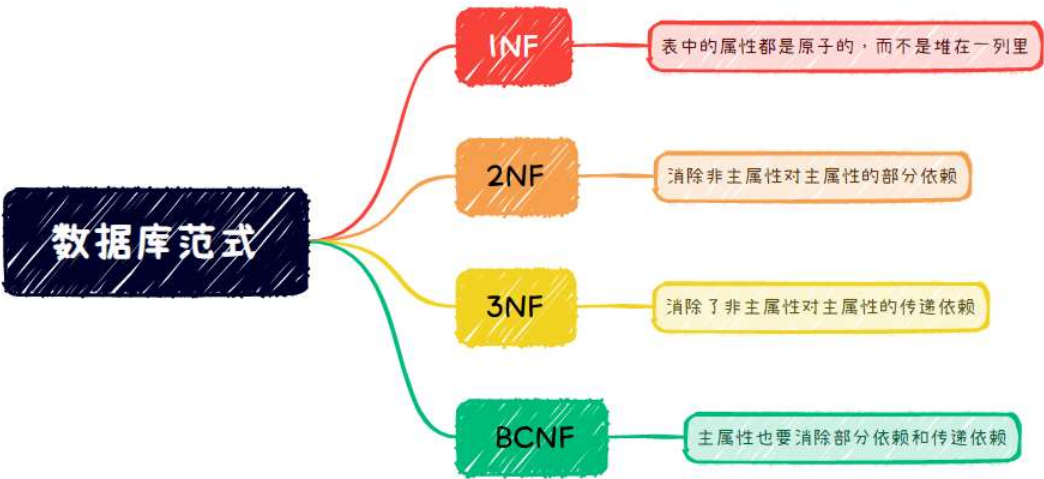


# 数据库范式1NF~BCNF超详细总结



CSDN @关山煮酒

省流图

## 目录

- 1.前置知识
  - 键
    - 超键
    - 候选键
    - 主键
    - 外键
  - 函数依赖
    - 完全函数依赖
    - 部分函数依赖
    - 传递依赖
  - 属性
    - 主属性
    - 非主属性
- 2.数据范式
  - 1NF
  - 2NF
  - 3NF
  - BCNF
- 3.总结

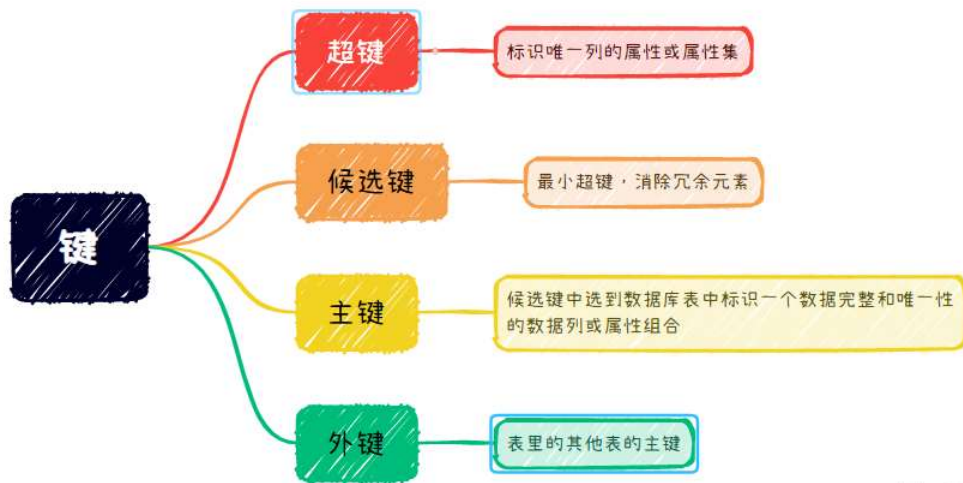
设计一个数据库系统时，我们实际上是在设计不同实体之间的关系。那么设计关系时所循行的标准，称为范式。具体有哪些标准呢？我们一般将数据库范式分为1NF，2NF，3NF，BCNF，4NF，5NF。级别由低到高。符合高一级别的范式，必定符合第一级别的范式。一般我们掌握到BCNF即可！

## 1.前置知识

学号	姓名	系名	系主任	班级	选课	分数
2000001	张三	电子系	钱峰	一年1班	高等数学	87
2000001	张三	电子系	钱峰	一年1班	大学英语	67
2000001	张三	电子系	钱峰	一年1班	线性代数	89
2000002	李四	电子系	钱峰	一年2班	高等数学	78
2000002	李四	电子系	钱峰	一年2班	大学英语	68
2000002	李四	电子系	钱峰	一年2班	计算机基础	90
2000003	王五	计算机系	李玲	一年3班	高等数学	85

图1

键



CSDN @关山煮酒

键省流图

键是一个属性集合

## 超键

关系中的某个或者某几个属性的集合，用于唯一地标识每一条数据（这里的每一条数据就是数据库中的每一条记录）。超键包含**候选键**和**主键**。

比如图1中知道了<学号，选课>这两个属性的集合作为的键，那么就可以唯一定位到一条记录，而不会定位到两条或更多条记录。就像身份证一样，每一条记录都有属于自己的唯一标识！

## 候选键

是最小超键，即没有**冗余**元素的超键。

比如图1中知道了<学号，姓名，选课>，也可以唯一确定一条记录，即这个集合也是一个超键，但是他有冗余元素姓名，所以不是一个候选键。<学号，选课>则是缺一不可的，只有学号或选课都可能对应到多条记录。

## 主键

数据库表中对储存数据对象予以**唯一和完整标识的数据列或属性的组合**。一个数据列只能有一个**主键**，且**主键的取值不能缺失**，即不能为空值（Null）。主键是从候选键中选择的。

比如我在图1中添加一列recordId作为自增主键

## 外键

在一个表中存在的另一个表的主键称此表的外键

比如我另外建一个表，里面有一列叫recordId，这一列的recordId和图1中的表关联，那么这个recordId就是这个表的外键。

## 函数依赖



CSDN @关山煮酒

函数依赖省流图

数学上的解释是： $y = f(x)$ ，输入一个  $X$ ，可以得到一个确定的  $Y$ 。

对应到一个表上就是，在属性X 确定的情况下，必定能够确定某个属性 Y 的值，这就能够称作 Y 函数依赖于 X，写作  $X \rightarrow Y$ 。

注意：X未必是一个属性，也可能是属性集合。

比如下面是图1中出现的部分函数依赖：

- (学号)  $\rightarrow$  (姓名)
- (学号, 课名)  $\rightarrow$  (分数)
- (系名)  $\rightarrow$  (系主任)

完全函数依赖

在一个表中，如果存在  $X \rightarrow Y$ ，那么对于 X 下的任何一个真子集 (X')， $X' \rightarrow Y$ 都不成立，那么我们就说 Y 对于 X 完成函数依赖。

通俗的说，必须通过键中的所有属性才可以唯一确定一个值。比如：

- (学号)  $\rightarrow F$  (姓名)
- (学号, 课名)  $\rightarrow F$  (分数)

部分函数依赖

在一个表中，如果存在  $X \rightarrow Y$ ，但是 Y 并不完全依赖于 X。存在一些 X 的子集 X'， $X' \rightarrow Y$ 成立，那么我们就说 Y 对于X 部分函数依赖。

记作:  $X \twoheadrightarrow P Y$

通俗的说，只需要键中的部分属性即可唯一确定一个值。比如：

- (学号, 课名)  $\rightarrow$  (姓名)，只需要根据键中的学号即可唯一确定姓名。

传递依赖

通俗的说，通过键可以唯一确定一个属性，然后通过该属性可以唯一确定另一个属性，所以就演变为可以通过键唯一确定一个无函数依赖的属性。

属性

属性就是我们在表中定义的每一个列。

主属性

在候选键中的所有属性（每一列）都称为主属性

非主属性

不包含在任一 候选 键 中的属性称为非主属性

2.数据范式

1NF

符合1NF的关系中的每个属性都不可再分，即表中的属性都是原子的。

反例就是就用一列，保存了所有的信息。

图1中仅仅满足了1NF，可能会出现下面的问题

学号	姓名	系名	系主任	班级	选课	分数
2000001	张三	电子系	钱峰	一年1班	高等数学	87
2000001	张三	电子系	钱峰	一年1班	大学英语	67
2000001	张三	电子系	钱峰	一年1班	线性代数	89
2000002	李四	电子系	钱峰	一年2班	高等数学	78
2000002	李四	电子系	钱峰	一年2班	大学英语	68
2000002	李四	电子系	钱峰	一年2班	计算机基础	90
2000003	王五	计算机系	李玲	一年3班	高等数学	85

图1

1. 数据冗余  
每一名学生的学号、姓名、系名、系主任这些数据重复多次。每个系与对应的系主任的数据也重复多次
2. 插入异常  
假如学校新建了一个系，但是暂时还没有招收任何学生，那么是无法将系名与系主任的数据单独地添加到数据表中。
3. 删除异常  
假如将某个系中所有学生相关的记录都删除，那么所有系与系主任的数据也就随之消失了（一个系所有学生都没有了，并不表示这个系就没有了）。
4. 修改异常  
假如李小明转系到法律系，那么为了保证数据库中数据的一致性，需要修改三条记录中系与系主任的数据

2NF

所有非主属性要完全依赖于主属性，不能是部分依赖。即消除非主属性的部分依赖。

比如图1（为了 便于阅读，我选择把图拿过来）

学号	姓名	系名	系主任	班级	选课	分数
2000001	张三	电子系	钱峥	一年1班	高等数学	87
2000001	张三	电子系	钱峥	一年1班	大学英语	67
2000001	张三	电子系	钱峥	一年1班	线性代数	89
2000002	李四	电子系	钱峥	一年2班	高等数学	78
2000002	李四	电子系	钱峥	一年2班	大学英语	68
2000002	李四	电子系	钱峥	一年2班	计算机基础	90
2000003	王五	计算机系	李玲	一年3班	高等数学	67

图1

候选键是<学号，课程>

主属性：学号、课程

非主属性：姓名、系名、系主任、分数

比如这里学号可以唯一确定姓名，不需要课程这个主属性，所以主属性的子集学号即可唯一确定姓名，所以出现了部分依赖。非主属性的依赖关系如下：

- 1. 姓名：有 学号->姓名，部分函数依赖
- 2. 系名：有 学号->系名，部分函数依赖
- 3. 系主任：有 学号->系主任，部分函数依赖  
注：显然，系主任也可以被系名唯一确定；这其实是第三范式的问题，我们稍后讨论。  
在此处，系主任也是可以被学号唯一确定的。
- 4. 分数：只有 {学号，课程}->分数，完全函数依赖

非主属性 中只有 分数 这一项完全依赖的， 所以可以按照主属性拆分表为学生表和选课表。

选课表		
学号	课名	分数
2000001	高等数学	87
2000001	大学英语	67
2000001	线性代数	89
2000002	高等数学	78
2000002	大学英语	68
2000002	计算机基础	90
2000003	高等数学	67

选课表

学生表				
学号	姓名	系名	系主任	班级
2000001	张三	电子系	钱峥	一年1班
2000002	李四	电子系	钱峥	一年2班

学生表

3NF

在2NF的基础上，消除了非主属性对主属性的传递依赖。

学生表中在系名可以由学号唯一确定，系主任也可以由学号唯一确定，可是我们总觉得有那里不对。更优雅的方式应该是：学号决定系名，而系名决定系主任。这里的关键点在于 传递函数依赖，事实上，当出现传递函数依赖的时候，我们可以将这一依赖涉及到的属性们再单独拉个表出来。

学生表			
学号	姓名	系名	班级
2000001	张三	电子系	一年1班
2000002	李四	电子系	一年2班

系表	
系名	系主任
电子系	钱峰

BCNF

所有列要消除部分依赖和传递依赖。即主属性也要消除部分依赖和传递依赖，而之前的范式要求都是针对非主属性的。以下图为例：

物品表			
仓库名	管理员	物品名	数量
上海仓	张三	iphone	30
上海仓	张三	huawei	40
北京仓	李四	iphone	30

- 候选键：（仓库名，物品名），（管理员，物品名）
  - 主属性：仓库名、物品名、管理员
  - 非主属性：数量
- 
- 依赖关系如下
    - 仓库名，物品名）-> 管理员，只要确定仓库名，即可确定管理员，所以管理员部分函数依赖于仓库名。
    - （管理员，物品名）-> 仓库名，同上。

可以拆分成物品表和仓库表

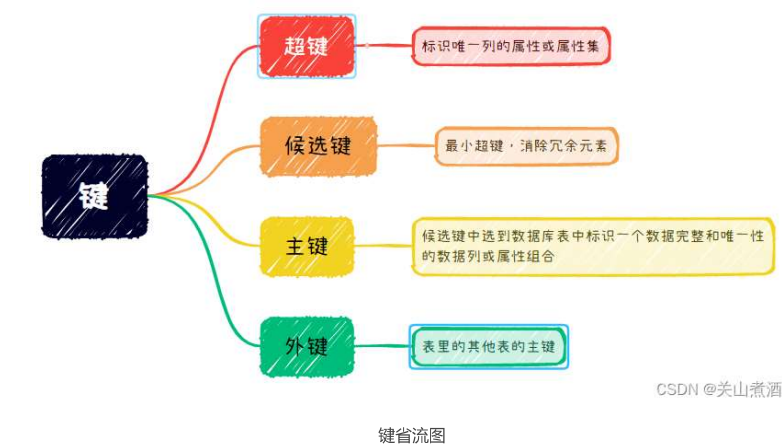
仓库表：

仓库名	管理员
上海仓	张三
北京仓	李四

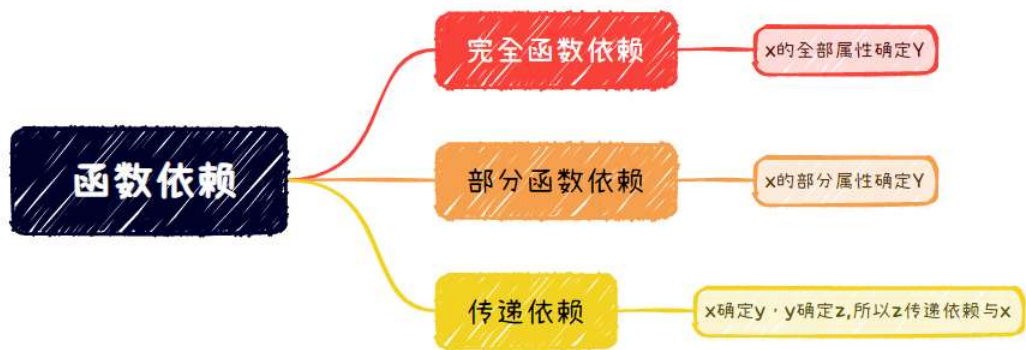
物品表：

仓库名	物品名	数量
上海仓	iphone	30
上海仓	huawei	40
北京仓	iphone	30

3.总结

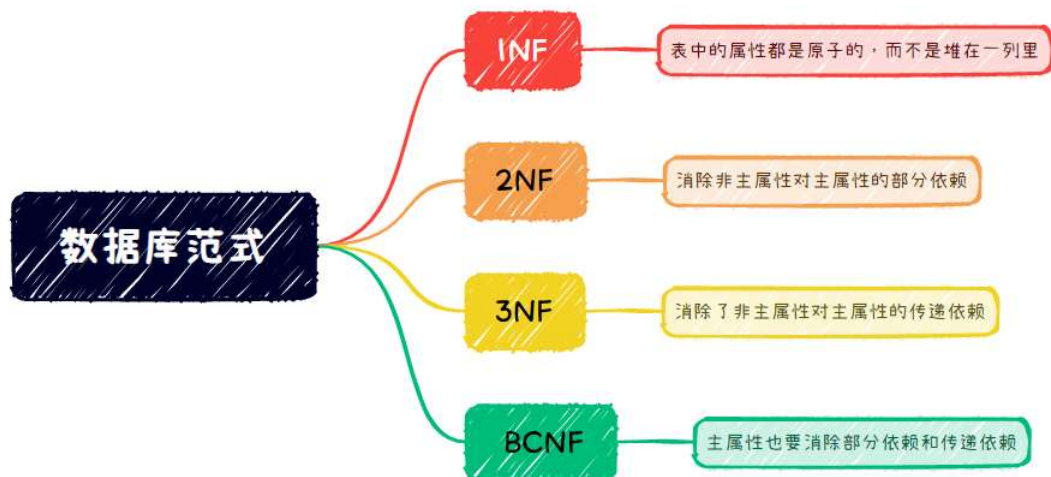






CSDN @关山煮酒

函数依赖省流图



CSDN @关山煮酒

#### 4.参考资料

- 1.超详细! 掌握数据库系统之范式 - 知乎 (zhihu.com)
- 2.Database System Concepts - 7th edition (db-book.com)