

基于三支概念格合并的决策背景规则提取

王一宾^{1,2}, 杨思春³

(1. 安庆师范大学计算机与信息学院 安徽 安庆 246011; 2. 安徽省智能感知与计算重点实验室 安徽 安庆 246011;

3. 安徽工业大学计算机科学与技术学院 安徽 马鞍山 243032)

【摘要】作为当前形式概念分析领域的研究热点,利用三支概念格可以实现更为有效的决策分析。该文在现有基于属性导出三支概念格的规则提取基础上,研究了基于对象导出三支概念格的规则提取,并与经典概念格及属性导出三支概念格下的规则提取进行了比较。然后通过对象导出三支概念格和属性导出三支概念格的合并,定义了对象/属性导出合并三支概念格,并提出了相应的规则提取算法。理论分析和实例结果表明,对象导出三支概念格和属性导出三支概念格的合并进一步改善了生成规则的质量。

关键词 形式概念分析; 决策形式背景; 规则提取; 三支概念格

中图分类号 TP18 **文献标志码** A **doi**:10.3969/j.issn.1001-0548.2018.06.018

Rules Extraction in Formal Decision Contexts Based on the Merging of Three-Way Concept Lattices

WANG Yi-bin^{1,2} and YANG Si-chun³

(1. School of Computer Science and Information, Anqing Normal University Anqing Anhui 246011;

2. The University Key Laboratory of Intelligent Perception and Computing of Anhui Province Anqing Anhui 246011;

3. School of Computer Science and Technology, Anhui University of Technology Maanshan Anhui 243032)

Abstract As one of the research focuses in the field of formal concept analysis, three-way concept lattice can make more effective decision analysis. Firstly, the rule extraction based on object-induced three-way concept are investigated and compared compared with those of the classical concept lattice and attribute-induced three-way concept lattice. Then, by merging the object-induced three-way concept lattice and the attribute-induced three-way concept lattice, a new three-way concept lattice, i.e. object-attribute-induced three-way concept lattice, and its rule extraction[0] algorithm are proposed. Theoretical analyses and results show that the merging of object-induced three-way concept lattice and attribute-induced three-way concept lattice can improve the quality of rules.

Key words formal concept analysis(FCA); formal decision context; rules extraction; three-way concept lattices

作为形式概念分析(FCA)的核心数据结构,概念格刻画了概念之间的泛化和例化关系,在数据挖掘、机器学习和信息检索等领域取得了广泛的应用^[1]。近年来,概念格研究取得了很多重要的成果,特别是文献[2]提出粗糙集框架下的概念格理论以来,一些学者借鉴粗糙集理论深入研究概念格的属性约简和规则提取等问题。文献[3-6]研究了粗糙集与信息系统的属性约简和规则提取问题,文献[7-10]探讨了粗糙集与概念格的属性约简理论以及基于概念格的决策形式背景规则获取问题。文献[11-12]针对不完备决策形式背景,研究了相应的近似概念构造、规

则提取以及知识约简。

作为二支决策的推广,三支决策(3-way decisions)是文献[13]提出的一种新的决策理论框架,它将决策问题分为三类:接收、拒绝和不承诺。传统的形式概念分析仅支持二支决策,实际构造形式概念时,仅仅考虑“某些对象共同具有哪些属性”或“某些属性为哪些对象所共同具有”,而没有同时考虑“某些对象共同不具有哪些属性”或“某些属性为哪些对象所共同不具有”。为了更好地支持实际应用中的复杂决策问题,文献[14-15]引入三支决策思想,将具有二支决策的形式概念分析推广为支持

收稿日期: 2017-08-06; 修回日期: 2017-11-15

基金项目: 安徽省自然科学基金(1808085MF178); 安徽省高校省级自然科学研究重点项目(KJ2013A177, KJ2016A098)

作者简介: 王一宾(1970-),男,教授,主要从事数据挖掘和概念格方面的研究。

三支决策的三支概念分析(3-way concept analysis), 提出了两种具体的三支概念格(对象导出三支概念格和属性导出三支概念格), 并研究了它们与经典概念格之间的关系。

自文献[14]提出三支概念格以来, 就引起了有关学者的关注, 并取得了一些初步的研究成果。文献[16]对概念格与三支决策相结合的研究历程、研究内容以及研究展望进行了系统的综述。目前, 这些工作主要集中在基于三支概念格的属性约简方面^[17], 在基于三支概念格的规则提取方面则偏少。文献[18]在利用三支概念格进行规则提取方面做了有益的尝试, 给出了基于属性导出三支概念格的决策背景规则提取方法, 并通过与经典概念格下的生成规则进行比较, 表明了基于属性导出三支概念格的决策背景规则提取方法的有效性和优越性。

本文在文献[18]的基础上进一步研究基于对象导出三支概念格的规则提取, 并与经典概念格及属性导出三支概念格下的生成规则进行比较。同时通过对对象导出三支概念格和属性导出三支概念格的合并, 进一步改善三支概念格下的生成规则质量。

1 基本概念和性质

先给出形式背景及其算子的定义。

定义 1^[1] (G, M, I) 为形式背景, 其中 $G = \{x_1, x_2, \dots, x_n\}$, 每个 $x_i (i \leq n)$ 称为对象; $M = \{a_1, a_2, \dots, a_m\}$, 每个 $a_j (j \leq m)$ 称为属性; 对于 $x \in G$ 与 $a \in M$, 如果 $(x, a) \in I$, 则称对象 x 拥有属性 a 。

定义 2^[1] (G, M, I) 为形式背景, $X \subseteq G, A \subseteq M$ 。一对算子定义如下:

$$*: P(G) \rightarrow P(M), X^* = \{a \in M \mid \forall x \in X, (x, a) \in I\};$$

$$*: P(M) \rightarrow P(G), A^* = \{x \in G \mid \forall a \in M, (x, a) \in I\}.$$

基于形式背景的算子, 可定义形式概念及其概念格。

定义 3^[1] (G, M, I) 为形式背景, 若 $X^* = A$ 且 $A^* = X$, 则称 (X, A) 是一个概念。 X 称为该概念的外延, A 称为该概念的内涵。

对于任意概念 $(X_1, A_1), (X_2, A_2)$, 定义偏序关系为 $(X_1, A_1) \leq (X_2, A_2) \Leftrightarrow X_1 \subseteq X_2 \Leftrightarrow A_2 \subseteq A_1$, 则形式背景 (G, M, I) 的所有概念在该偏序关系下是完备格, 称为概念格, 记为 $L(G, M, I)$ 。

以下给出概念格之间的细于关系以及决策形式背景及其协调性的定义。

定义 4^[8] $L(G, M, I)$ 与 $L(G, N, J)$ 是两个概念格, 若对 $\forall (Y, B) \in L(G, N, J)$, 存在 $(X, A) \in L(G, M, I)$, 使得

$X=Y$, 则称 $L(G, M, I)$ 细于 $L(G, N, J)$, 记作 $L(G, M, I) \leq L(G, N, J)$ 。

定义 5^[8] (G, M, I) 与 (G, N, J) 是两个形式背景, 则称 (G, M, I, N, J) 是决策形式背景。若 $L(G, M, I) \leq L(G, N, J)$, 则称该决策形式背景是协调的。

为了在形式概念分析中引入三支决策的思想, 给出形式背景负算子的定义:

定义 6^[14] (G, M, I) 是形式背景, 对于 $X \subseteq G, A \subseteq M$, 一对负算子定义如下:

$$*^-: P(G) \rightarrow P(M): X^{*-} = \{a \in M \mid \forall x \in X, \neg(xIa)\} = \{a \in M \mid \forall x \in X, xI^c a\}, \text{ 其中 } I^c = (G \times M) - I.$$

$$*^-: P(M) \rightarrow P(G): A^{*-} = \{x \in G \mid \forall a \in A, \neg(xIa)\} = \{x \in G \mid \forall a \in A, xI^c a\}.$$

在形式背景算子及负算子的基础上, 可定义形式背景的三支算子:

定义 7^[14] (G, M, I) 是形式背景, $X \subseteq G, A \subseteq M$, 一对三支算子定义如下:

$$\leq: P(G) \rightarrow DP(M): X^{\leq} = (X^*, X^{*-})$$

$$\leq: P(M) \rightarrow DP(G): A^{\leq} = (A^*, A^{*-})$$

进一步地, 可定义上述三支算子的逆运算:

$$\geq: DP(M) \rightarrow P(G): (A, B)^{\geq} = \{x \in G \mid x \in A^* \text{ 且 } x \in B^{*-}\} = A^* \cap B^{*-}$$

$$\geq: DP(G) \rightarrow P(M): (X, Y)^{\geq} = \{a \in M \mid a \in X^* \text{ 且 } a \in Y^{*-}\} = X^* \cap Y^{*-}$$

基于形式背景的三支算子, 可定义对象导出三支概念和对象导出三支概念格。

定义 8^[14] 设 (G, M, I) 是形式背景, $(X, (A, B))$ 称为对象导出三支概念, 简称 OE-概念, 这里 $X^{\leq} = (A, B)$ 且 $(A, B)^{\geq} = X$, 其中 $X \subseteq G, A, B \subseteq M$ 。 X 和 (A, B) 分别称为 OE-概念的外延和内涵。

对于 $(X, (A, B)), (Y, (C, D))$, 定义偏序关系 $(X, (A, B)) \leq (Y, (C, D)) \Leftrightarrow X \subseteq Y \Leftrightarrow (C, D) \subseteq (A, B)$, 形式背景 (G, M, I) 的所有 OE-概念在该偏序关系下是完备格, 称为对象导出三支概念格, 记为 $OE(L(G, M, I))$ 。

类似地, 可定义属性导出三支概念及属性导出三支概念格。

定义 9^[14] 设 (G, M, I) 是形式背景, 则称 $((X, Y), A)$ 为属性导出三支概念, 简称 AE-概念, 这里 $(X, Y)^{\geq} = A$ 且 $A^{\leq} = (X, Y)$, 其中 $X, Y \subseteq G, A \subseteq M$ 。 (X, Y) 和 A 分别称为 AE-概念的外延和内涵。

对于 $((X, Y), A), ((Z, W), B)$, 定义偏序关系 $((X, Y), A) \leq ((Z, W), B) \Leftrightarrow (X, Y) \subseteq (Z, W) \Leftrightarrow B \subseteq A$, 形式背景 (G, M, I) 的所有 AE-概念在该偏序关系下是完备格, 称为属性导出三支概念格, 记为 $AE(L(G, M, I))$ 。

文献[18]定义了属性导出三支概念格之间的细于关系, 以及决策背景基于属性导出三支概念格的协调性。

定义 10^[18] 设 $AEL(G,M,I)$ 与 $AEL(G,N,J)$ 是属性导出三支概念格。若对于任意 $((Z,W),B) \in AEL(G,N,J)$, 存在 $((X,Y),A) \in AEL(G,M,I)$, 使得 $X=Z$ 且 $Y=W$, 则称 $AEL(G,M,I)$ 细于 $AEL(G,N,J)$, 记作 $AEL(G,M,I) \leq AEL(G,N,J)$, 并称决策背景 (G,M,I,N,J) 是基于属性导出三支概念格协调的。

文献[18]比较了属性导出三支概念格和经典概念格下的规则提取。

定理 1^[18] 设 (G,M,I,N,J) 是决策形式背景, 且 $AEL(G,M,I) \leq AEL(G,N,J)$ 。R表示经典概念格下的生成规则集合, TR_a 表示属性导出三支概念格下的生成规则集合, 则 $R \subseteq TR_a$ 。

定理 2^[18] 设 (G,M,I,N,J) 是决策形式背景, 且 $AEL(G,M,I) \leq AEL(G,N,J)$, R表示经典概念格下的生成规则集合, TR_a 表示属性导出三支概念格下的生成规则集合, 则对于 $A \rightarrow B \in R$, 总存在 $C \rightarrow B \in TR_a$, 且 $C \subseteq A$ 。

2 基于对象导出三支概念格的决策背景规则提取

文献[18]给出基于属性导出三支概念格的决策背景规则提取方法, 本文在其基础上进一步研究基于对象导出三支概念格的规则提取, 并与经典概念格及属性导出三支概念格下的规则提取进行比较。

2.1 对象导出三支概念格下的生成规则

先定义对象导出三支概念格之间的细于关系, 以及决策背景基于对象导出三支概念格的协调性。

定义 11 设 $OEL(G,M,I)$ 与 $OEL(G,N,J)$ 是对象导出三支概念格, 若对于任意 $(Y,(C,D)) \in OEL(G,N,J)$, 存在 $(X,(A,B)) \in OEL(G,M,I)$, 使得 $X=Y$, 则称 $OEL(G,M,I)$ 细于 $OEL(G,N,J)$, 记作 $OEL(G,M,I) \leq OEL(G,N,J)$, 并称决策背景 (G,M,I,N,J) 是基于对象导出三支概念格协调的。

例1: 表1是一个决策形式背景^[18], $G=\{1,2,3,4\}$, $M=\{a,b,c,d,e,f\}$, $N=\{g,h,k\}$, 其对象导出三支概念格 $OEL(G,M,I)$ 和 $OEL(G,N,J)$ 分别如图1、图2所示。

表1 决策形式背景 (G,M,I,N,J)

| G | a | b | c | d | e | f | g | h | k |
|---|---|---|---|---|---|---|---|---|---|
| 1 | | | x | | | x | | | x |
| 2 | | | | x | x | x | | x | x |
| 3 | x | x | x | | x | x | | x | x |
| 4 | | | | x | x | | | x | |

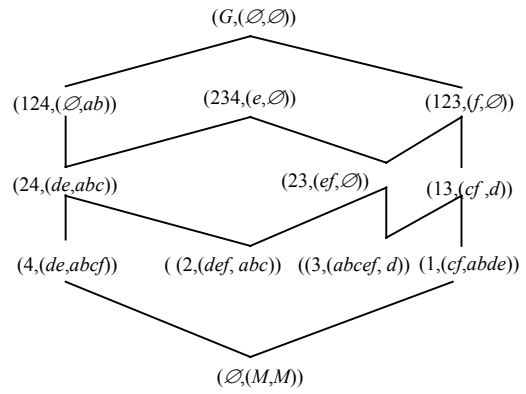


图1 对象导出三支概念格 $OEL(G,M,I)$

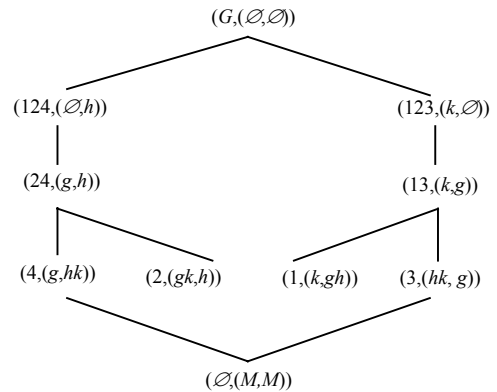


图2 对象导出三支概念格 $OEL(G,N,J)$

由图1、图2可以看出, $OEL(G,M,I) \leq OEL(G,N,J)$, 所以表1所示决策背景是基于对象导出三支概念格协调的。

接着定义对象导出三支概念格下的生成规则。

定义 12 设决策背景 (G,M,I,N,J) 是基于对象导出三支概念格协调的, 若对于 $X, Y \neq \emptyset, G$, 有 $(X,(A,B)) \in OEL(G,M,I)$, $(Y,(C,D)) \in OEL(G,N,J)$, 且 $X=Y$, 则称 $A \rightarrow C$ 和 $B \rightarrow D$ 分别是一个规则。

例2: 根据定义12, 表1所示决策背景在对象导出三支概念格下的生成规则集合 TR_o 为:

$d \rightarrow g; de \rightarrow g; ab \rightarrow h; abc \rightarrow h; f \rightarrow k; def \rightarrow gk; abde \rightarrow gh; abcf \rightarrow hk; abcef \rightarrow hk$

2.2 与经典概念格下的生成规则比较

2.2.1 经典概念格下的生成规则

定义 13 设决策背景 (G,M,I,N,J) 是协调的, 若对于 $X, Y \neq \emptyset, G$, 有 $(X,A) \in L(G,M,I)$, $(Y,B) \in L(G,N,J)$, 且 $X=Y$, 则称 $A \rightarrow B$ 是一个规则。

例3: 表1所示决策背景的概念格 $L(G,M,I)$ 和 $L(G,N,J)$ 分别如图3、图4所示。

根据定义13, 表1所示决策背景在经典概念格下的生成规则集合R为: $de \rightarrow g; f \rightarrow k; def \rightarrow gk; abcef \rightarrow hk$

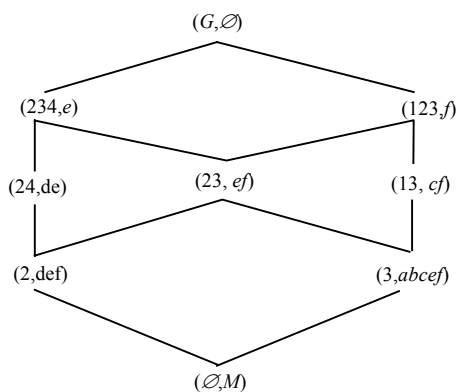


图3 概念格 $L(G, M, I)$

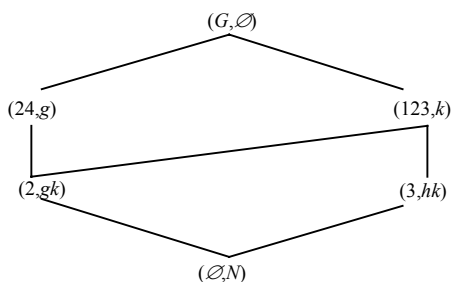


图4 概念格 $L(G, N, J)$

2.2.2 与经典概念格下的生成规则比较

定理 3 设 (G, M, I, N, J) 是决策形式背景, 且 $OEL(G, M, I) \leq OEL(G, N, J)$. R 表示经典概念格下的生成规则集合, TR_o 表示对象导出三支概念格下的生成规则集合, 则 $R \subseteq TR_o$.

证明: 对任意 $A \rightarrow B \in R$, 由定义13, 存在 $(X, A) \in L(G, M, I)$, $(Y, B) \in L(G, N, J)$, 且 $X=Y$. 又因为 $OEL(G, M, I) \leq OEL(G, N, J)$, 由定义11, 对于 $(Y, (B, Y^{*-I})) \in OEL(G, N, J)$, 存在 $(X, (C, X^{*-I})) \in OEL(G, M, I)$. 由于 $C=X^{*I}=A$, 所以 $(X, (C, X^{*-I}))$ 即 $(X, (A, X^{*-I}))$. 由定义12, $A \rightarrow B \in TR_o$, 因此 $R \subseteq TR_o$.

例4: 比较表1所示决策背景分别在对象导出三支概念格及经典概念格下的生成规则集合 TR_o 和 R , 可以看出, TR_o 不仅包含了 R 中每条规则, 还包含了 $ab \rightarrow h$; $abde \rightarrow gh$ 等其他更多规则.

定理 4 设 (G, M, I, N, J) 是决策形式背景, 且 $OEL(G, M, I) \leq OEL(G, N, J)$, R 表示经典概念格下的生成规则集合, TR_o 表示对象导出三支概念格下的生成规则集合, 则对于 $A \rightarrow B \in R$, 总存在 $C \rightarrow B \in TR_o$, 且 $C \subseteq A$.

证明: 对任意 $A \rightarrow B \in R$, 由定义13, 存在 $(X, A) \in L(G, M, I)$, $(Y, B) \in L(G, N, J)$, 且 $X=Y$. 由于存在保序嵌入 $^{[15]} \varphi: L(G, M, I) \rightarrow OEL(G, M, I)$, $\varphi: L(G, N, J) \rightarrow OEL(G, N, J)$, 使 $\varphi((X, A)) = (X, (A, X^{*-I}))$, $\varphi((Y, B)) = (Y, (B, Y^{*-I}))$, 且 $OEL(G, M, I) \leq OEL(G, N, J)$, 因此对于

$(Y, (B, Y^{*-I})) \in OEL(G, N, J)$, 存在 $(Y, (C, Y^{*-I})) \in OEL(G, M, I)$, 使得 $C \rightarrow B \in TR_o$. 又因为 $Y=C^{*I} \cap Y^{*-I}$, 所以 $Y \subseteq C^{*I}$, 进而 $C \subseteq Y^{*I}=X^{*I}=A$, 即 $C \subseteq A$.

例5: 比较表1所示决策背景分别在对象导出三支概念格及经典概念格下的生成规则集合 TR_o 和 R , 可以看出, 对于 R 中的每条规则 $A \rightarrow B$, 在 TR_o 中总存在在相应的规则 $C \rightarrow B$, 且 $C \subseteq A$.

与经典概念格下的生成规则比较, 对象导出三支概念格下的生成规则更为丰富也更为精炼. 但是, 由于对象导出三支概念格比经典概念格的概念数量更多, 结构也更为复杂, 因此构建对象导出三支概念格比构建经典概念格要付出更多的时间开销.

2.3 与属性导出三支概念格下的生成规则比较

2.3.1 属性导出三支概念格下的生成规则

定义 14 设决策背景 (G, M, I, N, J) 是基于属性导出三支概念格协调的, 若 $(X, Y), (Z, W) \neq (\emptyset, \emptyset), (G, G)$, 有 $((X, Y), A) \in AEL(G, M, I)$, $((Z, W), B) \in AEL(G, N, J)$, 且 $X=Z, Y \subseteq W$, 则称 $A \rightarrow B$ 是一个规则.

例6: 表1所示决策背景的属性导出三支概念格 $AEL(G, M, I)$ 和 $AEL(G, N, J)$ 如图5、图6所示.

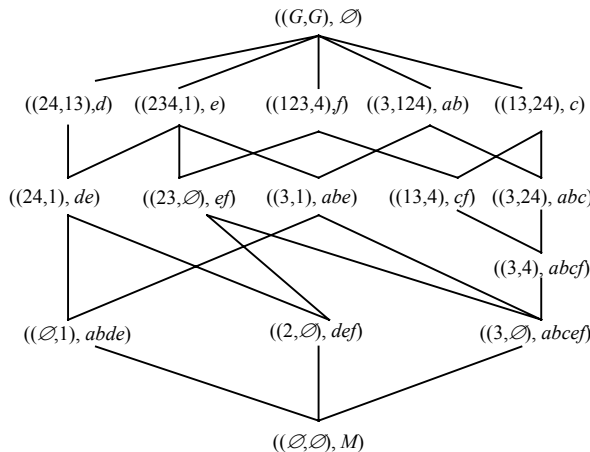


图5 属性导出三支概念格 $AEL(G, M, I)$

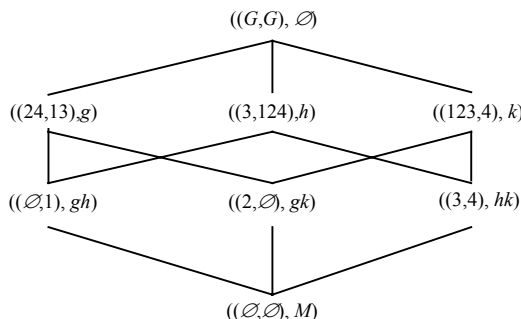


图6 属性导出三支概念格 $AEL(G, N, J)$

根据定义14, 表1所示决策背景在属性导出三支概念格下的生成规则集合 TR_a 为:

$d \rightarrow g$; $de \rightarrow g$; $ab \rightarrow h$; $abe \rightarrow h$; $abc \rightarrow h$; $abcf \rightarrow h$;

$abcef \rightarrow h; f \rightarrow k; def \rightarrow gk; abde \rightarrow gh; abcf \rightarrow hk;$
 $abcef \rightarrow hk$

2.3.2 与属性导出三支概念格下的生成规则比较

先定义对象导出三支概念格与属性导出三支概念格之间的细于关系, 以及决策背景基于对象导出三支概念格和属性导出三支概念格的协调性。

定义 15 设 (G, M, I) 是形式背景, 若对于任意 $(X, (A, B)) \in OEL(G, M, I)$, 存在 $((Y, Z), C) \in AEL(G, M, I)$, 使得 $X=Y$ 或 Z , 则称 $AEL(G, M, I)$ 细于 $OEL(G, M, I)$, 记作 $AEL(G, M, I) \leq OEL(G, M, I)$, 并称形式背景 (G, M, I) 是基于 $AEL(G, M, I)$ 和 $OEL(G, M, I)$ 协调的。

例7: 由定义15, 对于图1所示对象导出三支概念格 $OEL(G, M, I)$, 图5所示属性导出三支概念格 $AEL(G, M, I)$, 有 $AEL(G, M, I) \leq OEL(G, M, I)$, 所以形式背景 (G, M, I) 是基于 $AEL(G, M, I)$ 和 $OEL(G, M, I)$ 协调的; 类似地, 对于图2所示对象导出三支概念格 $OEL(G, N, J)$, 图4所示属性导出三支概念格 $AEL(G, N, J)$, 有 $AEL(G, N, J) \leq OEL(G, N, J)$, 所以形式背景 (G, N, J) 是基于 $AEL(G, N, J)$ 和 $OEL(G, N, J)$ 协调的。

接下来, 给出相应的规则定义。

定义 16 设形式背景 (G, M, I) 是基于 $AEL(G, M, I)$ 和 $OEL(G, M, I)$ 协调的, 若对于 $X \neq \emptyset, G, (Y, Z) \neq (\emptyset, \emptyset), (G, G)$, 有 $(X, (A, B)) \in OEL(G, M, I)$, $((Y, Z), C) \in AEL(G, M, I)$, 且 $X=Y$ 或 Z , 则称 $A \rightarrow B$ 是一个规则。

在上述定义的基础上, 以下比较对象导出三支概念格与属性导出三支概念格下的生成规则。

定理 5 设 (G, M, I, N, J) 是决策背景, 且 $AEL(G, M, I) \leq AEL(G, N, J)$, $OEL(G, M, I) \leq OEL(G, N, J)$, $AEL(G, M, I) \leq OEL(G, N, J)$, 则 $TR_o \subseteq TR_a$

证明: 对于任意 $A \rightarrow B \in TR_o$, 由定义12, 存在 $(X, (A, C)) \in OEL(G, M, I)$, $(X, (B, D)) \in OEL(G, N, J)$ 。因为 $AEL(G, M, I) \leq OEL(G, N, J)$, 因此对于 $(X, (B, D)) \in OEL(G, N, J)$, 存在 $((X, Z), A)$ 或 $((Y, X), A) \in AEL(G, M, I)$ 。又因为 $AEL(G, N, J) \leq OEL(G, N, J)$, 所以 $(X, (B, D)) \in OEL(G, N, J)$ 等价于 $((X, Z), A)$ 或 $((Y, X), A) \in AEL(G, N, J)$, 因此对于 $((X, Z), A)$ 或 $((Y, X), A) \in AEL(G, N, J)$, 存在 $((X, Z), A)$ 或 $((Y, X), A) \in AEL(G, M, I)$ 。因此 $A \rightarrow B \in TR_a$, 即 $TR_o \subseteq TR_a$

例8: 比较表1所示决策背景分别在对象导出三支概念格及属性导出三支概念格下的生成规则集合 TR_o 和 TR_a , 可以看出, TR_a 不仅包含了 TR_o 中每条规则, 还包含了 $abe \rightarrow h; abcf \rightarrow h; abcef \rightarrow h$ 等规则。

通过与属性导出三支概念格下的生成规则进行比较, 可以看出, 属性导出三支概念格下的生成规则更为丰富。但是, 属性导出三支概念格比对象导出三支概念格生成了更多的冗余规则, 如图5、图6所示属性导出三支概念格生成了 $de \rightarrow g, abc \rightarrow h, abe \rightarrow h, abcf \rightarrow h, abcef \rightarrow h, abcef \rightarrow hk$ 等6条冗余规则, 而图1、图2所示对象导出三支概念格只生成了 $de \rightarrow g, abc \rightarrow h, abcef \rightarrow hk$ 等3条冗余规则。因此, 在支持决策分析方面, 属性导出三支概念格的能力要稍弱于对象导出三支概念格。

3 基于三支概念格合并的决策背景规则提取

与属性导出三支概念格类似, 对象导出三支概念格下的生成规则中也引入了一些冗余规则。冗余规则是指^[18], 对规则 $B \rightarrow C$ 和 $B' \rightarrow C'$, 若 $B \subseteq B'$ 且 $C' \subseteq C$, 则称规则 $B' \rightarrow C'$ 是冗余的。例如, 由图5、图6, 属性导出三支概念格下的生成规则中 $de \rightarrow g, abc \rightarrow h, abe \rightarrow h, abcf \rightarrow h, abcef \rightarrow h, abcef \rightarrow hk$ 均为冗余规则; 由图1、图2, 对象导出三支概念格下的生成规则中 $de \rightarrow g, abc \rightarrow h, abcef \rightarrow hk$ 均为冗余规则。

进一步分析原因, 主要是因为基于对象导出三支概念格的规则提取时考虑的是对象特性, 而基于属性导出三支概念格的规则提取时考虑的是属性特性。因此单纯基于对象导出或属性导出三支概念格进行规则提取时, 就会生成一些冗余规则。

为进一步消除对象导出三支概念格或属性导出三支概念格下生成的冗余规则, 以下通过对象导出三支概念格与属性导出三支概念格的合并, 使得在进行规则提取时受双重特性(对象特性和属性特性)约束, 从而进一步减少冗余规则的生成。为此, 先给出对象/属性导出合并三支概念格的定义。

定义 17 (G, M, I, N, J) 是形式背景, 若 $((X, Y), A) \in AEL(G, M, I)$, $(X, (A, B))$ 或 $(Y, (B, A)) \in OEL(G, M, I)$, 则 $((X, Y), (A, B))$ 称为对象/属性导出合并三支概念, 简称 OAE-概念, 这里 $(X, Y) \supseteq A$ 且 $A \subseteq (X, Y)$, $X \subseteq (A, B)$ 且 $(A, B) \supseteq X$; 或 $(X, Y) \supseteq A$ 且 $A \subseteq (X, Y)$, $Y \subseteq (B, A)$ 且 $(B, A) \supseteq Y$ 。其中 $X, Y \subseteq G, A, B \subseteq M$ 。 (X, Y) 和 (A, B) 称为 OAE-概念的外延和内涵。

对于 $((X, Y), (A, B)), ((Z, W), (C, D))$, 定义其偏序关系 $((X, Y), (A, B)) \leq ((Z, W), (C, D)) \Leftrightarrow (X, Y) \subseteq (Z, W) \wedge (C, D) \subseteq (A, B)$, 形式背景 (G, M, I) 的所有 OAE-概念在上面定义的偏序关系下是完备格, 称为对象/属性导出合并三支概念格, 记为 $OAEL(G, M, I)$ 。

例9：合并图1的 $OEL(G,M,I)$ 和图5的 $AEL(G,M,I)$ ，以及图2的 $OEL(G,N,J)$ 和图6的 $AEL(G,N,J)$ ，所得对象/属性导出合并三支概念格 $OAEL(G,M,I)$ 和 $OAEL(G,N,J)$ 如图7、图8所示。

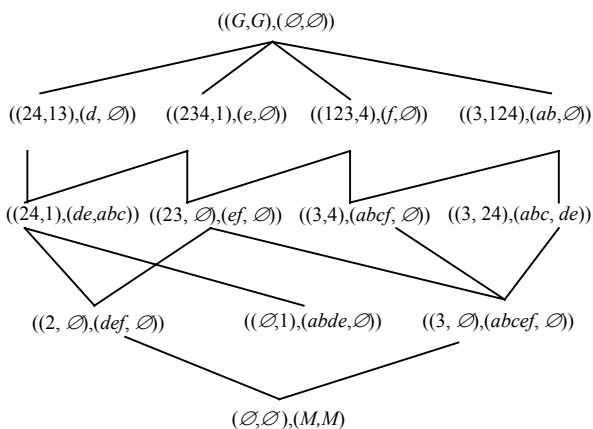


图7 对象/属性导出合并三支概念格 $OAEL(G,M,I)$

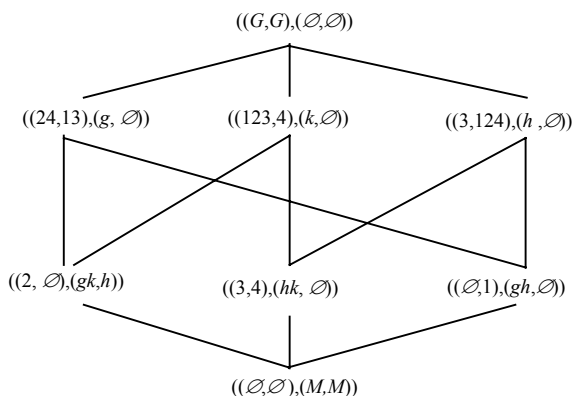


图8 对象/属性导出合并三支概念格 $OAEL(G,N,J)$

接下来，定义对象/属性导出合并三支概念格之间的细于关系，以及决策背景基于对象/属性导出合并三支概念格的协调性。

定义 18 设 $OAEL(G,M,I)$ 与 $OAEL(G,N,J)$ 是对象/属性导出合并三支概念格，若对于任意 $((Z,W),(C,D)) \in OAEL(G,N,J)$ ，存在 $((X,Y),(A,B)) \in OAEL(G,M,I)$ ，使得 $X=Z$ 且 $Y=W$ ，则称 $OAEL(G,M,I)$ 细于 $OAEL(G,N,J)$ ，记作 $OAEL(G,M,I) \leq OAEL(G,N,J)$ ，并称决策背景 (G,M,I,N,J) 是基于对象/属性导出合并三支概念格协调的。

例10：表1所示决策背景的对象/属性导出合并三支概念格 $OAEL(G,M,I) \leq OAEL(G,N,J)$ ，所以该决策背景是基于对象/属性导出合并三支概念格协调的。

在上述定义的基础上，进一步定义对象/属性导出合并三支概念格下的生成规则。

定义 19 设决策背景 (G,M,I,N,J) 是基于对象/属性导出合并三支概念格协调的，若对于 (X,Y) 、

$(Z,W) \neq (\emptyset, \emptyset)$ 、 (G,G) ，有 $((X,Y),(A,B)) \in OAEL(G,M,I)$ ， $((Z,W),(C,D)) \in OAEL(G,N,J)$ 且 $X=Z$ 、 $Y=W$ ，则称 $A \rightarrow C$ 、 $B \rightarrow D$ 分别是一个规则。

例11：由图7、图8可以看出，因为 $((\emptyset,1),(gh,\emptyset)) \in OAEL(G,N,J)$ ， $((\emptyset,1),(abde,\emptyset)) \in OAEL(G,M,I)$ ，根据定义19，可得规则 $abde \rightarrow gh$ ， $\emptyset \rightarrow \emptyset$ (舍去)。

以下给出基于对象/属性导出合并三支概念格的规则提取算法。

算法1 基于合并三支概念格的规则提取算法

输入：决策形式背景 (G,M,I,N,J) ；

输出：决策规则集合 C ；

- 1) $C = \emptyset$;
- 2) 构建对象导出三支概念格 $OEL(G,M,I)$ 和 $OEL(G,N,J)$;
- 3) 构建属性导出三支概念格 $AEL(G,M,I)$ 和 $AEL(G,N,J)$;

4) 合并 $OEL(G,M,I)$ 和 $AEL(G,M,I)$ ，构建对象/属性导出合并三支概念格 $OAEL(G,M,I)$ ；

① 对于 $AEL(G,M,I)$ 中每个概念 $((X,Y),A)$ ，若 $OEL(G,M,I)$ 中存在对应概念 $(X,(A,B))$ 或 $(Y,(B,A))$ ，则生成新概念 $((X,Y),(A,B))$ ；

② 由所有概念 $((X,Y),(A,B))$ 生成对象/属性导出合并三支概念格 $OAEL(G,M,I)$ ；

5) 合并 $OEL(G,N,J)$ 和 $AEL(G,N,J)$ ，构建对象/属性导出合并三支概念格 $OAEL(G,N,J)$ ；

① 对于 $AEL(G,N,J)$ 中每个概念 $((X,Y),A)$ ，若 $OEL(G,N,J)$ 中存在对应概念 $(X,(A,B))$ 或 $(Y,(B,A))$ ，则生成新概念 $((X,Y),(A,B))$ ；

② 由所有概念 $((X,Y),(A,B))$ 生成对象/属性导出合并三支概念格 $OAEL(G,N,J)$ ；

6) 对于 $OAEL(G,N,J)$ 中每个概念 $((X,Y),(C,D))$ ，重复下列操作：

若 $OAEL(G,M,I)$ 中存在对应概念 $((X,Y),(A,B))$ ，则 $C = C \cup \{A \rightarrow C, B \rightarrow D\}$ ；

7) 舍去 C 中所有前提或结论为 \emptyset 的规则，输出 C 。

该算法综合利用了对象导出三支概念格所包含的对象信息和属性导出三支概念格所包含的属性信息，这是因为OE概念 $(X,(A,B))$ 刻画了对象子集 X 共同具有的属性子集 A 和共同不具有的属性子集 B ，而AE概念 $((X,Y),A)$ 刻画了共同具有属性子集 A 的对象子集 X 和共同不具有属性子集 A 的对象子集 Y ，因此相对于单纯的对象导出三支概念格或属性导出三支概念格，所得的对象/属性导出合并三支概念格同时包含了对象和属性信息，进而也使得所生成的规则在

冗余性方面得到一定程度的降低。

另外, 该算法的实现效率主要取决于对象导出三支概念格和属性导出三支概念格中的三支概念构建算法。关于三支概念的构建, 文献[19]通过对形式背景及其补背景的变换和并置, 提出一种基于二支算子的三支概念构建算法。文献[20]针对三支概念分析理论中三支概念数量庞大、构建耗时的问题, 提出一种三支概念的并行构建算法。考虑到本文示例中的形式背景比较简单, 因此借鉴的是一般形式概念的构建方法, 在后续研究中将结合具体的应用, 尝试采用上述两类新的三支概念构建算法。

例12: 应用上述算法, 由图7、图8的对象/属性导出合并三支概念格 $OAEL(G, M, I)$ 和 $OAEL(G, N, J)$ 所得的生成规则集合如下:

$d \rightarrow g; ab \rightarrow h; f \rightarrow k; def \rightarrow gk; abde \rightarrow gh; abcf \rightarrow hk$

比较前面的属性导出三支概念格下的生成规则集合:

$d \rightarrow g; de \rightarrow g; ab \rightarrow h; abe \rightarrow h; abc \rightarrow h; abcf \rightarrow h;$

$abcef \rightarrow h; f \rightarrow k; def \rightarrow gk; abde \rightarrow gh;$

$abcf \rightarrow hk; abcef \rightarrow hk$

以及对象导出三支概念格下的生成规则集合:

$d \rightarrow g; de \rightarrow g; ab \rightarrow h; abc \rightarrow h; f \rightarrow k; def \rightarrow gk;$

$abde \rightarrow gh; abcf \rightarrow hk; abcef \rightarrow hk$

可以看出, 对属性导出三支概念格和对对象导出三支概念格进行合并以后, 进一步消除了原来的冗余规则。

4 结束语

在决策背景的规则提取方面, 现有文献基于属性导出三支概念格进行规则提取, 并取得了较好的效果。本文在其基础上进一步研究了基于对象导出三支概念格的规则提取, 并与经典概念格及属性导出三支概念格下的规则提取进行了比较; 同时通过对对象导出三支概念格和属性导出三支概念格的合并, 定义了对象/属性导出合并三支概念格, 并提出了基于对象/属性导出合并三支概念格的规则提取算法。理论分析和实例结果也表明了本文工作的有效性。下一步将本文提出的决策规则提取算法应用到地理课程自动解题研究, 以自动获取用于地理试题解答的隐式解题规则。

参 考 文 献

[1] GANTE R B, WILLE. Formal concept analysis[M]// Mathematical Foundations. New York: Springer-verlag, 1999.

- [2] YAO Y Y. Concept lattices in rough set theory[C]//The 2004 Annual Meeting of the North American Fuzzy Information Processing Society (NAFIPS 2004). [S.l.]: IEEE, 2004, 6: 796-801.
- [3] 张文修, 姚一豫, 梁怡. 粗糙集与概念格[M]. 西安: 西安交通大学出版社, 2006.
ZHANG Wen-xiu, YAO Yi-yu, LIANG Yi. Rough set and concept lattice[M]. Xi'an: Xi'an Jiaotong University Press, 2006.
- [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.
ZHANG Wen-xiu, QIU Guo-fang. Uncertain decision making based on rough sets[M]. Beijing: Tsinghua University Press, 2005.
- [5] 王国胤. 粗糙集理论与知识获取[M]. 西安: 西安交通大学出版社, 2001.
WANG Guo-yin. Rough set theory and knowledge acquisition[M]. Xi'an: Xi'an Jiaotong University Press, 2001.
- [6] 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, 2004.
ZHANG Wen-xiu, LIANG Yi, WU Wei-zhi. Information system and knowledge discovery[M]. Beijing: Science Press, 2004.
- [7] 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.
ZHANG Wen-xiu, MI Ju-sheng, WU Wei-zhi. Knowledge reductions in inconsistent information systems[J]. Chinese Journal of Computers, 2003, 26(1): 12-18.
- [8] 魏玲. 粗糙集与概念格的约简理论与方法[D]. 西安: 西安交通大学, 2005.
WEI Ling. The reduction theory and method of rough set and concept lattice[D]. Xi'an: Xi'an Jiaotong University, 2005.
- [9] 张文修, 魏玲, 祁建军. 概念格的属性约简理论与方法[J]. 中国科学E辑: 信息科学, 2005, 35(6): 628-639.
ZHANG Wen-xiu, WEI Ling, QI Jian-jun. The reduction theory and method of concept lattice[J]. Science in China Series E: Information Science, 2005, 35(6): 628-639.
- [10] 魏玲, 祁建军, 张文修. 决策形式背景的概念格属性约简[J]. 中国科学E辑: 信息科学, 2008, 38(2): 195-208.
WEI Ling, QI Jian-jun, ZHANG Wen-xiu. The reduction of concept lattice in formal decision context[J]. Science in China Series E: Information Science, 2008, 38(2): 195-208.
- [11] LI J H, MEI C L, LÜ Y J. Incomplete decision contexts: approximate concept construction, rule acquisition and knowledge reduction[J]. International Journal of Approximate Reasoning, 2013, 54(1): 149-165.
- [12] LI J H, MEI C L, WANG J H, et al. Rule-preserved object compression in formal decision contexts using concept lattices[J]. Knowledge-Based Systems, 2014, 71: 435-444.
- [13] YAO Y. An outline of a theory of three-way decisions [C]//The 2012 International Conference on Rough Sets and Current Trends in Computing(LNCS 7413). [S.l.]: Springer International Publishing, 2012: 1-17.

- [14] QI J J, WEI L, YAO Y Y. Three-way formal concept analysis[J]. Lecture Notes in Computer Science, 2014, 8818: 732-741.
- [15] QI J J, QIAN T, WEI L. The connections between three-way and classical lattices[J]. Knowledge-Based Systems, 2016, 91: 143-151.
- [16] 李金海, 邓硕. 概念格与三支决策及其研究展望[J]. 西北大学学报(自然科学版), 2017, 47(3): 321-329.
LI Jin-hai, DENG Shuo. Concept lattice, three-way decisions and their research outworks[J]. Journal of Northwest University(Natural Science Edition), 2017, 47(3): 321-329.
- [17] REN R S, WEI L. The attribute reductions of three-way concept lattices[J]. Knowledge-Based Systems, 2016, 99: 92-102.
- [18] 刘琳, 钱婷, 魏玲. 基于属性导出三支概念格的决策背景规则提取[J]. 西北大学学报(自然科学版), 2016, 46(4): 481-487.
- LIU Lin, QIAN Ting, WEI Ling. Rules extraction in formal decision contexts based on attribute-induced three-way concept lattices[J]. Journal of Northwest University(Natural Science Edition), 2016, 46(4): 481-487.
- [19] QIAN T, WEI L, QI J J. Constructing three-way concept lattices based on apposition and subposition of formal contexts[J]. Knowledge-Based Systems, 2016, 116: 39-48.
- [20] 祁建军, 汪文威. 多线程并行构建三支概念[J]. 西安交通大学学报(自然科学版), 2017, 51(3): 116-121.
QI Jian-jun, WANG Wen-wei. A multithreaded parallel algorithm for constructing three-way concepts[J]. Journal of Xi'an Jiaotong University, 2017, 51(3): 116-121.

编辑 蒋晓