

数据管理基础

第6章 关系数据理论

(范式与规范化)

智能软件与工程学院



❑ 定义6.6 2NF

- 不满足2NF可能产生的异常问题

❑ 定义6.7 3NF

- 不满足3NF可能产生的异常问题

❑ 定义6.8 BCNF

- 不满足BCNF可能产生的

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 $R \in 1NF$ 。
- ❑ 第一范式是对关系模式的最起码的要求。不满足第一范式的数据库模式不能称为关系数据库。
- ❑ 但是满足第一范式的关系模式并不一定是一个好的关系模式。

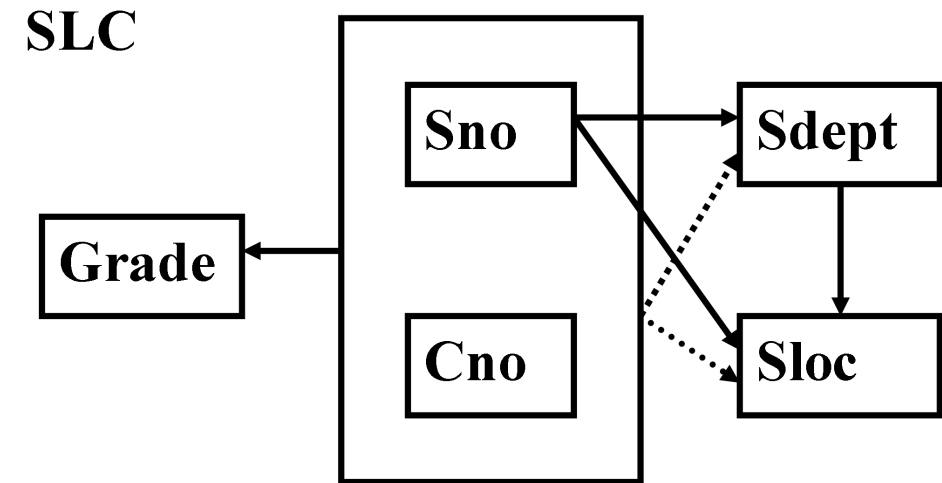
2NF (1)

- ❑ **定义6.6** 若关系模式 $R \in 1NF$ ，并且每一个非主属性都完全函数依赖于任何一个候选码，则 $R \in 2NF$
- ❑ 在一个满足2NF的关系 R 中，不允许存在如下的函数依赖：
 - A 是关系 R 的非主属性， K 是关系 R 的候选码，且 A 部分函数依赖于 K ，即 $K \xrightarrow{P} A$ 。
 - 即：不允许存在“非主属性对候选码的部分函数依赖”
- ❑ 判断一个关系 R 是否满足2NF，需要按照以下步骤检查：
 - ① 找到关系 R 的所有非主属性和所有的候选码；
 - ② 检查每一个非主属性 A 和每一个候选码 K 之间的函数依赖，判断是否存在‘非主属性对于候选码的部分函数依赖’。

❑ [例6.4] S-L-C(Sno,Sdept,Sloc,Cno,Grade), Sloc为学生的住处, 并且每个系的学生住在同一个地方。S-L-C的码为(Sno,Cno)。

➤ 函数依赖有

- $(Sno, Cno) \xrightarrow{F} Grade$
- $Sno \rightarrow Sdept, (Sno, Cno) \xrightarrow{P} Sdept$
- $Sno \rightarrow Sloc, (Sno, Cno) \xrightarrow{P} Sloc$
- $Sdept \rightarrow Sloc$



➤ 非主属性Sdept和Sloc部分函数依赖于码(Sno, Cno)

➤ 关系模式S-L-C不属于2NF

2NF (3)

□ 一个关系模式不属于2NF，会产生以下问题：

➤ 数据冗余

➤ 修改复杂，更新异常

- 如果一个学生选多门课，则Sdept, Sloc被存储了多次。如果该生转系，则需要修改所有相关的Sdept和Sloc，造成修改的复杂化

➤ 插入异常

- 如果插入一个新学生，但该生未选课，即该生无Cno，由于插入元组时，必须给定码值，因此插入失败。

➤ 删除异常

- 如果S4只选了一门课C3，现在他不再选这门课，在删除学生S4关于C3的选课元组后，元组中学生S4的信息也被删除了。

2NF (4)

❑ 出现这种问题的原因

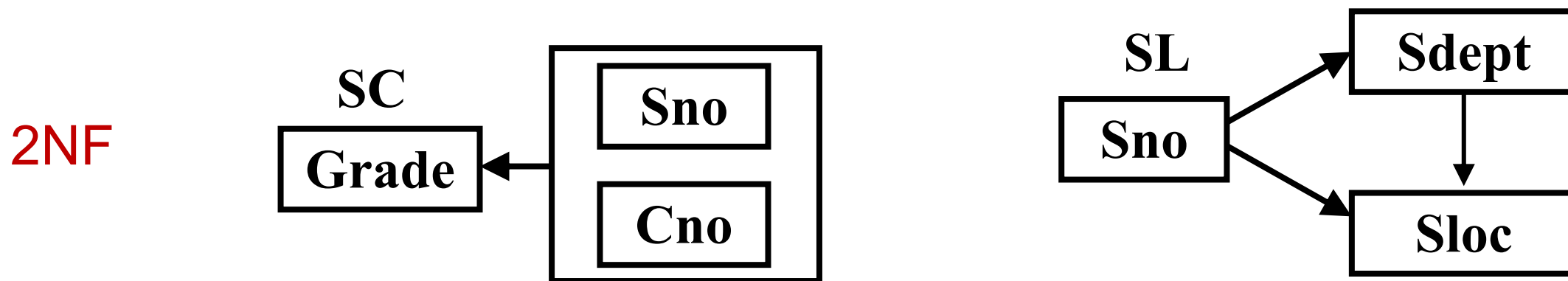
➤ 例子中有两类非主属性：

- 一类如**Grade**，它对码是完全函数依赖
- 另一类如**Sdept**、**Sloc**，它们对码不是完全函数依赖

❑ 解决方法：

➤ 用投影分解把关系模式**S-L-C**分解成两个关系模式

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



2NF (5)

❑ 规范到满足2NF后，模式设计问题的改进情况（以学生为例）

➤ 数据冗余

- 有改进

➤ 修改复杂，更新异常

- 有改进，如，修改学生院系，只需修改SL表格

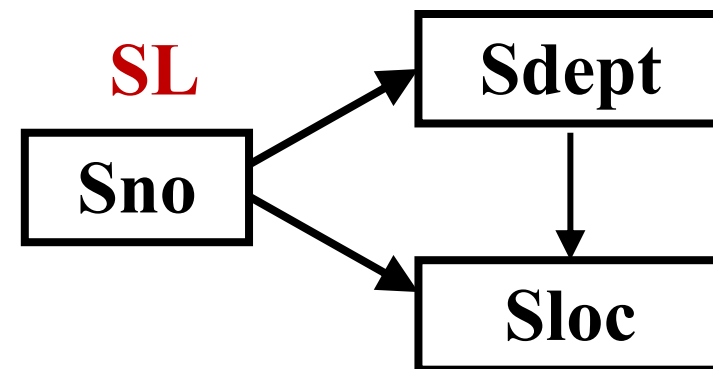
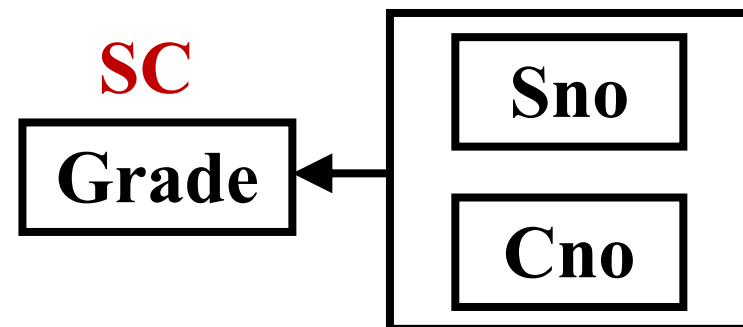
➤ 插入异常

- 仍然存在

➤ 删除异常

- 仍然存在

(2NF)

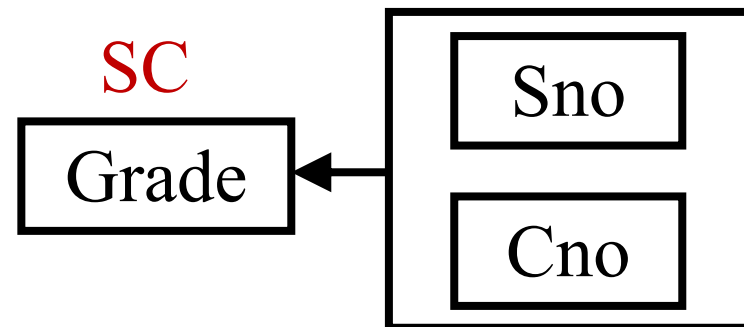


3NF (1)

❑ 定义6.7 设关系模式 $R(U, F) \in 1NF$ ，若 R 中不存在这样的码 X 、属性组 Y 及非主属性 Z ($Z \not\subseteq Y$)，使得 $X \rightarrow Y$ ， $Y \rightarrow Z$ 成立， $Y \nrightarrow X$ ，则称 $R(U, F) \in 3NF$ 。

❑ 以关系SC为例： $SC \in 3NF$

➤ SC中没有传递依赖



❑ 以关系S-L为例： $S-L \notin 3NF$

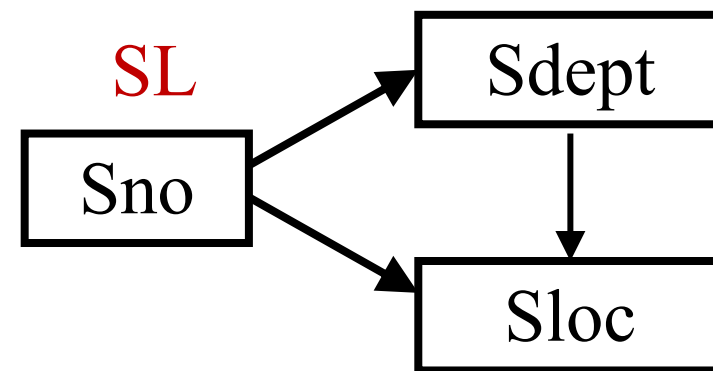
➤ S-L中有 $Sno \rightarrow Sdept$, $Sdept \nrightarrow Sno$, $Sdept \rightarrow Sloc$

➤ 可得： $Sno \xrightarrow{\text{传递}} Sloc$

➤ 解决的办法是将S-L分解成

• S-D ($Sno, Sdept$) $\in 3NF$

• D-L ($Sdept, Sloc$) $\in 3NF$



3NF (2)

□ 规范到满足2NF后，模式设计问题的改进情况（以学生为例）

➤ 数据冗余

● 有改进

➤ 修改复杂，更新异常

● 有改进

➤ 插入异常

● 有改进

➤ 删除异常

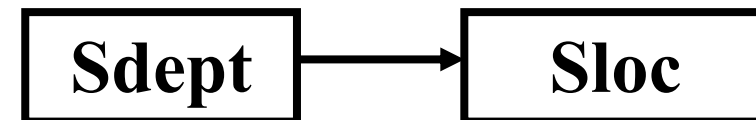
● 有改进

3NF

SD



DL



3NF的问题

| 仓库名 | 管理员 | 物品名 | 数量 |
|-----|-----|--------|----|
| 上海仓 | 张三 | iPhone | 30 |
| 上海仓 | 张三 | iPad | 40 |
| 北京仓 | 李四 | iPhone | 50 |
| 北京仓 | 李四 | iMac | 60 |

- ❑ 先新增加一个仓库，但尚未存放任何物品，是否可以为该仓库指派管理员？——不可以，因为物品名也是主属性，根据实体完整性的要求，主属性不能为空。
- ❑ 某仓库被清空后，需要删除所有与这个仓库相关的物品存放记录，会带来什么问题？——仓库本身与管理员的信息也被随之删除了。
- ❑ 如果某仓库更换了管理员，会带来什么问题？——这个仓库有几条物品存放记录，就要修改多少次管理员信息。

BCNF (1)

- ❑ **BCNF (Boyce Codd Normal Form)** 由Boyce和Codd提出，比3NF更进了一步。通常认为BCNF是修正的第三范式，有时也称为扩充的第三范式。
- ❑ **定义6.8** 设关系模式 $R(U, F) \in 1NF$ ，若 $X \rightarrow Y$ 且 $Y \not\subseteq X$ 时 X 必含有码，则 $R(U, F) \in BCNF$ 。
- ❑ 换言之，在关系模式 $R(U, F)$ 中，如果每一个决定属性集都包含候选码，则 $R \in BCNF$ 。

□ BCNF的关系模式所具有的性质

- 所有非主属性都完全函数依赖于每个候选码
- 所有主属性都完全函数依赖于每个不包含它的候选码
- 没有任何属性完全函数依赖于非码的任何一组属性

□ 如果一个关系数据库中的所有关系模式都属于BCNF，那么在函数依赖范畴内，它已实现了模式的彻底分解，达到了最高的规范化程度，消除了插入异常和删除异常。

□ [例6.5] 考察关系模式 $C(Cno, Cname, Pcnno)$

- 它只有一个码 Cno ，没有任何属性对 Cno 部分依赖或传递依赖，所以 $C \in 3NF$ 。
- 同时 C 中 Cno 是唯一的决定因素，所以 $C \in BCNF$ 。
- 对于关系模式 $SC(Sno, Cno, Grade)$ 可作同样分析。

□ [例6.6] 关系模式 $S(Sno, Sname, Sdept, Sage)$,

- 假定 $Sname$ 也具有唯一性，那么 S 就有两个码，这两个码都由单个属性组成，彼此不相交。
- 其他属性不存在对码的传递依赖与部分依赖，所以 $S \in 3NF$ 。
- 同时 S 中除 Sno , $Sname$ 外没有其他决定因素，所以 $S \in BCNF$ 。

BCNF (4)

❑ [例6.7] 关系模式 $SJP(S, J, P)$ 中， S 是学生， J 表示课程， P 表示名次。每一个学生选修每门课程的成绩有一个确定的名次，每门课程中每一名次只有一个学生（即没有并列名次）。

➤ 由语义可得到函数依赖： $(S, J) \rightarrow P$ 和 $(J, P) \rightarrow S$

➤ (S, J) 与 (J, P) 都可以作为候选码。

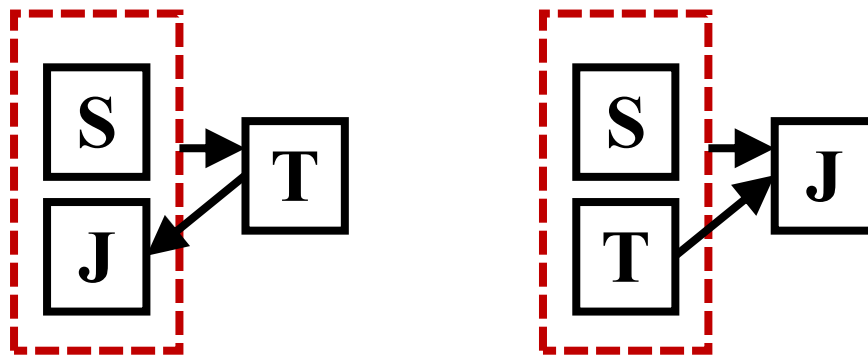
➤ 关系模式中没有属性对码传递依赖或部分依赖，所以 $SJP \in 3NF$

➤ 除 (S, J) 与 (J, P) 以外没有其他决定因素，所以 $SJP \in BCNF$

BCNF (5)

- ❑ [例6.8] 关系模式STJ(S,T,J)中，S表示学生，T表示教师，J表示课程。每一教师只教一门课。每门课有若干教师，某一学生选定某门课，就对应一个固定的教师。

➤ 由语义可得到函数依赖： $(S,J) \rightarrow T$ $T \rightarrow J$ $(S,T) \xrightarrow{P} J$



- 没有任何非主属性对码传递依赖或部分依赖，所以 $STJ \in 3NF$
- 因为T是决定因素，而T不包含码，所以 $STJ \notin BCNF$

BCNF (6)

❑ 非BCNF的关系模式也可以通过分解成为BCNF。

➤ 如例6.8中，STJ可分解为ST(S,T)与TJ(T,J)，它们都是BCNF。

❑ 3NF和BCNF是在函数依赖的条件下对模式分解所能达到的分离程度的测度。

➤ 一个模式中的关系模式如果都属于BCNF，那么在函数依赖范畴内，它已实现了彻底的分离，已消除了插入和删除的异常。

➤ 3NF的“不彻底”性表现在可能存在主属性对码的部分依赖和传递依赖。

6. 范式定义

- ① 请写出下列范式的定义：1NF, 2NF, 3NF, BCNF
- ② 简要说明，上述各级范式之间的相互关系。

7. 在一个关系模式中，会因为存在下列函数依赖而产生不合理的数据冗余存储和操作异常。请利用函数依赖、候选码等概念的定义来解释其原因。

- ① 存在非主属性A对候选码K的部分函数依赖： $K \xrightarrow{p} A$
- ② 存在非主属性A对候选码K的传递函数依赖： $K \rightarrow X, X \not\subset K, X \twoheadrightarrow A, X \rightarrow A$
- ③ 存在主属性B对候选码K的部分函数依赖： $K \xrightarrow{p} B$;
- ④ 存在主属性B对候选码K的传递函数依赖： $K \rightarrow X, X \not\subset K, X \twoheadrightarrow B, X \rightarrow B$