

# 形式概念分析及其应用进展

陈世祺

(大连海事大学 信息科学与技术学院 辽宁 大连 116026)

**摘 要** 形式概念分析是一种层次化的形式对象分析方法,能够从二元关系中挖掘出具有共同形式属性的一组形式对象的聚集。近十几年来,形式概念分析技术已在本体研究、软件工程、知识发现、Web 语义检索等领域取得成功。本文首先介绍了形式概念分析的相关概念,再从本体研究、软件工程、知识发现、Web 语义检索 4 个领域对形式概念分析的发展前沿和研究热点进行阐述。并在此基础上对未来做出展望。

**关键词** 形式概念分析, 本体研究, 软件工程, 知识发现, Web 语义检索

## Formal Concept Analysis and Its Application

Chen ShiQi

(Dalian Maritime University, Computer Science and Technology,  
Liaoning, Dalian, 116026, China)

**Abstract** Formal concept analysis is a hierarchical formal object analysis method, which can extract a set of formal objects with common form attributes from binary relations. Over the past decade, formal concept analysis technology has been successful in ontology research, software engineering, knowledge discovery, Web semantic retrieval and other fields. This paper first introduces the related concept of formal concept analysis, and then expounds the development foreground and research hotspot of formal concept analysis from four domains: ontology research, software engineering, knowledge discovery and Web semantic retrieval. And on the basis of the future to make prospects.

**Keywords** Formal concept analysis, Ontology research, Software Engineering, Knowledge Discovery, Web Semantic Retrieval

### 1. 绪论

在计算机与网络信息技术飞速发展的今天,各个领域的信息与数据急剧增加,并且由于人类的参与使数据与信息中的不确定性更加显著,信息与数据中的关系更加复杂。如何从大量杂乱无章和强干扰的数据中挖掘潜在的、新颖的、正确的、有利的价值知识,这给智能信息处理提出了严峻的挑战,形式概念分析(Formal Concept Analysis, FCA)对于处理复杂的信息不失为一种有效的方法。早在1940年Birkhoff就已为该方法提供较好的数学理论基础<sup>[1]</sup>;之后,Ganter等人将其作为一个较好的数据分析方法,深化、完善该理论基础,并将它们扩展到各种现实应用中<sup>[2]</sup>。形式概念分析提供了一种较好的层次化(形式)对象的分析方法,它能够识别那些具有共同(形式)属性的一组(形式)对象的组合<sup>[2]</sup>。一方面,形式概念分析已具备较完善的理论基础,在应用到其他领域中时,具有形式化方面的表达能力,从而具有较强的理论和技术说服

力;另一方面,形式对象和形式属性这种二元关系经常出现在软件世界中,对这种二元关系处理的方法也推动形式概念分析技术在其他领域应用的不断发展。由此,形式概念分析开始从应用角度走向理论形式化研究,而理论研究将进一步促进形式概念分析在各个领域中的应用。

### 2. 形式概念分析概念

#### 2.1 基本概念

形式概念分析中的概念来源于哲学。在哲学中,概念被理解为由外延和内涵两部分组成的思想单元。基于概念的这一哲学理解,Wille教授对概念进行了形式化描述,提出的形式概念分析可以用于概念的发现、排序和显示。在形式概念分析中,概念的外延被理解为属于这个概念的所有对象的集合,而内涵则被认为是所有这些对象所共有的特征(或属性)集。所有的概念连同他们之间的泛化与例化关系可以构成一个概念格。概念格结构模型是形式概念分析理论中的核心数据结构。其本质上描

述了对象和特征之间的联系，表明了概念之间的泛化与例化关系，其相应的 Hasse 图则实现了对数据的可视化。目前形式概念分析已被广泛地研究，并应用到机器学习、软件工程和信息获取等领域。

概念格是 FCA 的核心数据结构。概念格的每个节点是一个概念，由外延和内涵组成。外延是概念所覆盖的实例；而内涵是概念的描述，是该概念所覆盖实例的共同特征。另外，概念格可以通过其 Hasse 图生动简洁地体现概念之间的泛化和例化关系。

在形式概念分析中，数据是用形式背景来表示的，下面我们给出它的形式化定义<sup>[3]</sup>。

定义 1：一个形式背景  $K$  是一个三元组，设  $K = (G, M, I)$ ，其中  $G$  为所有对象的集合， $M$  为所有属性的集合， $I \subseteq G \times M$  为和中元素之间的关系集合。对于  $g \in G$ ， $m \in M$ ， $(g, m) \in I$  表示“对象  $g$  具有属性  $m$ ”，也写作  $gIm$ 。

FCA 形式背景是一个交叉表（见表 1），表格的行代表对象，列代表属性，行列交差处可以为“0”也可以为“1”，“1”表示相应的对象和属性之间具有关系，“0”表示相应的对象和属性之间没有关系。如表 1 就是一个典型的形式背景，包括 4 个对象和 4 个属性。

表 1 一个典型的形式背景

|       | Emale | Juvenil | Adult | Male |
|-------|-------|---------|-------|------|
|       | e     |         |       |      |
| Girl  | 1     | 1       | 0     | 0    |
| Woman | 1     | 0       | 1     | 0    |
| Boy   | 0     | 1       | 0     | 1    |
| Man   | 0     | 0       | 1     | 1    |

定义 2：设  $K = (G, M, I)$  为一形式背景。对于集合  $A \in G$ ，记：

$$A' = \{m \in M \mid (g, m)I, \forall g \in A\},$$

相应地，对于集合  $B \in M$ ，记：

$$B' = \{g \in G \mid (g, m)I, \forall m \in B\}$$

由定义可知，集合  $A'$  表示的是集合  $A$  中所有对象都具有的属性的集合；而集合  $B'$  则表示具有属性集合中属性的所有对象的集合。如表 1 中对于对象集合  $A = \{girl, woman\}$  来说，它们共同具有的属性集合为  $A' = \{female\}$ ；同理对于属性集合  $B = \{female, adult\}$  而言，具有  $B$  中属性的对象集合为  $B' = \{woman\}$ 。

为方便起见，我们用  $g'$  表示  $\{g\}$ ，用  $m'$  表示  $\{m\}$ 。

定义 3：设  $K = (G, M, I)$  为一形式背景， $A \in G$ ， $B \in M$ ，如果  $A' = B'$ ，则称  $C = (A, B)$  为  $K$  的一个概念。此时，称  $A$  为  $C$  的外延， $B$  为  $C$  的内涵。我们用  $\beta(K)$  记  $K$  的所有概念组成的集合。

如表 1 中令  $A = \{girl, woman\}$ ， $B = \{female\}$ ，那么  $A' = \{female\} = B'$ ， $B' = \{girl, woman\} = A$ ， $(\{girl, woman\}, \{female\})$  就称为一个概念。

定义 4：设  $K = (G, M, I)$  为一形式背景， $C_1 = (A_1, B_1)$ ， $C_2 = (A_2, B_2)$  是两个概念，规定：

$$C_1 \leq C_2 \Leftrightarrow A_1 \subseteq A_2 \Leftrightarrow \{B_1 \supseteq B_2\}$$

此时，称  $C_2$  为  $C_1$  的超概念，而称  $C_1$  为  $C_2$  的子概念。

显然，关系“ $\leq$ ”是集合  $\beta(K)$  上的一个偏序，它可诱导出  $\beta(K)$  上的一个格结构，可以证明，它

是一个 $\in$ 完备格<sup>[2]</sup>，相应的下确界和上确界定义为：

$$\wedge(A_i, B_i) = (\cap A_i, \cup B_i)$$

$$\vee(A_i, B_i) = ((\cup A_i), \cap B_i)$$

其中 $(A_i, B_i) \in \beta(K)$ ， $T$ 是指标集，此完备格称为形式背景 $K$ 的概念格，在没有歧义的情况下，仍然记为 $\beta(K)$ 。

形式概念分析具有以下性质：

命题 1：设 $K = (G, M, I)$ 为一形式背景，

$A, A_1, A_2 \in G$ 为对象子集， $B, B_1, B_2 \in M$ 为属性子集，下列结论成立：

$$A_1 \subseteq A_2 \Rightarrow A_2' \subseteq A_1' \quad (1)$$

$$B_1 \subseteq B_2 \Rightarrow B_2' \subseteq B_1' \quad (2)$$

$$A \subseteq A'', B \subseteq B'' \quad (3)$$

$$A' \subseteq A''', B \subseteq B''' \quad (4)$$

证明：一是如果 $m \in A_2'$ ，那么对于所有的对象 $g \in A_2$ 有 $gIm$ 成立；而 $A_1 \subseteq A_2$ ，所以关系 $gIm$ 对于任何 $g \in A_1$ 同样成立，那么必有 $g \in A_1'$ 成立，即 $A_2' \subseteq A_1'$ 。二是同理可证。三是如果 $g \in A'$ ，那么对于 $m \in A'$ 来说 $gIm$ 成立，反之，若 $m \in A'$ ，具有 $gIm$ 关系的对象属于的集合 $g \in A''$ ，所以 $A \subseteq A''$ 。四是由(2)可得 $A' \subseteq A'''$ ，而若 $A' \subseteq A''$ 由(1)又可得 $A''' \subseteq A'$ ，二者相互包含，所以 $A' = A'''$ 成立。

## 2.2 建格

建格的过程是一个不断寻找概念的外延和内涵

的过程<sup>[4]</sup>。为达到这一目的，用一个列表来保存概念的外延，首先进入这个列表的是对象全集 $G$ ，然后对属性集合的每一个属性依次执行以下相同步骤：设前一次进入列表对象集合记为 $A$ 。

步骤 1：对每个属性 $m \in M$ ，求出具有该属性的对象集合，记为 $m'$ 。

步骤 2：对 $m'$ 与先前进入列表的每一个对象集合 $A$ 求交集，记为 $A \cap m'$ 。

步骤 3：判断 $A \cap m'$ 是否已经出现在列表之中，若出现则忽略，若没有则加入。

依次执行，直到所有属性 $m \in M$ 都执行一遍，最终得到的列表 $\Phi$ 即所求的概念的所有外延的集合；再对每个外延求其内涵，则得到形式背景的所有概念。

用些方法对表 1 来求概念，一共得到以下 10 个概念：一是 $(\{girl, woman, boy, man\}, \Phi)$ ；二是 $(\{girl, woman\}, \{female\})$ ；三是 $(\{girl, boy\}, \{juvenile\})$ ；四是 $(\{girl\}, \{female, juvenile\})$ ；五是 $(\{woman, man\}, \{adult\})$ ；六是 $(\{woman\}, \{female, adult\})$ ；七是 $(\Phi, \{female, juvenile, adult, male\})$ ；八是 $(\{boy, man\}, \{male\})$ ；九是 $(\{boy\}, \{juvenile, male\})$ ；十是 $(\{man\}, \{adult, male\})$ 。

根据概念外延间的包含关系，可以得到形式背景的 Hasse 图（见图 1），同样由图也可读出概念之间的泛化和例化关系。

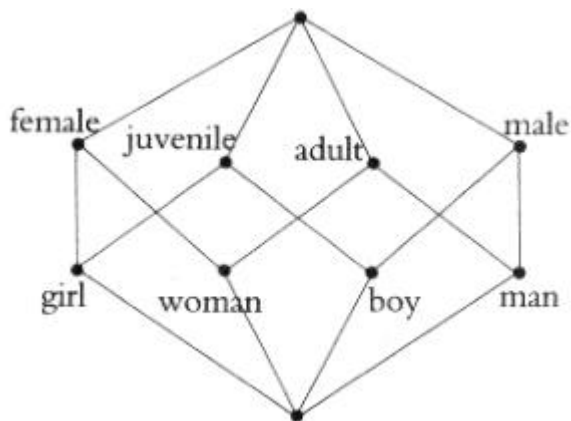


图1 建格举例

在形式背景中，此Hasse图也称为线图。图1中的一个节点代表一个概念。

### 3. 形式概念分析在各个领域的相关研究

#### 究

#### 3.1 形式概念分析与概念格在本体相关领域的应用研究

近年来，本体（Ontology）作为领域内共享概念模型的明确的形式化规范说明，凭借其在知识共享和知识重用方面的优势，在人工智能和知识工程领域的相关研究中越来越被人们所重视。传统的本体构建多采用手工方法，尽管研究人员一直致力于探寻自动或半自动构建本体的方法，但效果并不理想，直到Wille以形式概念分析重构概念格理论使得这一状况得到改变。

在本体构建相关研究中，人们普遍认为，以自然语言为基础的缺少形式化的本体尽管易于开发，但形式化的本体能够以“自动”的方式更好地被重用和共享。Obitko等认为，定义一个好的本体不仅需要语法结构，更需要语义描述。由于形式语义是实现基于本体的自动推理的重要环节，因此使得具有良好语义特征的本体更容易在不同领域中被共享和重用。在此基础上，Obitko等提出，利用形式概念分析能够通过构建概念格探寻潜在的对象和属性，并将现有的和潜在的实体以可视化的方式自动呈现。从而使得基于形式概念分析构建的本体在知识重用和知识共享等应用层面上比单纯的分类法具有更大的优势。Formica以本体重用为目的，将形式概念分析和概念格应用于对现有本体的分析改造。基于给定的形式背景，通过形式概念分析构建概念格，对同一背景下的概念进行比较，以及对不

同背景的概念格中的概念进行相似性评估，提出了概念的相似性推理，从而实现对现有本体的重用。Bendaoud等则在形式概念分析基础上扩展出关联概念分析（Relational Concept Analysis, RCA），将其应用于设计给定领域的真实世界本体，并以OWL编码，实现基于分类的推理。

在本体映射与本体合并的相关研究中，Choi等通过对本体映射的工具、系统、框架以及相关工作的考察，对多种不同类别的本体映射的特点和应用领域进行讨论。Kalfoglo等也对近年来本体映射与本体合并相关研究进行总结。这些文献都指向Stumme等提出的FCA-Merge[9]。FCA-Merge是一种基于形式概念分析和自然语言处理（NLP）的本体合并方法，它包括：抽取和产生待合并本体的形式背景；利用FCA-Merge核心算法计算生成概念格；在概念格基础上半自动地产生最终合并后的本体。FCA-Merge作为一种自底向上的本体合并方法，成为这一时期的经典。其后，Zhao等运用形式概念分析在形式化背景上生成概念格后，引入粗糙集理论对本体中节点的相似性进行测度，提出了一种本体快速映射的方法。

形式概念分析与概念格理论在本体构建、映射以及合并中的应用，不仅使本体构建由手工迈向自动化，而且还使新生成（包括构建、改造、映射、合并等）的本体更具有形式化的特征，提高了本体在跨异构系统和语义Web操作方面的灵活性。

#### 3.2 形式概念分析与概念格在软件工程领域的应用研究

最近10年来，软件工程成为国外形式概念分析与概念格应用研究中一个新的热点方向。相比其他方面的应用而言，软件工程更贴近于人们的社会经济生活，因此国际上关于形式概念分析与概念格在这一领域中应用的相关研究发展迅速而突出。Tilley等基于ISO12207软件工程标准对形式概念分析和概念格在软件工程领域中的应用进行分析，重点梳理形式概念分析和概念格在软件工程前期和后期两个阶段的应用情况。形式概念分析和概念格在软件编码之前的应用主要包括：需求分析、组件重用、形式规范等；形式概念分析和概念格在软件编码之后的应用包括：需求矫正、环境适应、完善功能等。

其中具有代表性的应用研究集中在：代码特征定位、类层次再造、模块结构调整、基于切片的影响分析等。Eisenbarth等将形式概念分析用于系统

需求识别的计算单元特征定位。通过形式概念分析,将由用户及用户行为触发的代码特征映射集合进行重构,然后在给定特征集中识别出全局和局部计算单元,通过动态与静态分析结合的方法迅速聚焦于相关的特定特征集,其效果优于早期的特征与场景的一对一的对应关系。Godin 等认为形式概念分析为设计和维护面向对象软件的类层次提供了天然的理论框架,他们基于形式概念分析的类层次的设计与维护保持类的特化关系,在类层次结构的生命周期开发中十分有效,实现了基于现有对象代码的程序重构。Al-Ekram 等整合概念分配、形式概念分析和程序切片三种技术,提出“概念切片格(Lattice of Concept Slices)”的思想对已有软件进行维护和重用。Al-Ekram 等将基于形式概念分析的概念格作为基础数据结构,以此来展示程序声明与概念域之间的关系,进而提高程序的模块化程度。Tonella 则把分解切片和概念格用于程序代码的细节呈现,提出“分解切片的概念格(Concept Lattice of Decomposition Slices)”的思想,用概念分析方法提供的计算模块的“依赖/冲突”关系,对自然数据结构进行影响分析处理。此外,形式概念分析还被应用于软件工程不同阶段的需求连续性问题,研究表明,用形式概念分析处理软件工程不同阶段的需求连续性问题,可以取得对软件需求连续性的系统化的精确验证。

事实上,国外关于形式概念分析和概念格理论在软件工程领域中的应用远不止于上述范围。在软件工程领域,形式概念分析与概念格还被应用于组件重构、故障跟踪、设计帮助等诸多方面。这些应用使得在软件开发过程中,客观成分逐渐加大,形式化的特征逐步增强,标准化趋势更加明显。另一方面,这些应用使得软件工程对个体认知的依赖程度逐步减弱,受研发人员个体素质的影响逐渐消失,进而大大提高了软件设计、开发、维护、重用的效率和效益。

### 3.3 形式概念分析与概念格在知识发现领域的应用研究

在国际学术界中,将形式概念分析和概念格理论应用于知识发现研究领域的杰出代表人物是德国学者 Gred Stumme。在最近 10 余年间,这一领域的核心成果与文献几乎都出自 Stumme 及其合作者。鉴于其在该领域研究的代表性和典型性,这一部分将以 Stumme 及其合作者的相关研究为主线进行阐述。

1998 年,德国达姆施塔特科技大学的 Stumme

等提出“数据库概念化知识发现(Conceptual Knowledge Discovery in Database, CKDD)”的思想,并将其建立在基于形式概念分析的概念化知识处理(Conceptual Knowledge Processing)的基础上。Stumme 等利用形式概念分析工具软件托斯卡纳系统(TOSCANA)对大型数据库中概念化知识进行探寻和挖掘,并验证其作为以人为中心的数据库知识发现手段是有效的。1999 年,Stumme 把概念格用于概念化信息系统(Conceptual Information System)的知识呈现,并以此构造关联规则的挖掘结果,简化概念化信息系统可视化的复杂程度。2000 年,Stumme 及其合作者将形式概念分析与概念格理论在知识发现领域中的应用进一步深化。他们利用形式概念分析展示“数据库概念化知识发现”中数据挖掘技术与传统数据分析技术的相互影响和相互整合,并基于 TOSCANA 阐释知识发现过程中由数据到知识的转变过程。同年,Stumme 等在基于概念格的关联规则挖掘中提出著名的 TITANIC 算法。研究表明,该算法同样可应用于概念聚类和本体工程,尤其对于弱相关的数据挖掘是有效的。而且,形式概念分析可以被用作蕴含规则和关联规则发现的形式框架,并且能够提高规则挖掘的响应效率。与此同时,Stumme 及其合作者针对关联规则挖掘中产生的大量冗余规则开始尝试利用形式概念分析提取闭合频繁项集,以此挖掘非冗余的关联规则。2001 年到卡尔斯鲁厄大学后,他继续致力于基于形式概念分析的冗余关联规则的缩减研究。2002 年,Stumme 在 TITANIC 算法的基础上提出用“冰山概念格(Iceberg Concept Lattices)”分析大型及特大型数据库。冰山概念格能够在不损失信息的情况下呈现关联规则的频繁模式,同时实现作为挖掘结果的规则的可视化。该方法很快被应用于数据分析、信息检索、知识发现等领域。此项研究在 Stumme 到卡塞尔大学后,仍然得以持续,并于 2005 年陆续发表了一些研究成果。2006 年开始,Stumme 及其合作者又将基于形式概念分析的关联规则挖掘应用于 Folksonmy(分众分类法)的分析和构建,并将形式概念分析的应用领域延伸到 Folksonmy 的语义识别、垃圾识别、系统评估等方面,开辟了形式概念分析应用的新领域。纵观 Stumme 及其合作者们的相关研究,也在一定程度上印证了知识发现领域中一个公认的事实,即关联规则挖掘是知识发现领域中最为重要的内容。

在这期间,其他研究者也纷纷加入形式概念分

析与概念格理论在知识发现领域的应用研究,这使得知识发现领域几乎成为形式概念分析与概念格理论应用中最为庞大的分支。但值得一提的是,该领域中相当一批研究成果建立在 Stumme 及其合作者的工作基础之上,或对其完善、或对其改进。对 Stumme 及其合作者在该领域的研究成果进行分析和梳理,几乎可以达到理清该领域发展全貌的效果。

### 3.4 形式概念分析与概念格在 Web 语义检索领域的应用研究

#### (1) Web 语义检索针对的领域

Kim 等从特定领域的信息检索出发,构建了一个以形式概念分析构造的概念格为基础的增量知识获取浏览机制。该浏览机制允许用户针对特定的领域动态升级文档的组织结构,随着时间的推移,多个用户可以相互协作建立和维护浏览方案,弹性地支持开放性的文档管理。而 Cigarrán 等则将基于形式概念分析与概念格的 Web 语义检索由特定领域扩展到自由文本,他们开发的检索系统不限于特定的领域或词库,其基于自动术语抽取的权重公式使得词库的构建无须手动分配描述符。Cigarrán 等的研究成果提供了多项措施用于评估概念格的质量,并将这些措施用于比较自动选择属性的若干策略,从而在返回的文档列表中动态决定最适当的属性。

#### (2) Web 语义检索的层级结构

Rome 等把形式概念分析应用于探索 Web 页面的链接结构和 Web 群落结构,并利用这种结构对基于概念格的知识库进行定位和组织,同时根据 Web 群落的相似性来降低知识库的复杂程度,进一步验证形式概念分析对 Web 语义检索的有效性。Messai 等从形式概念分析中形式背景的属性依赖关系着手,利用这些依赖关系定义了一个属性的层级结构,用于反映属性的重要性或兴趣度。属性的层级结构是一个基于属性重要程度的半序集合,能够生成由领域知识驾驭的导航信息。在基于概念格的查询中,属性的层级结构可用于定义包含不同重要程度的属性的复杂查询,重要的属性定义检索的焦点,次要的属性则反映次要的信息。复杂查询中属性间的关系隐含或明示了计算检索结果时需考虑的知识单元。

#### (3) Web 语义检索的困难处理

Carpineto 等针对传统检索系统无法轻易解决的信息检索任务,开发了一个允许用户查询 Web 文档并将检索结果以概念格方式组织后再呈现给用户的检索系统。该系统对于在模糊的查询结果中快速

定位文档以及对所涉及的文档内容的挖掘十分有效。针对实际应用中概念格的复杂程度可能随数据库规模而迅速增长的问题,Cheung 等在将概念格和向量空间模型进行比对的基础上,提出以“词文格(Term-Document Lattice)”作为信息检索的模型,并基于商格(Quotient Lattice)的思想对原始词文矩阵使用单值分解(Singular Value Decomposition)以降低词文格的复杂程度。Yadav 则提出了用户中心的高品质信息检索(UCQIR)的概念化模型,并以此设计了特殊领域高品质信息检索系统。

## 4 研究趋势与展望

形式概念分析在软件维护领域的应用在近十几年来取得长足的发展,尽管如此,我们认为如下一些方面仍值得国内外学者关注:

(1) 应用领域的深入。形式概念分析技术在程序理解和重构方面已取得大量的研究成果,而形式概念分析在修改影响分析、调试以及测试方面的研究工作目前还处于起步阶段,甚至在修改后的系统验证方面还没有找到相关研究文献,因而形式概念分析技术在这些方面的研究还有待深入。

(2) 应用层次之间的可跟踪化。虽然有部分研究人员将形式概念分析应用于需求层次或者设计层次,但各个层次使用形式概念分析技术所提取出来的概念是不同的,但这些概念之间从纵向的角度看是存在一定的可跟踪性,但目前还没有相关的从需求层次到设计层次以及规约层次之间的概念跟踪性研究。

(3) 应用工具的支撑。尽管形式概念分析技术已有一些工具支撑,但这些工具主要还是用于概念格的构造,当这些工具应用到某个具体的领域中还有一些不适应,需进一步修改和扩展,而这方面的工具目前还相当缺乏,所以开发适合某个具体领域形式概念分析的工具很有必要。

(4) 形式概念分析支持软件维护的统一化。形式概念分析可用于各种软件维护活动中,形式概念分析具有较完善的理论基础,研究如何利用该技术统一软件维护的各种活动:从修改(软件)理解到修改实施,修改后测试以及最后的系统验证,这样的研究将会很有意义。

## 5 结束语

事实上,形式概念分析与概念格理论在以上 4

个方面的应用并非是绝对孤立的，而是相互交叉、相互融通的。本体的构建、映射与合并，在本质上正是借助于形式概念分析与概念格理论在概念化知识呈现与处理方面的优势，而这一优势同时又为其在知识发现领域中的应用奠定了基础，软件工程中的代码定位、组件重构等应用一定程度上体现的就是面向软件开发与维护领域的知识重用，而 Web 语义检索则是利用形式概念分析与概念格在形式背景中定位概念、揭示概念间关系的优势，对语义 Web 层级结构的重新组织。

从最近 5 年（2006-2010）形式概念分析国际会（International Conference on Formal Concept Analysis, ICFA）的主题来看，软件工程与知识发现是最近 5 届都涉及到的主题内容。在第 4 届到第 6 届（2006-2008）会议中，与本体相关的研究内容主要集中在概念化知识呈现上，而第 7 届和第 8 届（2009、2010）会议则上升为概念化知识处理这一主题。在第 8 届（2010）会议中，首次出现了与网络和 Web 相关的主题。2010 年 10 月，第 9 届（ICFA2011）形式概念分析国际会议的信息已经发布。如果说从未来一届形式概念分析国际会议的主题可以窥得未来应用研究发展趋势的端倪的话，可以看到：在未来的一个时期内，软件工程和知识发现仍将是会议的主要议题和研究热点；本体相关的研究内容不再区分概念化知识呈现或概念化知识处理，而是笼统地以概念化知识作为主题，而概念化知识显然比本体或单纯强调呈现和处理具有更为广泛的外延，其研究范围也将会更广阔；网络和 Web 的相关主题中特别附加了语义 Web 这一主题内容，这明确表明，对于语义 Web 的分析和加工处理将成为未来研究中更为重要的热点方向之一。

随着对形式概念分析理论应用研究的不断深入，其发展前沿和研究热点也不会一成不变，本文所列的 4 个方面只是形式概念分析与概念格理论应用中诸多分支中的一部分，并不足以囊括国际上这一领域研究成果的全貌。在今后的研究中，仍然需要不断跟踪和把握国际学术界的发展前沿和研究热点。

形式概念分析的基本概念作为一种新型的数据挖掘工具，可应用于机器学习、软件工程和信息获取等领域。形式概念分析技术本身经过几十年的发展和完善，已成为具有较强形式化理论支撑的软件分析技术，该技术近年来被广泛应用到软件维护各种活动中。尽管形式概念分析在软件维护中有着广

泛的成功应用，但其在应用领域深化、工具支撑等方面还需进行深入的研究。

#### 参考文献

- [1]Birkhoff G. Lattice Theory [M]. USA: American Mathematical Society, 1940.
- [2]Ganter B, Wille R. Formal Concept Analysis: Mathematical Foundations [M]. Berlin: Springer-Verlag, 1996.
- [3]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): 32-33.
- [4]Wille R.Restructuring lattice theory: an approach based on hierarchies of concepts[M].In: Rival Id Ordered Sets Dordrecht: Reidel, 1982: 445-470.