1. 关系数据结构及形式化定义
2. 关系
3. 单一的数据结构——关系

现实世界的实体以及实体间的各种联系均用关系来表示

逻辑结构——二维表

从用户角度，关系模型中数据的逻辑结构是一张二维表

1. 域

域是一组**具有相同数据类型的值**的**集合**

例如：整数、{‘男’,’女’}、介于某个取值范围内的日期

1. 笛卡尔积

域上的一种集合运算，所有域的所有取值的一个组合

给定**一组域**D 1 ，D 2 ，…，D n ，允许其中某些域是相同的。

D 1 ，D 2 ，…，D n 的笛卡尔积为：

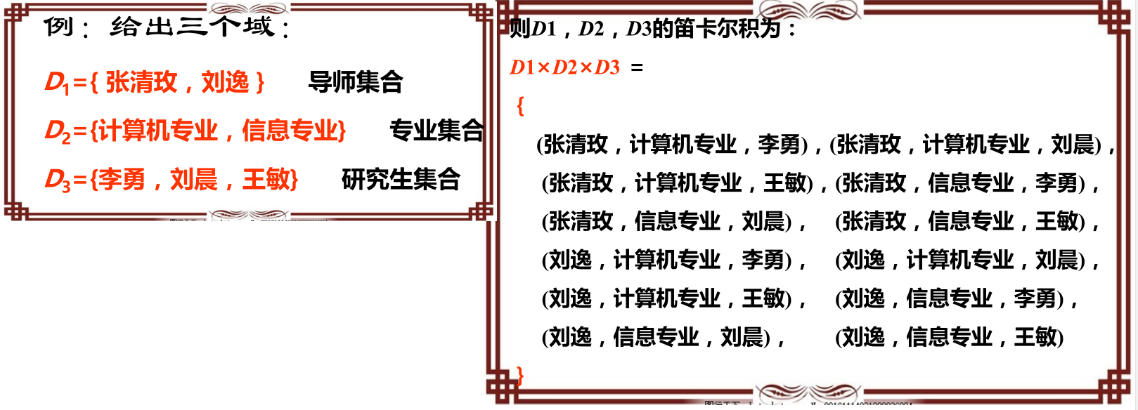
D 1 ×D 2 ×…×D n ＝

｛（d 1 ，d 2 ，…，d n ）｜d i  D i ，i＝1，2，…，n｝

其中每一个元素(d 1 ，d 2 ，…，dn)叫作一个n元组（n-tuple）或简称**元组**(Tuple)

元素中的每一个值d i 叫做一个**分量**

一个**域**允许的不同取值个数称为这个域的**基数**



（有的没有实际意义）笛卡尔积中取出有实际意义的元组来构造关系

笛卡尔积的表示方法：笛卡尔积可表示为一个二维表，表中的每行对应一个元组，表中的每列对应一个域

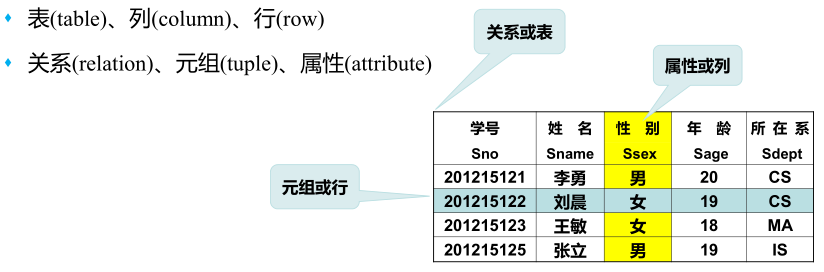
1. 关系

**关系是笛卡尔积的有限子集**

D 1 ×D 2 ×…×D n 的子集叫作在域D 1 ，D 2 ，…，D n 上的关系，表示为R（ （D 1 ， ，D 2 ， ，… ，D n ）

R：关系名；n：关系的目或度（Degree）

1. 关系的术语
2. 表（关系）、列（属性）、行（元组）



1. 码

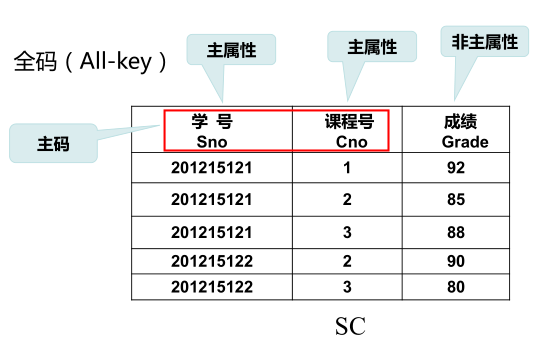
候选码：若关系中的**某一属性组的值能唯一地标识一个元组，而其子集不能，**则称该**属性组**为候选码

全码：**在最简单的情况下，候选码只包含一个属性。**在最极端的情况下，关系模式的所有属性是这个关系模式的候选码，称为全码

主码：**若一个关系有多个候选码，则选定一个为主码**

主属性：**候选码的诸属性成为主属性**

非主属性：不包含在任何候选码中的属性成为非主属性



注意：**Sno与Cno加一起是候选码/主码，单独出来不是，但都为主属性**

1. 三类关系

基本关系（基本表）：实际存在的表，是实际存储数据的逻辑表示

查询表：查询结果对应的表

视图表：由基本表或其他视图表导出的表，是虚表，不对应实际存储的数据

1. 基本关系的性质
2. 列是同质的（取自同一个域）
3. 不同的列可出自同一个域

其中的每一列称为一个属性

不同的属性要给予不同的属性名

1. 列的顺序无所谓, 列的次序可以任意交换
2. 任意两个元组的候选码不能相同
3. 行的顺序无所谓，行的次序可以任意交换
4. 分量必须取原子值（一个属性不能再分）
5. 关系模式
6. 关系模式是对关系的描述，是静态的、稳定的，是关系数据的型

关系是关系模式在某一时刻的状态或内容，是动态的、随时间不断变化的，是值



关系模式可以形式化地表示为：R （ U ， D ， DOM ， F ）

R 关系名；U 组成该关系的属性名集合；D 属性组U中属性所来自的域；DOM 属性向域的映象集合；F 属性间的数据依赖关系集合

关系模式通常可以简记为：R (U) 或 R (A1，A2，…，An)

R: 关系名；A1，A2，…，An : 属性名

1. 关系数据库

在一个给定的应用领域中，所有关系的集合构成一个关系数据库

关系数据库的型与值：关系数据库的型也称关系数据库模式，是对关系数据库的描述；关系数据库的值是关系模式在某一时刻对应的关系的集合，简称为关系数据库

1. 关系操作
2. 常见的关系操作

查询：选择、投影、连接、除、并、交、差

数据更新：插入、删除、修改

查询的表达能力是其中最主要的部分

**选择、投影、并、差、笛卡尔积**是5种基本操作

1. 关系操作的特点

集合操作方式：**操作的对象和结果都是集合**，一次一集合的方式

1. 关系的完整性
2. 关系模型的完整性规则是对关系的某种约束条件，也就是关系的值需要满足的约束条件

实体完整性

参照完整性

用户定义完整性

**实体完整性和参照完整性**：**关系模型必须满足**的完整性约束条件，**称为关系的两个不变性**，应该由关系系统自动支持

用户定义的完整性：应用领域需要遵循的约束条件，体现了具体领域中的语义约束

1. 实体完整性

**若属性A是基本关系R的主属性，则属性A不能取空值**

实体完整性规则是针对基本关系而言的。一个基本表通常对应现实世界的一个实体集

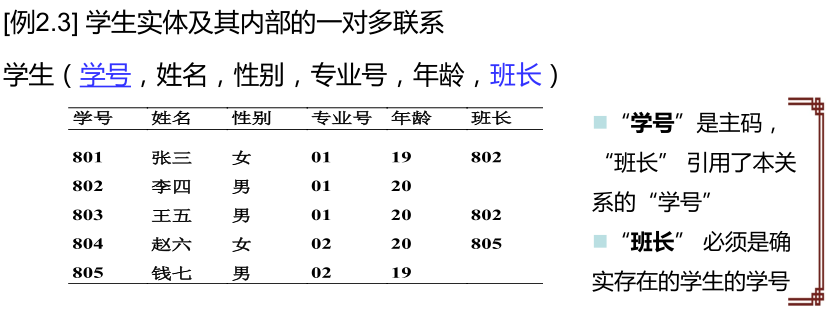
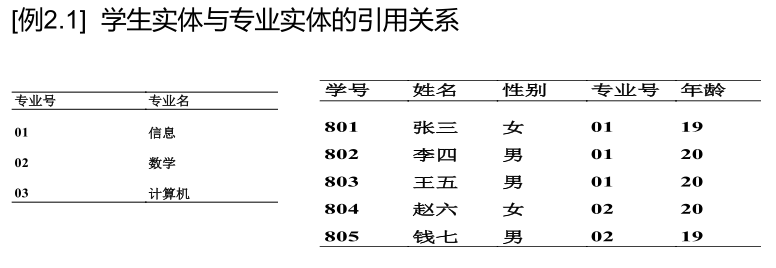
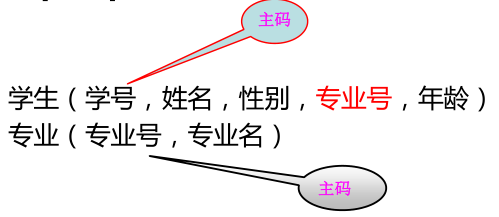
现实世界中的实体是可区分的，即它们具有某种唯一性标识

关系模型中以主码作为唯一性标识

~~主码中的属性即主属性不能取空值~~

1. 参照完整性
2. 关系间的引用

例如：学生实体与专业实体之间具有引用关系



1. 外码

设F是基本关系R的一个或一组属性，但**不是关系R的码**，K s 是基本关系S的**主码**。如果**F与K s** 相对应，则称F是基本关系R的外码

**在当前关系中不是主码**，在另一个关系中是主码

外码不能是主码，在SC关系中，主码是学号+课程号，外码是学号、课程号。

关系R和S可以是相同的关系，也可以是不同的关系

目标关系S的主码Ks 和参照关系的外码F必须定义在同一个（或一组）域上

外码并不一定要与相应的主码同名，当外码与相应的主码属于不同关系时，往往取相同的名字，以便于识别

1. 参照完整性规则

若属性（或属性组）F是基本关系R的外码，它与基本关系S的主码K s相对应（基本关系R和S不一定是不同的关系），则对于R中每个元组在F上的值必须为：或者取空值（F的每个属性值均为空值）、或者等于S中某个元组的主码值

1. 用户定义完整性

针对某一具体关系数据库的约束条件，反映某一具体应用所涉及的数据必须满足的语义要求

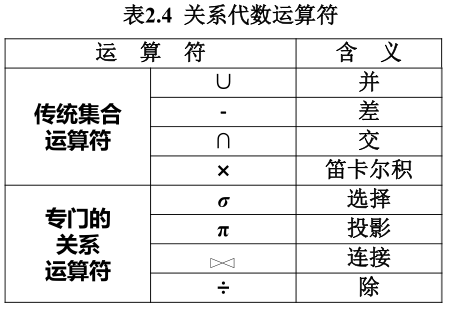
例如：“课程号”属性必须取唯一值、“学分”属性只能取值{1，2，3，4}

关系模型应提供定义和检验这类完整性的机制，以便用统一的数据库管理系统的方法处理它们，而不要由应用程序承担这一功能

1. 关系代数
2. 关系代数是一种抽象的查询语言，它用关系运算来表达查询

**运算对象是关系，运算结果亦为关系**

关系代数的运算符有两类：集合运算符和专门的关系运算符



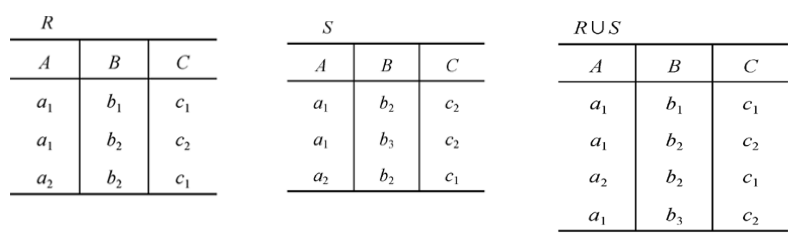
1. 并

前提： 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 ∪ S：仍为 n 目关系，由属于 R 或属于 S 的元组组成

R ∪ S = { t | t  R ∨ t  S }

重复的要去掉

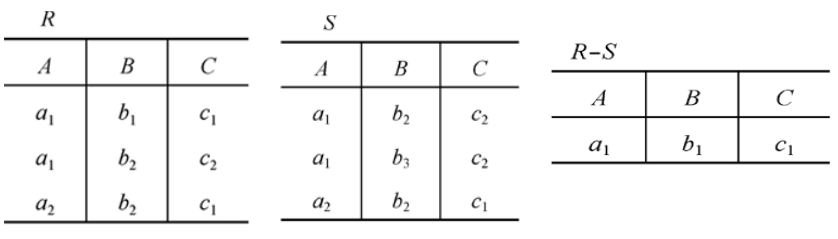


1. 差

前提： 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 – 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 | t  R ∧ t  S



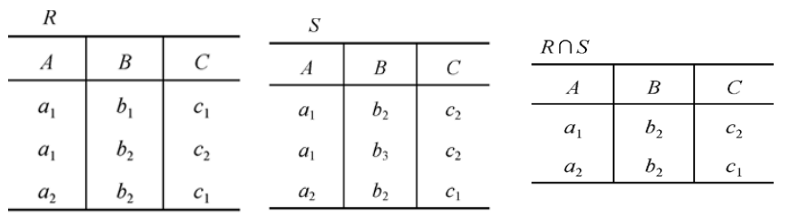
1. 交

前提： 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∩S：仍为n目关系，由既属于R又属于S的元组组成

R ∩ S = { t | t  R ∧ t  S }

R ∩ S = R – ( R – S



1. 笛卡尔积

严格地讲应该是广义的笛卡尔积

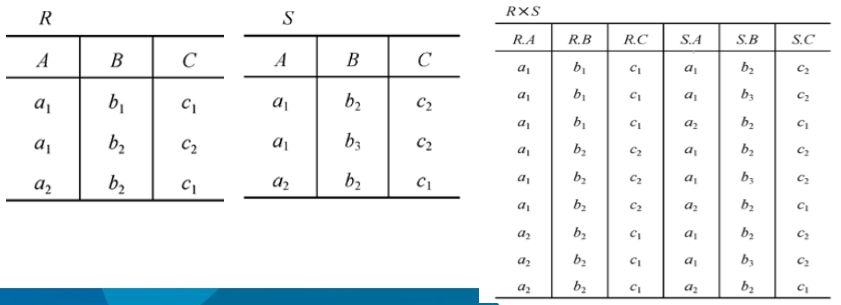
R: n目关系，k 1 个元组

S: m目关系，k 2 个元组

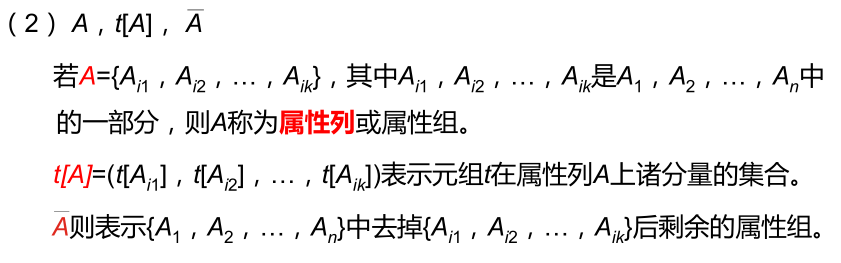
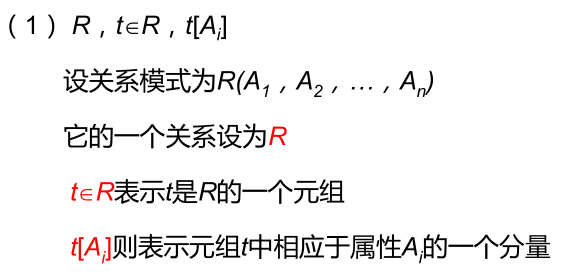
R×S：

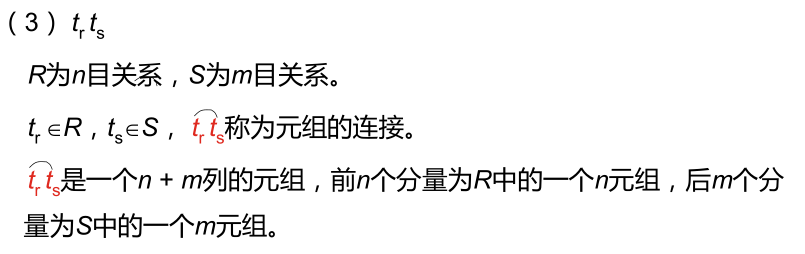
列：（n+m）列元组的集合，元组的前n列是关系R的一个元组，后m列是关系S的一个元组

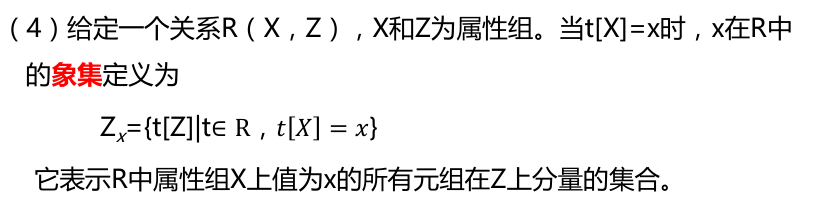
行：k 1 ×k 2 个元组，R×S = {t r t s |t r R ∧ t s S }

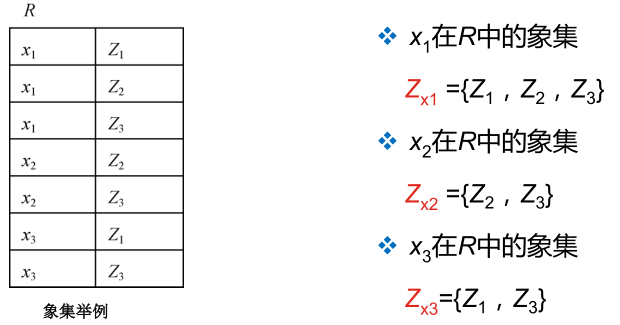


1. 专门的关系运算的基础符号









1. 选择

选择又称为限制

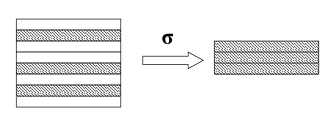
选择运算符的含义：在关系R中选择满足给定条件的**元组**

σ F (R) = {t|tR∧F(t)= True}

F：选择条件，是一个逻辑表达式，基本形式为： X 1 θY 1

Θ：表示比较运算符，它可以是＞，≥，＜，≤，＝或<

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**



Student关系上查询信息系（IS系）全体学生



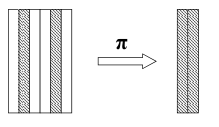
1. 投影

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

π A(R) = { t[A] | t R }

A：R中的属性列

**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）



求Student关系上学生姓名和所在系两个属性上的投影



1. 连接

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



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×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

1. 等值连接

等值连接（equijoin）：θ为“＝”的连接运算称为等值连接

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，等值连接为：

1. 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**

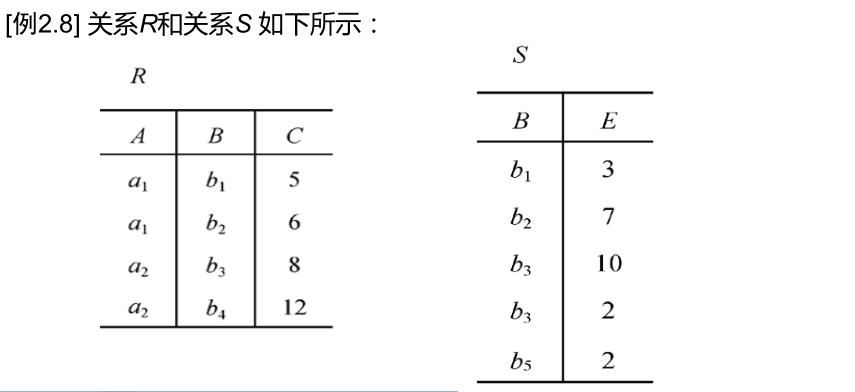
**在结果中把重复的属性列去掉**

自然连接的含义：R和S具有相同的属性组B

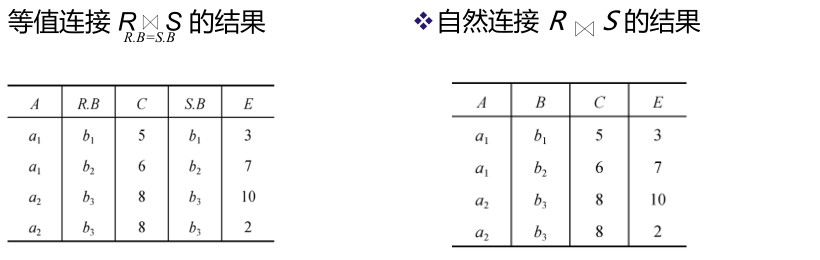


一般的连接操作是从行的角度进行运算，自然连接还需要取消重复列，所以是同时从行和列的角度进行运算。









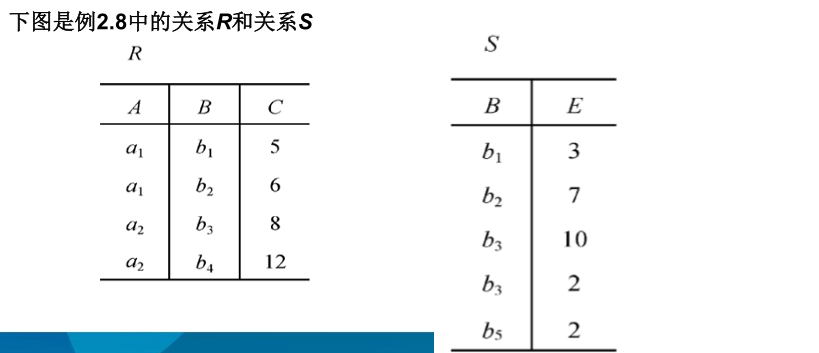
1. 外连接

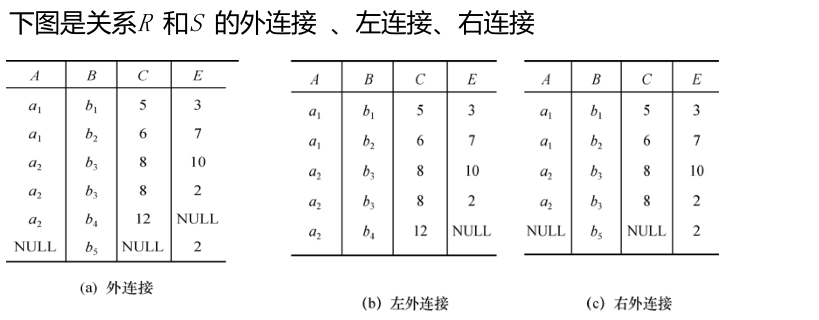
悬浮元组：两个关系R和S在做自然连接时，关系R中某些元组有可能在S中不存在公共属性上值相等的元组，从而造成R中这些元组在操作时被舍弃了，这些被舍弃的元组称为悬浮元组

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

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

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





1. 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



