 X→Y："X函数决定Y"，"Y函数依赖于X"。

 A⊆B   A包含于B，A小，B大，B→A

元组：二维表中的行

属性：二维表中的列

超键：能唯一标识元祖的属性集 （能唯一标识二维表中的行的列集）

候选键：不含多余属性的超建

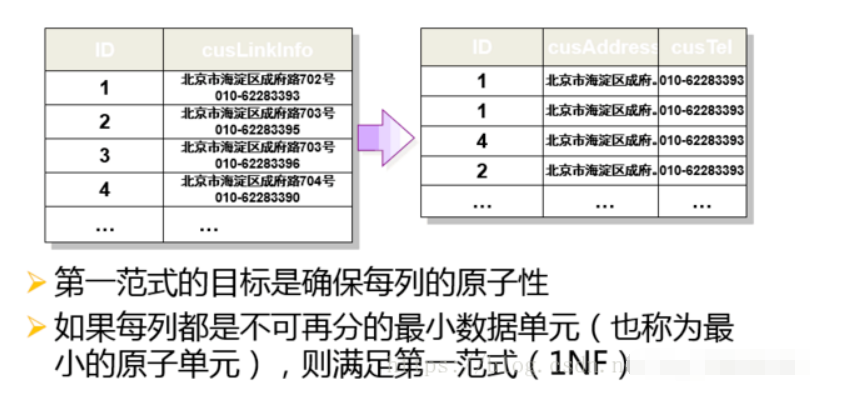
主属性：构成候选键的属性

主键：用户选作元组标识的候选键

外键：对于当前模式而言，是另一模式下的主键。

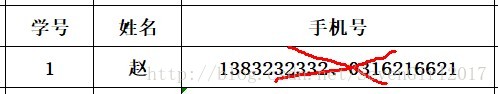
局部依赖与完全依赖：对于fd： W→A，如果存在X包含于W，有X→A成立，那么称W→A是局部依赖，否则成W→A是完全依赖。

**1NF：每个关系r的属性值为不可分的原子值（原子性）**



当赵同学有两个手机号时，他不能将两个手机号存储在一个属性框中，需要分开存放，如下表所示。

 错误：



       正确一：



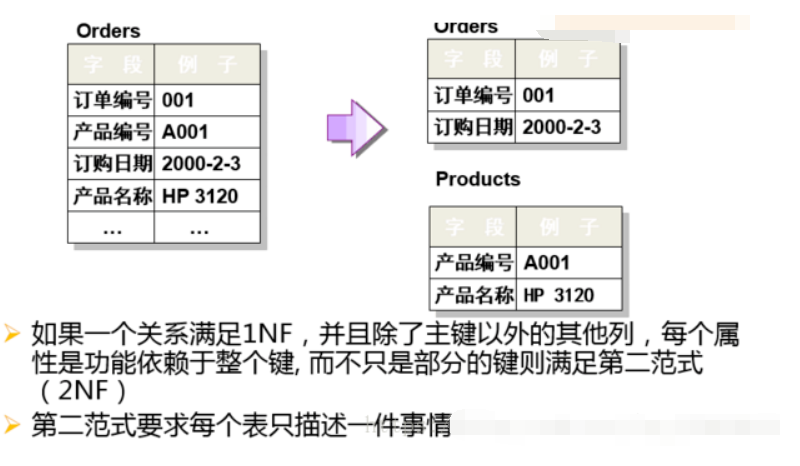
       正确二：

****

**2NF：满足1NF，非主属性完全函数依赖于候选键(左部不可约)**

要求表中的所有列，都必须依赖于主键，而不能有任何一列与主键没有关系，也就是说一个表只描述一件事情；

（sno, pno）->sname     sno->sname   pno->sname     =>sname部分函数依赖于（sno, pno）



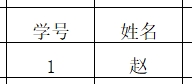
 若一张表的数据包括：“学号、姓名、课程号、授课老师”中，设“学号、课程号”为主键，其中，一门课程可以有多个老师进行授课。会存在如下关系：



(学号、课程号)→姓名

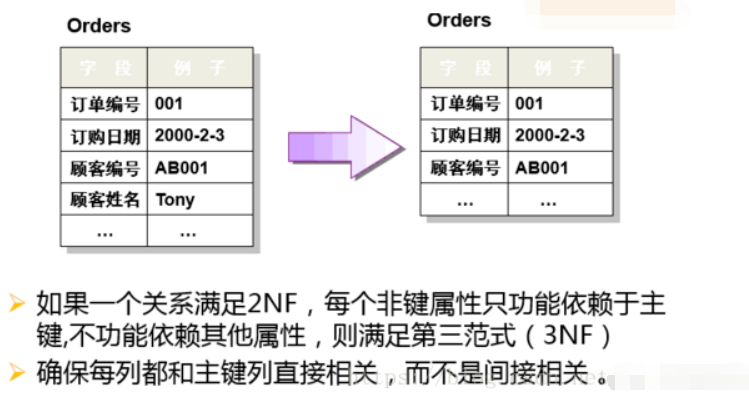
学号→姓名  ---------为局部依赖，即候选键的一部分可以推出非主属性系名

可分解为两个表，达到完全依赖：“学号、姓名”与“学号、课程号、授课老师”



**3NF： 满足2NF，消除非主属性对候选键的传递依赖**

在2NF基础上，任何非主[属性](https://baike.baidu.com/item/%E5%B1%9E%E6%80%A7)不依赖于其它非主属性（在2NF基础上消除传递依赖



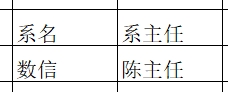
若一张表的数据包括：“学号、系名、系主任”，其中“学号”为主键，存在如下关系：



学号→系名→系主任

学号→系主任 ---------为传递依赖

同样可分解为两张表：“学号、系名”和“系名、系主任”

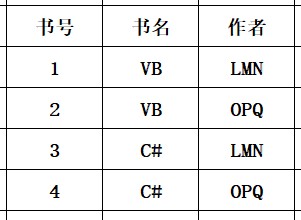
**          **

 对于第三范式，我们反过来理解也是可以的，在表1（学号、系名），表2（系名、系主任）中，学号和系名都是各自表中的主键，所以系名依赖于学号，系主任依赖于系名。当三个数据放置在一张表中时，学号是可以推出系主任的。你可以理解为通过看学生张小二的学号，是可以推理出他的系主任是谁的。

**BCNF：满足3NF，消除每一属性（主属性和非主属性）对候选键的传递依赖**

任何非主属性不能对主键子集依赖（在3NF基础上消除对主码子集的依赖）

若一张表的数据包括：“书号、书名、作者”其中，书号是唯一的，书名允许相同，一个书号对应一本书。一本书的作者可以多个，但是同一个作者所参与编著的书名应该是不同



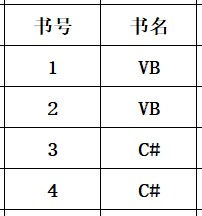
存在关系：书号→书名  (书名、作者)→书号

其中，每一个属性都为主属性，但是上述关系存在传递依赖，不能是BCNF。即：

                  (书名、作者)→书号→书名

                   (书名、作者)→书名

我们可以通过分解为两张表，实现BCNF。



**4NF：满足BCNF，消除非平凡且非FD的多值依赖(MVD)**

**多值依赖的定义：**

**设R(U)是一个属性集合U上的一个关系模式，X, Y, 和Z是U的子集，并且Z=U-X-Y，多值依赖X->->Y成立当且仅当对R的任一个关系r，r在(X,Z)上的每个值对应一组Y的值，这组值仅仅决定于X值而与Z值无关。**

**若X->->Y，而Z=空集，则称X->->Y为平凡的多值依赖。否则，称X->->Y为非平凡的多值依赖。**

可以看出，如果把上面的一组改为一个，那么多值依赖就变成了函数依赖。当然一个值组成的组也是组，所以说，函数依赖是多值依赖的特殊情况。

**课程C                                   教师T                              参考书B**

**数学                                      邓军                                数学分析**

**数学                                      邓军                                高等代数**

**数学                                      邓军                                微分方程**

表中，U = C+T+B，（C,T）确定一组B,但是这组B其实与T无关，仅由C确定，所以(C,T)->->B。又因为T不是空集，所以(C,T)->->B为非平凡多值依赖。

 要想消除多只依赖，可以分解为：(C,T), (C,B)及

表1：

                  课程C                  教师T

                  数学                     邓军

表2:

                课程C                    参考书B

                数学                      数学分析

                数学                      高等代数

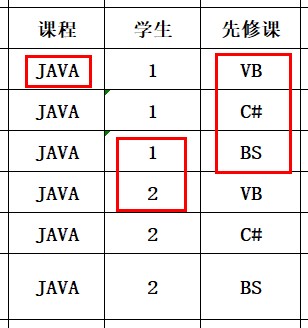
                数学                      微分方程

对于R中的每个非平凡多值依赖X->->Y（Y不属于X)，X都含有候选码，则R属于4NF。

分析：对于每一个非平凡多值依赖X->->Y，X若含有候选码，也就是X->Y，所以4NF所允许的非平凡多值依赖是函数依赖。

非形式说：只要两个独立的1：N联系出现在一个关系中，那么就可能出现多值依赖。举例说明。

 一个表中存在三个数据：“课程、学生、先修课”。假设2017级的计算机专业学生想要学习JAVA课程，那么他们需要先学习VB、C#、BS三门课，才可以选择进行JAVA课程。存在关系：



课程→学生

课程→先修课

两个均是1：N的关系，当出现在一张表的时候，会出现大量的冗余。所以就我们需要分解它，减少冗余。(Ps：该例子主要是为了说明概念帮助理解，具体应用中不会只是这样的简单粗暴的。)



三、E-R模型的基本要素

1) 实体 ：实体是E-R模型的基本对象，是现实世界中各种事物的抽象，凡是可以相互区别，并可以被识别的事、物概念等均可认为是实体。在一个单位中，具有共性的一类实体可以划分为一个实体集，例如，学生李明，黄颖等都是实体，为了便于描述，可以定义学生这样的一个实体集，所有学生都是这个集合的成员。

2)属性：每个实体都具有各种特征，称其为实体的属性，如学生有学号，姓名，年龄等属性。实体的属性值是数据库存储的主要数据。能唯一标识实体的属性或属性组称为实体键，如一个实体有多个键存在。则可从中选取一个作为主键。

3)联系：实体间会存在各种关系，如人与人之间可能存在领导与雇员关系等，实体间的关系被抽象为联系。

四、E-R图的四个组成的部分

矩形框:表示实体，在矩形框中写上实体的名字

椭圆形框:表示实体或联系的属性

菱形框:表示联系，在框中记入联系名

连线:实体与属性之间；实体与联系之间；联系与属性之间用直线相连，（对于一对一联系，要在两个实体连线方向各写1； 对于一对多联系，要在一的一方写1，多的一方写N；对于多对多关系，则要在两个实体连线方向各写N,M。)。

