

三支概念的恢复集

陈曜琦,徐伟华*,蒋宗颖

(西南大学人工智能学院,重庆 400715)

摘要:将三支形式概念分析这一工具引入到数据恢复领域,通过定义三支概念的恢复集和恢复度,研究三支概念间的隐藏信息,提出了一种有效的形式背景恢复算法。同时,针对三支概念恢复集问题,研究三支概念对形式背景二元关系的约束,设计了恢复集的合取范式化简(conjunctive normal form simplification, CNFS)算法,进一步给出了恢复集的动态更新算法,以适应形式背景的不断变化。最后,使用UCI机器学习数据库中的数据对CNFS算法进行了测试。实验结果表明,CNFS算法在形式背景恢复方面具有较高的准确性和有效性,同时也验证了不同概念对认知的重要程度是不同的。

关键词:合取范式;动态更新;三支概念;形式概念分析

中图分类号:TP181 **文献标志码:**A

引用格式:陈曜琦,徐伟华,蒋宗颖. 三支概念的恢复集[J]. 山东大学学报(理学版),2023,58(12):52-62.

Recovery set of three-way concept

CHEN Yaoqi, XU Weihua*, JIANG Zongying

(School of Artificial Intelligence, Southwest University, Chongqing 400715, China)

Abstract: The theory of three-way formal concept analysis is introduced into the field of data recovery in this paper. By defining the recovery set and recovery degree of three-way concept, it explores the hidden information from three-way concepts, and proposes an effective formal context recovery algorithm. Additionally, to solve the three-way concept recovery set problem, the constraints of three-way concept are considered on formal context binary relations, and a conjunctive normal form simplification algorithm is designed for the recovery set (CNFS). Furthermore, a dynamic update algorithm is provided for the recovery set to adapt to the continuous changes of the formal context. Finally, some numerical experiments on public datasets from the UCI perform the effectiveness of our proposed method. Experimental results indicate that the proposed algorithm has high accuracy and effectiveness in formal context recovery, and also verifies that different concepts have different important degrees in cognition.

Key words: conjunctive normal form; dynamic update; three-way concept; formal context analysis

0 引言

德国数学家 Wille^[1]给出了概念的形式化描述,继而产生了形式概念分析理论,该理论是知识发现、数据挖掘的有效数学理论。Qi等^[2]将三支决策思想^[3]引入形式概念分析理论中,提出了三支概念这一新概念。三支概念分析融合了三支决策理论与形式概念分析的优势,语义描述更加精确细致,使得形式概念分析研究有了新的发展。三支概念作为经典概念的另一种延伸可以更全面地描述概念信息。Zhan等^[4]首先在不完全模糊决策系统中构建了一个三支决策模型,并将其应用于不完全多属性决策问题的建模,为现实中的不完

收稿日期:2022-08-12;网络出版时间:2023-10-26 13:33:45

网络出版地址:https://link.cnki.net/urlid/37.1389.N.20231026.1005.002

基金项目:国家自然科学基金资助项目(62376229);重庆市研究生科研创新项目(CYS21133)

第一作者简介:陈曜琦(1998—),女,硕士研究生,研究方向为粒计算、认知计算。E-mail:1312941662@qq.com

*通信作者简介:徐伟华(1979—),男,教授,博士生导师,研究方向为粒计算、认知计算、信息融合、不确定性人工智能等。

E-mail:chxuwh@gmail.com

全多属性决策问题提供了一个新的视角,实验结果也证明了三支概念决策模型的有效性和优越性。Hao 等^[5-6]分析了三支概念的稳定,将其应用于自然语言生成,并进一步为社会网络中知识发现的三维概念格增量构建提出了新的问题和方法。Wei 等^[7]从三支概念格的角度对决策形式背景进行了规则提取。三支思想给数据挖掘提供了一种新的思路,近年来,三支概念分析在三支概念格构建^[8-12]、概念学习^[13-14]、模糊三支概念分析^[15-19]、不完备背景的三支概念获取问题^[20-23]等方面都取得了诸多成果,并被广泛应用于决策、规则提取、知识发现、数据挖掘^[24-29]。

在大数据时代,数据的收集和存储环节具有至关重要的意义;然而储存数据时可能会遇到数据丢失或被误删等风险,从而引发严重的后果,因此数据恢复策略的制定和实施成为了不可或缺的一部分。为了解决这个问题,形式概念分析这一强大的工具被引入到数据恢复领域。

众所周知,概念是人类认知的基本单位,人们通过概念进行判断和推理,科学抽象概念是人类理性思维对事物本质和客观世界发展过程的反映,使人们通过认识活动获得超越感官的认知。在认知过程中,概念是从感性认识到理性认识的抽象和概括,反映了所感知事物的共同本质特点,但随着时间的推移,原始样本可能已经不存在或难以采集,例如对于已经灭绝的动物恐龙,我们只能间接地了解关于它的一些非常少的概念。为了更深入地了解恐龙的详细特征,我们需要将概念重新转化为样本与属性特征的关系。通常我们可以找到一些已知的概念,通过挖掘这些概念间隐藏的信息,进一步有效地发现隐藏在其中的隐式概念及概念关系。目前,模拟人抽象概括能力的概念认知学习研究已引起许多学者的关注^[13-14,30-33]。

人脑不仅具有抽象概括的能力,还具有将抽象概念转化为具体事物的能力,但目前还没有模拟人具象能力的相关研究。三支概念是对形式背景中对象与属性之间关系的深刻反映,它们构成了形式背景的内在结构特征,在认知负荷、简化、流畅度和灵活性等方面具有显著的优势,使得它们在概念恢复集的构建中具有重要的应用价值^[34]。本文旨在通过研究三支概念间的隐藏信息,提出一种有效的形式背景恢复算法,在此基础上,探讨三支概念与形式背景之间的关系,并给出三支概念的等价合取范式,通过合取范式的简化,设计一种恢复集的生成算法,该算法能够有效地从部分三支概念中恢复形式背景的信息。此外,本文还进一步设计恢复集的动态更新算法,以适应形式背景的不断变化。最后,使用 UCI 机器学习数据库中的一些数据集对所提出的算法进行测试。实验结果表明,本文所提出的算法在形式背景恢复方面具有较高的准确性和有效性,同时也验证了不同概念对认知的重要程度是不同的,这一发现对于进一步丰富形式概念分析的理论研究具有重要的意义。

1 预备知识

1.1 三支概念分析基础知识

为了全文叙述方便,本小节给出三支概念分析的基础知识,包括正算子、负算子、对象导出三支算子、对象导出三支概念等。

定义 1^[35] 设 (OB, AT, I) 为形式背景,其中 $OB = \{o_1, o_2, \dots, o_n\}$ 为对象集, $AT = \{a_1, a_2, \dots, a_m\}$ 为属性集, $O \subseteq OB, A \subseteq AT, I = (OB \times AT) \setminus I$ 。若 $(o, a) \in I$, 则称 o 具有属性 a , 记为 oIa 。

定义算子 $*$: $\mathcal{P}(OB) \rightarrow \mathcal{P}(AT)$ 及 $*$: $\mathcal{P}(AT) \rightarrow \mathcal{P}(OB)$, 对于任意的 $O \subseteq OB$ 及 $A \subseteq AT$, 有:

$$\begin{aligned} O^* &= \{a \in AT \mid \forall o \in O(oIa)\} = \{a \in AT \mid O \subseteq Ia\}, \\ A^* &= \{o \in OB \mid \forall a \in A(oIa)\} = \{o \in OB \mid A \subseteq oI\}, \end{aligned}$$

其中: $Ia = \{o \in OB \mid (o, a) \in I\}$, $oI = \{a \in AT \mid (o, a) \in I\}$ 。

若 $O^* = A$ 且 $A^* = O$, 那么 (O, A) 被称为一个形式概念, 其中 O 是概念的外延, A 是概念的内涵。

定义算子 $\tilde{*}$: $\mathcal{P}(OB) \rightarrow \mathcal{P}(AT)$ 及 $\tilde{*}$: $\mathcal{P}(AT) \rightarrow \mathcal{P}(OB)$, 对于任意的 $O \subseteq OB$ 及 $A \subseteq AT$, 若 $(o, a) \in I^c$, 则称 o 不具有属性 a , 记为 $oI^c a$, 有:

$$\begin{aligned} O^{\tilde{*}} &= \{a \in AT \mid \forall o \in O(oI^c a)\} = \{a \in AT \mid O \subseteq I^c a\}, \\ A^{\tilde{*}} &= \{o \in OB \mid \forall a \in A(oI^c a)\} = \{o \in OB \mid A \subseteq oI^c\}, \end{aligned}$$

其中: $I^c a = \{o \in OB \mid (o, a) \in I^c\}$, $oI^c = \{a \in AT \mid (o, a) \in I^c\}$ ^[34]。

性质 1^[35]

- (1) $O_1 \subseteq O_2, O_2^* \subseteq O_1^*, A_1 \subseteq A_2, A_2^* \subseteq A_1^*$ 。
- (2) $O \subseteq O^{**}, A \subseteq A^{**}$ 。
- (3) $O^* = O^{***}, A^* = A^{***}$ 。
- (4) $O \subseteq A^* A \subseteq O^*$ 。
- (5) $(O_1 \cup O_2)^* = O_1^* \cap O_2^*, (A_1 \cup A_2)^* = A_1^* \cap A_2^*$ 。
- (6) $(O_1 \cap O_2)^* \supseteq O_1^* \cup O_2^*, (A_1 \cap A_2)^* \supseteq A_1^* \cup A_2^*$ 。

算子 $\tilde{*}$ 的性质与算子 $*$ 的性质相同。

形式背景 (OB, AT, I) 的概念可以用超概念与亚概念的关系来定义它们之间的序关系,

$$(O_1, A_1) \leq (O_2, A_2) \Leftrightarrow O_1 \subseteq O_2 (\Leftrightarrow A_1 \supseteq A_2)。$$

(OB, AT, I) 的所有概念的偏序集记为 $L(OB, AT, I)$, 称为概念格, 其中上确界与下确界的定义如下:

$$(O_1, A_1) \wedge (O_2, A_2) = (O_1 \cap O_2, (A_1 \cup A_2)^{**}),$$

$$(O_1, A_1) \vee (O_2, A_2) = ((O_1 \cup O_2)^{**}, A_1 \cap A_2)。$$

概念格是一个完备格, Ganter 等^[35] 给出定理, 说明一个形式背景可由它产生的所有概念重新组建。

性质 2^[35] 设 (OB, AT, I) 为形式背景, $L(OB, AT, I)$ 为其概念格, 则 $I = \cup \{O \times A \mid (O, A) \in L(OB, AT, I)\}$ 。

定义 2^[34] 设 (OB, AT, I) 为一个形式背景, 定义算子 $\triangleleft: \mathcal{P}(OB) \rightarrow \mathcal{D}\mathcal{P}(AT)$ 及 $\triangleright: \mathcal{D}\mathcal{P}(AT) \rightarrow \mathcal{P}(OB)$ 分别为: 对于任意的 $O \subseteq OB, A, B \subseteq AT, O^{\triangleleft} = (O^*, O^{\tilde{*}})$, 且 $(A, B)^{\triangleright} = \{o \in OB \mid o \in A^*, \text{ 且 } o \in B^{\tilde{*}}\} = A^* \cap B^{\tilde{*}}$, 称算子 \triangleleft 及 \triangleright 为对象导出三支算子, 简称 OE-算子。进一步, 若 $O^{\triangleleft} = (A, B)$, 且 $(A, B)^{\triangleright} = O$, 则称 $(O, (A, B))$ 为对象导出三支概念, 简称为 OE-概念, 其中称 O 为 OE-概念的外延, (A, B) 为 OE-概念的内涵。

性质 3^[34] 设 (OB, AT, I) 为形式背景, 则其概念格 $L(OB, AT, I)$ 、负概念格 $NL(OB, AT, I)$ 与对象导出的三支概念格 $OEL(OB, AT, I)$ 的内涵集 $(L_I(OB, AT, I), NL_I(OB, AT, I), OEL_I^+(OB, AT, I), OEL_I^-(OB, AT, I))$ 、外延集 $(L_E(OB, AT, I), NL_E(OB, AT, I), OEL_E(OB, AT, I))$ 之间有以下关系:

$$L_E(OB, AT, I) \subseteq OEL_E(OB, AT, I),$$

$$NL_E(OB, AT, I) \subseteq OEL_E(OB, AT, I),$$

$$L_I(OB, AT, I) = OEL_I^+(OB, AT, I),$$

$$NL_I(OB, AT, I) = OEL_I^-(OB, AT, I)。$$

例 1 表 1 给出了关于“生物与水”的形式背景 (OB, AT, I) ^[35]。对象集 OB 包含 8 个对象, 分别是 o_1 : leech、 o_2 : bream、 o_3 : frog、 o_4 : dog、 o_5 : spike-weed、 o_6 : reed、 o_7 : bean、 o_8 : maize。属性集 AT 包含 9 个属性, 分别是 a_1 : needs water to live、 a_2 : lives in water、 a_3 : lives no land、 a_4 : needs chlorophyll to produce foods、 a_5 : two seed leaves、 a_6 : one seed leaf、 a_7 : can move around、 a_8 : has limbs、 a_9 : suckles its offspring。

表 1 生物与水形式背景
Table 1 The context of living beings and water

OB	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9
o_1	1	1	0	0	0	0	1	0	0
o_2	1	1	0	0	0	0	1	1	0
o_3	1	1	1	0	0	0	1	1	0
o_4	1	0	1	0	0	0	1	1	1
o_5	1	1	0	1	0	1	0	0	0
o_6	1	1	1	1	0	1	0	0	0
o_7	1	0	1	1	1	0	0	0	0
o_8	1	0	1	1	0	1	0	0	0

表 1 中形式背景所对应对象导出的三支概念为:

$$(o_1 o_2 o_3 o_4 o_5 o_6 o_7 o_8, (a_1, \emptyset)), (o_1 o_2 o_3 o_4 o_5 o_6 o_8, (a_1, a_5)), (o_1 o_2 o_3 o_4 o_7, (a_1, a_6)),$$

$$\begin{aligned}
& (o_1o_2o_3o_7, (a_1, a_6a_9)), (o_1o_5o_6o_7o_8, (a_1, a_8a_9)), (o_1o_5o_6o_8, (a_1, a_5a_8a_9)), \\
& (o_1o_2o_5, (a_1a_2, a_3a_5a_9)), (o_1o_3o_6, (a_1a_2, a_5a_8a_9)), (o_1o_5, (a_1a_2, a_3a_5a_8a_9)), \\
& (o_3o_4o_6o_8, (a_1a_3, a_5)), (o_3o_4o_7, (a_1a_3, a_6)), (o_3o_6o_7o_8, (a_1a_3, a_9)), \\
& (o_3o_6o_8, (a_1a_3, a_5a_9)), (o_3o_7, (a_1a_3, a_6a_9)), (o_5o_6o_7o_8, (a_1a_4, a_7a_8a_9)), \\
& (o_1o_2o_3, (a_1a_2a_7, a_4a_5a_6a_9)), (o_1o_2, (a_1a_2a_7, a_3a_4a_5a_6a_9)), \\
& (o_1, (a_1a_2a_7, a_3a_4a_5a_6a_8a_9)), (o_5o_6o_8, (a_1a_3a_6, a_5a_7a_8a_9)), \\
& (o_2o_3o_4, (a_1a_7a_8, a_4a_5a_6)), (o_5o_6, (a_1a_2a_4a_6, a_5a_7a_8a_9)), \\
& (o_2, (a_1a_2a_7a_8, a_3a_4a_5a_6a_9)), (o_7, (a_1a_3a_4a_5, a_2a_6a_7a_8a_9)), \\
& (o_6o_8, (a_1a_3a_4a_6, a_5a_7a_8a_9)), (o_1o_2o_3o_5o_6o_7o_8, (a_1, a_9)), \\
& (o_1o_2o_3o_5o_6o_8, (a_1, a_5a_9)), (o_4o_8, (a_1a_3, a_2a_5)), (o_1o_7, (a_1, a_6a_8a_9)), \\
& (o_1o_2o_3o_5o_6, (a_1a_2, a_5a_9)), (o_1o_2o_3o_4, (a_1a_7, a_4a_5a_6)), (o_3o_4o_6o_7o_8, (a_1a_3, \emptyset)), \\
& (o_4o_7o_8, (a_1a_3, a_2)), (o_6o_7o_8, (a_1a_3a_4, a_7a_8a_9)), (o_4o_7, (a_1a_3, a_2a_6)), \\
& (o_5, (a_1a_2a_4a_6, a_3a_5a_7a_8a_9)), (o_2o_3, (a_1a_2a_7a_8, a_4a_5a_6a_9)), (o_3o_6, (a_1a_2a_3, a_5a_9)), \\
& (o_8, (a_1a_3a_4a_6, a_2a_5a_7a_8a_9)), (o_3o_4, (a_1a_3a_7a_8, a_4a_5a_6)), (o_6, (a_1a_2a_3a_4a_6, a_5a_7a_8a_9)), \\
& (o_3, (a_1a_2a_3a_7a_8, a_4a_5a_6a_9)), (o_4, (a_1a_3a_7a_8a_9, a_2a_4a_5a_6)), (o_7o_8, (a_1a_3a_4, a_2a_7a_8a_9))
\end{aligned}$$

注:根据形式概念分析理论的表示方法,除全集与空集外,其他集合用其元素序列表示。

1.2 合取范式基础知识

设 X 表示命题变量的集合,若 X 由 n 个变量 x_1, x_2, \dots, x_n 组成,则记 $X = \{x_1, x_2, \dots, x_n\}$ 。对于一组布尔变量 $X = \{x_1, x_2, \dots, x_n\}$,其真值赋值是一个 n 元布尔函数 $S(X): X \rightarrow \{0, 1\}^n$ 在 X 上存在 2^n 种不同的真值赋值。如果 $S(x_i) = 1$,则称 x_i 在 $S(x)$ 赋值下取真值,否则为假值。

定义 3^[36] 对任意变量 x_i ,符号 x_i 和 $\neg x_i$ 是其文字。正文字 x_i ,在真值赋值 $S(X)$ 下取真值当且仅当 $S(x_i) = 1$;负文字 $\neg x_i$,在真值赋值 $S(X)$ 下取真值当且仅当 $S(x_i) = 0$ 。 X 上的子句是 X 中有限个文字的析取,用 C 表示, $C = l_1 \vee l_2 \vee \dots \vee l_k$ 在真值赋值 $S(X)$ 下取真值(或称子句 C 在真值赋值 $S(X)$ 下是可满足的),当且仅当子句包含的文字中至少有一个在真值赋值 $S(X)$ 取真值。 $k = |C|$ 表示子句 C 中的文字数,称为子句长度。一个长度为 k 的子句是一个 k -子句,当 $k = 1$ 时,称为单元子句。一个子句中的文字不允许重复出现,因为重复出现的文字会被子句忽略,不会影响公式的可满足性。

X 上的合取范式 $F(X)$ 是 X 上的一些子句的合取, $F(X) = C_1 \wedge C_2 \wedge \dots \wedge C_m$ 。合取范式 $F(X)$ 在真值赋值 $S(X)$ 下取真值(或称 $F(X)$ 在真值赋值 $S(X)$ 下是可满足的),当且仅当 $F(X)$ 中包含的所有子句 C 在真值赋值 $S(X)$ 下都是取真值的。合取范式(conjunctive normal form, CNF)可简单描述为 $F(X) = \bigwedge_i^m = 1(\bigvee_{j=1}^k l_{i,j})$ 。

2 三支概念的恢复集理论

三支概念是由形式背景导出的,其中蕴含了形式背景中对象与属性的关系,揭示了形式背景中对象共同具有和共同缺失的属性,是对象共性的明确描述,因此可以通过三支概念来研究还原形式背景中的二元关系。由于对象导出的三支概念与属性导出的三支概念在形式背景恢复中有许多相似之处,因此本文只探讨对象导出的三支概念。

2.1 三支概念的恢复集

当已知一些三支概念时,可以得到一些对象与属性的关系。为了挖掘隐藏在概念间的信息,可以通过形式背景的二元关系来刻画隐藏在三支概念间的信息,因此恢复集定义如下。

定义 4 设 (OB, AT, I) 为形式背景, $OC(OB, AT, I)$ 为对象导出的三支概念集(简记为 OC), $\mathcal{P}(OC)$ 为 OC 的幂集, $\mathcal{F} \subseteq OC$ 。定义:

$$PI_{\mathcal{F}} = \{(o_i, a_j) \in OB \times AT \mid (o_i, a_j) \in I \Rightarrow \mathcal{F} \notin \mathcal{P}(OC)\}$$

$$NI_{\mathcal{F}} = \{(o_i, a_j) \in OB \times AT \mid (o_i, a_j) \in I \Rightarrow \mathcal{F} \in \mathcal{P}(OC)\},$$

称 $PI_{\mathcal{F}}$ 为 \mathcal{F} 的正恢复集, $NI_{\mathcal{F}}$ 为 \mathcal{F} 的负恢复集, $R_{\mathcal{F}} = PI_{\mathcal{F}} \cup NI_{\mathcal{F}}$ 为三支概念集 \mathcal{F} 的恢复集,简称 \mathcal{F} 恢复集。

由上面定义可知,若 $(o_i, a_j) \in PI_{\mathcal{F}}$,则 $(o_i, a_j) \in I \Rightarrow \mathcal{F} \notin \mathcal{P}(OC)$,又因为前提条件 $\mathcal{F} \subseteq OC$,所以可得 $(o_i, a_j) \in I$,由此可知正恢复集即为三支概念集 \mathcal{F} 所能确定的最大正二元关系集合。同理,负恢复集就是已知三支概念 \mathcal{F} 所能确定的最大负二元关系集合。

例 2 (续例 1) 在例 1 的形式背景中,给出对象集为 $OB = \{o_1, o_2, \dots, o_8\}$,属性集为 $AT = \{a_1, a_2, \dots, a_9\}$,且已知概念集 $\mathcal{F} = \{(o_1 o_2 o_3 o_4, (a_1 a_7, a_4 a_5 a_6))\} \subseteq OC$ 。对于集合 $OB \times AT$ 中的元素 (o_1, a_1) ,若 $(o_1, a_1) \in I$,则 $\{o_1, o_2, o_3, o_4\}^* \neq \{a_1, a_7\}, (o_1 o_2 o_3 o_4, (a_1 a_7, a_4 a_5 a_6)) \notin OC(OB, AT, I)$,根据定义 4 可知 $(o_1, a_1) \in PI_{\mathcal{F}^c}$

类似地,对集合 $OB \times AT$ 中其他元素进行判断,即可得到

$$PI_{\mathcal{F}} = \{(o_1, a_1), (o_1, a_7), (o_2, a_1), (o_2, a_7), (o_3, a_1), (o_3, a_7), (o_4, a_1), (o_4, a_7)\}。$$

因为人的推理过程实际上是信息传递的过程,是信息提取、变换、再生、组织的过程,所以在不确定推理中计算不确定性与信息量是十分重要的。如果能够对这种不确定性推理中的不确定性与信息量进行计算,那么便容易在计算机上得以实现这种不确定推理^[37]。我们发现不同的概念集合能够恢复形式背景二元关系的程度有可能不同,也就是所含的信息量不同。为了定量描述三支概念集对形式背景二元关系的恢复程度,下面给出恢复度的定义。

定义 5 设 (OB, AT, I) 为形式背景, $OC(OB, AT, I)$ 是对象导出的三支概念集, $\mathcal{F} \subseteq OC(OB, AT, I)$,称

$$E(\mathcal{F}) = \frac{|R_{\mathcal{F}}|}{|OB \times AT|}$$

为概念集 \mathcal{F} 的恢复度。

由以上定义可知,恢复度有以下性质。

性质 4 (单调性) 若 $\mathcal{G} \subseteq \mathcal{F} \subseteq OC(OB, AT, I)$,则 $E(\mathcal{G}) \leq E(\mathcal{F})$ 。

性质 5 (最大值) 若 $\mathcal{F} = OC(OB, AT, I)$,则概念集的恢复度达到最大值 1。除此之外,如果所选择的概念集为一个对象导出三支概念的概念约简或者概念协调集,其恢复度也为 1^[38-39]。

性质 6 (最小值) 若 $\mathcal{F} = \emptyset$ 或 $\mathcal{F} = \{(OB, (\emptyset, \emptyset)), (\emptyset, (AT, AT))\}$,则概念集的恢复度达到最小值 0。

性质 7 (有界性) 若 $\mathcal{F} \subseteq OC(OB, AT, I)$,则 $0 \leq E(\mathcal{F}) \leq 1$ 。

根据上述恢复度的定义,进一步给出 α 概念集的定义。

定义 6 设 $\mathcal{F} \subseteq OC(OB, AT, I)$,若 $E(\mathcal{F}) = \alpha$,则称 \mathcal{F} 是恢复度为 α 的形式背景恢复概念集,也称 α 概念集。若进一步对于任意的 $(O_i, (A_i, B_i)) \in \mathcal{F}, \mathcal{F}' = \mathcal{F} \setminus \{(O_i, (A_i, B_i))\}$ 有 $E(\mathcal{F}') < \alpha$,则称 \mathcal{F} 为恢复度为 α 的极小形式背景恢复概念集,简称极小 α 概念集。

2.2 三支概念恢复集的 CNFS 算法

为了更清晰地解释恢复集的求解过程,将逻辑公式引入恢复集的求解,给出三支概念与合取范式的关系,并利用范式化简的方法来计算三支概念的恢复集。

设 (OB, AT, I) 为一个形式背景, $OB = \{o_1, o_2, \dots, o_n\}$ 为对象集, $AT = \{a_1, a_2, \dots, a_m\}$ 为属性集,记命题变量 x_{ij} 当且仅当 $(o_i, a_j) \in I$ 时,下面性质显然成立。

性质 8 若 $(O, (A, B)) \in OC(OB, AT, I)$,则 $C_1 = \bigwedge_{(o_i, a_j) \in O \times A} x_{ij}$ 为真。

性质 9 若 $(O, (A, B)) \in OC(OB, AT, I)$,则 $C_2 = \bigwedge_{(o_i, a_j) \in O \times B} \neg x_{ij}$ 为真。

定理 1 若 $(O, (A, B)) \in OC(OB, AT, I)$,则 $C_3 = \bigwedge_{o_i \in O^c} (\bigvee_{a_j \in A} \neg x_{ij} \bigvee_{a_j \in B} x_{ij})$ 为真。

证明 若 C_3 为假,即 $\exists o_i \in O^c$ s.t. $\bigvee_{a_j \in A} \neg x_{ij} \bigvee_{a_j \in B} x_{ij}$ 为假,即 $\bigwedge_{a_j \in A} x_{ij} \bigwedge_{a_j \in B} \neg x_{ij}$ 为真, $o_i \in (A, B)^\triangleright = O$,矛盾,所以 C_3 为真。

定理 2 若 $(O, (A, B)) \in OC(OB, AT, I)$,则 $C_4 = \bigwedge_{a_j \in (A \cup B)^c} (\bigvee_{o_i \in O} x_{ij}) \wedge (\bigvee_{o_i \in O} \neg x_{ij})$ 为真。

证明 若 C_4 为假,即 $\exists a_j \in (A \cup B)^c$ s.t. $(\bigvee_{o_i \in O} x_{ij}) \wedge (\bigvee_{o_i \in O} \neg x_{ij})$ 为假,即 $(\bigwedge_{o_i \in O} x_{ij}) \vee (\bigwedge_{o_i \in O} \neg x_{ij})$ 为真, $a_j \in O^* \cup O^{\bar{*}} = A \cup B$,矛盾,所以 C_4 为真。

定理 3 $(O, (A, B)) \in OC(OB, AT, I)$ 当且仅当 $CNF_{(O, (A, B))} = C_1 \wedge C_2 \wedge C_3 \wedge C_4$ 为真。

证明如下。

必要性 由性质 8、性质 9、定理 1、定理 2 可知若 $(O, (A, B)) \in OC(OB, AT, I)$, 则 $CNF_{(O, (A, B))} = C_1 \wedge C_2 \wedge C_3 \wedge C_4$ 为真。

充分性 若 C_1 为真, 则 $O^* \supseteq A, A^* \supseteq O$; 若 C_2 为真, 则 $O^* \supseteq B, B^* \supseteq O$; 若 C_3 为真, 则 $A^* \cap B^* \subseteq O$; 若 C_4 为真, 则 $O^* \subseteq A, O^* \subseteq B$; 因此可得 $A^* \cap B^* = O, O^* = A, O^* = B$, 根据三支概念的定义可知 $(O, (A, B)) \in OC(OB, AT, I)$ 。

事实上, 由于 $CNF_{(O, (A, B))}$ 是三支概念 $(O, (A, B))$ 对形式背景二元关系的约束限制, 因此称 $CNF_{(O, (A, B))}$ 为三支概念 $(O, (A, B))$ 的恢复约束。

例 3 (续例 1) 在例 1 的形式背景中, 若取对象集为 $OB = \{o_1, o_2, \dots, o_8\}$, 属性集为 $AT = \{a_1, a_2, \dots, a_9, (o_1 o_2 o_3 o_4, (a_1 a_7, a_4 a_5 a_6))\} \in OC(OB, AT, I)$, 则它的恢复约束为

$$CNF_{(o_1 o_2 o_3 o_4, (a_1 a_7, a_4 a_5 a_6))} = x_{11} \wedge x_{17} \wedge x_{21} \wedge x_{27} \wedge x_{31} \wedge x_{37} \wedge x_{41} \wedge x_{47} \wedge \neg x_{14} \wedge \neg x_{15} \wedge \neg x_{16} \wedge \neg x_{24} \wedge \neg x_{25} \wedge \neg x_{26} \wedge \neg x_{34} \wedge \neg x_{35} \wedge \neg x_{36} \wedge \neg x_{44} \wedge \neg x_{45} \wedge \neg x_{46} \wedge (\neg x_{51} \vee \neg x_{57} \vee x_{54} \vee x_{55} \vee x_{56}) \wedge (\neg x_{61} \vee \neg x_{67} \vee x_{64} \vee x_{65} \vee x_{66}) \wedge (\neg x_{71} \vee \neg x_{77} \vee x_{74} \vee x_{75} \vee x_{76}) \wedge (\neg x_{81} \vee \neg x_{87} \vee x_{84} \vee x_{85} \vee x_{86}) \wedge (x_{12} \vee x_{13} \vee x_{18} \vee x_{19}) \wedge (\neg x_{12} \vee \neg x_{13} \vee \neg x_{18} \vee \neg x_{19}) \wedge (x_{12} \vee x_{13} \vee x_{18} \vee x_{19}) \wedge (\neg x_{12} \vee \neg x_{13} \vee \neg x_{18} \vee \neg x_{19}) \wedge (x_{22} \vee x_{23} \vee x_{28} \vee x_{29}) \wedge (\neg x_{22} \vee \neg x_{23} \vee \neg x_{28} \vee \neg x_{29}) \wedge (x_{32} \vee x_{33} \vee x_{38} \vee x_{39}) \wedge (\neg x_{32} \vee \neg x_{33} \vee \neg x_{38} \vee \neg x_{39}) \wedge (x_{42} \vee x_{43} \vee x_{48} \vee x_{49}) \wedge (\neg x_{42} \vee \neg x_{43} \vee \neg x_{48} \vee \neg x_{49})。$$

由上知, 下面推论显然成立。

推论 1 $\mathcal{F} \in \mathcal{P}(C)$ 当且仅当 $CNF_{\mathcal{F}}$ 为真, 其中 $CNF_{\mathcal{F}} = \bigwedge_{(O, (A, B)) \in \mathcal{F}} CNF_{(O, (A, B))} \circ$

由于求概念恢复集时已知 $\mathcal{F} \in \mathcal{P}(OC)$, 因此 $CNF_{\mathcal{F}}$ 一定是可满足的。要使得三支概念集 \mathcal{F} 的恢复约束 $CNF_{\mathcal{F}}$ 为真, 单子句是很重要的存在, 必须给其文字赋相应的值使得单子句满足。给单子句 x_{ij} 中的文字赋值后, 可化简公式 $CNF_{\mathcal{F}}$ 。

$CNF_{\mathcal{F}}$ 中的子句可分为以下 3 种情况。

第一种情况: 子句 C 中出现 $x_{ij}, x_{ij} = 1$ 时, 子句 C 一定满足, 对形式背景的二元关系将无约束效果, 可直接删除。

第二种情况: 子句 C 中出现 $\neg x_{ij}, x_{ij} = 0$ 对子句 C 的满足性不会有影响, 在 C 中可以删除 $\neg x_{ij}$ 。

第三种情况: 子句 C 中不出现 x_{ij} 或者 $\neg x_{ij}$, 这些子句在 $CNF_{\mathcal{F}}$ 中保持不变。

上述化简过程会导致新的单子句的产生, 其它的文字被赋值, 此过程即为单子句传播, 也称为布尔变元传播。如果当前 $CNF_{\mathcal{F}}$ 中存在单子句, 首先令 $P_{\mathcal{F}} = \{(o_i, a_j) \mid x_{ij} \text{ 是 } CNF_{\mathcal{F}} \text{ 中的单子句}\}$, 以及 $N_{\mathcal{F}} = \{(o_i, a_j) \mid \neg x_{ij} \text{ 是 } CNF_{\mathcal{F}} \text{ 中的单子句}\}$, 然后给单子句中的文字赋值使其满足 $CNF_{\mathcal{F}}$ 。如果化简后的 $CNF_{\mathcal{F}}$ 中还存在单子句, 则更新 $P_{\mathcal{F}}, N_{\mathcal{F}}$, 继续进行上述操作, 直到化简后的 $CNF_{\mathcal{F}}$ 中不存在单子句, 记化简后的 $CNF_{\mathcal{F}}$ 为 $SCNF_{\mathcal{F}}$ 。

定理 4 $P_{\mathcal{F}} = PI_{\mathcal{F}}, N_{\mathcal{F}} = NI_{\mathcal{F}}$ 。

证明 由推论 1 可知 $\mathcal{F} \in \mathcal{P}(C) \Rightarrow CNF_{\mathcal{F}}$ 为真, 所以 $P_{\mathcal{F}} \subseteq PI_{\mathcal{F}}, N_{\mathcal{F}} \subseteq NI_{\mathcal{F}}$ 。

若 $(o_i, a_j) \in PI_{\mathcal{F}}$, 由定义 1 可知当 $(o_i, a_j) \in I^c$ 时, $\mathcal{F} \notin \mathcal{P}(OC)$, 则 $CNF_{\mathcal{F}}$ 一定存在或可化简得到单子句 x_{ij} , 即 $(o_i, a_j) \in P_{\mathcal{F}}$ 。若 $(o_i, a_j) \in NI_{\mathcal{F}}$, 由定义 1 可知当 $(o_i, a_j) \in I$ 时, $\mathcal{F} \notin \mathcal{P}(OC)$, 则 $CNF_{\mathcal{F}}$ 一定存在或可化简得到单子句 $\neg x_{ij}$, 即 $(o_i, a_j) \in N_{\mathcal{F}}$, 所以 $P_{\mathcal{F}} = PI_{\mathcal{F}}, N_{\mathcal{F}} = NI_{\mathcal{F}}$ 。

定理 5 若 $(O, (A, B)) \in OC(OB, AT, I)$, 则 $SCNF_{\mathcal{F}} = C_3 \wedge C_4$ 。

定理 5 说明单子句传播时可以跳过产生此单子句的概念等价范式的子句, 在一定程度上可以减少算法复杂度。

根据定理 5, 算法 1 描述了概念集 F 的恢复集的算法框架。

算法 1 三支概念的恢复集 CNFS 算法

输入: 概念集 \mathcal{F} , 对象集 OB , 属性集 AT 。

输出: 正恢复集 $PI_{\mathcal{F}}$, 负恢复集 $NI_{\mathcal{F}}, SCNF_{\mathcal{F}}$ 。

1. $CNF_{\mathcal{F}} = \bigwedge_{(O, (A, B)) \in \mathcal{F}} CNF_{(O, (A, B))} \circ$

2. While $CNF_{\mathcal{F}}$ 中存在单子句 Do。

3. If $CNF_{\mathcal{F}}$ 中存在单子句 x_{ij} Then。

4. $PI_{\mathcal{F}}$ add (o_i, a_j) 并化简 $CNF_{\mathcal{F}}$ 。

5. If $CNF_{\mathcal{F}}$ 中存在单子句 $\neg x_{ij}$ Then。
6. $NI_{\mathcal{F}}$ add (o_i, a_j) 并化简 $CNF_{\mathcal{F}}$ 。
7. $SCNF_{\mathcal{F}} = CNF_{\mathcal{F}}$ 。
8. Return $PI_{\mathcal{F}}, NI_{\mathcal{F}}, SCNF_{\mathcal{F}}$ 。

利用上述算法,容易给出概念的恢复集。根据恢复约束的定义,最多可能有 $|OB \times AT|$ 个单子句出现,因此,该算法的时间复杂度为 $O(|OB \times AT|)$ 。

例 4 (续例 1) 在例 1 的形式背景中,已知 $\mathcal{F} = \{(o_3 o_7, (a_1 a_3, a_6 a_9)), (o_5 o_6 o_8, (a_1 a_3 a_6, a_5 a_7 a_8 a_9))\} \subseteq OC(OB, AT, I)$ 。恢复集求解过程如图 1 所示。

首先,可以根据给定的概念 $(o_3 o_7, (a_1 a_3, a_6 a_9))$ 、 $(o_5 o_6 o_8, (a_1 a_3 a_6, a_5 a_7 a_8 a_9))$ 分别得到对应的恢复约束,然后令单子句 x_{31} 满足,即 $x_{31} = 1, P_{\mathcal{F}} = \{(o_3, a_1)\}$ 。基于此,可以对 $\neg x_{31}$ 进行删除,且由定理 5 可知,只需要关注图中右框中的子句,因此图 1 右框第六行 $\neg x_{31}$ 被删除(图中用删除线表示)。令单子句 x_{56} 满足,即 $x_{56} = 1, P_{\mathcal{F}} = \{(o_3, a_1), (o_5, a_6)\}$,基于此可以对包含 x_{56} 的子句进行删除(图 1 左框第六行删除线)。类似地,对其他单子句进行此操作,直到不存在单子句,因此,根据定义 5 可知 $E(\mathcal{F}) = \frac{|R_{\mathcal{F}}|}{|OB \times AT|} = \frac{29}{72}$ 。

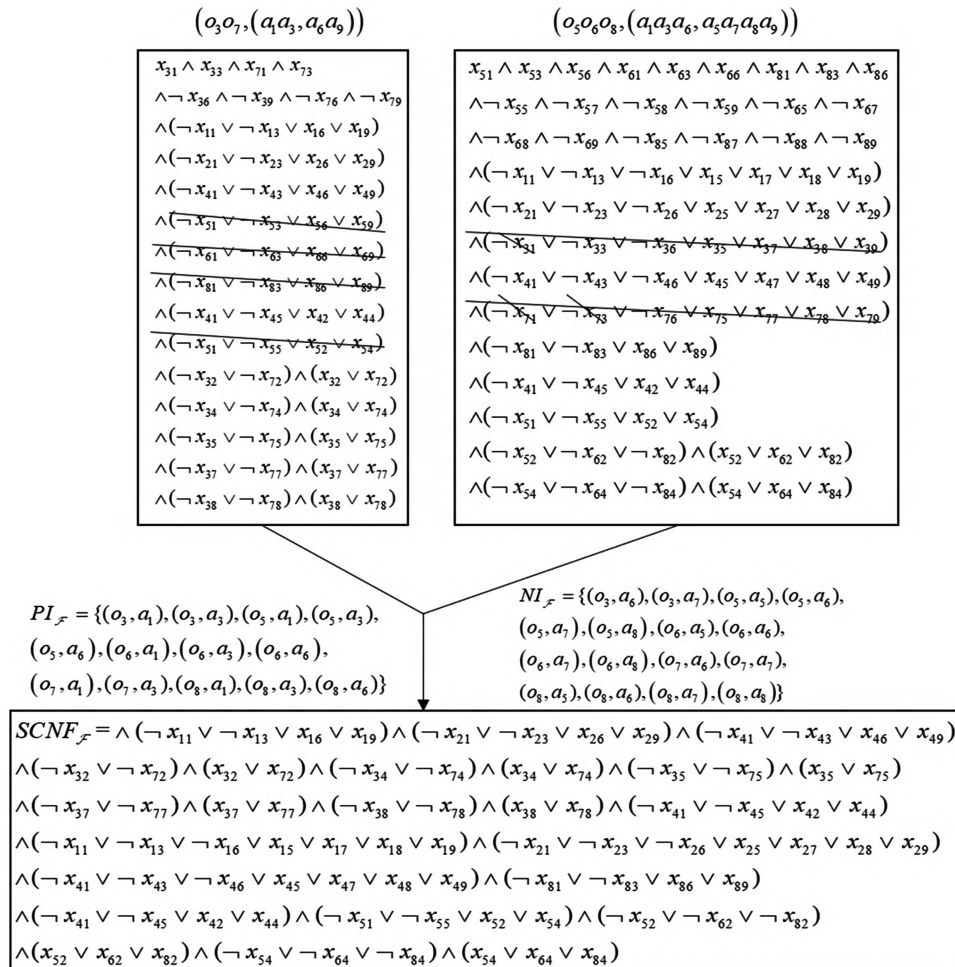


图 1 例 4 进行算法 1 的流程图
 Fig.1 Flowchart of example 4

2.3 三支概念集的动态算法

实际生活中,抽象得到或间接学习到的概念通常呈现出动态变化的趋势,在动态环境中,随着概念集中概念的增加,恢复集元素可能发生增加。除此之外,对于形式背景的局部缺失问题,也可以通过动态算法进行恢复。

设 \mathcal{F}_i 为 i 阶段的概念集,且 $(PI_{\mathcal{F}_{i-1}}, NI_{\mathcal{F}_{i-1}}, SCNF_{\mathcal{F}_{i-1}})$ 为第 $i-1$ 阶段的恢复状态,记 $\Delta \mathcal{F}_{i-1} = \mathcal{F}_i - \mathcal{F}_{i-1}$,则有如下性质成立。

性质 10 当 $PI_{\mathcal{F}_{i-1}} \subseteq I, NI_{\mathcal{F}_{i-1}} \subseteq I^c$ 时, $CNF_{\mathcal{F}_{i-1}}$ 为真当且仅当 $SCNF_{\mathcal{F}_{i-1}} \wedge CNF_{\Delta\mathcal{F}_{i-1}}$ 为真,因此根据性质 10 化简 $CNF_{\mathcal{F}_{i-1}}$ 并更新 $PI_{\mathcal{F}_{i-1}}, NI_{\mathcal{F}_{i-1}}$, 可得 $(PI_{\mathcal{F}_i}, NI_{\mathcal{F}_i}, SCNF_{\mathcal{F}_i})$ 。恢复集的动态更新算法如下。

算法 2 恢复集动态更新算法

输入: $(PI_{\mathcal{F}_{i-1}}, NI_{\mathcal{F}_{i-1}}, SCNF_{\mathcal{F}_{i-1}}), \Delta\mathcal{F}_{i-1}$ 。

输出: $(PI_{\mathcal{F}_i}, NI_{\mathcal{F}_i}, SCNF_{\mathcal{F}_i})$ 。

1. $CNF_{\Delta\mathcal{F}_{i-1}} = \bigwedge_{(O,(A,B)) \in \Delta\mathcal{F}_{i-1}} CNF_{(O,(A,B))}$ 。
2. 将 $PI_{\mathcal{F}_{i-1}}, NI_{\mathcal{F}_{i-1}}$ 中元素对应的命题变量分别赋值为 1 和 0, 化简公式 $CNF_{\Delta\mathcal{F}_{i-1}}$, 得到 $CNF1_{\Delta\mathcal{F}_{i-1}}$ 。
3. $CNF_{\mathcal{F}_i} = SCNF_{\mathcal{F}_{i-1}} \wedge CNF1_{\Delta\mathcal{F}_{i-1}}$ 。
4. While $CNF_{\mathcal{F}_i}$ 中存在单子句 Do。
5. If $CNF_{\mathcal{F}_i}$ 中存在单子句 x_{ij} Then。
6. $PI_{\mathcal{F}_i}$ add (o_i, a_j) 并化简 $CNF_{\mathcal{F}_i}$ 。
7. If $CNF_{\mathcal{F}_i}$ 中存在单子句 $\neg x_{ij}$ Then。
8. $NI_{\mathcal{F}_i}$ add (o_i, a_j) 并化简 $CNF_{\mathcal{F}_i}$ 。
9. $SCNF_{\mathcal{F}_i} = CNF_{\mathcal{F}_i}$ 。
10. Return $(PI_{\mathcal{F}_i}, NI_{\mathcal{F}_i}, SCNF_{\mathcal{F}_i})$ 。

由此可知,第 i 阶段的恢复状态可以通过 $(PI_{\mathcal{F}_{i-1}}, NI_{\mathcal{F}_{i-1}}, SCNF_{\mathcal{F}_{i-1}})$ 以及 $CNF_{\Delta\mathcal{F}_{i-1}}$ 得到,表明恢复系统可以对于新增三支概念信息通过更新而非重新恢复得到新的恢复集。利用上述算法,容易给出概念的恢复集,算法运行过程中最多可能有 $|OB \times AT| - |PI_{\mathcal{F}_{i-1}}| - |NI_{\mathcal{F}_{i-1}}|$ 个单子句出现,因此算法的时间复杂度为 $O(|OB \times AT| - |PI_{\mathcal{F}_{i-1}}| - |NI_{\mathcal{F}_{i-1}}|)$ 。

例 5 (续例 4) 在例 4 的基础上,已知 $\Delta\mathcal{F}_{i-1} = \{(o_1 o_2 o_3 o_4, (a_1 a_7, a_4 a_5 a_6))\} \subseteq OC(OB, AT, I)$, 则动态恢复的过程如图 2 所示。

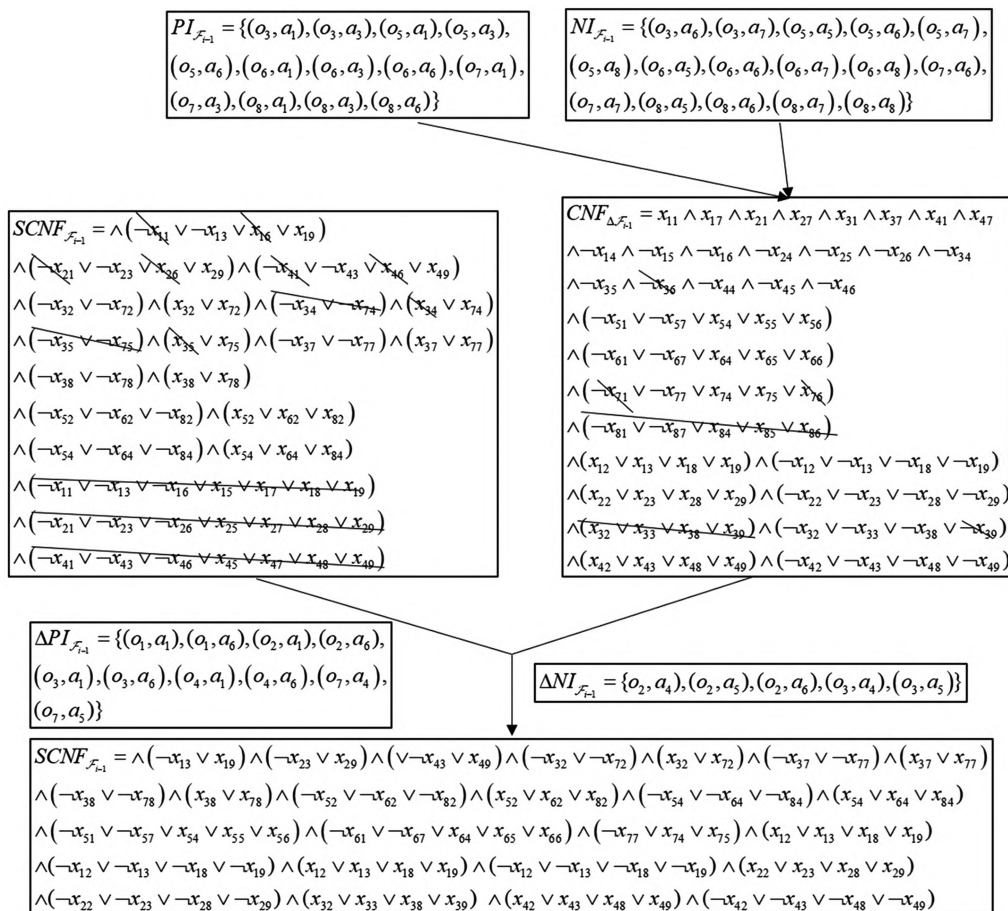


图 2 例 5 进行算法 2 的流程图
Fig.2 Flowchart of example 5

首先可以根据新增的三支概念 $(o_1o_2o_3o_4, (a_1a_7, a_4a_5a_6))$ 得到对应的恢复约束 $CNF_{\Delta\mathcal{F}_{i-1}}$, 然后对于集合 $PI_{\mathcal{F}_{i-1}}$ 中的元素 (o_3, a_3) , 令其对应的单子句 x_{33} 满足, 即 $x_{33} = 1$, 基于此可以对 $\neg x_{33}$ 进行删除。图 2 右框 $CNF_{\Delta\mathcal{F}_{i-1}}$ 第十行 $\neg x_{33}$ 被删除(图中用删除线表示)且对包含 x_{33} 的子句进行删除。图 2 右框 $CNF_{\Delta\mathcal{F}_{i-1}}$ 第十行 $x_{32} \vee x_{33} \vee x_{38} \vee x_{39}$ 被删除(图中用删除线表示), 类似地对 $PI_{\mathcal{F}_{i-1}}, NI_{\mathcal{F}_{i-1}}$ 中元素进行此操作。

进一步, 令单子句 x_{11} 满足, 即 $x_{11} = 1, P_{\mathcal{F}_{i-1}} = \{(o_3, a_1)\}$, 类似的对其他单子句进行此操作, 直到不存在单子句。

3 实验结果与分析

本节通过数值实验评估给出形式背景恢复的效果。用生物与水数据集^[35]和 UCI 机器学习数据库选出的 3 个不同数据集^[40]进行测试, 详细数据描述如表 3 所示。由于原数据不能直接用作形式背景, 因此先对这些数据集进行数据预处理, 使之转化为形式背景, 具体预处理方法: 西瓜数据 3.0 中将原有属性改为名义属性, 进一步转化为形式; Acute inflammations 数据集的第十三个属性 legs 大于等于 1 的属性均转化为 1, 第十七个属性小于等于 4 的属性值转化为 0 而其余值转化为 1。将数据 Shuttle_landing_control 中定量属性进行离散化操作, 生成新的属性描述, 构建一个形式背景, 对象集合是飞行器的 15 次安全飞行记录, $OB = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15\}$ 。属性集为各类因素, 包含飞行器状态是否稳定、故障类型、出现的迹象、视野是否可见、振动程度等多个因素, $AT = \{S_1, S_2, E_1, E_2, E_3, E_4, SI_1, SI_2, W_1, W_2, M_1, M_2, M_3, M_4, V_1, V_2\}$ 。

表 3 实验数据集描述
Table 3 Description of experimental data sets

编号	数据集	对象数	属性数	概念数
1	生物与水 ^[35]	8	9	43
2	西瓜数据集 3.0 ^[40]	14	7	878
3	Acuteinflammations ^[40]	120	6	48
4	Shuttle_landing_control ^[40]	15	16	2 752

由于目前没有相关可以比较的恢复算法, 因此, 为了测试算法的有效性, 同时为了进行对比分析, 本文在生成的所有概念集中随机选取 5%、10%、15%、20%、30%、50% 的概念进行恢复, 重复 50 次, 并计算平均恢复度及最大、最小恢复度, 见表 4—7。

表 4 生物与水形式背景概念数与恢复度
Table 4 Concept number and degree of recovery of living beings and water

概念个数 比例/%	概念 个数	平均 恢复度/%	最大 恢复度/%	最小 恢复度/%
5	2	32.14	47.22	19.44
10	4	56.53	77.78	37.50
15	6	73.28	93.06	50.00
20	9	90.83	100.00	70.83
30	13	96.56	100.00	87.50
50	21	99.58	100.00	94.44

表 5 西瓜数据集 3.0 概念数与恢复度
Table 5 Concept number and degree of recovery of watermelon data set

概念个数 比例/%	概念 个数	平均 恢复度/%	最大 恢复度/%	最小 恢复度/%
5	44	98.27	100.00	95.75
10	88	99.94	100.00	99.35
15	132	99.99	100.00	99.67
20	176	100.00	100.00	100.00
30	236	100.00	100.00	100.00
50	439	100.00	100.00	100.00

表 6 Acute inflammations 概念数与恢复度
Table 6 Concept number and degree of recovery of Acute inflammations

概念个数 比例/%	概念 个数	平均 恢复度/%	最大 恢复度/%	最小 恢复度/%
5	2	26.64	38.89	8.33
10	5	55.89	73.37	32.78
15	7	70.37	81.53	55.14
20	9	79.99	94.44	62.08
30	14	93.62	100.00	81.94
50	24	99.55	100.00	94.31

表 7 Shuttle_landing_control 概念数与恢复度
Table 7 Concept number and degree of recovery of Shuttle_landing_control

概念个数 比例/%	概念 个数	平均 恢复度/%	最大 恢复度/%	最小 恢复度/%
5	138	99.92	100.00	99.17
10	275	100.00	100.00	100.00
15	413	100.00	100.00	100.00
20	550	100.00	100.00	100.00
30	826	100.00	100.00	100.00
50	1 376	100.00	100.00	100.00

从表4—6中可以看出形式背景恢复的有效性,当概念比例达到50%时,恢复度接近100%;当生成的概念数比较多时,从表7中可以看出,概念数达5%时,平均恢复就可以达到99.92%,所以,事实上并不需要认识所有的概念就能对对象的具体属性有充分的了解,从表4—6中也能看出所含概念个数不同的概念集的恢复度差别很大,也说明在我们对事物具体认识时不同概念的重要程度是有所区别的。

4 结论

本文提出了恢复集理论,这一理论不仅扩展了三支概念的应用范围,还有效地丰富了其理论基础。为了解决三支概念的恢复集问题,提出了CNFS算法,该算法通过分析三支概念的定义,推导出三支概念对形式背景中对象与属性关系的等价约束条件,然后通过简化概念集的合取范式求解恢复集。此外,我们还进一步研究了恢复集的动态算法。实验结果表明,恢复集算法的有效性得到了验证。尽管本文的工作取得了一定进展,但仍需进一步研究一些关键问题,比如在寻求三支概念的恢复集时,依然存在概念冗余的问题。这种现象主要是由于已知概念集不一定能够实现概念的约简而造成的,因此在未来的研究中,将进一步考虑恢复过程中出现的概念冗余问题,以提升恢复集算法的准确性和应用效果。

参考文献:

- [1] WILLE R. Formal concept analysis[J]. *Electronic Notes in Discrete Mathematics*, 1999, 2(1):199-200.
- [2] QI J J, WEI L, YAO Y Y. Three-way formal concept analysis[C] // *Proceedings of the 9th International Conference on Rough Sets and Knowledge Technology*. Shanghai: Springer, 2014:732-741.
- [3] YAO Y Y. Three-way decision: an interpretation of rules in rough set theory[C] // *Proceedings of the 4th International Conference on Rough Sets and Knowledge Technology*. Gold Coast: Springer, 2009:14-16.
- [4] ZHAN J, YE J, DING W, et al. A novel three-way decision model based on utility theory in incomplete fuzzy decision systems[J]. *IEEE Transactions on Fuzzy Systems*, 2022, 30(7):2210-2226.
- [5] HAO F, GAO J, BISOGNI C, et al. Stability of three-way concepts and its application to natural language generation[J]. *Pattern Recognition*, 2021, 149:51-58.
- [6] HAO F, YANG Y, MIN G, et al. Incremental construction of three-way concept lattice for knowledge discovery in social networks[J]. *Information Science*, 2021, 578:257-280.
- [7] WEI L, LIU L, QI J, et al. Rules acquisition of formal decision contexts based on three-way concept lattices[J]. *Information Science*, 2020, 516: 529-544.
- [8] 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*, 2017, 51(3): 116-121.
QI Jianjun, WANG Wenwei. A multithreaded parallel algorithm for constructing three-way concepts[J]. *Journal of Xi'an Jiaotong University*, 2017, 51(3):116-121.
- [9] 钱婷. 经典概念格与三支概念格的构造及知识获取理论[J]. 西安:西北大学, 2016.
QIAN Ting. The constructions of classical concept lattices and three-way concept lattices and knowledge acquisition[J]. Xi'an: Northwest University, 2016.
- [10] QIAN T, WEI L, QI J. Constructing three-way concept lattices based on apposition and subposition of formal contexts[J]. *Knowledge-Based Systems*, 2017, 116:39-48.
- [11] 王明,魏玲. 基于K-Modes聚类的OE-概念格压缩[J]. *模式识别与人工智能*, 2018(8):704-714.
WANG Ming, WEI Ling. OE-Concept lattice compression based on K-Modes clustering[J]. *Pattern Recognition and Artificial Intelligence*, 2018(8):704-714.
- [12] HAO F, YANG Y, MIN G, et al. Incremental construction of three-way concept lattice for knowledge discovery in social networks[J]. *Information Sciences*, 2021, 578:257-280.
- [13] LI J H, HUANG C C, QI J J, et al. Three-way cognitive concept learning via multi-granularity[J]. *Information Sciences*, 2017, 378:244-263.
- [14] HUANG C, LI J, MEI C, et al. Three-way concept learning based on cognitive operators: an information fusion viewpoint[J]. *International Journal of Approximate Reasoning*, 2017, 84:1-20.
- [15] YUAN K H, XU W H, LI W T, et al. An incremental learning mechanism for object classification based on progressive fuzzy three-way concept[J]. *Information Sciences*, 2022, 584:127-147.
- [16] HE X, WEI L, SHE Y. L-fuzzy concept analysis for three-way decisions: basic definitions and fuzzy inference mechanisms[J]. *International Journal of Machine Learning and Cybernetics*, 2018, 9(11):1857-1867.

- [17] LIU X, MAO W, DAI J, et al. Comprehensive fuzzy concept-oriented three-way decision and its application[J]. *Information Sciences*, 2022, 593:233-270.
- [18] SUBHASHINI L, LI Y, ZHANG J, et al. Integration of semantic patterns and fuzzy concepts to reduce the boundary region in three-way decision-making[J]. *Information Sciences*, 2022, 595:257-277.
- [19] SINGH P. Three-way fuzzy concept lattice representation using neutrosophicset[J]. *International Journal of Machine Learning and Cybernetics*, 2016, 8(1):1-11.
- [20] YAO Y Y. Interval sets and three-way concept analysis in incomplete contexts[J]. *International Journal of Machine Learning and Cybernetics*, 2017, 8(1):3-20.
- [21] LI J H, MEI C L. Incomplete decision contexts: approximate concept construction, rule acquisition and knowledge reduction [J]. *International Journal of Approximate Reasoning*, 2013, 54(1):149-165.
- [22] LI M, WANG G. Approximate concept construction with three-way decisions and attribute reduction in incomplete contexts [J]. *Knowledge-Based Systems*, 2016, 91:165-178.
- [23] ZHI H, CHAO H. Three-way concept analysis for incomplete formal contexts[J]. *Mathematical Problems in Engineering*, 2018, 2018:1-11.
- [24] YAO Y Y. Three-way conflict analysis: reformulations and extensions of the Pawlakmodel[J]. *Knowledge-Based Systems*, 2019, 180:26-37.
- [25] FAN Y, QI J J, WEI L. A conflict analysis model based on three-way decisions[C]// *Proceedings of International Joint Conference. QuyNhon:Springer*, 2018:522-532.
- [26] SINGH P K. Medical diagnoses using three-way fuzzy concept lattice and their Euclidean distance[J]. *Computational and Applied Mathematics*, 2018, 37:3283-3306.
- [27] SUBRAMANIAN C M, CHERUKURI A K, CHELLIAH C. Role based access control design using three-way formal concept analysis[J]. *International Journal of Machine Learning and Cybernetics*, 2018, 9(11):1-31.
- [28] LONG B H, XU W H, ZHANG X Y, et al. The dynamic update method of attribute-induced three-way granular concept in formal contexts[J]. *International Journal of Approximate Reasoning*, 2020, 126:228-248.
- [29] NIU J, HUANG C, LI J, et al. Parallel computing techniques for concept-cognitive learning based on granular computing[J]. *International Journal of Machine Learning and Cybernetics*, 2018, 9(11):1785-1805.
- [30] SHI Y, MI Y L, LI J H, et al. Concept-cognitive learning model for incremental concept learning[J]. *IEEE Transactions on Systems, Man, and Cybernetics*, 2021, 51(2):809-821
- [31] SHI Y, MI Y L, LI J H, et al. Concurrent concept-cognitive learning model for classification[J]. *Information Sciences*, 2019, 496:65-81.
- [32] ZHANG T, LI H, LIU M, et al. Incremental concept-cognitive learning based on attribute topology[J]. *International Journal of Approximate Reasoning*, 2020, 118:173-189.
- [33] YUAN K H, XU W H, LI W T, et al. An incremental learning mechanism for object classification based on progressive fuzzy three-way concept[J]. *Information Sciences*, 2022, 584:127-147.
- [34] QI J J, QIAN T, WEI L. The connections between three-way and classical concept lattices[J]. *Knowledge-Based Systems*, 2016, 91:143-151.
- [35] GANTER B, WILLE R. *Formal concept analysis: mathematical foundations*[M]. New York: Springer, 1999.
- [36] 郭莹, 张长胜, 张斌. 求解 SAT 问题的算法的研究进展[J]. *计算机科学*, 2016, 43(3):8-17.
GUO Ying, ZHANG Changsheng, ZHANG Bin. Research advance of SAT solving algorithm[J]. *Computer Science*, 2016, 43(3):8-17.
- [37] 张文修, 梁怡, 徐萍. 基于包含度的不确定推理[M]. 北京:清华大学出版社, 2007.
ZHANG Wenxiu, LIANG Yi, XU Ping. *Uncertainty reasoning based on inclusion degree*[M]. Beijing: Tsinghua University Press, 2007.
- [38] 魏玲, 曹丽, 祁建军, 等. 形式概念分析中的概念约简与概念特征[J]. *中国科学(信息科学)*, 2020, 50:1817-1833.
WEI Ling, CAO Li, QI Jianjun, et al. Concept reduction and concept characteristics in formal concept analysis[J]. *Science China(Information Sciences)*, 2020, 50:1817-1833.
- [39] 曹丽, 魏玲, 祁建军. 保持二元关系不变的概念约简[J]. *模式识别与人工智能*, 2018, 31(6):516-524.
CAO Li, WEI Ling, QI Jianjun. Concept reduction preserving binary relations[J]. *Pattern Recognition and Artificial Intelligence*, 2018, 31(6):516-524.
- [40] 周志华. *机器学习*[M]. 北京:清华大学出版社, 2016.
ZHOU Zhihua. *Machine learning*[M]. Beijing: Tsinghua University Press, 2016.

(编辑:于善清)