

表示记号

- $\widehat{t_r t_s}$

- 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。它是一个 $n + m$ 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

$(n+m)$ 元组 $\widehat{t_r t_s}$



n 元组 t_r

m 元组 t_s



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



笛卡尔积(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

Relations r , s :

A	B
-----	-----

α	1
β	2

r

C	D	E
-----	-----	-----

α	10	a
β	10	a
β	20	b
γ	10	b

s

$r \times s$:

A	B	C	D	E
-----	-----	-----	-----	-----

α	1	α	10	a
α	1	β	19	a
α	1	β	20	b
α	1	γ	10	b
β	2	α	10	a
β	2	β	10	a
β	2	β	20	b
β	2	γ	10	b



2.4 关系代数

- 概述
- 传统的集合运算
- 专门的关系运算



专门的关系运算(续)

- 选择
- 投影
- 连接
- 除



➤图为学生-课程数据库中的student关系、Course关系、SC关系

SC:

Sno	Cno	Grade
200215121	1	92
200215121	2	85
200215121	3	88
200215122	2	90
200215122	3	80

Student:

Sno	Sname	Ssex	Sage	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

Course:

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

选择(Selection)

- **问题：**如何查询信息系(IS)的全体学生信息

选择满足条件：Sdept='IS'的元组

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

选择(续)

- 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

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

- F : 选择条件, 是一个逻辑表达式, 基本形式为:

$X_1 \theta Y_1$

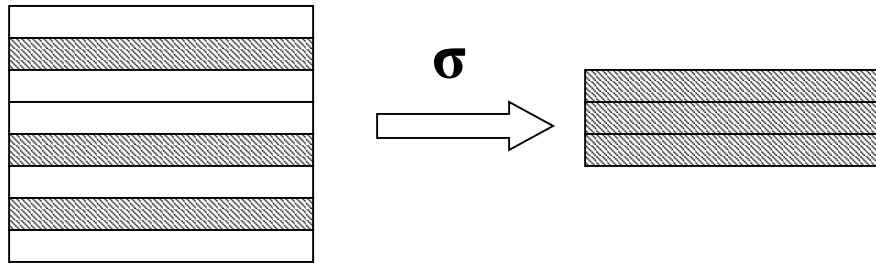
比较运算符, 如 $>$, \geq ,
 $<$, \leq , $=$, $<>$



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



• [例1] 查询信息系

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

或 $\sigma_5='IS' (Student)$

选

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

□ 结果:

属性名可以用属性序号代替

Sno	Sname	Ssex	Sage	Sdept
95002	刘晨	女	19	IS
95004	张立	男	19	IS

• [例2] 查询年龄小

$\sigma_{\text{Sage} < 20}(\text{Student})$
或 $\sigma_{4 < 20}(\text{Student})$

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

□ 结果:

Sno	Sname	Ssex	Sage	Sdept
95002	刘晨	女	19	IS
95003	王敏	女	18	MA
95004	张立	男	19	IS



• [例] 查询年龄小于

$\sigma_{\text{Sage} < 20 \wedge \text{Ssex} = \text{'女'}}$

或 $\sigma_{4 < 20 \wedge 3 = \text{'女'}}$ (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

□ 结果:

Sno	Sname	Ssex	Sage	Sdept
95002	刘晨	女	19	IS
95003	王敏	女	18	MA

——选择是在关系R中选择满足给定条件的诸元组

投影 (Projection)

- **问题：** 如何查询全部学生的姓名及其所在系信息

显示学生关系中的姓名和系属性列

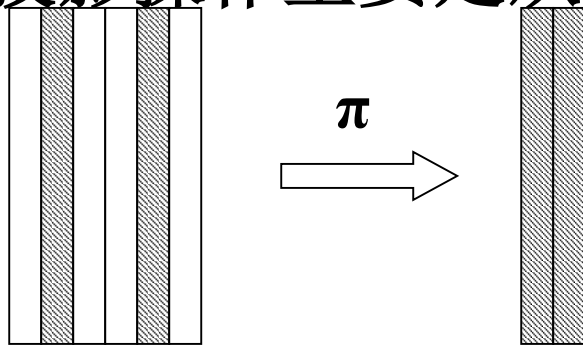
学号 Sno	姓名 Sname	性别 Ssex	年龄 Sage	所在系 Sdept
200215121	李勇	男	20	CS
200215122	刘晨	女	19	IS
200215123	王敏	女	18	MA
200215125	张立	男	19	IS

投影(续)

- **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

$$\pi_A(R) = \{ t[A] \mid t \in R \} \text{ ---其中} A \text{为} R \text{中的属性列}$$

- 投影操作主要是从列的角度进行运算



■但投影之后可能取消某些元组(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Sname和Sdept两个属性上的投影

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

或 $\pi_{2, 5}(\text{Student})$

属性名
可以用
属性序
号代替

□ 结果:

Sname	Sdept
李勇	CS
刘晨	IS
王敏	MA
张立	IS

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

• [例4] 查询学生关系

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

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

□ 结果:

Sdept
CS
IS
MA

投影之后取消了原关系中的某些元组
(避免重复行)



➤图为学生-课程数据库中的student关系、Course关系、SC关系

SC:

Sno	Cno	Grade
200215121	1	92
200215121	2	85
200215121	3	88
200215122	2	90
200215122	3	80

Student:

Sno	Sname	Ssex	Sage	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

Course:

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

选择投影运算复合使用示例

- 复合运用投影、选择、笛卡尔积运算，可以从任意n张表中截取满足条件的子表

例 查询选修了2号课程的学生学号。

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

例 查询选过“数据库”课程学生学号。

$$\pi_{Sno} (\sigma_{sc.cno=course.cno \wedge Cname = '数据库'} (SC \times Course))$$

解：需查询的数据要根据“选课”信息以及“课程”

信息来完成，因此，查询涉及SC和COURSE两张表

step1: 先将“选课”与“课程”表合并为一张表

SC × Course，结果为所有选过任意课程的信息

step2: 再从上一步结果中，对于任意一次选课，所选课

程存在sc.cno=course.cno，且选的课程是“数据库”即Cname=‘数据库’的元组

笛卡尔连接运算存在的问题

- 问题示例 观察上例中表达式 $SC \times Course$ ，其结果中包含所有的选课与所有课程的组合，而实际中，一个学生不可能选所有课程，这样结果元组中许多是无意义的。上例对200215121学生只有第一个元组有意义

SNO	CNO	GRADE	CNO	CNAME	CPNO	CREDIT
200215121	1	92	1	数据库	5	4
200215121	1	92	2	数学	null	2
200215121	1	92	3	信息系统	1	4
200215121	1	92	4	操作系统	6	3
200215121	1	92	5	数据结构	7	4
200215121	1	92	6	数据处理	null	2
200215121	1	92	7	C 语言	6	4
...
200215121	2	92	1	数据库	5	4

连接运算

- **问题分析** 笛卡尔乘运算为了保证数学上的**完整性**，将两张表的‘**所有**’内容合并，若该两张表之间存在关联关系，则合并后结果中将那些**无关联**关系的元组也合并了。

数学上的“完整”在使用中不便，甚至造成“信息丢失”

- **问题解决**
 - 1) 通过在 σ_P 运算中确定适当的P，去除那些无关联的元组。如 `sc.cno=course.cno`
 - 2) 定义扩展的笛卡尔乘运算，在**合并时去除**无关联元组



连接 (续)

- 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(续)

- 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

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

- 连接后，结果关系的关系模式由 R 和 S 的所有属性组成， R 的属性在前， S 的属性在后



连接(续)

- [例] 计算 SC  Course
SC. Cno < Course. Cno

关系SC

Sno	Cno	Grade
001	1	92
001	3	88
002	2	90
003	6	80

关系Course

Cno	Cname	Cpno	Ccredit
1	操作系统	6	3
3	数据结构	7	4
6	数据处理		2

----从SC × Course中选取SC关系中课程号小于Course关系中课程号值的元组



连接(续)

SC  Course
SC.Cno < Course.Cno

- 计算步骤

(1) 计算笛卡尔积 SC × Course, 结果为

Sno	SC.Cno	Grade	Course.Cno	Cname	Cpno	Ccredit
001	1	92	1	操作系统	6	3
001	1	92	3	数据结构	7	4
001	1	92	6	数据处理		2
001	3	88	1	操作系统	6	3
001	3	88	3	数据结构	7	4
001	3	88	6	数据处理		2
002	2	90	1	操作系统	6	3
002	2	90	3	数据结构	7	4
002	2	90	6	数据处理		2
003	6	80	1	操作系统	6	3
003	6	80	3	数据结构	7	4
003	6	80	6	数据处理		2

连接(续)

SC \bowtie Course
SC.Cno < Course.Cno

(2) 从SC \times Course中选择满足SC.Cno < Course.Cno的元组

Sno	SC.Cno	Grade	Course.Cno	Cname	Cpno	Ccredit
001	1	92	1	操作系统	6	3
001	1	92	3	数据结构	7	4
001	1	92	6	数据处理		2
001	6	88	1	操作系统	6	3
001	3	88	3	数据结构	7	4
001	3	88	6	数据处理		2
002	2	90	1	操作系统	6	3
002	2	90	3	数据结构	7	4
002	2	90	6	数据处理		2
003	6	80	1	操作系统	6	3
003	6	80	3	数据结构	7	4
003	6	80	6	数据处理		2

连接(续)

SC  Course
SC. Cno < Course. Cno

(3) 结果为

对关系SC和Course进行连接的结果

Sno	SC. Cno	Grade	Course. Cno	Cname	Cpno	Ccredit
001	1	92	3	数据结构	7	4
001	1	92	6	数据处理		2
001	3	88	6	数据处理		2
002	2	90	3	数据结构	7	4
002	2	90	6	数据处理		2



连接(续)----例

计算 $R \bowtie_{1>1} S$

R

S

A	B	D
2	5	8
5	4	1
4	5	8
3	4	9

B	C	D
3	5	8
4	4	1
4	1	8
6	4	1

$R \bowtie_{1>1} S$

A	R. B	R. D	S. B	C	S. D
5	4	1	3	5	8
5	4	1	4	4	1
5	4	1	4	1	8
4	5	8	3	5	8

- ❖ 注意连接运算的执行情况!!
- ❖ 满足条件的元组 ----匹配的元组(保留)
- ❖ 不满足条件的元组 ----不匹配的元组(舍去)



连接(续)

- 常用的两类连接运算
 - 等值连接(equijoin)
 - 自然连接(Natural join)



连接(续)----等值连接

- 等值连接

- θ 为“=”的连接运算

- 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

$$R \underset{A=B}{\bowtie} S = \{ \underset{r}{t_r} \underset{s}{t_s} \mid t_r \in R \wedge t_s \in S \wedge \underline{t_r[A] = t_s[B]} \}$$



等值连接(续)--例

计算 $R \bowtie_{1=1} S$

R

S

A	B	D
2	5	8
7	4	1
4	5	8
3	4	9

B	C	D
3	5	8
4	4	1
4	1	8
6	4	1

$R \bowtie_{1=1} S$

A	R. B	R. D	S. B	C	S. D
4	5	8	4	4	1
4	5	8	4	1	8
3	4	9	3	5	8

1. $R \times S$

2. $\sigma_{R.A=S.B}(R \times S)$



连接(续)----自然连接

❖ 自然连接：一种特殊的等值连接

■ 两个关系中进行比较的分量须是相同的属性组

■ 在结果中把重复的属性列去掉

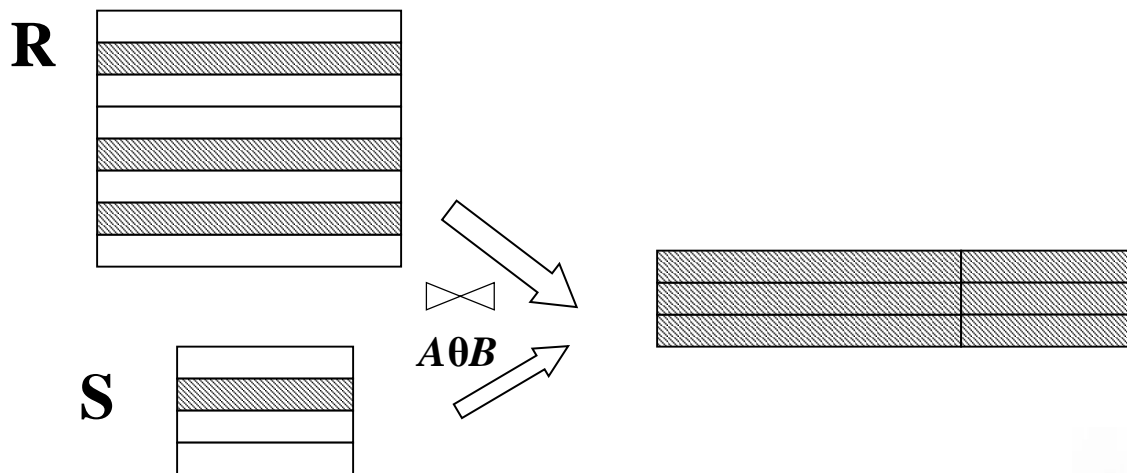
➤ 含义： R 和 S 具有相同的属性组 B

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



连接(续)

- 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。



自然连接(续)

- 自然连接运算的步骤
 - 计算笛卡尔积 $R \times S$
 - 从 $R \times S$ 中选择那些公共属性A的数值相同 ($\sigma_{R.A=S.A}$) 的元组
 - 去掉S.A(或R.A)，将留下来的R.A(或S.A)改为A，即得所要的结果



连接(续)

- 自然连接例：计算SC \bowtie Course
- 步骤：
 - (1) 计算SC \times Course

Sno	SC. Cno	Grade	Course. Cno	Cname	Cpno	Ccredit
001	1	92	1	操作系统	6	3
001	1	92	3	数据结构	7	4
001	1	92	6	数据处理		2
001	6	88	1	操作系统	6	3
001	3	88	3	数据结构	7	4
001	3	88	6	数据处理		2
002	2	90	1	操作系统	6	3
002	2	90	3	数据结构	7	4
002	2	90	6	数据处理		2
003	6	80	1	操作系统	6	3
003	6	80	3	数据结构	7	4
003	6	80	6	数据处理		2



自然连接例(续)

(2) 选择 $\sigma_{SC.Cno=Course.Cno}$ ($SC \times Course$)

Sno	SC. Cno	Grade	Course. Cno	Cname	Cpno	Ccredit
001	1	92	1	操作系统	6	3
001	3	88	3	数据结构	7	4
003	6	80	6	数据处理		2

自然连接例(续)

(3) 删除重复列SC.Cno，并将留下来的Course.Cno改为Cno，得到结果

进行SC ⋈ Course运算的结果

Sno	Grade	Cno	Cname	Cpno	Ccredit
001	92	1	操作系统	6	3
001	88	3	数据结构	7	4
003	80	6	数据处理		2



自然连接与等值连接

- 自然连接的两个关系要求有公共的属性组B，等值连接则不要求；
- 自然连接中等值的条件一定是公共属性组的值相等 ($R.B = S.B$)，而等值连接不一定；
- 自然连接中等值的条件隐含，不显式地写出来，而等值连接要写出来；
- 自然连接要在结果中去掉重复的一个属性组B，而等值连接则不。



自然连接运算----例

R和S有公共属性B, D

R

A	B	C	D
α	1	α	a
β	2	γ	a
γ	4	β	b
α	1	γ	a
δ	2	β	b

S

B	D	E
1	a	α
3	a	β
1	a	γ
2	b	δ
3	b	ϵ

结果中只有一个B, D

R \bowtie **S**

A	B	C	D	E
α	1	α	a	α
α	1	α	a	γ
α	1	γ	a	α
α	1	γ	a	γ
δ	2	β	b	δ



自然连接与笛卡尔积的区别

- 1) 两者的语义上都有合并两张表的作用

笛卡尔乘是数学意义上的所有可能组合的乘积，而自然连接则是将两张表中有关联关系的元组合并

例如 $SC \times Course$ 是所有可能的选课及所有课程信息

$SC \bowtie Course$ 则是所有的选课及所选课程的信息

- 2) 自然连接有选择 σ_p 的语义

例如 将两个表 SC 、 $Course$ 做自然连接 \bowtie ，可以理解为在 SC 中选择那些选过 “...课程” 的学生

示例 找出选过学分为4的学生号

$\sigma_{credit=4}(Course)$ // 找出学分为4的课程

$SC \bowtie \sigma_{credit=4}(Course)$ // 选出选过4分的选课

$\pi_{Sno} (SC \bowtie \sigma_{credit=4}(Course))$ // 选出选过4分的学号



假设 R_1 、 R_2 和 R_3 是三个关系

R_1

A	B	C
a ₁	b ₁	55
a ₂	b ₂	45
a ₄	b ₂	35

R_2

A	B	C
a ₁	b ₂	55
a ₂	b ₂	45
a ₃	b ₁	35
a ₄	b ₂	35

R_3

A	D
a ₁	101
a ₂	101
a ₂	102
a ₃	102

$R_1 \bowtie R_2$

$R_1.C < R_2.C$

$R_1 \bowtie R_2$

$R_1.C = R_2.C$

$R_1 \bowtie R_3$



外连接

例：列出教师的有关信息，包括姓名、工资、所教授的课程。

$\Pi_{Pno, PN, SAL, Cno, CN} (T \bowtie TC \bowtie C)$

教师表 (T)

Pno	PN	SAL
P01	赵明	800
P02	钱广	700
P03	孙立	600
P04	李三	500

任课表 (TC)

Cno	Pno
C01	P01
C02	P02
C02	P04

课程表 (C)

Cno	CN
C01	物理
C02	数学
C03	化学



==

Pno	PN	SAL	Cno	CN
P01	赵明	800	C01	物理
P02	钱广	700	C02	数学
P04	李三	500	C02	数学

问题：有关P03号教师的姓名和工资信息没有显示出来
——失配的元组

外连接(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)。

◆ **内连接**：即自然连接。连接结果只取匹配的元组，舍弃不匹配的元组。

◆ **外连接**：外连接=内连接+失配的元组



外连接

自然连接 $R \bowtie S$ 的结果

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

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

R中被舍去的元组

S中被舍去的元组

外连接(续)

■例：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


把舍弃的元组也保存在结果关系中，在其他属性上填空值(NULL)

(a) 外连接

外连接(续)

- 左外连接
 - 只把左边关系 R 中要舍弃的元组保留
- 右外连接
 - 只把右边关系 S 中要舍弃的元组保留
- 全外连接
 - 把左边 R 和右边关系 S 中要舍弃的元组都保留

左外连接 = 内连接 + 左侧表中失配的元组
右外连接 = 内连接 + 右侧表中失配的元组
全外连接 = 内连接 + 两侧表中失配的元组



左外连接运算：例

教师表(T)

Pno	PN	SAL
P01	赵明	800
P02	钱广	700
P03	孙立	600
P04	李三	500

任课表(TC)

Cno	Pno
C01	P01
C02	P02
C02	P04

课程表(C)

Cno	CN
C01	物理
C02	数学
C03	化学



=

Pno	PN	SAL	Cno	CN
P01	赵明	800	C01	物理
P02	钱广	700	C02	数学
P04	李三	500	C02	数学
P03	孙立	600	null	null

所有教师的信息包括匹配的和失配的

右外连接运算：例

教师表 (T)

Pno	PN	SAL
P01	赵明	800
P02	钱广	700
P03	孙立	600
P04	李三	500

任课表 (TC)

Cno	Pno
C01	P01
C02	P02
C02	P04

课程表 (C)

Cno	CN
C01	物理
C02	数学
C03	化学



==

Pno	PN	SAL	Cno	CN
P01	赵明	800	C01	物理
P02	钱广	700	C02	数学
P04	李三	500	C02	数学
null	null	null	C03	化学

所有课程的信息包括匹配的和失配的



全外连接运算：例

教师表(T)

Pno	PN	SAL
P01	赵明	800
P02	钱广	700
P03	孙立	600
P04	李三	500

任课表(TC)

Cno	Pno
C01	P01
C02	P02
C02	P04

课程表(C)

Cno	CN
C01	物理
C02	数学
C03	化学



==

Pno	PN	SAL	Cno	CN
P01	赵明	800	C01	物理
P02	钱广	700	C02	数学
P04	李三	500	C02	数学
P03	孙立	600	null	null
null	null	null	C03	化学

所有教师和课程的信息, 包括匹配的和失配的

表示记号：象集 Z_x

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：

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

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_x 表示 R 中属性组 X 上值为 x 的诸元组在 Z 上分量的集合。

R	X	Z
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} 表示 R 中属性列 X 上值为 x_1 的诸元组，在 Z 上分量的集合。

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



象集 Z_x 表示 R 中属性组 X 上值为 x 的诸元组在 Z 上分量的集合。

□ 象集举例

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

■ x_2 在 R 中的象集

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

■ x_3 在 R 中的象集

$$Z_{x_3} = \{Z_1, Z_3\}$$



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:

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

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



R

X	Y

S

Y	Z

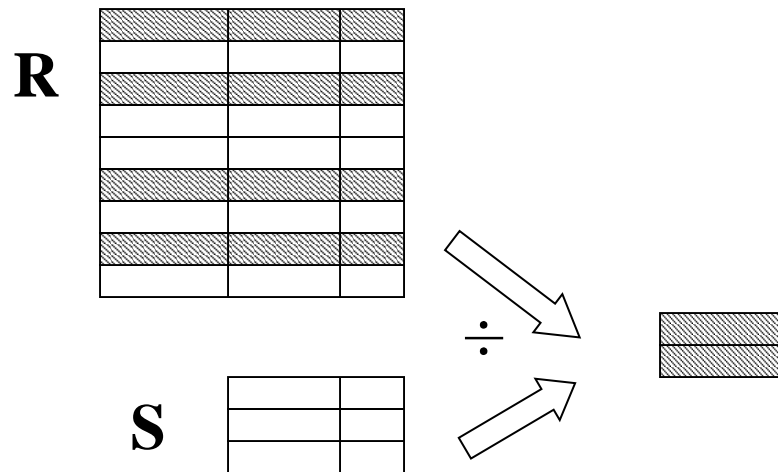
$\Pi_y(S)$

$R \div S \quad Y_x$



除(续)

- 2) 除操作是同时从行和列角度进行运算



除(续)

■ 计算 $R \div S$ 对于 A 中每一个 a , 其对应的像 (b, c) 若包含 S 中所有的 (b, c) , 则 (a) 是 $R \div S$ 的一个结果.

R	A	B	C
A	a1	b1	c2
a_1	a1	b2	c3
a_2	a1	b2	c1
a_3	a1	b2	c1
a_1	a2	b2	c7
a_4	a2	b2	c3
a_2	a2	b2	c3
a_1	a3	b4	c6
	a4	b6	c6

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

(b)

■ 显然只有 Z_{a_1} 包含了 S 在 (B, C) 属性组上的投影

■ 故: $R \div S = \{a_1\}$

分析

- 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 $R \div S = \{a_1\}$



除----练习

$$R(X, Y) \div S(Y) = R \div S$$

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

①

C	D
3	4
5	6

A	B
1	2
7	8

②

C	D
3	4
5	6
4	2

A	B
1	2

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 :

然后求: $\pi_{Sno, Cno}(SC) \div K$

Cno
1
3



综合举例(续)

- 例 7续 $\pi_{Sno,Cno}(SC)$

95001象集{1, 2, 3}

95002象集{2, 3}

$K=\{1, 3\}$

于是: $\pi_{Sno,Cno}(SC) \div K=\{95001\}$

Sno	Cno
95001	1
95001	2
95001	3
95002	2
95002	3

综合举例(续)

[例 8] 查询选修了2号课程的学生们的学号

$$\pi_{\text{Sno}} (\sigma_{\text{Cno}='2'} (\text{SC}))$$
$$= \{ 95001, 95002 \}$$



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学号和姓名。

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



思考

➤图为学生-课程数据库中的student关系、Course关系、SC关系

SC:

Sno	Cno	Grade
95001	1	92
95001	2	85
95001	3	88
95002	2	90
95002	3	

Student:

Sno	Sname	Ssex	Sage	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:

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

- 查2号课程的学生姓名和成绩
- 查选修“数学”的学生学号，姓名及该课程的成绩
- 查选修1号或2号课程的学生学号
- 查至少选修1号和2号课程的学生学号(不用除法)



- 查2号课程的学生姓名和成绩

$\Pi_{\text{Sname,Grade}}(\sigma_{\text{Cno}='2'}(\text{S} \bowtie \text{SC}))$



- 查选修“数学”的学生学号，姓名及该课程的成绩
 $\Pi_{\text{Sno}, \text{Sname}, \text{Grade}}(\sigma_{\text{Cname} = \text{'数学'}}(\text{S} \bowtie \text{SC} \bowtie \text{C}))$



- 查选修1号或2号课程的学生学号

$\Pi_{Sno}(\sigma_{Cno='1' \vee Cno='2'}(SC))$



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



元组关系演算

- 形式化定义

元组关系演算中，以元组为单位，通过谓词公式约束所要查找元组的条件，可以表示为：

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

其中： t 为元组变量，即查询的目的， φ 称为元组演算的谓词公式，即查询的条件。

$\{t \mid \varphi(t)\}$ 表示使 $\varphi(t)$ 为真的元组 t 的集合。



元组关系演算（续）

$\varphi(t)$ 可以通过原子公式、约束变量、自由变量、运算符构成

原子公式分3类：

$R(t)$ ： R 为关系名，表示 t 是 R 中的元组。

$t[i]\theta u[j]$ ：表示“元组 t 的第 i 个分量与元组 u 的第 j 个分量进行比较运算 θ ”，如 $t[2]<u[3]$ 。

$t[i]\theta C$ ：表示“元组 t 的第 i 个分量与常量 C 进行比较运算 θ ”，如 $t[3]>5$ 。



元组关系演算（续）

$\varphi(t)$ 约束变量与自由变量

若元组演算公式中的一个元组变量前有“全称量词”和“存在量词”，则称该变量为约束元组变量，否则称自由元组变量。

在公式 $(\exists t)\varphi(t)$ 和 $(\forall t)\varphi(t)$ 中， φ 称为是量词的辖域。 t 出现在 $(\forall t)$ 或 $(\exists t)$ 的辖域内， t 为约束元组变量，被量词所绑定。任何没有以这种方法显示绑定的变量都称为自由变量。



元组关系演算（续）

任意 $\varphi(t)$ 的递归定义



原子公式是公式



设 $\varphi_1(t_1)$ 和 $\varphi_2(t_2)$ 是公式，则 $\neg \varphi_1(t_1)$ ， $\varphi_1(t_1) \wedge \varphi_2(t_2)$ ， $\varphi_1(t_1) \vee \varphi_2(t_2)$ 也是公式



设 $\varphi(t)$ 是公式， t 是 $\varphi(t)$ 中的元组变量，
则 $(\exists t) \varphi(t)$ ， $(\forall t) \varphi(t)$ 也是公式



有限次使用上述规则得到的式子
都是公式



元组关系演算（续）

$\varphi(t)$ 中运算符的优先级

公式运算符

- ◆ 算术比较符： $<, >, \leq, \geq, \neq, =$
- ◆ 存在量词 \exists 和全称量词 \forall
- ◆ 逻辑运算符： $\neg, \wedge, \vee, \rightarrow$

高
↓
低



元组关系演算（续）

R

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

S

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

A	B	C
3	4	6
5	6	9

$\{ t \mid S(t) \wedge t[A] > 2 \}$

A	B	C
4	5	6
7	8	9

$\{ t \mid R(t) \wedge \neg S(t) \}$



元组关系演算（续）

R

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

S

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

$$\{ t \mid (\exists u) (S(t) \wedge R(u) \wedge t[C] < u[B]) \}$$

A	B	C
1	2	3
3	4	6



R 元组关系演算 (续) S

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

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

$$\{ t \mid (\forall u)(R(t) \wedge S(u) \wedge t[C] > u[A]) \}$$

A	B	C
4	5	6
7	8	9



元组关系演算（续）

R

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

S

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

$$\{ t \mid (\exists u)(\exists v)(R(u) \wedge S(v) \wedge u[A] > v[B] \wedge t[A] = u[B] \wedge t[B] = v[C] \wedge t[C] = u[A]) \}$$

R.B	S.C	R.A
5	3	4
8	3	7
8	6	7
8	9	7



元组关系演算（续）

- 表达式的安全性
 - 元组关系演算有可能会产生无限关系，这样的表达式是不安全的
如 $\{t \mid \neg(t \in R)\}$ ，求所有不在 R 中的元组
 - 引入公式 P 的域概念，用 $dom(P)$ 表示
 $dom(P) =$ 显式出现在 P 中的值 + 在 P 中出现的关系的元组中出现的值
（不必是最小集）
 - 如果出现在表达式 $\{t \mid P(t)\}$ 结果中的所有值均来自 $dom(P)$ ，则称 $\{t \mid P(t)\}$ 是安全的



元组关系演算（续）

R

A	B
A1	B1
A1	B2
A2	B3

$\{ t \mid \neg (t \in R) \}$

A	B
A1	B3
A2	B1
A2	B2

$dom(\neg (t \in R)) = \{\{A1, A2\}, \{B1, B2, B3\}\}$

S

A	B
A1	B1
A1	B2
A2	B3
A1	B3
A2	B1
A2	B2



思考

➤图为学生-课程数据库中的student关系、Course关系、SC关系

SC:

Sno	Cno	Grade
95001	1	92
95001	2	85
95001	3	88
95002	2	90
95002	3	

Student:

Sno	Sname	Ssex	Sage	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:

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

- 查询信息系(IS)的全体学生
- 查询年龄小于20岁的学生
- 查询学生的姓名和所在系



- 查询信息系(IS)的全体学生

$$\sigma_{Sdept='IS'}(S)$$
$$\{t \mid S(t) \wedge t[Sdept] = 'IS'\}$$

- 查询年龄小于20岁的学生

$$\sigma_{Sage < 20}(S)$$
$$\{t \mid S(t) \wedge t[Sage] < 20\}$$

- 查询学生的姓名和所在系

$$\Pi_{Sname, Sdept}(S) \text{ 或 } \Pi_{2,5}(S)$$
$$\{t \mid (\exists u)(S(u) \wedge t[1] = u[Sname] \wedge t[2] = u[Sdept])\}$$



- 查2号课程的学生学号和成绩

$\Pi_{\text{Sno, Grade}}(\sigma_{\text{Cno}='2'}(\text{SC}))$

$\{t \mid (\exists u)(\text{SC}(u) \wedge u[\text{Cno}] = '2' \wedge t[1] = u[\text{Sno}]$
 $\wedge t[2] = u[\text{Grade}])\}$

- 查选修2号课程的学生姓名和成绩



- 查选修2号课程的学生姓名和成绩

$$\Pi_{\text{Sname,Grade}}(\sigma_{\text{Cno}='2'}(S \bowtie SC))$$

$$\{t \mid (\exists u) (\exists v)(S(u) \wedge SC(v) \wedge v[\text{Cno}]='2' \\ \wedge u[\text{Sno}] = v[\text{Sno}] \\ \wedge t[1] = u[\text{Sname}] \\ \wedge t[2] = v[\text{Grade}])\}$$

- 查选修“数学”的学生学号、姓名及该课程的成绩



- 查选修“数学”的学生学号、姓名及该课程的成绩

$\Pi_{\text{Sno}, \text{Sname}, \text{Grade}} (\sigma_{\text{Cname} = \text{'数学'}} (S \bowtie SC \bowtie C))$

$\{t \mid (\exists u) (\exists v) (\exists w) (S(u) \wedge SC(v) \wedge C[w]$

$\wedge w[\text{Cname}] = \text{'数学'}$

$\wedge u[\text{Sno}] = v[\text{Sno}] \wedge v[\text{Cno}] = w[\text{Cno}]$

$\wedge t[1] = u[\text{Sno}] \wedge t[2] = u[\text{Sname}] \wedge t[3] = v[\text{Grade}]) \}$



- 查选修1号或2号课程的学生学号

$$\Pi_{\text{Sno}}(\sigma_{\text{Cno}='1' \vee \text{Cno}='2'}(\text{SC}))$$

$$\{t \mid (\exists u) (\text{SC}(u) \wedge (u[\text{Cno}]='1' \vee u[\text{Cno}] = '2') \\ \wedge t[1] = u[\text{Sno}]) \}$$

- 查至少选修1号和2号课程的学生学号



- 查至少选修1号和2号课程的学生学号

$\Pi_{Sno}(\sigma_{1=4 \wedge 2='1' \wedge 5='2'}(SC \times SC))$

$\{t \mid (\exists u) (\exists v) (SC(u) \wedge SC(v) \wedge u[Cno]='1' \wedge v[Cno]='2' \wedge u[Sno] = v[Sno] \wedge t[1] = u[Sno])\}$

SC

Sno	Cno	Grade
95001	1	92
95001	2	85
95001	3	88
95002	2	90
95002	3	80

SC

Sno	Cno	Grade
95001	1	92
95001	2	85
95001	3	88
95002	2	90
95002	3	80

Sno1	Cno1	Grade1	Sno2	Cno2	Grade2
95001	1	92	95001	1	92
95001	1	92	95001	2	85
95001	1	92	95001	3	88
95001	1	92	95002	2	90
95001	1	92	95002	3	80



域关系演算

域关系演算的定义

定义 以元组中的域为单位，按照谓词公式所约束的条件查询所需的元组，表示为：

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

其中 x_1, x_2, \dots, x_n 代表域变量，即元组的分量， R 代表由原子构成的公式。

R 的定义如元组关系演算，同样是反复由原子公式、自由变量、约束变量和运算符构成。



域关系演算（续）

R

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

S

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

W

D	E
7	5
4	8

$$R1 = \{ x \ y \ z \mid R(x, y, z) \wedge x < 5 \wedge y > 3 \}$$

A	B	C
4	5	6



域关系演算（续）

R			S			W	
A	B	C	A	B	C	D	E
1	2	3	1	2	3	7	5
4	5	6	3	4	6	4	8
7	8	9	5	6	9		

$$R2 = \{ x \ y \ z \mid (\exists x) (\exists y) (\exists z) (R(z, x, u) \wedge W(y, v) \wedge u > v) \}$$

RB	WD	RA
5	7	4
8	7	7
8	4	7



域关系演算（续）

R			S			W	
A	B	C	A	B	C	D	E
1	2	3	1	2	3	7	5
4	5	6	3	4	6	4	8
7	8	9	5	6	9		

$$R2 = \{xyz \mid (\exists x) (\exists y) (\exists z) R(x, y, z) \vee (S(x, y, z) \wedge y=4)\}$$

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



- 在八种关系代数运算中,并、差、笛卡儿积、投影和选择五种运算为基本的运算.其他三种运算,即交、连接和除,均可以用五种基本运算来表达.

交运算: $R \cap S = R - (R - S)$

连接运算: $R \bowtie_{A \theta B} S = \sigma_{A \theta B}(R \times S)$

除运算: $R(X,Y) \div S(Y,Z) = \pi_x(R) - (\pi_x(\pi_x(R) \times \pi_y(S)) - R)$



$$R \div S = \Pi_X(R) - (\Pi_X(\Pi_X(R) \times \Pi_Y(S) - R))$$

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

C	D
c	d
e	f

A	B
a	b
b	c
e	d

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

$$\Pi_{AB}(R) \times \Pi_{CD}(S) - R$$

A	B	C	D
b	c	c	d

$$R \div S =$$

A	B
a	b
b	c
e	d

-

A	B
a	b
b	c

=

A	B
a	b
e	d