数据库系统原理

李瑞轩
华中科技大学计算机学院
http://idc.hust.edu.cn/rxli/

2.1 概述

- E.F.Codd于70年代初提出**关系数据理论**，他因此获得1981年的ACM图灵奖

  1970年提出**关系数据模型**

  E.F.Codd, “A Relational Model of Data for Large Shared Data Banks”, Communications of the ACM, 1970

  之后提出了**关系代数**和**关系演算**的概念

  1972年提出了关系的**第一、第二、第三范式**

  1974年提出了关系的**BC范式**（BCNF）

2.1 概述

- 关系理论是建立在集合代数理论基础上的，有着坚实的数学基础
- 早期代表系统
  - System R： 由IBM研制
  - INGRES： 由加州Berkeley分校研制
- 目前主流的商业数据库系统
  - Oracle， Informix， Sybase， SQL Server， DB2
  - Access， Foxpro， Foxbase
  - 达梦（DM）， 金仓， 神州通用， 南大通用
  - 开源数据库：MySQL

2.1 概述

- 数据结构：二维表
- 关系操作：
  - 查询(Query):
    - 选择(select)、投影(project)、连接(join)
    - 除(divide)、并(union)、交(intersection)
    - 差(difference)
  - 增加(insert)、删除(delete)、修改(updated)
  - 关系代数，关系演算，SQL
- 关系的三类完整性约束：
  - 实体完整性、参照完整性、用户自定义的完整性

2.1 概述

- **关系代数**是用对关系的运算来表达查询要求的方式。

- **关系演算**是用谓词表达查询要求的方式。

  - 按谓词变元的基本对象是元组变量还是域变量分为**元组关系演算**和**域关系演算**

2.1 概述

具体系统中的实际语言

- **SQL**

  介于关系代数和关系演算之间，由IBM公司研制system R时提出

- **QUEL**

  基于Codd提出的元组关系演算语言ALPHA，在INGRES上实现

- **QBE**

  基于域关系演算，由IBM公司研制，在IBM370上实现

2.2 关系数据结构及形式化定义

- **域**（Domain）
  - 一组值的集合，这组值具有相同的数据类型

    如: 整数的集合、字符串的集合、全体学生的集合

- **笛卡尔积**（Cartesian Product）
  - 一组域D_1 , D_2 ,…, D_n的笛卡尔积为:

    $D_1 \times D_2 \times \ldots \times D_n = \{(d_1 , d_2 , \ldots , d_n) \mid d_i \in D_i , i=1,\ldots,n\}$
  - 笛卡尔积的每个元素(d_1 , d_2 , … , d_n)称作一个n-元组（n-tuple）
  - 元组的每一个值d_i叫做一个分量（Component）
  - 若D_i的基数为m_i，则笛卡尔积的基数为 $\prod_{i=1}^{n} m_i$

元组个数

2.2 关系数据结构及形式化定义(续)

- 例：设

  D_1为学生集合（T）= {张群，徐晶，王刚}

  D_2为性别集合（S）= {男，女}

  则$D_1 \times D_2$是个二元组集合，元组个数为3×2，是所有可能的（学生，性别）元组集合

$D_1 \times D_2 =$

T	S
张群	男
张群	女
徐晶	男
徐晶	女
王刚	男
王刚	女

2.2 关系数据结构及形式化定义(续)

- 例：设

  D_1为教师集合（T）= {t_1，t_2}

  D_2为学生集合（S）= {s_1，s_2，s_3}

  D_3为课程集合（C）= {c_1，c_2}

  则$D_1 \times D_2 \times D_3$是个三元组集合，元组个数为2×3×2，是所有可能的（教师，学生，课程）元组集合

- 笛卡尔积可表示为二维表的形式

域(课程集合)

T	S	C
t_1	s_1	c_1
t_1	s_1	c_2
t_1	s_2	c_1
…	…	…
t_2	s_3	c_2

元组

2.2 关系数据结构及形式化定义(续)

▪ 关系

- 笛卡尔积$D_1 \times D_2 \times \ldots \times D_n$的子集叫做在域$D_1$, D_2 ,…, D_n上的关系，用$R(D_1, D_2, \ldots, D_n)$表示
- R是关系的名字，n是关系的度或目
- 关系是笛卡尔积中有意义的子集
- 关系也可以表示为二维表

属性

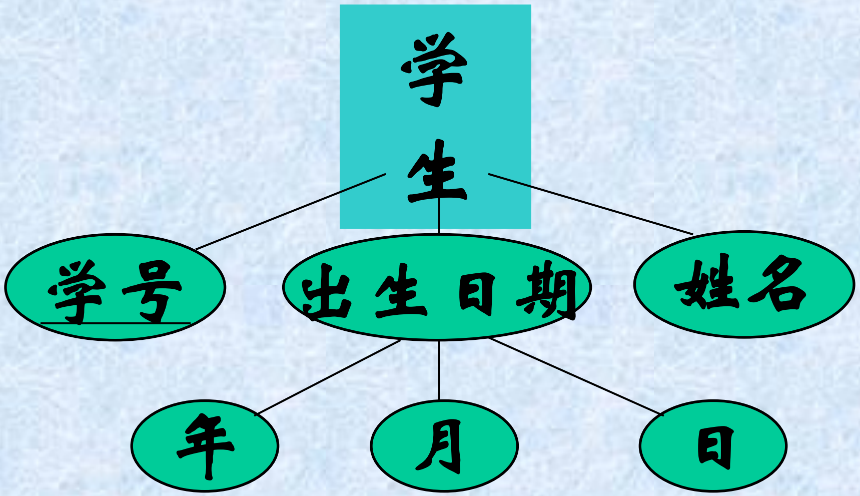
关系TEACH(T, S, C)

T	S	C
t_1	s_1	c_1
t_1	s_1	c_2
t_1	s_2	c_1
t_2	s_3	c_2

元组

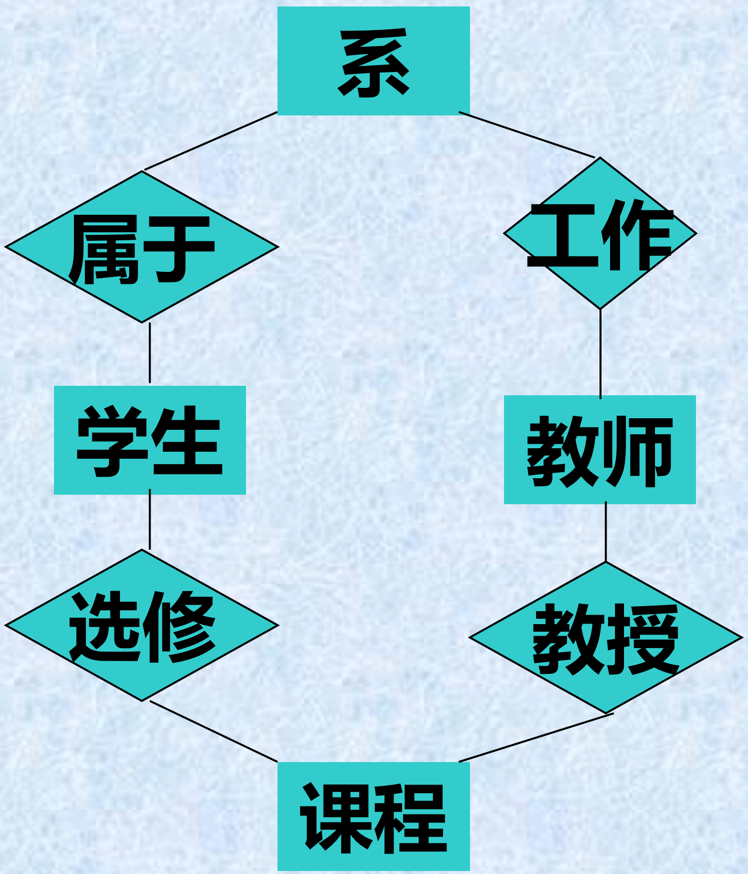
2.2 关系数据结构及形式化定义(续)

- **关系的性质**
  - 列是同质的
  - 行列的顺序无关紧要
  - 任意两个元组不能完全相同
  - 每一分量必须是不可再分的数据
  - 不同的属性，属性名不能相同



数据结构

- 单一的数据结构——关系
- 实体集、联系都表示成关系



DEPT(D# , DN , DEAN)

S(S# , SN , SEX , AGE , D#)

C(C# , CN , PC# , CREDIT)

PROF(P# , PN, D# , SAL)

SC(S# , C# , SCORE)

TEACH(P# , C#)

2.2 关系数据结构及形式化定义(续)

- **候选码**（Candidate Key）
  - 关系中的一个属性组，其值能唯一标识一个元组。若从属性组中去掉任何一个属性，它就不具有这一性质了，这样的属性组称作候选码

  如DEPT中的D#，DN都可作为候选码
  - 任何一个候选码中的属性称作**主属性**

  如SC中的S#，C#
- **主码**（Primary Key）
  - 若一个关系有多个候选码，则选定其中一个作为主码

  如可选定D#作为DEPT的主码
- **外码**（Foreign Key）
  - 关系R中的一个属性组，它不是R的码，但它与另一个关系S的码相对应，则称这个属性组为R的外码

  如S关系中的D#属性

2.2 关系数据结构及形式化定义(续)

- **关系模式**
  - 关系的描述称作关系模式，包括关系名、关系中的属性名、属性向域的映象、属性间的数据依赖关系等，记作$R(A_1, A_2, \ldots, A_n)$
  - 属性向域的映象一般直接说明为属性的类型、长度等
  - 某一时刻对应某个关系模式的内容(元组的集合)称作关系
  - 关系模式是型，是稳定的

  关系是某一时刻的值，是随时间不断变化的

2.2 关系数据结构及形式化定义(续)

- **关系数据库**
  - 其型是关系模式的集合，即数据库描述，称作数据库的内涵(Intension)，或关系数据库模式
  - 其值是某一时刻关系的集合，称作数据库的外延(Extension)，或关系数据库

关系模型的完整性约束

实体完整性

- 关系的**主码**中的属性值不能为空值
- **空值**：不知道或无意义
- 意义：关系对应到现实世界中的实体集，元组对应到实体，实体是相互可区分的，通过主码来唯一标识，若主码为空，则出现不可标识的实体，这是不容许的

学号	姓名	性别	系名
0101	张	男	CS
0102	李	女	CS
0203	赵	男	MA

2.3 关系的完整性(续)

- 参照完整性
  - 如果关系R_2的外码F_k与关系R_1的主码P_k相对应，则R_2中的每一个元组的F_k值或者等于R_1 中某个元组的P_k 值，或者为空值
  - 意义：如果关系R_2的某个元组t_2参照了关系R_1的某个元组t_1，则t_1必须存在
  - 例如，关系S在D#上的取值有两种可能
    - 空值，表示该学生尚未分到任何系中
    - 若非空值，则必须是DEPT关系中某个元组的D#值，表示该学生不可能分到一个不存在的系中

DEPT(<u>D#</u> , DN , DEAN)

S(<u>S#</u> , SN , SEX , AGE , D#)

2.3 关系的完整性(续)

学号	姓名	性别	系名
0101	张	男	CS
0102	李	女	CS
0203	赵	男	MA

学号	课号	成绩
0101	CS145	88
0101	CS148	90
0102	CS180	87
0203	CS145	78

课号	课名
CS145	数据库
CS148	操作系统
CS180	数据结构

2.3 关系的完整性(续)

供应商关系S（主码是“供应商号”）

供应商号	供应商名	所在城市
B01	红星	北京
S10	宇宙	上海
T20	黎明	天津
Z01	立新	重庆

零件关系P（主码是“零件号”，外码是“供应商号”）

零件号	颜色	供应商号
010	红	B01
312	白	S10
201	蓝	T20

今要向关系P中插入新行，新行的值分别列出如下。请问哪些行能够插入？为什么？

A、(‘037’, ‘绿’, null)

B、(null, ‘黄’, ‘T20’)

C、(‘201’, ‘红’, ‘T20’)

D、(‘105’, ‘蓝’, ‘B01’)

E、(‘101’, ‘黄’, ‘T11’)

2.3 关系的完整性(续)

- 用户定义的完整性
  - 用户针对具体的应用环境定义的完整性约束条件
  - 如S#要求是8位整数，SEX要求取值为“男”或“女”
- 系统支持
  - 实体完整性和参照完整性由系统自动支持
  - 系统应提供定义和检验用户定义的完整性的机制

2.4 关系代数

- 属于关系操作的一种
- 关系代数是一种抽象的查询语言
- 通过对关系的运算来表达查询操作
- 运算对象、结果均为关系
- 运算包括四类：
  - 集合运算、关系运算、比较运算、逻辑运算

2.4 关系代数

- 基本运算
  - 一元运算
    - 选择、投影
  - 多元运算
    - 笛卡尔积、并、差
- 其他运算
  - 交、连接、除、赋值
- 扩展运算
  - 广义投影、外连接、聚集
- 修改操作
  - 插入、删除、更新

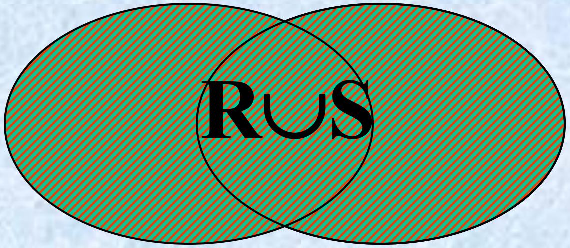
关系代数——运算符

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		$\geqslant$	大于等于
	$\cap$	交		$<$	小于
				$\leqslant$	小于等于
				$=$	等于
				$\neq$	不等于
专门的关系运算符	$\times$	广义笛卡尔积	逻辑运算符	$\neg$	非
	σ	选择		$\wedge$	与
	π	投影		$\vee$	或
	$\bowtie$	连接			
	$\div$	除			

并运算

- 定义
  - 所有至少出现在两个关系中之一的元组集合

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



- 两个关系R和S若进行并运算，则它们必须是相容的：
  - 关系R和S必须是同元的，即它们的属性数目必须相同
  - 对∀i，R的第i个属性的域必须和S的第i个属性的域相同

2.4 关系代数(续)

并运算

R

A	B	C
3	6	7
2	5	7
7	2	3
4	4	3

S

A	B	C
3	4	5
7	2	3

RuS

A	B	C
3	6	7
2	5	7
7	2	3
4	4	3
3	4	5

2.4 关系代数(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

$R \cup S$

A	B	C
a_1	b_1	c_1
a_1	b_2	c_2
a_2	b_2	c_1
a_1	b_3	c_2

2.4 关系代数(续)

R

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

S

A	B	C
b	g	a
b	a	f

T

D	E	F
g	h	i
j	k	l

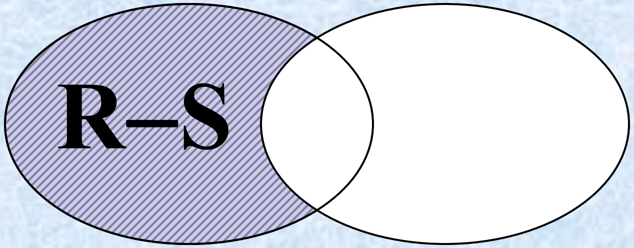
R ∪ S

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

差运算

- 定义
  - 所有出现在一个关系而不在另一关系中的元组集合

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



  - R和S必须是相容的

差运算

R

A	B	C
3	6	7
2	5	7
7	2	3
4	4	3

S

A	B	C
3	4	5
7	2	3

R—S

A	B	C
3	6	7
2	5	7
4	4	3

S—R

A	B	C
3	4	5

2.4 关系代数(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-S

A	B	C
a_1	b_1	c_1

R

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

S

A	B	C
b	g	a
b	a	f

T

D	E	F
g	h	i
j	k	l

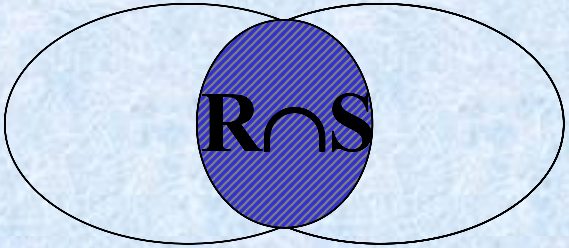
R－S

A	B	C
a	b	c
c	b	d

交运算

- 定义
  - 所有同时出现在两个关系中的元组集合

$$\mathbf{R \cap S = \{ r \mid r \in R \wedge r \in S \}}$$



  - 交运算可以通过差运算来重写

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

交运算

R

A	B	C
3	6	7
2	5	7
7	2	3
4	4	3

S

A	B	C
3	4	5
7	2	3

$R \cap S$

A	B	C
7	2	3

2.4 关系代数(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

$R \cap S$

A	B	C
a_1	b_2	c_2
a_2	b_2	c_1

2.4 关系代数(续)

R

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

S

A	B	C
b	g	a
b	a	f

T

D	E	F
g	h	i
j	k	l

$R \cap S$

A	B	C
b	a	f

广义笛卡尔积

- 元组的连串（Concatenation）
  - 若$r = (r_1, \ldots, r_n)$，$s = (s_1, \ldots, s_m)$，则定义r与s的连串为：

$$\widehat{rs} = (r_1, \ldots, r_n, s_1, \ldots, s_m)$$

- 定义
  - 两个关系R，S，其度分别为n，m，则它们的笛卡尔积是所有这样的元组集合：元组的前n个分量是R中的一个元组，后m个分量是S中的一个元组

$$\mathbf{R \times S = \{ \widehat{rs} \mid r \in R \wedge s \in S \}}$$

  - R×S的度为R与S的度之和， R×S的元组个数为R和S的元组个数的乘积

R

A	B
α	1
β	2

S

C	D	E
α	10	a
β	10	a
β	20	b
γ	10	b

Rx S

A	B	C	D	E
α	1	α	10	a
α	1	β	10	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(续)

R

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

S

A	B	C
b	g	a
b	a	f

T

D	E	F
g	h	i
j	k	l

R × S

R.A	R.B	R.C	S.A	S.B	S.C
a	b	c	b	g	a
a	b	c	b	a	f
b	a	f	b	g	a
b	a	f	b	a	f
c	b	d	b	g	a
c	b	d	b	a	f

2.4 关系代数(续)

- **记号说明**

  给定关系模式$R(A_1, A_2, \ldots, A_n)$，设R是它的一个具体的关系，$t \in R$是关系的一个元组

  - **分量**

    设$t \in R$，则$t[A_i]$表示元组t中相应于属性A_i的一个分量

  - **属性列（或属性组）**

    $A=\{A_{i1}, A_{i2}, \ldots, A_{ik}\} \subseteq \{A_1, A_2, \ldots, A_n\}$，称A为属性列

    $\bar{A}$表示$\{A_1, A_2, \ldots, A_n\}$中去掉A后剩余的属性组

    $t[A] = (t[A_{i1}], t[A_{i2}], \ldots, t[A_{ik}])$

  - **象集**

    给定关系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 关系代数(续)

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[X]:　　(女,CS)

属性组Z:　　{学号,姓名}

t[X] = x = (女,CS)

象集Z_x = ? CS系全部女生的学号,姓名

x = (女,CS),　Z_x = ?

X　　Z

姓名	课程
张蕊	物理
王红	数学
张蕊	数学

x=张蕊　Z_x=?

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

Z_x

课程
数学
物理

张蕊同学所选修的全部课程

选择运算

- 基本定义

在关系R中选择满足给定条件的元组（从行的角度）

$$\sigma_F(R)=\{t \mid t \in R \text{ , } F(t) = \text{‘真’} \}$$

F是选择的条件，$\forall t \in R$， F(t)要么为真，要么为假;

F的形式：由**逻辑运算符**连接**算术表达式**而成。

逻辑运算符：$\wedge$，$\vee$，$\neg$

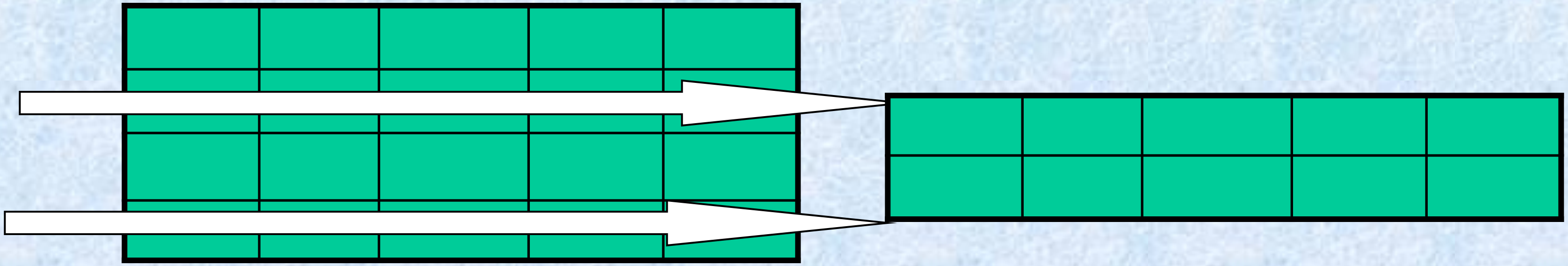
算术表达式：X θ Y

X，Y是属性名、常量、或简单函数

θ是比较算符，$\theta \in \{>, \geq, <, \leq, =, \neq\}$

2.4 关系代数（续）

- 选择运算



R

A	B	C
3	6	7
2	5	7
7	2	3
4	4	3

$\sigma_{A<5}(R)$

A	B	C
3	6	7
2	5	7
4	4	3

$\sigma_{A<5 \wedge C=7}(R)$

A	B	C
3	6	7
2	5	7

S

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

$\sigma_{4=\text{'CS'}}(S)$ 或

$\sigma_{Dept=\text{'CS'}}(S)$

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

2.4 关系代数(续)

S

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

- 示例：用关系表达式表达下列查询
  - 找关系S中计算机系（代号: ‘CS’）的全部男生

$$\sigma_{DEPT=‘CS’ \wedge SEX=‘男’}(S)$$

关系运算——选择(σ)

- 例: $\sigma_{Ssex = '男' \wedge Sdept = 'CS'}(Student)$

Student

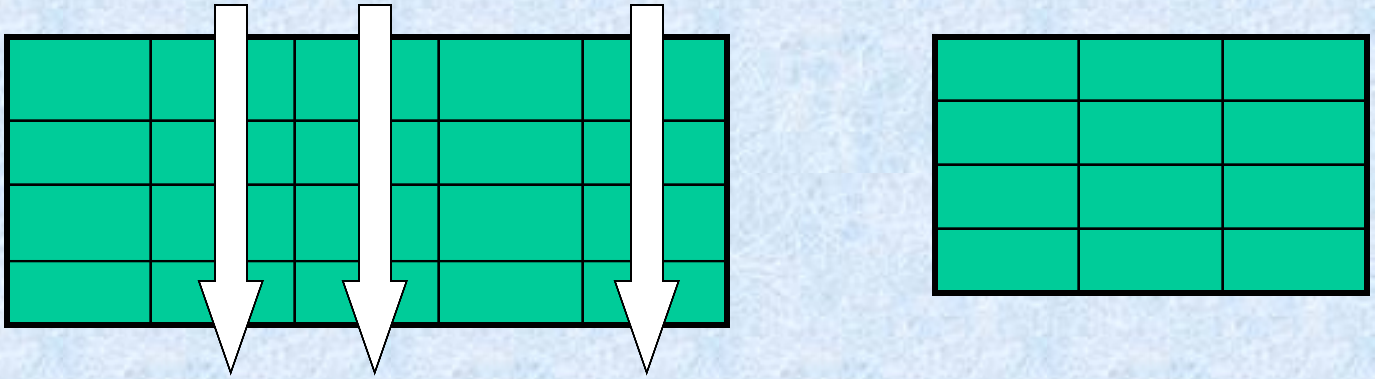
学 号 Sno	姓 名 Sname	性 别 Ssex	年 龄 Sage	所 在 系 Sdept
95001	李勇	男	20	CS
95002	刘晨	女	19	IS
95003	王敏	女	18	MA
95004	张立	男	19	IS

投影运算

- 定义
  - 从关系R中取若干列组成新的关系（从列的角度）

$$\Pi_A(R) = \{ t[A] \mid t \in R \}, A \subseteq R$$

  - 投影的结果中要去掉相同的行



投　影

R

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

$\Pi_{2,3}(R)$ 或

$\Pi_{B,C}(R)$

B	C
b	c
e	f

2.4 关系代数(续)

投影

Student(S#,SN,Age)

Course(C#,CN)

SC(C#,S#,Score)

- 示例

  列出所有学生的姓名和年龄

  $$\Pi_{SN,\ Age}\ (Student)$$

  找001号学生所选修的课程号

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

关系运算——投影(π)

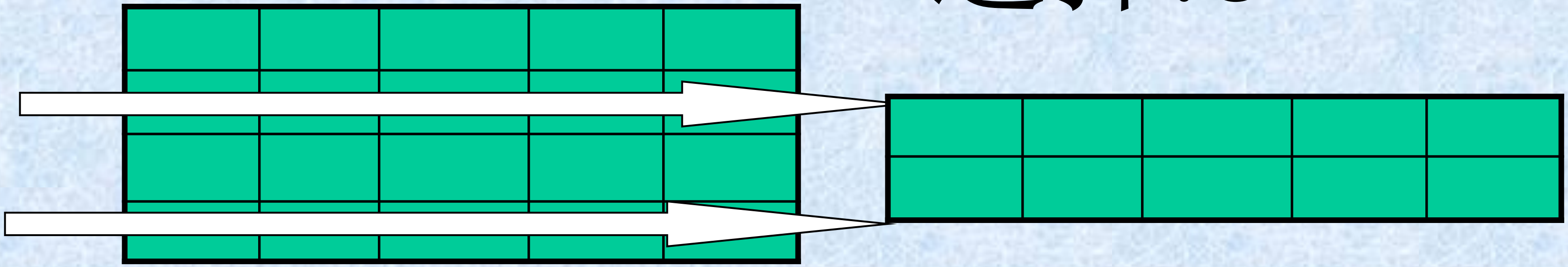
- 例: $\pi_{Ssex,\ Sage}(Student)$

Student

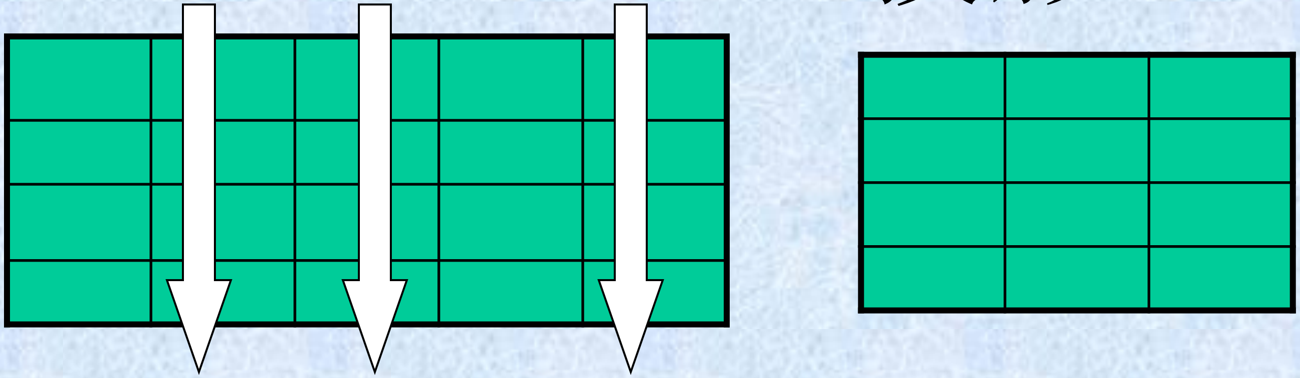
学 号 Sno	姓 名 Sname	性 别 Ssex	年 龄 Sage	所 在 系 Sdept
95001	李勇	男	20	CS
95002	刘晨	女	19	IS
95003	王敏	女	18	MA
95004	张立	男	19	IS

选择与投影的区别

选择:σ



投影:Π



S

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

- 示例
  - 列出CS系和MA系学生学号和姓名

方案1：

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

方案2：

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

θ连接

- 定义
  - 从两个关系的广义笛卡尔积中选取给定属性间满足一定条件的元组

$$R \underset{A\theta B}{\bowtie} S = \{ \widehat{t_r t_s} \mid t_r \in R \land t_s \in S \land t_r[A]\theta t_s[B] \}$$

A,B为R和S上度数相等且可比的属性列

θ为算术比较符，θ为等号时称为等值连接

- $R \underset{A\theta B}{\bowtie} S = \sigma_{R[A]\theta S[B]}(R \times S)$

2.4 关系代数(续)

R

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

S

D	E
3	1
6	2

$R \underset{B<D}{\bowtie} S$

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

自然连接

- 定义
  - 从两个关系的广义笛卡尔积中选取在相同属性列B上取值相等的元组，并去掉重复的列。

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

  - 自然连接与等值连接的不同
    - 自然连接中相等的分量必须是相同的属性组，并且要在结果中去掉重复的属性，而等值连接则不必。
  - 当R与S无相同属性时，$R \bowtie S = R \times S$

2.4 关系代数(续)

R

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

S

C	D
3	1
6	2

$R \bowtie S$

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

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	$\in$

$R \bowtie S$

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

- 外连接
  - 在自然连接的基础上，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)

2.4 关系代数(续)

R

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

S

C	D
3	1
6	2
8	5

$R \bowtie S$

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

外连接

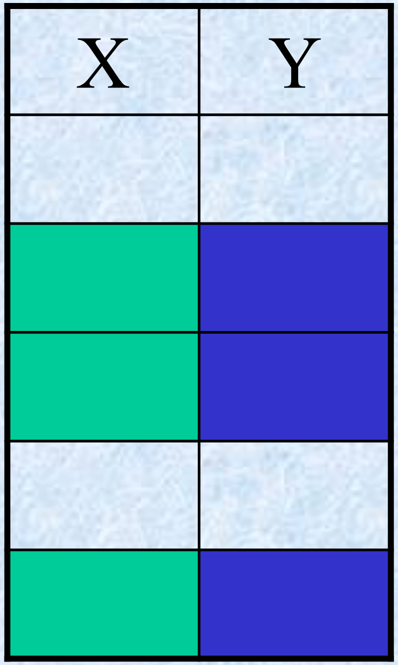
A	B	C	D
1	2	3	1
4	5	6	2
7	8	9	NULL
NULL	NULL	8	5

- 除
  - 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和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：

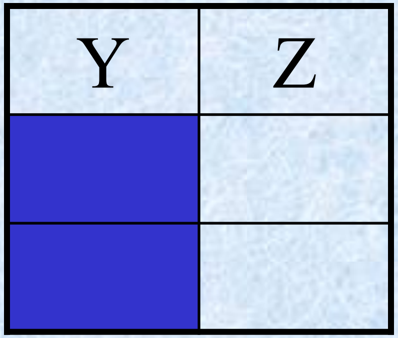
$$R \div S = \{t_r[X] \mid t_r \in R \wedge \Pi_y(S) \subseteq Y_x\}$$

其中，Y_x为x在R中的象集，$x=t_r[X]$

R



S



$\Pi_y(S)$

$R \div S$ $\quad Y_x$

R

A	B	C
a1	b1	c2
a2	b3	c7
a3	b4	c6
a1	b2	c3
a4	b6	c6
a2	b2	c3
a1	b2	c1

S

B	C	D
b1	c2	d1
b2	c1	d1
b2	c3	d2

$R \div S =$

A
a1

2.4 关系代数(续) $R \div S = \Pi_X(R) - \Pi_X(\Pi_X(R) \times \Pi_Y(S) - R)$

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

$\Pi_{AB}(R)$

A	B
a	b
b	c
e	d

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

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 ÷ S=

A	B
a	b
b	c
e	d

–

A	B
b	c

=

A	B
a	b
e	d

除法的现实意义：

姓名	课程
张军	物理
王红	数学
张军	数学

÷

课程
数学
物理

＝

姓名
张军

选修了全部课程的学生的姓名

2.4 关系代数(续) $R \div S = \Pi_X(R) - \Pi_X(\Pi_X(R) \times \Pi_Y(S) - R)$

姓名
张军
王红

×

课程
数学
物理

=

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

2.4 关系代数(续)

例题：

S

学号 Sno	姓名 Sname	性别 Ssex	年龄 Sage	系别 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

课程号 Cno	课程名 Cname	先行课 Cpno	学分 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	6	4

SC

学号 Sno	课程号 Cno	成绩 Grade
95001	1	92
95001	2	85
95001	3	88
95002	2	90
95002	3	80

2.4 关系代数(续)

- 查询信息系(IS)的全体学生

  $\sigma_{Sdept='IS'}(S)$ 或 $\sigma_{5='IS'}(S)$

- 查询年龄小于20岁的学生

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

- 查询学生的姓名和所在系

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

- 查选修了2号课程的学生的学号

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

- 查2号课程的学生学号和成绩

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

2.4 关系代数(续)

- 查2号课程的学生姓名和成绩

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

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

- 查选修“数学”的学生学号，姓名及该课程的成绩

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

- 查选修1号或2号课程的学生学号

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

2.4 关系代数(续)

- 查没学2号课程的学生姓名

$\Pi_{Sname}(\sigma_{Cno \neq '2'}(S \bowtie SC))$ ×

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

- 查询选修了全部课程的学生学号

$\Pi_{Sno,Cno}(SC) \div \Pi_{Cno}(C)$

2.4 关系代数(续)

- 5种基本运算
  - 并、差、笛卡尔积、投影、选择
- 其他运算
  - 交、连接、除
  - 均可用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
    - $R \cap S = R - (R - S)$
    - $R \bowtie S = \prod_{属性列表}(\sigma_{相同的属性列值相等}(R \times S))$
    - $R \div S = \Pi_X(R) - \Pi_X(\Pi_X(R) \times \Pi_Y(S) - R)$
- 关系代数中，这些运算经有限次复合后形成的式子称为关系代数表达式

2.6 小结

- 关系模型
- 关系数据结构及定义
- 完整性约束
- 关系代数:5个基本运算
- 关系演算:元组、 域

相近概念汇总

- 关系，表，实体集
- 元组，记录，实体
- 关系模式，记录型，实体型
- 关键字，码，标识符
- 主关键字，主码
- 候选关键字，候选码
- 外关键字，外码
- 分量，属性值，数据项
- 属性，字段
- 模式，外模式，视图
- 数据库，数据集，文件

下课了。。。

休息一会儿。。。