

文章编号: 1001-0920(2010)07-0987-06

混合概念格在案例相似性度量中的应用

鞠可一¹, 周德群¹, 吴君民²

(1. 南京航空航天大学 经济与管理学院, 南京 210016; 2. 江苏科技大学 经济与管理学院, 江苏 镇江 212003)

摘要: 传统的案例相似性度量方法无法很好地判断复杂案例库系统中某些特殊属性(如数值属性、模糊属性)的相似程度, 从而限制了案例推理(CBR)系统在判断、评估、决策、推理方面的进一步应用. 针对此问题提出混合概念格的方法, 对案例库中可能涉及到的数字、符号、模糊属性等进行相似性度量. 实验表明, 该方法较好地解决了混合案例属性的相似性度量问题, 提高了检索结果的查全率和查准率.

关键词: 混合概念格; 案例推理; 相似性度量

中图分类号: TP18; TH16.21

文献标识码: A

Multi-galois lattice for similarity measurement in case retrieving

JU Ke-yi¹, ZHOU De-qun¹, WU Jun-min²

(1. College of Economics and Management, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China; 2. College of Economics and Management, Jiangsu University of Science and Technology, Zhenjiang 212003, China. Correspondent: JU Ke-yi, E-mail: jky1983@163.com)

Abstract: The key issues for case retrieve are the similarity measurement and matching process. Classical methods for case retrieve can't meet the demands of those complex systems, which are of special features, such as textual or fuzzy attributes, so that case-based reasoning (CBR) system can't go further in judging, evaluating, deciding, and reasoning. A method, multi-galois lattice (MGL), is presented to do the similarity measurement for numerical, symbolic, fuzzy concept, and fuzzy interval attributes. Experiment results show that MGL can not only simplify the matching process, but also resolve the problems of similarity measurement for multi-attributes cases.

Key words: Multi-galois lattice; Case-based reasoning; Similarity measurement

1 引言

案例推理(CBR)是近年来人工智能界一直看好的一种方法, 在处理复杂问题和进行多属性决策时, CBR往往是优选方法. 基于案例的推理过程通常包括: 检索(Retrieve), 重用(Reuse), 修正(Revise)和保存(Retain)4个阶段(简称“4R”). 其中, 案例检索是CBR系统开展的关键, 没有有效的案例检索, CBR系统将一无所获, 后续工作便无法开展. 案例相似性度量是案例检索过程的核心, 传统的相似性度量方法无法很好地判断案例库中不同属性特征(如模糊属性、数字属性等)的相似度, 从而限制了CBR系统在判断、评估、决策、推理方面的进一步应用.

概念格, 又称形式概念分析(FCA), 也称 Galois 格, 是由 Wille 于 1982 年提出的^[1]. 它提供了一种从形式背景到建立概念格来进行数据分析的有效工具,

被广泛应用于知识管理、机器学习、软件工程和信息获取等许多领域^[2-6]. 概念格生动简洁地体现了概念之间的特化与泛化关系, 与其相对应的 Hasse 图实现了对数据的可视化操作.

随着属性类型的不断丰富, 传统二值属性概念格已不能满足实际应用. 作为扩展, 文献[7]提出了符号属性概念格, 随后很多学者在此基础上作了进一步的研究^[8-10]. 但是, 在当前的复杂大系统(如油价波动产生的影响)案例库中, 其属性具有多种不同的表现形式(如数字属性、符号属性、模糊概念属性和模糊区间值属性等), 因此使用概念格对复杂多值背景进行研究时, 需采取更有效的解决方法. 鉴于此, 本文提出了混合概念格(MGL)的方法, 既解决了传统概念格方法只能解决标准化形式背景的缺陷, 又很好地支持了案例检索过程, 提高了案例推理的查全率和查准率.

收稿日期: 2009-06-12; 修回日期: 2009-10-09.

基金项目: 国家社会科学基金重大项目(08&ZD046); 国家自然科学基金项目(70873058).

作者简介: 鞠可一(1983-), 女, 江苏靖江人, 博士, 从事决策支持系统、知识管理等研究; 周德群(1963-), 男, 江苏盐城人, 教授, 博士生导师, 从事能源系统工程等研究.

2 基本概念

2.1 案例推理

案例推理来源于人类认知和心理活动,属于一种类比推理的机器学习方法,是帮助人们获取实例样本性知识的手段^[11,12]. 现实世界中很多隐性知识很难用规则、公式等方式表达出来,但可用案例的形式对这些知识加以利用. 通常认为 Kolodner 对现实世界的 4 个假设是 CBR 方法的基础,它们分别是^[13]: 正则性、典型性、一致性(也称相对稳定性)和易适应性. CBR 的工作流程如图 1 所示:

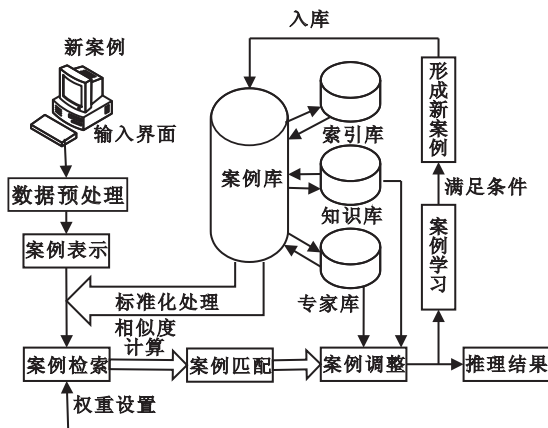


图 1 CBR 系统工作过程

案例检索是整个案例推理过程的关键. 它的主要任务是检索案例库, 计算案例库中的案例与待解决问题之间的匹配程度, 从而获得与新问题相似的案例. 案例检索的实质是案例的相似性度量, 常用的衡量方法包括最近邻法、粗糙集法和模糊集法. 其中最近邻法最为直观, 因而是普遍采用的方法. 但是, 这些方法的处理过程比较复杂, 且有各自的侧重点和适用范围, 不适用于某些特定复杂领域的问题求解. 因此, 寻找一个既适用于混合多属性特征, 同时具有较高的查全率和查准率的案例匹配方法便显得尤为重要.

2.2 概念格

概念格理论是以序理论和完备格理论为基础, 依据数据库中提供的基本信息建立的一种刻画对象与属性之间关系的数学结构^[14].

定义 1 假设给定形式背景 K 为三元组 $K = (G, M, I)$. 其中 G 为所有对象的集合, M 为所有属性的集合, $I(I \subseteq G \times M)$ 为 G 和 M 中元素之间关系的集合. 对于 $g \in G, m \in M, (g, m) \in I$, 表示“对象 g 具有属性 m ”.

定义 2 对于对象子集 $A \subseteq G$, 定义: $\forall A \in G: A' = f(A) = \{m \in M | \forall g \in A, (g, m) \in I\}$, 表示“ A 中全体对象所共有的属性集”; 相应地, 对于属性子集 $B \subseteq M$, 定义: $B' = h(B) = \{g \in G | \forall m \in B, (g, m) \in I\}$, 表示“同时具有 B 中所有属性的对象

的集合”.

定义 3 设形式背景 (G, M, I) 中的一个形式概念 $C = (A, B), A \subseteq G, B \subseteq M$. 若其满足 $A' = B, B' = A$, 则 A 和 B 分别称为形式概念 (A, B) 的外延和内涵. 用 $\text{Intent}(C)$ 表示节点 C 的内涵, 用 $\text{Extent}(C)$ 表示节点 C 的外延.

定义 4 设 $C_1 = (A_1, B_1)$ 和 $C_2 = (A_2, B_2)$ 是一个形式背景的两个形式概念, 如果 $A_1 \subseteq A_2$ (等同于 $B_1 \supseteq B_2$), 则 (A_1, B_1) 被称为 (A_2, B_2) 的子概念, (A_2, B_2) 被称为 (A_1, B_1) 的超概念, 记为 $(A_1, B_1) \leq (A_2, B_2)$. 关系 \leq 为形式概念之间的“序”.

定义 5 如果 C_1 是 C_2 的子概念且不存在其他的节点 C_3 , 满足 $C_1 \leq C_3 \leq C_2$, 则称 C_1 是 C_2 的子节点(直接后继), 而 C_2 是 C_1 的父节点(直接前驱). 按此方法构成的有序 (G, M, I) 所有形式概念的集合构成的完备格称为概念格, 表示为 $L(G, M, I)$. 根据序关系可生成概念格的 Hasse 图.

定义 6 关系 \leq 是集合 $L(K)$ 上的一个偏序, 其相应的下确界和上确界定义为

$$\inf(A_t, B_t) = \bigwedge_{t \in T} (A_t, B_t) = \left(\bigcap_{t \in T} A_t, \left(\bigcup_{t \in T} B_t \right)' \right),$$

$$\sup(A_t, B_t) = \bigvee_{t \in T} (A_t, B_t) = \left(\left(\bigcup_{t \in T} A_t \right)', \bigcap_{t \in T} B_t \right).$$

其中: $(A_t, B_t) \in L(K), T$ 是指标集.

不难发现, 概念格展现的是概念之间的泛化与特化的偏序关系, 每个概念以结点表示, 如果概念 A 在概念 B 之上并且二者有关联, 则称概念 A 是概念 B 的泛化. 从概念格的定义和性质可以看出, 概念格不仅精确定义了概念, 更重要的是它描述了概念之间的继承关系, 为度量概念的相似性提供了良好的数据结构和理论依据, 因此将其应用于 CBR 的案例检索过程无疑是可行的.

3 基于混合概念格理论的相似性度量

3.1 属性分类

在复杂大系统案例库中, 其包含的属性及其取值范围已有较大扩展. 将案例的属性按照其具有的属性值类型进行划分, 分为^[15]:

1) 确定数字属性: 属性值为数字型, 可以是连续的, 也可以是离散的;

2) 确定符号属性: 属性值通常为一个明确的术语;

3) 模糊数/区间属性: 属性值为一个模糊数或是一个没有确定边界的模糊区间;

4) 模糊概念属性: 属性值是一个概念变量, 所有的属性值构成一个项目集, 项目集中的每一个项目对应一个模糊概念.

3.2 混合概念格构造

概念格的构造算法可分为两类^[16]: 批处理构造算法^[17]和渐近式构造算法^[18]. 实践表明, 渐近式算法具有较优的性能^[19]. 本文也采用渐近式思想, 针对不同的属性类型, 构建不同类型的概念格.

渐近式构造概念格是在给定原始形式背景 $K = (G, M, I)$ 所对应的初始概念格 $L = (CS(K), \leq)$ 以及新增对象 x^* 的情况下, 求解形式概念背景 $K^* = (G \cup \{x^*\}, M, I)$ 所对应的概念格 $L^* = (CS(K^*), \leq)$ 的过程. 对初始概念格中的每个节点, 根据它和新增对象 x^* 及特征集 $f(x^*)$ 之间的关系, 定义它们的归属.

3.2.1 确定数字概念格构造

在已有的众多概念格构造方法中, 确定数字型属性往往会被舍弃, 原因是它们不可归纳, 且会使得概念格扁平化, 不利于发现概念间的泛化-特化关系^[20]. 但在案例检索过程中, 数字属性非常重要, 很多案例推理过程的最终目的就是要获得其中的某些数据结论(如油价涨跌幅度对某些重要经济数据的影响). 因此, 对确定数字属性概念格的构建, 在案例推理中是不能被忽略的. 确定数字属性背景的概念格构造算法设计如下:

Step 1 设初始 Galois 格图中包含两个元素 (G, \emptyset) 和 (\emptyset, M) . 假设需要添加进 Galois 格的形式概念为 $(\{x^*\}, \{f(x^*)\})$, L_1 为原格节点, L_2 为新增格节点, L_1^* 和 L_2^* 分别为变化后的形式概念, 下同.

Step 2 为了减少扫描次数, 提高算法效率, 首先要将形式概念按属性个数由少到多排列.

Step 3 扫描每个节点(为了叙述方便, 下文将不再区分“形式概念”与“Galois 格节点”, 统称为“节点”), 更改 Hasse 图的节点结构:

1) 如果已有节点中的属性全部出现在新增节点的属性中, 则该节点需要被修改:

- ① $L_m^*(H) = (\{G \cup \{x^*\}\}, M)$;
- ② 将待增节点作为其子节点, $L_m^*(H) \rightarrow L^*$;
- ③ 原有边保持不变, $L_m^*(H) \rightarrow L$;
- ④ 修改节点添加到形式概念集合 L_c 中, $L_c := L_m^*(H)$.

2) 如果已有节点中部分包含新增节点的属性值, 则将新节点增加到 Galois 中:

- ① 新节点的形式概念如下: $L_n^*(H) = (G(H) \cup \{x^*\}, G'(H) \cap f(\{x^*\}))$;
- ② 将待增节点作为其子节点, $L_n^*(H) \rightarrow L^*$;
- ③ 原有边保持不变, $L_n^*(H) \rightarrow L$;
- ④ 新增节点添加到形式概念集合 L_c 中, $L_c :=$

$L_n^*(H)$.

3) 如果新增概念的属性值已经在原格图中出现, 则不添加该节点.

Step 4 L_c 中存放的是 $f(\{x^*\})$ 所有子概念节点的形式概念. 概念仍然按照属性个数由少到多的顺序排列;

Step 5 更改 Hasse 图的边结构:

1) 扫描 L_c 中每一个格节点, 如果其子节点的属性被包含于 $f(\{x^*\})$ 中, 则修改边的走向: 断开 $L_c \rightarrow L^*$ 的连接, 同时创建 $L_c^{\text{child}} \rightarrow L^*$ 的连接;

2) 否则, 边结构不变.

3.2.2 确定符号概念格构造

确定符号概念格属性的值域范围一般小于确定数字属性的值域范围, 其概念格的构造过程与确定数字概念格构造过程相似, 在此不再赘述.

3.2.3 模糊数字概念格构造

常见的模糊数有两种: 三角模糊数和梯形模糊数. 由于三角模糊数可看作梯形模糊数的特例, 因此本文主要研究梯形模糊数概念格. 用 L-R 型函数表示其隶属度函数如下:

$$\mu_M(x) = \begin{cases} L(\frac{m-x}{p}), & x \leq m; \\ 1, & m \leq x \leq \bar{m}; \\ R(\frac{x-\bar{m}}{q}), & x \geq \bar{m}. \end{cases} \quad (1)$$

其中: \bar{m}, m, p, q 是参数, 三角模糊数 $\underline{m} = \bar{m}$, p 和 q 的大小随属性的不同而不同. 一般取 $p = c\underline{m}, q = c\bar{m}$, 默认 $c = 0.1$. 模糊数字概念格构造算法如下:

Step 1 初始状态同 3.2.1.

Step 2 形式概念仍按属性个数由少到多排列.

Step 3 扫描每个节点, 更改 Hasse 图的节点结构:

1) 如果已有节点中的属性全部出现在新增节点的属性中, 则该节点需要被修改:

- ① $L_m^*(H) = (\{G \cup \{x^*\}\}, M)$.
- ② 将待增节点作为其子节点, $L_m^*(H) \rightarrow L^*$.
- ③ 原有边保持不变, $L_m^*(H) \rightarrow L$.
- ④ 修改节点添加到形式概念集合 L_c 中, $L_c := L_m^*(H)$.

⑤ 计算节点之间的隶属度. 在此之前, 先要对模糊形式背景进行处理, 求出 L_1^* 和 L_2^* 模糊数属性的各个参数: \bar{m}, m, p, q ^[15].

$$\begin{cases} m(L_1^*) = \frac{qL_1 \underline{m}_{L_2} + pL_2 \bar{m}_{L_1}}{pL_2 + qL_1}, \\ m(L_2^*) = 1 - \frac{m(L_1^*) - \bar{m}_{L_1}}{qL_1}. \end{cases} \quad (2a)$$

接着, 计算 L_1^* 和 L_2^* 二者的隶属度

$$\text{membership}(L_1^*, L_2^*) = \begin{cases} 0, & m(L_2^*) \leq 0; \\ \frac{(\overline{m}_{L_1} + q_{L_1} - \underline{m}_{L_2} + p_{L_2})}{2} m(L_2^*), & m(L_2^*) \in (0, 1); \\ \frac{2\overline{m}_{L_1} + q_{L_1} - 2\underline{m}_{L_2} + p_{L_2}}{2}, & m(L_2^*) = 1. \end{cases} \quad (2b)$$

2) 如果已有节点中部分包含新增节点的属性值, 则将新节点增加到 Galois 中:

① 新节点的形式概念如下: $L_n^*(H) = (G(H) \cup \{x^*\}, G'(H) \cap f(\{x^*\}))$;

② 将待增节点作为其子节点, $L_n^*(H) \rightarrow L^*$;

③ 原有边保持不变, $L_n^*(H) \rightarrow L$;

④ 新增节点添加到形式概念集合 L_c 中, $L_c := L_n^*(H)$;

⑤ 计算节点之间的隶属度

$$\begin{cases} m(L_1^*) = \frac{q_{L_1} \underline{m}_{L_2} + p_{L_2} \overline{m}_{L_1}}{p_{L_2} + q_{L_1}}, \\ m(L_2^*) = 1 - \frac{m(L_1^*) - \overline{m}_{L_1}}{q_{L_1}}. \end{cases} \quad (3a)$$

$\text{membership}(L_1^*, L_2^*) =$

$$\begin{cases} 0, & m(L_2^*) \leq 0; \\ \frac{(\overline{m}_{L_1} + q_{L_1} - \underline{m}_{L_2} + p_{L_2})}{2} * m(L_2^*), & m(L_2^*) \in (0, 1); \\ \frac{2\overline{m}_{L_1} + q_{L_1} - 2\underline{m}_{L_2} + p_{L_2}}{2}, & m(L_2^*) = 1. \end{cases} \quad (3b)$$

3) 如果新增概念的属性值已经在原格图中出现, 则不添加该节点.

Step 4 L_c 中存放的是 $f(\{x^*\})$ 所有子概念节点的形式概念. 概念仍按属性个数由少到多的顺序排列, 以提高算法效率.

Step 5 更改 Hasse 图的边结构:

1) 扫描 L_c 中每一个格节点, 如果其子节点的属性被包含于 $f(\{x^*\})$ 中, 则修改边的走向: 断开 $L_c \rightarrow L^*$ 的连接, 同时创建 $L_c^{\text{child}} \rightarrow L^*$ 的连接;

2) 否则, 边结构不变.

3.2.4 模糊概念概念格构造

模糊概念概念格与模糊属性概念格的构造方式基本相似, 所不同的是, 模糊概念的梯形隶属度函数的参数 p 和 q 往往由领域专家确定. 鉴于篇幅限制, 模糊概念属性的概念格构造也不再详述.

3.3 混合概念格相似性度量

从直观上看, 相似度是概念间距离的单调递减函数, 距离越大, 其相似度越低; 反之, 两个概念背

景在 Hasse 图中的距离越小, 位置越接近, 其相似程度就越大. 另一方面, 概念所处的层次(即概念所处的深度)决定了概念间的相似程度, 概念之间的相似程度随着它所处深度的增加而增加. 因为层次越深, 表示分类越细致, 概念之间的相似程度越高^[21]. 在案例库系统中, 人们更为关注的是案例的属性, 反映在混合概念格中便是形式背景的属性. 鉴于此, 本文将混合概念格相似度定义如下: 设 $L_1(G_1, M_1, I)$ 和 $L_2(G_2, M_2, I)$ 为概念格中的两个形式背景, 其相似度为

$$\text{sim}(L_1, L_2) = \left(\frac{|G(L_1) \cap G(L_2)|}{\max(|G(L_1)|, |G(L_2)|)} + \sum_{i=1}^n \frac{|M(L_{1i}) \cap M(L_{2i})|}{\max(|M(L_{1i})|, |M(L_{2i})|)} * a_i \right) * (1 + b)^{\min(l_1, l_2)}. \quad (4)$$

式中

$$a_i = \begin{cases} 1, & M_{1i} = M_{2i}; \\ \text{membership}(L_1, L_2), & M_{1i} \cap M_{2i} \neq \emptyset; \\ 0, & M_{1i} \cap M_{2i} = \emptyset. \end{cases}$$

其中: l_1, l_2 表示形式概念 L_1, L_2 在 Hasse 图中所处的层次; b 表示节点所处层次对相似性的影响程度, 取 $b = 0.1$.

4 案例分析

本文构造了一个旅行案例, 其中包括旅游的人数(Person), 旅行的方式(Holiday), 价格区间(Price), 旅游持续时间(Duration)及食宿要求(Housing)等5项指标, 并各自包含其属性值, 如表1所示.

表 1 一个旅行案例

	Persons	Holiday	Price	Duration	Housing
Case1	2	Skiing	Expensive	About 3	*****
Case2	5	Active	Economical	6-7	****
Case3	3	Wandering	Cheap	1-2	(No Need)
Case4	8	Skiing	Economical	About 3	***
Case5	6	Active	Economical	About 8	**

这些指标及其各自的指标值基本涵盖了上述几类属性, 并借由此案例验证混合概念格在案例相似性度量中的应用.

4.1 传统概念格生成

根据传统概念格渐近式的构造方法, 得到其 Hasse 图如图2所示. 由传统概念格得到的相似性度量如表2所示.

4.2 混合概念格生成

图2表明, 传统的概念格构造方法无法涵盖模糊概念属性. 如 Case 2 和 Case 3, 在传统意义上, 没有共同包含的属性, 是不相似的. 但 Price 是一个模糊概念

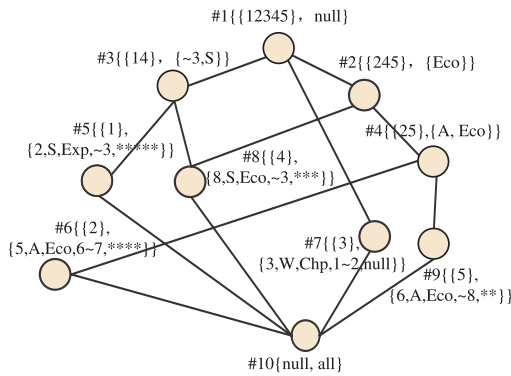


图2 传统 Galois 格的 Hasse 图

表2 传统概念格的案例相似度

	Case1	Case2	Case3	Case4	Case5
Case1	*	0	0	0.53	0
Case2	0	*	0	0.27	0.59
Case3	0	0	*	0	0
Case4	0.53	0.27	0	*	0.27
Case5	0	0.59	0	0.27	*

属性, 由图3可知, Eco和Chp有重叠部分, 因此Case2和Case3应当是有关联的, 其隶属度为 membership $(Eco \cap Chp) = 0.178$. 鉴于此, 将 Multi-Galois 格如图4所示构造, 其中新增格节点加黑, 其边为; 修改格节点由“下划线”表示, 边不变.

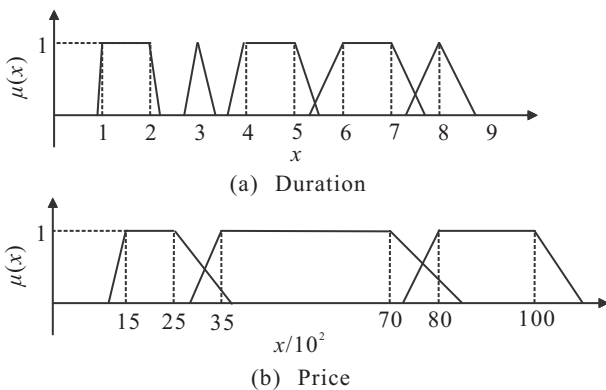


图3 模糊属性隶属度

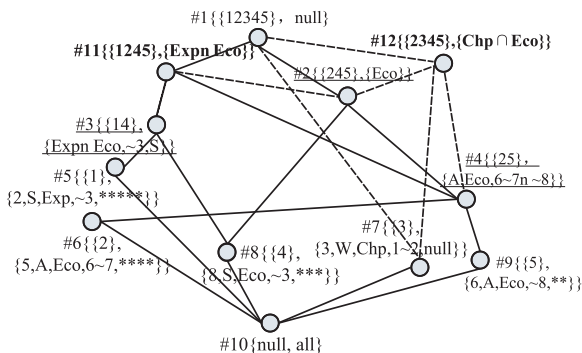


图4 混合 Galois 格 Hasse 图

由式(4)及混合概念格 Hasse 图可得混合概念格下案例间的相似度, 如表3所示.

以 Case1 行为例, 通过表2与表3的对比不难发

表3 混合概念格的案例相似度

	Case1	Case2	Case3	Case4	Case5
Case1	*	0.1	0	0.68	0.1
Case2	0.1	*	0.05	0.29	0.61
Case3	0	0.05	*	0.05	0.05
Case4	0.68	0.29	0.05	*	0.29
Case5	0.1	0.61	0.05	0.29	*

现, 混合概念格的使用使得 Case 1 与 Case 2, Case 4, Case 5 之间的相似度有了明显的增加. 因此, 可以认为混合概念格的提出有利于增加案例相似的匹配度, 并提高了案例的查全率和查准率.

4.3 基于混合概念格的案例匹配

现有待匹配新案例 Case6: $\{x^*\} = \{6\}, \{f(x^*)\} = \{9, W, Chp, 4-5, *\}$, 将其添加到混合概念格中, 得到新 Galois 格图如图5所示. 案例相似性度量如表4所示.

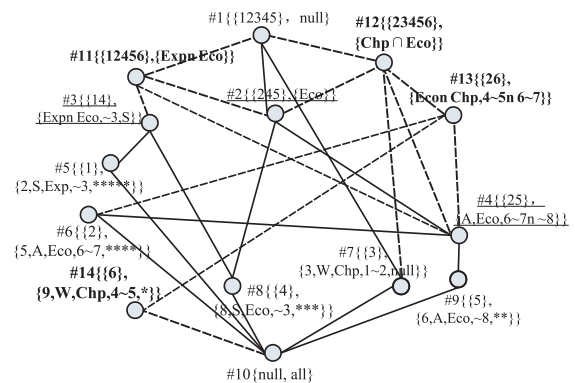


图5 Case 6 加入后的 Multi-Galois 格图

表4 案例相似性度量

	sim(6-i)(GL)	sim(6-i)(MGL)
Case1	0	0
Case2	0	0.049
Case3	0.53	0.53
Case4	0	0.049
Case5	0	0.049

以 #6, #14 节点为例, 在传统概念格中, 其对象、属性均无交集, 决策者无法发现他们两者之间的相关性. 但在混合概念格中, 它们是相似的, 其相似度为

$$\text{sim}(6, 2) = ((1/5) * 0.178 + (1/5) * 0.0045) * (1 + 0.1)^{\min(3,3)} = 0.0486.$$

由此可知, Case2 与 Case6 是相似案例, 在案例匹配过程中应当考虑, 而不能忽略或舍弃. 由此进一步表明, 混合概念格能够避免概念检索的缺失, 提高检索结论的准确性和精确度.

5 结论

案例检索是案例推理过程的核心,而案例检索的实质就是相似性度量.概念格是对概念之间泛化/特化关系的形象表示,将其用于案例检索过程,能够帮助系统准确直观地给出案例之间的依存关系.混合概念格是对经典概念格的扩展,它不仅能对0,1模式的标准形式背景进行研究,还可以应用到不确定型属性的表示中,更加符合实际情况和现实需求.实验表明,混合概念格不但是一种简便易行的案例检索方法,而且提高了案例检索过程的查全率和查准率,避免了案例检索的遗漏和缺失,符合案例匹配的要求.总之,混合概念格的思想为案例推理的检索过程提供了一种切实有效的新途径.

参考文献(References)

- [1] Wille R. Restructuring lattice theory: An approach based on hierarchies of concepts[C]. *Ordered Sets*. Dordrecht: Reidel, 1982: 445-470.
- [2] Krohn U, Davies N J, Weeks R. Concept lattices for knowledge management[J]. *BT Technology J*, 1999, 17(4): 108-116.
- [3] Zupa B, Bohance M. Learning by discovering concept hierarchies[J]. *Artificial Intelligence*, 1999, 109(1/2): 211-242.
- [4] Tonella. Using a concept lattice of decomposition slices for program understanding and impact analysis[J]. *IEEE Trans on Software Engineering*, 2003, 29(6): 495-509.
- [5] Valtchev P, Missaoui R, Godin R, et al. Generating frequent itemsets incrementally: Two novel approaches based on galois lattice theory[J]. *J of Experimental & Theoretical Artificial Intelligence*, 2002, 14(2/3): 115-142.
- [6] 曲开社, 翟岩慧. 偏序集、包含度与形式概念分析[J]. *计算机学报*, 2006, 29(2): 219-226.
(Qu K S, Zhai Y H. Posets, anclusion degree theory and FCA[J]. *Chinese J of Computers*, 2006, 29(2): 219-226.)
- [7] Brito P. Order structure of symbolic assertion objects[J]. *IEEE Trans on Knowledge and Data Engineering*, 1994, 6(5): 830-835.
- [8] Diday E, Emilion R. Maximal and stochastic galois lattices[J]. *Discrete Applied Mathematics*, 2003, 127(2): 271-284.
- [9] Wu Q, Liu Z T. Real formal concept analysis based on grey-rough set theory[J]. *Knowledge-Based Systems*, 2009, 22(1): 38-45.
- [10] Polaillon G, Diday E. Reduction of symbolic Galois lattices via hierarchies[C]. *Proc of the Conf on Knowledge Extraction and Symbolic Data Analysis*. Luxembourg, 1999: 137-143.
- [11] 郭艳红, 邓贵仕. 基于事例的推理(CBR)研究综述[J]. *计算机工程与应用*, 2004, 40(21): 1-5.
(Guo Y H, Deng G S. Review of case-based reasoning[J]. *Computer Engineering and Applications*, 2004, 40(21): 1-5.)
- [12] 史忠值. 知识发现[M]. 北京: 清华大学出版社, 2002.
(Shi Z Z. *Knowledge discovery*[M]. Beijing: Tsinghua University Press, 2002.)
- [13] Leake D B. *Case-based reasoning*[M]. Cambridge: AAI Press/MIT Press, 1996.
- [14] Ganter B, Willer R. *Formal concept analysis mathematical foundations*[M]. Berlin: Springer, 1999.
- [15] 张本生, 于永利. CBR系统案例搜索中的混合相似性度量方法[J]. *系统工程理论与实践*, 2002, 3(3): 131-136.
(Zhang B S, Yu Y L. Hybrid similarity measure for retrieval in case-based reasoning system[J]. *System Engineering-Theory & Practice*, 2002, 3(3): 131-136.)
- [16] 胡可云, 陆玉昌, 石纯一. 概念格及其应用进展[J]. *清华大学学报*, 2000, 40(9): 77-81.
(Hu K Y, Lu Y C, Shi C Y. Advances in concept lattice and its application[J]. *J of Tsinghua University*, 2000, 40(9): 77-81.)
- [17] Nourine L, Raynaud O. A fast algorithm for building lattices[J]. *Information Processing Letters*, 1999, 71(5): 199-204.
- [18] Godin R, Missaoui R, Alaoui H. Incremental concept formation algorithms based on calois (concept) lattices[J]. *Computational Intelligence*, 1995, 11(2): 246-267.
- [19] 刘宗田, 强宇, 周文, 等. 一种模糊概念格模型及其渐近式构造算法[J]. *计算机学报*, 2007, 30(2): 184-187.
(Liu Z T, Qiang Y, Zhou W, et al. A fuzzy concept lattice model and its incremental construction algorithm[J]. *Chinese J of Computers*, 2007, 30(2): 184-187.)
- [20] 王德兴, 胡学刚, 刘晓平. 量化扩展概念格的属性归纳及多粒度规则挖掘[J]. *系统工程学报*, 2009, 24(1): 54-61.
(Wang D X, Hu X G, Liu X P. Research on attribute-oriented induction based on quantitative extended concept lattice and multi-granule association rule mining[J]. *J of Systems Engineering*, 2009, 24(1): 54-61.)
- [21] 智慧来, 智东杰, 刘宗田, 等. 基于概念格的概念相似度计算[J]. *计算机科学*, 2008, 35(9): 156-157.
(Zhi H L, Zhi D J, Liu Z T, et al. Concept similarity based on concept lattice[J]. *Computer Science*, 2008, 35(9): 156-157.)