

DOI: 10.13718/j.cnki.xdzk.2021.10.002

# 基于三支概念簇的知识表示

智慧来, 徐彤, 李逸楠

河南理工大学 计算机科学与技术学院, 河南 焦作 454003

**摘要:** 概念簇刻画了现代范畴理论中具有家族相似性的概念集合. 概念簇中的概念包含了对象具有的共同属性, 但忽略了对对象都不具有的属性, 不利于解决某些特定的问题. 本文融合现代范畴理论与三支概念分析, 提出了用三支概念簇刻画具有家族相似性的三支概念集合; 其次, 研究了三支概念簇的若干重要性质与运算方法; 最后, 将三支概念簇用于满足特定约束条件下的相似对象检索. 分析结果表明, 三支概念簇相比较于概念簇, 能够检索到更加符合需求的对象.

**关键词:** 三支概念分析; 粒计算; 概念簇; 三支概念簇; 知识表示

**中图分类号:** TP18      **文献标志码:** A

**文章编号:** 1673-9868(2021)10-0010-09

开放科学(资源服务)标识码(OSID):



## Knowledge Representation Based on Three-Way Concept Cluster

ZHI Huilai, XU Tong, LI Yinan

*School of Computer Science & Technology, Henan Polytechnic University, Jiaozuo Henan 454003, China*

**Abstract:** A concept cluster describes a set of concepts with family resemblance in modern category theory. For a given concept cluster, where the concepts contain the common attributes that are possessed by the objects, the attributes that are not commonly possessed by the objects are ignored. As a consequence, it may not be effective enough to solve some specific problems. This paper combines the modern category theory with three-way concept analysis and proposes three-way concept clusters to characterize the sets of three-way concepts with family resemblance. Then, some important properties and operational methods of three-way concept clusters are investigated. Finally, the three-way concept clusters are used to retrieve similar objects satisfying some specific constraints. The results show that compared with concept clusters, three-way concept clusters can retrieve more suitable objects.

**Key words:** three-way concept analysis; granular computing; concept cluster; three-way concept cluster; knowledge representation

收稿日期: 2021-05-21

基金项目: 国家自然科学基金项目(61502150); 河南省高校基本科研业务费专项(NSFRF210318).

作者简介: 智慧来, 副教授, 硕士生导师, 博士, 主要从事形式概念分析、粗糙集理论、粒计算等研究.

概念是人类认知活动的基本单位,以概念为基础的粒计算方法是实现可解释性人工智能的途径之一<sup>[1-5]</sup>.粒计算通过对复杂数据的多粒度化处理得以实现复杂问题的简易求解<sup>[6]</sup>.多粒度化的数据呈现出大小不同的信息粒,从而构成了由小到大的粒度链条<sup>[7]</sup>.在粒度链条中,小粒是构成大粒的基础,粒度的变化体现着思维过程的泛化和特化,在泛化和特化的双向思维驱动下,人们得以解决许多看似无法下手的困难问题<sup>[8-9]</sup>.

形式概念分析(formal concept analysis, FCA)<sup>[10]</sup>是基于格论建立起来的数学方法,已经成为一种最有效的粒计算理论.在应用的驱动下,根据实际需求可以通过各种方法构造不同类型的形式概念,包括经典形式概念、面向对象概念、面向属性概念、对偶概念等<sup>[11-14]</sup>.最近十年,粗糙集、粒计算、模糊集、三支决策理论等推动了形式概念分析的发展,产生了许多有重要理论价值与应用价值的成果<sup>[15-20]</sup>.其中, QI 等<sup>[21]</sup>将三支决策理论(three-way decisions, 3WD)中的三分思想引入形式概念分析,提出了三支概念分析(three-way concept analysis, 3WCA).具体就是通过形式背景上定义负算子,刻画对象子集与属性子集间“共同不具有”的关系,并构造三支算子,将对象集或属性集三分,形成三支概念与三支概念格.三支概念分析自 2014 年提出以来,得到了广泛的认可<sup>[22-24]</sup>.

信息检索一直是人工智能的重要研究课题,但是,现有的信息检索方法广泛采用距离度量相似程度,不能完全客观精准地对具有家族相似性的原型范畴进行刻画.如果能够采用粒计算思想,从概念的角度对整个粒度链条进行深入研究,特别是对粒度链条顶端的子格进行研究,或许可以得到满意的解决.

现代范畴理论<sup>[25-26]</sup>强调相似的传递性,使得看似不相关的对象通过相似的传递性联系在一起,形成一个特定的范畴,在同一范畴内,成员共享“家族相似性”.显然,采用距离度量相似程度不能准确地刻画这样的情形.为了解决这一问题,在现代范畴理论的启发下,智慧来与李逸楠首次提出了概念簇,用以刻画一组具有家族相似性的概念,并应用于相似对象的检索<sup>[27]</sup>.概念簇中的概念包含了对象具有的共同属性,但忽略了对象都不具有的属性,不利于解决某些特定的问题.

为了解决上述问题,本文融合现代范畴理论与三支概念分析,提出了用三支概念簇刻画具有家族相似性的三支概念集合,并对三支概念簇的性质、运算以及应用进行了初步的探讨.

## 1 基本理论

本节介绍形式概念分析<sup>[10]</sup>、概念簇<sup>[27]</sup>以及三支概念分析<sup>[21]</sup>的相关基本概念.

### 1.1 形式概念分析与概念簇

形式概念分析<sup>[10]</sup>建立在格论基础之上,对哲学中的概念进行了高度的数学概括,具有良好的性质,已经广泛应用于知识表示、数据挖掘、近似推理等诸多领域.形式背景是领域本体的抽象描述,也是形式概念分析研究的具体对象和起点.

具体地,形式背景可以抽象为一个三元组  $K = (G, M, I)$ ,其中,  $G$  为对象集合,  $M$  为属性集合,且均非空有限,  $I$  为  $G$  和  $M$  之间的二元关系.进一步,使用  $I(x, a) = 1$  表示对象  $x$  具有属性  $a$ ,使用  $I(x, a) = 0$  表示对象  $x$  不具有属性  $a$ .在下文中,为了方便起见,对于单元素集合,只列出其中的元素.例如,用  $x$  表示集合  $\{x\}$ .

对于  $X \subseteq G$  和  $A \subseteq M$ ,分别定义对象集和属性集上的算子如下:

$$X^* = \{a \in M \mid \forall x \in X, I(x, a) = 1\} \quad A^* = \{x \in G \mid \forall a \in A, I(x, a) = 1\}$$

若  $X^* = A$  且  $A^* = X$ ,则称  $(X, A)$  是形式背景  $K$  的一个概念.已经证明,形式背景  $K$  的全体概念构成一个完全格,并称之为概念格,记做  $L(K)$ .

现代范畴理论<sup>[25-26]</sup>着重分析了家族相似概念,从理论上阐述了在同一个范畴中有可能存在一些不具有共同属性的对象. Wittgenstein 通过深入了解许多不同类型的游戏,发现当中的一些游戏并不具有共同点,但是体现出了不可忽视的紧密联系.经过分析得知,这些游戏是通过相似的传递性而密切联系在一起,利用这种密切联系, Wittgenstein 将这些游戏划分成了多个家族.语言学家 Labov 和 Rosch 在深入研究自然

范畴的基础上,将“家族相似性”解释为家族成员具有某些相似之处,并将“原型范畴”(prototype category)解释为具有“家族相似性”的自然范畴,从而奠定了现代范畴理论的基础.

现代范畴理论的主要观点包括:

- 1) 范畴内的各个成员通过相似的传递性聚集在一起,所有成员不一定存在相同的约束;
- 2) 范畴中可能存在一个或多个中心,形成单一中心结构或多中心结构;
- 3) 对于一个范畴,最具典型性的成员位于范畴的中心.

现代范畴理论的主要观点在粒计算理论中已经有大量的体现.在现代范畴理论的启发下,智慧来等提出了概念簇,用以刻画具有某些特定相似之处的概念,并应用于基于概念格的相似对象检索.

**定义 1**<sup>[27]</sup> 给定一个形式背景  $K = (G, M, I)$ ,  $X \subseteq G$  和  $A \subseteq M$ . 若

$$C(X, A) = \{(Y, B) \in L(K) \mid Y \supseteq X, B \supseteq A\} \neq \emptyset$$

则称  $C(X, A)$  是  $K$  的一个概念簇.

文献[27]已经证明,概念簇本质上是概念格的子格,位于粒度链条的最顶端,使用概念簇代替子格,是为了强调概念之间的家族相似性,以便于概念格理论的应用.

## 1.2 三支概念分析

三支概念分析<sup>[21-22]</sup> 相比较于经典形式概念分析,刻画了对象共同不具有的属性,能够提供经典形式概念分析所不能提供的某些细节.为了构造三支概念,首先需要定义对象集和属性集上的负算子.

给定一个形式背景  $K = (G, M, I)$ ,  $X \subseteq G$ ,  $B \subseteq M$ , 对象集和属性集上的负算子分别定义如下:

$$X^{\bar{\cdot}} = \{a \in M \mid \forall x \in X, I(x, a) = 0\}, B^{\bar{\cdot}} = \{x \in G \mid \forall a \in B, I(x, a) = 0\}$$

**定义 2** 给定一个形式背景  $K = (G, M, I)$ ,  $X \subseteq G$ ,  $A, B \subseteq M$ , 定义对象导出三支算子如下:

$$X^{\triangleleft} = (X^{\bar{\cdot}}, X^{\bar{\cdot}}), (A, B)^{\triangleright} = \{x \in G \mid x \in A^{\bar{\cdot}} \text{ 且 } x \in B^{\bar{\cdot}}\} = A^{\bar{\cdot}} \cap B^{\bar{\cdot}}$$

进一步,若  $X^{\triangleleft} = (A, B)$  且  $(A, B)^{\triangleright} = X$ , 则称  $(X, (A, B))$  为对象导出三支概念,简称为 OE-概念.已经证明,形式背景  $K$  的全体对象导出三支概念构成一个完全格,并称之为对象导出三支概念格,记做  $OEL(K)$ .

在  $OEL(K)$  中存在一类对象诱导的三支概念,两类属性诱导的三支概念,这些概念在基于三支概念分析的应用中发挥着重要作用,具体定义如下.

**定义 3** 给定一个形式背景  $K = (G, M, I)$ ,  $x \in G$ ,  $a \in M$ . 称  $(x^{\triangleleft \triangleright}, x^{\triangleleft})$  为对象诱导的三支概念,称  $(a^{\bar{\cdot}}, a^{\bar{\cdot} \triangleleft})$  为 I 型属性诱导的三支概念,称  $(a^{\bar{\cdot}}, a^{\bar{\cdot} \triangleleft})$  为 II 型属性诱导的三支概念.

**例 1** 给定一个形式背景  $K = (G, M, I)$ , 如表 1 所示,其中共包含 4 个对象和 5 个属性,用符号“ $\times$ ”表示某个对象具有某个属性.形式背景  $K$  的对象导出三支概念格  $OEL(K)$ , 如图 1 所示.

表 1 例 1 的形式背景  $K$

	$a$	$b$	$c$	$d$	$e$
1	$\times$		$\times$		$\times$
2		$\times$		$\times$	
3	$\times$	$\times$			$\times$
4	$\times$	$\times$	$\times$	$\times$	

形式背景  $K$  的对象导出三支概念格  $OEL(K)$  中包含 4 个对象诱导的三支概念,分别为:对象 1 诱导的  $(1, (ace, bd))$ , 对象 2 诱导的  $(2, (bd, ace))$ , 对象 3 诱导的  $(3, (abe, cd))$  和对象 4 诱导的  $(4, (abcd, e))$ .

此外,  $OEL(K)$  中包含 5 个 I 型属性诱导的三支概念,分别为:属性  $a$  诱导的  $(134, (a, \emptyset))$ , 属性  $b$  诱导的  $(234, (b, \emptyset))$ , 属性  $c$  诱导的  $(14, (ac, \emptyset))$ , 属性  $d$  诱导的  $(24, (bd, e))$  和属性  $e$  诱导的  $(13, (ae, d))$ .

再者,  $OEL(K)$  中包含 5 个 II 型属性诱导的三支概念, 分别为: 属性  $a$  诱导的  $(2, (bd, ace))$ , 属性  $b$  诱导的  $(1, (ace, bd))$ , 属性  $c$  诱导的  $(23, (b, c))$ , 属性  $d$  诱导的  $(13, (ae, d))$  和属性  $e$  诱导的  $(24, (bd, e))$ .

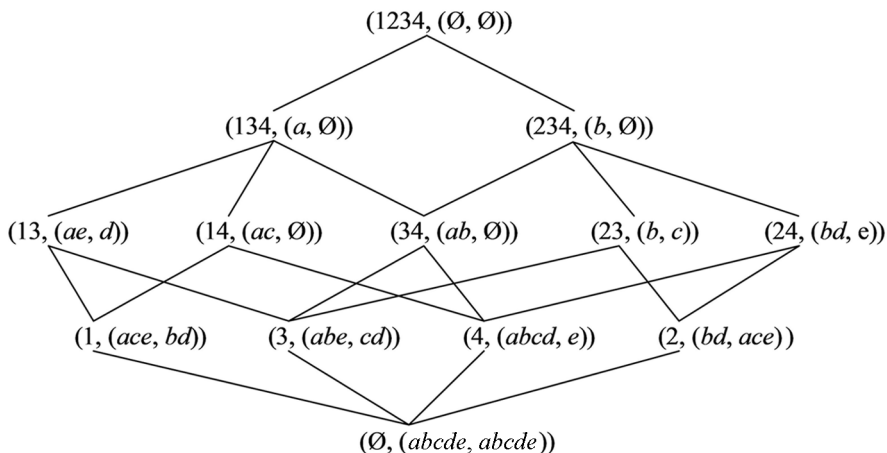


图 1 形式背景  $K$  的对象导出三支概念格  $OEL(K)$

## 2 三支概念簇

概念簇<sup>[27]</sup> 建立于经典概念格的基础之上, 只考虑了对象具有的共同属性. 然而, 在许多情况下, 不仅要考虑对象具有的共同属性, 还需要考虑对象都不具有的属性. 在现代范畴理论的启发下, 本文在三支概念分析的基础上提出三支概念簇.

### 2.1 三支概念簇的定义

本质上, 概念簇刻画了一组具有家族相似性的概念. 在此基础上, 我们提出三支概念簇, 用以描述具有家族相似性的三支概念集合.

**定义 4** 给定一个形式背景  $K = (G, M, I)$ ,  $x \in G, a \in M$ . 若

$$t(x, a) = \{(X, (A, B)) \in OEL(K) \mid (x^{\triangleleft \triangleright}, x^{\triangleleft}) \leq (X, (A, B)) \leq (a^*, a^{*\triangleleft})\} \neq \emptyset$$

则称  $t(x, a)$  为对象  $x$  和属性  $a$  诱导的 I 型三支概念簇.

**定义 5** 给定一个形式背景  $K = (G, M, I)$ ,  $x \in G, a \in M$ . 若

$$t(x, a^{\sim}) = \{(X, (A, B)) \in OEL(K) \mid (x^{\triangleleft \triangleright}, x^{\triangleleft}) \leq (X, (A, B)) \leq (a^{\sim*}, a^{\sim*\triangleleft})\} \neq \emptyset$$

则称  $t(x, a^{\sim})$  为对象  $x$  和属性  $a$  诱导的 II 型三支概念簇.

**性质 1** 给定一个形式背景  $K = (G, M, I)$ ,  $x \in G, a \in M$ . 三支概念簇  $t(x, a)$  和  $t(x, a^{\sim})$  均是  $OEL(K)$  的子格.

**证** 首先, 由形式背景  $K = (G, M, I)$  的有限性可知  $OEL(K)$  是有限的, 进而可知  $t(x, a)$  是有限的. 其次,  $t(x, a)$  中存在唯一的最大元  $(a^*, a^{*\triangleleft})$ , 且存在唯一的最小元  $(x^{\triangleleft \triangleright}, x^{\triangleleft})$ . 综上所述, 三支概念簇  $t(x, a)$  是  $OEL(K)$  的一个子格.

同理可证  $t(x, a^{\sim})$  也是  $OEL(K)$  的一个子格. 证毕.

**定义 6** 给定一个形式背景  $K = (G, M, I)$ ,  $x \in G, a \in M$ . 称  $t(x, \emptyset)$  为对象  $x$  诱导的三支概念簇, 称  $t(\emptyset, a)$  为属性  $a$  诱导的 I 型三支概念簇, 称  $t(\emptyset, a^{\sim})$  为属性  $a$  诱导的 II 型三支概念簇.

实际上, 对于对象  $x$  和属性  $a$  诱导的 I 型三支概念簇  $t(x, a)$ , 令  $a = \emptyset$ , 即可得到对象  $x$  诱导的三支概念簇  $t(x, \emptyset)$ ; 对偶地, 令  $x = \emptyset$ , 即可得到属性  $a$  诱导的 I 型三支概念簇  $t(\emptyset, a)$ .

对于对象  $x$  和属性  $a$  诱导的 II 型三支概念簇  $t(x, a^{\sim})$ , 令  $x = \emptyset$ , 即可得到属性  $a$  诱导的 II 型三支概念簇  $t(\emptyset, a^{\sim})$ .

**例 2** 对于例 1 中的形式背景  $K$ , 考察其上的三支概念簇如下.

形式背景  $K$  中包含 4 个对象诱导的三支概念簇, 分别为:

对象 1 诱导的  $t(1, \emptyset) = \{(1, (ace, bd)), (13, (ae, d)), (14, (ac, \emptyset)), (134, (a, \emptyset)), (1234, (\emptyset, \emptyset))\}$ ;

对象 2 诱导的  $t(2, \emptyset) = \{(2, (bd, ace)), (23, (b, c)), (24, (bd, e)), (234, (b, \emptyset)), (1234, (\emptyset, \emptyset))\}$ ;

对象 3 诱导的  $t(3, \emptyset) = \{(3, (abe, cd)), (13, (ae, d)), (23, (b, c)), (34, (ab, \emptyset)), (134, (a, \emptyset)), (234, (b, \emptyset)), (1234, (\emptyset, \emptyset))\}$ ;

对象 4 诱导的  $t(4, \emptyset) = \{(4, (abcd, e)), (14, (ac, \emptyset)), (24, (bd, e)), (34, (ab, \emptyset)), (134, (a, \emptyset)), (234, (b, \emptyset)), (1234, (\emptyset, \emptyset))\}$ .

形式背景  $K$  中包含 5 个属性诱导的 I 型三支概念簇, 分别为:

属性  $a$  诱导的  $t(\emptyset, a) = \{(134, (a, \emptyset)), (34, (ab, \emptyset)), (14, (ac, \emptyset)), (13, (ae, d)), (3, (abe, cd)), (1, (ace, bd)), (4, (abcd, e)), (\emptyset, (abcde, abcde))\}$ ;

属性  $b$  诱导的  $t(\emptyset, b) = \{(234, (b, \emptyset)), (23, (b, c)), (34, (ab, \emptyset)), (24, (bd, e)), (2, (bd, ace)), (3, (abe, cd)), (4, (abcd, e)), (\emptyset, (abcde, abcde))\}$ ;

属性  $c$  诱导的  $t(\emptyset, c) = \{(14, (ac, \emptyset)), (1, (ace, bd)), (4, (abcd, e)), (\emptyset, (abcde, abcde))\}$ ;

属性  $d$  诱导的  $t(\emptyset, d) = \{(24, (bd, e)), (2, (bd, ace)), (4, (abcd, e)), (\emptyset, (abcde, abcde))\}$ ;

属性  $e$  诱导的  $t(\emptyset, e) = \{(13, (ae, d)), (3, (abe, cd)), (1, (ace, bd)), (\emptyset, (abcde, abcde))\}$ .

形式背景  $K$  中包含 5 个属性诱导的 II 型三支概念簇, 分别为:

属性  $a$  诱导的  $t(\emptyset, a^\sim) = \{(2, (bd, ace)), (\emptyset, (abcde, abcde))\}$ ;

属性  $b$  诱导的  $t(\emptyset, b^\sim) = \{(1, (ace, bd)), (\emptyset, (abcde, abcde))\}$ ;

属性  $c$  诱导的  $t(\emptyset, c^\sim) = \{(23, (b, c)), (3, (abe, cd)), (2, (bd, ace)), (\emptyset, (abcde, abcde))\}$ ;

属性  $d$  诱导的  $t(\emptyset, d^\sim) = \{(13, (ae, d)), (1, (ace, bd)), (3, (abe, cd)), (\emptyset, (abcde, abcde))\}$ ;

属性  $e$  诱导的  $t(\emptyset, e^\sim) = \{(24, (bd, e)), (4, (abcd, e)), (2, (bd, ace)), (\emptyset, (abcde, abcde))\}$ .

## 2.2 三支概念簇的性质

三支概念簇是具有家族相似性的三支概念集合, 三支概念簇的运算在本质上是集合的运算, 首先给出三支概念簇的运算性质.

**性质 2** 给定一个形式背景  $K = (G, M, I)$ ,  $x \in G, a \in M$ . 以下命题成立:

1)  $t(x, a) = t(x, \emptyset) \cap t(\emptyset, a)$ ;

2)  $t(x, a^\sim) = t(x, \emptyset) \cap t(\emptyset, a^\sim)$ .

**证** 1) 一方面,  $t(x, \emptyset)$  由  $OEL(K)$  中所有大于等于  $(x^{\triangleleft}, x^{\triangleleft})$  且小于等于  $(G, G^{\triangleleft})$  的三支概念组成. 另一方面,  $t(\emptyset, a)$  由  $OEL(K)$  中所有大于等于  $(\emptyset, \emptyset^{\triangleleft})$  且小于等于  $(a^*, a^{*\triangleleft})$  的三支概念组成. 综合两个方面, 由  $t(x, a)$  的定义可知  $t(x, a) = t(x, \emptyset) \cap t(\emptyset, a)$ . 命题 1) 成立.

2) 同理可证命题 2 成立, 证毕.

性质 2 实际上给出了利用对象诱导和属性诱导的三支概念簇计算任意三支概念簇的方法. 例如, 考虑例 2 中的三支概念簇, 利用性质 2 及例 2 的计算结果可得:

$$t(1, a) = t(1, \emptyset) \cap t(\emptyset, a) = \{(1, (ace, bd)), (13, (ae, d)), (14, (ac, \emptyset)), (134, (a, \emptyset)), (1234, (\emptyset, \emptyset))\} \cap \{(134, (a, \emptyset)), (34, (ab, \emptyset)), (14, (ac, \emptyset)), (13, (ae, d)), (3, (abe, cd)), (1, (ace, bd)), (4, (abcd, e)), (\emptyset, (abcde, abcde))\} = \{(1, (ace, bd)), (13, (ae, d)), (14, (ac, \emptyset)), (134, (a, \emptyset))\}.$$

**性质 3** 给定一个形式背景  $K = (G, M, I)$ ,  $x \in G, a \in M$ . 以下命题成立:

- 1) 若  $I(x, a) = 0$ , 则  $t(x, a) = \emptyset$ ;
- 2) 若  $I(x, a) = 1$ , 则  $t(x, a^{\sim}) = \emptyset$ ;
- 3)  $OEL(K) = (\bigcup_{x \in G} t(x, \emptyset)) \cup (\emptyset, \emptyset^{\triangleleft})$ ;
- 4)  $OEL(K) = (\bigcup_{a \in M} t(\emptyset, a) \cup t(\emptyset, a^{\sim})) \cup (G, G^{\triangleleft})$ .

**证** 1) 若  $I(x, a) = 0$ , 则概念  $(x^{\triangleleft \triangleright}, x^{\triangleleft})$  与  $(a^*, a^{*\triangleleft})$  是不可比的, 所以  $OEL(K)$  中不存在大于等于  $(x^{\triangleleft \triangleright}, x^{\triangleleft})$  且小于等于  $(a^*, a^{*\triangleleft})$  的概念, 根据  $t(x, a)$  的定义可得  $t(x, a) = \emptyset$ . 命题 1) 成立.

2) 同理可证命题 2 成立.

3) 若  $t(x, \emptyset)$  中的最小元与最大元分别为  $(x^{\triangleleft \triangleright}, x^{\triangleleft})$  和  $(G, G^{\triangleleft})$ , 则  $\bigcup_{x \in G} t(x, \emptyset)$  即为  $OEL(K)$  中大于等于所有对象诱导的且小于等于  $(G, G^{\triangleleft})$  的三支概念. 因此易得  $OEL(K) = (\bigcup_{x \in G} t(x, \emptyset)) \cup (\emptyset, \emptyset^{\triangleleft})$ . 命题 3 成立.

4) 同理可证命题 4 成立, 证毕.

需要指出的是,  $OEL(K) = (\bigcup_{a \in M} t(\emptyset, a)) \cup (G, G^{\triangleleft})$  与  $OEL(K) = (\bigcup_{a \in M} t(\emptyset, a^{\sim})) \cup (G, G^{\triangleleft})$  均不一定成立.

在例 2 中,  $OEL(K) = (\bigcup_{a \in M} t(\emptyset, a)) \cup (G, G^{\triangleleft})$  成立, 但  $OEL(K) = (\bigcup_{a \in M} t(\emptyset, a^{\sim})) \cup (G, G^{\triangleleft})$  不成立. 实际上,  $OEL(K) = (\bigcup_{a \in M} t(\emptyset, a^{\sim})) \cup (G, G^{\triangleleft}) \cup (134, (a, \emptyset)) \cup (234, (b, \emptyset)) \cup (14, (ac, \emptyset)) \cup (34, (ab, \emptyset))$ .

### 3 基于三支概念簇的知识发现

在信息检索的过程中, 经常会请求系统推荐一类和某种已知对象类似, 且具有或不具有某些特殊限定的对象. 如在网上购物系统中, 顾客通常会要求客服进行产品的合理推荐, 如: “请给我推荐一款与 2 号模特上身穿的相似, 并且是深蓝色的, 但不含拉链的外套”, 这种需求可以形式化地表述为: 用  $G$  表示对象的集合,  $M$  表示属性的集合, 在  $G$  中查找与对象  $x$  相似, 且具有属性  $a$ , 但不具有属性  $b$  的对象.

基于以上讨论, 下面提出算法 1 以实现符合特定约束条件下的相似对象检索.

**算法 1** 基于三支概念簇的相似对象查找算法.

输入: 形式背景  $K = (G, M, I)$ , 对象  $x \in G$ , 属性  $a, b \in M$ .

输出: 与对象  $x$  相似, 且具有属性  $a$ , 但不具有属性  $b$  的所有对象.

步骤 1: 建立形式背景  $K$  的对象导出三支概念格  $OEL(K)$ ;

步骤 2: 计算三支概念簇  $t(\emptyset, a)$  与  $t(\emptyset, b^{\sim})$ ;

步骤 3: 计算  $t(\emptyset, a) \cap t(\emptyset, b^{\sim})$ , 并记其中全体对象诱导的三支概念的外延为集合  $E$ ;

步骤 4: 从概念  $(x^{\triangleleft \triangleright}, x^{\triangleleft})$  开始, 逐层遍历其父概念:

记当前遍历的概念为  $(X, (A, B))$ ;

若  $X \cap E \neq \emptyset$ , 则  $X \cap E$  中包含的对象即为所求.

步骤 5: 返回结果, 算法结束.

**例 3** 表 2 描述了一个由 8 种生物构成的形式背景, 也就是在形式概念分析理论研究中广为熟知的“生物和水”形式背景, 其对象导出三支概念格如图 2 所示. 由于所有的生物都需要水, 这里不再考虑“需要水”这个生物属性, 只考虑其余的 8 种生物属性. 用数字 1 到 8 分别代表蚂蝗、鱼、蛙、狗、水草、芦苇、豆与玉米. 另外, 用小写字母  $a$  到  $h$  表示 8 种生物属性, 即: 在水中生活、在陆地生活、有叶绿素、双子叶、单子叶、能运动、有四肢与会哺乳.

为了更好地说明上述提出的基于三支概念簇的相似对象查找算法, 我们将通过实际案例来进一步对基于概念簇和三支概念簇的相似对象查找算法进行分析对比.

首先,使用上述算法 1,查找和芦苇(对象 6)相似,且有叶绿素(具有属性  $c$ ),但不能在陆地生活(不具有属性  $b$ )的生物.

表 2 例 3 的形式背景

	$a$	$b$	$c$	$d$	$e$	$f$	$g$	$h$
1	×					×		
2	×					×	×	
3	×	×				×	×	
4		×				×	×	×
5	×		×		×			
6	×	×	×		×			
7		×	×	×				
8		×	×		×			

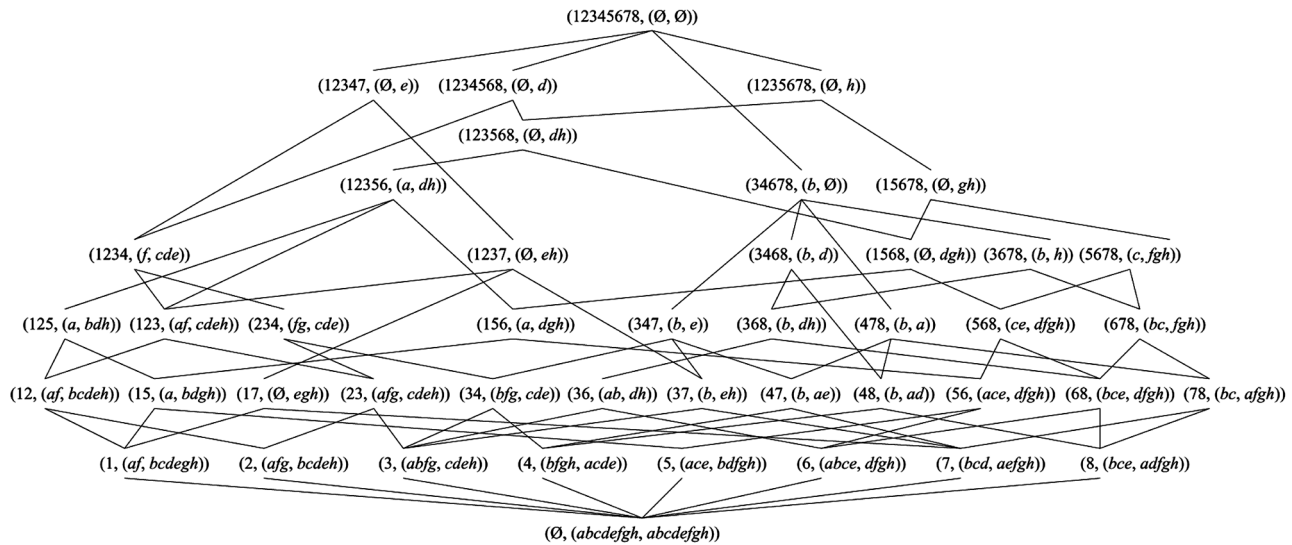


图 2 例 3 的对象导出三支概念格

输入:“生物和水”形式背景,对象  $6 \in G$ ,属性  $c, b \in M$ .

输出:与对象 6 相似,且具有属性  $c$ ,但不具有属性  $b$  的所有对象.

步骤 1:建立“生物和水”形式背景的对象导出三支概念格,如图 2 所示;

步骤 2:计算三支概念簇  $t(\emptyset, c) = \{(\emptyset, (abcdefgh, abcdefgh)), (5, (ace, bdfgh)), (6, (abce, dfg)), (7, (bcd, aefgh)), (8, (bce, adfgh)), (56, (ace, dfg)), (68, (bce, dfg)), (78, (bc, afg)), (568, (ce, dfg)), (678, (bc, fgh)), (5678, (c, fgh))\}$ ,  $t(\emptyset, b^{\sim}) = \{(\emptyset, (abcdefgh, abcdefgh)), (1, (af, bcdegh)), (2, (afg, bcdeh)), (5, (ace, bdfgh)), (12, (af, bcdeh)), (15, (a, bdgh)), (125, (a, bdh))\}$ ;

步骤 3:计算  $t(\emptyset, c) \cap t(\emptyset, b^{\sim}) = \{(\emptyset, (abcdefgh, abcdefgh)), (5, (ace, bdfgh))\}$ ,并记其中全体对象诱导的三支概念的外延为集合  $E = \{5\}$ ;

步骤 4:从概念  $(x^{\triangleleft}, x^{\triangleleft}) = (6, (abce, dfg))$  开始,逐层遍历其父概念:

遍历其第一层父概念得到  $(56, (ace, dfg))$  与  $(68, (bce, dfg))$ ,由  $\{5, 6\} \cap \{5\} = \{5\} \neq \emptyset$  可知对象 5 满足算法要求;由  $\{68\} \cap \{5\} = \emptyset$  可知不满足算法要求,舍弃;综上可得,对象 5 即为所求.

步骤 5:返回结果对象 5,即符合条件的生物是水草,算法结束.

接下来,使用文献[27]所提出的基于对象直观图与概念簇的相似对象查找算法,查找和芦苇(对象 6)相似,且有叶绿素(具有属性  $c$ ),但不能在陆地生活(不具有属性  $b$ )的生物.

输入: “生物和水”形式背景, 对象  $6 \in G$ , 属性  $c, b \in M$ .

输出: 与对象 6 相似且具有属性  $c$  的所有对象.

步骤 1: 建立“生物和水”形式背景的对象直观图与概念格, 见文献[27];

步骤 2: 在步骤 1 的对象直观图中查找大于等于  $(6, 6^*)$  的所有二元组, 可得最大二元组仍为  $(6, 6^*)$ ;

步骤 3: 计算  $AC(6, c) = C(6^{**}, c^{**}) = \{(6, abce), (56, ace), (68, bce), (568, ce), (678, bc), (5678, c)\}$ ;

步骤 4: 对于  $AC(6, c)$  中存在的概念  $(56, ace) = (5^{**}, 5^*)$  和  $(68, bce) = (8^{**}, 8^*)$ , 可得符合条件的对象为 5 和 8;

步骤 5: 返回结果对象 5 和对象 8, 即符合条件的生物是水草和玉米, 算法结束.

综上所述, 针对同一个案例, 对比两个算法可知, 使用基于对象直观图与概念簇的算法只能检索和芦苇(对象 6)相似且有叶绿素(具有属性  $c$ )的生物, 无法检索其中不能在陆地生活(不具有属性  $b$ )的生物, 即属性  $b$  无法被按需使用, 最终得到的结果是水草(对象 5)和玉米(对象 8). 虽然此算法得到两种结果, 但是根据常识可知, 算法 1 得到的结果水草和基于对象直观图与概念簇的算法得到的结果之一玉米相比较更加符合条件. 因此, 三支概念簇相比较于概念簇, 更加有利于检索到符合需求的对象.

通过深入思考发现, 三支概念簇之所以更加有利于检索到符合需求的对象, 原因在于三支概念不仅考虑了对象共同具有的属性, 而且还考虑了对象共同不具有的属性, 使得三支概念比经典形式概念能够捕捉到更加精确的对象细节, 从而发现更加细致的对象分类, 检索到更加贴切实际应用场景的相似对象.

## 4 结 论

本文融合现代范畴理论与三支概念分析, 提出了三支概念簇. 研究表明, 三支概念簇刻画了具有“家族相似性”的三支概念集合, 是检索符合特定约束条件下相似对象的有力工具. 三支概念簇作为三支概念分析中的重要粒度之一, 还有许多值得深入探讨的问题. 例如: 模糊形式背景下的模糊概念簇如何定义, 与经典二值形式背景下的三支概念簇有何区别与联系; 在动态环境下, 三支概念簇如何更新; 对于大规模形式背景, 如何提高三支概念簇的计算效率. 最后, 需要指出的是概念簇建立在形式概念的基础之上, 构造概念格的复杂性限制了其大规模应用的有效性. 因此, 如何采用新的思路, 从新的角度研究概念簇还需要付出更多的努力.

## 参考文献:

- [1] WU W Z, LEUNG Y, MI J S. Granular Computing and Knowledge Reduction in Formal Contexts [J]. IEEE Transactions on Knowledge and Data Engineering, 2009, 21(10): 1461-1474.
- [2] 仇国芳, 马建敏, 杨宏志, 等. 概念粒计算系统的数学模型 [J]. 中国科学(F辑: 信息科学), 2009, 39(12): 1239-1247.
- [3] 苗夺谦, 张清华, 钱宇华, 等. 从人类智能到机器实现模型——粒计算理论与方法 [J]. 智能系统学报, 2016, 11(6): 743-757.
- [4] 李金海, 吴伟志. 形式概念分析的粒计算方法及其研究展望 [J]. 山东大学学报(理学版), 2017, 52(7): 1-12.
- [5] 智慧来, 李金海. 基于必然属性分析的粒描述 [J]. 计算机学报, 2018, 41(12): 2702-2719.
- [6] ZADEH L A. Toward a Theory of Fuzzy Information Granulation and Its Centrality in Human Reasoning and Fuzzy Logic [J]. Fuzzy Sets and Systems, 1997, 90(2): 111-127.
- [7] QI J J, WEI L, WAN Q. Multi-Level Granularity in Formal Concept Analysis [J]. Granular Computing, 2019, 4(3): 351-362.
- [8] FUJITA H, GAETA A, LOIA V, et al. Resilience Analysis of Critical Infrastructures: A Cognitive Approach Based on Granular Computing [J]. IEEE Transactions on Cybernetics, 2019, 49(5): 1835-1848.
- [9] HONKO P. Recent Granular Computing Frameworks for Mining Relational Data [J]. Artificial Intelligence Review,



2019, 52(4): 2705-2742.

- [10] WILLE R. Restructuring Lattice Theory: An Approach Based on Hierarchies of Concepts [M] // RIVAL I. Ordered Set. Dordrecht-Boston: Reidel, 1982: 445-470.
- [11] 智慧来. 概念格构造与应用中的关键技术研究 [D]. 上海: 上海大学, 2010.
- [12] YAO YY. A Comparative Study of Formal Concept Analysis and Rough Set Theory in Data Analysis [C] // Rough Sets and Current Trends in Computing (LNCS: 3066), Uppsala, Sweden, 2004: 59-68.
- [13] YAO Y Y. Concept Lattices in Rough Set Theory [C] // Proceedings of the International Conference of the North American Fuzzy Information Processing Society, IEEE, Banff, Canada, 2004: 796-801.
- [14] ZHI H L, QI J J, QIAN T, et al. Three-Way Dual Concept Analysis [J]. International Journal of Approximate Reasoning, 2019, 114: 151-165.
- [15] YAO Y Y. Rough-Set Concept Analysis: Interpreting RS-Definable Concepts Based on Ideas from Formal Concept Analysis [J]. Information Sciences, 2016, 346/347: 442-462.
- [16] XU W H, LI W T. Granular Computing Approach to Two-Way Learning Based on Formal Concept Analysis in Fuzzy Datasets [J]. IEEE Transactions on Cybernetics, 2016, 46(2): 366-379.
- [17] SHAO M W, LEUNG Y, WANG X Z, et al. Granular Reducts of Formal Fuzzy Contexts [J]. Knowledge-Based Systems, 2016, 114: 156-166.
- [18] 李金海, 吴伟志, 邓 硕. 形式概念分析的多粒度标记理论 [J]. 山东大学学报(理学版), 2019, 54(2): 30-40.
- [19] ZHI H L, QI J J, QIAN T, et al. Conflict Analysis under One-Vote Veto Based on Approximate Three-Way Concept Lattice [J]. Information Sciences, 2020, 516: 316-330.
- [20] YAO Y Y. Three-Way Decision: An Interpretation of Rules in Rough Set Theory [M] // Rough Sets and Knowledge Technology Berlin, Heidelberg: Springer Berlin Heidelberg, 2009: 642-649.
- [21] QI J J, WEI L, YAO Y Y. Three-Way Formal Concept Analysis [M] // Rough Sets and Knowledge Technology Cham: Springer International Publishing, 2014: 732-741.
- [22] QI J J, QIAN T, WEI L. The Connections Between Three-Way and Classical Concept Lattices [J]. Knowledge-Based Systems, 2016, 91: 143-151.
- [23] 魏 玲, 高 乐, 祁建军. 三支概念分析研究现状与展望 [J]. 西北大学学报(自然科学版), 2019, 49(4): 527-537.
- [24] 魏 玲, 赵思雨. 三支概念分析中的粒与知识结构 [J]. 西北大学学报(自然科学版), 2020, 50(4): 537-545.
- [25] CROFT W, CRUSE D A. Cognitive Linguistics [M]. Cambridge: Cambridge University Press, 2004.
- [26] LAKOFF G. Women, Fire, and Dangerous Things: What Categories Reveal about the Mind [M]. Chicago: The University of Chicago Press, 1987.
- [27] 智慧来, 李逸楠. 基于概念簇的知识表示 [J]. 西北大学学报(自然科学版), 2020, 50(4): 529-536.

责任编辑 汤振金