区间概念格及其应用

张春英　刘保相　王立亚　著

科学出版社

区间概念格及其应用

张春英　刘保相　王立亚　著

科学出版社
北　京

内 容 简 介

概念格是根据数据集中对象与属性之间的二元关系建立的一种概念层次结构，是进行数据分析和知识处理的有效工具. 经过三十多年的发展，概念格已经从经典概念格拓展出了模糊概念格、加权概念格、约束概念格、量化概念格、区间值概念格、粗糙概念格等，本书在详细分析概念格的最新研究进展的同时，提出了一种新的概念格结构——区间概念格，详细讨论了区间概念格的结构与性质、构造算法、维护原理、压缩方法、动态合并、参数优化、规则提取及其在多个领域的应用方法.

本书可供信息科学技术、计算机科学与技术、智能科学与技术、自动化、控制科学与工程、管理科学与工程和应用数学等专业的教师、研究生、高年级本科生使用，也可供科研人员、工程技术人员参考.

图书在版编目（CIP）数据

区间概念格及其应用 / 张春英，刘保相，王立亚著. —北京：科学出版社，2017.5

ISBN 978-7-03-052572-7

Ⅰ. ①区… Ⅱ. ①张… ②刘… ③王… Ⅲ. ①区间-研究 Ⅳ. ①O174.1

中国版本图书馆 CIP 数据核字（2017）第 083062 号

责任编辑：王胡权 / 责任校对：邹慧卿
责任印制：张　伟 / 封面设计：陈　敬

科学出版社 出版
北京东黄城根北街 16 号
邮政编码：100717
http://www.sciencep.com

北京京华虎彩印刷有限公司 印刷

科学出版社发行　各地新华书店经销

*

2017 年 5 月第　一　版　开本：720 × 1000　B5
2017 年 5 月第一次印刷　印张：13 1/2
字数：270 000

定价：59.00 元

前　言

自从20世纪80年代德国Wille教授以形式概念分析(Formal Concept Analysis, FCA)重构格理论(Lattice Theory)之后，形式概念分析和基于形式概念分析的概念格相关理论与技术作为数据分析研究领域的重要方法和手段逐渐地受到国际学术界的广泛关注.

利用经典概念格提取的规则只能是确定性规则，导致在某些情况下挖掘成本过高、规则库过于庞大以及规则应用效率过于低下. 为此，国内外学者在经典概念格基础上进行了多种拓展研究，主要有模糊概念格、加权概念格、变精度概念格、区间值概念格、粗糙概念格等. 在一般的概念格中，外延是必须具备内涵中的所有属性来形成概念结点，进而进行格结构构建的，在概念格中查找具有部分属性的概念结点时必须要对概念进行合并，特别是对规模庞大的概念格来说时间代价太高. 粗糙概念格可以对具有内涵中部分属性的概念结点进行查找，然而会查找出大量仅具备概念内涵中一个属性的对象，由此挖掘出的关联规则的支持度和置信度都大大降低. 在实际应用中，管理者往往关心具备内涵中一定数量或比例的属性的对象集合，并由此进行关联规则挖掘，为制定有针对性的决策提供依据.

为了解决上述问题，作者提出了一种新型的概念层次结构——区间概念格，其概念外延是参数区间$[\alpha,\beta]$ $(0\leqslant\alpha\leqslant\beta\leqslant1)$内满足内涵中属性的对象集合. 本书从区间概念格的研究背景和意义出发，阐述分析了概念格的基本理论及其应用；从区间概念格的结构、性质开始，对区间概念格的建格算法、动态压缩原理、纵横向动态维护、多区间概念格的动态合并、参数优化、带参规则挖掘及其应用进行了详细的分析，并给出具体的应用实例，为读者提供新的数据分析思路和方法.

全书共分 9 章. 第 1 章绪论，从概念格的构建、扩展、约简、规则提取及概念格与粗糙集的结合等方面，详细分析了国内外有关概念格的研究进展，讨论了目前研究中存在的问题和不足，引出区间概念格的研究意义. 第 2 章区间概念格的结构与性质，给出了区间概念格的定义，并对其结构特征及性质进行了详细分析，特别针对决策形式背景提出决策区间概念格的相关概念. 第 3 章区间概念格的构造算法与实现，从经典概念格的构造原理出发，基于作者提出的属性链表方法构造关联规则格的思路，提出基于属性集合幂集的区间概念格构造算法，通过实例验证了算法的高效性. 第 4 章区间概念格的动态压缩，从经典概念格的属性

约简出发，在运用概念格同构方法对概念格约简的基础上，提出了基于覆盖的区间概念格动态压缩思想，并构建了动态压缩模型，实例证明了模型的有效性. 第 5 章区间概念格的动态维护，介绍了基于属性链表的概念格纵向和横向维护算法，详细阐述了区间概念格的横向和纵向维护原理，提出了基于属性集合的区间概念格动态维护算法. 第 6 章多区间概念格的动态合并，主要针对大量数据存在的情况，提出对分时产生的多区间概念格进行动态合并的思想，给出了纵向与横向合并的基本原理，设计了动态合并算法，为关联规则的进一步挖掘提供可操作性. 第 7 章区间概念格的带参数规则挖掘，针对不确定规则的度量标准——精度和不确定度,构建了基于区间概念格的带参数规则挖掘模型，研究了区间参数对区间关联规则的影响；进一步提出一种区间关联规则动态并行挖掘算法，并通过算法分析与实例证明说明该算法的正确性和高效性. 第 8 章区间概念格的参数优化，结合区间概念格结构特性，提出了基于参数变化的格结构更新算法，研究不同参数对区间概念及关联规则的影响度，分别构建了基于学习、基于遗传算法和基于信息熵的区间参数优化模型,对区间概念格的参数选取提供一种切实有效的方法. 第 9 章区间概念格的应用，主要讨论概念格、区间概念格等的应用问题，分别应用在群体决策、本体形式背景抽取、气象云图识别、中医喘证用药、三支决策、水库调度等领域，取得了较好的应用效果.

本书在编写过程中得到了很多专家的指点，特别是合肥工业大学的胡学钢教授提出了很多中肯的意见和建议，非常感谢专家们的点拨指导. 这本书能够展示给各位读者，离不开李明霞、张茹的认真编辑和校对，谢谢你们. 另外，本书引用了大量专家学者的文献，给了我们很大的启发，谢谢你们的无私奉献.

本书承蒙国家自然科学基金项目(课题名称：区间概念格理论及在粗糙控制规则动态优化中的应用研究，项目编号：61370168)资助，得到了河北省数据科学与应用重点实验室、华北理工大学河北省重点学科数学一级学科的大力支持，在此一并表示感谢.

由于作者水平有限，书中难免有疏漏之处，敬请读者批评指正. 若有问题探讨，可发邮箱 hblg_zcy@126.com，非常感谢！

作　者

2016 年 5 月

目　录

第1章 绪　　论

1.1 研究背景及目的意义

1.1.1 研究意义

概念格是根据数据集中对象与属性之间的二元关系建立的一种概念层次结构，是进行数据分析和知识处理的有效工具，现已被广泛应用于软件工程、知识工程、智能控制等领域. 特别地，由于概念格结点反映了概念内涵和外延的统一，结点间关系体现了概念之间的泛化和例化关系，因此适合于作为规则发现的基础性数据结构. 而规则提取是智能控制中至关重要的环节，规则挖掘的精度直接影响着控制的效率和准确度.

利用经典概念格提取的规则只能是确定性规则，这在某些情况下挖掘成本会过高、规则库会过于庞大、规则应用效率会过于低下. 为此，国内外学者在经典概念格基础上进行了多种拓展研究，主要有模糊概念格、加权概念格、约束概念格、量化概念格、区间值概念格、粗糙概念格等. 在模糊概念格中，基于隶属度阈值描述变精度概念，可查询一定隶属程度下的对象集，而隶属度阈值的选取始终具有一定的主观性；在加权概念格中，对概念的描述考虑了属性重要度，对概念格进行了一定程度的简化；不过二者依然是一定条件下基于外延完全具备内涵中所有属性进行构建，要查找具有部分属性的概念则要通过对概念格进行扫描，对概念进行合并，一般情况下会造成时空代价偏高，应用效率偏低；区间值概念格是基于对象的属性值在某一个区间$[a,b]$范围内的信息表提出的，比模糊形式背景表示的数据隶属度范围$[0,1]$更广泛，格上的每个概念仍然是要满足全部属性的，要进行部分检索同样存在着一定的困难；在粗糙概念格中，可以查找具有部分属性的概念，解决了确定-不确定概念描述问题，可同时提取确定性和不确定性规则，然而由于其概念的上近似外延只要求属性与内涵的交集非空即可，这样可能会存在大量仅具备内涵中一个属性的对象，由此构建的规则支持度和置信度大大降低. 因此使用粗糙集理论中的属性约简等方法对规则进行简化，或应用粗糙概念格等工具进行不确定性规则挖掘，虽部分缓解了规则库庞大等问题，但也带来了规则置信度难以测度与调控的新问题，产生了规则挖掘成本、应用效率与可靠性之间的矛盾. 将规则作为控制规则应用于智能控制特别是粗糙控制时，这一

矛盾成为粗糙控制实用化难以解决的瓶颈.

为了解决上述问题，作者提出并初步研究了一种新型概念层次结构——区间概念格[1]，其概念外延是区间$[\alpha,\beta]$ $(0\leqslant\alpha\leqslant\beta\leqslant 1)$范围内满足内涵属性的对象集. 通过研究区间概念格的理论、建格与带参数的规则挖掘算法，挖掘区间参数$[\alpha,\beta]$与规则置信度μ的互动规律，进而构建依μ调控区间参数$[\alpha,\beta]$的机理和方法；将其应用于粗糙控制，建立控制规则动态优化的数学模型，以期达到规则挖掘成本、应用效率与可靠性的整体最优. 基于区间概念格进行的规则挖掘具有动态优化的挖掘特性，对于规则提取及其质量控制有着重要的理论意义与应用价值.

1.1.2 国内外研究现状分析

区间概念格是经典概念格、粗糙概念格的拓展，重点研究区间概念格的构建算法、概念格约简、规则挖掘，以及与其他扩展概念格的比较，故重点从以下几个方面进行现状分析:

1. 概念格的构建

国内外学者和研究人员对此进行了深入研究，提出了一些有效算法来生成概念格，一般分为两类: 批处理算法和增量算法. 增量算法因其效率较高更受关注，最典型的为 Godin 算法[2]. 在此基础上，国内吉林大学的刘大有等[3]提出了一种基于搜索空间划分的概念生成算法；北京科技大学的杨炳儒等[4]通过构建树结构，缩小产生子格结点的范围，设计了增量式广义概念格构建算法；Lv 等[5]提出了一种通过增加标志对渐进式算法进行改进；合肥工业大学胡学钢等[6]提出了一个构造概念格的批处理和渐进式混合算法；上海大学的刘宗田等[7]通过对概念格进行纵向与横向合并形成一个新的概念格，提高了建格效率；山西大学梁吉业等提出了基于优势关系的概念格定义及构建方法[8]；作者给出了基于属性链表的概念格渐进式构造算法和纵横向维护算法[9]. 国外 Valerie 等[10]基于阈值和模糊闭合算子给出了模糊概念格的渐进式构造算法. 姚佳岷等[11]提出一种基于概念内涵、外延升降序的双序渐进式合并算法，同时实现子概念格的纵横向合并. 这些算法提高了概念格的构造效率，是概念格构建方面重要的研究成果.

2. 概念格的扩展

目前对概念格的扩展主要有：模糊概念格、实区间格、量化概念格、加权概念格、约束概念格、扩展概念格、粗糙概念格等，其中模糊概念格广受国内外学者关注，是目前的研究热点. Isabelle 基于双枝模糊集对概念格进行了拓展研究[12]. Michal 等[13]通过修改输入数据对模糊概念格进行分解；Pavel[14]详细分析了模糊概

念格中基于L-平等的L-序集和基于L-脆平等的L-序集之间的关系；刘宗田等[15, 16]提出了一种模糊概念格模型，定义了模糊概念的模糊参数σ和λ，给出了模糊概念格渐进式构造算法；仇国芳等[17]在两个完备格之间引入外延内涵算子与内涵外延算子，由两个完备格及两个算子组成的四元组构成了概念粒计算系统；作者对模糊概念格中参数选择及其优化进行了研究，给出了一种切实可行的选择优化方法[18]. 实区间格则是基于区间值$[a,b]$形式背景构建的，其区间值比模糊形式背景中的隶属度值$[0,1]$更广泛，刘宗田等[19]采用区间属性定标方法对区间数进行分解，将基于区间值的信息表转化为区间值形式背景，所构建的区间值概念格具有较合理的时空性能.

3. 概念格的约简

在概念格的渐进式构造过程中，采用剪枝的方法消除冗余信息，使概念格内涵的比较次数减少，提高了建格效率[7]. Aswanikumar 等[20]提出了一种基于K-均值模糊聚类方法对概念格进行剪枝的方法. Michal 等[13] 研究了通过分解减小模糊概念格的方法.

在结点外延不变的前提下，对概念格结点的内涵进行约简，减少基于概念格的规则提取时出现的冗余. Li[21]等基于模糊形式背景提出了 δ-约简概念，并给出了相关性质和定理. 张文修等[22]借鉴粗糙集约简理论提出了形式背景概念格的约简判定定理；仇国芳等[23]研究了外延集类上的等价关系和交的一致关系，得出了决策形式背景下规则提取与属性约简方法.

4. 概念格的规则提取

国内外学者在基于概念格的关联规则挖掘方面进行了深入的研究，Valtchev 等[24]提出利用概念格挖掘频繁闭项目集的算法；Zaki 等[25]利用概念格的闭合特性，提出了挖掘无冗余的关联规则算法；梁吉业等[26]提出了一种基于闭标记的渐进式规则提取算法；刘宗田等[27]提出利用容差关系建立广义概念格提取近似规则；胡学钢等[28]提出一种基于约简概念格的关联规则快速求解算法；李金海等提出基于概念格的决策形式背景属性约简及规则提取算法[29]；仇国芳、张文修等通过决策推理将变精度概念格生成的少数决策规则集拓展为所有方案集上的全部决策推理规则，获得了方案集的下近似决策推理规则和上近似决策推理规则[16,30]；Tang 等基于分类概念格提出了分类规则挖掘算法[31]；Greco 等基于模糊粗糙集给出了一种从粗糙集的上下近似中提取渐进决策规则的方法[32]；Fan 等基于粗糙集给出增量规则抽取方法，提高了工作效率[33]；Wu 等针对区间值信息系统提出了一种基于粗集方法的分类规则发现方法[34]；Tzung 等基于模糊粗糙集理

论提出一种从不完备系统中同时提取确定和不确定模糊规则的方法，并估计了学习过程中的遗漏值[35]；董威等基于可变精度粗糙集理论提出一种根据粗糙规则集的不确定性量度进行粗糙规则挖掘的算法[36]，该方法具有一定的容错能力，但是如何动态修正规则集以及阈值如何调整才能获得最适合的规则并未给出；王国胤等针对面向领域用户的决策规则挖掘问题,用属性序描述领域用户的需求和兴趣，提出了一种属性序下的分层递阶决策规则挖掘算法[37]；黄加增运用粗糙概念格给出了决策形式背景下的多属性约简与规则提取方法[38]；粗糙集已获得了一些成功的应用实例，但始终应用有限，究其原因应是其在技术上还存在着一些问题，如置信水平低、规则数量庞大等.

通过以上的分析可知，运用粗糙集进行属性约简、提取规则或者直接运用粗糙概念格挖掘规则，虽然能够同时提取确定性-不确定性规则，也部分地缓解了规则库过于庞大并提高了规则挖掘的实效性，但是其置信度和支持度仍然过于低下.于是，寻找一种更高效的规则表示模型和挖掘算法是当前需要迫切解决的问题.区间概念格是在粗糙概念格基础上，考虑概念外延为区间$[\alpha,\beta](0\leqslant\alpha\leqslant\beta\leqslant1)$范围内满足内涵属性的对象集而得到的一种新的概念层次结构，其能够描述决策中对符合一定条件范围的规则进行提取的实际问题.

5. 概念格与粗糙集的结合

采用粗糙集方法对概念格特征进行分析：Yao 等[39]在形式概念分析中引入粗糙近似概念，定义两种不同类型的近似算子，并对形式概念分析和粗糙集的异同进行了研究；张文修等在文献[40]中对概念格和粗糙集理论进行了详细分析；梁吉业等[41]利用形式概念分析中名义梯级背景的概念，证明了粗糙集理论中的划分、上下近似、独立、依赖、约简等核心概念都可以在相应的衍生背景中进行表示. 作者则探讨了概念格上的粗集、S-粗集和变异粗集特征[42].

(1) 采用粗糙集理论方法对概念格内涵进行处理：文献[43]利用粗糙集方法重新认识形式概念和概念格,为概念格的约简提供了一种新的思路和方法. 文献[44]中基于可变精度粗糙集的决策规则格，每一结点均可表示为相应的决策规则；文献[45]基于粗糙集理论给出一种多级属性约简方法和对象约简方法；文献[46]提出一种概念格约简的灰色粗集方法.

(2) 采用粗糙集理论对概念格进行扩展：近年来，为解决一些带有不确定性的现实问题，杨海峰等[47]融合粗糙集理论，定义概念格中内涵所拥有的上近似外延和下近似外延，形成新的格结构-粗糙概念格. 杨凌云等[48]对基于剩余格的 L-形式背景引入了 L-可定义集和 L-粗糙概念格，给出全体 L-可定义集恰为 L-粗糙概念外延的一个充要条件. 胡学钢等[49]基于变精度粗糙集建立了近似概念格模型，

对概念格结点通过近似关系进行合并.

粗糙概念格体现了对象与特征之间的确定与不确定关系. 利用粗糙概念格进行规则挖掘无疑扩大了概念格理论的应用范围，但由于不同概念结点的上、下近似集的比值不同且有可能差距过大，使得规则挖掘质量无法控制或者规则置信度过低，严重影响了获取的规则的可靠性. 作者提出了一种基于粗糙概念格的 $L(P)$ 关联规则挖掘算法，给出了一些特殊情况下提高和调控规则置信度的方法.

1.2　经典概念格

1.2.1　概念格的基本概念

定义 1.1　称 (U,A,R) 为一个形式背景，其中 $U=\{x_1,x_2,\cdots,x_n\}$ 为对象集，每个 $x_i(i\leqslant n)$ 称为一个对象； $A=\{a_1,a_2,\cdots,a_m\}$ 为属性集，每个 $a_j(j\leqslant m)$ 称为一个属性； R 为 U 和 A 之间的二元关系， $R\subseteq U\times A$. 若 $(x,a)\in R$ ，则称 x 具有属性 a ，记为 xRa .

定义 1.2　对于形式背景 (U,A,R) ，算子 f ， g 定义为

$\forall x\in U, f(x)=\{y\mid \forall y\in A, xRy\}$ ，即 f 是对象 x 与其具有的所有属性的映射；

$\forall y\in A, g(y)=\{x\mid \forall x\in U, xRy\}$ ，即 g 是属性 y 与其覆盖的所有对象的映射.

定义 1.3　对于形式背景 (U,A,R) ，若对于 $X\subseteq U$ ， $Y\subseteq A$ ，有 $f(x)=Y$ ， $g(y)=X$ ，则称序偶 (X,Y) 是一个形式概念，简称概念. 其中 X 称为概念的外延，Y 称为概念的内涵.

定义 1.4　用 $L(U,A,R)$ 表示形式背景 (U,A,R) 的全体概念，记:

$$(X_1,Y_1)\leqslant(X_2,Y_2)\Leftrightarrow X_1\subseteq X_2(\Leftrightarrow Y_1\supseteq Y_2)$$

则“$\leqslant$”是 $L(U,A,R)$ 上的偏序关系.

定义 1.5　若 $L(U,A,R)$ 中的所有概念满足“$\leqslant$”偏序关系，则可得到一个有序集 $\overline{CS(L)}=(L,\leqslant)$ ，形成一个完备格，称 $\overline{CS(L)}$ 是形式背景(U,A,R)的概念格.

定义 1.6　如果 $(X_1,Y_1)\leqslant(X_2,Y_2)$ ，且二者之间不存在概念 (X_3,Y_3) ，满足:

$$(X_1,Y_1)\leqslant(X_3,Y_3)\leqslant(X_2,Y_2)$$

则称 (X_1,Y_1) 是 (X_2,Y_2) 的子概念， (X_2,Y_2) 是 (X_1,Y_1) 的父概念.

定义 1.7　若只考虑 R 是对象和属性间的布尔关系的情形，且有 (X_1,Y_1) ， (X_2,Y_2) 是 $\overline{CS(L)}$ 上的两个概念，则定义二者的最大下界、最小上界为

$$(X_1,Y_1)\vee(X_2,Y_2)=(g(Y_1\cap Y_2),(Y_1\cap Y_2))$$

$$(X_1,Y_1)\wedge(X_2,Y_2)=((X_1\cap X_2),f(X_1\cap X_2))$$

可知，$\overline{CS(L)}$上的最大元为$(U,f(U))$，最小元为$(g(A),A)$；并且$X\subseteq U$，$Y\subseteq A$，则$(g(f(X)),f(X))$或$(g(Y),f(g(Y)))$是$\overline{CS(L)}$的概念.

表 1.1 所示为一个简单的形式背景，图 1.1 是其对应的概念格.

表 1.1　一个简单的形式背景

	a	*b*	*c*	*d*	*e*	*f*	*g*	*h*	*i*
1	1	0	1	0	0	1	0	1	0
2	1	0	1	0	0	0	1	0	1
3	1	0	0	1	0	0	1	0	1
4	0	1	1	0	0	1	0	1	0
5	0	1	0	0	1	0	1	0	0

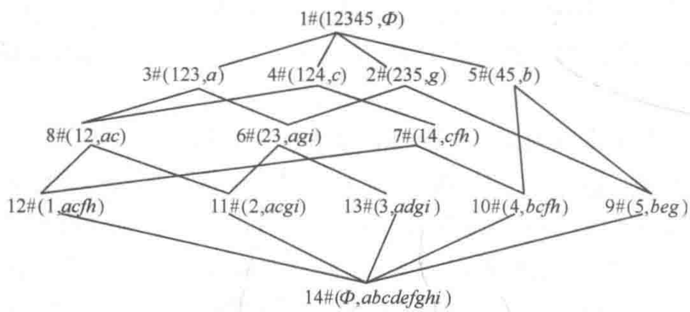

图 1.1　形式背景对应的概念格

1.2.2　概念格的结构特征

1. 概念格的分层特征

概念格有明显的分层特征[50]，基于此，可以进行快速建格.

定义 1.8　在概念格$\overline{CS(L)}$中，若概念(X,Y)到格中最大结点的极大链的长度为N，则称(X,Y)是第N层的格结点.

由此可得概念格的如下性质特征.

性质 1.1　同层格结点之间相互是不可比的.

证明　设C_1,C_2为同一层格结点，若C_1与C_2是可比较的，则C_1与C_2之间有链；因此，当$C_1\leqslant C_2$，则C_1到最大格结点的极大链的长度大于C_2到极大链的长度，这与它们在同一层的定义矛盾.

性质 1.2　任一个第$N+1$层的格结点，至少被一个第N层的格结点所覆盖.

证明 由定义，若一个第 $N+1$ 层的格结点，不被任何第 N 层的格结点所覆盖，则它只能被第 $N-1$ 层的格结点或 $N-1$ 层以上的结点覆盖. 这样，此结点到最大元的极大链的长度至多为 N，这与它是第 $N+1$ 层的格结点矛盾.

性质 1.3 若第 N 层的格结点 (X_N, Y_N) 覆盖了第 $N+1$ 层的格结点 (X_{N+1}, Y_{N+1})，则 $\forall y \in Y_{N+1} - Y_N$，有 $g(y) \cap g(Y_N) = X_{N+1}$.

证明 若存在 $\forall y \in Y_{N+1} - Y_N$，有 $g(y) \cap g(Y_N) \neq X_{N+1}$，则若 $g(y) \cap g(Y_N) \subset X_N$ 时，由概念的定义: $X_{N+1} = g(Y_{N+1}) = \bigcap_{y \in Y_{N+1}} g(y) \subset (g(y) \cap g(Y_N)) \supset X_N$，导出矛盾.

若 $g(y) \cap g(Y_N) \supset X_{N+1}$，则 $X_{N+1} = g(Y_{N+1}) = \bigcap_{y \in Y_{N+1}} g(y) \subset (g(y) \cap g(Y_N)) \supset X_N$，也导出矛盾，命题成立.

性质 1.4 在概念格中除去 $(g(A), A)$ 结点外，至多可分为 T_0 层，且

$$T_0 = \max\left(\sum_{j=1}^{n} R(x_1, y_j), \sum_{j=1}^{n} R(x_2, y_j), \cdots, \sum_{j=1}^{n} R(x_m, y_j)\right)$$

并且，对任意的格结点 (X,Y)，若 $|Y| = T_0$，则 (X,Y) 覆盖 $(g(A), A)$ 或者 $(X,Y) = (g(A), A)$.

证明 由于层数每增加 1 时，下 1 层格结点的属性集中的属性至少增加 1 个；所以，由此性质可知:

$$\begin{aligned} T_0 &= \max\left(\sum_{j=1}^{n} R(x_1, y_j), \sum_{j=1}^{n} R(x_2, y_j), \cdots, \sum_{j=1}^{n} R(x_m, y_j)\right) \\ &= \max\{|f(x)| \,|\, x \in I\} \geqslant \left|\bigcap_{x \in X} f(x)\right| \end{aligned}$$

这里 T_0 实际上是单个对象的属性个数的最大值.

2. 概念格上的粗集特征

根据概念格形式背景的二值性特点，由概念格结点划分等价类的时候只能根据属性值，将对象集分为两类，即具有该属性的对象的集合和不具有该属性的对象的集合. 然而，实际上，按照粗糙集的特点，对象集合在划分时可能出现部分具有属性的特点，因此，概念格上结点具有粗糙集特征[35].

定理 1.1 对于简单形式背景 (U, A, R)，设 $d \in A$，$x \in X \subseteq U$，令

$$C_d = \{\text{Concept} \,|\, d \in \text{Intent(Concept)}\}$$

其中，Intent(Concept) 表示概念 Concept 的内涵. 则 $[x]_d = \{\text{Extent}(\text{Sup}(C_d))\}$ 或者 $[x]_d = \{U - \text{Extent}(\text{Sup}(C_d))\}$ 即 $U / \text{IND}(d) = \{\{\text{Extent}(\text{Sup}(C_d))\}, \{U - \text{Extent}(\text{Sup}(C_d))\}\}$

定义 1.9 令 $C_d = \{\text{Concept} \,|\, d \in \text{Intent(Concept)}\}$，定义 $X \subset P(U)$ 的下近似: