

# 数据库原理

## The Theory of Database System

### 第二章 关系数据库(续)



中国矿业大学计算机学院



中国矿业大学数据库原理精品课程

# 第二章 关系数据库

2.1 关系

2.2 关系代数

2.3 查询优化

2.4 关系系统

2.5 关系演算



## 2.2 关系代数

### ➤ 概述

- 传统的集合运算
- 专门的关系运算



# 概述

1. 关系代数
2. 运算的三要素
3. 关系代数运算的三个要素
4. 关系代数运算的分类
5. 表示记号



# 概述

## 1.关系代数

一种抽象的查询语言

用对关系的运算来表达查询



## 2. 关系代数运算的三个要素

运算对象：关系

运算结果：关系

运算符：四类



# 概述(续)

## —集合运算符

- 将关系看成元组的集合
- 运算是从关系的“水平”方向即行的角度来进行

## —专门的关系运算符

- 不仅涉及行而且涉及列

## —算术比较符

- 辅助专门的关系运算符进行操作

## —逻辑运算符

- 辅助专门的关系运算符进行操作



# 概述(续)

表1 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
		广义笛卡尔积		$=$	等于
				$\neq$	不等于





# 概述(续)

表2 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	$\sigma$	选择	逻辑运算符	$\neg$	非
	$\pi$	投影		$\wedge$	与
	$\bowtie$	连接		$\vee$	或
	$\div$	除			



## 4. 关系代数运算的分类

传统的集合运算

并、差、交、广义笛卡尔积

专门的关系运算

选择、投影、连接、除



## 5. 表示记号

$R, t \in R, t[A_i]$

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



## 2.2 关系代数

- 概述
- 传统的集合运算
- 专门的关系运算



## 2.2.1 传统的集合运算

- 并
- 差
- 交
- 广义笛卡尔积



# 1. 并 (Union)

## ➤ $R$ 和 $S$ 相容

- 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

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



# 并(续)

$R$

$A$	$B$	$C$
$a1$	$b1$	$c1$
$a1$	$b2$	$c2$
$a2$	$b2$	$c1$

$S$

$A$	$B$	$C$
$a1$	$b2$	$c2$
$a1$	$b3$	$c2$
$a2$	$b2$	$c1$

$R \cup S$

$A$	$B$	$C$
$a1$	$b1$	$c1$
$a1$	$b2$	$c2$
$a1$	$b3$	$c2$
$a2$	$b2$	$c1$



## 2. 差 (Difference)

### ➤ $R$ 和 $S$ 相容

- 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

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





# 差(续)

*R*

<i>A</i>	<i>B</i>	<i>C</i>
<i>a1</i>	<i>b1</i>	<i>c1</i>
<i>a1</i>	<i>b2</i>	<i>c2</i>
<i>a2</i>	<i>b2</i>	<i>c1</i>

*R-S*

<i>A</i>	<i>B</i>	<i>C</i>
<i>a1</i>	<i>b1</i>	<i>c1</i>

*S*

<i>A</i>	<i>B</i>	<i>C</i>
<i>a1</i>	<i>b2</i>	<i>c2</i>
<i>a1</i>	<i>b3</i>	<i>c2</i>
<i>a2</i>	<i>b2</i>	<i>c1</i>



### 3. 交 (Intersection)

➤  $R$ 和 $S$ 相容

- 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)

$R$

$A$	$B$	$C$
$a1$	$b1$	$c1$
$a1$	$b2$	$c2$
$a2$	$b2$	$c1$

$R \cap S$

$A$	$B$	$C$
$a1$	$b2$	$c2$
$a2$	$b2$	$c1$

$S$

$A$	$B$	$C$
$a1$	$b2$	$c2$
$a1$	$b3$	$c2$
$a2$	$b2$	$c1$



## 4. 广义笛卡尔积 (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:  $(n+m)$  列的元组的集合

- 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

- $R \times S = \{t_r t_s \mid t_r \in R \wedge t_s \in S\}$



# 广义笛卡尔积 (续)

$R$

$A$	$B$	$C$
$a1$	$b1$	$c1$
$a1$	$b2$	$c2$
$a2$	$b2$	$c1$

$S$

$A$	$B$	$C$
$a1$	$b2$	$c2$
$a1$	$b3$	$c2$
$a2$	$b2$	$c1$

$R \times S$

$R.A$	$R.B$	$R.C$	$S.A$	$S.B$	$S.C$
$a1$	$b1$	$c1$	$a1$	$b2$	$c2$
$a1$	$b1$	$c1$	$a1$	$b3$	$c2$
$a1$	$b1$	$c1$	$a2$	$b2$	$c1$
$a1$	$b2$	$c2$	$a1$	$b2$	$c2$
$a1$	$b2$	$c2$	$a1$	$b3$	$c2$
$a1$	$b2$	$c2$	$a2$	$b2$	$c1$
$a2$	$b2$	$c1$	$a1$	$b2$	$c2$
$a2$	$b2$	$c1$	$a1$	$b3$	$c2$
$a2$	$b2$	$c1$	$a2$	$b2$	$c1$



练习：有两个关系R和S，分别包含15个和10个元组，则在 $R \cup S$ ， $R - S$ ， $R \cap S$ 中不可能出现的元组个数情况是（ ）。

A 15, 5, 10

B 18, 7, 7

C 21, 11, 4

D 25, 15, 0



## 2.2 关系代数

- 概述
- 传统的集合运算
- 专门的关系运算



## 2.2.2 专门的关系运算

- 选择
- 投影
- 连接
- 除





# 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

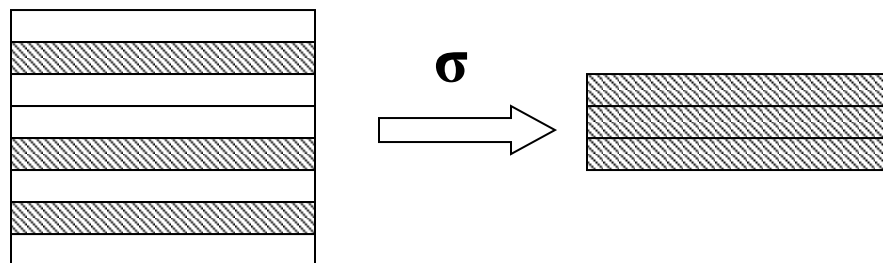
$$\sigma_F(R) = \{t | t \in R \wedge F(t) = \text{'真'}\}$$

–  $F$ : 选择条件, 是由比较运算符和/或逻辑运算符组合构成的表达式



# 选择（续）

## ➤ 3) 选择运算是从行的角度进行的运算



## ➤ 4) 举例

设有一个学生-课程数据库，包括学生关系、课程关系和选修关系。



# 学生

学号	姓名	性别	籍贯	出生年份	学院
090101	王英	女	河北	1989	计算机
090102	王小梅	女	江苏	1990	信电
090103	张小飞	男	宜昌	1990	计算机
090104	孙志鹏	男	海南	1991	计算机
090105	徐颖	女	江苏	1991	信电
090106	钱易蒙	男	河北	1990	外文
...	...	...	...	...	...



## [例1]查询计算机全体学生情况

$$\sigma_{\text{学院}=\text{“计算机”}}(\text{学生})$$
$$\sigma_6=\text{“计算机”}(\text{学生})$$

学号	姓名	性别	籍贯	出生年份	学院
090101	王英	女	河北	1989	计算机
090103	张小飞	男	宜昌	1990	计算机
090104	孙志鹏	男	海南	1991	计算机
...	...	...	...	...	...



## [例2]查询80年代出生的全体学生情况

$\sigma$  出生年份 $\geq 1980 \wedge$  出生年份 $\leq 1989$  (学生)

学号	姓名	性别	籍贯	出生年份	学院
090101	王英	女	河北	1989	计算机
090107	郭小娜	女	江苏	1988	机电
090108	黎明	男	北京	1986	采矿
....	...	...	...	...	...



### [例3]查询信电学院江苏籍全体学生情况

$\sigma$  学院="信电"  $\wedge$  籍贯="江苏" (学生)

学号	姓名	性别	籍贯	出生年份	学院
090102	王小梅	女	江苏	1990	信电
090105	徐颖	女	江苏	1991	信电
...	...	...	...	...	...



## [例4]查询江苏或者河北全体学生情况

$\sigma$  籍贯=“江苏”  $\vee$  籍贯=“河北” (学生)

学号	姓名	性别	籍贯	出生年份	学院
090101	王英	女	河北	1989	计算机
090102	王小梅	女	江苏	1990	信电
090105	徐颖	女	江苏	1991	信电
090106	钱易蒙	男	河北	1990	外文
090107	郭小娜	女	江苏	1988	机电
090109	徐明	男	河北	1987	体育
...	...	...	...	...	...



# 选择运算的关键问题

- 确定操作对象是哪个关系？
- 操作的条件是什么？
- 如何表示？





## 2. 投影 (Projection)

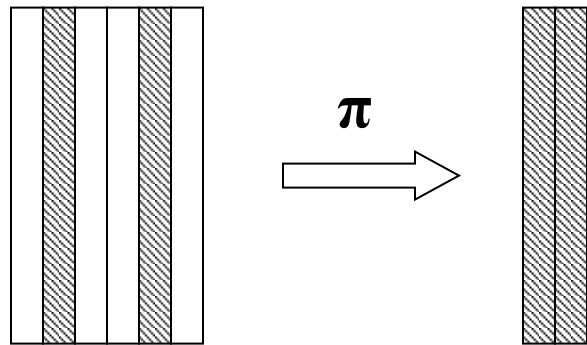
### 1) 投影运算符的含义

- 从  $R$  中选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）



# 学生表

学号	姓名	性别	籍贯	出生年份	学院
090101	王英	女	河北	1989	计算机
090102	王小梅	女	江苏	1990	信电
090103	张小飞	男	宜昌	1990	计算机
090104	孙志鹏	男	海南	1991	计算机
090105	徐颖	女	江苏	1991	信电
090106	钱易蒙	男	河北	1990	外文
...	...	...	...	...	...



## [例5]查询所有学生的姓名和籍贯

$\pi_{\text{姓名, 籍贯}}(\text{学生})$

$\pi_{2, 4}(\text{学生})$

姓名	籍贯
王英	河北
王小梅	江苏
张小飞	宜昌
孙志鹏	海南
徐颖	江苏
钱易蒙	河北
郭小娜	江苏
黎明	北京
徐明	河北



## [例6] 查询学生生源来自哪些省份？

$\pi_{\text{籍贯}}(\text{学生})$

$\pi_4(\text{学生})$

投影之后不仅取消了原关系中的某些列，而且还取消了某些元组。

籍贯

河北

江苏

宜昌

海南

...



[例7]查找出生年份在1980年以后（不含1980年）的学生的姓名、籍贯及其出生年份情况。

Π 姓名，籍贯，出生年份 (σ 出生年份>1980(学生))

姓名	籍贯	出生年份
王英	河北	1989
王小梅	江苏	1990
张小飞	宜昌	1990
孙志鹏	海南	1991
徐颖	江苏	1991
钱易蒙	河北	1990
郭小娜	江苏	1988
黎明	北京	1986
徐明	河北	1987
...	...	...



# 思考：

**C**（课程号，课程名，学分.....）

**SC**（学号，课程号，成绩）

查询选修了“数据库原理”的学生的学号。



# 3. 连接 (Join)

1) 连接也称为 $\theta$ 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

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

- $A$ 和 $B$ : 分别为 $R$ 和 $S$ 上度数相等且可比的属性组
- $\theta$ : 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的元组。



## 3) 两类常用连接运算

### – 等值连接 (equijoin)

- 什么是等值连接
  - $\theta$  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：

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





# 连接(续)

## — 自然连接 (Natural join)

- 什么是自然连接

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

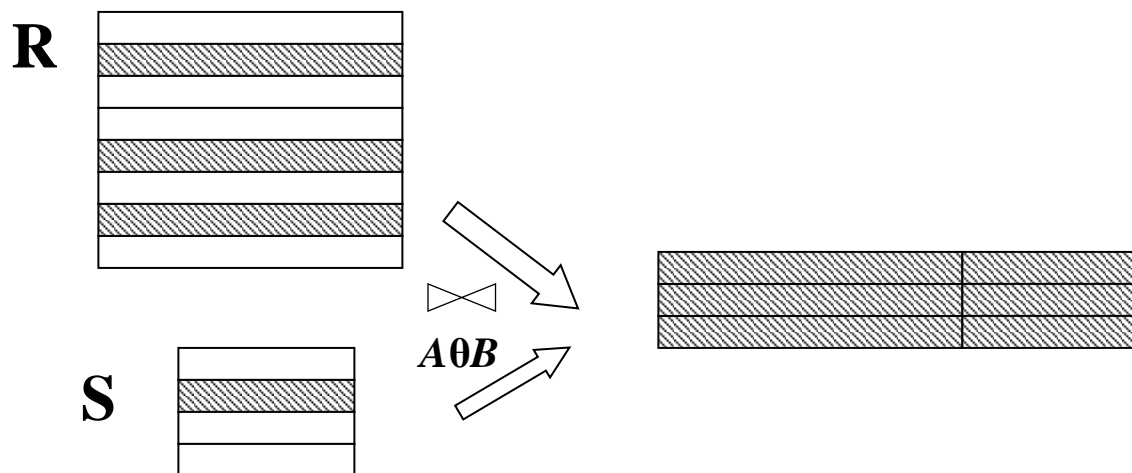
$R$ 和 $S$ 具有相同的属性组 $B$

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



# 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)

## 5) 举例

<i>A</i>	<i>B</i>	<i>C</i>
<i>a</i> <sub>1</sub>	<i>b</i> <sub>1</sub>	5
<i>a</i> <sub>1</sub>	<i>b</i> <sub>2</sub>	6
<i>a</i> <sub>2</sub>	<i>b</i> <sub>3</sub>	8
<i>a</i> <sub>2</sub>	<i>b</i> <sub>4</sub>	12

*R*

<i>B</i>	<i>E</i>
<i>b</i> <sub>1</sub>	3
<i>b</i> <sub>2</sub>	7
<i>b</i> <sub>3</sub>	10
<i>b</i> <sub>3</sub>	2
<i>b</i> <sub>5</sub>	2

*S*



# 连接(续)

$R \bowtie_{C < E} S$

<i>A</i>	<i>R.B</i>	<i>C</i>	<i>S.B</i>	<i>E</i>
<i>a</i> <sub>1</sub>	<i>b</i> <sub>1</sub>	5	<i>b</i> <sub>2</sub>	7
<i>a</i> <sub>1</sub>	<i>b</i> <sub>1</sub>	5	<i>b</i> <sub>3</sub>	10
<i>a</i> <sub>1</sub>	<i>b</i> <sub>2</sub>	6	<i>b</i> <sub>2</sub>	7
<i>a</i> <sub>1</sub>	<i>b</i> <sub>2</sub>	6	<i>b</i> <sub>3</sub>	10
<i>a</i> <sub>2</sub>	<i>b</i> <sub>3</sub>	8	<i>b</i> <sub>3</sub>	10



# 连接(续)

等值连接  $R \bowtie_{R.B=S.B} S$

$A$	$R.B$	$C$	$S.B$	$E$
$a_1$	$b_1$	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  $R \bowtie S$

<i>A</i>	<i>B</i>	<i>C</i>	<i>E</i>
<i>a</i> <sub>1</sub>	<i>b</i> <sub>1</sub>	5	3
<i>a</i> <sub>1</sub>	<i>b</i> <sub>2</sub>	6	7
<i>a</i> <sub>2</sub>	<i>b</i> <sub>3</sub>	8	10
<i>a</i> <sub>2</sub>	<i>b</i> <sub>3</sub>	8	2



# R × S

$R \bowtie S$   
R.B ≠ S.B

R.A	R.B	R.C	S.A	S.B	S.C
a <sub>1</sub>	b <sub>1</sub>	c <sub>1</sub>	a <sub>1</sub>	b <sub>2</sub>	c <sub>2</sub>
a <sub>1</sub>	b <sub>1</sub>	c <sub>1</sub>	a <sub>1</sub>	b <sub>3</sub>	c <sub>2</sub>
a <sub>1</sub>	b <sub>1</sub>	c <sub>1</sub>	a <sub>2</sub>	b <sub>2</sub>	c <sub>1</sub>
a <sub>1</sub>	b <sub>2</sub>	c <sub>2</sub>	a <sub>1</sub>	b <sub>3</sub>	c <sub>2</sub>
a <sub>1</sub>	b <sub>2</sub>	c <sub>2</sub>	a <sub>2</sub>	b <sub>2</sub>	c <sub>1</sub>
a <sub>2</sub>	b <sub>2</sub>	c <sub>1</sub>	a <sub>1</sub>	b <sub>2</sub>	c <sub>2</sub>
a <sub>2</sub>	b <sub>2</sub>	c <sub>1</sub>	a <sub>2</sub>	b <sub>2</sub>	c <sub>1</sub>

R.A	R.B	R.C	S.A	S.B	S.C
a <sub>1</sub>	b <sub>1</sub>	c <sub>1</sub>	a <sub>1</sub>	b <sub>2</sub>	c <sub>2</sub>
a <sub>1</sub>	b <sub>1</sub>	c <sub>1</sub>	a <sub>1</sub>	b <sub>3</sub>	c <sub>2</sub>
a <sub>1</sub>	b <sub>1</sub>	c <sub>1</sub>	a <sub>2</sub>	b <sub>2</sub>	c <sub>1</sub>
a <sub>1</sub>	b <sub>2</sub>	c <sub>2</sub>	a <sub>1</sub>	b <sub>3</sub>	c <sub>2</sub>
a <sub>2</sub>	b <sub>2</sub>	c <sub>1</sub>	a <sub>1</sub>	b <sub>3</sub>	c <sub>2</sub>

$R \bowtie S$   
R.B = S.B

R.A	R.B	R.C	S.A		S.C
a <sub>1</sub>	b <sub>2</sub>	c <sub>2</sub>	a <sub>1</sub>		c <sub>2</sub>
a <sub>1</sub>	b <sub>2</sub>	c <sub>2</sub>	a <sub>2</sub>		c <sub>1</sub>
a <sub>2</sub>	b <sub>2</sub>	c <sub>1</sub>	a <sub>1</sub>		c <sub>2</sub>
a <sub>2</sub>	b <sub>2</sub>	c <sub>1</sub>	a <sub>2</sub>		c <sub>1</sub>



# 关系SC和关系C的自然联接

关系SC

SNO	CNO	GRADE
S3	C3	87
S1	C2	88
S4	C3	79
S9	C4	83

关系C

CNO	CNAME	CDEPT	TNAME
C2	离散数学	计算机	汪宏伟
C3	高等数学	通讯	钱 红
C4	数据结构	计算机	马 良
C1	计算机原理	计算机	李 兵

第一步，计算 $SC \times C$ ；

第二步，计算满足 $SC.CNO=C.CNO$ 条件的元组；

第三步，去掉重复列，操作结果为：





# 关系SC和关系C的自然联接

SNO	CNO	GRADE	CNAME	CDEPT	TNAME
S3	C3	87	高等数学	通讯	钱 红
S1	C2	88	离散数学	计算机	汪宏伟
S4	C3	79	高等数学	通讯	钱 红
S9	C4	83	数据结构	计算机	马 良



# 自然连接和等值连接的区别

- ① 等值连接要求相等的分量，但不一定是公共属性，而自然连接要求相等的分量必须是公共属性
- ② 等值连接不做投影运算，而自然连接要把重复的属性去掉；
- ③ 自然连接一定是等值连接，但等值连接不一定是自然连接。



思考：

查询选修了全部课程的学生学号



# 象集Z

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：

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

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。



# 象集Z

$R$

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

$A=a_1$ 的象集

$B$	$C$
$b_1$	$c_2$
$b_2$	$c_3$
$b_2$	$c_1$



## 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。

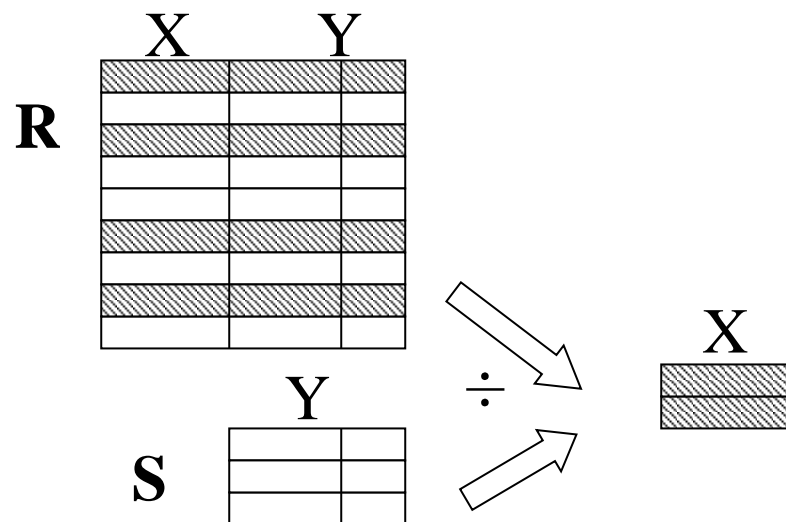
$$R \div S = \{t_r[X] \mid t_r \in R \wedge \pi_Y(S) \subseteq Y_x\}$$

$Y_x$ :  $x$  在  $R$  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



# $R \div S$

$R$

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

$S$

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





$R ( \textcolor{red}{X}, \textcolor{red}{Y} )$

$R$

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

$S ( \textcolor{red}{Y}, \textcolor{red}{Z} )$

$S$

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



# $R \div S$

$R$

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

$B$	$C$
$b_1$	$c_2$
$b_2$	$c_1$
$b_2$	$c_3$

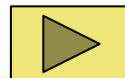
$R \div S$
$A$
$a_1$



# 5. 综合举例

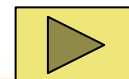
**Student**

学 号 Sno	姓 名 Sname	性 别 Ssex	年 龄 Sage	所在系 Sdept
95001	李勇	男	20	CS
95002	刘晨	女	19	IS
95003	王敏	女	18	MA
95004	张立	男	19	IS



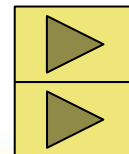
# Course

课程号	课程名	先行课	学分
Cno	Cname	Cpno	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



SC

学 号	课 程 号	成 绩
Sno	Cno	Grade
95001	1	92
95001	2	85
95001	3	88
95002	2	90
95002	3	80



# 以学生-课程数据库为例

**[例8] 查询至少选修1号课程和3号课程的学生学号**

构造小集合:

Cno
1
3

$$\pi_{Cno}(\sigma_{Cno='1' \vee Cno='3'}(Course))$$



## [例8] 查询至少选修1号课程和3号课程的学生学号

$$\pi_{Sno, Cno}(SC) \div \pi_{Cno}(\sigma_{Cno='1' \vee Cno='3'}(Course))$$

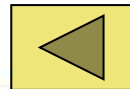


# 综合举例(续)

**[例9] 查询选修了全部课程的学生学号和姓名。**

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

思考：求被所有的学生都选修了的课程名？





# 练习题

- 已知关系模式

商场（商场号、商场名、所在城市）

商品（商品号、商品名、商品类别、商品颜色）

销售（商场号、商品号、售价）

- 求销售所有商品的商场号和商场名，用关系代数表示



# 除法的代数表示方法

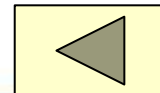
$$\mathbf{R \div S = \pi_x(R) - \pi_x((\pi_x(R) \times \pi_y(S)) - R)}$$



# 自连接查询

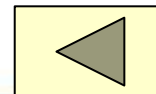
查询与刘晨在同一个学院学习的学生姓名

$$\Pi_7(\sigma_{5=10 \wedge 2='刘晨'}(\text{Student} \times \text{Student}))$$



# 练习：求每一门课程的间接先行课

$$\Pi_{1,7} (\sigma_{3=5} (\text{Course} \times \text{Course}))$$



查询至少选修了01号课程和03号课程的学生学号

$$\pi_{\text{Sno}, \text{Cno}}(\text{SC}) \div \pi_{\text{Cno}}(\sigma_{\text{Cno}='01' \vee \text{Cno}='03'}(\text{Course}))$$

思考： 还有其他两种表达方式，是什么？



# 外连接

- 外连接运算是自然连接运算的扩展，它可以处理缺失信息。
- 假设学生与选课信息如下：**求所有学生的全部信息（包括学号、姓名、性别、专业、课程号和成绩）**

学号	姓名	性别	专业
s01	Jack	m	CS
s02	Mary	f	CS
s03	Coco	f	IS

学号	课程号	成绩
s01	c01	85
s01	c02	80
s02	c03	75
s02	c02	85



# 外连接

- 使用外连接运算可以避免这样的信息丢失。  
外连接运算有三种形式：
  - 左外连接，用符号  $\bowtie$  来表示
  - 右外连接，用符号  $\ltimes$  来表示
  - 全外连接，用符号  $\Join$  来表示



# 左外连接

左侧关系的元组都要在连接结果中出现

学号	姓名	性别	专业	课程号	成绩
s01	Jack	m	CS	c01	85
s01	Jack	m	CS	c02	80
s02	Mary	f	CS	c03	75
s02	Mary	f	CS	c02	85
s03	Coco	f	IS	<b>null</b>	<b>null</b>





## 右外连接和全外连接

- 右外连接：右侧关系的元组都要在连接结果中出现
- 全外连接：左侧关系和右侧关系的元组都要在连接结果中出现



运算	单/双目	基本运算	复合运算	表示方法
并	双目	$\sqrt{\quad}$		<b>RUS</b>
差	双目	$\sqrt{\quad}$		<b>R-S</b>
交	双目		差	<b><math>R \cap S</math></b>
笛卡儿积	双目	$\sqrt{\quad}$		<b><math>R \times S</math></b>
选择	单目	$\sqrt{\quad}$		<b><math>\sigma_F(R)</math></b>
投影	单目	$\sqrt{\quad}$		<b><math>\pi_A(R)</math></b>
连接	双目		笛卡儿积、选择	<b><math>R \bowtie_{A \theta B} S</math></b>
自然连接	双目		笛卡儿积、选择、投影	<b><math>R \Join S</math></b>
除	双目		笛卡儿积、投影、差	<b><math>R \div S</math></b>

## 2.5 关系演算

- 与关系代数不同，关系演算是非过程化语言，它只描述所需要的信息，而不给出获得该信息的具体过程。
- 关系演算又分为元组关系演算和域关系演算两种。



# 1. 元组关系演算

➤ 元组关系演算的查询表达式为:

$$\{t \mid P(t)\}$$

它是所有使谓词P为真的元组t的集合。

➤ 是对关系的元组变量t进行运算，也就是说元组关系演算的结果是符合给定条件的元组的集合，也就是一个关系。

➤  $t[A]$  表示元组t在属性A上的取值，用  $t \in r$  表示元组t在关系r中。



# 例 题

(1)找出所有贷款额在1200元以上的贷款的分支机构、贷款号码和贷款金额，则其元组关系演算表达式为：

$$\{t \mid t \in \text{loan} \wedge t[\text{amount}] > 1200\}$$

**loan**

<i>branch-name</i>	<i>loan_number</i>	<i>amount</i>
Downtown	L-17	1000
Redwood	L-23	2000
Perryridge	L-15	1500
Downtown	L-14	1500
Mianus	L-93	500
Round Hill	L-11	900
Perryridge	L-16	1300



(2)找出所有贷款额在1200元以上的贷款的贷款号码，  
则其元组关系演算表达式为：

为了用元组关系演算来表达，需要为模式  
(loan-number) 上的关系写一个表达式。我们  
需要 (loan-number)上的、在loan中对应元组的  
属性amount>1200的那些元组。为了表达这样  
的要求，需要引入数理逻辑中的“存在”这一  
结构。记法  $\exists t \in r(Q(t))$

表示 “关系r中存在元组t使谓词Q (t) 为真”

$\{t | \exists s \in \text{loan}(t[\text{loan-number}] = s[\text{loan-number}]$   
 $\wedge s[\text{amount}] > 1200)\}$



# 找出从perryridge支行贷款的所有客户姓名

$$\{t \mid \exists s \in \text{borrower}(t[\text{customer-name}] = s[\text{customer-name}] \\ \wedge \exists u \in \text{loan}(u[\text{loan-number}] = s[\text{loan-number}] \\ \wedge u[\text{branch-name}] = \text{"Perryridge"}))\}$$

## loan

<i>branch-name</i>	<i>loan_number</i>	<i>amount</i>
Downtown	L-17	1000
Redwood	L-23	2000
Perryridge	L-15	1500
Downtown	L-14	1500
Mianus	L-93	500
Round Hill	L-11	900
Perryridge	L-16	1300

## borrower

<i>customer-name</i>	<i>loan_number</i>
Jones	L-17
Smith	L-23
Hayes	L-15
Jackson	L-14
Curry	L-93
Smith	L-11
Williams	L-17
Adams	L-16



找出在银行有贷款或有账户或二者兼有的所有客户

## borrower

<i>customer-name</i>	<i>account_number</i>
Johnson	A-101
Smith	A-215
Hayes	A-102
Turner	A-305
Johnson	A-201
Jones	A-217
Lindsay	A-222

## depositor

<i>customer-name</i>	<i>loan_number</i>
Jones	L-17
Smith	L-23
Hayes	L-15
Jackson	L-14
Curry	L-93
Smith	L-11
Williams	L-17
Adams	L-16

$$\{t \mid \exists s \in \text{borrower}(t[\text{customer-name}] = s[\text{customer-name}]) \vee \exists u \in \text{depositor}(t[\text{customer-name}] = u[\text{customer-name}])\}$$




找出在银行中有账户又有贷款的所有客户

$$\{t \mid \exists s \in \text{borrower}(t[\text{customer-name}] = s[\text{cusotmer-name}]) \\ \wedge \exists u \in \text{depositor}(t[\text{customer-name}] = u[\text{customer-name}])\}$$



## 2. 域关系演算

➤ 域关系演算的查询表达式为：

$$\{ \langle x_1, x_2, \dots, x_n \rangle \mid P(x_1, x_2, \dots, x_n) \}$$

其中 $x_1, x_2, \dots, x_n$ 代表域变量， $P$ 代表谓词。

- 是对域变量 $x_i$ 进行运算，域变量从属性的域中取值，而不是从整个元组中取值。
- 域关系演算的结果是符合给定条件的域变量值序列的集合，也就是一个关系。



# 例 题

(1)找出所有贷款额在1200元以上的贷款的分支机构、贷款号码和贷款金额，则其域关系演算的表达式为：

$$\{ \langle b, l, a \rangle \mid \langle b, l, a \rangle \in \text{loan} \wedge a > 1200 \}$$

**loan**

<i>branch-name</i>	<i>loan_number</i>	<i>amount</i>
Downtown	L-17	1000
Redwood	L-23	2000
Perryridge	L-15	1500
Downtown	L-14	1500
Mianus	L-93	500
Round Hill	L-11	900
Perryridge	L-16	1300



# 例 题

**(2)找出所有贷款额在1200元以上的贷款的贷款号码，则其域关系演算的表达式为：**

$$\{ \langle l \rangle \mid \exists b, a (\langle b, l, a \rangle \in loan \wedge a > 1200) \}$$

