

DOI: 10.11830/ISSN.1000-5013.202008033



知识基的布尔矩阵求解方法

谢小贤^{1,2}, 李进金^{1,3}, 陈东晓¹, 林荣德^{1,2}

(1. 华侨大学 数学科学学院, 福建 泉州 362021;
2. 华侨大学 计算科学福建省高校重点实验室, 福建 泉州 362021;
3. 闽南师范大学 数学与统计学院, 福建 漳州 363000)

摘要: 用布尔矩阵方法对知识空间的原子和知识基进行研究. 首先, 建立知识空间和(反)知识背景之间的联系; 其次, 用布尔矩阵表示(反)知识背景, 研究其对应的关系矩阵和对象关系矩阵的性质; 最后, 从知识状态、算子、布尔向量和布尔矩阵等角度判定原子的特征, 给出知识空间中原子和知识基的求解方法.

关键词: 知识空间; 知识基; 形式背景; 布尔矩阵

中图分类号: O 29; TP 182 **文献标志码:** A **文章编号:** 1000-5013(2021)03-0410-11

Boolean Matrix Method of Knowledge Base

XIE Xiaoxian^{1,2}, LI Jinjin^{1,3}, CHEN Dongxiao¹, LIN Rongde^{1,2}

(1. School of Mathematical Sciences, Huaqiao University, Quanzhou 362021, China;
2. Fujian Province University Key Laboratory of Computational Science, Huaqiao University, Quanzhou 362021, China;
3. School of Mathematics and Statistics, Minnan Normal University, Zhangzhou 363000, China)

Abstract: Atom and knowledge base on knowledge space are studied by Boolean matrix method. The connection between knowledge space and (anti-) knowledge context established, the (anti-) knowledge context is described by Boolean matrix, and the properties of there relation matrix and object relation matrix are studied. Finally, characteristic of the atom is judged by the perspective of knowledge state, operators, Boolean vector and Boolean matrix, and the method of getting the atom and the knowledge base is given.

Keywords: knowledge space; knowledge base; formal context; Boolean matrix

知识空间理论(KST)是美国数学心理学家 Doignon 等^[1]提出的, 通过分析学生对不同水平的一系列有关问题的解答情况, 获得学生的认知水平和学习路径, 进而指导学生学习 and 教师教学, 为教育教学提供有效的评价方法. Koppen 等^[2]用专家问询系统生成知识空间; Albert 等^[3]用问题系统构建知识空间; Dowling^[4]在有限知识空间中提出一种构建知识基的方法; Falmagne 等^[5]对 Dowling 生成知识空间的算法进行改进, 提高了效率. KST 作为自适应教学和测试系统中最有效的知识表示理论^[6-7], 已应用到教育领域中, 例如, 计算机知识诊断系统 ALEKS^[8]的开发与应用. 文献[9-10]用知识空间理论分析学习路径等问题, 并应用于化学教学.

形式概念分析(FCA)是德国数学家 Wille^[11]提出的, 刻画对象与属性之间的概念层次结构, 以及概念之间的泛化和特化关系, 主要研究属性约简^[12-20]、粒约简^[21-22]、概念约简^[23-25]和规则提取^[13-14]等, 作

收稿日期: 2020-08-22
通信作者: 李进金(1960-), 男, 教授, 博士, 博士生导师, 主要从事拓扑学与不确定性理论的研究. E-mail: jinjinli@mnnu.edu.cn.
基金项目: 国家自然科学基金资助项目(11871259); 福建省自然科学基金资助项目(2017J01114, 2016J01304); 福建省高校创新团队发展计划, 泉州市高层次人才团队项目(2017ZT012)

为一种有力的数据分析工具,已广泛应用于机器学习、知识评价等领域. Rusch 等^[26]建立了知识空间与形式背景之间的联系; Spoto 等^[27]将 KST 与 FCA 相结合,分析知识结构的有效性;李进金等^[28]分析知识空间和形式概念分析之间的联系,以及知识基和知识空间的构建方法.

知识基是知识空间的核心,是知识空间的最小生成组,蕴含了知识空间的所有信息. 因此,研究知识基具有非常重要的意义. 由于知识空间与形式背景关系密切,而形式背景可看作一个布尔关系矩阵,可用矩阵方法解决形式背景中的属性约简^[17-20]和概念约简^[24-25]等问题. 因此,本文将用矩阵方法研究知识空间中的原子和知识基. 首先,通过 KST 和 FCA 的联系,导出反知识背景和知识背景;其次,利用布尔矩阵运算,研究(反)知识背景的关系矩阵、对象关系矩阵及其相关性质;最后,从知识状态、算子、布尔向量和布尔矩阵等角度,讨论知识空间中原子和知识基的特征,并给出求解方法.

1 预备知识

1.1 知识空间

设 $q_i(1 \leq i \leq m)$ 表示问题,则称 $Q = \{q_1, q_2, \dots, q_m\}$ 为所讨论知识的问题域. 文中只考虑论域 Q 有限的情形.

定义 1^[1] 学生在理想条件下能正确回答问题域 Q 中的问题所构成的集合称为知识状态,记为 K .

定义 2^[2] 设 Q 为问题域, \mathcal{K} 是 Q 的知识状态集族,并且 \mathcal{K} 至少包含了空集 \emptyset 和全集 Q , 称 (Q, \mathcal{K}) 为知识结构,记

$$\mathcal{K} = \{\emptyset, K_1, K_2, \dots, Q\},$$

其中,每一个 $K_i \subseteq Q$.

对知识结构 (Q, \mathcal{K}) ,若 \mathcal{K} 对有限并封闭,即

$$\forall K_i, K_j \in \mathcal{K} \Rightarrow K_i \cup K_j \in \mathcal{K},$$

则称 (Q, \mathcal{K}) 为知识空间,或称 \mathcal{K} 为知识空间.

当 Q 有限时,称知识结构 (Q, \mathcal{K}) 是有限的. 当 \mathcal{K} 有限时,称知识结构 (Q, \mathcal{K}) 是实质有限的.

定义 3^[5] 若 \mathcal{G}' 包含 \mathcal{G} 中所有有限个元素的并组成的集合,则称集族 \mathcal{G}' 是 \mathcal{G} 的张成,记为 $S(\mathcal{G}) = \mathcal{G}'$,或称 \mathcal{G} 张成 \mathcal{G}' .

由 $S(\mathcal{G})$ 的定义可知, $S(\mathcal{G})$ 是并封闭的.

定义 4^[5] 设集族 \mathcal{F} 是并封闭的,若 \mathcal{B} 是张成 \mathcal{F} 的最小子集族,且 \mathcal{B} 中的任何一个集合 K ,均不能由 \mathcal{B} 中其他集合的并集表示,则称 \mathcal{B} 为 \mathcal{F} 的基,即 $S(\mathcal{B}) = \mathcal{F}$.

定义 4 中最小子集族是指对于任意 $\mathcal{H} \subseteq \mathcal{B}$ 且 $S(\mathcal{H}) = \mathcal{F}$,有 $\mathcal{H} = \mathcal{B}$. 称知识空间的基为知识基,并约定 $\emptyset \notin \mathcal{B}$. 另外,对知识空间 (Q, \mathcal{K}) ,文中仅讨论 $K_i = K_j \Rightarrow i = j (K_i, K_j \in \mathcal{K})$ 的情形.

引理 1^[5] 任何实质有限的知识空间有知识基.

知识基 \mathcal{B} 可生成唯一的知识空间 (Q, \mathcal{K}) ,知识空间 (Q, \mathcal{K}) 可确定唯一的知识基 \mathcal{B} ,两者一一对应.

1.2 形式概念分析

定义 5^[11-12] 称 (P, Q, I) 为形式背景,其中, $P = \{p_1, p_2, \dots, p_n\}$ 为对象集, $Q = \{q_1, q_2, \dots, q_m\}$ 为属性集, I 为 P 与 Q 间的二元关系, $I \subseteq P \times Q$. 若 $(p, q) \in I$,记为 pIq ,表示对象 p 具有属性 q ;若 $(p, q) \notin I$,表示对象 p 不具有属性 q . 设 $c_{i,j} = \begin{cases} 1, (p_i, q_j) \in I \\ 0, (p_i, q_j) \notin I \end{cases}$,则形式背景 (P, Q, I) 可以用布尔矩阵 $\mathbf{M}_I = (c_{i,j})_{n \times m}$ 表示,称 $\mathbf{M}_I = (c_{i,j})_{n \times m}$ 为该形式背景 (P, Q, I) 的关系矩阵.

设 $X \subseteq P, B \subseteq Q$,形式背景 (P, Q, I) 上的算子: $X^* = \{q \in Q \mid \forall p \in X, pIq\}, B^* = \{p \in P \mid \forall q \in B, pIq\}$,若满足 $X^* = B, B^* = X$,则称 (X, B) 为形式背景 (P, Q, I) 上的一个形式概念(简称概念),称 X 为概念的外延, B 为概念的内涵. 所有的概念构成的集合叫概念格,记为 $L(P, Q, I)$. 一般地,对 $p \in P, q \in Q$,记 $\{p\}^* = p^*, \{q\}^* = q^*$.

若 (X_1, B_1) 和 (X_2, B_2) 是概念,其上的偏序关系定义为 $(X_1, B_1) \leq (X_2, B_2) \Leftrightarrow X_1 \subseteq X_2 (B_1 \supseteq B_2)$. 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 $L(P, Q, I)$ 是完备格, 称为 (P, Q, I) 的概念格.

1.3 知识空间与形式背景的联系

文献[26]建立了知识空间与形式背景的联系, 由知识空间 (Q, \mathcal{K}) 可构造出对应的反知识背景 (P, Q, I_c) 和知识背景 (P, Q, I) .

定义 6^[26] 设有限集 $P = \{p_1, p_2, \dots, p_n\}$, 其中, $p_i (1 \leq i \leq n)$ 是被测试的对象, Q 是问题集, I 是 P 和 Q 之间的二元关系, pIq 表示对象 p 不能解决问题 q , 称三元组 (P, Q, I) 是知识背景.

定义 7^[26] 设有限集 $P = \{p_1, p_2, \dots, p_n\}$, 其中, $p_i (1 \leq i \leq n)$ 是被测试的对象, Q 是问题集, I_c 是 P 和 Q 之间的二元关系, pI_cq 表示对象 p 能解决问题 q , 则称三元组 (P, Q, I_c) 是相对于 (P, Q, I) 的反知识背景.

由知识空间 (Q, \mathcal{K}) 构造知识背景 (P, Q, I) 的过程如下.

设 $\mathcal{K} = \{K_1, K_2, \dots, K_n\}$, 其中, $K_1 = \emptyset, K_n = Q, K_i (1 \leq i \leq n)$ 表示对象 p_i 对应的知识状态, 由此就将知识空间中的每一个状态和被测试的对象之间建立一一对应关系. 于是有引理 2.

引理 2^[26] 知识背景 (P, Q, I) 与形式背景 (\mathcal{K}, Q, \notin) 是同构的, 其中, \mathcal{K} 是知识状态集族, Q 是问题集, \notin 是 $\mathcal{K} \times Q$ 上的二元关系, 表示某个问题 $q \notin K$, 其中, $q \in Q, K \in \mathcal{K}$.

设知识背景为 (P, Q, I) , 且 $X \subseteq P, B \subseteq Q$, 算子 $X^* = \{q \in Q \mid \forall p \in X, pIq\}, B^* = \{p \in P \mid \forall q \in B, pIq\}$, 若 $X^* = B$ 并且 $B^* = X$, 则称二元组 (X, B) 是一个知识概念, 简称概念, X 是概念的外延, B 是概念的内涵.

1.4 布尔矩阵的运算规律

定义在 $B_0 = \{0, 1\}$ 上的向量称为布尔向量, 记 n 维布尔行向量为 \mathbf{V}_n , 记 n 维布尔列向量为 \mathbf{V}^n . 通常把分量全为 0 的向量 $(0, 0, \dots, 0)$ 或 $(0, 0, \dots, 0)^T$ 记作 $\mathbf{0}$, 称为布尔零向量 (简称零向量); 否则, 称为布尔非零向量 (简称非零向量). 记 $m \times n$ 阶布尔矩阵为 $\mathbf{B}_{m \times n}$.

设 $\boldsymbol{\alpha} = (x_1, x_2, \dots, x_n), \boldsymbol{\beta} = (y_1, y_2, \dots, y_n) \in \mathbf{V}_n$, 则定义两个运算“+”与“ \cdot ”为: 1) $\boldsymbol{\alpha} + \boldsymbol{\beta} = (x_1 + y_1, x_2 + y_2, \dots, x_n + y_n)$; 2) $\boldsymbol{\alpha} \cdot \boldsymbol{\beta} = (x_1 \cdot y_1, x_2 \cdot y_2, \dots, x_n \cdot y_n)$, 其中, “+”表示取最大值, “ \cdot ”表示数量乘积, 取最小值.

对于行向量 $\boldsymbol{\alpha}$ 和 $\boldsymbol{\beta}$, 对于 $i \in \{1, 2, \dots, n\}$, 若 $x_i = 1 \Rightarrow y_i = 1$, 则记 $\boldsymbol{\alpha} \leq \boldsymbol{\beta}$. 如果 $\boldsymbol{\alpha} \leq \boldsymbol{\beta}$ 且 $\boldsymbol{\alpha} \neq \boldsymbol{\beta}$, 则记 $\boldsymbol{\alpha} < \boldsymbol{\beta}$. 列向量相关内容可类似定义^[29].

定义 8^[29] 设子空间 $W \subseteq \mathbf{V}_n, W$ 的一个无关的子集 V 称为 W 的一个基, 当且仅当 $W = \langle V \rangle$ (向量集合 V 的生成空间).

设 $\mathbf{A} \in \mathbf{B}_{m \times n}, \mathbf{B}_R(\mathbf{A})$ 表示由 \mathbf{A} 的所有行构成的集合生成的空间 $R(\mathbf{A})$ 的唯一基底, 称为 \mathbf{A} 的行基底, 其基数称为 \mathbf{A} 的行秩, 用 $\rho_R(\mathbf{A})$ 表示. 类似地, $\mathbf{B}_C(\mathbf{A})$ 表示由 \mathbf{A} 的所有列构成的集合生成的空间 $C(\mathbf{A})$ 的唯一基底, 称为 \mathbf{A} 的列基底, 其基数称为 \mathbf{A} 的列秩, 用 $\rho_C(\mathbf{A})$ 表示^[29].

设 $\mathbf{A} = (a_{i,j})_{m \times n}, \mathbf{B} = (b_{i,j})_{m \times n}, \mathbf{C} = (c_{i,j})_{n \times l}$ 为布尔矩阵, 用 \mathbf{A}^T 表示矩阵 \mathbf{A} 的转置. 规定矩阵运算律如下:

- 1) $\mathbf{A} \leq \mathbf{B}$ 当且仅当 $a_{i,j} \leq b_{i,j}, i = 1, 2, \dots, m; j = 1, 2, \dots, n$;
- 2) $\mathbf{A} \vee \mathbf{B} = (a_{i,j} \vee b_{i,j})_{m \times n}$;
- 3) $\mathbf{A} \wedge \mathbf{B} = (a_{i,j} \wedge b_{i,j})_{m \times n}$;
- 4) $\mathbf{A} - \mathbf{B} = (a_{i,j} \wedge (1 - b_{i,j}))_{m \times n}$;
- 5) $\sim \mathbf{A} = (1 - a_{i,j})_{m \times n}$;
- 6) $\mathbf{A} \cdot \mathbf{C} = (d_{i,j})_{m \times l}$, 其中, $d_{i,j} = \bigvee_{1 \leq k \leq n} (a_{i,k} \wedge c_{k,j})$, \vee 表示取最大值, \wedge 表示取最小值, 则称 $\mathbf{A} \cdot \mathbf{C}$ 是布尔矩阵 \mathbf{A} 与 \mathbf{C} 的乘积^[29].

设论域 $P = \{p_1, p_2, \dots, p_n\}, X \subseteq P$, 记子集 X 的特征向量函数 $\lambda_X = (\lambda_X(p_1), \lambda_X(p_2), \dots, \lambda_X(p_n))$, 其中, $\lambda_X(p_i) = \begin{cases} 1, & p_i \in X \\ 0, & p_i \notin X \end{cases}, i \in \{1, 2, \dots, n\}$.

定义 9^[17] 设 (P, Q, I) 是形式背景, $\mathbf{M}_I = (c_{i,j})_{n \times m}$ 为 I 的关系矩阵, 令 $\mathbf{M} = (m_{i,j})_{n \times n} = \sim(\mathbf{M}_I \cdot (\sim \mathbf{M}_I^T))$, 称 \mathbf{M} 为其对应的对象关系矩阵.

引理 3^[17] 设 (P, Q, I) 是形式背景,矩阵 M_I 为 I 的关系矩阵, M 为其对应的对象关系矩阵,则矩阵 M 的第 i 个行向量是集合 $\{p_j \mid p_i^* \subseteq p_j^*, j \in \{1, 2, \dots, n\}\}$ 的特征向量 $\lambda_{\{p_j \mid p_i^* \subseteq p_j^*, j \in \{1, 2, \dots, n\}\}}$.

2 原子的特征和知识基

知识基是知识空间的核心,是知识空间的最小生成组,可生成知识空间,而且知识基还能反应学生掌握的最基本的问题集族,为刻画知识空间和寻找学习路径提供依据.文献[4, 5, 28]对此作了一些研究.下面从知识状态、算子、布尔向量和布尔矩阵等角度讨论原子的特征和知识基.

2.1 基于知识状态的原子特征和知识基

定义 10^[4] 设 (Q, \mathcal{K}) 为知识空间,设 \mathcal{F} 是问题域 Q 的非空子集族, $q \in \bigcup \mathcal{F}$, \mathcal{F} 中包含 q 的最小集合,称为元素 q 的一个原子.

引理 4^[4] 如果知识空间 (Q, \mathcal{K}) 的知识基 \mathcal{B} 存在,则 \mathcal{B} 可由所有的原子组成的集族构成.

因此,寻找知识基的过程就是找出知识空间的所有问题的原子.

定理 1^[5] 设 (Q, \mathcal{K}) 为知识空间,设 \mathcal{F} 是问题域 Q 的任意非空子集族,知识状态 $K \in \mathcal{F}$ 是一个原子,当且仅当 $K \in \mathcal{F}$,其中, $K = \bigcup_{F \in \mathcal{F}} F$.

推论 1 设 (Q, \mathcal{K}) 为知识空间,设 \mathcal{F} 是问题域 Q 的非空子集族,若对知识状态 $K \in \mathcal{K}$,满足 $K = \bigcup_{F \in \mathcal{F}} F$,则 $F \subseteq K$.

推论 2 设 (Q, \mathcal{K}) 为知识空间,则知识状态 $K \in \mathcal{K}$ 是一个原子,当且仅当 $K \neq \bigcup_{\emptyset \neq F \subseteq K} F$.

证明:充分性.若 $K \in \mathcal{K}$ 是一个原子,设 $K = \bigcup_{\emptyset \neq F \subseteq K} F$.则对任意 $q \in K$,必存在 $F_q \neq \emptyset$,使 $q \in F_q \subset K$,则 K 不是 $q \in K$ 的原子,与条件矛盾,则 $K \neq \bigcup_{\emptyset \neq F \subseteq K} F$.

必要性.若 $K \neq \bigcup_{\emptyset \neq F \subseteq K} F$.取 $\mathcal{F} = \{F \mid \emptyset \neq F \subseteq K\} \cup \{K\}$,则仅当 $K \in \mathcal{F}$ 时,有 $K = \bigcup_{F \in \mathcal{F}} F$,由定理1得知识状态 $K \in \mathcal{K}$ 是一个原子.

推论 3 设 (Q, \mathcal{K}) 为知识空间,则知识状态 $K \in \mathcal{K}$ 不是一个原子,当且仅当 $K = \bigcup_{\emptyset \neq F \subseteq K} F$.

例 1 在知识空间

$$\mathcal{K} = \{\emptyset, \{a\}, \{a, b\}, \{b, c\}, \{a, b, c\}\}$$

中, $\{a, b, c\} = \{a\} \cup \{a, b\} \cup \{b, c\}$,则由推论3得 $\{a, b, c\}$ 不是原子.而 $\{a, b\}$ 不能由其非空真子集的并表示,由推论2得 $\{a, b\}$ 是原子;同理, $\{a\}, \{b, c\}$ 都是原子.因此,该知识空间的知识基 $\mathcal{B} = \{\{a\}, \{a, b\}, \{b, c\}\}$.或者由Dowling^[4]算法也可得出,问题 a 只有一个原子 $\{a\}$,问题 b 有两个原子 $\{a, b\}$ 和 $\{b, c\}$,问题 c 的原子为 $\{b, c\}$,该算法^[4]的计算复杂度为 $O(|\mathcal{K}|^2 |Q|)$.

2.2 在反知识背景中的原子特征和知识基

2.2.1 基于算子的原子特征和知识基 在知识空间 (Q, \mathcal{K}) 及其反知识背景 (P, Q, I_c) 中, $K_i \in \mathcal{K}$ 为对象 $p_i \in P$ 的知识状态,且满足 $K_i = p_i^*$.由知识空间的知识状态满足并封闭,则 $L_j(P, Q, I_c) = \{p^* \mid p \in P\} \setminus \{\emptyset\}$ 也满足并封闭,找出 $L_j(P, Q, I_c)$ 中最小的并式生成组,即知识基.下面定义并可约元,并研究其相关性质.

定义 11 设 (P, Q, I_c) 为反知识背景, $p \in P$,若满足

$$p^* = \bigcup_{p_i^* \neq p^*, p_i^* \neq \emptyset} p_i^*,$$

则称 p 为并可约元;否则,称 p 为并不可约元.

性质 1 若 $p^* = \bigcup p_i^*$,则 $p_i^* \subseteq p^*$.

由知识空间 (Q, \mathcal{K}) 满足并封闭,可得性质2.

性质 2 设 (Q, \mathcal{K}) 为知识空间, (P, Q, I_c) 为其反知识背景.在 (P, Q, I_c) 中,对 $\forall p_i, p_j \in P, \exists p_k \in P$,使得 $p_k^* = p_i^* \cup p_j^*$.

性质 3 设 (Q, \mathcal{K}) 为知识空间, (P, Q, I_c) 为其反知识背景.在 (P, Q, I_c) 中,对于 $p \in P$,若存在 $p_i, p_j \in P$,满足 $\emptyset \neq p_i^* \subset p^*, \emptyset \neq p_j^* \subset p^*$,使 $p^* = p_i^* \cup p_j^*$,则 $p^* = \bigcup_{p_i \in L(p)} p_i^*$,其中, $L(p) = \{p_i \in P \mid \emptyset \neq p_i^* \subset p^*\}$.

若 p^* 可由两个非空真子集 p_1^*, p_2^* 并运算得到,则 p^* 可由所有非空真子集 $p_i^* \subset p^*$ 并运算得到,

则由定义 11 得 p 为并可约元.

定理 2 设 (Q, \mathcal{K}) 为知识空间, (P, Q, I_c) 为其反知识背景, 知识状态 $\emptyset \neq K \in \mathcal{K}$ 与 $p \in P$ 对应. 则

- 1) 知识状态 K 不是一个原子, 当且仅当 p 是并可约元;
- 2) 知识状态 K 是一个原子, 当且仅当 p 是并不可约元.

由推论 2 和推论 3 易证.

推论 4 设 $L_J(P, Q, I_c) = \{p^* \mid p \in P\} \setminus \{\emptyset\}$, $O_J(P, Q, I_c) = \{p^* \mid p \text{ 是并可约元}\}$, 则 $L_J(P, Q, I_c) \setminus O_J(P, Q, I_c)$ 是原子的全体, 即为知识空间 (Q, \mathcal{K}) 的知识基.

例 2 由知识空间

$$\mathcal{K} = \{\emptyset, \{a\}, \{a, b\}, \{b, c\}, \{a, b, c\}\}$$

可导出反知识背景, 如表 1 所示. 因为 $p_5^* = \bigcup p_i^*$, $p_i^* \subset p_5^*$, $i = 2, 3, 4$, 所以, p_5 是并可约元, 故 $K_5 = p_5^* = \{a, b, c\}$ 不是原子. 又 $p_2^* = \{a\} \neq \bigcup_{\emptyset \neq p_i^* \subset p_2^*} p_i^*$, 由定理 2 得 p_2^* 是原子, 同理, p_3^*, p_4^* 也都是原子. 因此, $O_J(P, Q, I_c) = \{p_5^*\}$, 且 $L_J(P, Q, I_c) \setminus O_J(P, Q, I_c) = \{p_2^*, p_3^*, p_4^*\}$ 是所有原子的全体, 即为知识基.

表 1 由 \mathcal{K} 导出的反知识背景
Tab. 1 Anti-knowledge context derived by \mathcal{K}

P	a	b	c
p_1	0	0	0
p_2	1	0	0
p_3	1	1	0
p_4	0	1	1
p_5	1	1	1

2.2.2 基于布尔向量的原子特征和知识基 设 (Q, \mathcal{K}) 为知识空间, $(P,$

$Q, I_c)$ 为其反知识背景, 若记其关系矩阵 M_{I_c} , 则知识空间 (Q, \mathcal{K}) 中的 $K_i \in \mathcal{K}$ 为对象 $p_i \in P$ 的知识状态, 与 M_{I_c} 中的第 i 个行向量 α_i 对应, 即 $\alpha_i = M_{I_c}(i, :) = \lambda_{p_i^*}$. 记 $R_{I_c} = \{\alpha_i \mid \alpha_i \text{ 是对应于 } p_i \in P \text{ 的行向量}\} \setminus \{\mathbf{0}\}$, 则可得性质 4~6.

性质 4 设 (Q, \mathcal{K}) 为知识空间, (P, Q, I_c) 为其反知识背景. 对于 $\forall \alpha_i, \alpha_j \in R_{I_c}$, 则 $\exists \alpha_k \in R_{I_c}$, 得 $\alpha_k = \alpha_i + \alpha_j$.

由定义 2 可知, 知识空间 (Q, \mathcal{K}) 满足 \mathcal{K} 对有限并封闭, 易得性质 4, 则 R_{I_c} 满足对加法封闭. 找出集合 R_{I_c} 的一个基, 就可得到知识空间的原子和知识基.

性质 5 设 (Q, \mathcal{K}) 为知识空间, (P, Q, I_c) 为其反知识背景. 对于 $\alpha \in R_{I_c}$, 假设 $\alpha = \sum_{\alpha_i \in R_{I_c}} \alpha_i$, 则 $\alpha \geq \alpha_i$.

性质 6 设 (Q, \mathcal{K}) 为知识空间, (P, Q, I_c) 为其反知识背景. 对 $\alpha_i \in R_{I_c}$, 若 $\exists \alpha_j, \alpha_k \in R_{I_c}$, 且 $\alpha_i > \alpha_j > \mathbf{0}, \alpha_i > \alpha_k > \mathbf{0}$, 使得 $\alpha_i = \alpha_j + \alpha_k$, 则 $\alpha_i = \sum_{\alpha_l \neq \alpha_i, \alpha_l \in R_{I_c}} k_l \alpha_l$, 其中, $k_l = \begin{cases} 1, & \alpha_i > \alpha_l \\ 0, & \text{其他} \end{cases}$.

定义 12 在反知识背景 (P, Q, I_c) 中, 对 $\alpha_i \in R_{I_c}$, 若 $\alpha_i = \sum_{\alpha_l \neq \alpha_i, \alpha_l \in R_{I_c}} k_l \alpha_l$, 其中, $k_l = \begin{cases} 1, & \alpha_i > \alpha_l \\ 0, & \text{其他} \end{cases}$, 则称 α_i 为并可约元; 否则, 称 α_i 为并不可约元.

定理 3 设 (Q, \mathcal{K}) 为知识空间, (P, Q, I_c) 为其反知识背景, 知识状态 $\emptyset \neq K \in \mathcal{K}$ 与 $\alpha \neq \mathbf{0} \in R_{I_c}$ 对应. 则

- 1) 知识状态 K 不是一个原子, 当且仅当 α 是并可约元;
- 2) 知识状态 K 是一个原子, 当且仅当 α 是并不可约元.

推论 5 设 $R_J = \{\alpha \mid \alpha \in R_{I_c} \text{ 是并可约元}\}$, 则由 $R_{I_c} \setminus R_J$ 可找出所有原子, 即知识空间 (Q, \mathcal{K}) 的知识基.

例 3 续例 2, 由表 1 可得: $\alpha_5 = \alpha_2 + \alpha_3 + \alpha_4$, 可得出 α_5 为并可约元. 通过计算得 $R_J = \{\alpha_5\}$, 则 $R_{I_c} \setminus R_J = \{\alpha_2, \alpha_3, \alpha_4\}$, 从而得到所有原子及其知识基为 $\mathcal{B} = \{p_2^*, p_3^*, p_4^*\}$.

定理 4 设 (Q, \mathcal{K}) 为知识空间, (P, Q, I_c) 为其反知识背景及其关系矩阵 M_{I_c} , 则知识基所含元素的个数(即原子的总个数)为矩阵 M_{I_c} 的行秩.

证明: 设 $R_{I_c} = \{\alpha_i \mid \alpha_i \text{ 是对应于 } p_i \in P \text{ 的行向量}\} \setminus \{\mathbf{0}\}$, $R_J = \{\alpha \mid \alpha \in R_{I_c} \text{ 是并可约元}\}$, $R_B = R_{I_c} \setminus R_J$. 对于 $\forall \alpha_i \in R_B$, 有 $\alpha_i \neq \sum_{\alpha_j \neq \alpha_i, \alpha_j \in R_{I_c}} k_j \alpha_j$, 其中, $k_j = \begin{cases} 1, & \alpha_i > \alpha_j \\ 0, & \text{其他} \end{cases}$; 否则, 若 $\alpha_i =$

$\sum_{\alpha_j \neq \alpha_i, \alpha_j \in R_{I_c}} k_j \alpha_j$, 其中, $k_j = \begin{cases} 1, \alpha_i > \alpha_j \\ 0, \text{其他} \end{cases}$, 则 α_i 为并可约元. 进一步, $\alpha_i \neq \sum_{\alpha_j \neq \alpha_i, \alpha_j \in R_B} k_j \alpha_j$, 其中, $k_j = \begin{cases} 1, \alpha_i > \alpha_j \\ 0, \text{其他} \end{cases}$, 则 R_B 是一个无关的集合^[29]. 又对于 $\forall \alpha_i \in R_J \cup \{\emptyset\}$, 有 $\alpha_i = \sum_{\alpha_j \neq \alpha_i, \alpha_j \in R_B} k_j \alpha_j$, 则 $R_{I_c} \cup \{\emptyset\} = \langle R_{I_c} \rangle = \langle R_B \rangle$. 因此, R_B 是 R_{I_c} 的基, 则 R_{I_c} 的秩为 $|R_B|$. 由定理 3 可知, R_B 中的元素对应的知识状态是原子, 可组成知识空间的知识基, 所以, 知识基所含元素的个数为关系矩阵 M_{I_c} 的行秩.

例 4 续例 3, 由例 3 可得 $R_{I_c} = \{\alpha_2, \alpha_3, \alpha_4, \alpha_5\}$, $R_J = \{\alpha_5\}$, 则 $R_B = R_{I_c} \setminus R_J = \{\alpha_2, \alpha_3, \alpha_4\}$. 容易验证, 对于 $\forall \alpha \in R_B$, 都有 $\alpha \neq \sum_{\alpha_j \neq \alpha, \alpha_j \in R_B} k_j \alpha_j$, 其中, $k_j = \begin{cases} 1, \alpha > \alpha_j \\ 0, \text{其他} \end{cases}$, 则 R_B 是一个无关的集合. 又 $\alpha_5 = \alpha_2 + \alpha_3 + \alpha_4$, 且 $\emptyset = 0\alpha_2 + 0\alpha_3 + 0\alpha_4$, 则 $R_{I_c} \cup \{\emptyset\} = \langle R_{I_c} \rangle = \langle R_B \rangle$, 所以, R_B 是 R_{I_c} 的基, 则 R_{I_c} 的秩为 $|R_B| = 3$. 由定理 3 可知, R_B 中的元素对应的知识状态 $\{p_2^*, p_3^*, p_4^*\}$ 都是原子, 则知识基为 $\mathcal{B} = \{p_2^*, p_3^*, p_4^*\}$, 且知识基所含元素的个数为关系矩阵 M_{I_c} 的行秩.

2.2.3 基于布尔矩阵的原子特征和知识基 由知识空间 (Q, \mathcal{H}) 满足并封闭, 则其对应的反知识背景 (P, Q, I_c) 中, $\{p^* | p \in P\}$ 也满足并封闭. 此时, 反知识背景不是节 1.2 意义上的形式背景, 但仍然沿用其中的一些名称和记号, 例如, 关系矩阵和对象关系矩阵等. 从布尔矩阵的角度获取原子的特征, 有两种方式. 结合定义 12 和定理 3, 给出第一种矩阵运算方法, 具体过程如下.

定义 13 设 (Q, \mathcal{H}) 为知识空间, (P, Q, I_c) 为其反知识背景及其关系矩阵 M_{I_c} , $M = (m_{i,j})_{n \times n}$ 为其对应的对象关系矩阵. 令 $G = (k_{i,j})_{n \times n} = M \wedge (M - M^T)$, 称 G 为其对应的知识状态真包含关系矩阵. 设 α_i 是在 M_{I_c} 中对应于 $p_i \in P$ 的行向量, 不妨设 $\alpha_1 = \emptyset$, 对应于知识状态 $K_1 = p_1^* = \emptyset$. 设 $\alpha(p_i) = \sum_{\alpha_j \neq \alpha_i, \alpha_j \in R_{I_c}} k_l \alpha_l = k_1 \alpha_1 + \dots + k_{i-1} \alpha_{i-1} + k_{i+1} \alpha_{i+1} + \dots + k_n \alpha_n$, 其中 $k_l = \begin{cases} 1, \alpha_i > \alpha_l \\ 0, \text{其他} \end{cases}$. 令 $S = (s_{i,j})_{n \times m} = G^T \cdot M_{I_c}$, 称 S 为其对应的知识状态线性表示矩阵. 令 $S_d = (S_d(i))_{1 \times n} = (\sim (M_{I_c}(i, :)) \cdot (\sim S(i, :))^T)$, 称 S_d 为其对应的原子的特征矩阵.

定理 5 设 (Q, \mathcal{H}) 为知识空间, (P, Q, I_c) 为其反知识背景, G 为其对应的知识状态真包含关系矩阵, S 为其对应的知识状态线性表示矩阵, S_d 为其对应的原子的特征矩阵, 则

- 1) G 的第 i 个行向量是集合 $\{p_j | p_i^* \subset p_j^*, j \in \{1, 2, \dots, n\}\}$ 的特征向量 $\lambda_{\{p_j | p_i^* \subset p_j^*, j \in \{1, 2, \dots, n\}\}}$;
- 2) S 中的第 i 个行向量为 $\alpha(p_i)$;
- 3) S_d 为集合 $\{p_i | \alpha_i = \alpha(p_i), i \in \{1, 2, \dots, n\}\}$ 的特征向量 $\lambda_{\{p_i | \alpha_i = \alpha(p_i), i \in \{1, 2, \dots, n\}\}}$. 当 $S_d(i) = 1 (i \neq 1)$ 时, 即 $M_{I_c}(i, :) = S(i, :)$ 时, p_i^* 是一个并可约元, 则 p_i^* 不是原子; 当 $S_d(i) = 0 (i \neq 1)$ 时, 即 $M_{I_c}(i, :) \neq S(i, :)$ 时, p_i^* 不是并可约元, 则 p_i^* 是一个原子.

证明: 1) 令 $G = M \wedge (M - M^T)$, 若 $p_i \in \{p_j | p_i^* \subset p_j^*\}$, 即 $p_i^* \subset p_i^*$, 有 $p_i \in p_i^{**} \subseteq p_i^{**}$, 则 $m_{i,i} = 1$, 但 $m_{i,i} = 0$ (反证, 假设 $m_{i,i} \neq 0$, 即 $p_i \in p_i^{**}$, 则有 $p_i^* = p_i^{**} \subseteq p_i^{**}$, 与 $p_i^* \subset p_i^*$ 相矛盾). 所以, $m_{i,i} \wedge (m_{i,i} \wedge (1 - m_{i,i})) = 1$, 有 $p_i \in \{p_j | p_j \in P, g_{i,j} = 1\}$, 则矩阵 G 的第 i 个行向量是集合 $\{p_j | p_i^* \subset p_j^*, j \in \{1, 2, \dots, n\}\}$ 的特征向量 $\lambda_{\{p_j | p_i^* \subset p_j^*, j \in \{1, 2, \dots, n\}\}}$.

2) 因为 $k_l = \begin{cases} 1, \alpha_i > \alpha_l \\ 0, \text{其他} \end{cases} = \begin{cases} 1, p_i^* \supset p_l^* \\ 0, \text{其他} \end{cases} = G^T(i, l), i, l \in \{1, 2, \dots, n\}$, 则 $(k_1, \dots, k_{i-1}, 0, k_{i+1}, \dots, k_n) = G^T(i, :)$. 所以, $\alpha(p_i) = k_1 \alpha_1 + \dots + k_{i-1} \alpha_{i-1} + k_{i+1} \alpha_{i+1} + \dots + k_n \alpha_n = (k_1, \dots, k_{i-1}, 0, k_{i+1}, \dots, k_n) \cdot (\alpha_1^T, \dots, \alpha_{i-1}^T, \alpha_i^T, \alpha_{i+1}^T, \dots, \alpha_n^T)^T = G^T(i, :) \cdot M_{I_c}$, 即 $S(i, :) = \alpha(p_i)$.

3) 因为 $S(i, :) = \alpha(p_i) = k_2 \alpha_2 + \dots + k_{i-1} \alpha_{i-1} + k_{i+1} \alpha_{i+1} + \dots + k_n \alpha_n$, 其中, $k_l = \begin{cases} 1, \alpha_i > \alpha_l \\ 0, \text{其他} \end{cases}, i, l \in \{1, 2, \dots, n\}$, 显然有 $S(i, :) \leq \alpha_i$. 若 $S_d(i) = \sim(\alpha_i \cdot (\sim S(i, :))^T) = \sim(M_{I_c}(i, :) \cdot (\sim S(i, :))^T) = 1$, 则有 $S(i, :) \geq \alpha_i$, 从而 $S(i, :) = \alpha_i$, 即 $\alpha_i = \alpha(p_i)$, 又当 $i \neq 1$ 时, $p_i^* \neq \emptyset$, 则 p_i^* 是一个并可约元, p_i^* 不是原子. 若 $S_d(i) = 0$, 则 $S(i, :) \geq \alpha_i$ 不成立, 又 $S(i, :) \leq \alpha_i$, 从而 $S(i, :) < \alpha_i$, 所以 $\alpha_i \neq \alpha(p_i)$, 又当 $i \neq 1$ 时, $p_i^* \neq \emptyset$, 则 p_i^* 不是并可约元, p_i^* 是一个原子, 且 S_d 为集合 $\{p_i | \alpha_i = \alpha(p_i), i \in \{1, 2, \dots, n\}\}$ 的特征向

量 $\lambda_{\{p_i \mid \alpha_i = \alpha(p_i), i \in \{1, 2, \cdots, n\}\}}$.

定理 5 说明,若 $S_d(i) = 1(i \neq 1)$,则 $\alpha_i = \sum_{\alpha_l \neq \alpha_i, \alpha_l \in R_{I_c}} k_l \alpha_l$,其中, $k_l = \begin{cases} 1, \alpha_i > \alpha_l \\ 0, \text{其他} \end{cases}$,则 α_i 是并可约元,不是原子,且 α_i 可由 $\alpha_1, \cdots, \alpha_{i-1}, \alpha_{i+1}, \cdots, \alpha_n$ 加法运算表示,其中表示的系数为 $G^T(i, :)$.

例 5 续例 2,通过定理 5,计算可得 $G = \begin{pmatrix} 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$, $S_d = (1, 0, 0, 0, 1)$. 由于 $S_d(1) = 1$ 对应

$\alpha_1 = \theta$,对应于知识状态 $K_1 = \emptyset$,而 $\emptyset \notin \mathcal{B}$,故不加以讨论. 由 $S_d(5) = 1$ 知 α_5 是一个并可约元,且 α_5 可由 $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ 加法运算表示,其中表示的系数为 $G^T(5, :)$,即 $\alpha_5 = \alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 = \alpha_2 + \alpha_3 + \alpha_4$. 由 $S_d(2) = S_d(3) = S_d(4) = 0$ 得并不可约元为 $\alpha_2, \alpha_3, \alpha_4$,可找出所有原子为 p_2^*, p_3^*, p_4^* ,则知识基为

$$\mathcal{B} = \{p_2^*, p_3^*, p_4^*\} = \{\{a\}, \{a, b\}, \{b, c\}\}.$$

算法 1 在反知识背景 (P, Q, I_c) 中获取原子和知识基的矩阵算法 1

输入:反知识背景 (P, Q, I_c) ,二元关系矩阵 M_{I_c} ,并且 $|P| = n, |Q| = m$.

输出:原子和知识基.

步骤 1: 计算对象关系矩阵 $M_{n \times n}$.

步骤 2: 计算矩阵 S 和 S_d .

步骤 3: 输出原子和知识基: $\mathcal{B} = \{p_i^* \mid S_d(i) = 0, p_i \in P\}$.

该方法通过布尔矩阵运算可得到所有并可约元及其表示的系数矩阵,同时,也可得到并不可约元,找到所有原子和知识基. 若 $p_i^* \in \mathcal{B}$,则可知 p_i^* 是原子,但无法判别 p_i^* 是哪个问题 q 的原子.

因为问题 q 的原子是在包含关系下的状态集族的极小集,而这种关系也可用布尔矩阵运算实现,即第二种矩阵方法,具体过程如下.

定义 14 设 (Q, \mathcal{H}) 为知识空间, (P, Q, I_c) 为其反知识背景及关系矩阵 M_{I_c} , G 为其对应的知识状态真包含关系矩阵. 称 $C^j = (\alpha^j, \alpha^j, \cdots, \alpha^j)_{n \times n}$ 为复制矩阵,其中, $\alpha^j = M_{I_c}(:, j), j \in \{1, 2, \cdots, m\}$. 称 $C_{\min}^j = (\sim C^j) \vee (\sim G)$ 为知识状态集族的极小运算矩阵. 称 $\alpha_{\min}^j = ((\alpha^j)^T \wedge C_{\min}^j(1, :)) \wedge ((\alpha^j)^T \wedge C_{\min}^j(2, :)) \wedge \cdots \wedge ((\alpha^j)^T \wedge C_{\min}^j(n, :))$ 为知识状态集族的极小向量,其中, n 为知识状态的个数.

定理 6 设 (Q, \mathcal{H}) 为知识空间, (P, Q, I_c) 为其反知识背景, G 为其对应的知识状态真包含关系矩阵, C^j 为复制矩阵, C_{\min}^j 为知识状态集族的极小运算矩阵, α_{\min}^j 为知识状态集族的极小向量,则 α_{\min}^j 为集合 $q_j^* \setminus \cup_l \{p_l \mid p_l^* \subset p_i^*, p_i, p_l \in q_j^*\}$ 的特征向量 $\lambda_{q_j^* \setminus \cup_l \{p_l \mid p_l^* \subset p_i^*, p_i, p_l \in q_j^*\}}$,其中, $i, l \in \{1, 2, \cdots, n\}, j \in \{1, 2, \cdots, m\}$.

证明:由定理 5 的证明过程,可知 G 的第 l 个行向量为集合 $\{p_i \mid p_l^* \subset p_i^*, p_l \in q_j^*\}$ 的特征向量 $\lambda_{\{p_i \mid p_l^* \subset p_i^*, p_l \in q_j^*\}}$.

矩阵 $C^j = (\alpha^j, \alpha^j, \cdots, \alpha^j)_{n \times n}$ 的第 l 个行向量为 $\alpha^j(l)$ 对应的复制向量. 当 $\alpha^j(l) = 1$ 时, $p_l \in q_j^*$,且 $C_{\min}^j = (\sim C^j) \vee (\sim G)$ 的第 l 个行向量为集合 $\sim \{p_i \mid p_l^* \subset p_i^*, p_l \in q_j^*\}$ 的特征向量 $\lambda_{\sim \{p_i \mid p_l^* \subset p_i^*, p_l \in q_j^*\}}$,所以 $(\alpha^j)^T \wedge C_{\min}^j(l, :)$ 为集合 $q_j^* \cap \{\sim \{p_i \mid p_l^* \subset p_i^*, p_i, p_l \in q_j^*\}\} = q_j^* \setminus \{p_i \mid p_l^* \subset p_i^*, p_i, p_l \in q_j^*\}$ 对应的特征向量 $\lambda_{q_j^* \setminus \{p_i \mid p_l^* \subset p_i^*, p_i, p_l \in q_j^*\}}$. 当 $\alpha^j(l) = 0$ 时, $C_{\min}^j = (\sim C^j) \vee (\sim G)$ 的第 l 个行向量的分量全为 1,所以, $(\alpha^j)^T \wedge C_{\min}^j(l, :)$ 仍为 $(\alpha^j)^T$,即为 q_j^* 对应特征向量 $\lambda_{q_j^*}$. 因此, $\alpha_{\min}^j = ((\alpha^j)^T \wedge C_{\min}^j(1, :)) \wedge ((\alpha^j)^T \wedge C_{\min}^j(2, :)) \wedge \cdots \wedge ((\alpha^j)^T \wedge C_{\min}^j(n, :))$ 为集合 $q_j^* \setminus \cup_l \{p_i \mid p_l^* \subset p_i^*, p_i, p_l \in q_j^*\}$ 的特征向量 $\lambda_{q_j^* \setminus \cup_l \{p_i \mid p_l^* \subset p_i^*, p_i, p_l \in q_j^*\}}$.

定理 6 说明,可通过 $\{p_i \mid p_l^* \subset p_i^*, p_i, p_l \in q_j^*\}$ 找到所有含有问题 q_j 的状态 $p_l^* (p_l \in q_j^*)$,再利用 $q_j^* \setminus \cup_l \{p_i \mid p_l^* \subset p_i^*, p_i, p_l \in q_j^*\}$ 删除包含关系下的较大的知识状态集,得出极小向量 α_{\min}^j ,从而获得含有问题 q_j 的极小知识状态集,即含有 q_j 的原子.

例 6 续例 2,由表 1 可得 $\alpha^1 = (0, 1, 1, 0, 1)^T$,利用定理 6 进行运算得 $\alpha_{\min}^1 = (0, 1, 0, 0, 0)$,即含有 a

的原子为 $p_2^* = \{a\}$. 进一步地, 利用定理 6 把表 1 中的原始二元关系更新为极小关系矩阵, 如表 2 所示.

由表 2 可知: 含有含有 b 的原子为 p_3^*, p_4^* , 含有 c 的原子为 p_4^* , 从而可得该知识空间的知识基 $\mathcal{B} = \{p_2^*, p_3^*, p_4^*\}$.

算法 2 在反知识背景 (P, Q, I_c) 中获取原子和知识基的矩阵算法 2

输入: 形式背景 (P, Q, I_c) , 二元关系矩阵 \mathbf{M}_{I_c} , 并且 $|P| = n$,
 $|Q| = m$.

输出: 形式背景的所有原子和知识基.

步骤 1: 计算对象关系矩阵 $\mathbf{M}_{n \times n}$.

步骤 2: 计算知识状态集族的极小集

for $j = 1: |Q|$

给出 $\alpha^j = \mathbf{M}_{I_c}(:, j), C^j$; 计算 $\mathbf{C}_{\min}^j, \alpha_{\min}^j, \mathbf{M}_{I_c}(:, j) = (\alpha_{\min}^j)^T$.

end for

$\mathbf{M}_{\min} = \mathbf{M}_{I_c}$.

步骤 3: 输出所有原子和知识基 $\mathcal{B} = \{p_i^* \mid \mathbf{M}_{\min}(i, j) = 1, p_i \in P, q_j \in Q\}$.

2.3 在知识背景中的原子特征和知识基

设 (Q, \mathcal{K}) 为知识空间, (P, Q, I) 为其知识背景. 由于知识背景 (P, Q, I) 与形式背景 $(\mathcal{K}, Q, \subseteq)$ 是同构的, 因此, 知识空间 (Q, \mathcal{K}) 中的 $K_i \in \mathcal{K}$ 与对象 $p_i \in P$ 的知识状态的补集对应, 即 $K_i = Q \setminus p_i^*$. 由知识空间满足并封闭, 则在知识背景 (P, Q, I) 中 $\{p_i^* \mid p_i \in P\}$ 满足交封闭.

类似地, 在知识背景中也可从算子和布尔矩阵的角度刻画原子特征, 构建知识基.

2.3.1 基于算子的原子特征和知识基

定义 15 设 (P, Q, I) 为形式背景, $p \in P$, 若满足

$$p^* = \bigcap_{p_i^* \neq p^*, p_i^* \neq Q} p_i^*,$$

则称 p 为交可约元; 否则, 称 p 为交不可约元.

性质 7 若 $p^* = \bigcap p_i^*$, 则 $p^* \subseteq p_i^*$.

性质 8 设 (Q, \mathcal{K}) 为知识空间, (P, Q, I) 为其知识背景. 在 (P, Q, I) 中, 对 $p \in P$, 若存在 $p_i, p_j \in P$, 满足 $p^* \subset p_i^* \neq Q, p^* \subset p_j^* \neq Q$, 使得 $p^* = p_i^* \cap p_j^*$, 则 $p^* = \bigcap_{p_i \in H(p)} p_i^*$, 其中, $H(p) = \{p_i \in P \mid p_i^* \supset p^*, p_i^* \neq Q\}$. 由定义 15 可得, p 为交可约元.

定理 7 设 (Q, \mathcal{K}) 为知识空间, (P, Q, I) 为其知识背景, 知识状态 $\emptyset \neq K \in \mathcal{K}$ 与 $p \in P$ 对应. 则

- 1) 知识状态 K 不是一个原子, 当且仅当 p 是交可约元;
- 2) 知识状态 K 是一个原子, 当且仅当 p 是交不可约元.

推论 6 设 $L_M(P, Q, I) = \{p^* \mid p \in P\} \setminus \{Q\}, O_M(P, Q, I) = \{p^* \mid p \text{ 是交可约元}\}$, 则 $L_M(P, Q, I) \setminus O_M(P, Q, I)$ 是原子的全体, 即为知识空间 (Q, \mathcal{K}) 的知识基.

2.3.2 基于布尔矩阵的原子特征和知识基

在知识背景 (P, Q, I) 中, 可通过算子 “ \swarrow ”, “ \nearrow ” 和 “ \nearrow ” 获取原子^[26].

定义 16^[26] 在知识背景 (P, Q, I) 中, 定义算子 “ \swarrow ”, “ \nearrow ” 和 “ \nearrow ”. 其中, $p \swarrow q$ 表示 $p I_c q$ 并且 p^* 是不包含 q 的极大集; $p \nearrow q$ 表示 $p I_c q$ 并且 q^* 是不包含 p 的极大集; $p \nearrow q$ 表示同时满足 $p \swarrow q$ 和 $p \nearrow q$.

定理 8 设 (Q, \mathcal{K}) 为知识空间, (P, Q, I) 为其知识背景. 在 (P, Q, I) 中, 每一个问题 $q \in Q$ 对应的 “ \swarrow ” 的知识状态就是它的原子.

证明: 若 $p \swarrow q$, 则 $p I_c q$, 且 p^* 是不包含 q 的极大集. 因此, $Q \setminus p^*$ 是包含 $q \in Q$ 的知识状态中的极小集. 故 $Q \setminus p^*$ 为 q 的一个原子.

定理 9^[26] 设 (Q, \mathcal{K}) 为知识空间, (P, Q, I) 为其知识背景. 在 (P, Q, I) 中, 每一个问题 $q \in Q$ 对应的 “ \nearrow ” 的知识状态是一个原子.

例 7 由知识空间

$$\mathcal{K} = \{\emptyset, \{a\}, \{a, b\}, \{b, c\}, \{a, b, c\}\}$$

表 2 由表 1 更新得到的极小关系矩阵

Tab. 2 Minimal matrix obtained by updating table 1

P	a	b	c
p_1	0	0	0
p_2	1	0	0
p_3	0	1	0
p_4	0	1	1
p_5	0	0	0

可导出知识背景,如表 3 所示.由表 3 可知:不包含问题 a 的有 p_2^*, p_3^*, p_5^* ,且 $p_2^* \supset p_3^* \supset p_5^*$,则 p_2^* 为不包含问题 a 的极大集,有 $p_2 \not\prec a$,则 p_2 对应的知识状态 $Q \setminus p_2^* = \{a\}$ 为问题 a 的原子;不包含 p_2 的只有 a^* ,则 $p_2 \nearrow a$,从而 $p_2 \nearrow a$.类似地,由 $p_4 \not\prec b$ 可得 $Q \setminus p_4^*$ 是问题 b 的原子,虽然此时不满足 $p_4 \nearrow b$.

因此,通过研究知识背景中的运算“ $\not\prec$ ”,“ \nearrow ”和“ \nearrow ”,可找出问题 q 的原子.由于 $p \not\prec q$ 表示 $pI_c q$,且 p^* 是不包含 q 的极大集,即不包含 q 的极大对象内涵的补集.为了实现“ $\not\prec$ ”的极大集运算,引入定义 17.

定义 17 设 (Q, \mathcal{K}) 为知识空间, (P, Q, I) 为其知识背景. $M_I = (c_{i,j})_{n \times m}$ 为 I 的关系矩阵, $M = (m_{i,j})_{n \times n}$ 为对应的对象关系矩阵.令 $T = (t_{i,j})_{n \times n} = [M \wedge (M - M^T)]^T$,则称 T 为其对应的对象真包含关系矩阵.称 $C^j = (\alpha^j, \alpha^j, \dots, \alpha^j)_{n \times n}$ 为复制矩阵,其中, $\alpha^j = M_I(:, j), j \in \{1, 2, \dots, m\}$.称 $C_{\max}^j = C^j \vee (\sim T)$ 为对象内涵的极大运算矩阵.称 $\alpha_{\max}^j = \sim(((\sim \alpha^j)^T \wedge C_{\max}^j(1, :)) \wedge ((\sim \alpha^j)^T \wedge C_{\max}^j(2, :)) \wedge \dots \wedge ((\sim \alpha^j)^T \wedge C_{\max}^j(n, :)))$ 为对象内涵的极大向量,其中, n 为对象个数.

定理 10 设 (Q, \mathcal{K}) 为知识空间, (P, Q, I) 为其知识背景. T 为其对象真包含关系矩阵, C^j 为复制矩阵, C_{\max}^j 为其对象内涵极大运算矩阵, α_{\max}^j 为对象内涵的极大向量,则

- 1) T 的第 l 个行向量为集合 $\{p_i | p_i^* \subset p_l^*\}$ 的特征向量 $\lambda_{\{p_i | p_i^* \subset p_l^*\}}$;
- 2) α_{\max}^j 为集合 $\sim\{(q_j^*)^c \setminus \bigcup_l \{p_i | p_i^* \subset p_l^*, p_i, p_l \notin q_j^*\}\}$ 的特征向量 $\lambda_{\sim\{(q_j^*)^c \setminus \bigcup_l \{p_i | p_i^* \subset p_l^*, p_i, p_l \notin q_j^*\}\}}$,其中, $i, l \in \{1, 2, \dots, n\}, j \in \{1, 2, \dots, m\}$.

定理 10 说明可以通过 $\bigcup_l \{p_i | p_i^* \subset p_l^*, p_i, p_l \notin q_j^*\}$ 找出不包含问题 q_j 的基数较小的知识状态,再由 $\sim\{(q_j^*)^c \setminus \bigcup_l \{p_i | p_i^* \subset p_l^*, p_i, p_l \notin q_j^*\}\}$ 删去不包含问题 q_j 的知识状态集中基数小的集合,获得极大向量 α_{\max}^j ,从而获得不包含问题 q_j 的极大集.

同理,也可以用矩阵运算实现算子“ \nearrow ”的极大集运算,在此不再进行描述.

例 8 Korossy^[26]选取了初等几何学中 with 毕达哥拉斯定理有关的 5 个问题,记为 $Q = \{1, 2, 3, 4, 5\}$,通过实验测试并分析得到知识空间 (Q, \mathcal{K}) ,其中,

$$\mathcal{K} = \{\emptyset, \{1\}, \{3\}, \{1, 2\}, \{1, 3\}, \{1, 4\}, \{2, 3\}, \{3, 4\}, \{1, 2, 3\}, \{1, 2, 4\}, \{1, 3, 4\}, \{2, 3, 4\}, \{1, 2, 3, 4\}, \{1, 2, 3, 5\}, \{1, 2, 3, 4, 5\}\}.$$

由该知识空间导出的知识背景,如表 4 所示.

表 4 由 \mathcal{K} 导出的知识背景

Tab. 4 Knowledge context derived by \mathcal{K}

P	\mathcal{K}					P	\mathcal{K}				
	1	2	3	4	5		1	2	3	4	5
p_1	1	1	1	1	1	p_9	0	0	0	1	1
p_2	0	1	1	1	1	p_{10}	0	0	1	0	1
p_3	1	1	0	1	1	p_{11}	0	1	0	0	1
p_4	0	0	1	1	1	p_{12}	1	0	0	0	1
p_5	0	1	0	1	1	p_{13}	0	0	0	0	1
p_6	0	1	1	0	1	p_{14}	0	0	0	1	0
p_7	1	0	0	1	1	p_{15}	0	0	0	0	0
p_8	1	1	0	0	1						

问题 1 的外延集 1^* 与 $\lambda_1^* = (\alpha^1)^T = (1, 0, 1, 0, 0, 0, 1, 1, 0, 0, 0, 1, 0, 0, 0)$ 对应,用定理 10 计算得 $\alpha_{\max}^1 = (1, 0, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1)$,则由 α_{\max}^1 可得 不包含问题 1 的极大集为 p_2^* ,即 $p_2 \not\prec 1$,所以, $Q \setminus p_2^* = \{1\}$ 是问题 1 的一个原子.

利用定理 10 对知识背景的关系矩阵 M_I 中的所有列进行对象内涵的极大运算,最后,可以得到的新的布尔矩阵,记为 M_{\max} . M_{\max} 中所有二元对(取值为零)对应着 $p \not\prec q$ 关系下的极大集,在二元对取值为零处的问题 q 对应的知识状态 $K = Q \setminus p^*$ 为问题 q 的原子,所有原子组成该知识背景所对应的知识空间的知识基.由 M_{\max} 可得出例 7 中知识空间的所有原子对应的二元对(取值为零)和原子,如表 5 所示.

所以,该知识空间 (Q,\mathcal{K}) 的知识基为

$$\mathcal{B} = \{\{1\},\{3\},\{1,2\},\{2,3\},\{1,4\},\{3,4\},\{1,2,3,5\}\}.$$

算法 3 在知识背景 (P,Q,I) 中获取原子和知识基的
矩阵算法

输入：知识背景 (P,Q,I) ,二元关系矩阵 \mathbf{M}_I ,并
且 $|P|=n,|Q|=m$.

输出：知识背景的所有原子和知识基.

步骤 1：计算对象关系矩阵 $\mathbf{M}_{n\times n}$.

步骤 2：计算对象内涵的极大集

for $j=1:|Q|$

给出 $\alpha^j=\mathbf{M}_I(:,j),C^j$;计算 $C_{\max}^j,\alpha_{\max}^j,\mathbf{M}_I(:,j)=(\alpha_{\max}^j)^T$.

end for

$\mathbf{M}_{\max}=\mathbf{M}_I$.

步骤 3：输出原子和知识基 $\mathcal{B}=\{Q\setminus p_i^* \mid \mathbf{M}_{\max}(i,j)=0,p_i\in P,q_j\in Q\}$.

文中提出的矩阵算法 1、算法 2 和算法 3 都可以得到知识空间的原子和知识基,计算的时间复杂均为 $O(|Q||\mathcal{K}|^2)$.

3 结束语

先建立知识空间与反知识背景、知识背景之间的关系,再从并不可约元、交不可约元、集合的包含关系等方面判定原子特征,进而从知识状态、算子、布尔向量和布尔矩阵等不同角度,获取原子和知识基.在 3 种矩阵算法中,算法 1 通过矩阵方法计算并不可约元,获得原子和知识基,该算法还可以得到并可约元的并式表达式;算法 2 用矩阵方法计算出集合在包含关系下的极小集,从而得到各个问题的原子,构建知识基;算法 3 用矩阵方法计算出集合在包含关系下的极大集,从而获得 $p\blacktriangleleft q$,进而得到各个问题的原子和知识基.文中提出的矩阵方法为知识空间的研究拓展了计算方法,后续可进一步研究形式背景与知识空间的关系,并将它们相互结合应用到教育教学等领域.

参考文献：

[1] DOIGNON J P,FALMAGNE J C. Spaces for the assessment of knowledge[J]. International Journal of Man-Machine Studies,1985,23(2):175-196. DOI:10. 1016/S0020-7373(85)80031-6.

[2] KOPPEN M,DOIGNON J P. How to build a knowledge space by querying an expert[J]. Journal of Mathematical Psychology,1990,34(3):311-331. DOI:10. 1016/0022-2496(90)90035-8.

[3] ALBERT D,HELD T. Establishing knowledge spaces by systematical problem construction[M]//ALBERT D. Knowledge structures. Berlin-Heidelberg:Springer,1994:81-115. DOI:10. 1007/978-3-642-52064-8.

[4] DOWLING C E. Applying the basis of a knowledge space for controlling the questioning of an expert[J]. Journal of Mathematical Psychology,1993,37(1):21-48. DOI:10. 1006/jmps. 1993. 1002.

[5] FALMAGNE J C,DOIGNON J P. Learning space: Interdisciplinary applied mathematics[M]. New York:Springer Verlag,2011. DOI:10. 1007/978-3-642-01039-2.

[6] ALBERT D,LUKAS J. Knowledge spaces: Theories, empirical research, and applications[M]. London:Psychology Press,1999. DOI:10. 4324/9781410602077.

[7] DOIGNON J P,FALMAGNE J C. Knowledge spaces[M]. New York:Springer-Verlag,1999. DOI:10. 1007/978-3-642-58625-5.

[8] COSYN E,DOBLE C,FALMAGNE J C,et al. Assessing mathematical knowledge in a learning space[M]//FALMAGNE J C,ALBERT D,DOBLE C,et al. Knowledge Spaces: Applications in education. New York:Springer,2013:27-50. DOI 10. 1007/978-3-642-35329-1.

[9] 麦裕华,何庆辉,肖信. 基于知识空间理论的高中生科学原理学习分析:以氧化还原反应为例[J]. 化学教育(中英文),2018,39(19):34-40. DOI:10. 13884/j. 1003-3807hxjy. 2017090006.

[10] 何庆辉,麦裕华. 基于知识空间理论的高一学生离子反应关键学习路径[J]. 化学教学,2018(7):12-17. DOI:10.3969/j. issn. 1005-6629. 2018. 07. 004.

[11] WILLE R. Restructuring lattice theory: An approach based on hierarchies of concepts[C]// International Conference on Formal Concept Analysis. Berlin;Springer,1982;314-339. DOI:10. 1007/978-94-009-7798-3_15.

[12] 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. DOI:10. 1360/112004-104.

[13] 徐伟华,李金海,魏玲,等. 形式概念分析理论与应用[M]. 北京:科学出版社,2016.

[14] 李进金,李克典,吴端恭. 基于粗糙集与概念格的知识系统模型[M]. 北京:科学出版社,2013.

[15] 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.

[16] 李进金,张燕兰,吴伟志,等. 形式背景与协调决策形式背景属性约简与概念格生成[J]. 计算机学报,2014,37(8):1768-1774. DOI:10. 3724/SP. J. 1016. 2014. 01768.

[17] 张清新. 基于布尔矩阵的决策形式背景协调集判断方法[J]. 漳州师范学院学报,2012,75(1):22-25. DOI:10.3969/j. issn. 1008-7826. 2012. 01. 004.

[18] BELOHLAVEK R,TRNECKA M. From-below approximations in boolean matrix factorization: Geometry and new algorithm[J]. Journal of Computer and System Sciences,2015,81:1678-1697. DOI:10. 1016/j. jcss. 2015. 06. 002.

[19] TRNECKA M,TRNECKOVA M. Data reduction for boolean matrix factorization algorithms based on formal concept analysis[J]. Knowledge-Based Systems,2018,158:75-80. DOI:10. 1016/j. knosys. 2018. 05. 035.

[20] 林艺东,李进金,张呈玲. 基于矩阵的模糊-经典概念格属性约简[J]. 模式识别与人工智能,2020,33(1):21-31. DOI:10. 16451/j. cnki. issn1003-6059. 202001003.

[21] WU Weizhi,LEUNG Y,MI Jusheng. Granular computing and knowledge reduction in formal contexts[J]. IEEE Transactions on Knowledge and Data Engineering,2009,21(10):1461-1474. DOI:10. 1109/TKDE. 2008. 223.

[22] 陈东晓,李进金. 形式背景的下近似协调与粒协调的关系[J]. 华侨大学学报(自然科学版),2020,41(1):130-136. DOI:10. 11830/ISSN. 1000-5013. 201907035.

[23] 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. DOI:10. 16451/j. cnki. issn1003-6059. 201806004.

[24] 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),2019,50(12):1817-1833. DOI:10. 1360/N112018-00272.

[25] 谢小贤,李进金,陈东晓,等. 基于布尔矩阵的保持二元关系不变的概念约简[J]. 山东大学学报(理学版),2020,55(5):32-45. DOI:10. 6040/j. issn. 1671-9352. c. 2020. 004.

[26] RUSCH A,WILLE R. Knowledge spaces and formal concept analysis[M]//BOCK H H,POLASEK W. Data analysis and information systems. Berlin-Heidelberg:Springer,1996:427-436. DOI:10. 1007/978-3-642-80098-6_36.

[27] SPOTO A,STEFANUTTI L,VIDOTTO G. Knowledge space theory, formal concept analysis, and computerized psychological assessment[J]. Behavior research methods,2010,42(1):342-350. DOI:10. 3758/BRM. 42. 1. 342.

[28] 李进金,孙文. 知识空间、形式背景和知识基[J]. 西北大学学报(自然科学版),2019,49(4):517-526. DOI:10. 16152/j. cnki. xdxbr. 2019-04-004.

[29] KIM K H. 布尔矩阵理论及其应用[M]. 何善培,等译. 北京:知识出版社,1987.

(责任编辑:黄晓楠 英文审校:黄心中)