

数据库系统概论

An Introduction to Database System

第二章 关系数据库

中国人民大学信息学院

关系代数

(专门的集合运算)



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
201215121	李勇	男	20	CS
201215122	刘晨	女	19	CS
201215123	王敏	女	18	MA
201215125	张立	男	19	IS

(a)

学生-课程数据库

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

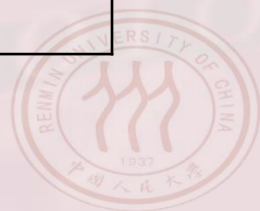
(b)

学生-课程数据库

SC

学号 Sno	课程号 Cno	成绩 Grade
201215121	1	92
201215121	2	85
201215121	3	88
201215122	2	90
201215122	3	80

(c)



专门的关系运算

1. 选择
2. 投影
3. 连接
4. 除运算

中国人民大学
数据库系统概论



1. 选择 (Selection)

❖ 选择又称为限制 (**Restriction**)

❖ 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, 取值为“真”或“假”

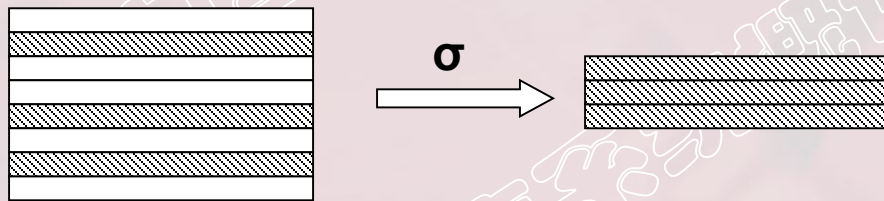
● 基本形式为: $X_1 \theta Y_1$

● θ 表示比较运算符, 它可以是 $>$, \geq , $<$, \leq , $=$ 或 \neq



选择（续）

❖ 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2.4] 查询信息系（IS系）全体学生。

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

结果：

Sno	Sname	Ssex	Sage	Sdept
201215125	张立	男	19	IS

选择（续）

[例2.5] 查询年龄小于20岁的学生。

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

结果:

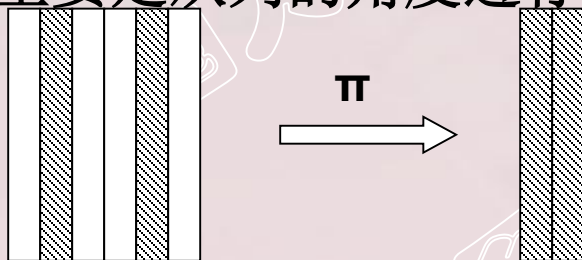
Sno	Sname	Ssex	Sage	Sdept
201215122	刘晨	女	19	IS
201215123	王敏	女	18	MA
201215125	张立	男	19	IS

2. 投影 (Projection)

- 从 R 中选择出若干属性列组成新的关系

$$\pi_A(R) = \{ t[A] \mid t \in R \}, \quad A: R \text{ 中的属性列}$$

- 投影操作主要是从列的角度进行运算



- 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2.6] 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

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

结果：

Sname	Sdept
李勇	CS
刘晨	CS
王敏	MA
张立	IS



投影（续）

[例2.7] 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)

❖ 连接也称为 θ 连接

❖ 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

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

- 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 属性组上的值满足比较关系 θ 的元组

连接（续）

❖ 两类常用连接运算

■ 等值连接（equijoin）

- θ 为 “=” 的连接运算称为等值连接
- 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

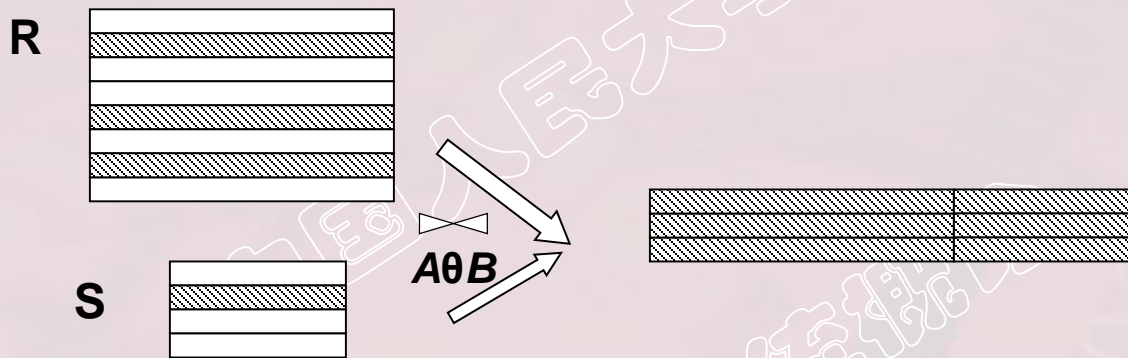
R 和 S 具有相同的属性组 B

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



连接（续）

❖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2.8]关系 R 和关系 S 如下所示：

R

A	B	C
a1	b1	5
a1	b2	6
a2	b3	8
a2	b4	12

S

B	E
b1	3
b2	7
b3	10
b3	2
b5	2



连接(续)

一般连接 $R \bowtie_{C < E} S$ 的结果如下:

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_4	2
b_5	2

$R \bowtie_{C < E} S$				
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
a1	b1	5	b1	3
a1	b2	6	b2	7
a2	b3	8	b3	10
a2	b3	8	b3	2

连接（续）

自然连接 $R \bowtie S$ 的结果如下：

A	B	C	E
a1	b1	5	3
a1	b2	6	7
a2	b3	8	10
a2	b3	8	2



连接（续）

❖ 悬浮元组（Dangling tuple）

- 两个关系 R 和 S 在做自然连接时，关系 R 中某些元组有可能在 S 中不存在公共属性上值相等的元组，从而造成 R 中这些元组在操作时被舍弃了，这些被舍弃的元组称为悬浮元组。



连接（续）

❖ 外连接（Outer Join）

- 如果把悬浮元组也保存在结果关系中，而在其他属性上填空值(Null)，就叫做外连接
- 左外连接(LEFT OUTER JOIN或LEFT JOIN)
 - 只保留左边关系 R 中的悬浮元组
- 右外连接(RIGHT OUTER JOIN或RIGHT JOIN)
 - 只保留右边关系 S 中的悬浮元组



连接（续）

下图是例2.8中关系***R***和关系***S***的外连接

A	B	C	E
a1	b1	5	3
a1	b2	6	7
a2	b3	8	10
a2	b3	8	2
a2	b4	12	NULL
NULL	b5	NULL	2

连接（续）

图(b)是例2.8中关系**R**和关系**S**的左外连接,图(c)是右外连接

A	B	C	E
a1	b1	5	3
a1	b2	6	7
a2	b3	8	10
a2	b3	8	2
a2	b4	12	NULL

图(b)

A	B	C	E
a1	b1	5	3
a1	b2	6	7
a2	b3	8	10
a2	b3	8	2
NULL	b5	NULL	2

图(c)



4. 除运算 (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 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：

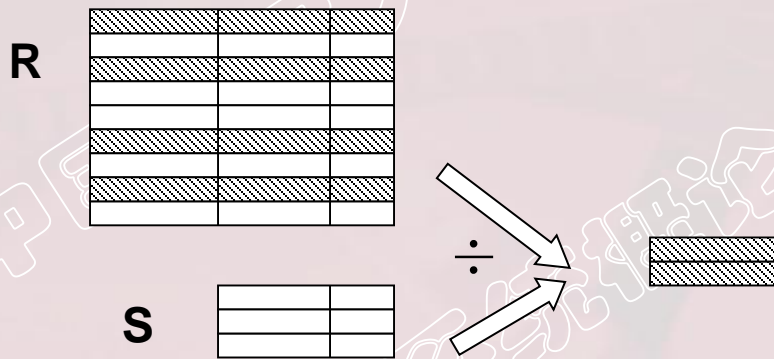
$$R \div S = \{t_r[X] \mid t_r \in R \wedge \pi_Y(S) \subseteq Y_x\}$$

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



除运算（续）

❖ 除操作是同时从行和列角度进行运算



除运算（续）

[例2.9] 设关系 R 、 S 分别为下图的(a)和(b),
 $R \div S$ 的结果为图(c)

R

A	B	C
a1	b1	c2
a2	b3	c7
a3	b4	c6
a1	b2	c3
a4	b6	c6
a2	b2	c3
a1	b2	c1

S

B	C	D
b1	c2	d1
b2	c1	d1
b2	c3	d2

$R \div S$

A
a1



除运算（续）

[例2.9] 设关系 R 、 S 分别为下图的(a)和(b),
 $R \div S$ 的结果为图(c)

R

A	B	C
a1	b1	c2
a1	b2	C1
a1	b2	C3
a2	b3	C7
a2	b2	C3
a3	b4	C6
a4	b6	c6

S

B	C	D
b1	c2	d1
b2	c1	d1
b2	c3	d2

$R \div S$

A
a1



除运算（续）

❖ 在关系R中，A可以取四个值{a1, a2, a3, a4}

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

$\{(b1, c2), (b2, c1), (b2, c3)\}$

❖ 只有 a_1 的象集包含了S在(B, C)属性组上的投影

所以 $R \div S = \{a_1\}$



