

数据库系统概论



计算机与信息学院
人工智能学院

数据模型 (Data Model)

➤定义：对数据结构，数据语义、数据之间的关系、数据约束等特征的抽象描述

A collection of tools for describing

Data

Data relationships

Data semantics

Data constraints

DBMS采用的，便于应用程序和开发人员使用的数据模型：关系模型

1) 概念模型：数据库设计人员采用的模型

2) **逻辑模型：DBMS采用的模型**

3) 物理模型：实际存储采用的模型

✓ **关系模型**

✓ 层次模型、网状模型

✓ 面向对象模型

✓ 对象关系模型

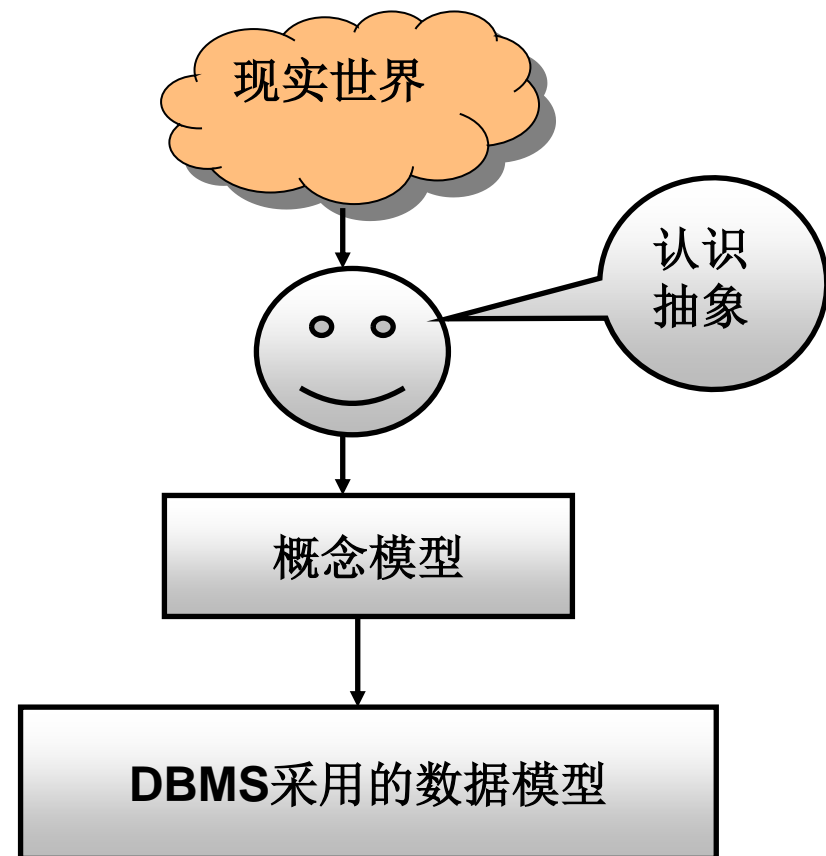
✓ 半结构化模型：XML

✓ 图模型

✓ 矩阵向量模型

信息世界

机器世界



1 关系模型 (Relational Model)

- 在数据库发展的早期：层次模型和网状模型
- 1970年，IBM公司的Edgar Codd首次提出了关系模型的概念

关系型数据库：采用关系模型作为数据模型

关系代数
图灵奖：Edgar Codd

Edgar Frank
"Ted" Codd
Inventor of
Relational Model for
Database

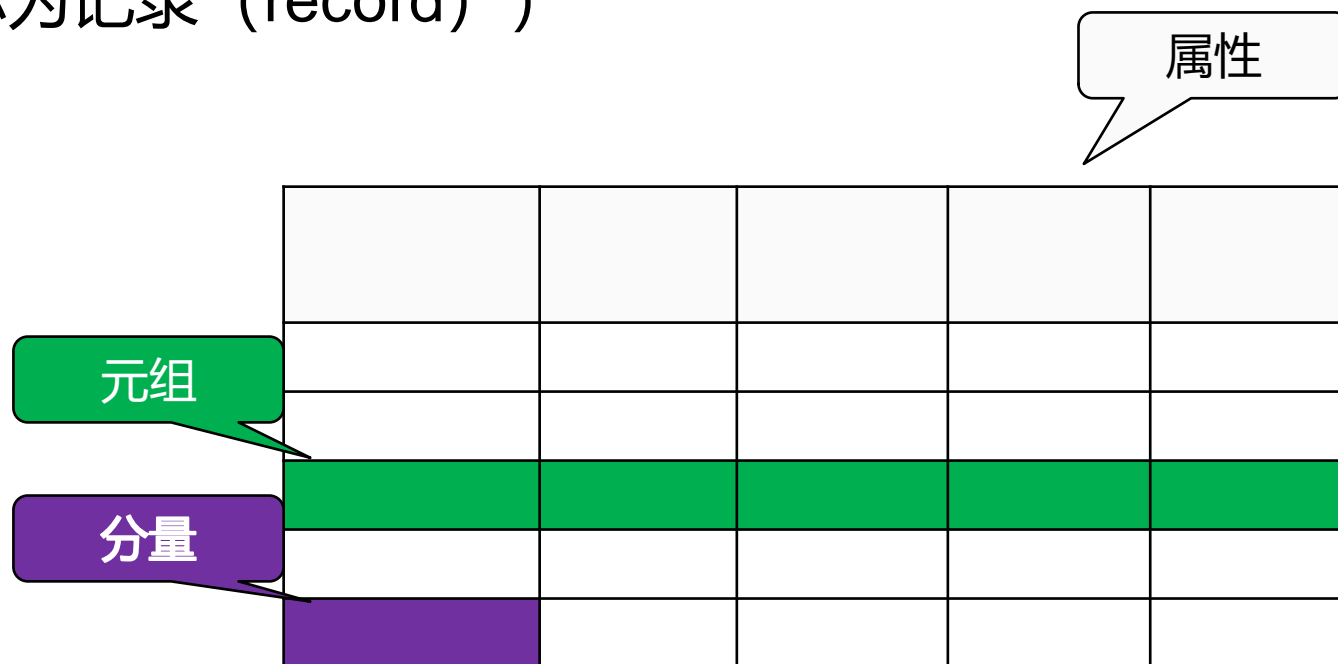


关系模型（逻辑模型）：三个要素

- **数据结构**：描述数据库的组成对象（类型，内容），以及对象之间的联系
- 数据操作
- 完整性约束

• 关系模型的数据结构

- Relation: 表 (table) , 关系理论中称为关系
- Tuple: 行, (元组, 也称为记录 (record))
- Attribute: 列 (属性)



Student表

<u>Sno</u> (学号)	<u>Sname</u> (姓名)	<u>Sgender</u> (性别)	<u>Sage</u> (年龄)	<u>Sdept</u> (所在系)
2021310721	李博	男	17	CS
2021310722	赵宇	男	19	CS
2021310723	张敏	女	18	CS
2021310724	王勇	男	18	MA
2021310725	刘佳	女	17	MA

SC表

<u>Sno</u> (学号)	<u>Cno</u> (课程号)	<u>Grade</u> (成绩)
2021310721	5	98
2021310722	1	87
2021310723	1	92
2021310723	5	76
2021310724	7	84
2021310725	4	95

Course表

<u>Cno</u> (课程号)	<u>Cname</u> (课程名)	<u>Cpno</u> (先修课)	<u>Ccredit</u> (学分)
1	数据库	2	4
2	数据结构与算法	6	4
3	操作系统	2	3
4	高等数学		4
5	软件工程	6	2
6	程序设计		3
7	数值分析	4	2

- ✓超码 (Superkey) :唯一标识元组的属性或属性集合
- ✓候选码 (Candidate key) : 超码, 其任意真子集都不是超码
- 主码 (Primary key) : 数据库设计的时候, 由设计者选定一个候选码, 作为主码

Student表

Sno (学号)	Sname (姓名)	Sgender (性别)	Sage (年龄)	Sdept (所在系)
2021310721	李博	男	17	CS
2021310722	赵宇	男	19	CS
2021310723	张敏	女	18	CS
2021310724	王勇	男	18	MA
2021310725	刘佳	女	17	MA

关系模型（逻辑模型）：三个要素

- 数据结构：描述数据库的组成对象（类型，内容），以及对象之间的联
- **数据操作：对关系模式以及关系，允许执行的操作的集合（增删修、查）**
- 完整性约束

关系模型（逻辑模型）：三个要素

- 数据结构：描述数据库的组成对象（类型，内容），以及对象之间的联
- 数据操作：对关系模式以及关系，允许执行的操作的集合（增删修、查）
- **完整性约束：一组规则，对关系进行约束，保证数据正确，有意义**

完整性约束

1. 实体完整性（主码约束）
2. 参照完整性（外码约束）
3. 用户定义完整性（属性约束）

实体完整性

- 实体完整性约束
 - 关系数据库中的每个元组都应该是唯一且可区分的
 - 主码不能重复，也不能为空

Student表

<u>Sno</u> (学号)	<u>Sname</u> (姓名)	<u>Sgender</u> (性别)	<u>Sage</u> (年龄)	<u>Sdept</u> (所在系)
2021310721	李博	男	17	CS
2021310722	赵宇	男	19	CS
2021310723	张敏	女	18	CS
2021310724	王勇	男	18	MA
2021310725	刘佳	女	17	MA

Student

<u>Sno</u> (学号)	Sname (姓名)	Sgender (性别)	Sage (年龄)	Sdept (所在系)
2021310721	李博	男	17	CS
2021310722	赵宇	男	19	CS
2021310723	张敏	女	18	CS
2021310724	王勇	男	18	MA
2021310725	刘佳	女	17	MA

Course

<u>Cno</u> (课程号)	Cname (课程名)	Cpno (先修课)	Ccredit (学分)
1	数据库	2	4
2	数据结构与算法	6	4
3	操作系统	2	3

SC

<u>Sno</u> (学号)	<u>Cno</u> (课程号)	Grade (成绩)
2021310721	5	98
2021310722	1	87
2021310723	1	92
2021310723	5	76

参照完整性

- 外码定义:

设 F 是基本关系 R 的一个或一组属性，但不是关系 R 的码。如果 F 与基本关系 S 的主码 K_s 相对应，则称 F 是基本关系 R 的**外码（FK）**， R 为**参照关系**， S 为**被参照关系**

- 1) **专业** (专业编号, 专业名称, 负责人, 成立时间)
- 2) **学生** (学号, 学生姓名, 出生日期, 籍贯, 专业编号)

参照完整性

- F 是基本关系 R 的外码，它与基本关系 S 的主码相对应，则 F 的值

- ✓ 空值

- ✓ 或者等于 S 中某个元组的主码值

- 1) **专业** (专业编号, 专业名称, 负责人, 成立时间)

- 2) **学生** (学号, 学生姓名, 出生日期, 籍贯, 专业编号)

- student (Sno, Sname, Sgender, Sage, Sdept)
- course (Cno, Cname, Cpno, Ccredit)
- Sc (Sno, Cno, Grade)

Student

<u>Sno</u> (学号)	Sname (姓名)	Sgender (性别)	Sage (年龄)	Sdept (所在系)
2021310721	李博	男	17	CS
2021310722	赵宇	男	19	CS
2021310723	张敏	女	18	CS
2021310724	王勇	男	18	MA
2021310725	刘佳	女	17	MA

Course

<u>Cno</u> (课程号)	Cname (课程名)	Cpno (先修课)	Ccredit (学分)
1	数据库	2	4
2	数据结构与算法	6	4
3	操作系统	2	3

SC

<u>Sno</u> (学号)	<u>Cno</u> (课程号)	Grade (成绩)
2021310721	5	98
2021310722	1	87
2021310723	1	92
2021310723	5	76

- 外码定义：关系R和S可以为同一个关系
- 参照完整性可以在一个表上进行约束

Course(Cno,Cname,Ccredit,Cpno)

课程号 Cno	课程名 Cname	学分 Ccredit	先修课 Cpno
81001	程序设计基础与C语言	4	
81002	数据结构	4	81001
81003	数据库系统概论	4	81002
81004	信息系统概论	4	81003
81005	操作系统	4	81001
81006	Python语言	3	81002
81007	离散数学	4	
81008	大数据技术概论	4	81003

- 1) 专业表中, PK [填空1], FK [填空2]
- 2) 教师表中, PK [填空3], FK [填空4]
- 3) 学生表, PK [填空5], FK [填空6]

- 1) **专业** (专业编号, 专业名称, 负责人工号, 成立时间)
- 2) **教师** (工号, 教师姓名, 入职时间, 出生日期, 专业编号)
- 3) **学生** (学号, 学生姓名, 出生日期, 籍贯, 专业编号)

作答

自定义完整性：定义属性的取值约束

- ✓唯一：unique
- ✓能否为空：null/not null
- ✓取值范围：check
- ✓默认值：default
- ✓自增：auto_increment

完整性约束

1. 实体完整性（主码约束）
2. 参照完整性（外码约束）
3. 用户定义完整性（属性约束）

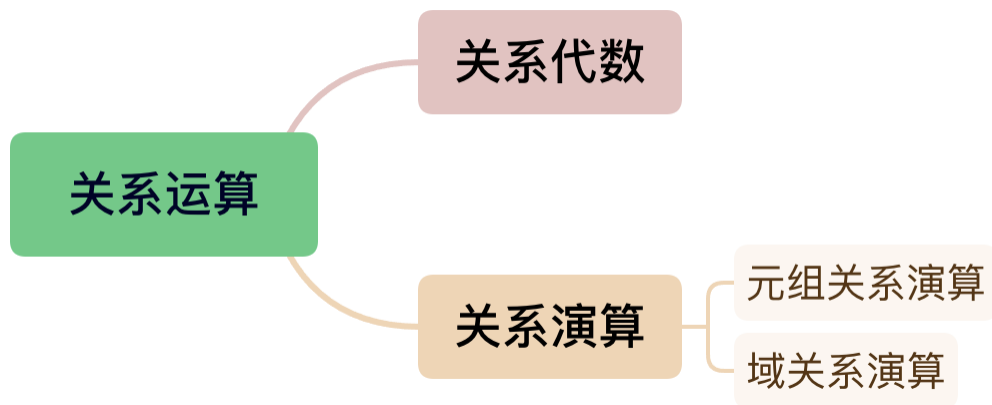
实体完整性和参照完整性：称为关系的两个不变性，由关系型DBMS自动支持

关系模型（逻辑模型）：三个要素

- 数据结构：描述数据库的组成对象（类型，内容），以及对象之间的联
- 数据操作：对关系模式以及关系，允许执行的操作的集合（增删修、查）
- 完整性约束：一组规则，对关系进行约束，保证数据正确，有意义

2 关系操作

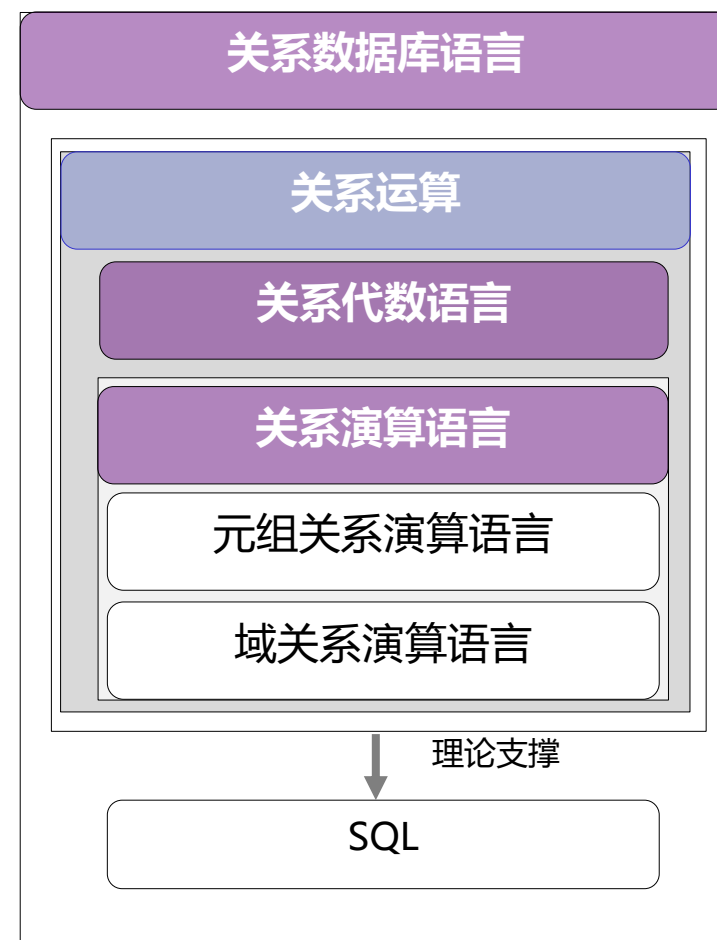
- 关系模型上的数据操作——关系操作，包括：**查询操作（Query）**和数据更新操作
- 关系操作可以通过**关系运算**进行表示。
- 关系运算是埃德加·科德博士所提出的关系模型的一部分，本质上是一种形式化的关系查询语言，可以用于从数据库中查询数据。



- 关系代数 (relational algebra) : 通过描述对关系的运算 (集合运算符和专门的关系运算符) 来表达查询、获取数据
- 关系演算: 对关系的谓词演算 (元组/分量作为谓词变元) 表达查询

结构化查询语言SQL

- ✓ 方便用户操作
- ✓ 理论基础: 关系代数
- ✓ 执行SQL需要通过关系代数理论进行优化



3 关系代数

➤ 关系的形式化定义

1. 域 (domain)
2. 笛卡儿积 (Cartesian product)
3. 关系 (relation)

- 域 (Domain) : 一组具有相同数据类型的值的集合

- ✓ 整数
- ✓ 实数
- ✓ 介于某个取值范围的整数
- ✓ 长度小于25B的变长字符串集合
- ✓ {男, 女}
- ✓

- 笛卡尔积

给定一组域 D_1, D_2, \dots, D_n , 某些域可以相同

D_1, D_2, \dots, D_n 的笛卡尔积为:

$$D_1 \times D_2 \times \dots \times D_n =$$

$$\{ (d_1, d_2, \dots, d_n) \mid d_i \in D_i, i=1, 2, \dots, n \}$$

- 所有域的所有取值的一个组合
- 不能重复

- ❖ D_1 =导师集合**SUPERVISOR**= { 张清玫, 刘逸 }
- ❖ D_2 =专业集合**MAJOR**= { 计算机科学与技术, 信息管理与信息系统 }
- ❖ D_3 =研究生集合**POSTGRADUATE**= { 李勇, 刘晨, 王敏 }

SUPERVISOR 导师	MAJOR 专业	POSTGRADUATE 研究生
张清玫	计算机科学与技术	李勇
张清玫	计算机科学与技术	刘晨
张清玫	计算机科学与技术	王敏
张清玫	信息管理与信息系统	李勇
张清玫	信息管理与信息系统	刘晨
张清玫	信息管理与信息系统	王敏
刘逸	计算机科学与技术	李勇
刘逸	计算机科学与技术	刘晨
刘逸	计算机科学与技术	王敏
刘逸	信息管理与信息系统	李勇
刘逸	信息管理与信息系统	刘晨
刘逸	信息管理与信息系统	王敏

➤关系：形式化定义

$D_1 \times D_2 \times \dots \times D_n$ 的**子集**叫作在域 D_1, D_2, \dots, D_n 上的关系

关系的列对应域，域可以相同，由属性名加以区分

n：关系的度（degree）/目

SUPERVISOR	MAJOR	POSTGRADUATE
张清玫	计算机科学与技术	李勇
张清玫	计算机科学与技术	刘晨
刘逸	信息管理与信息系统	王敏

SUPERVISOR	MAJOR	POSTGRADUATE
张清玫	计算机科学与技术	李勇
张清玫	计算机科学与技术	刘晨
刘逸	信息管理与信息系统	王敏

关系模式: $R(A_1, A_2, \dots, A_n)$

关系: 具体的一张表 (实例)

关系的6个性质:

- ① 列是同质的 (homogeneous)
- ② 不同的列可出自同一个域
- ③ 列的顺序无所谓,列的次序可以任意交换
- ④ 任意两个元组的码不能相同
- ⑤ 行的顺序无所谓, 行的次序可以任意交换
- ⑥ **分量必须取原子值, 即每一个分量都必须是不可分的数据项**

分量: 元组在每个属性上的取值

关系的6个性质:

- ⑥ 分量必须取原子值，即每一个分量都必须是不可分的数据项

SUPERVISOR	SPECIALITY	POSTGRADUATE	
		PG1	PG2
张清玫	计算机科学与技术	李勇	刘晨
刘逸	信息管理与信息系统	王敏	

Sno (学号)	Sname (姓名)	Contact (联系方式)
2021310721	李博	137-XXXX-1790, Lib@uni.edu.cn
2021310722	赵宇	130-XXXX-1110, Zhaoy@uni.edu.cn
2021310723	张敏	139-XXXX-9999, Zhangm@uni.edu.cn

Sno (学号)	Sname (姓名)	Phone (手机号)	Email (邮箱)
2021310721	李博	137-XXXX-1790	Lib@uni.edu.cn
2021310722	赵宇	130-XXXX-1110	Zhaoy@uni.edu.cn
2021310723	张敏	139-XXXX-9999	Zhangm@uni.edu.cn

➤关系代数

- 运算对象：关系（元组的集合）
- 运算结果：关系（元组的集合）
- 运算符：关系运算符

集合 运算符	\cup	并
	$-$	差
	\cap	交
	\times	笛卡尔积

专门 的关系 运算符	σ	选择
	π	投影
	\bowtie	连接
	\div	除

集合运算

1. 并(union): $R \cup S = \{ t | t \in R \vee t \in S \}$
2. 差(except): $R - S = \{ t | t \in R \wedge t \notin S \}$
3. 交(intersection): $R \cap S = \{ t | t \in R \wedge t \in S \}$

<i>R</i>		
<i>A</i>	<i>B</i>	<i>C</i>
<i>a</i> ₁	<i>b</i> ₁	<i>c</i> ₁
<i>a</i> ₁	<i>b</i> ₂	<i>c</i> ₂
<i>a</i> ₂	<i>b</i> ₂	<i>c</i> ₁

<i>S</i>		
<i>A</i>	<i>B</i>	<i>C</i>
<i>a</i> ₁	<i>b</i> ₂	<i>c</i> ₂
<i>a</i> ₁	<i>b</i> ₃	<i>c</i> ₂
<i>a</i> ₂	<i>b</i> ₂	<i>c</i> ₁

<i>R</i> \cup <i>S</i>		
<i>A</i>	<i>B</i>	<i>C</i>
<i>a</i> ₁	<i>b</i> ₁	<i>c</i> ₁
<i>a</i> ₁	<i>b</i> ₂	<i>c</i> ₂
<i>a</i> ₂	<i>b</i> ₂	<i>c</i> ₁
<i>a</i> ₁	<i>b</i> ₃	<i>c</i> ₂

4. 笛卡尔积: $R \times S = \{ \overbrace{t_r t_s} \mid t_r \in R \wedge t_s \in S \}$

<i>R</i>		
<i>A</i>	<i>B</i>	<i>C</i>
<i>a</i> ₁	<i>b</i> ₁	<i>c</i> ₁
<i>a</i> ₁	<i>b</i> ₂	<i>c</i> ₂
<i>a</i> ₂	<i>b</i> ₂	<i>c</i> ₁

<i>S</i>		
<i>A</i>	<i>B</i>	<i>C</i>
<i>a</i> ₁	<i>b</i> ₂	<i>c</i> ₂
<i>a</i> ₁	<i>b</i> ₃	<i>c</i> ₂
<i>a</i> ₂	<i>b</i> ₂	<i>c</i> ₁

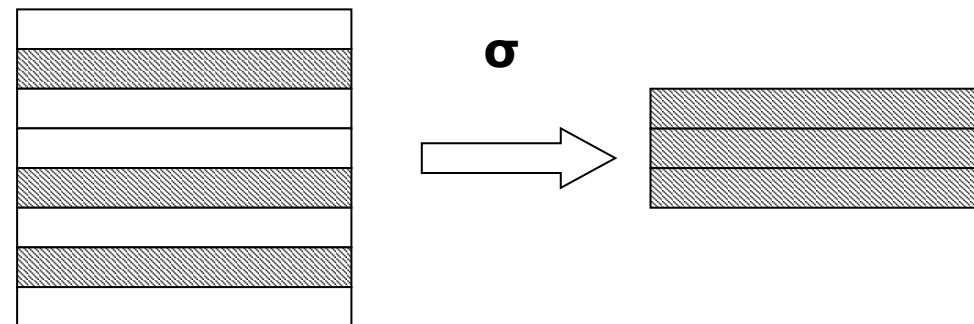
<i>R</i> × <i>S</i>					
<i>R.A</i>	<i>R.B</i>	<i>R.C</i>	<i>S.A</i>	<i>S.B</i>	<i>S.C</i>
<i>a</i> ₁	<i>b</i> ₁	<i>c</i> ₁	<i>a</i> ₁	<i>b</i> ₂	<i>c</i> ₂
<i>a</i> ₁	<i>b</i> ₁	<i>c</i> ₁	<i>a</i> ₁	<i>b</i> ₃	<i>c</i> ₂
<i>a</i> ₁	<i>b</i> ₁	<i>c</i> ₁	<i>a</i> ₂	<i>b</i> ₂	<i>c</i> ₁
<i>a</i> ₁	<i>b</i> ₂	<i>c</i> ₂	<i>a</i> ₁	<i>b</i> ₂	<i>c</i> ₂
<i>a</i> ₁	<i>b</i> ₂	<i>c</i> ₂	<i>a</i> ₁	<i>b</i> ₃	<i>c</i> ₂
<i>a</i> ₁	<i>b</i> ₂	<i>c</i> ₂	<i>a</i> ₂	<i>b</i> ₂	<i>c</i> ₁
<i>a</i> ₂	<i>b</i> ₂	<i>c</i> ₁	<i>a</i> ₁	<i>b</i> ₂	<i>c</i> ₂
<i>a</i> ₂	<i>b</i> ₂	<i>c</i> ₁	<i>a</i> ₁	<i>b</i> ₃	<i>c</i> ₂
<i>a</i> ₂	<i>b</i> ₂	<i>c</i> ₁	<i>a</i> ₂	<i>b</i> ₂	<i>c</i> ₁

专门的关系运算

5. 选择: $\sigma_F(R) = \{t | t \in R \wedge F(t) = \text{'True'}\}$

$F: X \theta Y$

$\theta: >, \geq, <, \leq, = \text{或} <>$



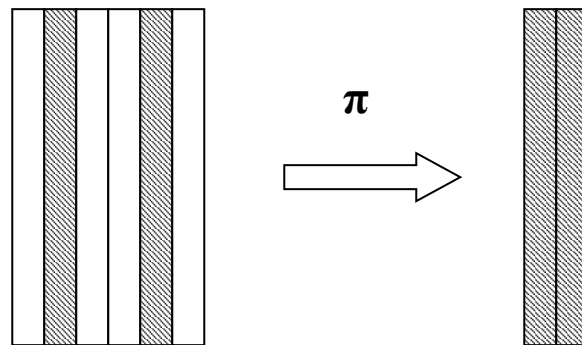
$\sigma_{Sdept="CS"} (Student)$

$\sigma_{(Sdept="CS" \wedge Sage \geq 18) \vee Sdept="MA"} (Student)$

<u>Sno</u> (学号)	Sname (姓名)	Sgender (性别)	Sage (年龄)	Sdept (所在系)
2021310721	李博	男	17	CS
2021310722	赵宇	男	19	CS
2021310723	张敏	女	18	CS
2021310724	王勇	男	18	MA
2021310725	刘佳	女	17	MA

Student

6. 投影: $\pi_A(R) = \{ t[A] \mid t \in R \}$



$\Pi_{Sno, Sname} (Student)$

<u>Sno</u> (学号)	Sname (姓名)	Sgender (性别)	Sage (年龄)	Sdept (所在系)
2021310721	李博	男	17	CS
2021310722	赵宇	男	19	CS
2021310723	张敏	女	18	CS
2021310724	王勇	男	18	MA
2021310725	刘佳	女	17	MA

Student

- 选择和投影结合

<u>Sno</u> (学号)	Sname (姓名)	Sgender (性别)	Sage (年龄)	Sdept (所在系)
2021310721	李博	男	17	CS
2021310722	赵宇	男	19	CS
2021310723	张敏	女	18	CS
2021310724	王勇	男	18	MA
2021310725	刘佳	女	17	MA

Student

- 查询student中cs系学生的学号、姓名

$\Pi_{Sno, Sname} (\sigma_{Sdept='CS'} (Student))$

<u>Sno</u> (学号)	Sname (姓名)	Sgender (性别)	Sage (年龄)	Sdept (所在系)
2021310721	李博	男	17	CS
2021310722	赵宇	男	19	CS
2021310723	张敏	女	18	CS
2021310724	王勇	男	18	MA
2021310725	刘佳	女	17	MA

Student

1. 查询学号为0722的学生姓名及所在系
2. 查询计算机系未满18岁的学生学号和姓名

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

A, B : R 和 S 上可比的属性 (组)

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

$R \bowtie_{C < E} S$				
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

- 等值连接：

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

- A 、 B ： R 和 S 上可比的属性（组）

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

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

$$R \bowtie_{R.B=S.B} S$$

- **自然连接**： $R \bowtie S$

特殊的等值连接，连接的两个属性（组），**属性名相同**

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

$R \bowtie S$

$$R$$

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

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

$R \bowtie_{R.B=S.B} S$

• 查询每位学生的选课信息

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	CS
200215123	王敏	女	18	MA
200515125	张立	男	19	IS

Student

学号 Sno	课程号 Cno	成绩 Grade
200215121	1	92
200215121	2	85
200215121	3	88
200215122	2	90
200215122	3	80

SC

- 查询每位学生的选课信息
- 查询200215121学生的选课信息

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	CS
200215123	王敏	女	18	MA
200515125	张立	男	19	IS

Student

学号 Sno	课程号 Cno	成绩 Grade
200215121	1	92
200215121	2	85
200215121	3	88
200215122	2	90
200215122	3	80

SC

- 查询每位学生的选课信息
- 查询200215121学生的选课信息
- 查询2002151学生的选课信息，给出其选修的课程号以及成绩

学号 <u>Sno</u>	姓名 Sname	性别 Ssex	年龄 Sage	所在系 Sdept
200215121	李勇	男	20	CS
200215122	刘晨	女	19	CS
200215123	王敏	女	18	MA
200515125	张立	男	19	IS

Student

学号 Sno	课程号 Cno	成绩 Grade
200215121	1	92
200215121	2	85
200215121	3	88
200215122	2	90
200215122	3	80

SC

- 外连接 $R \bowtie S$ (左外 $R \bowtie S$ / 右外连接 $R \bowtie S$)

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

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

- **关系代数表达式**：关系运算经过有限次的复合
- 各运算符的计算顺序为从左到右，括号具有最高优先级

• 多表连接：2个及以上的表，连接查询

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	CS
200215123	王敏	女	18	MA
200515125	张立	男	19	IS

Student

学号 Sno	课程号 Cno	成绩 Grade
200215121	1	92
200215121	2	85
200215121	3	88
200215122	2	90
200215122	3	80

SC

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语言	6	4

Course

查询每位学生的选课信息

练习一

- 查询每位学生的选课信息（3表连接）
- 查询刘晨的选课信息，给出学号、课程号、课程名、成绩
- 查询选修了“数据库”课程的学生学号、姓名和成绩

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	CS
200215123	王敏	女	18	MA
200515125	张立	男	19	IS

Student

学号 Sno	课程号 Cno	成绩 Grade
200215121	1	92
200215121	2	85
200215121	3	88
200215122	2	90
200215122	3	80

SC

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语言	6	4

Course

- 自连接：一个关系与自身做笛卡尔积

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语言	6	4

Course

例：查询每门课程的先修课的课程名


例：查询每门课程的先修课的课程名

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语言	6	4

Course

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语言	6	4

集合运算符	\cup $-$ \cap \times	并 差 交 笛卡尔积
-------	-------------------------------------	---------------------

专门的关系运算符	σ π  \div	选择 投影 连接 除
----------	--	---------------------

8. 除: $R \div S$

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

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

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_1 在 R 中的象集

$$Z_{x1} = \{Z_1, Z_2, Z_3\},$$

❖ x_2 在 R 中的象集

$$Z_{x2} = \{Z_2, Z_3\},$$

❖ x_3 在 R 中的象集

$$Z_{x3} = \{Z_1, Z_3\}$$

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

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 $P(X)$ ，

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

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

$R \div S$		
A		
a_1		

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

查询至少选修1号课程和3号课程的学生号

学号 Sno	课程号 Cno	成绩 Grade
200215121	1	92
200215121	2	85
200215121	3	88
200215122	2	90
200215122	3	80

首先建立一个临时关系K:

Cno
1
3

然后求: $\pi_{\text{Sno.Cno}}(\text{SC}) \div K$

练习二

- 查询选修了全部课程的学生学号
- 查询选修了3号和5号课程的学生学号、姓名及所在系

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	CS
200215123	王敏	女	18	MA
200515125	张立	男	19	IS

学号 Sno	课程号 Cno	成绩 Grade
200215121	1	92
200215121	2	85
200215121	3	88
200215122	2	90
200215122	3	80

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语言	6	4

关系代数

```
graph TD; A[关系代数] --> B[基本关系代数运算]; A --> C[附加关系代数运算]; A --> D[扩展关系代数运算]; B --> B1[选择]; B --> B2[投影]; B --> B3[并]; B --> B4[差]; B --> B5[笛卡尔积]; B --> B6[重命名]; C --> C1[交]; C --> C2[连接]; C --> C3[赋值]; C --> C4[除]; D --> D1[去重]; D --> D2[广义投影]; D --> D3[聚集]; D --> D4[分组]; D --> D5[排序];
```

基本关系代数运算

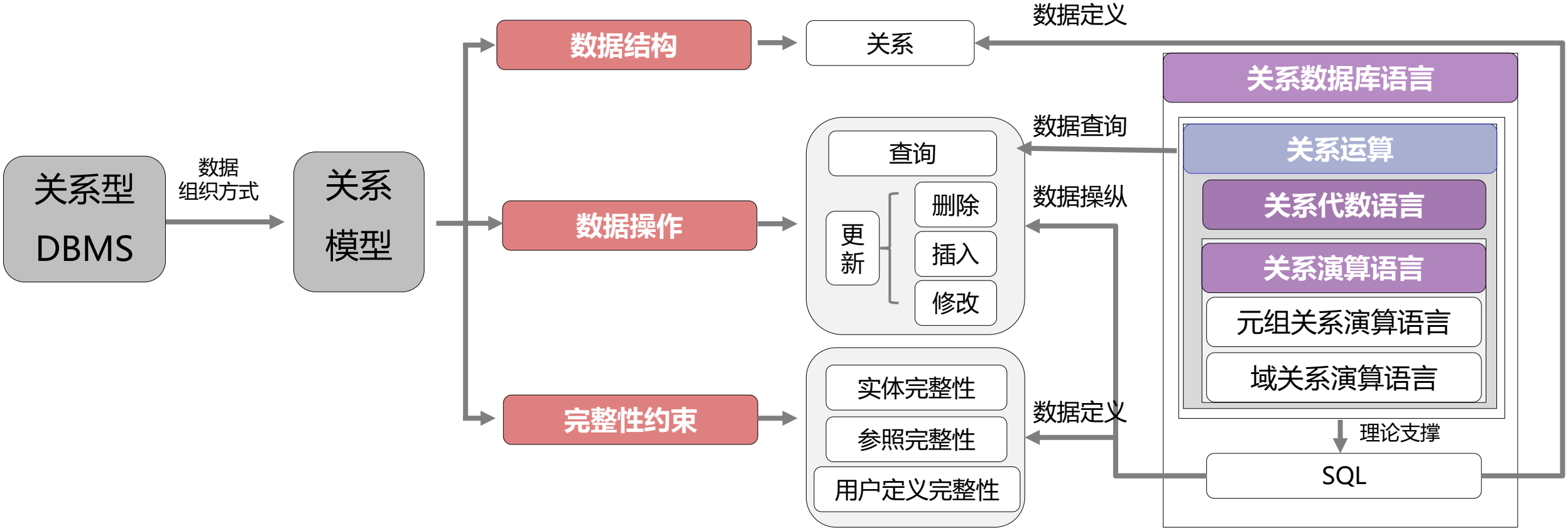
选择 投影 并 差 笛卡尔积 重命名

附加关系代数运算

交 连接 赋值 除

扩展关系代数运算

去重 广义投影 聚集 分组 排序



作业

6. 设有一个 SPJ 数据库，包括 S、P、J 及 SPJ 4 个关系模式：

S (SNO, SNAME, STATUS, CITY) ;

P (PNO, PNAME, COLOR, WEIGHT) ;

J (JNO, JNAME, CITY) ;

SPJ (SNO, PNO, JNO, QTY) 。

供应商表 S 由供应商代码 (SNO)、供应商姓名 (SNAME)、供应商状态 (STATUS)、供应商所在城市 (CITY) 组成。

零件表 P 由零件代码 (PNO)、零件名 (PNAME)、颜色 (COLOR)、重量 (WEIGHT) 组成。

工程项目表 J 由工程项目代码 (JNO)、工程项目名 (JNAME)、工程项目所在城市 (CITY) 组成。供应情况表 SPJ 由供应商代码 (SNO)、零件代码 (PNO)、工程项目代码 (JNO)、供应数量 (QTY) 组成，表示某供应商供应某种零件给某工程项目的数量为 QTY。

今有若干数据如下：

作业

S 表

SNO	SNAME	STATUS	CITY
S1	精益	20	天津
S2	盛锡	10	北京
S3	东方红	30	北京
S4	丰泰盛	20	天津
S5	为民	30	上海

P 表

PNO	PNAME	COLOR	WEIOHT
PI	螺母	红	12
P2	螺栓	绿	17
P3	螺丝刀	蓝	14
P4	螺丝刀	红	14
P5	凸轮	蓝	40
P6	齿轮	红	30

J表

JNO	JNAME	CITY
J1	三建	北京
J2	一汽	长春
J3	弹簧厂	天津

J4	造船厂	天津
J5	机车厂	唐山
J6	无线电厂	常州
J7	半导体厂	南京

SPJ表

SNO	PNO	JNO	QTY
S1	P1	J1	200
S1	P1	J3	100
S1	P1	J4	700
S1	P2	J2	100
S2	P3	J1	400
S2	P3	J2	200
S2	P3	J4	500
S2	P3	J5	400
S2	P5	J1	400
S2	P5	J2	100
S3	P1	J1	200
S3	P3	J1	200
S4	P5	J1	100
S4	P6	J3	300
S4	P6	J4	200
S5	P2	J4	100
S5	P3	J1	200
S5	P6	J2	200
S5	P6	J4	500

作业

试用关系代数、ALPHA 语言、QBE 语言完成如下查询：

- (1) 求供应工程 J1 零件的供应商号码 SNO;
- (2) 求供应工程 J1 零件 P1 的供应商号码 SNO;
- (3) 求供应 T 程 J1 零件为红色的供应商号码 SNO;
- (4) 求没有使用天津供应商生产的红色零件的工程项目代码 JNO;
- (5) 求至少用了供应商 S1 所供应的全部零件的工程项目代码 JNO。

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	CS
200215123	王敏	女	18	MA
200515125	张立	男	19	IS

Student

学号 Sno	课程号 Cno	成绩 Grade
200215121	1	92
200215121	2	85
200215121	3	88
200215122	2	90
200215122	3	80

SC

Course

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语言	6	4

练习：关系代数

- ① 查询计算机系学生的选课信息，给出学号、姓名、课程号、成绩
- ② 查询选修了“数据库”课程的学生信息，给出学号，系，成绩
- ③ 查询选修了3号和5号课程的学生学号、姓名及所在系