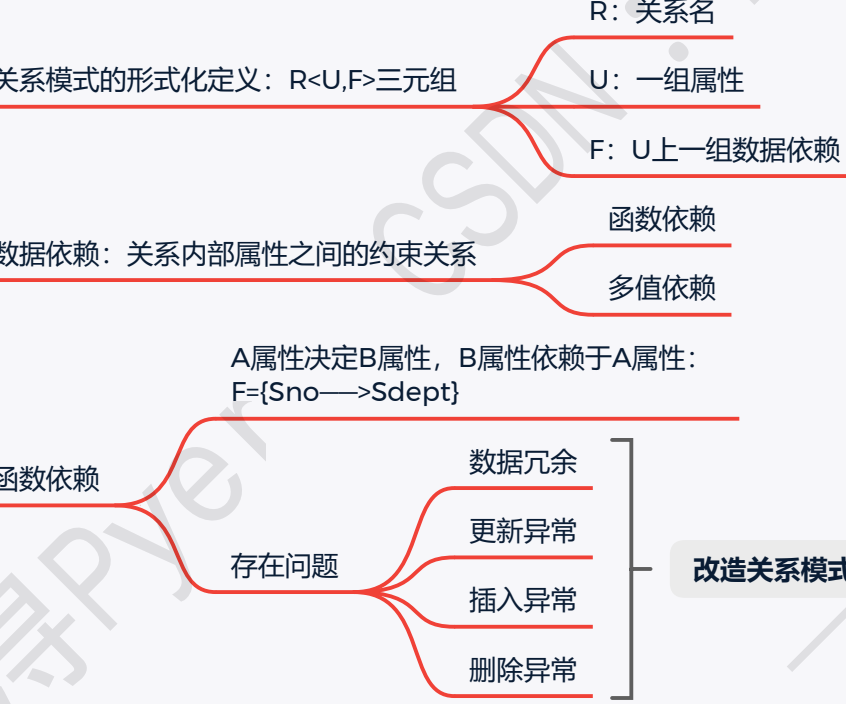


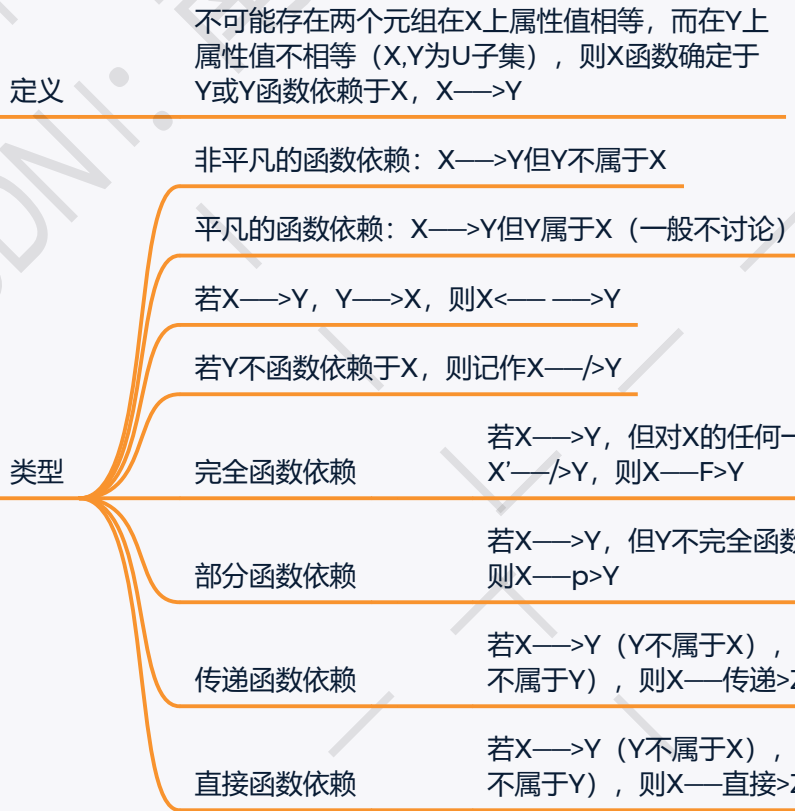
关系数据库理论

如何构造一个好的数据库模式

应该构造几个关系模式？
每个关系由哪些属性组成？

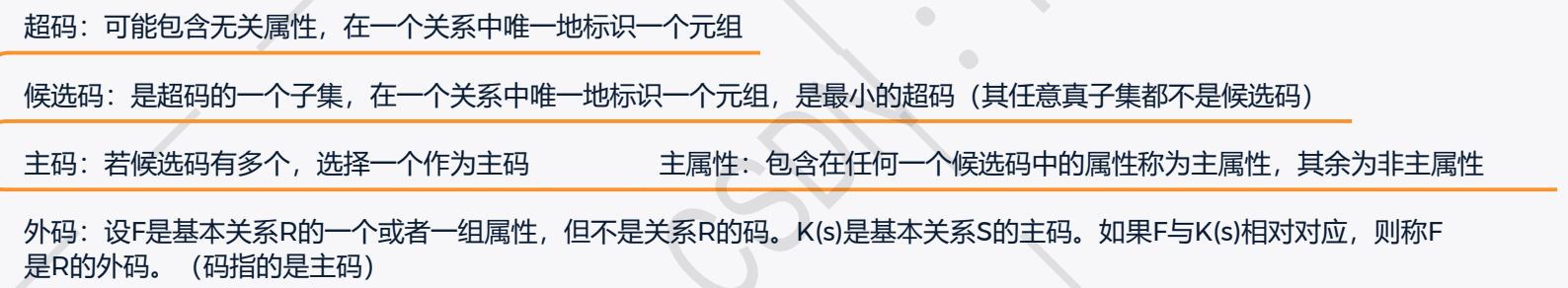


函数依赖



若 $X \rightarrow Y$ ，则称X为这个函数依赖的决定属性组，也称决定因素

码



均简称为码

规范化

一个关系属性间不同的依赖情况

如何根据依赖情况判定关系是否具有不合适的性质

按依赖情况区分关系规范化程度为1NF、2NF、3NF、4NF

如何将不合适性质的关系转为更合适的性质

范式：某一级别关系模式的集合（低级范式包含高级范式）

规范化：一个低一级的范式的关系模式通过模式分解可以转换为若干个高一级的范式的关系模式，这种过程称为规范化

基本思想：逐步消除数据依赖中不合适的部分，使模式中的各关系模式达到某种程度的“分离”。

1NF：如果一个关系模式R的所有属性都是不可分的基本数据项，则R属于1NF（关系数据库模式最基本要求）

- 插入异常：学生无选课则不能插入
- 删除异常：删了选课其他都删了
- 修改复杂：学生转系要修改Sloc

S-L-C (Sno,Sdept,Sloc,Cno,Grade)
码 (Sno, Cno) Sdept→Sloc

2NF：R属于1NF、且每一个非主属性完全函数依赖于任何一个候选码

- 数据冗余度大：学生选了k门课，Sdept、Sloc重复存储k次
- 插入异常：无学生时S-L不能插入系数据
- 删除异常：删除学生信息也将系信息删除
- 修改复杂：系调整住处要修改所有Sloc
- 数据冗余度大：相同系的学生住处信息重复存储

SC(Sno,Cno,Grade)、S-L(Sno,Sdept,Sloc)
码 (Sno, Cno) - (Sno)

3NF：R属于1NF，若R中不存在这样的码X，属性组Y及非主属性Z（Z不属于Y），使得 $X \rightarrow Y$ ， $Y \twoheadrightarrow Z$ 成立， $Y \not\rightarrow X$ ，则R属于3NF

每一个非主属性既不传递依赖于码，也不部分依赖于码，R属于3NF，则R属于2NF。

SC属于3NF

S-L中， $Sno \rightarrow Sdept$ ($Sdept \not\rightarrow Sno$)， $Sdept \twoheadrightarrow Sloc$ ，所以Sloc传递依赖于Sno

SC(Sno,Cno,Grade)、S-D(Sno,Sdept)、D-L(Sdept,Sloc)
码 (Sno, Cno) - (Sno) - (Sdept)