带有对象结构的三元概念分析研究

信息2101班：秦文文 指导教师：钱婷 论文类型：毕业论文

摘要：三元概念是三元概念分析理论的研究基础，它由外延、内涵及条件三部分构成，能够更精确的反映对象、属性及条件之间的内在联系。然而在实际应用中仍面临数据获取复杂及计算成本高等诸多挑战。本文基于带有对象结构信息形式概念，通过引入连通性将对象结构加入到三元背景，提出了对象全局和对象局部结构信息三元概念。进一步研究了带有对象结构的形式概念分析与带有对象结构的三元概念分析之间的内在联系，从概念定义、分析维度等层面进行比较，揭示了对象结构在二元与三元分析框架中的关系及转化方法，为不同知识框架间的信息交互与知识发现提供了全新的理论思路。

关键词：形式概念；三元概念；三元背景；结构信息三元概念

1 研究背景

形式概念分析（Formal Concept Analysis，FCA）由德国数学家Wille提出，其核心是根据对象与属性之间的二元关系构造形式概念，每一个形式概念由外延和内涵两部分构成。形式概念分析在处理二元关系方面具有很多优势，但随着网络上三维数据的日益增多，传统二元形式概念分析逐渐显现出局限性。1995年，Wille等人受Peirce C. 关于三个论域范畴的实用主义哲学的启发，把形式概念分析推广到三元结构，提出了三元概念分析(Triadic concept analysis TCA)。三元概念分析，一方面继承了形式概念分析的方法与核心思想；另一方面通过引入“条件”这个维度，定义了三元背景、三元概念等相关知识，以处理更加复杂的数据分析。

三元概念分析虽然已经初步建立了理论体系，但由于不能充分利用对象之间的关联关系，故在实际应用中仍面临数据获取处理复杂及计算成本高等诸多挑战。而在形式概念分析中，李金海等人将对象关系中引入连通性提出全局结构信息形式概念和局部结构信息形式概念，进而围绕基于结构信息形式概念的知识发现问题展开讨论，这一研究既完善了理论体系，又推动该理论在多学科交叉融合中的发展。通过整合对象间结构关联与传统属性-条件分析框架，为解决形式概念分析应用中的复杂数据处理提供了新的理论思路和方法。

受李金海等人的启发，本文将对象结构引入三元背景，拓展三元概念分析的研究。

2 发展现状

随着形式概念分析领域研究的不断深入拓展，其应用范围也逐渐扩大。目前，

TCA 研究涵盖理论拓展与应用创新等多个方面。在基础理论层面，围绕概念三元格的构造算法持续优化，如 Ganter 等学者提出多种高效形式概念生成算法，大幅度提高了大规模数据的处理效率；从代数结构、拓扑结构等数学视角分析三元概念格的性质，为理解概念内在结构提供了理论依据。在应用方法层面，研究范畴拓展到概念三元格的建立、三元蕴含及关联规则挖掘、三元模态算子、概念聚类、多粒度分析、三元概念分类问题、背景因子分析及模糊化处理等领域，形成了完整的方法体系。

带有对象结构的形式概念分析作为一个新的课题，蕴含着广阔的探索空间。目前，在形式概念分析领域，李金海等人通过合理地调整对象结构或优化二元关系，实现对二元概念数量的精简以及结构的优化。能否借助带有对象结构的形式概念分析的知识体系与方法路径，类比解决的三元概念分析中的难题，充分挖掘研究对象间的结构信息，用已知去求解未知，是我们研究的主要内容。

3 相关分析

3.1 带有对象结构的三元概念分析概念

定义 3.1 称五元组 (K_1,K_2,K_3,A,C) 为带有对象结构的三元背景，其中 $K_1=\{g_1,g_2,...,g_n\}$ 为非空有限对象集，$K_2=\{m_1,m_2,...,m_n\}$ 为非空有限属性集，$K_3=\{b_1,b_2,...,b_n\}$ 为非空有限条件集，A 是 K_1 上的结构矩阵，即对任意的 $g_i,g_j\in K_1$，若 $a_{ij}=1$，则表示对象 g_i，g_j 是相关联的；若 $a_{ij}=0$，则表示对象 g_i，g_j 是不相关联的；C 是 $K_1\times K_2\times K_3$ 上的连接矩阵，即任意的 $g_i\in K_1$、$m_j\in K_2$ 和 $b_k\in K_3$，用 $\left(g_i,m_j,b_k\right)=1$ 表示对象 g_i 在条件 b_k 下具有属性 m_j，用 $\left(g_i,m_j,b_k\right)=0$ 表示对象 g_i 在条件 b_k 下不具有属性 m_j。

定义 3.2 设五元组 (K_1,K_2,K_3,A,C) 为带有对象结构的三元背景，$X\subseteq K_1$ 包含的所有节点之间的结构信息诱导出的子图是连通的，称 X 是连连通的。

定义 3.3 设五元组 (K_1,K_2,K_3,A,C) 为带有对象结构的三元背景，$A_i\subseteq K_i$，i=1, 2, 3。(1)如果满足 $A_i=(A_j\times A_k)^{(i)}$，其中 $\{i,j,k\}=\{1,2,3\}$ 且 A_1 是连通，则称 (A_i,A_j,A_k) 是连通，则称 (A_i,A_j,A_k) 为对象全局结构信息三元概念。所有对象全局结构信息三元概念构成的集合记为 $N(K_1,K_2,K_3,A,C)$。(2)如果满足 $A_1^{(1)}=A_2\times A_3$，A_1 连通且不存在 $g_i\in(A_2\times A_3)^{(1)}-A_1$ 使得 $A_1\cup\{g_i\}$ 连通，则称 (A_1,A_j,A_k) 序对为对象局部结构信息三元概念。所有对象局部结构信息三元概念构成的集合记为 $N_L(K_1,K_2,K_3,A,C)$。

显然我们易得 $N_L(K_1,K_2,K_3,A,C)\subseteq N(K_1,K_2,K_3,A,C)$。

3.2 对象结构矩阵重构背景下结构信息三元概念的演化

当只有对象结构矩阵 A 发生变化时，将其更新后的对象结构矩阵记为 A'，根据定义 3.3 可得到如下结论。

定理 3.1 设 (K_1,K_2,K_3,A,C) 为带有对象结构的三元背景，$a_{ij}=a_{ji}$ 从 1 变为 0，那么 $(A_1,A_2,A_3)\in N(K_1,K_2,K_3,A_1,C)$ 满足：

（1）如果 A_1 仍然连通，则 $(A_1,A_2,A_3)\in N(K_1,K_2,K_3,A_1',C)$；

（2）如果 A_1 不连通，A_1 的任一连通子集 B_1，对于新的序对 $(A_1,B_1^{(1)})$，如果 $B_1^{(1)}\supset A_2\times A_3$ 且 $B_1^{(1)(1)}=B_1^{(1)}$，那么，$(B_1,B_1^{*})\in N(K_1,K_2,K_3,A_1',C)$，否则 $(B_1,B_1^{*})\notin N(K_1,K_2,K_3,A_1',C)$。

注：B_1^* 是 B_1 作 (1) 算子后成对包含两个子集 B_{11} 和 B_{12}，B_{11} 是 B_1^* 第一部分分量的并集，B_{12} 是 B_1^* 第二部分分量的并集。

定理 3.2 设 (K_1,K_2,K_3,A,C) 为带对象结构的三元背景，$(A_1,A_2,A_3)\in N(K_1,K_2,K_3,A,C)$，若 $a_{ij}=a_{ji}$ 从 0 变为 1，那么 $(A_1,A_2,A_3)\in N(K_1,K_2,K_3,A_1',C)$。

接下来研究对象结构矩阵改变后，对象局部结构三元概念的变化情况。

定理 3.3 设 (K_1,K_2,K_3,A,C) 为带有对象结构的三元背景，$a_{ij}=a_{ji}$ 从 1 变为 0，那么 $(A_1,A_2,A_3)\in N_L(K_1,K_2,K_3,A_1,C)$ 满足：

（1）如果 A_1 在 (K_1,K_2,K_3,A_1',C) 中连通，则 $(B_1,B_1^{*})\in N(K_1,K_2,K_3,A_1',C)$；

（2）如果 A_1 在 (K_1,K_2,K_3,A_1',C) 中不连通，则 $(A_1,A_2,A_3)\notin N_L(K_1,K_2,K_3,A_1',C)$ 对 A_1 的最大连通子集 B_0，有 $(B_0,B_0^{*})\in N_L(K_1,K_2,K_3,A_1',C)$。

定理 3.5 设 (K_1,K_2,K_3,A,C) 为带有对象结构的三元背景，$a_{ij}=a_{ji}$ 从 0 变为 1，那么 $(A_1,A_2,A_3)\in N_L(K_1,K_2,K_3,A_1,C)$ 满足：

（1）如果存在 $g_i\in(A_j\times A_k)^{(i)}-A_i$ 在 (K_1,K_2,K_3,A_1',C) 中与 A_1 连通，那么 $(A_1,A_2,A_3)\notin N_L(K_1,K_2,K_3,A_1',C)$；

（2）如果不存在 $g_i\in(A_j\times A_k)^{(i)}-A_i$ 在 (K_1,K_2,K_3,A_1',C) 中与 A_1 连通，那么 $(A_1,A_2,A_3)\in N_L(K_1,K_2,K_3,A_1,C)$。

定理 3.6 设 (K_1,K_2,K_3,A,C) 为带有对象结构的三元背景，$a_{ij}=a_{ji}$ 从 0 变为 1，$(A_1,A_2,A_3)\in L(K_1,K_2,K_3,C)-N(K_1,K_2,K_3,A,C)$ 满足：

（1）若 A_1 在 (K_1,K_2,K_3,A_1',C) 中连通，则 $(A_1,A_2,A_3)\in N(K_1,K_2,K_3,A_1',C)$；

（2）若 A_1 在 (K_1,K_2,K_3,A_1',C) 中不连通，则 $(A_1,A_2,A_3)\notin N(K_1,K_2,K_3,A_1',C)$。

若$(A_1, A_2, A_3) \in L(K_1, K_2, K_3, C) - N_L(K_1, K_2, K_3, A, C)$则有定理 3.6 的类似结论。

3.3 带有对象结构的形式概念分析和三元概念分析之间的关系

王啸等人基于外延待选集研究了形式概念与三元概念之间的相互关系。受此启发，我们研究了带有对象结构的形式概念与带有对象结构的三元概念之间的关系。

定理 3. 7 设 T 为三元背景 $K=(K_1, K_2, K_3, Y)$ 下的对象全局结构信息三元概念集合，F 为某个诱导的形式背景 $K^{(1)}=(K_1, K_2 \times K_3, Y^{(1)})$ 下的连通形式概念集合。对于任意的 $(A_1, A_2, A_3) \in T$，若满足连通形式概念外延的条件，则存在 $(X, B) \in F$，使得 $A_1 = X$。

定理 3. 8 设形式背景为 (G, M, I)，带有对象结构矩阵 A，二元连通概念通过 $X' = B$，$B' = X$ 以及连通性确定。三元背景为 $K=(K_1, K_2, K_3, Y)$，带有对象结构矩阵 A，三元连通概念通过 $A_i = (A_j \times A_k)^{(i)}$，其中 $\{i, j, k\} = \{1, 2, 3\}$ 且 A_i 是连通确定。

4 研究结论及对策建议

本文主要的创新点及结论为：将图论中的连通性融入到三元背景中进行研究，提出了基于连通性的全局与局部结构信息三元概念。

本文后续研究建议可从以下几个方面展开：1）动态带有结构信息的三元概念分析；2）融合不同类型结构信息的三元概念分析研究；3）基于带有结构信息三元概念明确的语义解释研究。