



数据库系统概论

An Introduction to Database System

第二章 关系数据库（续）

中国人民大学信息学院

第二章 关系数据库



2.1 关系模型概述

2.2 关系数据结构

2.3 关系的完整性

2.4 关系代数

2.5 关系演算

2.6 小结

2.4 关系代数



- ❖ 概述
- ❖ 传统的集合运算
- ❖ 专门的关系运算

概述



表 2.4 关系代数运算符

运算符		含义	运算符		含义
集合运算符	\cup	并	比较运算符	$>$	大于
	$-$	差		\geq	大于等于
	\cap	交		$<$	小于
	\times	笛卡尔积		\leq	小于等于
				$=$	等于
				$<>$	不等于

概述 (续)



表 2.4 关系代数运算符 (续)

运算符	含义		运算符	含义	
专门的关系运算符	σ	选择	逻辑运算符	\neg	非
	π	投影		\wedge	与
	\bowtie	连接		\vee	或
	\div	除			

2.4 关系代数



- ❖ 概述
- ❖ 传统的集合运算
- ❖ 专门的关系运算

1. 并 (Union)



❖ R 和 S

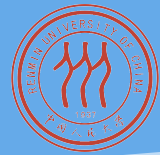
- 具有相同的目 n (即两个关系都有 n 个属性)
- 相应的属性取自同一个域

❖ $R \cup S$

- 仍为 n 目关系, 由属于 R 或属于 S 的元组组成

$$R \cup S = \{ t | t \in R \vee t \in S \}$$

并 (续)



R

<i>A</i>	<i>B</i>	<i>C</i>
<i>a</i> ₁	<i>b</i> ₁	<i>c</i> ₁
<i>a</i> ₁	<i>b</i> ₂	<i>c</i> ₂
<i>a</i> ₂	<i>b</i> ₂	<i>c</i> ₁

S

<i>A</i>	<i>B</i>	<i>C</i>
<i>a</i> ₁	<i>b</i> ₂	<i>c</i> ₂
<i>a</i> ₁	<i>b</i> ₃	<i>c</i> ₂
<i>a</i> ₂	<i>b</i> ₂	<i>c</i> ₁

RUS

<i>A</i>	<i>B</i>	<i>C</i>
<i>a</i> ₁	<i>b</i> ₁	<i>c</i> ₁
<i>a</i> ₁	<i>b</i> ₂	<i>c</i> ₂
<i>a</i> ₂	<i>b</i> ₂	<i>c</i> ₁
<i>a</i> ₁	<i>b</i> ₃	<i>c</i> ₂

2. 差 (Difference)



❖ R 和 S

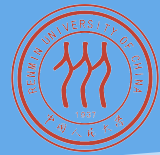
- 具有相同的目 n
- 相应的属性取自同一个域

❖ $R - S$

- 仍为 n 目关系，由属于 R 而不属于 S 的所有元组组成

$$R - S = \{ t | t \in R \wedge t \notin S \}$$

差 (续)



R

<i>A</i>	<i>B</i>	<i>C</i>
a_1	b_1	c_1
a_1	b_2	c_2
a_2	b_2	c_1

S

<i>A</i>	<i>B</i>	<i>C</i>
a_1	b_2	c_2
a_1	b_3	c_2
a_2	b_2	c_1

R-S

<i>A</i>	<i>B</i>	<i>C</i>
a_1	b_1	c_1

3. 交 (Intersection)



❖ R 和 S

- 具有相同的目 n
- 相应的属性取自同一个域

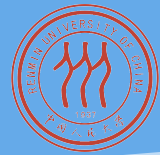
❖ $R \cap S$

- 仍为 n 目关系，由既属于 R 又属于 S 的元组组成

$$R \cap S = \{ t | t \in R \wedge t \in S \}$$

$$R \cap S = R - (R - S)$$

交 (续)



R

A	B	C
a_1	b_1	c_1
a_1	b_2	c_2
a_2	b_2	c_1

S

A	B	C
a_1	b_2	c_2
a_1	b_3	c_2
a_2	b_2	c_1

$R \cap S$

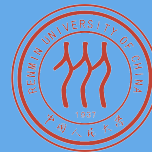
A	B	C
a_1	b_2	c_2
a_2	b_2	c_1

4. 笛卡尔积 (Cartesian Product)



- ❖ 严格地讲应该是广义的笛卡尔积 (Extended Cartesian Product)
- ❖ R : n 目关系, k_1 个元组
- ❖ S : m 目关系, k_2 个元组
- ❖ $R \times S$
 - 列: ($n+m$) 列元组的集合
 - 元组的前 n 列是关系 R 的一个元组
 - 后 m 列是关系 S 的一个元组
 - 行: $k_1 \times k_2$ 个元组
 - $R \times S = \{t_r t_s \mid t_r \in R \wedge t_s \in S\}$

交 (续)



R		
A	B	C
a_1	b_1	c_1
a_1	b_2	c_2
a_2	b_2	c_1

S		
A	B	C
a_1	b_2	c_2
a_1	b_3	c_2
a_2	b_2	c_1

$R \times S$					
$R.A$	$R.B$	$R.C$	$S.A$	$S.B$	$S.C$
a_1	b_1	c_1	a_1	b_2	c_2
a_1	b_1	c_1	a_1	b_3	c_2
a_1	b_1	c_1	a_2	b_2	c_1
a_1	b_2	c_2	a_1	b_2	c_2
a_1	b_2	c_2	a_1	b_3	c_2
a_1	b_2	c_2	a_2	b_2	c_1
a_2	b_2	c_1	a_1	b_2	c_2
a_2	b_2	c_1	a_1	b_3	c_2
a_2	b_2	c_1	a_2	b_2	c_1

2.4 关系代数



- ❖ 概述
- ❖ 传统的集合运算
- ❖ 专门的关系运算

2.4.2 专门的关系运算



先引入几个记号

(1) R , $t \in R$, $t[A_i]$

设关系模式为 $R(A_1, A_2, \dots, A_n)$

它的一个关系设为 R

$t \in R$ 表示 t 是 R 的一个元组

$t[A_i]$ 则表示元组 t 中相应于属性 A_i 的一个分量

专门的关系运算（续）



(2) A , $t[A]$, \overline{A}

若 $A = \{A_{i_1}, A_{i_2}, \dots, A_{i_k}\}$, 其中 $A_{i_1}, A_{i_2}, \dots, A_{i_k}$ 是 A_1, A_2, \dots, A_n 中的一部分, 则 A 称为属性列或属性组。

$t[A] = (t[A_{i_1}], t[A_{i_2}], \dots, t[A_{i_k}])$ 表示元组 t 在属性列 A 上诸分量的集合。

\overline{A} 则表示 $\{A_1, A_2, \dots, A_n\}$ 中去掉 $\{A_{i_1}, A_{i_2}, \dots, A_{i_k}\}$ 后剩余的属性组。

专门的关系运算（续）



(3) $\widehat{t_r t_s}$

R 为 n 目关系， S 为 m 目关系。

$t_r \in R$ ， $t_s \in S$ ， $\widehat{t_r t_s}$ 称为元组的连接。

$\widehat{t_r t_s}$ 是一个 $n + m$ 列的元组，前 n 个分量为 R 中的一个 n 元组，后 m 个分量为 S 中的一个 m 元组。

专门的关系运算（续）



（4）象集 Z_x

给定一个关系 R （ X ， Z ）， X 和 Z 为属性组。

当 $t[X]=x$ 时， x 在 R 中的象集（Images Set）为：

$$Z_x = \{t[Z] | t \in R, t[X]=x\}$$

它表示 R 中属性组 X 上值为 x 的诸元组在 Z 上分量的集合

专门的关系运算（续）



R

x_1	Z_1
x_1	Z_2
x_1	Z_3
x_2	Z_2
x_2	Z_3
x_3	Z_1
x_3	Z_3

象集举例

❖ x_1 在 R 中的象集

$$Z_{x_1} = \{Z_1, Z_2, Z_3\}$$

,

❖ x_2 在 R 中的象集

$$Z_{x_2} = \{Z_2, Z_3\},$$

❖ x_3 在 R 中的象集

专门的关系运算（续）



- ❖ 选择
- ❖ 投影
- ❖ 连接
- ❖ 除

专门的关系运算 (续)



4) 学生 - 课程数据库 :

学生关系 Student、课程关系 Course 和选修关系 SC

Student

学号 Sno	姓名 Sname	性别 Ssex	年龄 Sage	所在系 Sdept
200215121	李勇	男	20	CS
200215122	刘晨	女	19	IS
200215123	王敏	女	18	MA
200215125	张立	男	19	IS

专门的关系运算（续）

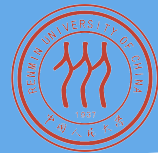


Course

课程号 Cno	课程名 Cname	先行课 Cpno	学分 Ccredit
1	数据库	5	4
2	数学		2
3	信息系统	1	4
4	操作系统	6	3
5	数据结构	7	4
6	数据处理		2
7	PASCAL 语言	6	4

(b)

专门的关系运算（续）



SC

学号 Sno	课程号 Cno	成绩 Grade
200215121	1	92
200215121	2	85
200215121	3	88
200215122	2	90
200215122	3	80

(c)

1. 选择 (Selection)



❖ 1) 选择又称为限制 (Restriction)

❖ 2) 选择运算符的含义

- 在关系 R 中选择满足给定条件的诸元组

$$\sigma_F(R) = \{t | t \in R \wedge F(t) = \text{'真'}\}$$

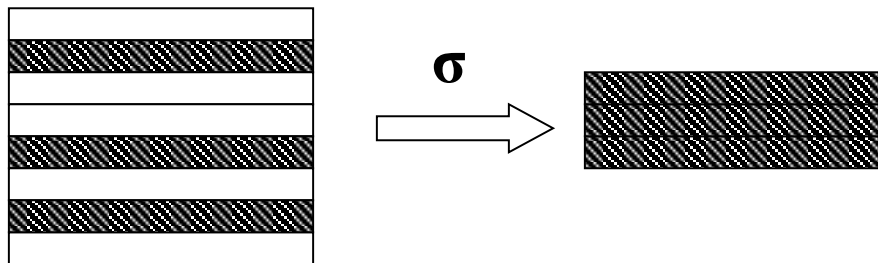
- F : 选择条件, 是一个逻辑表达式, 基本形式为:

$$X_1 \theta Y_1$$

选择（续）



- ❖ 3) 选择运算是从关系 R 中选取使逻辑表达式 F 为真的元组，是从行的角度进行的运算



选择（续）



[例 1] 查询信息系（IS 系）全体学生

$\sigma_{Sdept = 'IS'}(Student)$

或 $\sigma_{5 = 'IS'}(Student)$

结果：

Sno	Sname	Ssex	Sage	Sdept
200215122	刘晨	女	19	IS
200215125	张立	男	19	IS

选择（续）



[例 2] 查询年龄小于 20 岁的学生

$\sigma_{\text{Sage} < 20}(\text{Student})$

或 $\sigma_{4 < 20}(\text{Student})$

结果：

Sno	Sname	Ssex	Sage	Sdept
200215122	刘晨	女	19	IS
200215123	王敏	女	18	MA
200215125	张立	男	19	IS

2. 投影 (Projection)



❖ 1) 投影运算符的含义

- 从 R 中选择出若干属性列组成新的关系

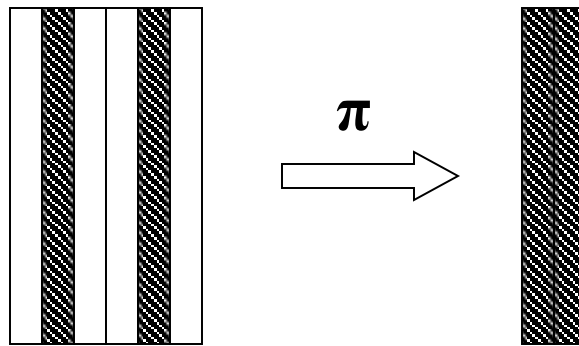
$$\pi_A(R) = \{ t[A] \mid t \in R \}$$

A : R 中的属性列

2. 投影 (Projection)



❖ 2) 投影操作主要是从列的角度进行运算



- 但投影之后不仅取消了原关系中的某些列，而且还可能取消某些元组（避免重复行）

投影（续）



❖ [例 3] 查询学生的姓名和所在系

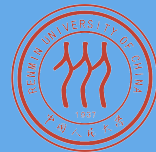
即求 Student 关系上学生姓名和所在系两个属性上的投影

$\pi_{\text{Sname}, \text{Sdept}}(\text{Student})$

或 $\pi_{2, 5}(\text{Student})$

结果：

投影（续）



Sname	Sdept
李勇	CS
刘晨	IS
王敏	MA
张立	IS

投影（续）



[例 4] 查询学生关系 Student 中都有哪些系

$\pi_{\text{Sdept}}(\text{Student})$

结果：

Sdept
CS
IS
MA

3. 连接 (Join)



❖ 1) 连接也称为 θ 连接

❖ 2) 连接运算的含义

从两个关系的笛卡尔积中选取属性间满足一定条件的元组

$$R \underset{A\theta B}{\bowtie} S = \{ \widehat{t_r t_s} \mid t_r \in R \wedge t_s \in S \wedge t_r[A] \theta t_s[B] \}$$

➤ A 和 B : 分别为 R 和 S 上度数相等且可比的属性组

➤ θ : 比较运算符

- 连接运算从 R 和 S 的广义笛卡尔积 $R \times S$ 中选取 (R 关系) 在 A 属性组上的值与 (S 关系) 在 B 属性组上值满足比较关系 θ 的元组

连接（续）



❖ 3) 两类常用连接运算

■ 等值连接（equijoin）

➤ 什么是等值连接

θ 为 “=” 的连接运算称为等值连接

➤ 等值连接的含义

从关系 R 与 S 的广义笛卡尔积中选取 A 、 B 属性值相等的那些元组，即等值连接为：

$$R \bowtie_{A=B} S = \{ \widehat{t_r t_s} \mid t_r \in R \wedge t_s \in S \wedge t_r[A] = t_s[B] \}$$

连接（续）



■ 自然连接（Natural join）

- 自然连接是一种特殊的等值连接
 - 两个关系中进行比较的分量必须是相同的属性组
 - 在结果中把重复的属性列去掉
- 自然连接的含义

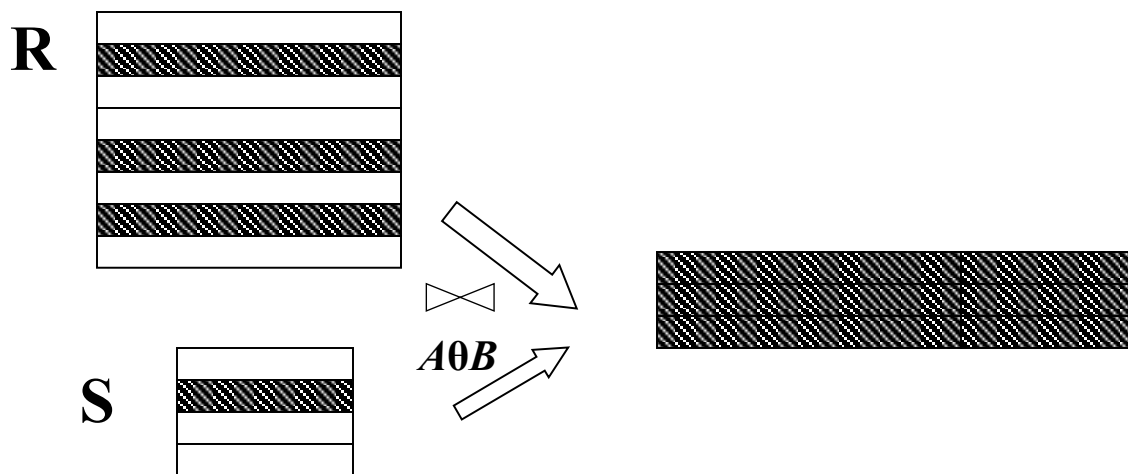
R 和 S 具有相同的属性组 B

$$R \bowtie S = \{ \widehat{t_r t_s} \mid t_r \in R \wedge t_s \in S \wedge t_r[B] = t_s[B] \}$$

连接（续）



❖ 4) 一般的连接操作是从行的角度进行运算。



自然连接还需要取消重复列，所以是同时从行和列的角度进行运算。

连接（续）



❖ [例 5] 关系 R 和关系 S 如下所示：

R		
A	B	C
a_1	b_1	5
a_1	b_2	6
a_2	b_3	8
a_2	b_4	12

S	
B	E
b_1	3
b_2	7
b_3	10
b_3	2
b_5	2

连接（续）



一般连接 $R \bowtie_{C < E} S$ 的结果如下：

$R \bowtie_{C < E} S$

A	$R.B$	C	$S.B$	E
a_1	b_1	5	b_2	7
a_1	b_1	5	b_3	10
a_1	b_2	6	b_2	7
a_1	b_2	6	b_3	10
a_2	b_3	8	b_3	10

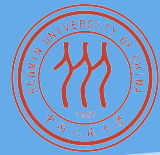
连接（续）



等值连接 $R \bowtie_{R.B=S.B} S$ 的结果如下：

A	$R.B$	C	$S.B$	E
a_1	b_1	5	b_1	3
a_1	b_2	6	b_2	7
a_2	b_3	8	b_3	10
a_2	b_3	8	b_3	2

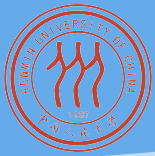
连接（续）



自然连接 $R \bowtie S$ 的结果如下：

A	B	C	E
a_1	b_1	5	3
a_1	b_2	6	7
a_2	b_3	8	10
a_2	b_3	8	2

连接（续）



❖ 外连接

- 如果把舍弃的元组也保存在结果关系中，而在其他属性上填空值 (Null)，这种连接就叫做外连接（OUTER JOIN）。

❖ 左外连接

- 如果只把左边关系 R 中要舍弃的元组保留就叫做左外连接 (LEFT OUTER JOIN 或 LEFT JOIN)

❖ 右外连接

- 如果只把右边关系 S 中要舍弃的元组保留就叫做右外连接 (RIGHT OUTER JOIN 或 RIGHT JOIN)。

连接（续）



下图是例 5 中关系 R 和关系 S 的外连接

A	B	C	E
a_1	b_1	5	3
a_1	b_2	6	7
a_2	b_3	8	10
a_2	b_3	8	2
a_2	b_4	12	NULL
NULL	b_5	NULL	2

(a) 外连接

连接（续）



图 (b) 是例 5 中关系 R 和关系 S 的左外连接，图 (c) 是右外连接

A	B	C	E
a_1	b_1	5	3
a_1	b_2	6	7
a_2	b_3	8	10
a_2	b_3	8	2
a_2	b_4	12	NULL

(b) 左外连接

A	B	C	E
a_1	b_1	5	3
a_1	b_2	6	7
a_2	b_3	8	10
a_2	b_3	8	2
NULL	b_5	NULL	2

(c) 右外连接

4. 除 (Division)



给定关系 $R(X, Y)$ 和 $S(Y, Z)$ ，其中 X, Y, Z 为属性组。

R 中的 Y 与 S 中的 Y 可以有不同的属性名，但必须出自相同的域集。

R 与 S 的除运算得到一个新的关系 $P(X)$ ，

P 是 R 中满足下列条件的元组在 X 属性列上的投影：

元组在 X 上分量值 x 的象集 Y_x 包含 S 在 Y 上投影的集合，记作：

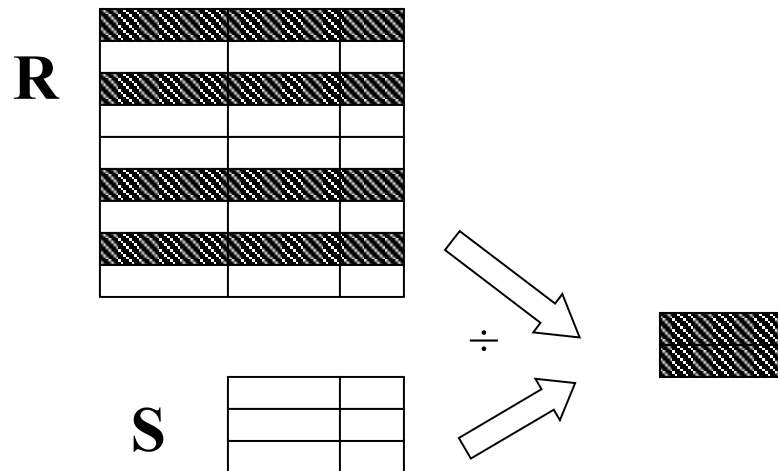
$$R \div S = \{t_r[X] \mid t_r \in R \wedge \pi_Y(S) \subseteq Y_x\}$$

$$Y_x: x \text{ 在 } R \text{ 中的象集, } x = t_r[X]$$

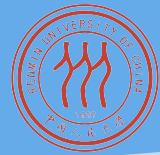
除 (续)



❖ 2) 除操作是同时从行和列角度进行运算



除 (续)



[例 6] 设关系 R 、 S 分别为下图的 (a) 和 (b)， $R \div S$ 的结果为图

(c)

A	B	C
a_1	b_1	c_2
a_2	b_3	c_7
a_3	b_4	c_6
a_1	b_2	c_3
a_4	b_6	c_6
a_2	b_2	c_3
a_1	b_2	c_1

(a)

B	C	D
b_1	c_2	d_1
b_2	c_1	d_1
b_2	c_3	d_2

(b)

$R \div S$
A
a_1

(c)

分析



❖ 在关系 R 中， A 可以取四个值 $\{a_1, a_2, a_3, a_4\}$

a_1 的象集为 $\{(b_1, c_2), (b_2, c_3), (b_2, c_1)\}$

a_2 的象集为 $\{(b_3, c_7), (b_2, c_3)\}$

a_3 的象集为 $\{(b_4, c_6)\}$

a_4 的象集为 $\{(b_6, c_6)\}$

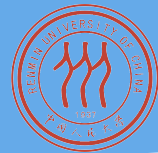
❖ S 在 (B, C) 上的投影为

$\{(b_1, c_2), (b_2, c_1), (b_2, c_3)\}$

❖ 只有 a_1 的象集包含了 S 在 (B, C) 属性组上的投影

所以 $R \div S = \{a_1\}$

5 . 综合举例



以学生 - 课程数据库为例 (P56)

[例 7] 查询至少选修 1 号课程和 3 号课程的学生号码

首先建立一个临时关系 K :

Cno
1
3

然后求: $\pi_{Sno, Cno}(SC) \div K$

综合举例（续）



❖ 例 7 续 $\pi_{\text{Sno,Cno}}(\text{SC})$

200215121 象集 {1, 2, 3}

200215122 象集 {2, 3}

$K=\{1, 3\}$

于是：

$\pi_{\text{Sno,Cno}}(\text{SC}) \div K = \{200215121\}$

Sno	Cno
200215121	1
200215121	2
200215121	3
200215122	2
200215122	3

综合举例（续）



[例 8] 查询选修了 2 号课程的学生们的学号。

$$\pi_{\text{Sno}} \left(\sigma_{\text{Cno}='2'} \left(\text{SC} \right) \right) \\ = \{ 200215121, 200215122 \}$$

综合举例（续）



[例 9] 查询至少选修了一门其直接先行课为 5 号课程的学生姓名

$\pi_{\text{Sname}}(\sigma_{\text{Cpno}='5'}(\text{Course} \bowtie \text{SC} \bowtie \text{Student}))$

或

$\pi_{\text{Sname}}(\sigma_{\text{Cpno}='5'}(\text{Course}) \bowtie \text{SC} \bowtie \pi_{\text{Sno}, \text{Sname}}(\text{Student}))$

或

$\pi_{\text{Sname}}(\pi_{\text{Sno}}(\sigma_{\text{Cpno}='5'}(\text{Course}) \bowtie \text{SC}) \bowtie \pi_{\text{Sno}, \text{Sname}}(\text{Student}))$

综合举例（续）



[例 10] 查询选修了全部课程的学生号码和姓名。

$$\pi_{Sno, Cno} (SC) \div \pi_{Cno} (Course) \bowtie \pi_{Sno, Sname} (Student)$$

小结



❖ 关系代数运算

- 关系代数运算

并、差、交、笛卡尔积、投影、选择、连接、除

- 基本运算

并、差、笛卡尔积、投影、选择

- 交、连接、除

可以用 5 种基本运算来表达

引进它们并不增加语言的能力，但可以简化表达

小结（续）



❖ 关系代数表达式

- 关系代数运算经有限次复合后形成的式子

❖ 典型关系代数语言

- ISBL （ Information System Base Language ）
 - 由 IBM United Kingdom 研究中心研制
 - 用于 PRTV （ Peterlee Relational Test Vehicle ） 实验系统