中国科学: 信息科学 2020 年 第 50 卷 第 12 期: 1817–1833
SCIENTIA SINICA Informationis
《中国科学》杂志社
SCIENCE CHINA PRESS

论文

形式概念分析中的概念约简与概念特征

魏玲[1,4*], 曹丽[1,4], 祁建军[2,4], 张文修[3]

1. 西北大学数学学院, 西安 710127
2. 西安电子科技大学计算机科学与技术学院, 西安 710071
3. 西安交通大学数学与统计学院, 西安 710049
4. 西北大学概念认知与智能研究中心, 西安 710127
* 通信作者. E-mail: wl@nwu.edu.cn

收稿日期: 2018–10–14; 修回日期: 2019–03–26; 接受日期: 2019–08–09; 网络出版日期: 2020–11–18

国家自然科学基金 (批准号: 61772021, 11371014) 资助项目

摘要 形式概念分析是以形式背景及其概念格为基础的一种数据分析方法, 其中的形式概念明确反映了数据信息中对象与属性间的关系, 是哲学中 "概念" 这一名词的形式化描述. 文章提出在形式概念分析框架下进行概念约简的思想, 研究保持形式背景中二元关系不变的概念约简的相关理论; 针对概念约简理论中作用不同的 3 种概念类型, 分别从算子角度以及布尔矩阵角度分析了对象 (属性) 概念的特征, 并给出了求解概念约简的方法.

关键词 形式背景, 形式概念分析, 二元关系, 约简, 特征

1 引言

1982 年, 德国数学家 Wille 为重建格理论提出了形式概念分析 (formal concept analysis, FCA) [1,2]. 自 20 世纪 90 年代至今, 该理论得到充分发展, 已经成为一种进行数据分析和规则获取的有力工具, 应用到诸如信息检索、语义 web、知识工程等很多领域, 成为了一个新的有影响力的研究方向.

形式概念分析的数据基础是形式背景, 即一个由对象集、属性集, 以及二者间二元关系形成的三元组. 以形式背景为基础, Wille 定义了对象集与属性集之间的两个诱导算子, 并由此生成形式概念, 实现了对哲学中 "概念" 这一名词的形式化刻画. 进而, 所有的概念按照一定的偏序关系和上下确界的定义, 形成一个格结构, 称之为概念格. 概念格是形式概念分析理论最基础也是最核心的数据结构, 形式概念分析则是在形式背景和概念层次分析的框架下来研究格的结构与表示理论.

因为概念是人类思维的基本单元, 因此概念的形式化也为人类思维、智能的形式化打开了一个小小的窗口. 这使得形式概念分析与人工智能、知识发现等的结合研究逐渐增多.

引用格式: 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. 中国科学: 信息科学, 2020, 50: 1817–1833, doi: 10.1360/N112018-00272
Wei L, Cao L, Qi J J, et al. Concept reduction and concept characteristics in formal concept analysis (in Chinese). Sci Sin Inform, 2020, 50: 1817–1833, doi: 10.1360/N112018-00272

近年来, 形式概念分析在自身框架下的研究成果颇丰, 比如: 概念格构建理论与方法 [3~7]、属性约简 [8~14], 以及规则获取 [15~18] 等. 并且, 形式概念分析与其他理论间的关系研究与结合研究 [3, 19~39], 也产生了丰富的研究成果以及很多有意义的新的研究分支.

属性约简是近年来形式概念分析研究领域中的一个热点问题, 作者于 2005 年提出的概念格约简是国内最早的属性约简理论. 概念格约简是在保持形式背景中对象集不变的前提下, 寻找极小属性子集, 使得由该极小属性子集确定的概念格与由属性集全体确定的概念格同构 [8, 9]. 该理论涉及属性协调集与属性约简的基本定义以及判定定理、利用差别矩阵求解属性约简的方法、属性分类及不同类型属性特征分析、利用属性特征构造属性约简的方法等较为完整的一系列结果.

继概念格约简理论提出之后, Wu 等 [12] 于 2009 年提出了粒约简的概念; Wang 等 [10] 研究了保持交不可约元外延集不变的属性约简问题; Li 等 [11] 研究了保持并不可约元外延集不变的属性约简问题. 他们的研究思路与方法都沿用了文献 [8] 的分析过程. 在近年来的研究过程中, 对属性约简这一研究主题也逐渐形成了一个较为普遍的认识: 所谓属性约简, 就是寻找能够保持形式背景或者概念格的某种特性不变的极小属性子集的过程; 而获取到的极小属性子集则被称为约简. 我们还将属性约简的思想推广至决策形式背景, 讨论了决策形式背景的条件属性约简问题 [9, 13] 以及基于 AE – 概念格的决策形式背景属性约简 [29]. Shao 等 [14] 讨论了广义单边形式背景上的属性约简, Ren 等 [40] 在三支概念分析中讨论了 4 种属性约简以及这 4 种属性约简之间的关系, Wang 等 [41] 在不完备背景中讨论了基于一类区间集概念格的属性约简.

将形式背景视为一个布尔 (Boolean) 矩阵, Keprt 和 Snasel [42] 于 2004 年提出用形式概念作为因子的方法, 用以解决布尔因子分析问题. 在此基础上, Belohlavek, Vychodil 及 Trnecka 等做了许多工作, 并得到了很好的成果, 这些工作对我们有很大的启发 [43~46].

结合 FCA 中的属性约简以及因子分析的思想, 我们提出了保持二元关系不变的概念约简这一研究主题 [47], 从概念角度出发, 设法减少概念格中概念的个数, 却能够保留全部原始信息, 即保留下的概念可以使得原始数据集中的二元关系完全不变. 本文在回顾文献 [47] 工作的基础上, 主要从算子角度和布尔矩阵角度对对象概念与属性概念的 3 类概念特征进行了研究, 并探讨获得了保持二元关系不变的概念约简的计算方法.

2 基础知识

2.1 形式概念分析基础知识

本小节回顾形式概念分析的基本定义.

定义1 ([1]) 设 (G, M, I) 是形式背景, 其中 $G = \{g_1, \dots, g_p\}$ 为对象集, 每一个 g_i $(i \leqslant p)$ 称为一个对象; $M = \{m_1, \dots, m_q\}$ 为属性集, 每个 m_j $(j \leqslant q)$ 称为一个属性; I 为 G 与 M 之间的二元关系, $I \subseteq G \times M$. 若 $(g, m) \in I$, 则表示对象 g 拥有属性 m, 记为 gIm.

本文为了将形式背景与布尔矩阵对应, 将 $(g, m) \in I$ 记为 1, $(g, m) \notin I$ 记为 0.

对任意的 $X \subseteq G$, $B \subseteq M$, Wille [1] 给出了两者之间的一对诱导算子, 如下所示:

$$X^* = \{m \mid m \in M, \forall g \in X, gIm\}, \quad B^* = \{g \mid g \in G, \forall m \in B, gIm\},$$

X^* 表示 X 中所有对象所共同具有的属性集合, B^* 表示拥有 B 中所有属性的对象集合.

这对诱导算子的具体性质见文献 [2].

表 1　生物与水形式背景 (G, M, I)

Table 1　The context of living beings and water (G, M, I)

G	a	b	c	d	e	f	g	h	i
1	1	1	0	0	0	0	1	0	0
2	1	1	0	0	0	0	1	1	0
3	1	1	1	0	0	0	1	1	0
4	1	0	1	0	0	0	1	1	1
5	1	1	0	1	0	1	0	0	0
6	1	1	1	1	0	1	0	0	0
7	1	0	1	1	1	0	0	0	0
8	1	0	1	1	0	1	0	0	0

表 2　例 1 形式背景 (G, M, I) 对应的概念

Table 2　The concept lattice corresponding to the formal context in example 1

Label	Concept	Label	Concept	Label	Concept	Label	Concept	Label	Concept
c_1	$(\emptyset, M)$	c_5	$(6, abcdf)$	c_9	$(36, abc)$	c_{13}	$(234, agh)$	c_{17}	$(12356, ab)$
c_2	$(4, acghi)$	c_6	$(34, acgh)$	c_{10}	$(678, acd)$	c_{14}	$(568, adf)$	c_{18}	$(5678, ad)$
c_3	$(3, abcgh)$	c_7	$(23, abgh)$	c_{11}	$(68, acdf)$	c_{15}	$(1234, ag)$	c_{19}	(G, a)
c_4	$(7, acde)$	c_8	$(123, abg)$	c_{12}	$(56, abdf)$	c_{16}	$(34678, ac)$		

如果一个二元组 (X, B) $(X \subseteq G, B \subseteq M)$ 满足 $X^* = B$, $B^* = X$, 则称 (X, B) 是一个形式概念, 简称概念. 其中, X 称为概念的外延, B 称为概念的内涵.

在概念中, 有两类特殊的概念: 由单个对象生成的对象概念 (g^{**}, g^*) $(g \in G)$ 和由单个属性生成的属性概念 (m^*, m^{**}) $(m \in M)$. 我们记对象概念全体为 $\mathcal{O}(G, M, I) = \{(g^{**}, g^*) \mid g \in G\}$, 属性概念全体为 $\mathcal{A}(G, M, I) = \{(m^*, m^{**}) \mid m \in M\}$ (为方便起见, 本文将 $\{g\}^*$ 记为 g^*, 将 $\{m\}^*$ 记为 m^*).

用 $L(G, M, I)$ 表示形式背景 (G, M, I) 的全体概念, 若 (X_1, B_1) 和 (X_2, B_2) 是概念, 其上的偏序关系定义为 $(X_1, B_1) \leqslant (X_2, B_2) \Leftrightarrow X_1 \subseteq X_2 (B_1 \supseteq B_2)$. 定义下确界为 $(X_1, B_1) \wedge (X_2, B_2) = (X_1 \cap X_2, (B_1 \cup B_2)^{**})$, 上确界为 $(X_1, B_1) \vee (X_2, B_2) = ((X_1 \cup X_2)^{**}, B_1 \cap B_2)$, 二者也是概念, 从而 $L(G, M, I)$ 是完备格, 称为 (G, M, I) 的概念格.

文献 [2] 给出了以下引理, 说明一个形式背景可以由它所产生的所有概念重新组建.

引理1　设 (G, M, I) 是形式背景, $L(G, M, I)$ 是概念格, 则 $I = \bigcup\{A \times B \mid (A, B) \in L(G, M, I)\}$.

例1 ([2])　表 1 是关于 "生物与水" 的形式背景 (G, M, I), 其中, 对象集 G 包含 8 个对象, 它们分别是: 1: leech, 2: bream, 3: frog, 4: dog, 5: spike-weed, 6: reed, 7: bean, 8: maize; 属性集 M 包含 9 个属性, 它们分别是: a: needs water to live, b: lives in water, c: lives no land, d: needs chlorophyll to produce foods, e: two seed leaves, f: one seed leaf, g: can move around, h: has limbs, i: suckles its offspring. 表 1 中的形式背景所对应的概念如表 2 所示, 对应的概念格如图 1 所示 (注: 按 FCA 理论的习惯表示, 除全集与空集外, 其他集合用其元素序列表示).

2.2　布尔矩阵与布尔向量基础知识

布尔代数 [48] 由集合 $B_0 = \{0, 1\}$ 及定义在其上的 3 种运算 "+", "·" 和 "c" 构成, 运算规则如下:

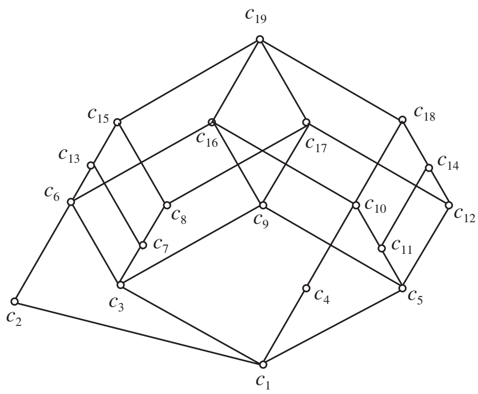

图 1　例 1 形式背景对应的概念格

Figure 1　The concept lattice corresponding to example 1

(1) $0+0=0\cdot 0=1\cdot 0=0\cdot 1=0$,

(2) $1+0=0+1=1+1=1\cdot 1=1$,

(3) $0^c=1, 1^c=0$.

称定义在 B_0 上的向量为布尔向量, 记所有的 n 维布尔行向量为 $\boldsymbol{V}_n$, 所有的 n 维布尔列向量为 $\boldsymbol{V}^n$, 所有的 $m\times n$ 阶布尔矩阵为 $\boldsymbol{B}_{m\times n}$.

如果 $\boldsymbol{\alpha}_1=(x_1,x_2,\dots,x_n)$, $\boldsymbol{\alpha}_2=(y_1,y_2,\dots,y_n)$, 则定义布尔行向量上的两个运算 "$+$" 与 "$\cdot$" 为: (1) $\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2=(x_1+y_1,x_2+y_2,\dots,x_n+y_n)$, (2) $\boldsymbol{\alpha}_1\cdot\boldsymbol{\alpha}_2=(x_1\cdot y_1,x_2\cdot y_2,\dots,x_n\cdot y_n)$.

对于两个行向量 $\boldsymbol{\alpha}_1$ 和 $\boldsymbol{\alpha}_2$, 若 $x_i=1\Rightarrow y_i=1$, 则记 $\boldsymbol{\alpha}_1\leqslant\boldsymbol{\alpha}_2$. 如果 $\boldsymbol{\alpha}_1\leqslant\boldsymbol{\alpha}_2$ 但 $\boldsymbol{\alpha}_1\neq\boldsymbol{\alpha}_2$, 则记 $\boldsymbol{\alpha}_1<\boldsymbol{\alpha}_2$. 列向量相关内容可类似定义 [48].

定义2 ([48])　设子空间 $W\subset\boldsymbol{V}_n$, W 的一个子集合 B 称为 W 的一个基当且仅当 $W=\langle B\rangle$ (向量集合 B 的生成空间), 而且 B 是一个无关的集合.

设 $A\in\boldsymbol{B}_{m\times n}$. $B_R(A)$ 表示由 A 的所有行构成的集合生成的空间 $R(A)$ 的唯一基底, 称为 A 的行基底, 其基数称为 A 的行秩, 用 $\rho_R(A)$ 表示. 类似地, $B_C(A)$ 表示由 A 的所有列构成的集合生成的空间 $C(A)$ 的唯一基底, 称为 A 的列基底, 其基数称为 A 的列秩, 用 $\rho_C(A)$ 表示 [48].

设 $\boldsymbol{A}_{n\times k}$, $\boldsymbol{B}_{k\times m}$ 为布尔矩阵, 若定义 $\boldsymbol{A}\circ\boldsymbol{B}=\bigvee_{l=1}^{k}A_{il}\cdot B_{lj}$, 其中 $\bigvee$ 表示最大值 (逻辑或的真值函数, 下文均用 $+$ 表示), $\cdot$ 表示数量乘积 (逻辑与的真值函数), 则称 $\boldsymbol{A}\circ\boldsymbol{B}$ 是布尔矩阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 的乘积 [48].

2.3　形式概念作为因子处理布尔因子分析问题

布尔因子分析要解决的问题是: 对一个布尔矩阵 $\boldsymbol{W}_{n\times m}$, 通过寻找 k ($k<m$) 个隐藏的因子 (k 应该尽可能小), 使得 $\boldsymbol{W}$ 可表示为两个布尔矩阵 $\boldsymbol{A}_{n\times k}$, $\boldsymbol{B}_{k\times m}$ 的布尔乘积 [42, 43].

Glodeanu [49] 于 2010 年用形式概念分析的语言给出了因子的定义.

定义3 ([49])　设 (G,M,I) 是一个形式背景, $L(G,M,I)$ 是概念格, $\mathcal{F}\subseteq L(G,M,I)$. 若 $I=\bigcup_{(A,B)\in\mathcal{F}}(A\times B)$, 则称 $\mathcal{F}$ 是基于经典概念格的因子分解. 进一步, 若 $|\mathcal{F}|$ 最小, 则称 $\mathcal{F}$ 为理想因子分解. $\mathcal{F}$ 中的元素称作 (理想) 因子.

Belohlavek 和 Vychodil [43] 给出了如下的用概念构造两个布尔矩阵的具体方法.

对于形式背景 (G,M,I), 记对象集 $G=\{x_1,x_2,\ldots,x_n\}$ 中每一个对象 x_i $(i\leqslant n)$ 为 i, 属性集 $M=\{a_1,a_2,\ldots,a_m\}$ 中每一个属性 a_j $(j\leqslant m)$ 为 j. 若 $\mathcal{F}=\{(A_1,B_1),\ldots,(A_t,B_t)\}\subseteq L(G,M,I)$, 构造布尔矩阵 $\boldsymbol{A}_{\mathcal{F}}$ 和 $\boldsymbol{B}_{\mathcal{F}}$ 如下:

$$(\boldsymbol{A}_{\mathcal{F}})_{il}=\begin{cases}1, & i\in A_l,\\ 0, & i\notin A_l,\end{cases}\qquad (\boldsymbol{B}_{\mathcal{F}})_{lj}=\begin{cases}1, & j\in B_l,\\ 0, & j\notin B_l.\end{cases}$$

则存在 $\mathcal{F}$, 使得形式背景 (G,M,I) 所对应的布尔矩阵 $\boldsymbol{W}$ 可分解为如上方法所构造的两个布尔矩阵乘积的形式 $\boldsymbol{A}_{\mathcal{F}}\circ\boldsymbol{B}_{\mathcal{F}}$.

定理1 ([43]) 设 (G,M,I) 是一个形式背景, $L(G,M,I)$ 是概念格, 如果存在 $\mathcal{F}\subseteq L(G,M,I)$ 使得 $\boldsymbol{W}=\boldsymbol{A}_{\mathcal{F}}\circ\boldsymbol{B}_{\mathcal{F}}$, 则 $\mathcal{O}(G,M,I)\cap\mathcal{A}(G,M,I)\subseteq\mathcal{F}$.

文献 [43] 称满足定理 1 的形式概念为强制性因子.

例2 例 1 中形式背景 (G,M,I) 的概念标号如表 2 所示, $\boldsymbol{W}$ 是该形式背景对应的布尔矩阵. 令 $\mathcal{F}=\{c_2,c_3,c_4,c_7,c_8,c_{11},c_{12}\}$, 则按照上述方式构造的矩阵 $\boldsymbol{A}_{\mathcal{F}}$ 与 $\boldsymbol{B}_{\mathcal{F}}$ 如下:

$$\boldsymbol{W}=\begin{pmatrix}1&1&0&0&0&0&1&0&0\\1&1&0&0&0&0&1&1&0\\1&1&1&0&0&0&1&1&0\\1&0&1&0&0&0&1&1&1\\1&1&0&1&0&1&0&0&0\\1&1&1&1&0&1&0&0&0\\1&0&1&1&1&0&0&0&0\\1&0&1&1&0&1&0&0&0\end{pmatrix},\quad \boldsymbol{A}_{\mathcal{F}}=\begin{pmatrix}1&0&0&0&0&0&0\\1&1&0&0&0&0&0\\1&1&1&0&0&0&0\\0&0&0&1&0&0&0\\0&0&0&0&1&0&0\\0&0&0&0&1&0&1\\0&0&0&0&0&1&0\\0&0&0&0&0&0&1\end{pmatrix},\quad \boldsymbol{B}_{\mathcal{F}}=\begin{pmatrix}1&1&0&0&0&0&1&0&0\\1&1&0&0&0&0&1&1&0\\1&1&1&0&0&0&1&1&0\\1&0&1&0&0&0&1&1&1\\1&1&0&1&0&1&0&0&0\\1&0&1&1&1&0&0&0&0\\1&0&1&1&0&1&0&0&0\end{pmatrix}.$$

因为 $\boldsymbol{W}=\boldsymbol{A}_{\mathcal{F}}\circ\boldsymbol{B}_{\mathcal{F}}$, 则 $\mathcal{F}$ 是因子. 而此概念格中, $\mathcal{O}(G,M,I)=\{c_2,c_3,c_4,c_5,c_7,c_8,c_{11},c_{12}\}$, $\mathcal{A}(G,M,I)=\{c_2,c_4,c_{13},c_{14},c_{15},c_{16},c_{17},c_{18},c_{19}\}$, 易见 $\mathcal{O}(G,M,I)\cap\mathcal{A}(G,M,I)=\{c_2,c_4\}\subseteq\mathcal{F}$.

3 概念约简理论与概念特征

在文献 [47] 中, 我们提出了保持二元关系不变的概念约简问题, 给出了该问题的相关研究成果. 本节在回顾保持二元关系不变的概念约简基本理论的基础上, 从二元关系、算子, 以及布尔矩阵角度来分析获得概念特征的不同刻画.

定义4 ([47]) 设 (G,M,I) 是一个形式背景, $L(G,M,I)$ 是概念格, $\mathcal{F}\subseteq L(G,M,I)$. 如果 $I=\bigcup_{(A,B)\in\mathcal{F}}(A\times B)$, 则称 $\mathcal{F}$ 为保持二元关系不变的概念协调集; 若对任意的 $(A_i,B_i)\in\mathcal{F}$, $\mathcal{F}'=\mathcal{F}\setminus(A_i,B_i)$, 有 $I\neq\bigcup_{(A,B)\in\mathcal{F}'}(A\times B)$, 则称 $\mathcal{F}$ 为保持二元关系不变的概念约简.

为简单起见, 下文将保持二元关系不变的概念协调集 (概念约简) 统称为概念协调集 (概念约简).

定理2 ([47]) 对象概念的全体 $\mathcal{O}(G,M,I)$ 与属性概念的全体 $\mathcal{A}(G,M,I)$ 均为概念协调集.

通过概念在概念约简中的作用, 一个形式背景所对应的所有概念可以被分为如下 3 类.

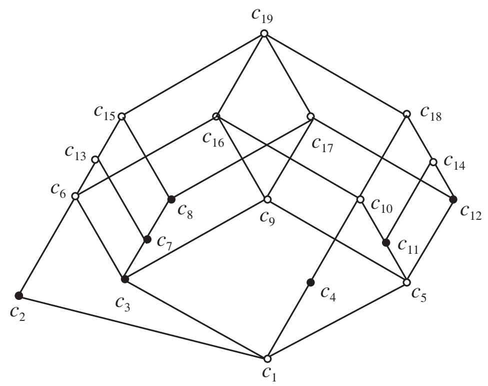

图 2　例 1 的一个概念约简 $\mathcal{F}_1$

Figure 2　A concept reduct $\mathcal{F}_1$ of example 1

表 3　例 1 的概念分类

Table 3　The classification of example 1

Type	Concept
Core concept	c_2, c_4
Relatively necessary concept	c_3, c_6, c_7, c_8, c_9, c_{10}, c_{11}, c_{12}, c_{13}, c_{14}, c_{15}, c_{16}, c_{17}
Unnecessary concept	c_1, c_5, c_{18}, c_{19}

定义5 ([47])　设 (G,M,I) 是一个形式背景, $L(G,M,I)$ 是概念格. 设 $\mathcal{R}=\{\mathcal{F}_i \mid i\in\tau,\tau$ 为指标集$\}$为 (G,M,I) 的所有概念约简的集合, 我们可将 $L(G,M,I)$ 中的概念分为 3 类:

(1) 核心概念集 $K=\bigcap_{i\in\tau}\mathcal{F}_i$;

(2) 相对必要概念集 $N=\bigcup_{i\in\tau}\mathcal{F}_i\setminus\bigcap_{i\in\tau}\mathcal{F}_i$;

(3) 不必要概念集 $U=L(G,M,I)\setminus\bigcup_{i\in\tau}\mathcal{F}_i$.

例3　例 1 的形式背景 (G,M,I) 存在许多概念约简, 例如, $\mathcal{F}_1=\{c_2,c_3,c_4,c_7,c_8,c_{11},c_{12}\}$ 即为一个概念约简, 在图 2 中用黑点标记. 也就是说, 通过 $\mathcal{F}_1$ 中的这些概念, 就可以得到原形式背景中的所有二元关系.

不过我们也容易从例 3 发现, 这个概念约简不是原始概念格的子格. 根据定义 5, 表 2 中的概念可以被分为 3 类, 具体如表 3 所示.

3.1　基于二元关系的概念特征

文献 [47] 从二元关系角度对形式背景 (G,M,I) 中所有概念的概念特征进行了研究, 此处列出后文所需的研究结果.

从二元关系角度得到的不必要概念的概念特征如下.

定理3 ([47])　设 (G,M,I) 是形式背景, $L(G,M,I)$ 是其概念格, $\mathcal{C}=\{\mathcal{F}\mid\mathcal{F}$ 为概念协调集, $\mathcal{F}\subseteq L(G,M,I)\}$. 则 $(A,B)\in L(G,M,I)$ 为不必要概念当且仅当 $A\times B\subseteq R((A,B))$, 其中, $R((A,B))=\bigcap_{\mathcal{F}\in\mathcal{C}}\{\bigcup_{(A_i,B_i)\in\mathcal{F}\setminus\{(A,B)\}}(A_i\times B_i)\}$.

从二元关系角度得到的核心概念的概念特征如下.

定理4 ([47])　设 (G,M,I) 是形式背景, $L(G,M,I)$ 是其概念格, $\mathcal{R}=\{\mathcal{F}_i\mid\mathcal{F}_i$ 为概念约简, $i\in\tau,\tau$

为指标集}. 则以下命题等价.

(1) (A,B) 是核心概念;

(2) $\bigcup_{(A_i,B_i)\in L(G,M,I)\setminus\{(A,B)\}}(A_i\times B_i)\neq\bigcup_{(A_i,B_i)\in L(G,M,I)}(A_i\times B_i)$;

(3) 存在 $(g,m)\in I\cap(A\times B)$, 对任意的 $(A_i,B_i)\in L(G,M,I)$, $(A_i,B_i)\neq(A,B)$, 有 $(g,m)\notin A_i\times B_i$;

(4) $(A,B)\in\mathcal{O}(G,M,I)\cap\mathcal{A}(G,M,I)$;

(5) (A,B) 是强制性因子.

从二元关系角度得到的相对必要概念的概念特征如下.

定理5 ([47])　设 (G,M,I) 是形式背景, $L(G,M,I)$ 是其概念格, 记

$$R((A,B))=\bigcap_{\mathcal{F}\in\mathcal{C}}\left\{\bigcup_{(A_i,B_i)\in\mathcal{F}\setminus\{(A,B)\}}(A_i\times B_i)\right\},$$

其中 $\mathcal{C}=\{\mathcal{F}\mid\mathcal{F}$ 为概念协调集, $\mathcal{F}\subseteq L(G,M,I)\}$. 对任意的 $(A,B)\in L(G,M,I)$, (A,B) 为相对必要概念当且仅当 $\bigcup_{(A_i,B_i)\in L(G,M,I)\setminus\{(A,B)\}}(A_i\times B_i)=\bigcup_{(A_i,B_i)\in L(G,M,I)}(A_i\times B_i)$ 且 $A\times B\nsubseteq R((A,B))$.

3.2　基于算子的对象 (属性) 概念的概念特征

本小节主要从 FCA 理论中导出算子 “$*$” 的角度对不同类别概念中的对象概念和属性概念的概念特征进行研究.

3.2.1　不必要概念的概念特征

定理6　设 (G,M,I) 是形式背景, $L(G,M,I)$ 是其概念格. 设 $g\in G$, 记 $G_0=\{g_i\mid g_i^*\neq g^*\}$, 则 (g^{**},g^*) 是一个不必要概念当且仅当存在 $G_{\rm sub}\subseteq G_0$, 使得 $g^*=\bigcup_{g_i\in G_{\rm sub}}g_i^*$.

证明　必要性. 假设对任意的 $G_{\rm sub}\subseteq G_0$, 有 $g^*\neq\bigcup_{g_i\in G_{\rm sub}}g_i^*$, 则存在 m, 使得 $(g,m)\in g^{**}\times g^*$, 且对于任意的 $G_{\rm sub}\subseteq G_0$, $g_i\in G_{\rm sub}$, 有 $(g,m)\notin g_i^{**}\times g_i^*$. 由定理 2 知, 对象概念全集 $\mathcal{O}(G,M,I)$ 是一个概念协调集, 又由上可知, 存在 $G_1\subseteq G_0$, 有 $\bigcup_{g_i\in G_1\cup\{g\}}g_i^{**}\times g_i^*=\bigcup_{(g_i^{**},g_i^*)\in\mathcal{O}(G,M,I)}g_i^{**}\times g_i^*=\bigcup_{(A,B)\in L(G,M,I)}A\times B$ 且 $\bigcup_{g_i\in G_1}g_i^{**}\times g_i^*\neq\bigcup_{(g_i^{**},g_i^*)\in\mathcal{O}(G,M,I)}g_i^{**}\times g_i^*$. 因此, 由定理 3 知, (g^{**},g^*) 不是一个不必要概念, 与已知条件矛盾. 故而, 一定存在 $G_{\rm sub}\subseteq G_0$, 使得 $g^*=\bigcup_{g_i\in G_{\rm sub}}g_i^*$.

充分性. 这里只针对 $|G_{\rm sub}|=2$ 的情况进行证明, $|G_{\rm sub}|\geqslant 3$ 时的证明类似. 即证明: 对任意的 $g\in G$, 存在 g_1, $g_2\in G$ ($g_1^*\neq g^*$, $g_2^*\neq g^*$), 满足 $g^*=g_1^*\cup g_2^*$, 则 (g^{**},g^*) 是不必要概念. 为证 (g^{**},g^*) 是不必要概念, 只需证 $g^{**}\times g^*\subseteq R((g^{**},g^*))$, 即对任意概念协调集 $\mathcal{F}$, 有 $g^{**}\times g^*\subseteq\bigcup_{(A_i,B_i)\in\mathcal{F}\setminus\{(g^{**},g^*)\}}A_i\times B_i$. 因此, 只需证以下两点: (1) 对任意概念协调集 $\mathcal{F}$, 存在 $\mathcal{F}_{\rm sub}\subseteq\mathcal{F}$, 使 $(g_1^{**}\times g_1^*)\cup(g_2^{**}\times g_2^*)\subseteq\bigcup_{(A_j,B_j)\in\mathcal{F}_{\rm sub}}(A_j\times B_j)$, $(A_j,B_j)\neq(g^{**},g^*)$; (2) $g^{**}\times g^*\subseteq(g_1^{**}\times g_1^*)\cup(g_2^{**}\times g_2^*)$. 下证 (1), 首先证: $g_1\notin g^{**}$, $g_2\notin g^{**}$. 假设 $g_1\in g^{**}$, 又 $g^*\neq g_1^*$, 则 $g_1^*\supset g^*$, 这与 $g^*=g_1^*\cup g_2^*$ 矛盾, 则 $g_1\notin g^{**}$; 同理, $g_2\notin g^{**}$. 因此, $g_1\times g_1^*\nsubseteq g^{**}\times g^*$ 且 $g_2\times g_2^*\nsubseteq g^{**}\times g^*$. 又若存在概念 (A,B), 使得 $g_1\in A$, 则 $g_1^{**}\subseteq A$, 则 $g_1\times B\subseteq A\times B$ 且 $g_1^{**}\times B\subseteq A\times B$. 因为存在概念 (A_i,B_i), $i=1,\dots,n$, 使得 $g_1\times g_1^*\subseteq\bigcup(A_i\times B_i)$, 所以, $g_1^{**}\times g_1^*\subseteq\bigcup(A_i\times B_i)$ 且 $(A_i,B_i)\neq(g^{**},g^*)$. 同理, 若 $g_2^{**}\times g_2^*\subseteq\bigcup(A_j\times B_j)$ 且 $(A_j,B_j)\neq(g^{**},g^*)$. 因此, (1) 得证. 下证 (2). 由 $g^*=g_1^*\cup g_2^*$ 可知 $g_1^*\subset g^*$, $g_2^*\subset g^*$, 因此, $g^{**}\subset g_1^{**}$, $g^{**}\subset g_2^{**}$, 则有 $g^{**}\times g_1^*\subset g_1^{**}\times g_1^*$, $g^{**}\times g_2^*\subset g_2^{**}\times g_2^*$. 于是, $(g^{**}\times g_1^*)\cup(g^{**}\times g_2^*)=g^{**}\times g^*\subset(g_1^{**}\times g_1^*)\cup(g_2^{**}\times g_2^*)$, 因此, (2) 得证.

由定理 6 的证明过程, 可以得到与其等价的定理如下.

定理6′ 设 (G,M,I) 是形式背景, $L(G,M,I)$ 是其概念格. 对于任意的 $g \in G$, 记 $G_1 = \{g_i \mid g_i^* \subset g^*\}$, (g^{**}, g^*) 是一个不必要概念当且仅当 $g^* = \bigcup_{g_i \in G_1} g_i^*$.

通过定理 6, 可以得到如下推论.

推论1 设 (G,M,I) 是形式背景, $L(G,M,I)$ 是其概念格. 对于任意的 $g \in G$, (g^{**}, g^*) 不是不必要概念当且仅当对任意的 $G_{\rm sub} \subseteq G_0$, $g^* \neq \bigcup_{g_i \in G_{\rm sub}} g_i^*$, 其中, $G_0 = \{g_i \mid g_i^* \neq g^*\}$.

推论2 设 $\mathcal{O}_u(G,M,I) = \{(g^{**}, g^*) \mid (g^{**}, g^*)$ 是不必要概念$\}$, 则 $\mathcal{O}(G,M,I) \setminus \mathcal{O}_u(G,M,I)$ 是概念约简.

证明 因为 $(g^{**}, g^*) \in \mathcal{O}(G,M,I) \setminus \mathcal{O}_u(G,M,I)$ 不是一个不必要概念, 则由推论 1 知, 对任意的 $G_{\rm sub} \subseteq G_0$, 有 $g^* \neq \bigcup_{g_i \in G_{\rm sub}} g_i^*$. 若 $g_i^* \subset g^*$, 有 $g^{**} \subset g_i^{**}$, 又 $g \in g^{**}$, 则 $g \in g_i^{**}$, 然而, 因为 $g^* \neq \bigcup_{g_i \in G_{\rm sub}} g_i^*$, 所以, 一定存在 $m \in M$, 使得 $(g,m) \in g^{**} \times g^*$ 且 $(g,m) \notin g_i^{**} \times g_i^*$; 若 $g_i^* \not\subseteq g^*$, 则 $g \notin g_i^{**}$, 则对任意的 $m \in g^*$, 有 $(g,m) \notin g_i^{**} \times g_i^*$. 综上可知, 对任意的 $(g^{**}, g^*) \in \mathcal{O}(G,M,I) \setminus \mathcal{O}_u(G,M,I)$, $\bigcup_{(A,B) \in \mathcal{O}(G,M,I) \setminus \{\mathcal{O}_u(G,M,I) \bigcup (g^{**}, g^*)\}} A \times B \neq \cup_{(A,B) \in L(G,M,I)} A \times B$. 故而, $\mathcal{O}(G,M,I) \setminus \mathcal{O}_u(G,M,I)$ 是一个概念约简.

类似地, 对于从属性角度出发考虑的属性概念及属性概念子集, 上述命题的类似结论亦成立.

定理7 设 (G,M,I) 是形式背景, $L(G,M,I)$ 是其概念格. 对于任意的 $m \in M$, 记 $M_0 = \{m_j \mid m_j^* \neq m^*\}$, 则 (m^*, m^{**}) 是一个不必要概念当且仅当存在 $M_{\rm sub} \subseteq M_0$, 使得 $m^* = \bigcup_{m_j \in M_{\rm sub}} m_j^*$.

定理7′ 设 (G,M,I) 是形式背景, $L(G,M,I)$ 是其概念格. 对于任意的 $m \in M$, 记 $M_0 = \{m_j \mid m_j^* \subset m^*\}$, (m^*, m^{**}) 是一个不必要概念当且仅当 $m^* = \bigcup_{m_j \in M_0} m_j^*$.

推论3 设 (G,M,I) 是形式背景, $L(G,M,I)$ 是其概念格. 对于任意的 $m \in M$, (m^*, m^{**}) 不是不必要概念当且仅当对任意的 $M_{\rm sub} \subseteq M_0$, $m^* \neq \bigcup_{m_j \in M_{\rm sub}} m_j^*$, 其中, $M_0 = \{m_j \mid m_j^* \neq m^*\}$.

推论4 设 $\mathcal{A}_u(G,M,I) = \{(m^*, m^{**}) \mid (m^*, m^{**})$ 是不必要概念$\}$, 则 $\mathcal{A}(G,M,I) \setminus \mathcal{A}_u(G,M,I)$ 是概念约简.

例4 以例 1 中概念 c_5 和 c_{18} 为例. 由例 3 知, 对象概念 $c_5 = (6, abcdf)$ 是个不必要概念, 而事实上, $6^* = 5^* \cup 8^*$, 符合定理 6; 由例 3 又知, 属性概念 $c_{18} = (5678, ad)$ 是个不必要概念, 而 $d^* = e^* \cup f^*$, 符合定理 7.

3.2.2 核心概念的概念特征

定理8 设 (G,M,I) 是形式背景, $L(G,M,I)$ 是其概念格. 对于任意的 $g \in G$, 记 $G_1 = \{g_i \mid g^* \not\subseteq g_i^*\}$, 则 (g^{**}, g^*) 是核心概念当且仅当 $g^* \not\subseteq \bigcup_{g_i \in G_1} g_i^*$.

证明 必要性. 假设 $g^* \subseteq \bigcup_{g_i \in G_1} g_i^*$. 则 $g^* \cap (\bigcup_{g_i \in G_1} g_i^*) = g^*$, 即 $\bigcup_{g_i \in G_1} (g^* \cap g_i^*) = g^*$. 又由 $g^* \cap g_i^* \subseteq (g^* \cap g_i^*)^{**}$ 可得 $\bigcup_{g_i \in G_1} (g^* \cap g_i^*) \subseteq \bigcup_{g_i \in G_1} (g^* \cap g_i^*)^{**}$, 于是 $g^* \subseteq \bigcup_{g_i \in G_1} ((g^* \cap g_i^*)^{**})$. 又 $g^* \not\subseteq g_i^* \Rightarrow g_i^{**} \not\subseteq g^{**}$, 则 $g^{**} \subset g^{**} \cup g_i^{**} \subseteq (g^* \cap g_i^*)^*$, 于是, $g^{**} \times g^* \subseteq g^{**} \times (\bigcup_{g_i \in G_1} ((g^* \cap g_i^*)^{**})) = \bigcup_{g_i \in G_1} g^{**} \times ((g^* \cap g_i^*)^{**}) \subset \bigcup_{g_i \in G_1} (g^* \cap g_i^*)^* \times (g^* \cap g_i^*)^{**}$, 因此, $\bigcup_{(A,B) \in L(G,M,I) \setminus \{(g^{**}, g^*)\}} A \times B = \bigcup_{(A,B) \in L(G,M,I)} A \times B$, 这与 (g^{**}, g^*) 是核心概念矛盾. 因此, 必要性得证.

充分性. 由定理 4 知, 要证 (g^{**}, g^*) 为核心概念, 即要证存在 $(g,m) \in g^{**} \times g^*$, 对任意的 $(A,B) \in L(G,M,I) \setminus \{(g^{**}, g^*)\}$, 有 $(g,m) \notin A \times B$. 由于对任意概念 $(A,B) \in L(G,M,I)$, 有 $(A,B) = \bigvee_{g_i \in A} (g_i^{**}, g_i^*)$, 其中的 g_i 有两种情况: $g^* \not\subseteq g_i^*$, $g^* \subseteq g_i^*$. 因此, 我们可将对象集 G 分为 2 个集合: $G_1 = \{g_i \mid g^* \not\subseteq g_i^*\}$, $G_2 = \{g_j \mid g^* \subseteq g_j^*\}$, 则, $G = G_1 \cup G_2$. 记 $G_{1\rm sub} \subseteq G_1$, $G_{2\rm sub} \subseteq G_2$,

则 A 有 3 种情况: $A=G_{1\text{sub}}$, $A=G_{2\text{sub}}$, $A=G_{1\text{sub}}\cup G_{2\text{sub}}$. 因此, 下面从这 3 种情况证明: 存在 $(g,m)\in g^{**}\times g^{*}$, 对任意的 $(A,B)\in L(G,M,I)\setminus\{(g^{**},g^{*})\}$, 有 $(g,m)\notin A\times B$.

(1) 若 $A=G_{1\text{sub}}$, 则 $(A,B)=\bigvee_{g_i\in A}(g_i^{**},g_i^{*})=\bigvee_{g_i\in G_{1\text{sub}}}(g_i^{**},g_i^{*})=((\bigcup_{g_i\in G_{1\text{sub}}}g_i^{**})^{**},\bigcap_{g_i\in G_{1\text{sub}}}g_i^{*})$. 因为 $g^{*}\not\subseteq\bigcup_{g_i\in G_1}g_i^{*}$, 则对任意的 $m\in g^{*}\setminus(\bigcup_{g_i\in G_1}g_i^{*})$, 有 $m\in g^{*}$ 且 $m\notin g_i^{*}$, 进而, $m\notin\bigcap_{g_i\in G_{1\text{sub}}}g_i^{*}$. 因此, $(g,m)\notin A\times B$.

(2) 若 $A=G_{2\text{sub}}$, 则 $(A,B)=\bigvee_{g_j\in A}(g_j^{**},g_j^{*})=\bigvee_{g_j\in G_{2\text{sub}}}(g_j^{**},g_j^{*})=((\bigcup_{g_j\in G_{2\text{sub}}}g_j^{**})^{**},\bigcap_{g_j\in G_{2\text{sub}}}g_j^{*})$. 因为对任意的 $g_j\in G_{2\text{sub}}$, 有 $g^{*}\subseteq g_j^{*}$, 所以, $g^{*}\subseteq\bigcap_{g_j\in G_{2\text{sub}}}g_j^{*}$. 若 $g^{*}\subset\bigcap_{g_j\in G_{2\text{sub}}}g_j^{*}$, 则存在 $m\in\bigcap_{g_j\in G_{2\text{sub}}}g_j^{*}$, $(g,m)\notin I$, 所以有 $g\notin(\bigcup_{g_j\in G_{2\text{sub}}}g_j^{**})^{**}$, 因此, 对任意的 $m\in M$, $(g,m)\notin A\times B$. 若 $g^{*}=\bigcap_{g_j\in G_{2\text{sub}}}g_j^{*}$, 则有 $(g^{**},g^{*})=((\bigcup_{g_j\in G_{2\text{sub}}}g_j^{**})^{**},\bigcap_{g_j\in G_{2\text{sub}}}g_j^{*})$.

(3) 若 $A=G_{1\text{sub}}\cup G_{2\text{sub}}$, 则 $(A,B)=\bigvee_{g_k\in A}(g_k^{**},g_k^{*})=(\bigvee_{g_i\in G_{1\text{sub}}}(g_i^{**},g_i^{*}))\vee(\bigvee_{g_j\in G_{2\text{sub}}}(g_j^{**},g_j^{*}))=(((\bigcup_{g_i\in G_{1\text{sub}}}g_i^{**})\cup(\bigcup_{g_j\in G_{2\text{sub}})}g_j^{**}))^{**},(\bigcap_{g_i\in G_{1\text{sub}}}g_i^{*})\cap(\bigcap_{g_j\in G_{2\text{sub}}}g_j^{*}))$. 由于 $g^{*}\not\subseteq\bigcup_{g_i\in G_{1\text{sub}}}g_i^{*}$, 则对任意的 $m\in g^{*}\setminus(\bigcup_{g_i\in G_1}g_i^{*})$, 有 $m\in g^{*}$ 且 $m\notin(\bigcap_{g_i\in G_{1\text{sub}}}g_i^{*})\cap(\bigcap_{g_j\in G_{2\text{sub}}}g_j^{*})$. 因此, $(g,m)\notin A\times B$.

综上可得, 对任意的 $m\in g^{*}\setminus(\bigcup_{g_i\in G_1}g_i^{*})$, 都满足 $(g,m)\in g^{**}\times g^{*}$ 且对任意的 $(A,B)\in L(G,M,I)\setminus\{(g^{**},g^{*})\}$, 我们有 $(g,m)\notin A\times B$. 因而, (g^{**},g^{*}) 是核心概念. 充分性得证.

从属性角度考虑的类似结论如下.

定理9 设 (G,M,I) 是形式背景, $L(G,M,I)$ 是其概念格. 对任意的 $m\in M$, 记 $M_1=\{m_j\mid m^{*}\not\subseteq m_j^{*}\}$, 则 (m^{*},m^{**}) 是一个核心概念当且仅当 $m^{*}\not\subseteq\bigcup_{m_j\in M_1}m_j^{*}$.

例5 以例 1 中概念 $c_2=(4,acghi)$ 说明定理 8 和 9. 从对象角度考虑, c_2 是对象 4 导出的对象概念. 对任意的 $g\in G\setminus\{4\}$, 有 $4^{*}\not\subset g^{*}$, 且 $4^{*}\not\subset\bigcup_{g\in G\setminus\{4\}}g^{*}$, 因此, c_2 是一个核心概念. 从属性角度考虑, c_2 是属性 i 导出的属性概念. 对任意的 $m\in M\setminus\{i\}$, 有 $i^{*}\not\subset m^{*}$, 且 $i^{*}\not\subset\bigcup_{m\in G\setminus\{i\}}m^{*}$, 因此, c_2 是一个核心概念. 这与例 3 给出结果一致.

3.2.3 相对必要概念的概念特征

由于概念共有 3 种类型, 而定理 6∼9 已给出对象 (属性) 概念为不必要概念与核心概念这两种概念的充分必要条件. 因此, 易得对象概念或属性概念为相对必要概念的充分必要条件.

定理10 设 (G,M,I) 是形式背景, $L(G,M,I)$ 是其概念格, 设 $g\in G$, 则对象概念 (g^{**},g^{*}) 是相对必要概念当且仅当

(1) 对任意的 $G_{\text{sub}}\subseteq G_0$, 有 $g^{*}\neq\bigcup_{g_i\in G_{\text{sub}}}g_i^{*}$, 其中 $G_0=\{g_i\mid g_i^{*}\neq g^{*}\}$;

(2) $g^{*}\subseteq\bigcup_{g_i\in G_1}g_i^{*}$, 其中 $G_1=\{g_i\mid g^{*}\not\subseteq g_i^{*}\}$.

定理11 设 (G,M,I) 是形式背景, $L(G,M,I)$ 是其概念格, 设 $m\in M$, 则属性概念 (m^{*},m^{**}) 是相对必要概念当且仅当

(1) 对任意的 $M_{\text{sub}}\subseteq M_0$, 有 $m^{*}=\bigcup_{m_j\in M_{\text{sub}}}m_j^{*}$, 其中 $M_0=\{m_j\mid m_j^{*}\neq m^{*}\}$;

(2) $m^{*}\subseteq\bigcup_{m_j\in M_1}m_j^{*}$, 其中 $M_1=\{m_j\mid m^{*}\not\subseteq m_j^{*}\}$.

3.3 基于布尔矩阵的概念约简

不难看出, 形式背景 (G,M,I) 的表达方式类似于一个布尔矩阵. 我们记第 i 个对象 g_i 所共同拥有的属性 g_i^{*} 为一个 $1\times|M|$ 的行向量 $\boldsymbol{\alpha}_i$, 记第 j 个属性 m_j 所共同拥有的对象 m_j^{*} 为一个 $|G|\times 1$ 的列向量 $\boldsymbol{\alpha}^j$. 则可以用布尔向量的语言对概念特征进行描述.

3.3.1 3 种类型对象 (属性) 概念的概念特征

利用定理 6 与 7 可从布尔矩阵角度得到不必要对象概念和属性概念的概念特征如下.

命题1 设 $\boldsymbol{W}$ 是对应于形式背景 (G,M,I) 的布尔矩阵, $\boldsymbol{\alpha}_i$ 是对应于 $g_i\in G$ 的行向量, 则 (g_i^{**},g_i^*) 是一个不必要概念当且仅当 $\boldsymbol{\alpha}_i=k_1\boldsymbol{\alpha}_1+\cdots+k_{i-1}\boldsymbol{\alpha}_{i-1}+k_{i+1}\boldsymbol{\alpha}_{i+1}+\cdots+k_p\boldsymbol{\alpha}_p$, 其中, $k_s=\begin{cases}1, & \boldsymbol{\alpha}_i>\boldsymbol{\alpha}_s\\ 0, & 否则\end{cases}$, $s\in\{1,2,\dots,p\}$.

命题2 设 $\boldsymbol{W}$ 是对应于形式背景 (G,M,I) 的布尔矩阵, $\boldsymbol{\alpha}^j$ 是对应于 $m_j\in M$ 的列向量, 则 (m_j^*,m_j^{**}) 是一个不必要概念当且仅当 $\boldsymbol{\alpha}^j=k_1\boldsymbol{\alpha}^1+\cdots+k_{j-1}\boldsymbol{\alpha}^{j-1}+k_{j+1}\boldsymbol{\alpha}^{j+1}+\cdots+k_q\boldsymbol{\alpha}^q$, 其中, $k_t=\begin{cases}1, & \boldsymbol{\alpha}^j>\boldsymbol{\alpha}^t\\ 0, & 否则\end{cases}$, $t\in\{1,2,\dots,q\}$.

利用定理 8 与 9 可从布尔矩阵角度得到核心概念的概念特征如下.

命题3 设 $\boldsymbol{W}$ 是对应于形式背景 (G,M,I) 的布尔矩阵, $\boldsymbol{\alpha}_i$ 是对应于 $g_i\in G$ 的行向量, (g_i^{**},g_i^*) 是一个核心概念当且仅当 $\boldsymbol{\alpha}_i\nleqslant k_1\boldsymbol{\alpha}_1+\cdots+k_{i-1}\boldsymbol{\alpha}_{i-1}+k_{i+1}\boldsymbol{\alpha}_{i+1}+\cdots+k_p\boldsymbol{\alpha}_p$, 其中, $k_s=\begin{cases}1, & \boldsymbol{\alpha}_i\nleqslant\boldsymbol{\alpha}_s\\ 0, & 否则\end{cases}$, $s\in\{1,2,\dots,p\}$.

命题4 设 $\boldsymbol{W}$ 是对应于形式背景 (G,M,I) 的布尔矩阵, $\boldsymbol{\alpha}^j$ 是对应于 $m_j\in M$ 的列向量, (m_j^*,m_j^{**}) 是一个核心概念当且仅当 $\boldsymbol{\alpha}^j\nleqslant k_1\boldsymbol{\alpha}^1+\cdots+k_{j-1}\boldsymbol{\alpha}^{j-1}+k_{j+1}\boldsymbol{\alpha}^{j+1}+\cdots+k_q\boldsymbol{\alpha}^q$, 其中, $k_t=\begin{cases}1, & \boldsymbol{\alpha}^j\nleqslant\boldsymbol{\alpha}^t\\ 0, & 否则\end{cases}$, $t\in\{1,2,\dots,q\}$.

利用定理 10 与 11 可从布尔矩阵角度得到相对必要对象概念和属性概念的概念特征如下.

命题5 设 $\boldsymbol{W}$ 是对应于形式背景 (G,M,I) 的布尔矩阵, $\boldsymbol{\alpha}_i$ 是对应于 $g_i\in G$ 的行向量, (g_i^{**},g_i^*) 是一个相对必要概念当且仅当

(1) $\boldsymbol{\alpha}_i\neq k_1\boldsymbol{\alpha}_1+\cdots+k_{i-1}\boldsymbol{\alpha}_{i-1}+k_{i+1}\boldsymbol{\alpha}_{i+1}+\cdots+k_p\boldsymbol{\alpha}_p$, 其中, $k_s=\begin{cases}1, & \boldsymbol{\alpha}_i>\boldsymbol{\alpha}_s\\ 0, & 否则\end{cases}$, $s\in\{1,2,\dots,p\}$.

(2) $\boldsymbol{\alpha}_i\leqslant k_1\boldsymbol{\alpha}_1+\cdots+k_{i-1}\boldsymbol{\alpha}_{i-1}+k_{i+1}\boldsymbol{\alpha}_{i+1}+\cdots+k_p\boldsymbol{\alpha}_p$, 其中, $k_s=\begin{cases}1, & \boldsymbol{\alpha}_i\nleqslant\boldsymbol{\alpha}_s\\ 0, & 否则\end{cases}$, $s\in\{1,2,\dots,p\}$.

命题6 设 $\boldsymbol{W}$ 是对应于形式背景 (G,M,I) 的布尔矩阵, $\boldsymbol{\alpha}^j$ 是对应于 $m_j\in M$ 的列向量, (m_j^*,m_j^{**}) 是一个相对必要概念当且仅当

(1) $\boldsymbol{\alpha}^j\neq k_1\boldsymbol{\alpha}^1+\cdots+k_{j-1}\boldsymbol{\alpha}^{j-1}+k_{j+1}\boldsymbol{\alpha}^{j+1}+\cdots+k_q\boldsymbol{\alpha}^q$, 其中, $k_t=\begin{cases}1, & \boldsymbol{\alpha}^j>\boldsymbol{\alpha}^t\\ 0, & 否则\end{cases}$, $t\in\{1,2,\dots,q\}$.

(2) $\boldsymbol{\alpha}^j\leqslant k_1\boldsymbol{\alpha}^1+\cdots+k_{j-1}\boldsymbol{\alpha}^{j-1}+k_{j+1}\boldsymbol{\alpha}^{j+1}+\cdots+k_q\boldsymbol{\alpha}^q$, 其中, $k_t=\begin{cases}1, & \boldsymbol{\alpha}^j\nleqslant\boldsymbol{\alpha}^t\\ 0, & 否则\end{cases}$, $t\in\{1,2,\dots,q\}$.

例6 表 1 所示的形式背景中, 对象 6 所对应的行向量为 $\boldsymbol{\alpha}_6=(1,1,1,1,0,1,0,0,0)$, 其中, $\boldsymbol{\alpha}_5<\boldsymbol{\alpha}_6$, $\boldsymbol{\alpha}_7<\boldsymbol{\alpha}_6$, 且 $\boldsymbol{\alpha}_5+\boldsymbol{\alpha}_7=\boldsymbol{\alpha}_6$, 故 $(6^{**},6^*)$ 为一个不必要概念; 对象 2 所对应的行向量为 $\boldsymbol{\alpha}_2=(1,1,0,0,0,0,1,1,0)$, 其中, $\boldsymbol{\alpha}_2>\boldsymbol{\alpha}_1$, 且 $\boldsymbol{\alpha}_2\neq\boldsymbol{\alpha}_1$; $\boldsymbol{\alpha}_2\nleqslant\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_4,\boldsymbol{\alpha}_5,\boldsymbol{\alpha}_6,\boldsymbol{\alpha}_7,\boldsymbol{\alpha}_8$ 且 $\boldsymbol{\alpha}_2\leqslant\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_4+\boldsymbol{\alpha}_5+\boldsymbol{\alpha}_6+\boldsymbol{\alpha}_7+\boldsymbol{\alpha}_8$, 故 $(2^{**},2^*)$ 为一个相对必要概念; 对象 4 所对应的行向量为 $\boldsymbol{\alpha}_4=(1,0,1,0,0,0,1,1,1)$, 其中, $\boldsymbol{\alpha}_4\nleqslant\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_5,\boldsymbol{\alpha}_6,\boldsymbol{\alpha}_7,\boldsymbol{\alpha}_8$ 且 $\boldsymbol{\alpha}_4\nleqslant\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2+\boldsymbol{\alpha}_3+\boldsymbol{\alpha}_5+\boldsymbol{\alpha}_6+\boldsymbol{\alpha}_7+\boldsymbol{\alpha}_8$, 故 $(4^{**},4^*)$ 为一个核心概念. 这均与例 3 一致.

3.3.2 概念约简与布尔矩阵的秩之间的关系

定理12 对于一个形式背景 (G,M,I), 至少存在一个概念约简, 其概念个数等于与该形式背景对应的布尔矩阵的行秩.

证明 记形式背景所对应的矩阵中行向量的集合为 W, 由其行向量所形成的空间为 $R(W)$, 记 $\mathcal{O}(G,M,I)\setminus\mathcal{O}_u(G,M,I)$ 所对应的行向量的集合为 W_v, $\mathcal{O}_u(G,M,I)$ 所对应的行向量的集合为 W_u. 由于对任意的 $\boldsymbol{\alpha}_i\in W_v$, 有 $\boldsymbol{\alpha}_i\neq\bigcup_{\boldsymbol{\alpha}_j\neq\boldsymbol{\alpha}_i,\boldsymbol{\alpha}_j\in W}k_j\boldsymbol{\alpha}_j$, 其中 $k_j=\begin{cases}1, & \boldsymbol{\alpha}_i>\boldsymbol{\alpha}_j\\ 0, & 否则\end{cases}$; 进一步, 有 $\boldsymbol{\alpha}_i\neq$

$\bigcup_{\boldsymbol{\alpha}_j \neq \boldsymbol{\alpha}_i, \boldsymbol{\alpha}_j \in W_v} k_j \boldsymbol{\alpha}_j$, 其中 $k_j = \begin{cases} 1, & \boldsymbol{\alpha}_i > \boldsymbol{\alpha}_j \\ 0, & \text{否则} \end{cases}$. 则 W_v 是一个无关的集合. 又对任意的 $\boldsymbol{\alpha}_i \in W_u$, 有 $\boldsymbol{\alpha}_i = \bigcup_{\boldsymbol{\alpha}_j \neq \boldsymbol{\alpha}_i, \alpha_j \in W_v} k_j \boldsymbol{\alpha}_j$, 因此, $R(W) = \langle W \rangle = \langle W_v \rangle$. 因此, W_v 为 $R(W)$ 的基, 则 W 的行秩为 $|W_v|$. 于是, 由推论 2, 命题得证.

类似地, 可以得到如下结论.

定理13　对于一个形式背景 (G, M, I), 至少存在一个概念约简, 其概念个数等于与该形式背景对应的布尔矩阵的列秩.

4　概念约简的计算方法

根据概念协调集的定义, 一个协调集应该能够体现出形式背景中具有确定二元关系的所有 (对象, 属性) 二元对. 为了得到形式背景 (G, M, I) 的概念约简, 可以把背景所确定的二元关系中的每一个二元对对应的概念罗列, 并在包含该二元对的概念中分别拿出一个再进行组合, 就会得到一个概念协调集, 进而再设法得到概念约简即可. 具体方法由以下命题给出.

命题7　设 (G, M, I) 是形式背景, $L(G, M, I)$ 是其概念格, 对任意的 $(g_i, m_j) \in I$, 记 $\mathrm{CS}_{ij} = \{(A, B) \mid (g_i, m_j) \in A \times B\}$, $\mathrm{DC} = \{\mathrm{CS}_{ij} \mid (g_i, m_j) \in I\}$, $g(\wedge) = \bigwedge_{\mathrm{CS}_{ij} \in \mathrm{DC}} (\bigvee_{(A,B) \in \mathrm{CS}_{ij}} (A, B))$, 则 $g(\wedge)$ 所对应的最小析取范式中的所有合取式为该形式背景所对应的概念约简的全体.

事实上, 某些二元对所对应的概念包含于其他二元对所对应的概念, 因此, 在进行合取过程中实质上是不需要被考虑的. Belohlavek 和 Trnecka [45] 在用形式概念作为因子研究布尔因子分解时, 用到了一个很好的思想, 即 $L(G, M, I)$ 中的区间. 具体内容可简单描述为: $L(G, M, I)$ 中的区间 $\mathcal{I}_{ij} = [(g_i^{**}, g_i^*), (m_j^*, m_j^{**})]$ 涵盖所有体现二元对 (g_i, m_j) 的概念 (对于二元对 (g, m), 记其所对应的区间为 $\mathcal{I}_{gm}$). 因此, 只需找出每个二元对所对应的区间, 并找出这些区间在包含关系下的极小区间即可. 具体方法由如下命题给出.

命题8　对任意的 $(g_i, m_j) \in I$, $\mathcal{I}_{ij} = [(g_i^{**}, g_i^*), (m_j^*, m_j^{**})]$, 若 $\mathcal{I}_{\min}$ 是包含关系下的极小区间集, 记 $\mathcal{I}_{\min_m}$ 为 $\mathcal{I}_{\min}$ 中的第 m 个元素, $g(\wedge) = \bigwedge_{m=1}^{|\mathcal{I}_{\min}|} (\bigvee_{(A,B) \in \mathcal{I}_{\min_m}} (A, B))$, 则 $g(\wedge)$ 所对应的最小析取范式中的所有合取式为该形式背景所对应的概念约简的全体.

例7　由于 $\mathcal{I}_{8f} = [(8^{**}, 8^*), (f^*, f^{**})] = \{c_{11}, c_{14}\}$, $\mathcal{I}_{8d} = [(8^{**}, 8^*), (d^*, d^{**})] = \{c_{11}, c_{10}, c_{14}, c_{18}\}$, $\mathcal{I}_{8a} = [(8^{**}, 8^*), (a^*, a^{**})] = \{c_{11}, c_{10}, c_{14}, c_{16}, c_{18}, c_{19}\}$, 又 $\mathcal{I}_{8f}, \mathcal{I}_{8d}, \mathcal{I}_{8a}$ 间有包含关系, 将其所对应的极小区间记为 $\mathcal{I}_{\min_1}$, 则有 $\mathcal{I}_{\min_1} = \mathcal{I}_{8f} = \{c_{11}, c_{14}\}$.

同理, 可得到 $\mathcal{I}_{\min}$ 中的其他元素, 则按照命题 8, 可得到所有的概念约简.

命题 7 与 8 的方法都是从二元关系所对应的所有二元对出发来考虑的. 结合概念特征, 可从部分二元对出发去得到所有的概念约简. 其方法由定理 14 保证.

定理14　设 (G, M, I) 是形式背景, $L(G, M, I)$ 是其概念格, K 是其核心概念集, $\mathcal{O}_u$ 为其不必要对象概念集, $\mathcal{A}_u$ 为其不必要属性概念集. 记 $I_{\mathrm{considered}} = \{(g_i, m_j) \mid (g_i, m_j) \in I \setminus (\bigcup_{K \cup \mathcal{O}_u \cup \mathcal{A}_u} A \times B)\}$, $\mathcal{I}_{\min}$ 是 $I_{\mathrm{considered}}$ 中所有二元关系所对应的区间在包含关系下的极小区间的集合. $\mathcal{I}_{\min_m}$ 是 $\mathcal{I}_{\min}$ 中的第 m 个极小区间, $g(\wedge) = K \wedge (\bigwedge_{m=1}^{|\mathcal{I}_{\min}|} (\bigvee_{(A,B) \in \mathcal{I}_{\min_m}} (A, B)))$, 则 $g(\wedge)$ 所对应的最小析取范式中的所有合取式为该形式背景所对应的概念约简的全体.

证明　首先, 若 $(A, B) \in K$, 则其所对应的二元对所形成的区间在包含关系下的极小区间为 $\{(A, B)\}$, 即 $\{(A, B)\} \in \mathcal{I}_{\min}$. 由定理 4 知, 若 $(A, B) \in K$, 则存在 $(g, m) \in I$, $(g, m) \in A \times B$, 对任意的 $(A_p, B_p) \in L(G, M, I)$, $(A_p, B_p) \neq (A, B)$, 有 $(g, m) \notin A_p \times B_p$. 则一定存在一个二元对 (g, m), 使得其所

表 4　部分二元对及其所对应的概念

Table 4　A part of pairs and concepts

Serial number k	Pair	Concept set CS_k	Serial number k	Pair	Concept set CS_k
1	$(1,b)$	$\{c_8,c_{17}\}$	8	$(3,g)$	$\{c_3,c_6,c_7,c_8,c_{13},c_{15}\}$
2	$(1,g)$	$\{c_8,c_{15}\}$	9	$(3,h)$	$\{c_3,c_6,c_7,c_{13}\}$
3	$(2,b)$	$\{c_7,c_8,c_{17}\}$	10	$(5,b)$	$\{c_{12},c_{17}\}$
4	$(2,g)$	$\{c_7,c_8,c_{13},c_{15}\}$	11	$(5,f)$	$\{c_{12},c_{14}\}$
5	$(2,h)$	$\{c_7,c_{13}\}$	12	$(8,c)$	$\{c_{10},c_{11},c_{16}\}$
6	$(3,b)$	$\{c_3,c_7,c_8,c_9,c_{17}\}$	13	$(8,f)$	$\{c_{11},c_{14}\}$
7	$(3,c)$	$\{c_3,c_6,c_9,c_{16}\}$			

表 5　极小概念区间所对应的二元对及概念

Table 5　Pairs and concepts corresponding to the minimum concept intervals

Serial number k	Minimum concept interval	Pair	Concept set CS_k
1	$\mathcal{I}_{1b}$	$(1,b)$	$\{c_8,c_{17}\}$
2	$\mathcal{I}_{1g}$	$(1,g)$	$\{c_8,c_{15}\}$
5	$\mathcal{I}_{2h}$	$(2,h)$	$\{c_7,c_{13}\}$
7	$\mathcal{I}_{3c}$	$(3,c)$	$\{c_3,c_6,c_9,c_{16}\}$
10	$\mathcal{I}_{5b}$	$(5,b)$	$\{c_{12},c_{17}\}$
11	$\mathcal{I}_{5f}$	$(5,f)$	$\{c_{12},c_{14}\}$
12	$\mathcal{I}_{8c}$	$(8,c)$	$\{c_{10},c_{11},c_{16}\}$
13	$\mathcal{I}_{8f}$	$(8,f)$	$\{c_{11},c_{14}\}$

对应的区间为 $\{(A,B)\}$, 并且, 对任意的 $(g_i,m_j)\in A\times B$, 其所对应的区间为 $\mathcal{I}_{ij}$, 则有 $\{(A,B)\}\subseteq\mathcal{I}_{ij}$. 因此, 极小区间为 $\{(A,B)\}$.

其次, 若 $(A,B)\in\mathcal{O}_u$, 则对任意的 $(g_i,m_j)\in A\times B$, 一定有 $\mathcal{I}_{ij}\notin\mathcal{I}_{\min}$. 对于任意的 $(g^{**},g^*)\in\mathcal{O}_u$, $(g_i,m_j)\in g^{**}\times g^*$. 因为 $\mathcal{I}_{ij}=[(g_i^{**},g_i^*),(m_j^*,m_j^{**})]$, 又 $g_i\in g^{**}$, 则 $g_i^*\supseteq g^*$, 于是, $(g_i^{**},g_i^*)\leqslant(g^{**},g^*)$. 又由定理 6 知, 存在 $G_{\mathrm{sub}}\subseteq G_0$, 有 $g^*=\bigcup_{g_k\in G_{\mathrm{sub}}}g_k^*$, 则 $g_k^*\subset g^*$, 因此 $(g^{**},g^*)<(g_k^{**},g_k^*)$, 且由于 $m_j\in g^*$, 则存在 $g_k\in G_{\mathrm{sub}}$, 有 $m_j\in g_k^*$, 则 $\mathcal{I}_{kj}=[(g_k^{**},g_k^*),(m_j^*,m_j^{**})]$, 于是, $\mathcal{I}_{kj}\subset\mathcal{I}_{ij}$, 因此, $\mathcal{I}_{ij}$ 一定不是极小区间.

同理, 若 $(A,B)\in\mathcal{A}_u$, 则对任意的 $(g_i,m_j)\in A\times B$, 一定有 $\mathcal{I}_{ij}\notin\mathcal{I}_{\min}$. 综上, 定理得证.

也就是说, 在进行概念约简的计算中, 若结合概念特征, 可以只通过不存在于核心概念及不必要对象/属性概念中的二元对得到所有的概念约简.

例8　由例 3 知, 例 1 形式背景 (G,M,I) 所对应的核心概念集合 $K=\{c_2,\ c_4\}$, 所对应的不必要对象概念集合 $\mathcal{O}_u=\{c_5\}$, 所对应的不必要属性概念集合 $\mathcal{A}_u=\{c_{18},c_{19}\}$, 且 c_1 中无任何二元关系, 因此, 通过定理 14, 只考虑以下二元对所对应的概念集合 $L(G,M,I)\setminus(K\cup\mathcal{O}_u\cup\mathcal{A}_u)$ 中的概念, 如表 4 所示 (记序号 k 所对应的概念的集合为 CS_k).

进一步, $\mathcal{I}_{\min}=\{\mathcal{I}_{1b},\ \mathcal{I}_{1g},\ \mathcal{I}_{2h},\ \mathcal{I}_{3c},\ \mathcal{I}_{5b},\ \mathcal{I}_{5f},\ \mathcal{I}_{8c},\ \mathcal{I}_{8f}\}$, 其所对应的具体概念如表 5 所示. 则 $g(\wedge)=K\wedge(\bigwedge_{k=1}^{8}(\bigvee_{c_j\in\mathrm{CS}_k}c_j))$ 为该形式背景的所有概念约简的集合. 然而, 本例中 $g(\wedge)$ 中元素的个数高达 52 个, 这正好给我们提供了诸多有趣的可进一步研究的问题: 这些概念约简分别具有什么特

点, 何时该使用哪个约简, 某种特定的约简如何获得等.

求解保持二元关系不变的概念约简的具体算法如算法 1 所示.

Algorithm 1 The algorithm for finding concept reducts in the formal context (G, M, I)

Input: (G, M, I), $L(G, M, I)$, set of row vectors Row, set of column vectors Col;

Output: Concept reducts.

// Type of object concepts;
$\mathcal{O}_u = \emptyset$, $\mathcal{A}_u = \emptyset$, $K = \emptyset$, $R = \emptyset$;
for $i = 1$ to $|\mathrm{Col}|$ **do**
  $\boldsymbol{\alpha} \Leftarrow \mathbf{0}$, $\boldsymbol{\beta} \Leftarrow \mathbf{0}$;
  for $j = 1$ to $|\mathrm{Col}|$ **do**
    if $\mathrm{Col}[j] < \mathrm{Col}[i]$ **then**
      $\boldsymbol{\alpha} \Leftarrow \boldsymbol{\alpha} + \mathrm{Col}[i]$;
    else if $\mathrm{Col}[j] \geqslant \mathrm{Col}[i]$ **then**
      $\boldsymbol{\beta} \Leftarrow \boldsymbol{\beta}$;
    else
      $\boldsymbol{\beta} \Leftarrow \boldsymbol{\beta} + \boldsymbol{\alpha}$;
    end if
  end for
  if $\boldsymbol{\alpha} = \mathrm{Col}[i]$ **then**
    Add (g_i^{**}, g_i^{*}) to $\mathcal{O}_u$;
  else if $\mathrm{Col}[i] \nleqslant \boldsymbol{\beta}$ **then**
    Add (g_i^{**}, g_i^{*}) to K;
  else
    Add (g_i^{**}, g_i^{*}) to R;
  end if
end for
// Type of unnecessary attribute concepts;
for $p = 1$ to $|\mathrm{Row}|$ **do**
  $\boldsymbol{\alpha} \Leftarrow \mathbf{0}$;
  for $q = 1$ to $|\mathrm{Row}|$ **do**
    if $\mathrm{Row}[q] < \mathrm{Row}[p]$ **then**
      $\boldsymbol{\alpha} \Leftarrow \boldsymbol{\alpha} + \mathrm{Row}[q]$;
    else
      $\boldsymbol{\alpha} \Leftarrow \boldsymbol{\alpha}$;
    end if
  end for
  if $\mathrm{Row}[p] = \alpha$ **then**
    Add (m^{*}, m^{**}) to $\mathcal{A}_u$;
  else
    $\mathcal{A}_u \Leftarrow \mathcal{A}_u$;
  end if
end for
// Calculate $\mathcal{I}_{\text{considered}}$;
$\mathcal{I}_{\text{considered}} = \emptyset$;
$L_{\text{sub}}(G, M, I) = L(G, M, I) \setminus (K \cup \mathcal{O}_u \cup \mathcal{A}_u)$;
$I_{\text{considered}} = I \setminus (\bigcup_{(A,B) \in L_{\text{sub}}(G,M,I)} A \times B)$;
for $i = 1$ to $|I_{\text{considered}}|$ **do**
  $\mathcal{I}_i = \emptyset$;
  for every $(A_j, B_j) \in L_{\text{sub}}(G, M, I)$ **do**

        if $I_{\text{considered}}[i] \subseteq A_j \times B_j$ **then**
            Add (A_j, B_j) to $\mathcal{I}_i$;
        end if
    end for
    Add $\mathcal{I}_i$ *to* $\mathcal{I}_{\text{considered}}$;
end for
// Minimum interval sets;
$\mathcal{I}_{\min} = \mathcal{I}_{\text{considered}}$;
for $i = 1$ to $\mathcal{I}_{\text{considered}}$ **do**
    for $j = 1$ to $\mathcal{I}_{\text{considered}}$ **do**
        if $i \neq j$ **then**
            if $\mathcal{I}_{\text{considered}}[i] \subseteq \mathcal{I}_{\text{considered}}[j]$ **then**
                Remove $\mathcal{I}_{\text{considered}}[j]$ from $\mathcal{I}_{\min}$;
            end if
        end if
    end for
end for
// Obtain concept reducts by conjunction and disjunction;
$\mathrm{CS} = K \wedge (\wedge_{n=1}^{\mathcal{I}_{\min}} (\vee_{(A,B)\in(\mathcal{I}_{\min}(A,B))}))$;
return CS.

在算法 1 中, 求解对象概念类型的时间复杂度为 $O(|M||G|^2)$, 求解不必要属性概念的时间复杂度为 $O(|G||M|^2)$, 计算 $\mathcal{I}_{\text{considered}}$ 的时间复杂度为 $\max_{j\leqslant|L(G,M,I)\setminus(K\cup\mathcal{O}_u\cup\mathcal{A}_u)|}O(|A_j\times B_j||L(G,M,I)\setminus(K\cup\mathcal{O}_u \cup \mathcal{A}_u)|^2)$, 求解极小区间集的时间复杂度为 $\max_{i,j\leqslant|\mathcal{I}_{\text{considered}}|}O(|\mathcal{I}_{\text{considered}}[i]||\mathcal{I}_{\text{considered}}[j]||\mathcal{I}_{\text{considered}}|^2)$, 通过析取、合取得到所有概念约简的时间复杂度为 $\max_{n\leqslant|\mathcal{I}_{\min}|-1}O(|\mathcal{I}_{\min}[n]||\mathcal{I}_{\min}[n+1]||\mathcal{I}_{\min}|)$, 因此, 整体的时间复杂度主要来自于求解极小区间集, 为 $\max_{i,j\leqslant|\mathcal{I}_{\text{considered}}|}O(|\mathcal{I}_{\text{considered}}[i]||\mathcal{I}_{\text{considered}}[j]||\mathcal{I}_{\text{considered}}|^2)$.

5 结论

本文基于文献 [47], 对保持二元关系不变的概念约简进行了更深一步的研究. 我们从算子与布尔矩阵角度对对象概念和属性概念的概念特征进行了研究并对保持二元关系不变的概念约简的计算方法进行了讨论. 这些结论与文献 [47] 的结论结合起来, 展示了从二元关系、算子、布尔矩阵 3 个角度研究保持二元关系不变的概念约简的一个理论框架和成果.

概念约简从概念而非属性的角度出发, 删除部分概念, 使得所留下的概念一定包含了原有的二元关系. 可以确定的是, 所留下的概念一定可以形成一个偏序集, 但不一定是格, 也不一定是原概念格的子格. 我们将在此基础上挖掘更多有意义的知识, 并探索这些新结论与既有知识之间的联系.

在概念约简的研究过程中, 我们也发现很多新的问题, 比如, (1) 从概念约简的数量而言, 一个形式背景可能会产生许多的概念约简, 因此, 这些概念约简是否还有各自独特的性质? 如, 概念约简中概念个数是否最小. (2) 从概念的完备性而言, 一个概念约简内部的元素间是否存在独立性, 一个概念约简是否具有完备性? (3) 从概念构造与概念格生成角度而言, 每一个概念约简是否都能设法生成所有的概念, 也就是说, 概念约简是否是概念格的一组 “基”? 这些问题都是很有趣且具有研究价值的. 我们会在后续的工作中进一步深入探索.

参考文献

1 Wille R. Restructuring lattice theory: an approach based on hierarchies of concepts. In: Proceedings of the NATO Advanced Study Institute, 1982. 445–470
2 Ganter B, Wille R. Formal Concept Analysis: Mathematical Foundations. Berlin: Springer, 1999
3 Yao Y Y. Concept lattices in rough set theory. In: Proceedings of Annual Meeting of the North American Fuzzy Information Processing Society, New York, 2004. 796–801
4 Fu H, Nguifo E M. A parallel algorithm to generate formal concepts for large data. In: Proceedings of International Conference on Formal Concept Analysis, Berlin, 2004. 394–401
5 Wang D X, Hu X G, Liu X P. Novel attribute induction algorithm based on quantized concept lattice. J Xi'an Jiaotong Univ, 2007, 41: 176–179 [王德兴, 胡学钢, 刘晓平. 一种新颖的基于量化概念格的属性归纳算法. 西安交通大学学报, 2007, 41: 176–179]
6 Krajca P, Outrata J, Vychodil V. Advancesin algorithms based on CbO. In: Proceedings of the 7th International Conference on Concept Lattices and Their Applications, Sevilla, 2010. 325–337
7 Andrews S. A 'best-of-breed' approach for designing a fast algorithm for computing fixpoints of Galois connections. Inf Sci, 2015, 295: 633–649
8 Zhang W X, Wei L, Qi J J. Attribute reduction theory and approach to concept lattice. Sci China Ser F-Inf Sci, 2005, 48: 713–726
9 Wei L. Reduction theory and approach to rough set and concept lattice. Dissertation for Ph.D. Degree. Xi'an: Xi'an Jiaotong University. 2005 [魏玲. 粗糙集与概念格约简理论与方法. 博士学位论文. 西安: 西安交通大学. 2005]
10 Wang X, Ma J M. A novel approach to attribute reduction in concept lattices. In: Proceedings of the 1st International Conference on Rough Sets and Knowledge Technology, Chongqing, 2006. 522–529
11 Li T J, Li M Z, Gao Y. Attribute reduction of concept lattice based on irreducible elements. Int J Wavelets Multiresolut Inf Process, 2013, 11: 1350046
12 Wu W Z, Leung Y, Mi J S. Granular computing and knowledge reduction in formal contexts. IEEE Trans Knowl Data Eng, 2009, 21: 1461–1474
13 Wei L, Qi J J, Zhang W X. Attribute reduction theory of concept lattice based on decision formal contexts. Sci China Ser F-Inf Sci, 2008, 51: 910–923
14 Shao M W, Li K W. Attribute reduction in generalized one-sided formal contexts. Inf Sci, 2017, 378: 317–327
15 Zhang Q H, Wang G Y, Liu X Q. Rule acquisition algorithm based on maximal granule. Pattern Recogn Artif Intell, 2012, 25: 388–396 [张清华, 王国胤, 刘显全. 基于最大粒的规则获取算法. 模式识别与人工智能, 2012, 25: 388–396]
16 Li J H, Mei C L, Wang J H, et al. Rule-preserved object compression in formal decision contexts using concept lattices. Knowl-Based Syst, 2014, 71: 435–445
17 Zhu Z C, Wei L. Two-way rules acquisition based on class contexts. J Northwest Univ (Nat Sci Edit), 2015, 45: 517–524 [朱治春, 魏玲. 基于类背景的双向规则的获取. 西北大学学报 (自然科学版), 2015, 45: 517–524]
18 Liu L, Wei L, Qian T. Three-way rules extraction in formal decision contexts with confidence. J Shandong Univ (Nat Sci), 2017, 52: 101–110 [刘琳, 魏玲, 钱婷. 决策形式背景中具有置信度的三支规则提取. 山东大学学报 (理学版), 2017, 52: 101–110]
19 Yao Y Y. Rough-set concept analysis: interpreting RS-definable concepts based on ideas from formal concept analysis. Inf Sci, 2016, 346: 442–462
20 Qi J J, Wei L, Li Z Z. A partitional view of concept lattice. In: Proceedings of the 10th International Conference on Rough Sets, Fuzzy Sets, Data Mining, and Granular Computing, Regina, 2005. 74–83
21 Fuentes-González R, Burusco A. The study of the L-fuzzy concept lattice. Mathware Soft Comput, 1994, 1: 209–218
22 Bělohlávek R. Concept lattices and order in fuzzy logic. Ann Pure Appl Logic, 2004, 128: 277–298
23 Yao Y Y. An outline of a theory of three-way decisions. In: Proceedings of Rough Sets and Current Trends in Computing, Chengdu, 2012. 1–17
24 Qi J J, Wei L, Yao Y Y. Three-way formal concept analysis. In: Proceedings of Rough Sets and Knowledge Technology, Shanghai, 2014. 732–741
25 Qi J J, Qian T, Wei L. The connections between three-way and classical concept lattices. Knowl-Based Syst, 2016, 91: 143–151
26 Huang C C, Li J H, Mei C L, et al. Three-way concept learning based on cognitive operators: an information fusion

viewpoint. Int J Approx Reason, 2017, 83: 218–242
27 Qi J J, Wang W W. A multithreaded parallel algorithm for constructing three-way concepts. J Xi'an Jiaotong Univ, 2017, 51: 116–121 [祁建军, 汪文威. 多线程并行构建三支概念. 西安交通大学学报, 2017, 51: 116–121]
28 Qian T, Wei L, Qi J J. Constructing three-way concept lattices based on apposition and subposition of formal contexts. Knowl-Based Syst, 2017, 116: 39–48
29 Chen X, Wei L, Qian T. Attribute reduction in formal decision contexts based on AE-concept lattices. J Shandong Univ (Nat Sci), 2017, 52: 95–103 [陈雪, 魏玲, 钱婷. 基于 AE – 概念格的决策形式背景属性约简. 山东大学学报(理学版), 2017, 52: 95–103]
30 Li J H, Mei C L, Xu W Z, et al. Concept learning via granular computing: a cognitive viewpoint. Inf Sci, 2015, 298: 447–467
31 Zhi H L, Li J H. Granule description based on formal concept analysis. Knowl-Based Syst, 2016, 104: 62–73
32 Qi J J, Wei L, Wan Q. Multi-level granularity in formal concept analysis. Granul Comput, 2019, 4: 351–362
33 Li J H, Mei C L, Lv Y J. Incomplete decision contexts: approximate concept construction, rule acquisition and knowledge reduction. Int J Approx Reason, 2013, 54: 149–165
34 Djouadi Y, Dubois D, Prade P. Graduality, uncertainty and typicality in formal concept analysis. In: 35 Years of Fuzzy Set Theory. Berlin: Springer, 2010. 127–147
35 Yao Y Y. Interval sets and three-way concept analysis in incomplete contexts. Int J Mach Learn Cyber, 2017, 8: 3–20
36 Ren R S, Wei L, Yao Y Y. An analysis of three types of partially-known formal concepts. Int J Mach Learn Cyber, 2018, 9: 1767–1783
37 Li M Z, Wang G Y. Approximate concept construction with three-way decisions and attribute reduction in incomplete contexts. Knowl-Based Syst, 2016, 91: 165–178
38 Zhi H L. Knowledge representation on incomplete formal context. Comput Sci, 2015, 42: 276–278 [智慧来. 不完备形式背景上的知识表示. 计算机科学, 2015, 42: 276–278]
39 Wu W Z, Qian Y H, Li T J, et al. On rule acquisition in incomplete multi-scale decision tables. Inf Sci, 2017, 378: 282–302
40 Ren R S, Wei L. The attribute reductions of three-way concept lattices. Knowl-Based Syst, 2016, 99: 92–102
41 Wang Z, Wei L. Attribute reduction of partially-known formal concept lattices for incomplete contexts. Comput Sci, 2018, 45: 73–78 [王振, 魏玲. 基于单边区间集概念格的不完备形式背景的属性约简. 计算机科学, 2018, 45: 73–78]
42 Keprt A, Snasel V. Binary factor analysis with help of formal concepts. In: Proceedings of International Workshop on Concept Lattices and Their Applications, Ostrava, 2004. 90–101
43 Belohlavek R, Vychodil V. On Boolean factor analysis with formal concept as factors. SCIS & ISIS, 2006. 1054–1059
44 Belohlavek R, Vychodil V. Discovery of optimal factors in binary data via a novel method of matrix decomposition. J Comput Syst Sci, 2010, 76: 3–20
45 Belohlavek R, Trnecka M. From-below approximations in Boolean matrix factorization: geometry and new algorithm. J Comput Syst Sci, 2015, 81: 1678–1697
46 Trnecka M, Trneckova M. Data reduction for Boolean matrix factorization algorithms based on formal concept analysis. Knowl-Based Syst, 2018, 158: 75–80
47 Cao L, Wei L, Qi J J. Concept reduction preserving binary relations. Pattern Recogn Artif Intell, 2018, 31: 516–524 [曹丽, 魏玲, 祁建军. 保持二元关系不变的概念约简. 模式识别与人工智能, 2018, 31: 516–524]
48 Kim K H. He S Y, Kong D Y, Huang Z L, et al. Boolean Matrix Theory and Applications. Beijing: Knowledge Publishing House. 1987 [Kim K H, 著. 何善堉, 孔德涌, 黄正篱, 等译. 布尔矩阵理论及其应用. 北京: 知识出版社, 1987]
49 Glodeanu C. Triadic factor analysis. In: Proceedings of the 7th International Conference on Concept Lattices and Their Applications, Sevilla, 2010. 127–138

Concept reduction and concept characteristics in formal concept analysis

Ling WEI[1,4*], Li CAO[1,4], Jianjun QI[2,4] & Wenxiu ZHANG[3]

1. *School of Mathematics, Northwest University, Xi'an* 710127*, China*;
2. *School of Computer Science and Technology, Xidian University, Xi'an* 710071*, China*;
3. *School of Mathematics and Statistics, Xi'an Jiaotong University, Xi'an* 710049*, China*;
4. *Institute of Concepts, Cognition and Intelligence, Northwest University, Xi'an* 710127*, China*

* Corresponding author. E-mail: wl@nwu.edu.cn

Abstract Formal concept analysis, which is based on formal contexts and concept lattices, is a data analysis method. Formal concepts reflect the relationship between objects and attributes and realize a formal description of concepts in philosophy. To simplify the information description, this paper proposes concept reduction in the formal concept analysis framework and studies the related theories on concept reduction preserving binary relations. In addition, characteristics analysis of three types of concepts that play different roles in reduction processing is performed from the perspective of operators and a Boolean matrix. Finally, a method to calculate concept reduction is provided.

Keywords formal context, formal concept analysis, binary relations, reduction, characteristic

Ling WEI was born in 1972. She received her Ph.D. degree in mathematics from the Xi'an Jiaotong University, Xi'an, China in 2005. She is currently a professor at Northwest University. Her research interests include formal concept analysis, rough set theory, three-way decisions, and granular computing.

Li CAO was born in 1995. She received her master's degree in applied mathematics from Northwest University, Xi'an, China in 2019. Her research interest includes formal concept analysis.

Jianjun QI was born in 1970. He received his Ph.D. degree in computer science from Xi'an Jiaotong University, Xi'an, China in 2005. He is currently an associate professor at Xidian University. His research interests include formal concept analysis, three-way decisions, and intelligent data analysis.

Wenxiu ZHANG was born in 1940. He received his master's degree in information theory from Nankai University, Tianjin in 1967. He is currently a professor at Xi'an Jiaotong University. His research interests include uncertainty analysis, information theory, rough set theory, formal concept analysis, and granular computing.