq}$ 。因此该区间值模糊序决策信息系统是不协调的。

对于表 1 给出的关于风险投资的区间值模糊序决策信息系统, 求分布约简。

情形 1 利用定义 6、定义 7 求解。

在该系统中记:

$$D_1 = [x_1]_d^{\geq} = [x_5]_d^{\geq}$$

$$D_2 = [x_2]_d^{\geq} = [x_4]_d^{\geq}$$

$$D_3 = [x_3]_d^{\geq} = [x_6]_d^{\geq}$$

由分布函数 $\mu_A(x)$ 及其最大分布函数 $\gamma_A(x)$ 定义可得:

$$\mu_A(x_1) = \left(\frac{1}{3}, \frac{1}{2}, \frac{2}{3}\right), \gamma_A(x_1) = \frac{2}{3}$$

$$\mu_A(x_2) = \left(\frac{1}{6}, \frac{1}{3}, \frac{1}{2}\right), \gamma_A(x_2) = \frac{1}{2}$$

$$\mu_A(x_3) = \left(\frac{1}{6}, \frac{1}{2}, \frac{5}{6}\right), \gamma_A(x_3) = \frac{5}{6}$$

$$\mu_A(x_4) = \left(0, \frac{1}{6}, \frac{1}{3}\right), \gamma_A(x_4) = \frac{1}{3}$$

$$\mu_A(x_5) = \left(\frac{1}{6}, \frac{1}{6}, \frac{1}{6}\right), \gamma_A(x_5) = \frac{1}{6}$$

$$\mu_A(x_6) = \left(0, 0, \frac{1}{6}\right), \gamma_A(x_6) = \frac{1}{6}$$

当取 $B = \{a_2, a_3\}$ 时, 容易验证对于 $\forall x \in U$, 有 $[x]_A^{\geq} = [x]_B^{\geq}$, 因此有 $\mu_B(x) = \mu_A(x)$ 和 $\gamma_B(x) = \gamma_A(x)$ 。故 $B = \{a_2, a_3\}$ 是个分布协调集, 也是个最大分布协调集。进一步可以计算 $\{a_2\}, \{a_3\}$ 均不是分布协调集和最大分布协调集, 因此 $B = \{a_2, a_3\}$ 是分布约简, 是一个最大分布约简。

容易验证 $\{a_1, a_3\}, \{a_1, a_2\}$ 和 $\{a_1\}$ 都不是分布协调集, 也不是最大分布协调集, 因此该区间值模糊决策信息系统只有一个分布约简, 也只有一个最大分布约简, 即 $\{a_2, a_3\}$ 。

情形 2 利用定理 5 求解。

该信息系统的分布辨识矩阵如表 2 所示。

Table 2 Distribution discernibility matrix of interval-valued fuzzy ordered decision information system

表 2 区间值模糊序决策信息系统的分布辨识矩阵

$Dis_{\geq A}^{\mu}$	x_1	x_2	x_3	x_4	x_5	x_6
x_1	\emptyset	\emptyset	a_2	a_2	\emptyset	\emptyset
x_2	A	\emptyset	A	A	\emptyset	\emptyset
x_3	A	\emptyset	\emptyset	\emptyset	\emptyset	\emptyset
x_4	A	a_1, a_3	A	\emptyset	a_1, a_3	\emptyset
x_5	A	A	A	A	\emptyset	a_2, a_3
x_6	A	A	A	A	a_1, a_3	\emptyset

由定义 9 可得:

$$M_{\geq AT}^{\mu} = a_2 \wedge (a_1 \vee a_3) \wedge (a_1 \vee a_2 \vee a_3) \wedge (a_2 \vee a_3) = a_2 \wedge a_3$$

因此 $\{a_2, a_3\}$ 是该区间值模糊序决策信息系统的分布约简, 而且也是最大分布约简。

这里针对情形 2 提出如下算法:

算法 区间值模糊序决策信息系统分布约简

输入: 区间值模糊序决策信息系统 $F^{\geq} = (U, AT \cup \{d\}, F, G)$

输出: 分布约简

1. Begin
2. for each $x \in U$
3. compute: $[x]_{AT}^{\geq}$ //求解论域 U 中对象 x 的优势类。
4. end
5. for each $x_i \in U$ //循环辨识矩阵的所有行。
6. for each $x_j \in U$ //循环辨识矩阵的所有列。
7. $Dis_{\geq AT}^{\mu}(x_i, x_j) \leftarrow \emptyset$ //初始化对象 x_i 和 x_j 辨识项为空集。
8. for each $a \in AT$ //循环所有条件属性集求辨识矩阵。
9. if $a^l(x_i) > a^l(x_j)$ or $a^u(x_i) > a^u(x_j)$
10. $Dis_{\geq AT}^{\mu}(x_i, x_j) \leftarrow Dis_{\geq AT}^{\mu}(x_i, x_j) \cup \{a\}$
11. end
12. end
13. end
14. end
15. $M_{\geq AT}^{\mu} = \bigwedge \{a \mid a \in Dis_{\geq AT}^{\mu}(x_i, x_j)\} \mid \forall x_i, x_j \in U\}$
16. compute: $M_{\geq AT}^{\mu}$ 的极小析取范式 $M_{\geq \min}^{\mu} = \bigvee_{k=1}^p \left(\bigwedge_{s=1}^{q_k} a_s \right)$
17. return: $M_{\geq \min}^{\mu}$ 的所有极小项 //每一个极小项都是一个约简。

end

算法的时间复杂度如表3所示。

Table 3 Time complexity on case 2

表3 情形2算法的时间复杂度

计算步骤	时间复杂度
步骤2~步骤4	$O(U^2)$
步骤5~步骤14	$O(AT \times U^2)$
步骤15~步骤16	$O(U^2)$
总时间复杂度	$O((AT+2) \times U^2)$

上述情形1和情形2所求得的结果是一致的,显然在该决策问题中技术和管理风险因子是不可缺少的,但是其时间复杂性不一样。情形1从定义出发,时间复杂度较高,不太具有可行性;而情形2的时间复杂度为 $O((AT+2) \times U^2)$ 。在这个例子中体现了求分布约简时,利用情形2求解具有明显的时间优势。

6 结束语

本文在不协调的区间值模糊序决策信息系统中

通过分析分布约简的性质得到了对应的判定定理以及辨识矩阵,建立了获取这种约简的具体方法,并且用两种情形对实例进行了对比分析。通过比较可以知道,本文对分布约简进行了更精确的刻画,可以简化在时间上的求解过程。

References:

- [1] Pawlak Z. Rough sets: theoretical aspects of reasoning about data[M]. Boston: Kluwer Academic Publishers, 1991.
- [2] Pawlak Z, Grzymalabusse J, Slowinski R, et al. Rough sets [J]. Communication of the ACM, 1995, 38(11): 88-95.
- [3] Wang Yu, Miao Duoqian, Zhou Yujian. A summary of the theory and application on rough set[J]. Pattern Recognition and Artificial Intelligence, 1996, 9(4): 337-344.
- [4] Miao Duoqian, Wang Yu. Rough sets based approach for multivariate decision tree construction[J]. Journal of Software, 1997, 8(6): 425-431.
- [5] Zhang Xiaohong, Pei Daowu, Dai Jianhua. Fuzzy mathematics and the rough set theory[M]. Beijing: Tsinghua University Press, 2013.
- [6] Xu Weihua, Zhang Xiantao, Wang Qiaorong. Experimental computing on attribute by Matlab in dominance-based variable precision rough set[J]. Journal of Chongqing University of Technology, 2013, 27(1): 107-115.
- [7] Xu Weihua, Zhang Wenxiu. Knowledge reductions in inconsistent information systems based on dominance relations [J]. Computer Science, 2006, 33(2): 182-184.
- [8] Mi Jusheng, Wu Weizhi, Zhang Wenxiu. Comparative studies of knowledge reductions in inconsistent systems[J]. Fuzzy Systems and Mathematics, 2003, 17(3): 54-60.
- [9] Zhang Wenxiu, Mi Jusheng, Wu Weizhi. Knowledge reductions in inconsistent information systems[J]. Chinese Journal of Computers, 2003, 26(1): 12-18.
- [10] Zhang Wenxiu, Liang Yi, Wu Weizhi. Information system and knowledge discovery[M]. Beijing: Science Press, 2003.
- [11] Yu Yiingying, Zeng Xuelan, Sun Xingxing. Interval-valued information systems based on dominance relation and its attribute reduction[J]. Computer Engineering and Applications, 2011, 47(35): 122-124.
- [12] Zhang Nan, Miao Duoqian, Yue Xiaodong. Approaches to knowledge reduction in interval-valued information system

- [J]. Computer Research and Development, 2010, 47(8): 1362-1371.
- [13] Greco S, Matarazzo B, Slowinski R. Rough approximation of a preference relation by dominance relations[J]. European Journal of Operational Research, 1999, 117(1): 63-68.
- [14] Xu Weihua, Zhang Wenxiu. Consistent approximation spaces based on dominance relations[J]. Computer Science, 2005, 32(9): 164-165.
- [15] Xu Weihua. Ordered information systems and rough sets theory[M]. Beijing: Science Press, 2013.
- [16] Yu Jianhang, Xu Weihua. Rough set based on logical disjoint operation of variable precision and grade in ordered information system[J]. Journal of Frontiers of Computer Science and Technology, 2015, 9(1): 112-118.
- [17] Sun Wenxin, Zhuo Chunying, Wang Guodong, et al. Generalized multi-granulation rough set in ordered information system[J]. Journal of Frontiers of Computer Science and Technology, 2015, 9(3): 376-384.
- [18] Xu Weihua, Zhang Wenxiu. Distribution reduction in inconsistent information systems based on dominance relations[J]. Fuzzy Systems and Mathematics, 2007, 21(4): 124-131.

附中文参考文献:

- [5] 张小红, 裴道武, 代建华. 模糊数学与 Rough 集理论[M]. 北京: 清华大学出版社, 2013.
- [9] 张文修, 米据生, 吴伟志. 不协调目标信息系统的知识约简[J]. 计算机学报, 2003, 26(1): 12-18.
- [10] 张文修, 梁怡, 吴伟志. 信息系统与知识发现[M]. 北京: 科学出版社, 2003.
- [15] 徐伟华. 序信息系统与粗糙集[M]. 北京: 科学出版社, 2013.
- [16] 余建航, 徐伟华. 序信息系统下变精度与程度的“逻辑或”粗糙集[J]. 计算机科学与探索, 2015, 9(1): 112-118.
- [17] 孙文鑫, 卓春英, 王国栋, 等. 序信息系统的一般多粒度粗糙集[J]. 计算机科学与探索, 2015, 9(3): 376-384.



SHI Derong was born in 1991. She is an M.S. candidate at Chongqing University of Technology. Her research interest is the mathematical foundation of artificial intelligence.

史德容(1991—),女,重庆人,重庆理工大学硕士研究生,主要研究领域为人工智能的数学基础。



XU Weihua was born in 1979. He received the Ph.D. degree from Xi'an Jiaotong University in 2007. Now he is the vice-dean, professor and M.S. supervisor at School of Sciences, Chongqing University of Technology, and the director of Chongqing Mathematical Society. His research interests include artificial intelligence, granular computing, fuzzy mathematics and rough set, etc.

徐伟华(1979—),男,山西浑源人,2007年于西安交通大学获得博士学位,现为重庆理工大学理学院副院长、教授、硕士生导师,主要研究领域为人工智能,粒计算,模糊数学,粗糙集等。