



基于信息粒的不协调决策形式背景的分布约简

王霞, 李俊余, 吴伟志

(浙江海洋大学 信息工程学院, 浙江 舟山 316022)

摘要 粒计算和知识约简是知识发现和数据挖掘的两个重要课题。基于信息粒研究不协调决策形式背景的属性约简的定义和方法。首先,利用对象概念的内涵定义对象集上的拟序关系,并研究其相关性质。然后,利用拟序类定义分布函数和最大分布函数,进而提出不协调决策形式背景的(最大)分布协调集和(最大)分布约简的定义。最后,定义不协调决策形式背景的(最大)分布辨识矩阵及(最大)分布辨识公式,基于辨识矩阵给出(最大)分布协调集的判定定理,并提出计算分布约简和最大分布约简的方法。

关键词 信息粒; 不协调决策形式背景; 分布约简; 最大分布约简; 辨识矩阵

中图分类号: TP301 **DOI**: 10.16152/j.cnki.xdxbzr.2024-04-011

Distribution reduction of inconsistent formal decision contexts based on information granules

WANG Xia, LI Junyu, WU Weizhi

(School of Information Engineering, Zhejiang Ocean University, Zhoushan 316022, China)

Abstract Granular computing and knowledge reduction are two important topics in knowledge discovery and data mining. Based on information granules, this paper studies the definition and method of attribute reduction for inconsistent formal decision contexts. First, a quasi-ordering relation on the object set is defined and its related properties are studied too. Then, a distribution function and a maximum distribution function are defined using the quasi-ordering classes. Moreover, a distribution reduct and a maximum distribution reduct are proposed for the inconsistent formal decision context. Finally, a (maximum) distribution discernibility matrix and the corresponding distribution discernibility functions are introduced into the inconsistent formal decision context. And the judgment theorems of the (maximum) distribution consistent set are given to calculate all the distribution reducts and the maximum distribution reducts.

Keywords information granule; inconsistent formal decision context; distribution reduct; maximum distribution reduct; discernibility matrix

粒度计算是知识表示和数据挖掘中的一个基本问题。粒是粒计算的一个基本概念,它是以不可区分性、相似性或功能性标准聚集在一起的一簇对象^[1]。而粒计算的主要研究方向有粒的构

建、解释及粒的表示。文献[2]将粒的概念引入到形式概念分析中,构建了形式背景的信息粒和对象粒。形式概念分析^[3]是由德国数学 Wille 于 1982 年提出的,其基本概念为形式背景和形式概

念。形式背景是一个三元组,它由对象集、属性集和二元关系构成,其中二元关系指定了哪些对象具有哪些属性。而形式概念是由一个对象集(称为外延)和一个属性集(称为内涵)构成,其中内涵恰好是由外延中的对象共同具有的那些属性组成,外延恰好是由共同具有内涵中所有属性的那些对象组成。特别地,由单个对象生成的形式概念称为对象概念,而所有对象概念的集合是反映概念格结构的信息粒,并称对象概念的外延集为概念格的对象粒^[2]。基于对象粒,文献[2]还提出了形式背景的粒约简。一个粒约简是保持形式背景的所有对象粒不变的一个极小属性子集。

文献[2-25]从不同的角度或根据不同的目的研究了形式背景的属性约简。其中文献[16-22]在带有决策的形式背景中引入了形式背景的几种协调性的概念,并进一步研究了协调决策形式背景的不同属性约简方法。文献[23]基于不协调决策形式背景提出了分布约简和最大分布约简的概念。文献[24]在多粒度决策形式背景中引入协调粒度层的信息熵及条件信息熵,提出协调粒度约简方法、最粗协调粒度约简方法、最细协调粒度约简方法及其实现算法。文献[25]定义了多源决策形式背景及其属性约简。在实际问题中,由于预测能力、数据中的噪声等各种因素导致很多决策形式背景是不协调的。由于不协调性,从决策形式背景中提取有用的信息会变得更加复杂和困难。而属性约简可以使隐藏在决策形式背景中的知识更加清晰,使决策形式背景的知识表示更加简洁,使规则集对决策形式背景的适应性更强。Wu 等^[2]基于对象粒提出了协调决策形式背景的属性约简方法,但是并未研究不协调决策形式背景的属性约简的相关问题。在不协调决策形式背景下通常需要构造恰当的函数借此对形式背景进行约简。

本文首先利用对象粒定义决策形式背景上的分布函数和最大分布函数,进而研究对象粒下不协调决策形式背景的(最大)分布约简的定义及方法。

1 预备知识

本节介绍形式概念分析的一些基本概念和性质。

定义 1^[3] 形式背景 (U,A,I) 由两个集合 U

和 A ,以及关系 $I \subseteq U \times A$ 构成。称 U 中的元素为形式背景的对象, A 中的元素为形式背景的属性。

设 (U,A,I) 为形式背景, $\forall X \subseteq U, \forall B \subseteq A$, 定义如下算子:

$$X^* = \{a \in A \mid \forall x \in X, (x,a) \in I\},$$

$$B^* = \{x \in U \mid \forall a \in B, (x,a) \in I\}.$$

定义 2^[3] 设 (U,A,I) 为形式背景且 $B \subseteq A$, 称 (U,B,I_B) 为 (U,A,I) 的子形式背景,其中 $I_B = I \cap (U \times B)$ 。

$\forall B \subseteq A$, 记 ^{*B} 为定义在子形式背景 (U,B,I_B) 上的算子。显然, $X^{*B} = X^{*A} \cap B$ 。

定义 3^[3] 设 (U,A,I) 为形式背景。 $\forall X \subseteq U, B \subseteq A$, 若 $X = B^*$ 且 $B = X^*$, 则称 (X,B) 为形式背景 (U,A,I) 的形式概念。称 X 和 B 分别为形式概念 (X,B) 的外延和内涵。

$\forall x \in U, a \in A$, 记 $x^* = \{x\}^*, a^* = \{a\}^*$, 分别称 (x^{**}, x^*) 和 (a^*, a^{**}) 为对象概念和属性概念。

对任意两个形式概念 (X_1, B_1) 和 (X_2, B_2) , 定义其下确界和上确界:

$$\begin{aligned} (X_1, B_1) \wedge (X_2, B_2) &= \\ & (X_1 \cap X_2, (B_1 \cup B_2)^{**}); \\ (X_1, B_1) \vee (X_2, B_2) &= \\ & ((X_1 \cup X_2)^{**}, B_1 \cap B_2). \end{aligned}$$

记 $L(U,A,I) = \{(X,B) \mid X = B^*, B = X^*\}$ 称其为形式背景 (U,A,I) 的概念格^[3]。

容易验证, $\forall (X,B) \in L(U,A,I)$, 当 $X \neq \emptyset$ 时, $(X,B) = \bigvee_{x \in X} (x^{**}, x^*)$; 当 $B \neq \emptyset$ 时, $(X,B) = \bigwedge_{a \in B} (a^*, a^{**})$ 。文献[2] 称所有的对象概念 $\{(x^{**}, x^*) \mid x \in U\}$ 为概念格 $L(U,A,I)$ 的信息粒, 并称所有对象概念的外延 $\{x^{**} \mid x \in U\}$ 为概念格 $L(U,A,I)$ 的对象粒。

性质 1^[3] 设 (U,A,I) 为形式背景, X, X_1, X_2 为任意的对象子集, B, B_1, B_2 为任意的属性子集, 则有以下性质:

- 1) 若 $X_1 \subseteq X_2$, 则 $X_1^* \supseteq X_2^*$;
若 $B_1 \subseteq B_2$, 则 $B_1^* \supseteq B_2^*$ 。
- 2) $X \subseteq X^{**}; B \subseteq B^{**}$ 。
- 3) $X^* = X^{***}; B^* = B^{***}$ 。
- 4) $(X_1 \cup X_2)^* = X_1^* \cap X_2^*$;
 $(B_1 \cup B_2)^* = B_1^* \cap B_2^*$ 。
- 5) $(X_1 \cap X_2)^* \supseteq X_1^* \cup X_2^*$;
 $(B_1 \cap B_2)^* \supseteq B_1^* \cup B_2^*$ 。

定义 4^[4] 设 (U,A,I) 和 (U,C,J) 是两个形式背景,有相同的对象集。称 (U,A,I,C,J) 为决策形式背景,其中 $I \subseteq U \times A, J \subseteq U \times C, A \cap C = \emptyset$ 。分别称 A 和 C 为条件属性集和决策属性集。

定义 5^[2] 设 (U,A,I,C,J) 为决策形式背景。若 $x^{*A * A} \subseteq x^{*C * C}, \forall x \in U$, 则称 (U,A,I,C,J) 为粒协调的。否则,称 (U,A,I,C,J) 为粒不协调的。

定义 6^[2] 设 (U,A,I,C,J) 为粒协调决策形式背景, $B \subseteq A$ 。 $\forall x \in U$, 若 $x^{*B * B} \subseteq x^{*C * C}$, 则称 B 为 (U,A,I,C,J) 的粒协调集。

2 主要结果

设 (U,A,I,C,J) 为决策形式背景。 $\forall B \subseteq A$, 定义对象集 U 上的二元关系

$$R_B^* = \{(x,y) \in U \times U \mid x^{*B} \subseteq y^{*B}\}。$$

显然, R_B^* 满足自反性和传递性, 但不一定满足对称性, 即 R_B^* 为对象集 U 上的拟序关系。记 $[x]_{R_B^*} = \{y \in U \mid (x,y) \in R_B^*\}$, 称其为 x 关于 R_B^* 的条件拟序类。

性质 2 设 (U,A,I) 为形式背景, 则有

- 1) $\forall B \subseteq A$, 若 $B \neq \emptyset$, 则 $R_B^* = \bigcap_{a \in B} R_a^*$ 。
- 2) 若 $B_1 \subseteq B_2 \subseteq A$, 则 $R_{B_2}^* \subseteq R_{B_1}^*$ 。
- 3) $\forall x,y \in U$, 若 $y \in [x]_{R_B^*}$, 则 $[y]_{R_B^*} \subseteq [x]_{R_B^*}$ 。
- 4) $\forall x \in U, [x]_{R_B^*} = x^{*B * B}$ 。

证明 1) 由拟序关系的定义知, $\forall (x,y) \in R_B^* \Leftrightarrow x^{*B} \subseteq y^{*B} \Leftrightarrow \forall a \in B, (x,y) \in R_a^* \Leftrightarrow (x,y) \in \bigcap_{a \in B} R_a^*$, 即 $R_B^* = \bigcap_{a \in B} R_a^*$ 。

- 2) 由性质 2 中的条件 1) 可得条件 2)。
- 3) 由拟序类的定义知, 若 $y \in [x]_{R_B^*}$, 则 $x^{*B} \subseteq y^{*B}$ 。因为 $\forall z \in [y]_{R_B^*}$, 有 $y^{*B} \subseteq z^{*B}$ 。因此有 $x^{*B} \subseteq y^{*B} \subseteq z^{*B}$, 即 $z \in [x]_{R_B^*}$ 。所以 $[y]_{R_B^*} \subseteq [x]_{R_B^*}$ 。

4) 由拟序类的定义知, $\forall y \in [x]_{R_B^*}, x^{*B} \subseteq y^{*B} \Leftrightarrow y \in x^{*B * B}$, 即 $[x]_{R_B^*} = x^{*B * B}$ 。

由性质 2 中条件 4) 可知 $[x]_{R_A^*} = x^{*A * A}$, 即形式背景 (U,A,I) 的每个对象概念的外延是该对象的拟序类。根据定义 5 和性质 2 可得引理 1 和引理 2。

引理 1 设 (U,A,I,C,J) 为决策形式背景, 则以下条件等价:

- 1) (U,A,I,C,J) 为粒协调的;
- 2) $\forall x \in U, [x]_{R_A^*} \subseteq [x]_{R_C^*}$;
- 3) $R_A^* \subseteq R_C^*$ 。

引理 2 设 (U,A,I,C,J) 为粒协调决策形式背景, $B \subseteq A$, 则以下条件等价:

- 1) B 为 (U,A,I,C,J) 的粒协调集;
- 2) $\forall x \in U, [x]_{R_B^*} \subseteq [x]_{R_C^*}$;
- 3) $R_B^* \subseteq R_C^*$ 。

例 1 表 1^[2] 是一个决策形式背景 (U,A,I,C,J) , 其中 $U = \{1,2,3,4,5\}, A = \{a,b,c,d,e\}, C = \{f,g,h\}$ 。

表 1 决策形式背景 (U,A,I,C,J)

Tab. 1 Formal decision context (U,A,I,C,J)

U	a	b	c	d	e	f	g	h
1	0	1	0	1	0	1	0	0
2	1	0	1	0	1	0	1	1
3	1	1	0	0	1	0	1	0
4	0	1	1	1	0	1	0	0
5	1	0	0	0	1	0	1	0

由于 $1^{*A} = \{b,d\}, 2^{*A} = \{a,c,e\}, 3^{*A} = \{a,b,e\}, 4^{*A} = \{b,c,d\}, 5^{*A} = \{a,e\}$, 所以条件拟序类为 $[1]_{R_A^*} = \{1,4\}, [2]_{R_A^*} = \{2\}, [3]_{R_A^*} = \{3\}, [4]_{R_A^*} = \{4\}, [5]_{R_A^*} = \{2,3,5\}$ 。

又因为 $1^{*C} = \{f\}, 2^{*C} = \{g,h\}, 3^{*C} = \{g\}, 4^{*C} = \{f\}, 5^{*C} = \{g\}$, 所以决策拟序类为 $[1]_{R_C^*} = \{1,4\}, [2]_{R_C^*} = \{2\}, [3]_{R_C^*} = \{2,3,5\}, [4]_{R_C^*} = \{1,4\}, [5]_{R_C^*} = \{2,3,5\}$ 。

显然, $\forall x \in U$, 有 $[x]_{R_A^*} \subseteq [x]_{R_C^*}$, 即 $R_A^* \subseteq R_C^*$ 。因此, 表 1 的决策形式背景 (U,A,I,C,J) 是协调的。

令 $B = \{a,c,d\}$, 则 $1^{*B} = \{d\}, 2^{*B} = \{a,c\}, 3^{*B} = \{a\}, 4^{*B} = \{c,d\}, 5^{*B} = \{a\}$, 即 $[1]_{R_B^*} = \{1,4\}, [2]_{R_B^*} = \{2\}, [3]_{R_B^*} = \{2,3,5\}, [4]_{R_B^*} = \{4\}, [5]_{R_B^*} = \{2,3,5\}$ 。显然, $\forall x \in U$, 有 $[x]_{R_B^*} \subseteq [x]_{R_C^*}$, 因此, B 为 (U,A,I,C,J) 的一个粒协调集。

2.1 不协调决策形式背景的分布约简定义

设 (U,A,I,C,J) 为不协调决策形式背景。记 $R_C^* = \{(x,y) \in U \times U \mid x^{*C} \subseteq y^{*C}\}, U/R_C^* =$

$\{D_1, D_2, \dots, D_t\}$, 其中 $D_j, j = 1, 2, \dots, t$ 为决策拟序类. $B \subseteq A$, 定义函数 $\mu_B: U \rightarrow [0, 1]^t$ 为

$$\mu_B(x) = \left(\frac{|D_1 \cap [x]_{R_B^*}|}{|U|}, \frac{|D_2 \cap [x]_{R_B^*}|}{|U|}, \dots, \frac{|D_t \cap [x]_{R_B^*}|}{|U|} \right),$$

$\forall x \in U$, 称 $\mu_B(x)$ 为对象 x 基于条件属性子集 B 的分布函数。

$\forall x \in U$, 定义 $\gamma_B(x) = \max_{j \leq t} \frac{|D_j \cap [x]_{R_B^*}|}{|U|}$, 称 $\gamma_B(x)$ 为对象 x 基于条件属性子集 B 的最大分布函数。

例 2 设决策形式背景 (U, A, I, C, J) 如表 2 所示, 其中 $U = \{1, 2, 3, 4, 5\}$, $A = \{a, b, c, d, e\}$, $C = \{f, g, h\}$ 。

表 2 决策形式背景 (U, A, I, C, J)

Tab. 2 Formal decision context (U, A, I, C, J)

U	a	b	c	d	e	f	g	h
1	0	1	0	1	0	0	0	1
2	1	0	1	0	1	0	1	1
3	1	1	0	0	1	0	1	0
4	0	1	1	1	0	1	0	0
5	1	0	0	0	1	0	1	0

条件拟序类计算如下:

$$[1]_{R_A^*} = \{1, 4\}, [2]_{R_A^*} = \{2\}, [3]_{R_A^*} = \{3\}, [4]_{R_A^*} = \{4\}, [5]_{R_A^*} = \{2, 3, 5\}.$$

决策拟序类计算如下:

$$D_1 = [1]_{R_C^*} = \{1, 2\}, D_2 = [2]_{R_C^*} = \{2\}, D_3 = [3]_{R_C^*} = [5]_{R_C^*} = \{2, 3, 5\}, D_4 = [4]_{R_C^*} = \{4\}.$$

显然有 $[1]_{R_A^*} \not\subseteq [1]_{R_C^*}$, 即 $R_A^* \not\subseteq R_C^*$, 所以表 2 的决策形式背景 (U, A, I, C, J) 是不协调的。分布函数和最大分布函数计算如下:

$$\mu_A(1) = \{1/5, 0, 0, 1/5\}, \mu_A(2) = \{1/5, 1/5, 0, 0\}, \mu_A(3) = \{0, 0, 1/5, 0\}, \mu_A(4) = \{0, 0, 0, 1/5\}, \mu_A(5) = \{1/5, 1/5, 3/5, 0\};$$

$$\gamma_A(1) = 1/5, \gamma_A(2) = 1/5, \gamma_A(3) = 1/5, \gamma_A(4) = 1/5, \gamma_A(5) = 3/5.$$

设 (U, A, I, C, J) 为不协调决策形式背景, $B \subseteq A, x, y \in U, \forall D_i \in U/R_C^*$, 若 $\frac{|D_i \cap [x]_{R_B^*}|}{|U|} =$

$\frac{|D_i \cap [y]_{R_B^*}|}{|U|}$, 则称 $\mu_B(x)$ 等于 $\mu_B(y)$, 记为

$$\mu_B(x) = \mu_B(y); \forall D_i \in U/R_C^*, \text{ 若 } \frac{|D_i \cap [x]_{R_B^*}|}{|U|}$$

$$\leq \frac{|D_i \cap [y]_{R_B^*}|}{|U|}, \text{ 则称 } \mu_B(x) \text{ 小于等于 } \mu_B(y), \text{ 记为 } \mu_B(x) \leq \mu_B(y);$$

若 $\mu_B(x) \leq \mu_B(y)$ 且 $\exists D_j \in U/R_C^*$ 使得 $\frac{|D_j \cap [x]_{R_B^*}|}{|U|} < \frac{|D_j \cap [y]_{R_B^*}|}{|U|}$, 则称

$\mu_B(x)$ 小于 $\mu_B(y)$, 记为 $\mu_B(x) < \mu_B(y)$; 若 $\exists D_j \in U/R_C^*$ 使得 $\frac{|D_j \cap [x]_{R_B^*}|}{|U|} > \frac{|D_j \cap [y]_{R_B^*}|}{|U|}$, 则

记为 $\mu_B(x) \sim \mu_B(y)$ 。

引理 3 设 (U, A, I, C, J) 为不协调决策形式背景。若 $B_1 \subseteq B_2 \subseteq A$, 则

$$1) \forall x \in U, \mu_{B_2}(x) \leq \mu_{B_1}(x);$$

$$2) \forall x \in U, \gamma_{B_2}(x) \leq \gamma_{B_1}(x).$$

证明 由性质 2 中条件 2) 知, 若 $B_1 \subseteq B_2 \subseteq A$, 则 $R_{B_2}^* \subseteq R_{B_1}^*$, 于是 $\forall x \in U, \mu_{B_2}(x) \leq \mu_{B_1}(x), \gamma_{B_2}(x) \leq \gamma_{B_1}(x)$ 。

定义 7 设 (U, A, I, C, J) 为不协调决策形式背景。 $\forall B \subseteq A$, 有

1) $\forall x \in U$, 若 $\mu_B(x) = \mu_A(x)$, 则称 B 为 (U, A, I, C, J) 的分布协调集; $\forall x \in U$, 若 $\gamma_B(x) = \gamma_A(x)$, 则称 B 为 (U, A, I, C, J) 的最大分布协调集。

2) 设 $B' \subset B$, 若 B 为分布协调集, 而 B' 不是分布协调集, 则称 B 为 (U, A, I, C, J) 的分布约简; 同样地, 若 B 为最大分布协调集, 而 B' 不是最大分布协调集, 则称 B 为 (U, A, I, C, J) 的最大分布约简。

结合分布函数、最大分布函数的定义以及定义 7 可知, 若 $B \subseteq A$ 为 (U, A, I, C, J) 的分布协调集, 则 B 为 (U, A, I, C, J) 的最大分布协调集。

定理 1 设 (U, A, I, C, J) 为不协调决策形式背景, $\forall B \subseteq A$ 。

1) B 是分布协调集, 则下列结论成立: $\forall x, y \in U$, 若 $[y]_{R_B^*} \subseteq [x]_{R_B^*}$, 则有 $\mu_A(y) \leq \mu_A(x)$ 。

2) B 是最大分布协调集, 则下列结论成立: $\forall x, y \in U$, 若 $[y]_{R_B^*} \subseteq [x]_{R_B^*}$, 则有 $\gamma_A(y) \leq \gamma_A(x)$ 。

证明 1) 假设 B 是分布协调集, 则 $\mu_A(x) = \mu_B(x), \forall x \in U. \forall x, y \in U$, 若 $[y]_{R_B^*} \subseteq [x]_{R_B^*}$,

则 $\mu_B(y) \leq \mu_B(x)$ 。于是 $\mu_A(y) = \mu_B(y) \leq \mu_B(x) = \mu_A(x)$ 。

2) 类似 1) 的证明。

推论 1 设 (U,A,I,C,J) 为不协调决策形式背景, $B \subseteq A$ 。

1) B 是分布协调集,则下列结论成立: $\forall x,y \in U$,若 $\mu_A(y) \leq \mu_A(x)$,则有 $[y]_{R_B^*} \not\subset [x]_{R_B^*}$ 。

2) B 是最大分布协调集,则下列结论成立: $\forall x,y \in U$,若 $\gamma_A(y) > \gamma_A(x)$,则有 $[y]_{R_B^*} \not\subset [x]_{R_B^*}$ 。

2.2 不协调决策形式背景分布约简方法

定义 8 设 (U,A,I,C,J) 为不协调决策形式背景。 $\forall B \subseteq A$,记

$D_B^\mu = \{(x,y) \in U \times U \mid \mu_B(y) \leq \mu_B(x)\},$

$D_B^\gamma = \{(x,y) \in U \times U \mid \gamma_B(y) > \gamma_B(x)\}。$

$\forall x,y \in U$,定义

$D_B^\mu(x,y) = \begin{cases} \{a \in B \mid (x,y) \notin R_a^*\}, & (x,y) \in D_B^\mu; \\ \emptyset, & \text{其他。} \end{cases}$

$D_B^\gamma(x,y) = \begin{cases} \{a \in B \mid (x,y) \notin R_a^*\}, & (x,y) \in D_B^\gamma; \\ \emptyset, & \text{其他。} \end{cases}$

分别称 $D_B^\mu(x,y)$ 和 $D_B^\gamma(x,y)$ 为对象 x 和 y 关于条件属性子集 B 和决策属性集 C 的分布辨识属性集和最大分布辨识属性集。记 $\Delta_B^\mu = (D_B^\mu(x,y) \mid x,y \in U)$, $\Delta_B^\gamma = (D_B^\gamma(x,y) \mid x,y \in U)$,称 Δ_B^μ 为 (U,A,I,C,J) 基于条件属性子集 B 的分布辨识矩阵,并称 Δ_B^γ 为最大分布辨识矩阵。

引理 4 设 (U,A,I,C,J) 为不协调决策形式背景, $\forall x,y \in U, \forall B \subseteq A$ 。

1) $D_B^\mu(x,y) \neq \emptyset$ 当且仅当 $\mu_B(y) \leq \mu_B(x)$;

2) $D_B^\gamma(x,y) \neq \emptyset$ 当且仅当 $\gamma_B(y) > \gamma_B(x)$ 。

证明 1) $\Rightarrow$ 根据分布辨识属性集的定义可得, $\forall x,y \in U$,若 $D_B^\mu(x,y) \neq \emptyset$,则 $\mu_B(y) \leq \mu_B(x)$ 。

$\Leftarrow \forall x,y \in U$,若 $\mu_B(y) \leq \mu_B(x)$,则 $\exists D_j \in U/R_C^*$ 使得 $\frac{|D_j \cap [y]_{R_B^*}|}{|U|} > \frac{|D_j \cap [x]_{R_B^*}|}{|U|}$,即 $|D_j \cap [y]_{R_B^*}| > |D_j \cap [x]_{R_B^*}|$,于是 $[y]_{R_B^*} \not\subset [x]_{R_B^*}$,从而 $\exists a \in B$ 使得 $(x,y) \notin R_a^*$,即 $a \in D_B^\mu(x,y) \neq \emptyset$ 。

2) 类似 1) 的证明。

定理 2 设 (U,A,I,C,J) 为不协调决策形式背景, $\forall B \subseteq A, B \neq \emptyset$,有

1) B 是分布协调集 $\Leftrightarrow$ 若 $D_A^\mu(x,y) \neq \emptyset$,则 $B \cap D_A^\mu(x,y) \neq \emptyset$;

2) B 是最大分布协调集 $\Leftrightarrow$ 若 $D_A^\gamma(x,y) \neq \emptyset$,则 $B \cap D_A^\gamma(x,y) \neq \emptyset$ 。

证明 1) $\Rightarrow$ 由引理 4 条件 1) 知, $D_A^\mu(x,y) \neq \emptyset$ 等价于 $\mu_A(y) \leq \mu_A(x)$ 。又因为 B 是分布协调集,再结合推论 1 条件 1) 可得, $[y]_{R_B^*} \not\subset [x]_{R_B^*}$,即 $\exists a \in B$,使得 $(x,y) \notin R_a^*$,从而 $a \in B \cap D_A^\mu(x,y) \neq \emptyset$ 。

$\Leftarrow$ 由引理 4 条件 1) 知, $D_A^\mu(x,y) \neq \emptyset$ 等价于 $\mu_A(y) \leq \mu_A(x)$ 。又因为若 $D_A^\mu(x,y) \neq \emptyset$,则 $B \cap D_A^\mu(x,y) \neq \emptyset$ 。于是,若 $\mu_A(y) \leq \mu_A(x)$,则 $B \cap D_A^\mu(x,y) \neq \emptyset$ 。因此, $\exists a \in B$,使得 $(x,y) \notin R_a^*$,即 $[y]_{R_B^*} \not\subset [x]_{R_B^*}$ 。从而由推论 1 条件 1) 可知, B 是分布协调集。

2) 类似于 1) 的证明。

例 3(续例 2) 根据分布辨识矩阵的定义可计算表 2 的分布辨识矩阵,如表 3 所示。再由定理 2 条件 1) 可得, (U,A,I,C,J) 的两个分布约简为 $B_1 = \{a,b,c,d\}$ 和 $B_2 = \{b,c,d,e\}$ 。

表 3 分布辨识矩阵 Δ_A^μ

Tab. 3 Distribution matrix Δ_A^μ

	1	2	3	4	5
1	$\emptyset$	$\{b,d\}$	$\{d\}$	$\emptyset$	$\{b,d\}$
2	$\{a,c,e\}$	$\emptyset$	$\emptyset$	$\{a,e\}$	$\{c\}$
3	$\{a,e\}$	$\{b\}$	$\emptyset$	$\{a,e\}$	$\{b\}$
4	$\{c\}$	$\{b,d\}$	$\{c,d\}$	$\emptyset$	$\{b,c,d\}$
5	$\{a,e\}$	$\emptyset$	$\emptyset$	$\{a,e\}$	$\emptyset$

根据定义 8 计算表 2 的最大分布辨识矩阵,如表 4 所示。

表 4 最大分布辨识矩阵 Δ_A^γ

Tab. 4 Maximum distribution matrix Δ_A^γ

	1	2	3	4	5
1	$\emptyset$	$\emptyset$	$\emptyset$	$\emptyset$	$\{b,d\}$
2	$\emptyset$	$\emptyset$	$\emptyset$	$\emptyset$	$\{c\}$
3	$\emptyset$	$\emptyset$	$\emptyset$	$\emptyset$	$\{b\}$
4	$\emptyset$	$\emptyset$	$\emptyset$	$\emptyset$	$\{b,c,d\}$
5	$\emptyset$	$\emptyset$	$\emptyset$	$\emptyset$	$\emptyset$

根据定理 2 条件 2) 和表 4 可知, $B_3 = \{b, c\}$ 是 (U, A, I, C, J) 的最大分布约简。

定义 9 设 (U, A, I, C, J) 为不协调决策形式背景。记

$$M_A^\mu = \bigwedge \{ \bigvee \{ a \in A \mid a \in D_A^\mu(x, y) \} \mid x, y \in U \},$$
$$M_A^\gamma = \bigwedge \{ \bigvee \{ a \in A \mid a \in D_A^\gamma(x, y) \} \mid x, y \in U \}。$$

称 M_A^μ 为 (U, A, I, C, J) 的分布辨识公式, M_A^γ 为 (U, A, I, C, J) 的最大分布辨识公式。

定理 3 设 (U, A, I, C, J) 为不协调决策形式背景, 并记 M_A^μ, M_A^γ 的极小析取范式为 $M_{\min}^\mu = \bigvee_{k=1}^p$

$(\bigwedge_{i=1}^{q_k} a_i)$ 和 $M_{\min}^\gamma = \bigvee_{l=1}^m (\bigwedge_{j=1}^{n_l} a_j)$ 。记 $B_k^\mu = \{a_i \mid i = 1,$

$2, \cdots, q_k\}, B_l^\gamma = \{a_j \mid j = 1, 2, \cdots, n_l\}$, 则

- 1) $\{B_k^\mu \mid k = 1, 2, \cdots, p\}$ 包含所有分布约简;
- 2) $\{B_l^\gamma \mid l = 1, 2, \cdots, m\}$ 包含所有最大分布约简。

证明 1) $\forall (x, y) \in D_A^\mu$, 显然有 $B_k^\mu \cap D_A^\mu(x, y) \neq \emptyset$ 。结合由定理 2 条件 1) 可知, B_k^μ 是 (U, A, I, C, J) 的分布协调集。记 $B_k^{\mu'} = B_k^\mu \setminus \{a\}, a \in B_k^\mu$, 由于 $M_{\min}^\mu = \bigvee_{k=1}^p B_k^\mu$, 则 $\exists (x_0, y_0) \in D_A^\mu$ 使得 $B_k^{\mu'} \cap D_A^\mu(x_0, y_0) = \emptyset$, 因此 $B_k^{\mu'}$ 不是分布协调集, 所以 B_k^μ 是分布约简。又因为 M_A^μ 包含了所有的 $D_A^\mu(x, y)$, 于是不存 $\{B_k^\mu \mid k = 1, 2, \cdots, p\}$ 之外的分布约简。

2) 类似 1) 的证明。

例 4 根据表 3 分布辨识矩阵 Δ_A^μ 可得极小析取范式为 $M_{\min}^\mu = (a \wedge b \wedge c \wedge d) \vee (b \wedge c \wedge d \wedge e)$, 所以表 2 的决策形式背景的所有分布约简为 $B_1 = \{a, b, c, d\}$ 和 $B_2 = \{b, c, d, e\}$ 。

根据表 4 最大分布辨识矩阵 Δ_A^γ 可得极小析取范式为 $M_{\min}^\gamma = b \wedge c$, 所以表 2 的决策形式背景的最大分布约简只有一个, $B_3 = \{b, c\}$ 。

3 结语

由对象概念可以生成所有的形式概念, 因此, 对象概念构成的集合称为概念格的信息粒, 它包含了形式背景的本质信息。基于信息粒可以对形式背景以及带有决策的协调形式背景进行属性约简, 从而达到简化形式背景的目的。同样地, 利用

信息粒也可以对不协调决策形式背景进行简化。本文考虑不协调的决策形式背景的约简问题。通过对对象集上的拟序关系引入了对象粒下不协调决策形式背景的(最大)分布约简, 提出了两种计算约简的方法以及(最大)分布协调集的判定定理。

参考文献

[1] ZADEH L A. Toward a theory of fuzzy information granulation and its centrality in human reasoning and fuzzy logic [J]. Fuzzy Sets and Systems, 1997, 90 (2): 111-127.

[2] WU W Z, LEUNG Y, MI J S. Granular computing and knowledge reduction in formal contexts[J]. IEEE Transactions on Knowledge and Data Engineering, 2009, 21(10): 1461-1474.

[3] GANTER B, WILLE R. Formal concept analysis: Mathematical foundations[M]. New York: Springer-Verlag, 1999.

[4] 张文修, 仇国芳. 基于粗糙集的不确定决策[M]. 北京: 清华大学出版社, 2005.

[5] ZHANG W X, WEI L, QI J J. Attribute reduction theory and approach to concept lattice[J]. Science in China Series F: Information Sciences, 2005, 48(6): 713-726.

[6] WANG X, MA J M. A novel approach to attribute reduction in concept lattices[J]. Lecture Notes in Artificial Intelligence, 2006, 4062: 522-529.

[7] WANG X. Approaches to attribute reduction in concept lattices based on rough set theory[J]. International Journal of Hybrid Information Technology, 2012, 5(2): 67-80.

[8] WAN Q, WEI L. Attribute reduction based on property pictorial diagram [J]. Scientific World Journal, 2014: 109706.

[9] DIAS S M, VIEIRA N J. Concept lattices reduction: Definition, analysis and classification[J]. Expert Systems with Applications, 2015, 42(20): 7084-7097.

[10] 贺明利, 魏玲. 基于优势关系的序形式背景约简[J]. 计算机科学, 2015, 42(6): 46-49.

HE M L, WEI L. Reduction of ordered formal context based on dominance relation[J]. Computer Science, 2015, 42(6): 46-49.

[11] LI T J, LI M Z, GAO Y. Attribute reduction of concept lattice based on irreducible elements[J]. International Journal of Wavelets, Multiresolution and Information Processing, 2013, 11(6): 1350046.

[12] 马文胜, 侯锡林. 形式概念分析中的同效关系与概念约简[J]. 计算机科学, 2023, 50(4): 63-76.

- MA W S, HOU X L. Same effect relation and concept reduction in formal concept analysis [J]. Computer Science, 2023, 50(4): 63-76.
- [13] AKRAM M, NAWAZ H S, DEVECI M. Attribute reduction and information granulation in Pythagorean fuzzy formal contexts[J]. Expert Systems with Applications, 2023, 222: 119794.
- [14] 李同军, 徐珍珍, 吴明瑞, 等. 经典-经典变精度概念格的一种属性约简[J]. 西北大学学报(自然科学版), 2022, 52(5): 765-773.
- LI T J, XU Z Z, WU M R, et al. An attribute reduction of crisp-crisp variable threshold concept lattices [J]. Journal of Northwest University (Natural Science Edition), 2022, 52(5): 765-773.
- [15] LI J H, MEI C L, LV Y J. Knowledge reduction in decision formal contexts [J]. Knowledge-Based Systems, 2011, 24(5): 709-715.
- [16] LI L J, MI J S, XIE B. Attribute reduction based on maximal rules in decision formal context[J]. International Journal of Computational Intelligence Systems, 2014, 7(6): 1044-1053.
- [17] WEI L, QI J J, ZHANG W X. Attribute reduction theory of concept lattice based on decision formal contexts[J]. Science in China Series F: Information Sciences, 2008, 51(7): 910-923.
- [18] PEI D, MI J S. Attribute reduction in decision formal context based on homomorphism [J]. International Journal of Machine Learning and Cybernetics, 2011, 2: 289-293.
- [19] LI J H, MEI C L, LV Y J. Knowledge reduction in formal decision contexts based on an order-preserving mapping [J]. International Journal of General Systems, 2012, 41(2): 143-161.
- [20] 李进金, 张燕兰, 吴伟志, 等. 形式背景与协调决策形式背景属性约简与概念格生成[J]. 计算机学报, 2014, 37(8): 1768-1774.
- LI J J, ZHANG Y L, WU W Z, et al. Attribute reduction for formal context and consistent decision formal context and concept lattice generation [J]. Chinese Journal of Computers, 2014, 37(8): 1768-1774.
- [21] WANG X, WU W Z. Approximate reduction in inconsistent formal decision contexts [C] // 2012 IEEE International Conference on Granular Computing. Hangzhou, China: IEEE, 2012: 1-6.
- [22] WANG X. Construction of a unified model for formal contexts and formal decision contexts [J]. International Journal of Database Theory and Application, 2014, 7(2): 81-90.
- [23] LI J Y, WANG X, WU W Z, et al. Attribute reduction in inconsistent formal decision contexts based on congruence relations [J]. International Journal of Machine Learning and Cybernetics, 2017, 8: 81-94.
- [24] 李金海, 周新然. 多粒度决策形式背景的属性约简 [J]. 模式识别与人工智能, 2022, 35(5): 387-400.
- LI J H, ZHOU X R. Attribute reduction in multi-granularity formal decision contexts [J]. Pattern Recognition and Artificial Intelligence, 2022, 35(5): 387-400.
- [25] 魏玲, 王振, 钱婷, 等. 多源决策形式背景的属性约简 [J]. 陕西师范大学学报(自然科学版), 2019, 47(5): 57-63.
- WEI L, WANG Z, QIAN T, et al. The attribute reduction of multi-source formal decision contexts [J]. Journal of Shaanxi Normal University (Natural Science Edition), 2019, 47(5): 57-63.

(编辑 张欢)