

DOI: 10.11992/tis.201703026

网络出版地址: <http://kns.cnki.net/kcms/detail/23.1538.tp.20180423.1019.008.html>

利用二部图生成概念格

窦林立, 展正然

(中国地质大学长城学院 基础课教学部, 河北 保定 071002)

摘要: 概念格作为一种有效的知识发现与数据处理的工具, 在许多领域得到了广泛应用, 概念格的构造在其应用中具有重要的意义。每个概念格的形式背景都可以对应一个二部图, 本文通过二部图的极大完全子图的概念来生成概念格, 给出了基于二部图的深度优先的概念格的迭代算法。首先, 对形式背景进行必要的约简; 其次, 利用二部图的极大完全子图得到顶层概念的直接子概念; 最后, 通过求二部图的导出子图来简化形式背景, 并得出每个概念的直接子概念和所有子概念, 从而生成概念格。

关键词: 形式背景; 概念格; 二部图; 极大完全子图; 直接子概念; Hasse 示图; 图论; 导出子图

中图分类号: TP18 **文献标志码:** A **文章编号:** 1673-4785(2018)05-0687-06

中文引用格式: 窦林立, 展正然. 利用二部图生成概念格[J]. 智能系统学报, 2018, 13(5): 687-692.

英文引用格式: DOU Linli, ZHAN Zhengran. Constructing concept lattice using bipartite graph[J]. CAAI transactions on intelligent systems, 2018, 13(5): 687-692.

Constructing concept lattice using bipartite graph

DOU Linli, ZHAN Zhengran

(Basic Teaching Department, The Great Wall College, China University of Geosciences, Baoding 071002, China)

Abstract: As an effective tool for knowledge discovery and data processing, concept lattices have been widely applied in many fields, and in these applications, it is important to efficiently construct concept lattice. The formal context of each concept lattice corresponds to a bipartite graph. In this paper, the maximum complete subgraph of a bipartite graph is used to generate a concept lattice, and then where an iterative algorithm with depth priority is proposed based on a bipartite graph. The process is as follows: first, a formal context is reduced; then, the direct subconcepts of the top concept are obtained using maximal complete-subgraph of bipartite graph; finally, through the derivation of the induced subgraph of bipartite graph to reduce the formal context, and find direct subconcepts and all subconcepts of every concept, then the concept lattice was established.

Keywords: formal context; concept lattice; bipartite graph; maximum complete subgraph; direct subconcept; Hasse diagram; graph theory; induced subgraph

形式概念分析又称概念格^[1-2], 是由德国 Wille 教授在 20 世纪 80 年代首次提出的, 它提供了一种支持数据分析和知识处理的数学工具^[3-4]。近些年来, 国内外各位学者提出了许多的构造算法。例如: Godin 等^[5]提出了概念格的渐进生成算法; Ho^[6]提出了基于概念格的概念聚类算法; 张文修等^[7]给出了在保持格同构的条件下建立概念格

的属性约简的方法; Elloumi 等^[8]基于模糊形式背景建立了一个多层次的约简理论。

近几年许多学者从图论方面对概念格进行研究。例如: Amilhastre 等^[9]与 Berry 等^[10-12]将二部图与概念格进行了交叉研究; 李立峰等^[13-14]利用弦二部图和链图对概念格进行表示; 张涛等^[15-17]利用属性树和属性拓扑图来表示形式背景, 简化了形式背景的表示结构, 并提出了基于树图的属性约简算法。本文通过二部图的极大完全子图的

收稿日期: 2017-03-21. 网络出版日期: 2018-04-24.

基金项目: 河北省高校科研基金项目 (Z2015137).

通信作者: 窦林立. E-mail: 1321407258@qq.com.

概念来生成概念格,给出了基于二部图的深度优先的概念格的迭代算法。

1 基本概念

本节主要介绍概念格与二部图的基本知识。关于概念格的更多内容参见文献[1,18],有关图论的详细内容参见文献[19-20]。

1.1 概念格

定义 1 1) 在形式概念分析中,一个形式背景是一个三元数组 $K = (O, M, I)$, 其中 O 是对象集合, M 是属性集合, I 是 O 与 M 之间的一个二元关系。用 oIm 表示“对象 o 有属性 m ”。对于 $A \subseteq O$ 及 $B \subseteq M$, 定义: $A' = \{m \in M | \forall o \in A, oIm\}$, $B' = \{o \in O | \forall m \in B, oIm\}$ 。则形式背景 $K = (O, M, I)$ 的概念定义为元素对 (A, B) , 其中 $A \subseteq O$, $B \subseteq M$, $A' = B$, $B' = A$ 。概念 (A, B) 的外延是 A , 内涵是 B 。

2) 在形式背景 $K = (O, M, I)$ 的概念集合间建立一种偏序关系: 给定两个概念 $C_1 = (A_1, B_1)$ 和 $C_2 = (A_2, B_2)$, 则 $C_1 < C_2 \Leftrightarrow A_1 \subset A_2$ (或 $B_1 \supset B_2$), 那么 C_1 是 C_2 的子概念, C_2 是 C_1 的父概念。根据这种偏序关系, 形式背景 $K = (O, M, I)$ 的概念节点的集合产生一种格结构, 称它为 $K = (O, M, I)$ 的概念格, 记为 $L(K)$ 。并且由这种偏序关系可生成 $L(K)$ 的 Hasse 图: 对于 $L(K)$ 中的两个不同的节点 C_1 和 C_2 , 如果 $C_1 < C_2$, 并且不存在其他的节点 C_3 , 使得 $C_1 < C_3 < C_2$, 则 C_1 与 C_2 之间就存在一条边。此时 C_1 称为 C_2 的直接子概念, C_2 称为 C_1 的直接父概念。

形式背景 $K = (O, M, I)$ 中集合 O 和 M 中所含元素的个数直接影响到生成概念格的复杂程度, 所以我们有必要对形式背景进行约简。

定义 2 1) 在形式背景 $K = (O, M, I)$ 中, 若存在 $m \in M$, 对任意的 $o \in O$, 都有 oIm 成立, 即属性 m 与对象集中每个对象都满足关系 I , 则可以从形式背景中去掉这个属性 m , 所得概念格 $L(K^*)$ 与原形式背景所生成的概念格 $L(K)$ 同构。要得到 $L(K)$, 只需在 $L(K^*)$ 的每个概念的内涵中都添加属性 m 即可。

2) 若 $i, j \in M$ 且 $i \neq j$, $P = \{o \in O | oIi\}$, $Q = \{o \in O | oIj\}$, 若 $P = Q$, 则可以将两个属性 i 和 j 缩写成一个属性 ij 。显然, 缩写后的形式背景与原形式背景生成的概念格相同。

通过上述两种方法约简形式背景后, 就可以有效地减少形式背景中属性集合中的元素个数, 从而降低生成概念格的难度, 更迅速地生成概念格。

例 1 表 1 给出了形式背景 $K = (O, M, I)$, 其中

$O = \{1, 2, 3, 4, 5, 6, 7\}$, $M = \{a, b, c, d, e, f, g\}$, 表 2 为约简后的形式背景。其中属性的个数由 7 个约简成 5 个。

表 1 形式背景 (O, M, I)
Table 1 Formal context (O, M, I)

对象	a	b	c	d	e	f	g
1	×	×	×		×		×
2				×	×	×	×
3		×		×		×	×
4		×	×				×
5		×	×	×		×	×
6	×	×	×				×
7					×		×

表 2 约简后的形式背景 (O, M, I)
Table 2 Reduced formal context (O, M, I)

对象	a	b	c	df	e
1	×	×	×		×
2				×	×
3		×		×	
4		×	×		
5		×	×	×	
6	×	×	×		
7					×

1.2 图论

定义 3 1) 设 V_1 和 V_2 是图 G 的顶点子集, 使 $V_1 \cup V_2 = V(G)$, $V_1 \cap V_2 = \emptyset$ 且 G 的每一条边都是一个端点在 V_1 中另一个端点在 V_2 中, 则称 G 为二部图, 记作 $G = (V_1, V_2; E)$ 。如果 V_1 中的顶点与 V_2 中的每一个顶点都相邻, 则称为完全二部图。

2) 设 V_0 是图 G 的任一顶点子集, G 中与 V_0 的顶点邻接的所有顶点的集合称为 V_0 的邻集, 记作 $N(V_0)$ 。

3) 设 $v \in V(G)$, G 中与顶点 v 关联的边的数目称为 v (在 G 中) 的度。

4) G 与 S 是两个图, 且 $V(S) \subseteq V(G)$, $E(S) \subseteq E(G)$, 则 S 是 G 的子图, 记作 $S \subseteq G$ 。设 V^* 是 $G = (V, E)$ 的顶点集的一个非空子集, 以 V^* 作为顶点集, 以两个端点均在 V^* 中的边的全体为边的子图称为由 V^* 导出的 G 的子图, 记为 $G[V^*]$, 则 $G[V^*]$ 是 G 的导出子图。

由概念格和二部图的定义可知, 每个形式背景都对应一个二部图, 其中二部图的顶点集为 $O \cup M$, 顶点 $o \in O$ 与 $m \in M$ 之间有一条边当且仅当 o 与 m 满足关系 I 。

下面用具体实例来说明以上定义。

例2 图1是例1中约简后的形式背景表2对应的二部图, 图2是表2对应的概念格, 原形式背景表1对应的概念格即为图3。

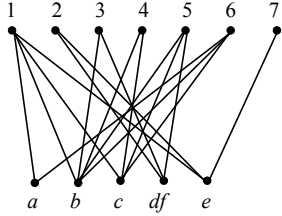


图1 表2对应的二部图
Fig. 1 Bipartite graph for table 2

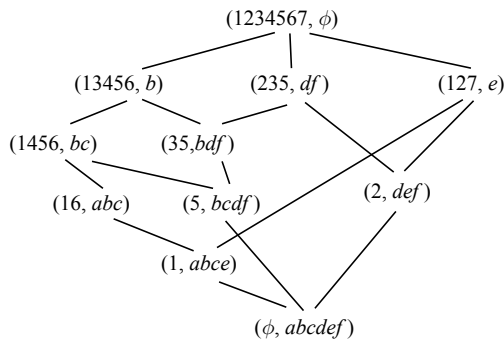


图2 表2对应的概念格
Fig. 2 Concept lattice for table 2

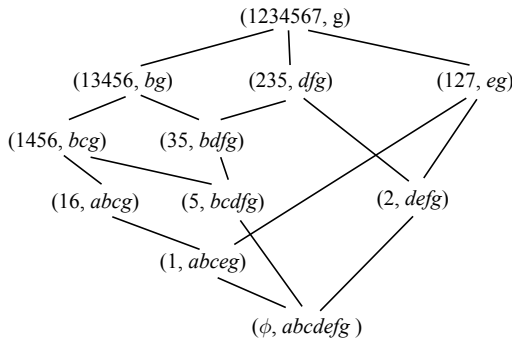


图3 表1对应的概念格
Fig. 3 Concept lattice for table 1

2 利用二部图生成概念格

形式背景中的数据, 在很多情况下, 对象的个数较大, 而属性的个数则相对较少, 在这种情况下, 基于属性来生成概念格, 可以有效地减少算法的执行时间。本文基于属性, 利用二部图子图和邻集的知识来生成概念格。

本节中的形式背景是指约简后的形式背景, 二部图也指约简后的形式背景所对应的二部图。

首先给出二部图的极大完全子图的定义。

定义4 若 S 是二部图 G 的子图, 且 S 是完全二部图, 对任意的 $v \in V(G) - V(S)$, G 由 $V(S) \cup v$ 导出的子图 $G[V(S) \cup v]$ 不是完全二部图, 则 S 称为二部图 G 的极大完全子图。

下面给出二部图的极大完全子图与概念的对应关系。

定理1 $S = (A, B; E)$ (其中 $A \subseteq O, B \subseteq M$) 是二部图 $G = (O, M; E)$ 的极大完全子图的充要条件为 (A, B) 是概念。

证明 由概念格的定义知充分性显然成立。

下证必要性, 用反证法。假设 (A, B) 不是概念, 则有 $A' \neq B$ 或 $B' \neq A$ 。如果 $A' \neq B$, 因为 $S = (A, B; E)$ 是完全二部图, 所以 B 中的顶点与 A 中的每个顶点都相邻, 故至少存在一个顶点 $m \in M$, 有 m 与 A 中的每个顶点都相邻, 即 $A' = B \cup m$ 。这样可以得到 $S_1 = (A, B \cup m; E^*)$ 为完全二部图, 与 S 是二部图 G 的极大完全子图矛盾, 故 $A' = B$ 。同理可证 $B' = A$ 。因此, (A, B) 是概念。

下面的定理2和定理3就是利用二部图的极大完全子图得到顶层概念 (O, ϕ) 的直接子概念。

定理2 形式背景 $K = (O, M, I)$ 对应的二部图 $G = (O, M; E)$, i 为图 G 的顶点集 M 中度数最大的顶点, $N(i) = P$, 则 i 与 P 构成的 G 的导出子图 $G[P \cup i]$ 是 G 的极大完全子图, 即 (P, i) 为概念, 且 (P, i) 为顶层概念 (O, ϕ) 的直接子概念。

证明 用反证法。假设 $G[P \cup i]$ 不是 G 的极大完全子图, 即存在 $o \in O - P$, 有 $S_1 = (P \cup o, i; E_1)$ 为完全二部图, 或存在 $m \in M$, 有 $S_2 = (P, i \cup m; E_2)$ 为完全二部图。若有第一种情况, 则顶点 i 与 $P \cup o$ 相邻, 此与 $N(i) = P$ 矛盾。若有第二种情况, 由完全二部图的定义知 $i \cup m$ 与 P 中每个顶点都相邻, 如果 m 的邻集恰好为 P , 则 i 与 m 应约简为 im , 与 i 和 m 是两个顶点矛盾; 如果 m 的邻集包含 P , 即至少存在 $o_1 \in O$, 使 m 与 $P \cup o_1$ 相邻, 则 m 的度数大于 i 的度数, 此与 i 为 M 中度数最大的顶点矛盾。因此, $G[P \cup i]$ 是 G 的极大完全子图, 即 (P, i) 为概念。且其为顶层概念 (O, ϕ) 的直接子概念, 因为不存在概念 C , 使 $(P, i) < C < (O, \phi)$ 。

定理3 若 j 是形式背景 $K = (O, M, I)$ 对应的二部图 $G = (O, M; E)$ 的顶点集 M 中除 i 外度数最大的顶点, $N(j) = P^*$ 。若 $P^* \not\subseteq P$, 则 (P^*, j) 为顶层概念 (O, ϕ) 的直接子概念; 若 $P^* \subset P$, 则以 P^* 为外延的概念应为 (P, i) 的子概念。

证明 同定理2。

由定理2和定理3可以得到顶层概念 (O, ϕ) 的直接子概念。

以下的定理4给出了利用二部图的子图来求 (O, ϕ) 的直接子概念 (P, i) 的直接子概念及其所有子概念的依据。

定理4 形式背景 $K = (O, M, I)$ 对应的二部图 $G = (O, M; E)$, i 是图 G 的属性集 M 中度数最大的

顶点, $N(i) = P$, 则 (P, i) 为形式背景 K 的概念。 G 的导出子图记为 $S = G[P \cup (M - i)]$, 与二部图 S 相对应的形式背景约简后记为 $K^* = (P, (M - i), I^*)$, 若 j 是 S 的顶点集 $M - i$ 中度数最大的顶点, $N(j) = Q$, 则 (Q, j) 为形式背景 K^* 的概念, $(Q, i \cup j)$ 为形式背景 K 的概念, 且 $(Q, i \cup j)$ 是 (P, i) 的直接子概念。

证明 (P, i) 为形式背景 K 的概念, (Q, j) 为形式背景 K^* 的概念, 前文已证。 下证 $(Q, i \cup j)$ 为形式背景 K 的概念, 即需证 $G[Q, i \cup j]$ 为二部图 G 的极大完全子图。

用反证法。 假设 $G[Q, i \cup j]$ 不是二部图 G 的极大完全子图, 那么可能有下列两种情况: 1) 至少存在 $o \in P - Q$, 使 $G[Q \cup o, i \cup j]$ 为完全二部图; 2) 至少存在 $m \in M - i - j$, 使 $G[Q, i \cup j \cup m]$ 为完全二部图。 如果出现第一种情况, 由完全二部图的定义知在图 S 中顶点 j 与 $Q \cup o$ 中的每个顶点相邻, 此与在形式背景 K^* 中 $N(j) = Q$ 矛盾。 如果出现第二种情况, 由完全二部图定义知在图 S 中 m 与集 Q 中的每个顶点相邻, 那么 m 在 S 中的邻集必包含 j 在 S 中的邻集, 这与形式背景已约简且 j 是 S 的顶点集 $M - i$ 中度数最大的顶点相矛盾。 所以 $(Q, i \cup j)$ 为形式背景 K 的概念, 并且不能找到集合 B , 使 $\{i\} \subset B \subset \{i \cup j\}$, 因此 $(Q, i \cup j)$ 是 (P, i) 的直接子概念。

根据定理 4 可以通过求 G 的导出子图的概念来简化形式背景, 并求出 (P, i) 的直接子概念。 这样经过反复的分解和约简, 使形式背景越来越简化, 同时也能求出 (P, i) 的所有子概念。

因此便能生成顶层概念的所有直接子概念以及它们各自的子概念。 这便是基于二部图的深度优先的概念格的迭代算法。

基于二部图的深度优先的概念格的迭代算法:

1) 最顶层概念为 (O, ϕ) , 标号为 1。

2) 形式背景 $K = (O, M, I)$ (约简后), 其对应的二部图为 $G = (O, M; E)$, 取属性 M 中度数最大的顶点 i , $N(i) = P$, 则 (P, i) 为顶层概念 (O, ϕ) 的直接子概念, 标号为 21。

3) 画出 G 的导出子图 $S_1 = G[P \cup (N(P) - i)]$ (约简后), 其对应的形式背景为 K_1 。 取二部图 S_1 的属性集 $N(P) - i$ 中度数最大的顶点 i_1 , $N(i_1) = P_1$, 则 (P_1, i_1) 为形式背景 K_1 的概念, $(P_1, i \cup i_1)$ 为形式背景 K 中概念 (P, i) 的直接子概念, 标号为 31。

4) 画出 S_1 的导出子图 $S_2 = S_1[P_1 \cup (N(P_1) - i_1)]$ (约简后), 其对应的形式背景为 K_2 , 取二部图 S_2 的属性集 $N(P_1) - i_1$ 中度数最大的顶点 i_2 , $N(i_2) = P_2$, 则 (P_2, i_2) 为形式背景 K_2 的概念, 故 $(P_2, i \cup i_1 \cup i_2)$ 为形式背景 K 中概念 $(P_1, i \cup i_1)$ 的直接子概念, 标号为 41。 继续下去, 直到 P_i 为单点集, 则 $(P_i, i \cup i_1 \cup i_2 \cdots \cup i_i)$

为形式背景 K 的倒数第二层概念, 它由导出子图 S_i 生成, 其直接子概念为最底层概念 (ϕ, M) , 标号为 $(i+2)1$, 最底层概念 (ϕ, M) 的标号为 $i+3$ 。

5) 从标号为 $r1$ 的概念回到生成它的导出子图 S_{r-2} , 考察二部图 S_{r-2} 的属性集中未考虑过的度数最大的顶点 i_{r-2}^* 。 如果 $N(i_{r-2}^*) = P_{r-2}^*$ 是已生成的概念的外延, 且当此已生成概念的内涵包含 $i \cup i_1 \cup i_2 \cdots \cup i_{r-2}^*$ 时, 说明以 P_{r-2}^* 为外延的概念及其子概念已生成, 则不必再考虑此顶点; 而当此已生成概念的内涵等于 $i \cup i_1 \cup i_2 \cdots \cup i_{r-2}^*$ 时, 则 (P_{r-2}^*, i_{r-2}^*) 是 (P_{r-1}, i_{r-1}) 的直接子概念, 且此概念及其子概念已生成 (已标号), 此顶点也不必再往下考虑。 如果 P_{r-2}^* 不是已生成概念的外延, 则 (P_{r-2}^*, i_{r-2}^*) 是 (P_{r-1}, i_{r-1}) 的直接子概念, 标号为 $(r-1)2$, 然后按步骤 3) 和步骤 4) 的方法画出 S_{r-2} 的导出子图 $S_{r-1}^* = S_{r-2}[P_{r-1}^* \cup (N(P_{r-1}^*) - i_{r-1}^*)]$, 接着再取未考虑过的最大度数顶点生成概念, 直到倒数第二层概念。

6) 从标号为 $(i+2)1$ 的概念起, 重复 5), 便可得到所有概念以及它们之间的父子关系。

3 实例

下面通过例 3 来具体说明此方法如何产生概念格。

例 3 形式背景同例 1, 其中的对象 $O = \{1, 2, 3, 4, 5, 6, 7\}$ 表示 1 班至 7 班共 7 个班级, 属性 $M = \{a, b, c, d, e, f, g\}$ 表示班级特色, 其中 a 表示团结, b 表示纪律严明, c 表示学风端正, d 表示卫生环境好, e 表示干部队伍过硬, f 表示英语四六级通过率高, g 表示课余活动丰富。

下面用基于二部图的深度优先的概念格的迭代算法来生成概念格。

1) 最顶层概念为 $(1234567, \phi)$, 标号为 1。

2) 画出形式背景 (约简后) 所对应的二部图 G , 其属性中度数最大的顶点为 b , $N(b) = \{1, 3, 4, 5, 6\}$, 故 $(13456, b)$ 为 $(1234567, \phi)$ 的直接子概念, 标号为 21。

3) 画出 G 的由 $P = \{1, 3, 4, 5, 6\}$ 和 $N(P) - b = \{a, c, d, f, e\}$ 导出的子图 S_1 , 二部图 S_1 的属性集中度数最大的顶点 c , $N(c) = \{1, 4, 5, 6\}$, 故 $(1456, bc)$ 为 $(13456, b)$ 的直接子概念, 标号为 31。

4) 画出 S_1 的由 $P_1 = \{1, 4, 5, 6\}$ 和 $N(P_1) - c = \{a, d, f, e\}$ 导出的子图 S_2 , 二部图 S_2 的属性集中度数最大的顶点 a , $N(a) = \{1, 6\}$, 故 $(16, abc)$ 为 $(1456, bc)$ 的直接子概念, 标号为 41。 画出 S_2 的由 $P_2 = \{1, 6\}$ 和 $N(P_2) - a = \{e\}$ 导出的子图 S_3 , 二部图 S_3 的属性集中度数最大的顶点 e , $N(e) = \{1\}$, 故 $(1, abce)$ 为 $(16, abc)$ 的直接子概念, 标号为 51。 此时 $P_3 = \{1\}$ 为单点

集, 故 $(1, abce)$ 为倒数第二层概念, 其直接子概念只有 $(\phi, abcdef)$, $(\phi, abcdef)$ 为最底层概念, 标号为6。

5) 从标号为51的概念 $(1, abce)$ 开始返回到生成它的二部图 S_3 , 其属性集中除 e 外无其他顶点。所以返回到标号为41的概念 $(16, abc)$ 的二部图 S_2 , 其属性集中除 a 外还有顶点 df 和 e 。先看 df , $N(df) = \{5\}$, 未出现过以5为外延的概念, 故 $(5, bcdf)$ 为 $(1456, bc)$ 的直接子概念, 标号为42, 此时 $P_2^* = \{5\}$ 为单点集, 故 $(5, bcdf)$ 为倒数第二层概念, 其直接子概念只有 $(\phi, abcdef)$ 。再看 e , $N(e) = \{1\}$, 已出现过以1为外延的概念 $(1, abce)$, 且其内涵 $abce$ 真包含 bce , 故此处不再考虑。

现在返回到生成标号为31的概念 $(1456, bc)$ 的二部图 S_1 , 其属性集中除 c 外还有顶点 a, df, e 。其中 $N(a) = \{1, 6\}$, $N(e) = \{1\}$, 已出现过以16和1为外延的概念 $(16, abc)$ 和 $(1, abce)$, 且其内涵真包含 ab 和 be , 故此二顶点不再考虑。而 $N(df) = \{3, 5\}$, 未出现过以35为外延的概念, 故 $(35, bdf)$ 为 $(13456, b)$ 的直接子概念, 标号为32。然后接着画出 S_1 的由 $P_1^* = \{3, 5\}$ 和 $N(P_1^*) - df = \{c\}$ 导出的子图 S_2^* , 二部图 S_2^* 的属性集中度数最大的顶点为 c , $N(c) = \{5\}$, 已出现过以5为外延的概念 $(5, bcdf)$, 此概念的内涵恰好为 $bcdf$, 则 $(5, bcdf)$ 也为 $(35, bdf)$ 的直接子概念, 此顶点不再考虑。

现在返回到生成标号为21的概念 $(13456, b)$ 的二部图 G , 其属性集中除 b 外, 还有 a, c, df 和 e 。其中 $N(a) = \{1, 6\}$, $N(c) = \{1, 4, 5, 6\}$, 已出现过以16和

1456为外延的概念 $(16, abc)$ 和 $(1456, bc)$, 且其内涵真包含 a 和 c , 故此二顶点不再考虑。 $N(df) = \{2, 3, 5\}$, 未出现过以235为外延的概念, 故 $(235, df)$ 为 $(1234567, \phi)$ 的直接子概念, 标号为22。然后接着画出 G 的由 $P^* = \{2, 3, 5\}$ 和 $N(P^*) - df = \{b, c, e\}$ 导出的子图 S_1^* , 二部图 S_1^* 的属性集中度数最大的顶点为 b , $N(b) = \{3, 5\}$, 已出现过以35为外延的概念 $(35, bdf)$, 此概念的内涵恰好为 bdf , 则 $(35, bdf)$ 也为 $(235, df)$ 的直接子概念, 此顶点不再考虑。 S_1^* 的属性集中除 b 外, 还有 c 和 e , $N(c) = \{5\}$, 已出现外延为5的概念 $(5, bcdf)$, 且其内涵真包含 cdf , 故此顶点不再考虑。 $N(e) = \{2\}$, 未出现过以2为外延的概念, 故 $(2, def)$ 为 $(235, df)$ 的直接子概念, 标号为33。此时其外延 $P_1^* = \{2\}$ 为单点集, 其直接子概念只有 $(\phi, abcdef)$ 。

返回二部图 G 中的属性 e , $N(e) = \{1, 2, 7\}$, 未出现过以127为外延的概念, 故 $(127, e)$ 为 $(1234567, \phi)$ 的直接子概念, 标号为23。接着画出 G 的由 $P^{**} = \{1, 2, 7\}$ 和 $N(P^{**}) - e = \{a, b, c, df\}$ 导出的子图 S_1^{**} , 二部图 S_1^{**} 的属性集中 a, b, c 的邻集相等, 约简后属性集中度数最大的顶点为 abc 和 df , $N(abc) = \{1\}$, $N(df) = \{2\}$, 已出现过以1和2为外延的概念 $(1, abce)$ 和 $(2, def)$, 其内涵恰好为 $abce$ 和 def , 则 $(1, abce)$ 和 $(2, def)$ 也为 $(127, e)$ 的直接子概念, 此二顶点不再考虑。

到此已生成所有概念, 见图4。其概念格见图3。

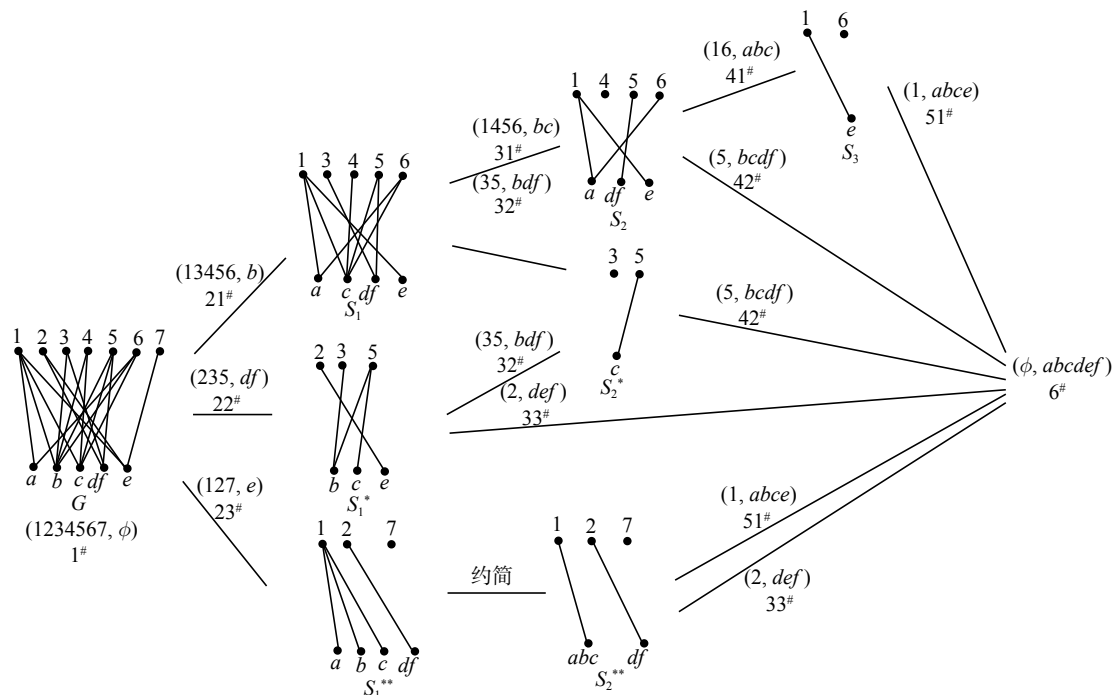


图4 生成概念格的过程

Fig. 4 Process of generating concept lattice

4 结束语

本文结合图论的内容,将概念格的形式背景和一个二部图相对应,利用二部图的极大完全子图来寻找概念,并且同时得到概念之间的父子关系,最终构造出概念格。此方法同时生成 Hasse 图,简单直观,能够快速生成概念格。基于概念格的形式背景与图论内容的高度关联性,二者之间其余理论的相互应用是我们下一进努力的方向。

参考文献:

- [1] GANTER B, WILLE R. Formal concept analysis: mathematical foundations[M]. FRANZKE C, trans. Berlin Heidelberg: Springer-Verlag, 1999.
- [2] WILLE R. Restructuring lattice theory: an approach based on hierarchies of concepts[M]//RIVAL I. Ordered Sets. Berlin Heidelberg: Springer, 1982: 445–470.
- [3] BELOHLAVEK R, SIGMUND E, ZACPAL J. Evaluation of IPAQ questionnaires supported by formal concept analysis[J]. Information sciences, 2011, 181(10): 1774–1786.
- [4] NGUYEN T T, HUI S C, CHANG Kuiyu. A lattice-based approach for mathematical search using formal concept analysis[J]. Expert systems with applications, 2012, 39(5): 5820–5828.
- [5] GODIN R, MISSAOUI R, ALAOUI H. Incremental concept formation algorithms based on Galois (concept) lattices[J]. Computational intelligence, 1995, 11(2): 246–267.
- [6] HO T B. Incremental conceptual clustering in the framework of Galois lattice[C]//LU H, MOTODA H, LIU H. KDD: Techniques and Applications. Singapore: World Scientific, 1997: 49–64.
- [7] 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.
ZHANG Wenxiu, WEI Ling, QI Jianjun. Attribute reduction theory and approach to concept lattice[J]. Science in China series E: information sciences, 2005, 35(6): 628–639.
- [8] ELLOUMI S, JAAM J, HASNAH A, et al. A multi-level conceptual data reduction approach based on the Lukasiewicz implication[J]. Information sciences, 2004, 163(4): 253–262.
- [9] AMILHASTRE J, VILAREM M C, JANSSEN P. Complexity of minimum biclique cover and minimum biclique decomposition for bipartite domino-free graphs[J]. Discrete applied mathematics, 1998, 86(2/3): 125–144.
- [10] BERRY A, SIGAYRET A. Representing a concept lattice by a graph[J]. Discrete applied mathematics, 2004, 144(1/2): 27–42.
- [11] BERRY A, MCCONNELL R M, SIGAYRET A, et al. Very fast instances for concept generation[M]//MIS-
SAOUI R, SCHMIDT J. Formal Concept Analysis. Berlin, Heidelberg: Springer, 2006, 3874: 119–129.
- [12] BERRY A, SANJUAN E, SIGAYRET A. Generalized domination in closure systems[J]. Discrete applied mathematics, 2006, 154(7): 1064–1084.
- [13] 李立峰, 刘三阳, 罗清君. 弦二部图的概念格表示[J]. 电子学报, 2013, 41(7): 1384–1388.
LI Lifeng, LIU Sanyang, LUO Qingjun. Representing chordal bipartite graph using concept lattice theory[J]. Acta electronic sinica, 2013, 41(7): 1384–1388.
- [14] 李立峰. 链图的概念格表示[J]. 计算机科学, 2014, 41(2): 264–266.
LI Lifeng. Chain graph and their concept lattice representation[J]. Computer science, 2014, 41(2): 264–266.
- [15] 张涛, 洪文学, 路静. 形式背景的属性树表示[J]. 系统工程理论与实践, 2011, 31(S2): 197–202.
ZHANG Tao, HONG Wenxue, LU Jing. Attribute tree representation for formal context[J]. Systems engineering-theory and practice, 2011, 31(S2): 197–202.
- [16] 张涛, 任宏雷. 形式背景的属性拓扑表示[J]. 小型微型计算机系统, 2014, 35(3): 590–593.
ZHANG Tao, REN Honglei. Attribute topology of formal context[J]. Journal of Chinese computer systems, 2014, 35(3): 590–593.
- [17] 张涛, 路静, 任宏雷. 一种基于树图的属性约简算法[J]. 小型微型计算机系统, 2014, 35(1): 177–180.
ZHANG Tao, LU Jing, REN Honglei. An algorithm based on free graph for calculation for attribute reduction[J]. Journal of Chinese computer systems, 2014, 35(1): 177–180.
- [18] 黄天民, 徐扬, 赵海良, 等. 格、序引论及其应用[M]. 成都: 西南交通大学出版社, 1998.
- [19] 王树禾. 图论[M]. 北京: 科学出版社, 2009.
- [20] 肖位枢. 图论及其算法[M]. 北京: 航空工业出版社, 1993.

作者简介:



窦林立, 女, 1975 年生, 硕士研究生, 讲师, 主要研究方向为计算机数学、离散数学。参与完成多项省级和市级课题。发表学术论文 6 篇。



展正然, 女, 1981 年生, 硕士研究生, 讲师, 主要研究方向为微分方程。发表学术论文 9 篇, 被 EI 检索 2 篇, SCI 检索 1 篇, 出版教材 1 部。