MIANXIANG DUIXIANG
GAINIANGE

面向对象概念格

陈永平◎著

合肥工业大学出版社
HEFEI UNIVERSITY OF TECHNOLOGY PRESS

MIANXIANG DUIXIANG

面向对象概念格

陈永平◎著

合肥工业大学出版社

内容提要

本书全面介绍了面向对象概念格的属性约简、压缩、合成背景的面向对象概念格的生成和属性约简技术以及决策背景下面向对象概念格的属性约简的方法，配以算法设计或求解步骤，并对各种方法进行性能分析；同时，通过实例来验证本书所提出的各方法的有效性。本书可作为数据分析师、软件工程师、项目规划师等工程技术人员的技术参考用书，也可以作为高等院校计算机相关专业的教学用书或参考用书。

图书在版编目(CIP)数据

面向对象概念格/陈永平著．—合肥：合肥工业大学出版社，2016．4
ISBN 978-7-5650-2669-0

Ⅰ．①面…　Ⅱ．①陈…　Ⅲ．①人工智能—应用　Ⅳ．①TP18

中国版本图书馆 CIP 数据核字(2016)第 034446 号

面向对象概念格

陈永平　著

责任编辑	张择瑞
出版发行	合肥工业大学出版社
地　　址	(230009)合肥市屯溪路 193 号
网　　址	www. hfutpress. com. cn
电　　话	理工教材编辑部：0551－62903204 市 场 营 销 部：0551－62903198
开　　本	710 毫米×1010 毫米　1/16
印　　张	10
字　　数	151 千字
版　　次	2016 年 4 月第 1 版
印　　次	2016 年 4 月第 1 次印刷
印　　刷	安徽联众印刷有限公司
书　　号	ISBN 978-7-5650-2669-0
定　　价	20.00 元

前　　言

形式概念分析和粗糙集理论从不同侧面来研究和表现数据中隐含的知识。概念格结构模型是形式概念分析的核心数据结构，形式背景是概念格理论的数据表现方式，概念格是其研究基础，概念格本质上描述对象与属性之间的联系，表明概念之间的泛化、特化关系，其相应的 Hasse 图实现了数据的可视化。而粗糙集理论的数据表现方式是信息系统，对象间的等价关系是其研究基础。虽然形式概念分析和粗糙理论是两种不同的理论，但二者有许多相似之处，而且有许多学者将二者结合起来研究，并且也取得了很好的成果。其中面向对象概念格就是将概念格和粗糙集理论结合起来研究的产物之一。本书主要针对形式背景中面向对象概念格的属性约简、面向对象概念格的压缩、合成背景的面向对象概念格的生成和属性约简方法和决策背景下的面向对象概念格的属性约简算法进行研究，从而更好地帮助读者了解面向对象概念格的属性约简、压缩、合成背景的概念格的生成和约简，还有决策背景的约简的方法和算法及算法的性能分析，进一步扩大了概念格的理论。

本书共分五章，内容安排如下：

第 1 章是基础理论部分。主要介绍了关系和格、经典概念格、粗糙集和面向对象概念格的定义和相关基本性质和定理，同时也通过面向对象概念的定义及相关知识，介绍了经典概念格、粗糙集和面向对象概念格之间的关系。

第 2 章是介绍面向对象概念格的属性约简方法。首先介绍面向对象概念格的协调集的定义及判定定理，面向对象概念格的约简的定义及属性特征；然后重点介绍了四种面向对象概念格的属性约简方法，其中基于属性特

征和基于可辨识属性矩阵的面向对象概念格的属性约简方法和经典概念格的属性约简方法基本一样；而基于并不可约元和基于并可约元的面向对象概念格属性约简方法和经典概念格约简方法有很大区别；书中同时也对各种属性约简方法进行分析比较，并通过例子帮助读者理解面向对象概念格的属性约简方法。

第 3 章是面向对象概念格的压缩。总的来说可以分为两个部分：第一部分是基于形式背景中的对象、属性和二元关系对的相似度，来实现对面向对象概念格的压缩和简化；第二部分是基于面向对象概念格中的外延、内涵和概念的相似度实现对面向对象概念格的压缩。从时间效率分析，前一部分要优于后一部分，但后一部分可以有目的地删除一些概念节点。

第 4 章是合成背景的面向对象概念格的生成和属性约简。本章可以分为三个部分：第一部分分别研究基于外延并运算和基于内涵交运算的横向和纵向合成背景的面向对象概念格的生成。第二部分介绍了通过属性扩展或对象扩展，使待双向合成的形式背景变成和横向合成或纵向合成相一致的形式背景，然后利用第一部分方法来实现双向合成背景的面向对象概念格的生成；此外在第二部分也研究了根据面向对象概念格的特征，直接通过待双向合成的各个子面向对象概念格的外延和内涵分别进行并运算得到合成背景的面向对象概念格的外延和内涵，从而进一步得到合成背景的面向对象概念格；第三部分介绍合成背景的面向对象概念格的属性约简，通过定义面向对象概念格的属性近似算子真子集，并给出了各子形式背景和合成形式背景的属性特征间的关系，在此基础上，给出了合成形式背景的属性近似算子真子集、子形式背景的属性近似算子真子集和属性等价类间的关系，并利用它们间的关系给出了合成背景的面向对象概念格的约简集。

第 5 章是决策形式背景的面向对象概念格的属性约简。重点介绍了基于并不可约元的决策形式背景属性约简，通过求出决策概念格的外延集在条件属性集中执行相关运算，并利用这些运算结果的并运算得到决策形式背景的面向对象概念格的属性约简集。

本书力求使读者全面了解面向对象概念格的相关理论知识，并可针对自己面临的面向对象概念格的问题找到相关的可行的解决方案。

由于时间仓促，书中存在的错误和疏漏，恳请专家和读者不吝赐教。

目　　录

第1章 绪 论

1.1 引 言

人类在认识世界过程中，将所感知到的事物的共同特点抽取出来，加以总结、概括和抽象，称之为概念。在哲学上说，概念是被理解为由外延与内涵所组成的思想单元，内涵是对一切外延的概括，外延是内涵表达的具体化。概念作为人的思想和知识的基本单元，一直以来，深受哲学界和科学界的关注和重视，很自然地，也就成为了人工智能学科的重要研究对象，这主要体现在知识表示（语义网络、概念图、描述逻辑）和机器学习（概念聚类、概念学习、认知计算）等领域。再有一个新兴的领域，那就是形式概念分析，形式概念分析是研究知识表示的领域，它的核心数据结构 —— 概念格是对概念以及概念之间关系的描述，在一定程度上是对客观世界的一种高度简化的描述形式。这种简化的最大优点是具有良好的数学性质。

基于对概念的哲学思想，德国数学家 Wille. R 于 1982 年提出来形式概念分析理论[1]，是一种表达和处理概念与概念层次结构的应用数学理论，作为形式概念分析的两个基本概念 —— 形式背景和形式概念。其中，形式背景是包含对象集、属性集以及对象集与属性集之间的一个二元关系构成的一个三元组；形式概念通过一对伽罗瓦算子进行定义，用外延和内涵构成的

序对来刻画。一个形式背景的所有形式概念连同它们之间的泛化和特化关系共同构成一个完备格;在相关的著作中第一次提到关于这种按集合的包含关系排序的集合,对组成的结构满足完备格的性质,起初被定义为 Galois 格并引起了广泛的研究[2,3],关于形式概念分析的术语是由 Wille 首先提出的。对形式概念研究主要对象是概念与概念格;概念是由一对外延集与内涵集组成的统一体,外延表示这个概念所有对象的集合,内涵表示所有这些对象所共同拥有的属性集合。所有概念以及它们的泛化与例化层次关系构成了概念格,其相应的 Hasse 图实现了对数据的可视化。作为数据分析和知识发现的有力工具,概念格理论已经被广泛运用到机器学习、软件工程、知识管理、信息检索和规则提取等众多领域[4,5,6]。

1982 年,德国数学家 Wille. R 发表的《Restrueting lattice theorybase on hierarchies of concepts》标志着概念格理论的建立。《Formal concept analysis》使概念格获得进一步的发展,讨论它的合成与分解等若干性质。经过 30 多年的研究发展,概念格已经从最初的单纯数学位置转移到计算机科学方面。随后,又有学者研究概念格的同构性;近些年来,又有人提出邻域、闭包、随机概念格、粒概念格、决策概念格等概念。总的来说,概念格的研究主要集中在概念格的构造、概念格的约简和多概念格的合成和决策形式背景的比较上,概念格的构造过程实际是概念的聚类过程,在同一个形式背景下,所生成的格也是唯一的。目前,概念格的构造算法可以分为两类:批处理算法和增量算法。批处理算法的思想是先找出所有概念,再根据它们之间的直接前驱-后继关系生成边并确定其结构。增量算法是将当前要插入的对象与概念格中的所有概念作交运算,由交的结果把概念格中的节点分为不变节点、更新节点与新增节点,根据交的结果采取不同的方法。其中,概念格的维护算法也属于概念格增量算法,主要是对象的插入、修改、删除和属性的增加与删除。概念格的约简,就是在保持形式背景所有概念的外延集不变的基础下,删除不必要属性,寻找最小属性子集,以该属性子集为属性集的子形式背景的概念格能够完全确定形式背景上的原有概念,并保持它们之间的原有层次结构关系,这使得形式背景中隐含知识的发现变得更容易。多概念格的合成就是将分布在不同地理位置上的多个形式背景合并一个大的形式背景,也可以是将一个大的形式背景分布存储于不同

地理位置的多台计算机上，目前主要研究多概念格的合成时的概念格的生成和合成背景的属性约简。此外，也有学者直接从形式背景出发，通过属性间的相似度，提出了基于对象覆盖压缩、属性覆盖压缩和基于属性融合压缩的压缩方法，对概念格进行压缩，简化了概念格。目前关于决策形式背景的研究主要集中在属性约简上和决策形式背景的压缩和简化。通过对概念格相关理论知识的研究，使得概念格的应用取得良好的经济效益和社会效益。

由于我们日常生活中能够接触到的大多数信息都是带有不确定性因素的，而且在采集样本的过程中往往还带有一些其他的干扰，否则的话我们就能够很好的根据这些已给的信息做出准确的抉择。为了从已知的各种不完备的信息中发现并获取隐含在其中的知识和规律，人们开始探索这些不确定性的问题。在 20 世纪 70 年代，波兰数学家 Z. Pawlak 就着手这方面的研究工作，并在 1982 年首次提出了一个分析数据的数学理论，即粗糙集理论（Rough Set，Rough 集）。它主要是研究不完整性以及不确定性的问题，从已知的各种不完备的信息中发现并获取隐含在其中的知识和规律。粗糙集理论是数据挖掘中的一项重要的理论，主要用于发现不确定数据或者噪声数据之间的联系，它能够用于不同见解程度表示的数据要求以及在处理不确定信息时不需要添加额外的数据信息。粗糙集用于定义两个类，上近似集和下近似集，用于解决边界不清晰的模糊问题。粗糙集的主要应用在于它的两个近似集。在我们的生活中，并不是非黑即白的世界，所以，在聚类的时候，可能有的对象不一定属于某一个类，它可能同时属于多个类。在这种情况下，粗糙集就有了利用价值。粗糙集被广大学者们研究了很多年，粗糙集理论研究的主要问题包含信息系统或者决策表的属性约简问题、不完备决策表的处理情况、连续属性的离散化实现以及与其他方法相结合处理的问题。粗糙集理论主要是在保持原有数据的划分能力和决策能力不发生改变的前提下，消除冗余的信息，从而提取出有价值的知识。其获得的信息就会更加具有客观性，从而被广泛地应用于数据挖掘、机器学习、专家系统等相关领域，并获得了一定的成功[7,8,9]。

概念格理论和粗糙理论是两种不同的数据分析方法，它们从不同的角度研究数据集合中所隐含的知识，都已被成功地应用于许多领域。概念格

理论的数据表现方式是形式背景，其研究的基础是概念格；粗糙集理论的数据表现方式是信息系统，其研究基础是等价关系。概念格理论和粗糙理论虽然是两种不同的理论，但二者有许多相似之处，而且有许多学者将二者结合起来研究，并且也取得了很好的成果。Y. Y. Yao 利用粗糙集近似算子定义了两个外延和内涵平衡的条件，并提出了面向对象概念格[10,11]，D. untsh 和 Gediga 构造了另外一种新的概念格 —— 面向属性概念格[12,13]，从而得到了两种新的概念格：面向对象概念格和面向属性概念格，丰富了概念格理论，为数据分析提供了新的理论依据。

1.2 关系和格[14,15,16]

序偶的定义 设 a,b 是一对被规定了有次序之分的事物，由 a,b 组成一个新的元素，用 $<a,b>\in R$ 表示，称 $<a,b>$ 为序偶，其中 a 称为序偶 $<a,b>$ 的第一个成员，b 称为 $<a,b>$ 的第二个成员。如：＜上，下＞表示一对序偶。序偶是一个整体元素；序偶中的成员之间是有序的；一般，如果 $a\neq b$，那么 $<a,b>\neq<b,a>$。

有序 m 元组的定义 设 $m(m\geqslant 3)$ 是给定的自然数，一个序偶 a，如果它的第一个成员为 $m-1$ 元组，第二个成员为单独元素，那么 a 被称为有序 m 元组，记作 $<a_1,a_2,\cdots,a_{m-1},a_m>$。

关系是一个基本概念，在我们日常生活中，关系无处不在；如兄弟关系、同学关系、师生关系、同事关系、同乡关系等。在数学上，可用关系表示集合中的元素间的联系，如 9 大于 8，点 x 在 y 和 z 之间等。又因为序偶是表达两个元素、三个元素，或 m 个元素之间的联系，因此用序偶表达关系这个概念是非常适合的。

定义 1.2.1 任一序偶的集合确定了一个二元关系 R，R 中任一序偶 $<a,b>$ 可记作 $<a,b>\in R$，或 aRb；不在 R 中的任一序偶 $<a,b>$ 可记作 $<a,b>\notin R$。

如：整数集中关系“$<$”，可记作 $<=\{(a,b)\mid a,b$ 是整数且 $a<b\}$。

由于关系是序偶的集合，如果序偶的第一个成员和第二个成员分别来自于不同的集合，那么关系就是两个集合直积的子集。

定义 1.2.2 令 A 和 B 是任意两个集合，直积 $A\times B$ 的子集 R，称作 A 到 B 的关系，$R\subseteq A\times B$。从关系的定义中可以了解到，在集合上的二元关系 R 的一些特殊性质。

定义 1.2.3 设 R 为定义在集合 A 上的二元关系，如果对于每个 $a\in A$，有 $<a,a>\in R$，则称二元关系 R 是自反的，即 R 在 A 上自反的 $\Leftrightarrow$ 任意 $a(a\in A\Rightarrow <a,a>\in R)$。

如：在实数集合中"$\geqslant$"是自反的，平面上三角形的全等关系也是自反的。

定义 1.2.4 设 R 为定义在集合 A 上的二元关系，如果对于每个 $a,b\in A$，当 $<a,b>\in R$ 就有 $<b,a>\in R$，则称集合 A 上的关系 R 是对称的，即 R 在 A 上是对称的 $\Leftrightarrow$ 任意 a，任意 $b(a\in A\wedge b\in A\wedge <a,b>\in R)\Rightarrow <b,a>\in R)$。

如：平面上三角形集合中的相似关系就是对称关系。

定义 1.2.5 设 R 为定义在集合 A 上的二元关系，如果对于任意 $a,b,c\in A$，当 $<a,b>\in R$，$<b,c>\in R$ 时，就有 $<a,c>\in R$，称关系 R 在 A 上是传递的。R 在 A 上传递，则对于任意 a,b,c $(a\in A\wedge b\in A\wedge c\in A\wedge <a,b>\in R,\wedge <b,c>\in R)\Rightarrow <a,c>\in R)$。

如：实数集合中的关系"$\leqslant$"就是传递关系。

定义 1.2.6 设 R 为定义在集合 A 上的二元关系，如果对于每一个 $a\in A$，都有 $<a,a>\notin R$，则称关系 R 是反自反的，即：R 在 A 上是反自反的 $\Leftrightarrow$ 任意 $a(a\in A\Rightarrow <a,a>\notin R)$。

如：日常生活中的母女关系就是反自反的。

注意：一个关系不是自反的，也不一定就是反自反的。

例如：$A=\{1,2,3\}$，$R=\{<1,1>,<1,2>,<3,2>,<2,3>,<3,3>\}$。因为 $2\in A$，而 $<2,2>\notin A$，所以 R 不是自反的也不是反自反的。

定义 1.2.7 设 R 为定义在集合 A 上的二元关系，如果对于每个 $a,b\in A$，当 $<a,b>\in R$ 和 $<b,a>\in R$，必有 $a=b$，则称 R 在 A 上是反对称的，即对于任意 $a,b(a\in A\wedge b\in A\wedge <a,b>\in R\wedge <b,a>\in R)\Rightarrow a=b$。

如：集合的“$\subseteq$”关系就是反对称的。有时可能有的关系既是对称，也是反对称的。

如：$A=\{1,2,3\}$，$R=\{<1,1>,<2,2>,<3,3>\}$，则关系 R 既是对称的，也是反对称的。但也有的关系既不是对称的，也不是反对称。如：$A=\{a,b,c\}$，$R=\{<a,b>,<a,c>,<c,a>\}$，则关系 R 既不是对称的，也不是反对称的。

下面介绍具有意义的等价关系和偏序关系。

定义 1.2.8 设 R 为定义在集合 A 上的二元关系，如果关系 R 是自反的、对称的和传递的，则称 R 为等价关系。

如：平面上三角形集合中，三角形的全等关系是等价关系。

定义 1.2.9 设 R 为定义在集合 A 上的等价关系，对于任意 $a\in A$，集合中所有和 a 是等价关系的元素构成的集合，称为 a 的等价类，记为$[a]_R$。由定义可知$[a]_R=\{b\mid b\in A,<a,b>\in R\}$。

定义 1.2.10 设 A 是一个集合，如果 A 上的一个关系 R，满足自反性。反对称性和传递性，则称 R 是 A 上的一个偏序关系，记为“$\leqslant$”，序偶 $<A,\leqslant>$ 称为偏序集。

如实数集合中的“$\geqslant$”关系是偏序关系。

偏序关系可以用图形来描述，该图形称为哈斯图，下面先介绍“盖住”的定义。

定义 1.2.11 设偏序集 $<A,\leqslant>$ 中，如果 $a,b\in A$，$a\leqslant b$，$a\neq b$ 且没有其他元素 c，满足 $a\leqslant c$，$c\leqslant b$，则称元素 b 盖住元素 a，记为：$COV(A)=\{<a,b>\mid a,b\in A:b$ 盖住 $a\}$。

由于给定的偏序集 $<A,\leqslant>$，它的盖住关系是唯一的，所以可以用盖住的性质画出偏序集合图，也称哈斯图，其画图规则为：

第一步：先画哈斯图的结点，将 A 上的每一个元素作为哈斯图上的结点，任取 $a\in A$，依据以下条件来画结点 a 的位置。

① 若 a 不盖住任何元素，那么 a 被画在最下方(图的最低层)。

② 若 b 盖住 a，则将 b 画在 a 的直接上方(a 的上一层)。

第二步：画哈斯图的边，如果 b 盖住 a，那么在 b 与 a 之间画一条边。

注意：有直接上下层关系的结点间不一定有边相连，有边相连的结点

间，必须是有盖住关系的上下层元素。

例：$A=\{1,2,3,4,6,12\}$，偏序集 $<A,/>$，其中“/”为整除关系。$/=\{<1,1>,<1,2>,<1,3>,<1,4>,<1,6>,<1,12>,<2,2>,<2,4>,<2,6>,<2,12>,<3,3>,<3,6>,<3,12>,<4,4>,<4,12>,<6,6>,<6,12>,<12,12>\}$。其中 $COV(A)=\{<1,2>,<1,3>,<2,4>,<2,6>,<3,6>,<4,12>,<6,12>\}$，其对应的哈斯图如图 1.2.1 所示。

下面讨论偏序集中一些特殊元素，上确界和下确界：

定义 1.2.12 设偏序集 $<A,\leqslant>$ 中，集合 B 是集合 A 的非空子集，对于 $b\in B$，如果不存在 $c\in B$，$b\neq c$，使 $b\leqslant c$，则称 b 为 B 的极大元。

定义 1.2.13 设偏序集 $<A,\leqslant>$ 中，集合 B 是集合 A 的非空子集，对于 $b\in B$，如果不存在 $c\in B$，$b\neq c$，使 $c\leqslant b$，则称 b 为 B 的极小元。

定义 1.2.14 设偏序集 $<A,\leqslant>$ 中，集合 B 是集合 A 的非空子集，若存在 $b\in B$，对于一切 $c\in B$，都有 $c\leqslant b$，则称 b 为 B 的最大元。

定义 1.2.15 设偏序集 $<A,\leqslant>$ 中，集合 B 是集合 A 的非空子集，若存在 $b\in B$，对于一切 $c\in B$，都有 $b\leqslant c$，则称 b 为 B 的最小元。

定理 1.2.1 设 $<A,\leqslant>$ 是一个偏序集合且 B 是 A 的非空子集，若 B 有最大(最小)元，则必是唯一的。

证明 假定 a 和 b 都是 B 的最大元素，则 $a\leqslant b$ 和 $b\leqslant a$，又因为“$\leqslant$”具有反对称性，所以得到 $a=b$。所以 B 的最大元是唯一的，同理可证 B 的最小元也是唯一的。

定义 1.2.16 设 $<A,\leqslant>$ 是一个偏序集合，对于 $B\subseteq A$，如果有 $a\in A$ 且对 B 的任意元素 b，都有 $b\leqslant a$ 成立，则称 a 为子集 B 的上界。同样的，对于 B 中任意元素 b，都有 $a\leqslant b$ 成立，则称 a 为 B 的下界。

定义 1.2.17 设 $<A,\leqslant>$ 是一个偏序集合，且 $B\subseteq A$，a 为 B 的任一上界，若对 B 的所有上界 c，均有 $a\leqslant c$，则称 a 为 B 的最小上界(上确界)。同样，若 b 为 B 的任一下界，若对 B 的所有下界 c，均有 $c\leqslant b$，则称 b 为 B 的最大下界(下确界)。

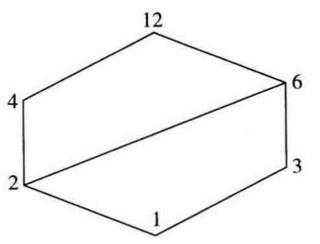

图 1.2.1 哈斯图

定义 1.2.18 设 $< A, \leqslant >$ 是一个偏序集合，如果对任意的 $a, b \in A$，下确界 $a \wedge b$ 和上确界 $a \vee b$ 都存在，则称 $< A, \leqslant >$ 是一个格。

如：设 $P(S)$ 是给定集合 S 的幂集，$< P(S), \subseteq >$ 是一个偏序集。由于 $P(S)$ 中的任意两个元素 T, H，它们的最大下界为 $T \cap H$，最小上界为 $T \cup H$，所以 $< P(S), \subseteq >$ 是一个格。

定义 1.2.19 设 A 是一个非空集合，如果 A 上有二元运算 $\wedge$，$\vee$，集合 A 中的任意元素 a, b, c，满足：

① 交换律：$a \wedge b = b \wedge a, a \vee b = b \vee a$

② 结合律：$(a \wedge b) \wedge c = a \wedge (b \wedge c), (a \vee b) \vee c = a \vee (b \vee c)$

③ 等幂律：$a \wedge a = a, a \vee a = a$

④ 吸收律：$(a \wedge b) \vee a = a, (a \vee b) \wedge a = a$ 且 $a \leqslant b \Leftrightarrow a \wedge b = a, a \leqslant b \Leftrightarrow a \vee b = b$

则称$(A, \wedge, \vee)$是一个格，简称 A 是格。

由格和二元运算 $\wedge$ 与 $\vee$ 的关系可知，上述两个定义是等价的。

定义 1.2.20 设 $< A, \leqslant >$ 是一个格，B 是 A 的非空子集，如果对 B 中的任意元素 a, b，都有 $a \wedge b \in B, a \vee b \in B$，则称 $< B, \leqslant >$ 是 $< A, \leqslant >$ 的子格。

定义 1.2.21 设 $< A, \leqslant >$ 是一个格，如果 $< A, \leqslant >$ 的任一非空子集都有上确界和下确界，则称格 $< A, \leqslant >$ 是完备格。

1.3 经典概念格

概念格是根据数据集合中的对象与属性之间的二元关系建立的一种概念层次结构，概念格的结构模型是形式概念分析理论中的核心数据结构，生动简洁地体现了概念之间的泛化和例化关系。所有的概念同它们之间的泛化及例化关系构成一个概念格，概念格的每个结点是一个形式概念，其相应的 Hasse 图则实现了数据的可视化。所以，概念格是进行数据分析的重要工具之一，概念格理论已经被广泛运用到人工智能等众多领域中。

定义 1.3.1[1,17,18] 设一个三元组(G,M,R)是一个形式背景，其中$G=\{g_1,g_2,g_3,\cdots,g_n\}$为一个对象集，每个$g_i(i\leqslant n)$是一个对象，$M=\{m_1,m_2,m_3,\cdots,m_q\}$为属性集，每个$m_j(j\leqslant q)$一个属性，$R$是$G$和$M$之间的二元关系，即$R\subseteq G\times M$。若$(g,m)\in R$，则称$g$具有属性$m$，用1表示，若$(g,m)\notin R$，则称$g$不具有属性$m$，用0表示。

对于形式背景(G,M,R)，在对象集$A\subseteq G$和属性集$B\subseteq M$上分别定义“$'$”和“$*$”运算：

$$A'=\{m\in M:\forall g\in A,(g,m)\in R\}$$

$$B^*=\{g\in G:\forall m\in B,(g,m)\in R\}$$

A'表示对象集A中所有对象共同具有的属性集合，而B^*表示拥有属性B的对象集合，任意$g\in G$，任意$m\in M$，分别记$\{g\}'$为g'，$\{m\}^*$为m^*称形式背景(G,M,R)是正则的，若对任意$g\in G$，$g'\neq\varphi$，$g'\neq M$，且对任意$m\in M$，$m^*\neq\varphi$，$m^*\neq G$。以下假定形式背景都是正则的。

定理 1.3.1[17,18] 对于形式背景(G,M,R)，任意$A_1,A_2,A_3\subseteq G$，任意$B_1,B_2,B_3\subseteq M$，则有以下基本性质：

(1) 如果$A_1\subseteq A_2$，$B_1\subseteq B_2$，则有：$A'_2\subseteq A'_1$，$B_2^*\subseteq B_1^*$

(2)$A\subseteq A'^*$，$B\subseteq B^{*'}$

(3)$A'=A'^{*'}$，$B^*=B^{*'*}$

(4)$A\subseteq B^*\Leftrightarrow B\subseteq A'$

(5)$(A_1\cup A_2)'=A'_2\cap A'_1$，$(B_1\cup B_2)^*=B_2^*\cap B_1^*$

(6)$A'_2\cup A'_1\subseteq(A_1\cap A_2)'$，$B_2^*\cup B_1^*\subseteq(B_1\cap B_2)^*$

证明 (1)若$m\in A'_2$，对于$\forall g\in A_1$，因为$A_1\subseteq A_2$，所以对于任意g也属于A_2，所以有$(g,m)\in R$，所以$m\in A'_1$，所以$A'_2\subseteq A'_1$。同理可证$B_2^*\subseteq B_1^*$。

(2) 如果$g\in A$，$m\in A'$，则$(g,m)\in R$，又因为A'表示对象集A中所有对象共同具有的属性集合，所以对于任意$m\in A'$，都有$g\in A$使$(g,m)\in R$，所以由定义1.3.1知，$g\in A'^*$，所以有$A\subseteq A'^*$，同理可证$B\subseteq B^{*'}$。

(3) 由(2)的$B\subseteq B^{*'}$可知，$A'\subseteq A'^{*'}$，又由(2)的$A\subseteq A'^*$及(1)知，$A'^{*'}\subseteq A'$，所以$A'=A'^{*'}$。同理可证$B^*=B^{*'*}$。

(4) 因为 $A\subseteq B^*$，所以由(1)得 $B^{*'}\subseteq A'$，又由(2)得 $B\subseteq B^{*'}$，所以 $B\subseteq B^{*'}\subseteq A'$，所以 $B\subseteq A'$；同理 $B\subseteq A'$，所以 $A'^*\subseteq B^*$，又 $A\subseteq A'^*$，所以 $A\subseteq A'^*\subseteq B^*$，所以 $A\subseteq B^*$，所以 $A\subseteq B^*\Leftrightarrow B\subseteq A'$。

(5) 因为 $A_1\subseteq A_1\cup A_2$，由(1)得 $(A_1\cup A_2)'\subseteq A'_1$，同理 $(A_1\cup A_2)'\subseteq A'_2$，所以有 $(A_1\cup A_2)'\subseteq A'_1\cap A'_2$；又因为 $\forall m\in A'_1\cap A'_2$，所以有：$m\in A'_1$ 且 $m\in A'_2$，又 $\forall g_1\in A_1, g_2\in A_2$，有 $g_1\subseteq m^*$，$g_2\subseteq m^*$，所以 $g_1\cup g_2\subseteq m^*$，由(1)知 $m\in(g_1\cup g_2)'$，由于 g_1 和 g_2 的任意性，所以 $m\in(A_1\cup A_2)'$，所以 $A'_1\cap A'_2\subseteq(A_1\cup A_2)'$。所以有 $(A_1\cup A_2)'=A'_1\cap A'_2$，同理可证 $(B_1\cup B_2)^*=B_1^*\cap B_2^*$。

(6) 因为 $A_1\cap A_2\subseteq A_1$，$A_1\cap A_2\subseteq A_2$，由(1)知 $A'_1\subseteq A'_1\cap A'_2$，$A'_2\subseteq A'_1\cap A'_2$，所以 $A'_1\cup A'_2\subseteq A'_1\cap A'_2$，同理可证得 $B_1^*\cup B_2^*\subseteq(B_1\cap B_2)^*$。

定义 1.3.2[17,18] 设三元组 (G,M,R) 是一个形式背景，如果对于任意 $A\subseteq G, B\subseteq M$，二元组 (A,B) 满足 $A'=B, B^*=A$，则称 (A,B) 是一个形式概念，其中 A 称为概念的外延，B 称为概念的内涵。

定理 1.3.2[17,18] 形式背景 (G,M,R)，对于任意 $A\subseteq G, B\subseteq M$，则 (A'^*,A') 和 $(B^*,B^{*'})$ 都是形式背景 (G,M,R) 的概念。

证明 因为 $(A'^*)'=A'^{*'}$，又由 $A'^{*'}=A'$，所以 $(A'^*)'=A'$，又 $(A')^*=A'^*$，所以 (A'^*,A') 是形式背景 (G,M,R) 的概念，同理可证 $(B^*,B^{*'})$ 也是一个概念。

用 $L(G,M,R)$ 表示形式背景 (G,M,R) 的全体概念，称为概念格，在 $L(G,M,R)$ 上定义序关系如下：

$$(A_1,B_1)\leqslant(A_2,B_2)\Leftrightarrow A_1\subseteq A_2(\Leftrightarrow B_2\subseteq B_1)$$

其中的上确界和下确界定义如下：

$$(A_1,B_1)\wedge(A_2,B_2)=(A_1\cap A_2,(B_1\cup B_2)^{*'})$$

$$(A_1,B_1)\vee(A_2,B_2)=((A_1\cup A_2)'^*,B_1\cap B_2)$$

事实上概念格 $L(G,M,R)$ 是一个完备格。

例 1 在形式背景 (G,M,R) 中，$G=\{1,2,3,4\}$，$M=\{a,b,c,d,e\}$，任意 $(g,m)\in R$ 时，用 1 表示，任意 $(g,m)\notin R$ 用 0 表示，其对应的概念格用