

数据库系统原理

引用中国人民大学信息学院原版PPT

华中科技大学计算机学院左琼修改版

School of Computer Science and Technology , HUST
2020



第二章 关系数据库

2.1 关系数据结构及形式化定义

2.2 关系操作

2.3 关系的完整性

2.4 关系代数

2.5 关系演算

2.6 小结

案例：教学数据库有三个关系

问题：如何表示对该数据库的各种操作？

- 查询选修了“数据库”课程的学生姓名。
- 查询学习了1号课程但没学5号课程的学生学号和姓名。
- 查询选修了全部课程的学生学号。

S

<u>Sno</u>	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

C

<u>Cno</u>	Cname	先行课号 <u>Cpno</u>	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

SC

<u>Sno</u>	<u>Cno</u>	Grade
95001	1	92
95001	2	85
95001	3	88
95002	2	90
95002	3	80

2.4 关系代数

- **关系代数**——将关系作为运算单位(操作数), 用关系代数表达式表示的运算方法。

运算对象: 关系

运算结果: 关系

关系操作: 按运算符的不同主要分为两类:

- **传统的集合运算**: 把关系看成元组的集合, 从行的角度进行运算, 包括并、差、交和笛卡尔积等。
- **专门的关系运算**: 不仅从行的角度, 也从列的角度进行运算, 是为数据库的应用而引进的特殊运算。包括选择、投影、连接和除法等。

2.4.1 传统的集合运算

传统的集合运算是**二目**运算。

并不是任意的两个关系都能进行这种集合运算。

- 除笛卡尔积外，要求参加运算的关系必须具备**相容性**。

定义：设给定两个关系R、S，若满足：

- (1) 具有相同的**度n**，
- (2) R中第*i*个属性和S中第*i*个属性来自**同一个域**，则说关系R、S是**相容的**。

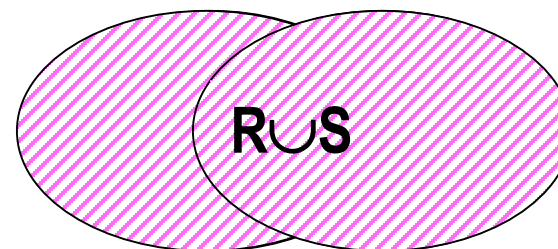
表 关系代数运算符

运算符		含义		运算符		含义	
集合运算符	U	并 差 交 笛卡尔积	比较运算符	>		大于	
	-			≥		大于等于	
	∩			<		小于	
	×			≤		小于等于	
				=		等于	
				<		不等于	
				>			

1. 并 (Union)

- 设两个关系 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 和 S 的并 (记为: $R \cup S$)
 - 仍为 n 目关系, 由属于 R 或属于 S 的元组组成

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



1. 并

例:

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

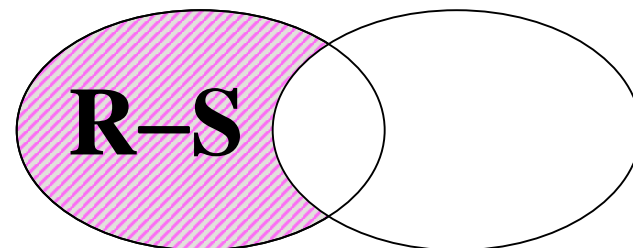
RUS	A	B	C
	a1	b1	c1
	a1	b2	c2
	a2	b2	c1
	a1	b3	c2

实例:

- 选修了1号或者2号课程的学生选课记录。

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



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

思考题 如何用差运算求补集？

已知所有选过课的学生和全校学生，查询那些没有选过课的学生？

2. 差

■ 例:

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-S	A	B	C
	a1	b1	c1

实例:

——选修了1号课程但没选2号课程的学生选课记录。

3. 交 (Intersection)

- R和S

- 具有相同的目 n
- 相应的属性取自同一个域

- $R \cap S$

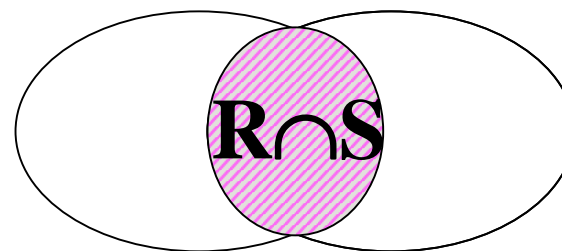
- 仍为 n 目关系，由既属于 R 又属于 S 的元组组成

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

- 而交运算为**非**基本运算，**不属于**最小操作完备集中的操作，可以用差运算来表示：

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

或 $R \cap S = S - (S - R)$



3. 交

例:

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

$R \cap S$	A	B	C
	a1	b2	c2
	a2	b2	c1

实例:

——既选修了1号课程又选修了2号课程的学生选课记录。

课堂练习:

- 设R和S同为相容的k元关系，R有m个元组，S有n个元组，则关于 $R \cap S$ ，以下论述错误的是（ ）
 - A. 等于 $R - (R - S)$
 - B. 等于 $S - (S - R)$
 - C. 最多有m个元组
 - D. 最少有0个元组

4. 笛卡尔积 (Cartesian Product)

- **应用需求** 应用程序中需查询来自两张表的信息，系统如何解决？将两张表合并为一张表。
- 严格地讲是广义笛卡尔积 (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**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**
 $R \times S = \{t_r t_s \mid t_r \in R \wedge t_s \in S\}$
- **作用**: 将两个关系无条件的连接成一个新关系，可用于两关系的连接操作。

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

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。

2.4.2 专门的关系运算

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

若 $A = \{A_{i1}, A_{i2}, \dots, A_{ik}\}$, 其中 $A_{i1}, A_{i2}, \dots, A_{ik}$ 是 A_1, A_2, \dots, A_n 中的一部分, 则 A 称为属性列或属性组。

\overline{A} 则表示 $\{A_1, A_2, \dots, A_n\}$ 中去掉 $\{A_{i1}, A_{i2}, \dots, A_{ik}\}$ 后剩余的属性组。

$t[A] = (t[A_{i1}], t[A_{i2}], \dots, t[A_{ik}])$ 表示元组 t 在属性列 A 上诸分量的集合。

2.4.2 专门的关系运算

(3) 元组的连接 $\overbrace{t_r t_s}$

R 为 n 目关系, S 为 m 目关系,

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

$\overbrace{t_r t_s}$ 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。

2.4.2 专门的关系运算

(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：

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

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。

2.4.2 专门的关系运算

R

x1	z1
x1	z2
x1	z3
x2	z2
x2	z3
x3	z1
x3	z3

象集举例

- x_1 在 R 中的象集:

$$Z_{x1} = \{z1, z2, z3\},$$

- x_2 在 R 中的象集:

$$Z_{x2} = \{z2, z3\},$$

- x_3 在 R 中的象集:

$$Z_{x3} = \{z1, z3\}$$

从R中选出在X上取值为x的元组，去掉X上的分量，只留Z上的分量

例1

Z

X

学号	姓名	性别	系别
0101	张	男	CS
0102	李	女	CS
0203	赵	男	MA
0103	吴	女	CS

关系模式: 学生(学号,姓名,性别,系别)

元组t: (0102,李,女,CS)

t[性别]: 女

属性列X: {性别,系别}

t[性别,系别]: (女,CS)

属性组Z: {学号,姓名}

t[X] = (女,CS)

X = (女,CS), Zx = ?

CS系全部女生的学号,姓名

例2

X

Z

姓名	课程
张蕊	物理
王红	数学
张蕊	数学

x=张蕊

Z_x

课程
数学
物理

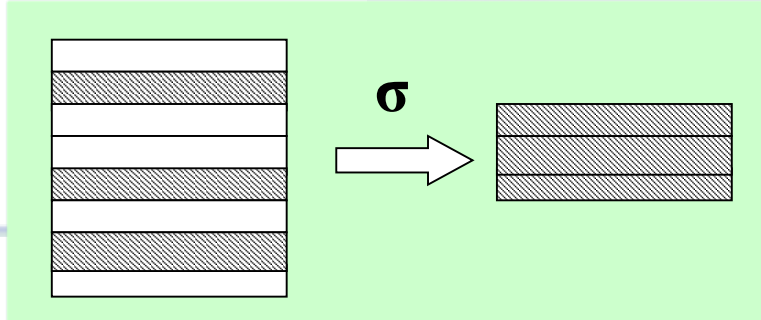
张蕊同学所选修的全部课程

2.4.2 专门的关系运算

表 关系代数运算符

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			

1. 选择 (Selection)



- 选择操作是根据某些条件对关系做**水平分割**，即选取符合条件的元组构成结果关系，又称为**限制 (Restriction)**。
- 关系R关于公式F的选择记作：

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

其中：σ为选择运算符，F为条件表达式，σ_F(R)表示从R中挑选满足公式F的元组所构成的关系。

F：是一个逻辑表达式，基本形式为： **$x_1 \theta y_1$**

F的组成：

- ❖ 运算对象：属性，常数(如数字)
- ❖ 运算符：算术运算符 ($>, \geq, =, <, \leq, \neq$) ,
逻辑运算符 (\wedge, \vee, \neg)

1. 选择

[例1]

R

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

$\sigma_{A>1 \wedge B=2}(R)$

A	B	C
2	2	3

上式也可写作: $\sigma_{[1]>1 \wedge [2]=2}(R)$

[例2] 在S(Sno,Sname,Ssex,Sage,Sdept)上查询年龄小于20岁的学生

$\sigma_{Sage < 20}(S)$ 或 $\sigma_{[4] < 20}(S)$

Sno	Sname	Ssex	Sage	Sdept
201415122	刘晨	女	19	IS
201415123	王敏	女	18	MA
201415125	张立	男	19	IS

1. 选择

例3：用关系表达式表达下列查询

找前页关系S中计算机系(代号: 'CS')全部的男生

$\sigma_{Sdept='CS' \wedge Ssex='男'}(S)$

Sno	Sname	Ssex	Sage	Sdept
201515121	张晨	男	19	CS
201515123	李敏	男	18	CS
201515127	何立	男	19	CS

2. 投影 (Projection)

- 关系R上的投影是从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为R中的属性列，可用列的属性名或列在关系中的序号表示。

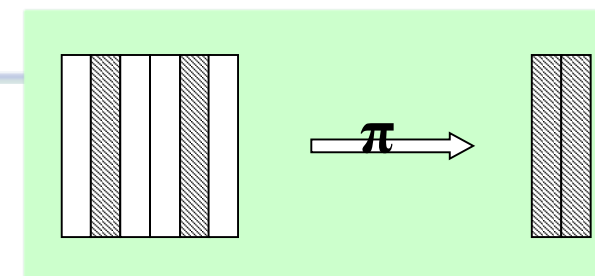
- 特征：

- 1) 在单个关系上进行
- 2) 从列的角度进行运算
- 3) 投影的列可按自己的要求的顺序排列

- 作用：在关系中选择某些需要的列，并按要求组成一个新关系。

2. 投影

- 投影操作主要是从列的角度进行运算。
- 投影的结果中要去掉相同的行（避免重复行）。Why?
 - 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。



例：求关系 R 在 A、C 两列上的投影。

解：关系代数表达式为： $\pi_{A,C}(R)$ 或 $\pi_{[1],[3]}(R)$

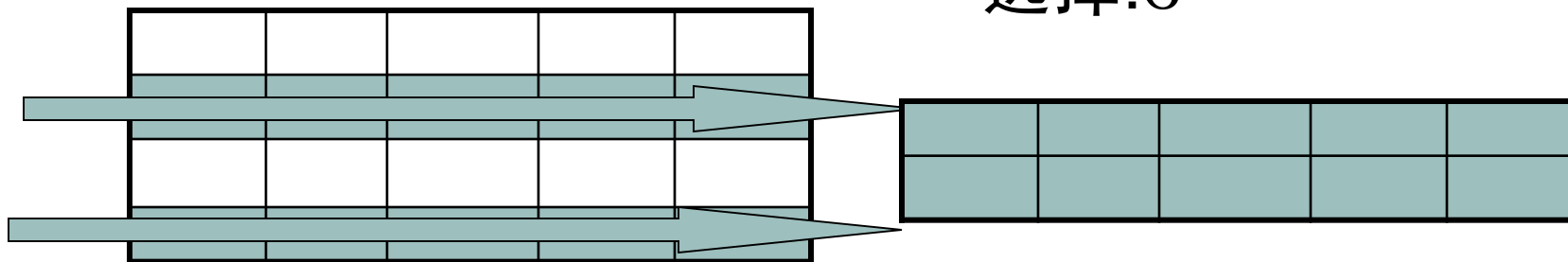
R		
A	B	C
1	2	3
4	5	6
2	2	3

$\pi_{A,C}(R)$	
A	C
1	3
4	6
2	3

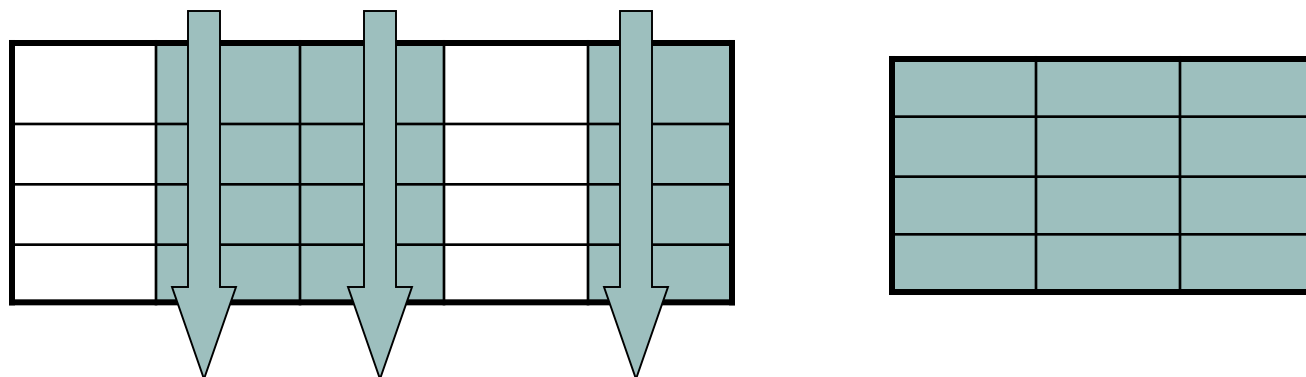
$\pi_{B,C}(R)$	
B	C
2	3
5	6

选择与投影的区别

选择: σ



投影: Π



投影与选择

■ 例1:

■ 给出所有学生的姓名和年龄

$\Pi_{\text{Sname, Sage}}(\text{S})$

■ 找001号学生所选修的课程号

$\Pi_{\text{C\#}}(\sigma_{\text{S\#} = '001'}(\text{SC}))$

S(S#,Sname,Sage)

Course(C#,Cname)

SC(C#,S#,Score)

投影与选择

- 复合运用投影、选择、笛卡尔运算，可以从任意n张表中截取满足条件的子表

例2:

列出CS系和MA系学生的学号和姓名。

方案1:

$$\Pi_{SNO, SNA}(\sigma_{DEPT = 'CS' \vee DEPT = 'MA'}(S))$$

方案2:

$$\Pi_{SNO, SNA}(\sigma_{DEPT = 'CS'}(S)) \cup \Pi_{SNO, SNA}(\sigma_{DEPT = 'MA'}(S))$$

S

SNO	SNA	SEX	DEPT
0101	张	男	CS
0102	李	女	CS
0203	赵	男	MA
0103	吴	女	CS

3. 连接 (Join)

- 连接也称为 θ 连接;
- 连接运算是从两个关系的笛卡尔积中选取满足连接条件的元组, 记作:

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

- 其中:
 - A和B分别为R和S上度数相等且可比的属性组。
 - θ 是比较运算符 ($> \geq = < \leq \neq$) 。
 - 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

3. 连接

- 2类常用连接运算:
- 等值连接(equijoin): θ 为 “=” 的连接运算。

两个关系参加运算,
不一定有公共属性

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

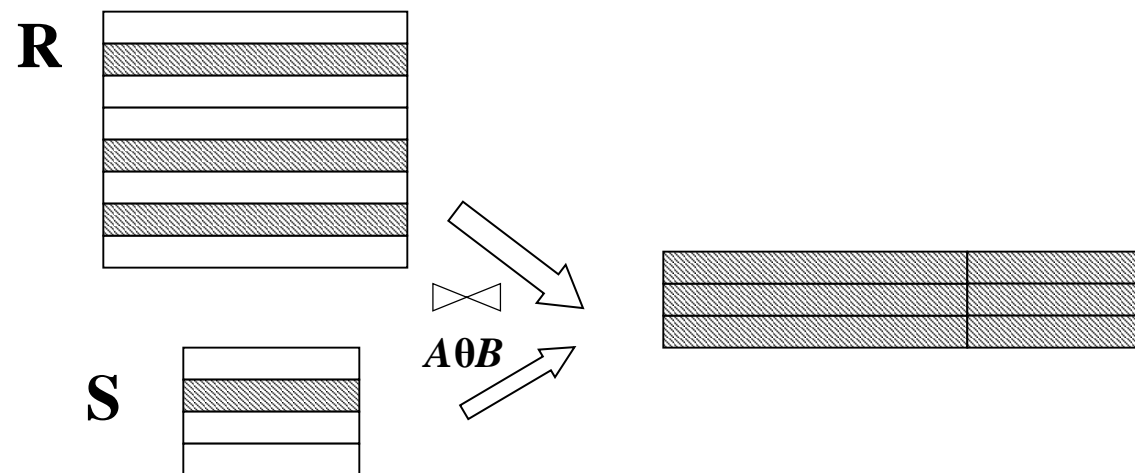
- 含义: 从关系 R 与 S 的广义笛卡尔积中选取 A 、 B 属性值相等的那些元组。
- 自然连接 (natural join): 是一种特殊的等值连接。
 - 两个关系中进行比较的分量必须是相同的属性组
 - 在结果中把重复的属性列去掉

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

- R 和 S 具有相同的属性组 B
- 当 R 与 S 无相同属性时, $R \bowtie S = R \times S$

3. 连接

- 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。
- 自然连接的本质是将两张有关联的表，按照元组之间在**属性B（外码）上的等值关系**，合并为一张表。

3. 连接

运算步骤:

1) 求笛卡尔积 $R \times S$

2) 选择其中满足 $A \theta B$ 的元组 $R \bowtie S$

R

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

S

C	D
a	7
b	8

A	B	R.C	S.C	D
1	2	3	a	7
1	2	3	b	8
4	5	6	a	7
4	5	6	b	8
7	8	9	a	7
7	8	9	b	8

$R \bowtie_{[3]>[2]} S$

或

$R \bowtie_{C>D} S$

A	B	R.C	S.C	D
7	8	9	a	7
7	8	9	b	8

自然连接

R

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

S

B	E	F
b	c	f
g	h	i

1) 计算 $R \times S$

A	R. B	C	S.B	E	F
a	b	c	b	c	f
a	b	c	g	h	i
b	a	f	b	c	f
b	a	f	g	h	i
c	b	d	b	c	f
c	b	d	g	h	i

A	R. B	C	S.B	E	F
a	b	c	b	c	f
c	b	d	b	c	f

2) 选择 $R.B = S.B$ 的元组

3) 去掉重复属性

A	B	C	E	F
a	b	c	c	f
c	b	d	c	f

3. 连接 (续)

- 问题：自然连接会丢失信息，需引入新的连接运算
例如：student \bowtie sc会将一个未选课的学生丢失

- 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**)。

3. 连接例子

一般连接 $R \bowtie_{C < E} S$ 的结果如下:

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

$R \bowtie_{C < E} S$

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

3. 连接例子

等值连接 $R \bowtie_{R.B=S.B} S$ 的结果
如下:

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

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

3. 连接例子

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

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

3. 连接例子

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

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

3. 连接例子

关系 R 和关系 S 的左、右外连接:

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

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

外连接运算

- 外连接 外连接运算是扩展运算，可以用其它运算代数表达式表示。

- 例如 左外连接可以写成

$$(r \bowtie s) \cup (r - \pi_R(r \bowtie s)) \times \{(\text{null}, \dots, \text{null})\}$$

- 分析 问题的关键是如何求r与s的自然连接后丢失的r中元组？

利用减法运算求补集

表达式 $r - \pi_R(r \bowtie s)$ 为所需丢失元组

- 思考题 写出右外连接和全连接的代数式

关系运算综合举例

- 常用的代数思维解决方法:

- 1) 整体法

首先分析所需信息来自哪些表;

其次用适当的连接运算合并表;

再用选择运算 σ_P 行分解表, 通过P去除无用元组

最后用投影运算 π_A 列分解表, 通过A选择所需结果

- 2) 分步法

首先将问题分解为多个简单步骤 (可用单表解决)

其次对最里层的问题用一个代数表达式表示结果

再将结果作为已知值, 代入上一层步骤中

注 分步法也可以是从外层向里层的迭代过程

案例：教学数据库有三个关系：

C

<u>Cno</u>	Cname	先行课号 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

S

<u>Sno</u>	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

<u>Sno</u>	<u>Cno</u>	Grade
95001	1	92
95001	2	85
95001	3	88
95002	2	90
95002	3	80

问题：如何表示对该数据库的各种操作？

- 查询选修了“数据库”课程的学生姓名。
- 查询学习了1号课程但没学5号课程的学生学号和姓名。
- 查询选修了全部课程的学生学号。

关系运算综合举例

例 查询选修了一门其直接先行课为5号课程的学生姓名

整体法

问题分析 该查询是查找选过...课程的学生，所以先将学生表、选修表、课程表合并

$\text{Course} \bowtie \text{SC} \bowtie \text{Student}$

再确定选择运算谓词P为 $\text{cpno}=5$

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

最后投影所需的学生姓名即可

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

关系运算综合举例

解法2：分步法

首先找出先修课为5的课程

$\sigma_{Cpno='5'}(Course)$

再找出选过上表达式结果的选课元组中学号

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

最后从学生表中找出学号为上一步结果的学生

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

注 上述例子中我们通过自然连接实现了选择运算的功能能否用
嵌套方法实现选择? 例如: 第二步改为:

$\sigma_{cno = \pi_{cno}(\sigma_{Cpno='5'}(Course))}(SC) \quad ?$

Student (Sno, Sname, Ssex, Sage, Sdept)

Course (Cno, Cname, Cpno, Ccredit)

SC (sno, cno, grade)

[上例] 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

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

4. 除 (Division)

- **引入动机** 在查询中，经常需要查询包含短语“**所有的**”这样的查询。
- 例如，找出**选过**学分为3分的**所有**课程的学生？
- **如何解决：**
 - 1) 找出学分为3分的**所有**课程；
 - 2) 从选课表中，找出学生，其所选课程**包含**1) 中结果；

解 令 $S = \pi_{cno} (\sigma_{Ccredit=3}(Course))$

对选课表按照Sno分组 $_{sno}G(SC)$

令 $S' =$ 分组以后的由Cno分组构成的表；若 S' 包含 S ，则将这样的学生放入结果集

记作 $SC \div \pi_{cno} (\sigma_{Ccredit=3}(Course))$

- **语义** $R \div S$ 是指从 R 中**去除**哪些不包含 S 的元组，即：从 R 中查找“**选过所有的 S** ”的查询

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 在 R 中的象集, $x = t_r[X]$

- 除操作是同时从行和列角度进行运算

4. 除

为方便起见，我们假设 S 的属性为 R 中后 s 个属性。

$R \div S$ 的具体计算过程如下：

$$1) T = \pi_{1,2,\dots,r-s}(R)$$

$$2) W = (T \times S) - R$$

$$3) V = \pi_{1,2,\dots,r-s}(W)$$

$$4) R \div S = T - V$$

$$\text{即 } R \div S = \pi_{1,2,\dots,r-s}(R) - \pi_{1,2,\dots,r-s}((\pi_{1,2,\dots,r-s}(R) \times S) - R)$$

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

S

C	D
c	d
e	f

1) $T = \pi_{A,B}(R)$

A	B
a	b
b	c
e	d

2) $W = (T \times S) - R$

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

3) $V = \pi_{A,B}(W)$

A	B
b	c

4) $R \div S = T - V$

A	B
a	b
e	d

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

课程
数学
物理

×

姓名
张军
王红

=

姓名	课程
张军	物理
王红	数学
张军	数学
王红	物理

所有学生选修全部课程

姓名	课程
张军	物理
王红	数学
张军	数学
王红	物理

—

姓名	课程
张军	物理
王红	数学
张军	数学

=

姓名
王红

没有选修全部课程的学生

姓名
王红
张军

—

姓名
王红

=

姓名
张军

选修了全部课程的学生

除——分析

例：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
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)

S

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)

除——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, $R \div S = \{a_1\}$

姓名	课程	成绩
张军	物理	88
王红	数学	80
张军	数学	90

÷

课程
数学
物理

=

姓名
张军

选修了全部课程
的学生的姓名

除

$\Pi_{\text{姓名, 课程}}(R) \div S$

R

姓名	课程	成绩
张军	物理	88
王红	数学	80
张军	数学	90

S

课程
数学
物理

÷

$\Pi_y(S) = \{\text{课程}\}, Y_x = \{\text{课程}\},$
 $X = \{\text{姓名, 成绩}\}$

$R \div S = ?$

X	Y

Y	Z

$\Pi_y(S) \subseteq Y_x$

$\{\text{张军}, 88\} = \{\text{物理}\},$

$\{\text{王红}, 80\} = \{\text{数学}\},$

$\{\text{张军}, 90\} = \{\text{数学}\}$

没有哪个X的象集包含了
 $\{\text{物理}, \text{数学}\}$, 所以答
案应该是空集!

除运算示例

例 查询选修了全部课程的学生号码和姓名。

解 与上题类似

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

是选过全部课程的学生号

将其与Student自然合并，为所需结果

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

例 查询选过学分为3分的所有课程的学生姓名？

$$\pi_{sno, sname, cno} (Student \bowtie SC) \div \pi_{cno} (\sigma_{credit=3} (Course))$$