组合数学论丛

超图的理论基础

王建方 编著

高等教育出版社
Higher Education Press

● 组合数学论丛

超图的理论基础

■ 王建方 编著

高等教育出版社
Higher Education Press

内容简介

本书介绍源于数据库理论的超图理论。主要内容为无圈超图理论和超图的圈结构理论。超图的圈公理构成了该理论的基础。这是全新的理论，且在信息科学、生命科学、经济学、计算机科学等领域有重要应用。该书适合数学和上述领域的研究人员、高校教师、研究生参考使用。

图书在版编目（CIP）数据

超图的理论基础 / 王建方编著：—北京：高等教育出版社，2006.7

ISBN 7-04-018810-4

Ⅰ.超…　Ⅱ.王…　Ⅲ.超图-基础理论　Ⅳ.O157.5

中国版本图书馆 CIP 数据核字（2006）第 043234 号

策划编辑　郭　伟　　**责任编辑**　郭　伟　　**封面设计**　王凌凌
责任绘图　朱　静　　**版式设计**　王　莹　　**责任校对**　朱惠芳
责任印制　韩　刚

出版发行	高等教育出版社	购书热线	010-58581118
社　　址	北京市西城区德外大街 4 号	免费咨询	800-810-0598
邮政编码	100011	网　　址	http://www.hep.edu.cn
总　　机	010-58581000		http://www.hep.com.cn
		网上订购	http://www.landraco.com
经　　销	蓝色畅想图书发行有限公司		http://www.landraco.com.cn
印　　刷	北京鑫丰华彩印有限公司	畅想教育	http://www.widedu.com
开　　本	787×960　1/16	版　　次	2006 年 7 月第 1 版
印　　张	10.25	印　　次	2006 年 7 月第 1 次印刷
字　　数	160 000	定　　价	23.00 元

物料号　18810-00

序 言

2002 年, 我在中国科学院数学与系统科学研究院开设了一门博士生学位课 —— 图与超图, 主要讲述源于数据库理论的超图理论。后来在浙江师范大学、中国矿业大学讲述过部分内容, 基本东西是近些年的研究成果。听讲的朋友都建议能写一本书。在讲课中, 我发现分散在论文中的结果很零散, 有些东西写得粗糙, 甚至有毛病, 很有必要进行整理、修正、系统化。所以决定写这本小册子, 供有兴趣的朋友参阅。

我非常感谢香港中文大学李东讲座教授, 是他把我引入这个领域的。他总是有非常深刻的思想, 善于抓住事物的本质。他引入超图的交闭半格的概念和由此产生的一个参数, 这对研究超图理论非常重要。我们进行了长时间的合作, 合作非常愉快, 成效很大。

我要感谢 Beeri,Fagin,Maier Yannakakis, 他们的系统而富创见性的论文 "On the desirability of acyclic database schemes" 给了我很大启发。

我要感谢李海珠博士, 我与他进行了 2 年多的富有成效的合作, 他有很深刻的思想。感谢闫桂英博士、许宝光博士、吉日木图教授, 他们在这一领域的研究中都作出了自己的贡献。

我还要感谢程郁琨、李培之、任伟, 他们帮我完成了书稿的打印。

这个领域研究的历史很短, 还不成熟, 再加上作者水平有限, 错误缺点在所难免, 诚恳欢迎读者批评指正。

王建方

2005

目 录

第一章　基本概念和术语

引言

超图是有限集合的子集系统, 是离散数学中最一般的结构. 早期的定理有 Sperner 定理和 Ramsey 定理等. 于 20 世纪 60 年代,“超图” 这个词正式提出来, 是作为普通图的推广,基 本概念和定义都是图的相应概念和定义的平移与推广, 业已取得了一些重要结果, 如 Erdös-Ko-Rado 定理等. Berge 写了一本专著“Hypergraphs”[6] 对其做了系统的总结.

进入信息时代, 信息科学技术对人类社会各个领域都产生着巨大的影响, 也为创建发展新的数学理论提供了机遇、源泉和动力. 由于信息科技、生命科技的不断发展, 人们要研究处理的系统也越来越庞大, 越来越复杂. 集成化就成了一个重要方向. 就是要把一个大系统化为子系统的集成. 反映在数据库理论中, 就是把大数据库化为小数据库的联合. 首先把数据库的属性集合化为其子集合的并, 形成数据库图式. 信息科学的发展, 特别是数据库理论的发展为超图理论的发展注入了新的活力, 赋予了新的内涵, 给予了巨大动力.

20 世纪 80 年代, 信息科学家研究数据库理论时, 就发现超图与数据库密切相关, 而超图圈的传统定义与数据库的性质相差甚远, 在 [7~10] 中他们引入了无圈超图的概念. 这不是一个直观定义, 而是由运算过程来界定. 我们这里将称之为超图的无圈公理. 他们证明了由无圈公理界定的无圈超图在数据理论中十分有用. 后来的理

论成果 [23~27] 也显示了无圈公理是科学的. 无圈公理构成了无圈超图理论的基础.

超图圈的传统定义与数据库的性质相差太远. 那么什么是超图的圈呢? Wang 和 Lee[11] 给出了超图圈的一个新定义. Wang[38] 将这个定义做了修正, 形成一个公理, 称之为圈公理. 超图的圈由圈公理来界定. Wang[38] 给出了基本定理, 证明了圈公理覆盖无圈公理.

从数据库理论来看, 圈结构在超图理论中是最本质、最基本的. 圈公理构成了超图理论的基础.

定义 1.1 设 $V = \{x_1, x_2, \cdots, x_n\}$ 是一个有限集合, $\varepsilon = \{e_1, e_2, \cdots, e_m\}$ 是 V 的一个子集合族, 即 $\varepsilon \subseteq 2^V$. 我们称 $\mathcal{H} = (V, \varepsilon)$ 是在 V 上的一个超图, 如果 $\forall i \in \{1, 2, \cdots, m\}, e_i \neq \varnothing$ 和 $\cup_{i=1}^{m} e_i = V$. 我们也简单地称 ε 是在 V 上的一个超图. V 的元素称为顶点, ε 的元素称为边. $|V|$ 称为 ε 的阶, $|\varepsilon|$ 称为 ε 的规模.

如果 ε 的每条边 e, 都有 $|e| = k$, 则称 ε 是一个 k–匀齐超图. 如 k=2, 则 ε 就是一个通常的图.

定义 1.2 一个超图 ε 称为不可约超图, 或简单超图, 如果它的任意两条边都互不包含, 即 $\forall e, e' \in \varepsilon, e \neq e'$, 都有 $e \setminus e' \neq \varnothing$. 下面除特别说明, 所说的超图都是简单超图.

设 ε 和 ε' 是两个超图, 如果 $\varepsilon' \subset \varepsilon$, 则称 ε' 是 ε 的部分超图.

设 ε 是 V 上的一个超图, $S \subset V$, $\varepsilon[S] = \{e \in \varepsilon : e \subseteq S\}$ 称为 ε 的由 S 导出的子超图.

令 $V^{(k)} = \{S \subseteq V : |S| = k\}$.

设 ε 是 V 上的一个超图, 定义 $\varepsilon^{(k)} = \{S \in V^{(k)}$: 对某个 $e \in \varepsilon, S \subseteq e\}$.

例如: $V = \{a, b, c, d, e\}, \varepsilon = \{abcd, abce, bcde\}$, 则

$$V^{(3)} = \{abc, abd, abe, acd, ace, ade, bcd, bce, bde, cde\}.$$

$$\varepsilon^{(3)} = \{abc, abe, acd, ace, bcd, bce, bde, abd, cde\}.$$

定义 1.3 一个超图 ε 称为是保形的 (conformal), 如果 $\varepsilon^{(2)}$ 的所有极大团的集合等于 ε 的所有边的集合.

例如 $\varepsilon_1 = \{abc, bcd, cde\}$ 是保形的, 而 $\varepsilon_2 = \{abc, cde, efa\}$ 不是保形的, 因为 $\{ac, ce, ea\}$ 是 $\varepsilon^{(2)}$ 的一个极大团, 而 ace 不是 ε 的一条边.

定义 1.4 设 ε 是在 V 上的一个超图, $\bar{\varepsilon} = \{\bar{e} = V \setminus e : e \in \varepsilon\}$ 称为 ε 的伴随超图.

例如:$V = \{a, b, c, d, e, f, g\}, \varepsilon = \{abc, cde, def, efg\}$, 则

$$\bar{\varepsilon} = \{defg, abfg, abcg, abcd\}.$$

设 ε 是 V 上的一个超图. 一个顶点 $u \in V$, 被称为 ε 的一个孤立顶点, 如果 u 仅属于 ε 的一条边. 一条边 $e \in \varepsilon$ 称为 ε 的一个耳朵, 如果存在另外一条边 $e' \in \varepsilon$ 使得在 $e \setminus e'$ 中每个顶点都是孤立的. 如果 e 的所有顶点都是孤立的, 则 e 也是一个耳朵.

例如:$\varepsilon = \{abc, cde, efa, ace\}$, b, d, f 都是 ε 的孤立顶点, abc, cde, efa 都是 ε 的耳朵.

信息科学家在研究数据库理论时引入了无圈超图的概念[7~10], 他们使用称之为 Graham 约简的来界定一个超图是否是无圈的. 我们现在概括成下面的公理.

无圈公理 (acyclic-axiom) 一个超图 ε 可以通过重复使用下面两种运算化为空集:

(a_1) 若 x 是孤立顶点, 去掉 x.

(a_2) 若 $e_i \subseteq e_j, i \neq j$, 去掉 e_i.

一个等价的表述:ε 可以通过重复移去耳朵变为空集.

满足无圈公理的超图被称为无圈 (acyclic) 超图, 否则称为有圈的 (cyclic).

一个无圈超图的部分超图可能是有圈的, 例如 $\varepsilon = \{abc, cde, efa, ace\}$ 是无圈的, 如图 1.1 所示. 而 $\varepsilon' = \{abc, cde, efa\}$ 是 ε 的部分超图, 但 ε' 是有圈的.

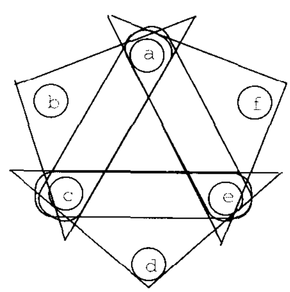

图 1.1

超图的这种性质是普通图所没有的.

一个超图 ε 的 Graham 约简是由 ε 通过重复运算 (a_1) 和 (a_2) 直到不能所得的超图, 或者通过重复移除耳朵直到不能移除所得到的超图.

定理 1.1 一个不少于两条边的无圈超图至少有两个耳朵.

证明 对边数进行归纳. 易见对只有两条边的无圈超图, 定理成立. 假设边数为 $k \geqslant 2$, 定理成立. 现在设 ε 有 $k+1$ 条边, 且是无圈的. 因此, 它有一个耳朵, 记为 e. 设 $e' \in \varepsilon$ 使得 $e \setminus e'$ 中每个顶点是孤立的. 令 $\varepsilon' = \varepsilon - e$, 则 ε' 有 k 条边. 由定义, ε' 是无圈的, 因为由 ε' 重复去耳朵可变空. 由归纳假设 ε' 有两个耳朵. 因为从 ε 移去耳朵 e 时, 除 e' 外, 不会产生新的耳朵, 所以 ε' 中的两个耳朵, 至少有一个是 ε 的耳朵. 完成了证明.

定义 1.5 我们说超图 ε 有移动交性质, 如果存在 ε 的边排序 $e_1, e_2, \cdots, e_m$, 使得对任意 i, $2 \leqslant i \leqslant m$, 都存在 $j_i < i$ 使 $e_i \cap (e_i \cup e_2 \cup \cdots \cup e_{i-1}) \subseteq e_{j_i}$, 即 e_i 与它前面边的并集的交被包含在这些边之一中.

由定义, 我们有下面的定理.

定理 1.2 一个超图 ε 是无圈的, 当且仅当它有移动交性质.

证明由定义直接得到, 这里略.

定义 1.6 设 ε 是一个超图, ε 的连接树 T 是一棵树, T 的顶点为 ε 的边, 它具有下述性质:

(1) 每条边 (e_i, e_j) 由集合 $e_i \cap e_j$ 标号.

(2) 对每一对 $e_k, e_l, e_k \cap e_l$ 中每个顶点 u, 在 T 中沿 e_k 和 e_l 间的唯一路上的每条边的标号包含 u.

例子: 设 $\varepsilon = \{abcd, bcde, cdf, bceg, uvw, uvx, uwy\}$. 图 1.2 所示树是 ε 的一个连接树.

定理 1.3[7] 一个超图是无圈的, 当且仅当 ε 有一个连接树.

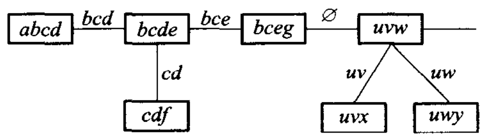

图 1.2 超图的连接树

证明 假设 ε 有一个连接树 T, 则 T 至少有两个叶子. 设 e_1 是 T 的一个叶子, e_2 是 e_1 在 T 中的邻点. 因为, 对任何 $e \neq e_1, e_2$, e_1 和 e 间在 T 中的路, 都包含边 (e_1, e_2), 所以 $e_1 \cap e \subseteq e_1 \cap e_2$, 故 e_1 是 ε 的一个耳朵. $T - e_1$ 是 $\varepsilon \setminus \{e_1\}$ 的连接树, 这个树的一个叶子对应 $\varepsilon \setminus \{e_1\}$ 的一个耳朵, 清除这个耳朵, 一直到 ε 空为止. 因此, ε 是无圈的.

现在假设 ε 是无圈的, 我们要证明 ε 有一个连接树. 在边数上进行归纳. 当 $|\varepsilon| = 2$ 时, 结论显然成立. 假设当 $|\varepsilon'| < m$ 时, 结论成立, 并设 $|\varepsilon| = m \geqslant 3$. 因为 ε 是无圈的, ε 有一个耳朵 e_1. 令 $\varepsilon' = \varepsilon \setminus \{e_1\}$ 和 $e_2 \in \varepsilon'$ 使得 $e_1 \setminus e_2$ 中每个顶点是孤立的. 由归纳假设, ε' 有一个连接树 T'. 增加一个顶点 e_1 和边 (e_1, e_2) 到 T', 形成树 T, (e_1, e_2) 的标号为 $e_1 \cap e_2$. 我们证明 T 是 ε 的一个连接树. 对任意 e_i, e_j, 如果 $e_i \neq e_1$ 和 $e_j \neq e_1$ 同时成立, 则 e_i 和 e_j 在 T 中的路也是 T' 中的唯一路. 因为 T' 是 ε' 的连接树, 所以沿 e_i 和 e_j 间的唯一路上的每条边的标号包含 $e_i \cap e_j$. 因而仅需证明, 对任意 $e_i \in \varepsilon'$, e_1 和 e_i 间在 T 中的唯一路上每条边的标号包含 $e_1 \cap e_i$ 的每个顶点. 因为 $e_1 \setminus e_2$ 中的每个顶点是孤立的. $e_1 \cap e_i \subseteq e_2 \cap e_i$. 设 $u \in e_1 \cap e_i$, 则 $u \in e_2 \cap e_i$. 由归纳假设, 沿 T' 中 e_2 和 e_i 间的唯一路上每条边的标号包含 u, 边 (e_1, e_2) 包含 u. $P' + (e_i, e_2)$ 是 e_i 和 e_1 间在 T 中唯一路. 因此 T 是 ε 的一个连接树. 完成了证明.

例 1.1 $\varepsilon = \{abc, cde, efa, ace\}$, 易见 ε 是一个无圈超图, 它的连接树由图 1.3 所示. $abc \cap efa = a$ 在 T 中 abc 和 efa 间的唯一路上每条弧的标号包含 a.

例 1.2 $\varepsilon = \{abcd, bcde, cdef, cdh, defg\}$, $abcd \cap cdh = cd$, 沿 T 中 $abcd$ 和 cdh 间的唯一路上每条边的标号包含 $\{c, d\}$, 如图 1.4

所示.

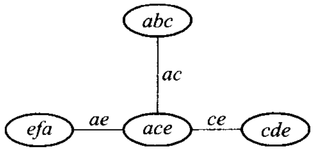

图 1.3　超图的连接图 T

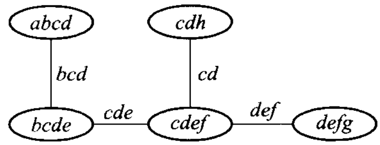

图 1.4　连接图 T

第二章　关系数据库

数据库在信息存储、信息传输、信息处理等领域起中心角色的作用. 数据库理论是信息科学的重要组成部分, 数据库技术在信息产业中占重要地位. 关系数据库是一类数据库模型. 无圈超图的概念即超图的无圈公理, 是信息科学家在关系数据库研究中引进的. 为了便于读者了解本书所涉及的超图内容与数据库之间的渊源, 本章简要地介绍关系数据库的有关概念及理论. 读者要对这方面的内容作进一步的了解, 可参阅文献 [7~10].

设 $\Omega = \{A_1, A_2, \cdots, A_n\}$ 是属性集合,$\mathrm{Dom}(A_i)$ 表示属性 A_i 的定义域, 即 A_i 的所有可能取值的集合. Ω 上的一个关系, 记为 $R[\Omega]$, 是这些定义域的笛卡儿积的一个子集合.

令 $D_i = \mathrm{Dom}(A_i), i = 1, 2, \cdots, n$, 它们的笛卡儿积记为 $D_1 \times D_2 \times \cdots \times D_n$, $R[\Omega] \subseteq D_1 \times D_2 \times \cdots \times D_n$. 我们称 n 为关系 $R[\Omega]$ 的度. 一个关系就是一个表, 每个行称为一个 tuple, 每列对应一个属性. 一个关系的属性集合称为关系图式. $D_1 \times D_2 \times \cdots \times D_n$ 是所有 n-tuple$(a_1, a_2, \cdots, a_n)$ 的集合, 其中 $a_i \in D_i$. 例如, 若 $n=2$, $D_1 = \{0, 1\}, D_2 = \{a, b, c\}$, 则 $D_1 \times D_2$ 是 $\{(0,a), (0,b), (0,c), (1,a), (1,b), (1,c)\}$. $\{(0,a), (0,c), (1,b)\}$ 是一个关系. 这个关系写成表的形式为:

A_1	A_2
0	a
0	c
1	b

Tuple$(a_1, a_2, \cdots, a_n)$ 有 n 个分量, 第 i 个分量是 a_i. 通常我们用 $a_1a_2\cdots a_n$ 表示 $(a_1, a_2, \cdots, a_n)$, 或简写为 a.

设 Ω 是属性集合, $\Omega_i \subseteq \Omega, i = 1, 2, \cdots, m$. 如果 $\forall i \in \{1, 2, \cdots, m\}$, $\Omega_i \neq \varnothing$ 且 $\cup_{i=1}^{m}\Omega_i = \Omega$, 则我们称 $D = \{\Omega_1, \Omega_2, \cdots, \Omega_n\}$ 是 Ω 上的一个数据库图式. 若 R_i 是 Ω_i 上的一个关系, 我们称 $R = \{R_1, R_2, \cdots, R_m\}$ 是在 D 上的数据库.

2.1 关系运算和算子

1. **并** 关系 R 和 S 的并, 记为 $R \cup S$, 是在 R 中或 S 中或同时在 R 和 S 中的 tuple 的集合, R 与 S 需有相同的度.

2. **差** 关系 R 和 S 的差, 记为 $R - S$, 是在 R 而不在 S 中的 tuple 的集合. R 和 S 要有相同的度.

3. **笛卡儿积** 设 R 和 S 是度分别为 k_1 和 k_2 的两个关系, 则 R 和 S 的笛卡儿积 $R \times S$ 是前 k_1 个分量为 R 中的一个 tuple 和后 k_2 个分量为 S 的一个 tuple 而形成的 $(k_1 + k_2)$-tuple 的集合.

4. **射影** 对一个关系 R, 移去某些分量和重新排列剩下的分量. 如果 R 是一个 k 度的关系, 令 $\Pi_{i_1, i_2, \cdots, i_m}(R)$ 表示 R 到分量指标 $i_1, i_2, \cdots, i_m$ 的射影, 是 m-tuples $a_1a_2\cdots a_m$ 的集合使得存在 R 中的一个 k-tuple $b_1b_2\cdots b_k$ 有 $a_j = b_{i_j}$, 其中 $j = 1, 2, \cdots, m$, i_j 是 1 到 k 中的一个数. 例如 $\Pi_{3,1}(R)$ 是取 R 中每个 tuple, 使该 tuple 的第 3 分量和第 1 分量依此顺序形成的 2-tuple 来运算的集合. 如果用属性命名关系的列, 我们可以使同样的属性命名射影关系中的列. 例如, $R[A, B, C, D]$, $\Pi_{C,A}$ 由 C 命名第一列, A 命名第二列.

5. **选择** 设 F 是一个涉及下列算子的公式:

(1) 常数, 分枝数;

(2) 算术比较算子 $<, =, >, \leqslant, \neq, \geqslant$;

(3) 逻辑算子 $\wedge, \vee, -$.

则 $\sigma_F(R)$ 是 R 中的具有下述性质的 tuple t 的集合: 当对所有 i, 在 F 中的任意出现的 i 都用 t 的第 i 个分量代替, 公式 F 变为真. 例如, $\sigma_{1=Smith \vee 1=Jones}(R)$ 表示 R 中的第一个分量是 Smith 或 Jones 的所有 tuples 的集合. 如果一个关系的列被命名了, 则选择的公式

可用命名代替数.

6. **交** $R\cap S$ 表示 $R-(R-S)$.

例 2.1 设 R,S 如图 2.1,

$R=$

A	B	C
a	b	c
d	a	f
c	b	d

$S=$

D	E	F
b	g	a
d	a	f

图 2.1

则 (a) $R\cup S$,(b)$R-S$,(c)$R\times S$,(d)$\Pi_{A\cdot C}(R)$,(e)$\sigma_{B=b}(R)$ 的运算如图 2.2 所示.

a	b	c
d	a	f
c	b	d
b	g	a

(a)

a	b	c
c	b	d

(b)

A	B	C	D	E	F
a	b	c	b	g	a
a	b	c	d	a	f
d	a	f	b	g	a
d	a	f	d	a	f
c	b	d	b	g	a
c	b	d	d	a	f

(c)

A	C
a	c
d	f
c	d

(d)

A	B	C
a	b	c
c	b	d

(e)

图 2.2 某些运算结果

7. **商**(quotient) 令 R 和 S 分别有度为 r 和 s 的两个关系, $r>s,S\neq\varnothing$. 商 $R\div S$ 是具有下述性质的 $(r-s)$-tuple t 的集合: 对所有在 S 中的 s-tuple u,tu 在 R 中.

易见, $R\div S=\Pi_{1,2,\cdots,r-s}(R)-\Pi_{1,2,\cdots,r-s}[(\Pi_{1,2,\cdots,r-s}(R)\times S)-R]$, 图 2.3 是一个商运算 $R\div S$ 的例子.

a	b	c	d
a	b	e	f
b	c	e	f
e	d	c	d
e	d	e	f
a	b	d	e

(a)

c	d
e	f

(b)

a	b
e	d

(c)

图 2.3 商运算的例子

(a) R (b) S (c) $R\div S$

8. **连接** (join) 设 θ 是一个算术比较算子 ($=, <$ 等), R 和 S 的 θ 连接, 记为 $R \bowtie_{i\theta j} S$, 是 $\sigma_{i\theta(r+j)}(R \times S)$ 的简写, 其中 r 为 R 的度. 即 R 和 S 的 θ 连接是 $R \times S$ 的那些 tuple 的集合, 使得 R 的第 i 分量由 θ 联系到 S 的第 j 分量. 如果 θ 是 "=" 运算, 通常称为等连接.

例 2.2 R 和 S 如图 2.4(a) 和 (b) 所示, $R \bowtie_{B<D} S$ 由 2.4(c) 给出.

A	B	C
1	2	3
4	5	6
7	8	9

(a)

D	E
3	1
6	2

(b)

A	B	C	D	E
1	2	3	3	1
1	2	3	6	2
4	5	6	6	2

(c)

图 2.4

(a) 关系 R (b) 关系 S (c)$R \bowtie_{B<D} S$

9. **自然连接** (nature join) 自然连接记为 $R \bowtie S$. 仅当 R 和 S 的列由属性命名时可用.

如下计算 $R \bowtie S$.

(1) 计算 $R \times S$.

(2) 对每个同在 R 的一个列和 S 中的一个列的属性 A, 从 $R \times S$ 中选出那些 tuple, 它们在 $R \cdot A$ 和 $S \cdot A$ 一致, 其中 $R \cdot A$ 是 $R \times S$ 中对应 R 的列 A 的命名, $S \cdot A$ 类似的定义.

(3) 对上述的每个属性 A, 除去列 $S \cdot A$.

如果 $A_1, A_2, \cdots, A_k$ 是所有同在 R 和 S 中使用的属性名. $R \bowtie S$ 是 $\Pi_{i_1, i_2, \cdots, i_m} \sigma_{R \cdot A_1 = S \cdot A_1 \wedge \cdots \wedge R \cdot A_k = S \cdot A_k}(R \times S)$, 其中, $i_1, i_2, \cdots, i_m$ 是除去 $S \cdot A_1, \cdots, S \cdot A_k$ 后所有列依次的目录.

例 2.3 设 R 和 S 是图 2.5(a) 和 (b) 所示的关系, $R \bowtie S$ 表示 $\Pi_{A, R \cdot B, R \cdot C, D} \sigma_{R \cdot B = S \cdot B \wedge R \cdot C = S \cdot C}(R \times S)$.

A	B	C
a	b	c
d	b	c
b	b	f
c	a	d

(a)

B	C	D
b	c	d
b	c	e
a	d	b

(b)

A	B	C	D
a	b	c	d
a	b	c	e
d	b	c	d
d	b	c	e
c	a	d	b

(c)

图 2.5 自然联接

(a) R (b) S (c) $R \bowtie S$

10. **半连接** 关系 R 和 S 的半连接, 记为 $R \ltimes S$, 是 $\Pi_R(R \bowtie S)$. $R \rtimes S$ 是 $\Pi_S(R \bowtie S)$.

例 2.4 R 和 S 分别为图 2.5(a) 和 2.5(b) 所示, 则 $R \ltimes S$ 和 $R \rtimes S$ 分别表示在图 2.6(a) 和 2.6(b) 中.

A	B	C
a	b	c
d	b	c
c	a	d

(a)

B	C	D
b	c	d
b	c	e
a	d	b

(b)

图 2.6 半连接

(a)$R \ltimes S$ (b)$R \rtimes S$

半连接是不对称的. 我们用 Ω_R 和 Ω_S 分别表示关系 R 和关系 S 的属性集合. $R \ltimes S$ 是射影 S 到属性集合 $\Omega_R \cap \Omega_S$ 上, 并且从 R 中去掉所有使 $t[\Omega_R \cap \Omega_S]$ 不在这个射影中的 tuple t.

定理 2.1[8] $(R \ltimes S) \bowtie S = R \bowtie S$.

证明 因为 $R \ltimes S = \Pi_R(R \bowtie S)$, 我们有 $R \ltimes S \subseteq R$, 于是 $(R \ltimes S) \bowtie S \subseteq R \bowtie S$. 反过来, 假设 t 是 $R \bowtie S$ 中的一个 tuple, 则 $u = t[R]$ 一定在 R 中, $v = t[S]$ 一定在 S 中, 这就得到 u 在 $R \ltimes S$ 中, 因而 t 在 $(R \ltimes S) \bowtie S$ 中. 于是 $R \bowtie S \subseteq (R \ltimes S) \bowtie S$. 完成证明.

例 2.5 $R[A, B]$ 和 $S[B, C]$ 分别如图 2.7(a) 和 2.7(b) 所示: