



# 第六章 关系数据库理论

## ■ 引言

- \* 关系数据库系统由三级结构 / 两级映象的体系结构构成。关系模式是指关系数据库三级结构中的模式，即数据库的逻辑结构。
- \* 模式是数据库的整体结构，由所有的基本表的结构构成的整体就是模式。视图是子模式，它是模式的子集，是面向某一具体应用的。
- \* 内模式是对数据的存储形式，它由一组物理文件构成，每一个物理文件对应一个或多个基本表。
- \* 从存储角度看，数据要求冗余度小、共享性高。解决办法是进行**模式分解**，得到一组关系模式。使用时将多个关系模式进行自然连接，构成完整的**关系模式**。
- \* 怎样的分解才是合理的？答案是**关系数据库理论**，它用来指导关系数据库模式设计，判断数据库设计的好坏。



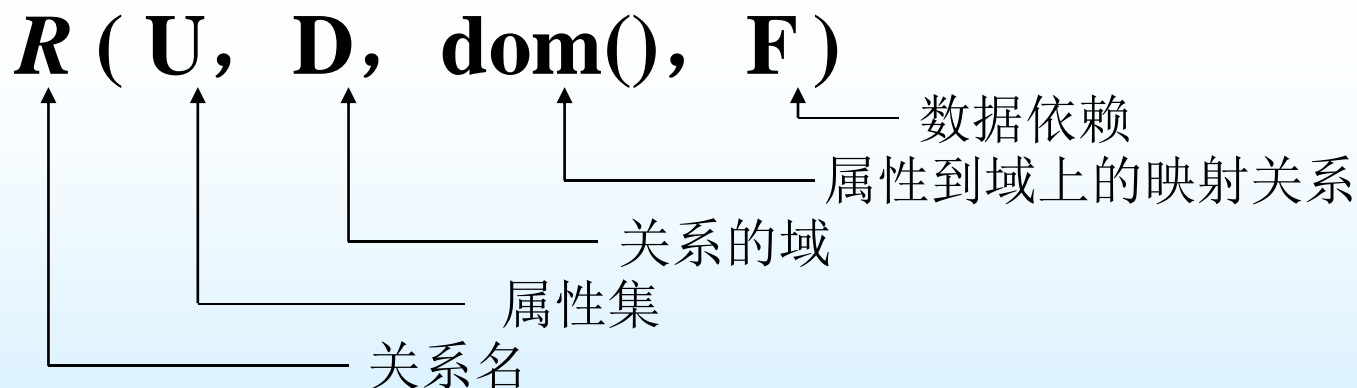
## 6.1 关系模式设计引论

### 👉 关系的外延和内涵

外延：关系模型的值，即关系模型中的数据，称为关系、实例或当前值。它随着时间的推移而变化，主要由元组的插入、删除和修改而引起。外延是动态的。

内涵：是对关系、属性、域的定义和说明，即关系模型的型的定义。

一个完整的关系模型内涵的定义：



关系的内涵称为关系模式。它是静态的，与时间无关。

关系模式通常简记为： $R(U, F)$ 。



# 6.1 关系模式设计引论

## ☞ 关系模式的存储异常

例：描述学校的数据库有如下属性：

学生的学号(Sno)、所在系(Sdept)、系主任姓名(Mname)、课程名(Cname)、成绩(Grade)

关系模式：Student ( Sno, Sdept, Mname, Cname, Grade )

语义：

1. 一个系有若干学生， 一个学生只属于一个系；
2. 一个系只有一名主任；
3. 一个学生可以选修多门课程， 每门课程有若干学生选修；
4. 每个学生所学的每门课程都有一个成绩。



## 6.1 关系模式设计引论

关系模式 Student (Sno, Sdept, Mname, Cname, Grade)

存在的问题：

1. 数据冗余太大

浪费大量的存储空间。

例：每一个系主任的姓名重复出现

2. 更新异常 (Update Anomalies)

数据冗余，更新数据时，维护数据完整性代价大。

例：某系更换系主任后，系统必须修改与该系学生有关的每一个元组

3. 插入异常 (Insertion Anomalies)

该插的数据无法插入到表中。

例：如果一个系刚成立，尚无学生，我们就无法把这个系及其系主任的信息存入数据库。



## 6.1 关系模式设计引论

### 4. 删除异常 (Deletion Anomalies)

不该删除的数据不得不删。

例：如果某个系的学生全部毕业了，在删除该系学生信息的同时，把这个系及其系主任的信息也丢掉了。

结论：

- Student关系模式不是一个好的模式。
- 所谓“好”的模式：不会发生插入异常、删除异常、更新异常，数据冗余应尽可能少。

原因：由存在于模式中属性间的**某些依赖关系**引起的。

如Mname的取值依赖于Sdept的值，通过**模式分解**来消除其中不合适的依赖：两个关系模式Student ( Sno, Sdept, Cname, Grade )和 Dept ( Sdept, Mname )。