

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



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:

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

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



R

X	Y

S

Y	Z

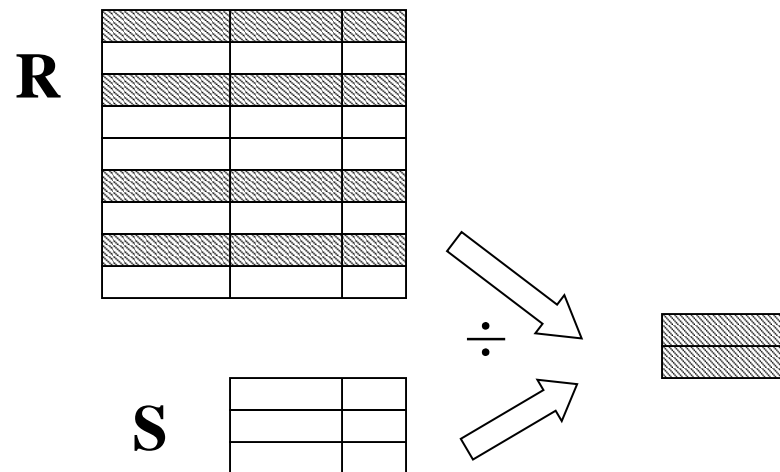
$\Pi_y(S)$

$R \div S \quad Y_x$



除(续)

- 2) 除操作是同时从行和列角度进行运算



除(续)

■ 计算 $R \div S$

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

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 :

然后求: $\pi_{Sno, Cno}(SC) \div K$

Cno
1
3



综合举例(续)

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

95001象集{1, 2, 3}

95002象集{2, 3}

$K=\{1, 3\}$

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

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

- 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)

SC:

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

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

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



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



- 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

$$\Pi_{Sname, Sdept}(S) \text{ 或 } \Pi_{2,5}(S)$$

$$\{t \mid (\exists u)(S(u) \wedge t[1] = u[Sname] \wedge t[2] = u[Sdept])\}$$

SC:

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

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

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

SC:

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

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

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

- 查2号课程的学生学号和成绩

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

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

- 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

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

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

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

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

- 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

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

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

- 查至少选修1号和2号课程的学生学号

SC:

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

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

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



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: $\mathbf{R} \cap \mathbf{S} = \mathbf{R} - (\mathbf{R} - \mathbf{S})$

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

除运算: $\mathbf{R}(\mathbf{X}, \mathbf{Y}) \div \mathbf{S}(\mathbf{Y}, \mathbf{Z}) = \pi_{\mathbf{x}}(\mathbf{R}) - (\pi_{\mathbf{x}}(\pi_{\mathbf{x}}(\mathbf{R}) \times \pi_{\mathbf{y}}(\mathbf{S}) - \mathbf{R}))$



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

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

补充举例

- 职工关系EMP(E# ,ENAME,AGE,SEX,ECITY)
 - 工作关系WORKS(E#,C#,SALARY)
 - 公司关系COMP(C#,CNAME,CITY)
-
- 1检索超过50岁的男职工的工号和姓名
 - 2检索不在联华公司工作的职工工号和姓名
 - 3检索居住城市和公司所在城市相同的职工工号和姓名
 - 4工号为E6的职工在多个公司工作,试检索至少在E6职工兼职的所有公司工作的职工工号

