



计算机科学与技术学院

数据库原理

王兴梅

计算机科学与技术学院

Email: wangxingmei@hrbeu.edu.cn





第二章 关系数据库

2.1 关系模型概述

2.2 关系数据结构

2.3 关系操作

2.4 关系的完整性

2.5 关系代数

2.6 小结



2.5 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



概述

2. 关系代数运算的三个要素

运算对象：关系

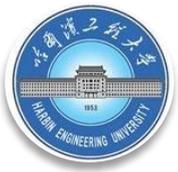
运算结果：关系

运算符：四类



概述

- 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

表2.4 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
				$=$	等于
				\neq	不等于



概述

表2.4 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	σ	选择	逻辑运算符	\neg	非
	π	投影		\wedge	与
	\bowtie	连接		\vee	或
	\div	除			



概述

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



2.4 关系代数

- 概述
- 传统的集合运算
- 专门的关系运算



2.4.1 传统的集合运算

- 并
- 差
- 交
- 广义笛卡尔积



1. 并 (Union)

■ 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

■ $R \cup S$

- 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

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
$a1$	$b1$	$c1$
$a1$	$b2$	$c2$
$a2$	$b2$	$c1$
$a1$	$b3$	$c2$



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 且不属于 S 的所有元组组成

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



差

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

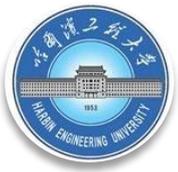
- 具有相同的目 n
- 相应的属性取自同一个域

- $R \cap S$

- 仍为 n 目关系，由既属于 R 又属于 S 的元组组成

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

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



交

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



广义笛卡尔积

R

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

$R \times S$

S

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

RA	RB	RC	SA	SB	SC
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



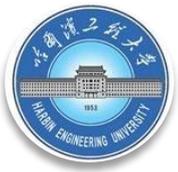
2.4 关系代数

- 概述
- 传统的集合运算
- 专门的关系运算



2.4.2 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, 是一个逻辑表达式, 基本形式为:

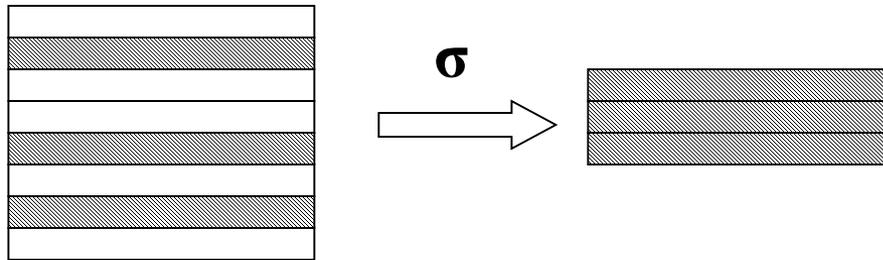
$$[\neg ([X_1 \theta Y_1 [])] [\varphi [\neg ([X_2 \theta Y_2 [])]]] \dots$$

- θ : 比较运算符 ($>$, \geq , $<$, \leq , $=$ 或 $<>$)
- X_1, Y_1 等: 属性名、常量、简单函数; 属性名也可以用它的序号来代替;
- φ : 逻辑运算符 (\neg , \wedge 或 \vee)
- $[]$: 表示任选项
- \dots : 表示上述格式可以重复下去



选择

- 3) 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。



- 4) 举例



有职工表 S，试找出满足条件性别 = “女” 的元组集。

职工编号	姓名	性别	所在部门
1001	张力	女	物资资源
1002	胡爱军	男	市场营销
1003	阿法	女	企划部
1004	曾玲	女	公关部
1005	刘逸飞	男	市场营销

选择条件为：性别 = ‘女’，用选择表示为：

$$\sigma_{\text{性别}='女'}(S)$$



结果如下：

职工编号	姓名	性别	所在部门
1001	张力	女	物资资源
1003	阿法	女	企划部
1004	曾玲	女	公关部

选择

- 对元组进行操作，行变、列一定不变！
- 只作用于单一关系之上！



2. 投影 (Projection)

■ 1) 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

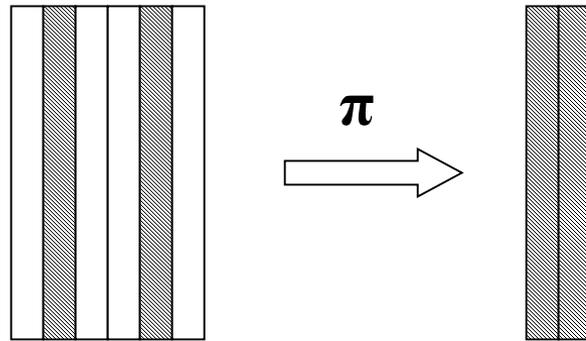
$$\pi_A(R) = \{ t[A] \mid t \in R \}$$

A : R 中的属性列



2. 投影 (Projection)

- 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）



3) 举例

学生表student，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

Student

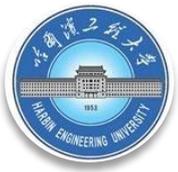
学号 Sno	姓名 Sname	性别 Ssex	年龄 Sage	所在系 Sdept
200215121	李勇	男	20	CS
200215122	刘晨	女	19	IS
200215123	王敏	女	19	MA
200215125	张立	男	18	IS

投影表示为： $\pi_{Sname, Sdept}(Student)$
或 $\pi_{2, 5}(Student)$



结果如下：

Sname	Sdept
李勇	CS
刘晨	IS
王敏	MA
张立	IS



3. 连接 (Join)

- 1) 连接也称为 θ 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

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

- 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 属性组上值满足比较关系的元组。



连接

■ 3) 两类常用连接运算

➤ 等值连接 (equijoin)

■ 什么是等值连接

➤ θ 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：

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



连接

■ 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

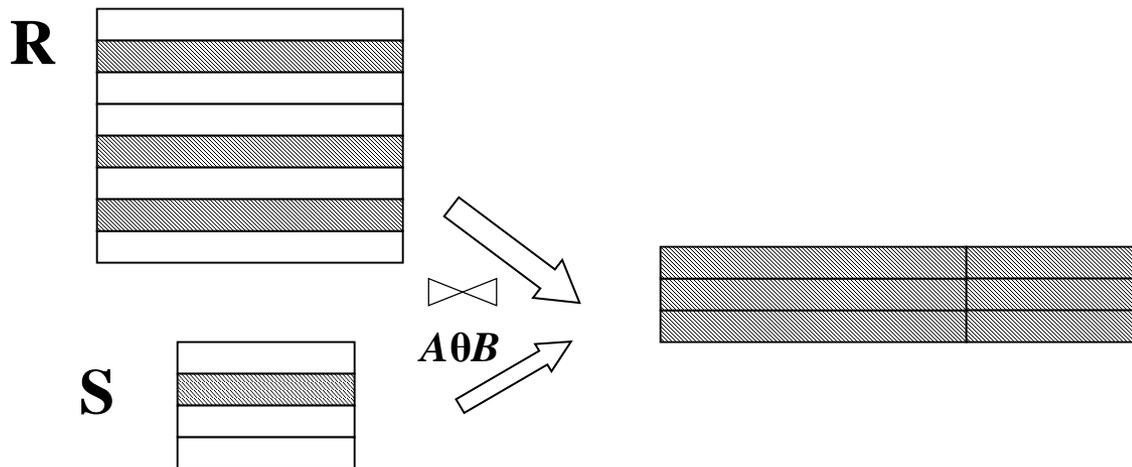
R 和 S 具有相同属性组 B ， U 为 R 和 S 的全体属性集合。

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



连接

- 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

■ 5) 举例

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

R

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

S



连接

$R \underset{C < E}{\bowtie} S$

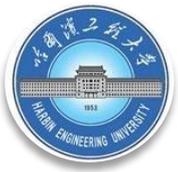
<i>A</i>	<i>R.B</i>	<i>C</i>	<i>S.B</i>	<i>E</i>
<i>a</i> ₁	<i>b</i> ₁	5	<i>b</i> ₂	7
<i>a</i> ₁	<i>b</i> ₁	5	<i>b</i> ₃	10
<i>a</i> ₁	<i>b</i> ₂	6	<i>b</i> ₂	7
<i>a</i> ₁	<i>b</i> ₂	6	<i>b</i> ₃	10
<i>a</i> ₂	<i>b</i> ₃	8	<i>b</i> ₃	10



连接

等值连接 $R \bowtie S$
 $R.B=S.B$

<i>A</i>	<i>R.B</i>	<i>C</i>	<i>S.B</i>	<i>E</i>
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

自然连接 $R \bowtie S$

<i>A</i>	<i>B</i>	<i>C</i>	<i>E</i>
a_1	b_1	5	3
a_1	b_2	6	7
a_2	b_3	8	10
a_2	b_3	8	2



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。

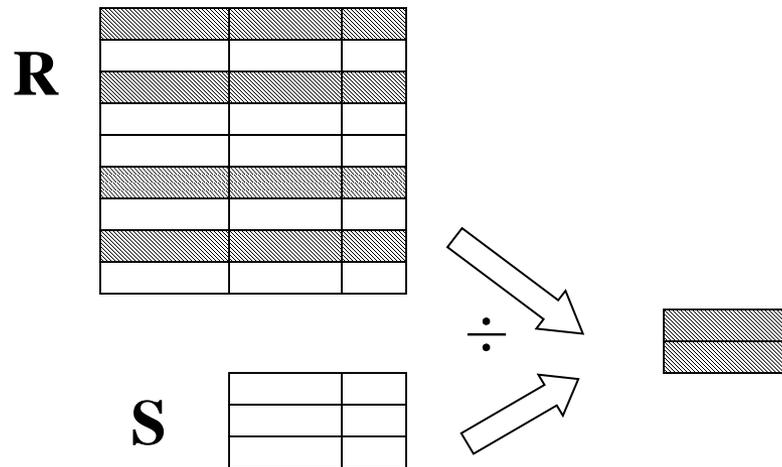
$$R \div S = \{t_r[X] \mid t_r \in R \wedge \pi_Y(S) \subseteq Y_x\}$$

Y_x : x在R中的象集, $x = t_r[X]$

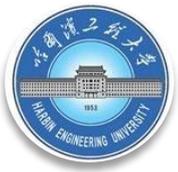


除

- 2) 除操作是同时从行和列角度进行运算



- 3) 举例



除

R

<i>A</i>	<i>B</i>	<i>C</i>
<i>a</i>₁	<i>b</i>₁	<i>c</i>₂
<i>a</i>₂	<i>b</i>₃	<i>c</i>₇
<i>a</i>₃	<i>b</i>₄	<i>c</i>₆
<i>a</i>₁	<i>b</i>₂	<i>c</i>₃
<i>a</i>₄	<i>b</i>₆	<i>c</i>₆
<i>a</i>₂	<i>b</i>₂	<i>c</i>₃
<i>a</i>₁	<i>b</i>₂	<i>c</i>₁

S

<i>B</i>	<i>C</i>	<i>D</i>
<i>b</i>₁	<i>c</i>₂	<i>d</i>₁
<i>b</i>₂	<i>c</i>₁	<i>d</i>₁
<i>b</i>₂	<i>c</i>₃	<i>d</i>₂

<i>R</i> ÷ <i>S</i>
<i>A</i>
<i>a</i>₁



分析:

在关系R中, **A**可以取四个值 $\{a_1, a_2, a_3, a_4\}$

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

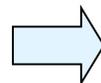
$\{(b_1, c_2), (b_2, c_1), (b_2, c_3)\}$

只有 a_1 的象集包含了**S**在**(B, C)**属性组上的投影

所以 $R \div S = \{a_1\}$



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	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



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	D
b_1	c_2	d_1
b_2	c_1	d_1
b_2	c_3	d_2

$R \div S$

A
a_1



小结

■ 关系代数运算

➤ 关系代数运算

并、差、交、笛卡尔积、投影、选择、连接、除

➤ 基本运算

并、差、笛卡尔积、投影、选择

➤ 交、连接、除

可以用**5**种基本运算来表达

引进它们并不增加语言的能力，但可以简化表达



小结(续)

- 关系代数表达式
 - 关系代数运算经有限次复合后形成的式子
- 典型关系代数语言
 - ISBL (Information System Base Language)
 - 由IBM United Kingdom研究中心研制
 - 用于PRTV (Peterlee Relational Test Vehicle) 实验系统



- 已知三个关系R1, R2, S, 如下图:

X	Y	Z
x	5	k
x	6	y
y	5	k
z	6	y

X	Y	Z
x	3	c
y	5	k
z	6	y
i	5	f
j	6	f

Y	Z
5	k
6	y

- 试求:
- $R1 \times S = ?$
 - $R1 - R2 = ?$
 - $R1 \div S = ?$
 - $\sigma_{Y>5 \vee Z \neq 'f'}(R2) = ?$
 - $\pi_{Y,Z}(R2) = ?$



- 设有数据库 S(Sno,Sname,Sage,Ssex)
- C(Cno,Cname, Cpno) SC(Sno,Cno,Grade)
- 试用关系代数表达式表示一下查询要求：
- 1.查询选修了课程号为C2的学生号；
■ $\pi_{Sno} (\sigma_{Cno='2'} (SC))$
- 2.查询选修数据库课程的女学生的学生号；
■ $\pi_{Sno} (\sigma_{Cname='DB' \wedge Ssex='女'} (S \bowtie C \bowtie SC))$
- 3.查询至少选修C1和C3课程的学生号；
■ $\pi_{Sno, Cno} (SC) \div \pi_{Cno} (\sigma_{Cno='C1' \vee Cno='C3'} (C))$
- 4.查询至少选修一门先行课号为C5的学生姓名；
■ $\pi_{Sname} (\sigma_{Cpno='C5'} (C) \bowtie SC \bowtie \pi_{Sno, Sname} (S))$



设有数据库

EMP(Eno, Ename, age, sex) 职工表

WORK(Eno, Cno, Salary) 工作表

COMP(Cno, Cname, Caddr) 公司表

试用关系代数表达式表示以下查询：

1. 查询超过50岁的男性职工的工号和姓名；
2. 假设每个职工只能在一个公司工作，查询工资超过1000元的男性职工的工号和姓名；
3. 查询在“联华公司”工作的职工工号和姓名；
4. 假设每个职工可在多个公司工作，查询在编号为C4和C8的公司兼职的职工工号和姓名；
5. 查询在“联华公司”工作，工资超过1000元男性的职工工号和姓名。



本章小结

基本内容:

关系数据库的基本概念，关系模型及其描述，关系数据库语言，关系代数。

重要知识点:

- 1.关系模型的三类完整性规则
- 2.关系代数的九个操作

本章题型:

选择，填空，简答题，应用题