

数据库系统原理

第6章 关系数据理论1



分级通关平台,QQ交流



□ 基础篇

第1章 绪论

第2章 关系数据库*3

第3章 标准语言SQL*3

第4章 数据库安全性

第5章 数据库完整性

实验1

□ 设计与应用开发篇

第6章 关系数据理论*2

第7章 数据库设计

实验2

第8章 数据库编程

Ⅲ 系统篇

第9章 *关系查询处理和优化 实验3

第10章 数据库恢复技术 第11章 并发控制 实验4







内容提要

- ✓ 范式
- ✓ 2NF
- ✓ 3NF
- ✓ BCNF
- ✓ 多值依赖
- ✓ 4NF
- ✓ 规范化总结





数据库维护中出现的一些问题

[例6.1] 试建立一个描述学校教务的数据库。涉及的对象包括:

- 学生的学号(Sno),所在系(Sdept),系主任姓名(Mname),课程号(Cno),成绩(Grade)
 - 关系模式R<U, F>
 - U ={Sno, Sdept, Mname, Cno, Grade}
 - 一个系有若干学生, 但一个学生只属于一个系;
 - 一个系只有一名(正职)负责人;
 - 一个学生可以选修多门课程,每门课程有若干学生选修;
 - 每个学生学习每一门课程有一个成绩。



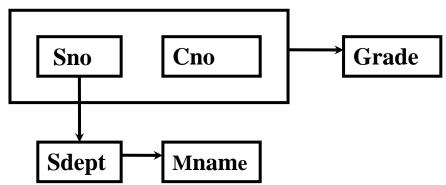


关系模式

- U ={Sno, Sdept, Mname, Cno, Grade}
 - 可得到属性组U上的一组<mark>函数依赖F</mark>:

F={Sno→ Sdept, Sdept→ Mname, (Sno, Cno)→ Grade}







什么是函数依赖?

关系模式R<U,F>中存在哪些问题?





数据依赖

- * 什么是数据依赖
 - 是一个关系内部属性与属性之间的一种约束关系
 - 通过属性间值的相等与否体现出来的数据间相互联系
 - 是数据属性内在的性质,是语义的体现
- 数据依赖的主要类型
 - 函数依赖(Functional Dependency,简记为FD)
 - 多值依赖 (Multi-Valued Dependency, 简记为MVD)





函数依赖

- 描述一个学生关系,可以有学号、姓名、系名等属性。
 - 一个学号只对应一个学生,一个学生只在一个系中学习
 - "学号"值确定后,学生的姓名及所在系的值就被唯一确定。

- Sname=f(Sno), Sdept=f(Sno)
 - Sno函数决定Sname
 - Sno函数决定Sdept
 - 记作Sno→Sname, Sno→Sdept







Grade

关系模式存在的问题

- 1) 数据冗余 重复存储
 - 每一个系主任的姓名重复出现,重复次数与该系所有学生的所有课程成绩出现次数相同。
- 2) 更新异常(Update Anomalies) 冗余数据维护不一致
 - 某系更换系主任后,必须修改与该系学生有关的每一个元组。

Sno

Sdept

Cno

Mname

- 3) 插入异常(Insertion Anomalies)
 - 如果一个系刚成立,尚无学生,则无法记录系主任的信息。
- 4) 删除异常(Deletion Anomalies)
 - 如果某个系的学生全部毕业删除了,则系主任的信息也丢失了。



Grade

Cno

Mname

Sno

Sdept

关系模式存在的问题

- * 结论
 - 关系模式R不是一个好的模式。

* 原因

- 由存在于模式中的某些数据依赖引起的。
- 表现形式: **数据冗余** 重复存储

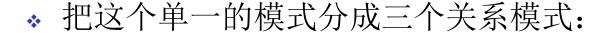
*解决方法

用规范化理论改造关系模式来消除其中不合适的数据依赖。



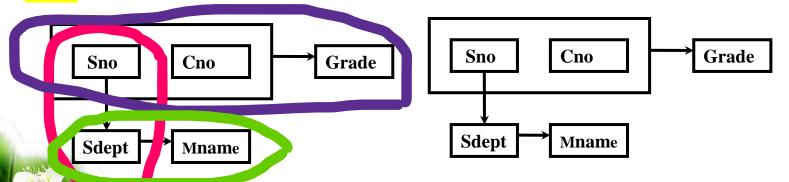


关系模式的改造





- S(Sno,Sdept, Sno → Sdept);
- SC(Sno,Cno,Grade, (Sno,Cno) → Grade);
- DEPT(Sdept,Mname, Sdept → Mname);
- 这三个模式都不会发生插入异常、删除异常的问题,数据的 冗余也得到了控制。



华中科技大学网络空间安全学院



范式

- ❖ 范式是符合某一种级别的关系模式的集合。
- ❖ 关系数据库中的关系必须满足一定的要求。满足不同程度 要求的为不同范式。
- * 范式的种类:

第一范式(1NF)

第二范式(2NF)

第三范式(3NF)

BC范式(BCNF)

第四范式(4NF)

第五范式(5NF)







全民

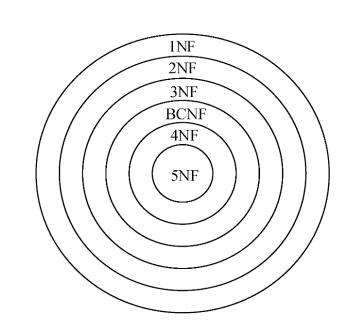
阅读

范式

* 各种范式之间存在联系:

 $1NF \supset 2NF \supset 3NF \supset BCNF \supset 4NF \supset 5NF$

- 某一关系模式R为第n范式,可简记为R∈nNF。
- ❖ 一个低一级范式的关系模式,通过模式分解(schema decomposition)可以转换为若干个高一级范式的关系模式的集合。
- ❖ 过程就叫规范化(normalization)







范式—INF

- * 各种范式之间存在联系: $1NF \supset 2NF \supset 3NF \supset BCNF \supset 4NF \supset 5NF$
 - 关系模式作为二维表,要符合一个最基本的条件:
 - ▶ 每个分量必须是<mark>不可分开的数据项</mark>。满足了这个条件的关系 模式就属于第一范式(**1NF**)。

| SUPERVISO | SPECIALITY | POSTGRADUATE | |
|-----------|------------|--------------|-----|
| R | | PG1 | PG2 |
| 张清玫 | 计算机专业 | 李勇 | 刘晨 |
| 刘逸 | 信息专业 | 王敏 | |





范式一2NF

- ❖ 定义6.6 若关系模式R∈1NF,并且每一个非主属性都<mark>完</mark> 全函数依赖于任何一个<mark>候选码</mark>,则R∈2NF
- * 定义6.4 设K为R<U,F>中的属性或属性组合。 若 $K \rightarrow U$,则K称为R的一个候选码(Candidate Key)。
 - 1.函数依赖
 - 2.平凡函数依赖与非平凡函数依赖
 - 3.完全函数依赖与部分函数依赖
 - 4.传递函数依赖

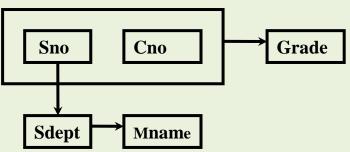




1. 函数依赖

- * 定义**6.1** 设R(U)是一个属性集U上的关系模式,X和Y 是U的子集。
- * 若对于R(U)的任意一个可能的关系r, r中不可能存在两个元组在X上的属性值相等,而在Y上的属性值不等,则称"X函数确定Y"或"Y函数依赖于X",记作 $X \rightarrow Y$ 。







函数依赖

❖ [例] Student(Sno, Sname, Ssex, Sage, Sdept),

假设<mark>不允许重名</mark>,则有:

Sno \rightarrow Ssex, Sno \rightarrow Sage

Sno \rightarrow Sdept, Sno $\leftarrow \rightarrow$ Sname

Sname \rightarrow Ssex, Sname \rightarrow Sage

Sname \rightarrow Sdept

但Ssex \Sage, Ssex \Sdept

 $若X\rightarrow Y$,并且 $Y\rightarrow X$,则记为 $X\leftarrow\rightarrow Y$ 。 若Y不函数依赖于X,则记为 $X\leftarrow Y$ 。





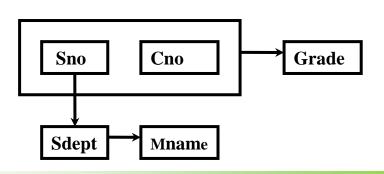
2. 平凡函数依赖与非平凡函数依赖

- * $X \rightarrow Y$, 但 $Y \nsubseteq X$ 则称 $X \rightarrow Y$ 是非平凡的函数依赖。
- * $X \rightarrow Y$,但 $Y \subseteq X$ 则称 $X \rightarrow Y$ 是平凡的函数依赖。

对于任一关系模式,平凡函数依赖都是必然成立的,它不反映新的语义。 若不特别声明, 我们总是讨论<mark>非平凡函数依赖</mark>。









3. 完全函数依赖与部分函数依赖

- * 定义6.2 在R(U)中,如果 $X \rightarrow Y$,并且对于X的任何一个真子集X',都有 $X' \nrightarrow Y$,则称Y对X完全函数依赖,记作 $X \stackrel{F}{\rightarrow} Y$
- * 若 $X \rightarrow Y$,但Y不完全函数依赖于X,则称Y对X部分函数依赖,记作 $X \stackrel{P}{\rightarrow} Y$

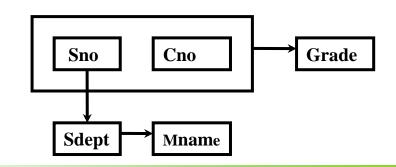
[例] 在关系SC(Sno, Cno, Grade)中,有:

曲于: Sno →Grade, Cno → Grade,

因此: (Sno, Cno) $\xrightarrow{\mathbf{F}}$ Grade

 $(Sno, Cno) \xrightarrow{P} Sno$

(Sno, Cno) P→Cno



华中科技大学网络空间安全学院



4. 传递函数依赖

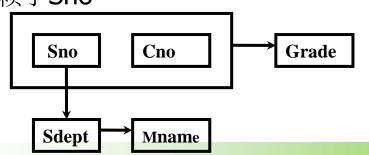
- ❖ 定义6.3 在R(U)中,如果 $X \rightarrow Y(Y \not\subseteq X)$, $Y \nrightarrow X$, $Y \rightarrow Z$, $Z \not\subseteq Y$,则称Z 对X 传递函数依赖(transitive functional dependency)。记为: X ^隻 Z。
 - 注: 如果 $Y \rightarrow X$, 即 $X \leftarrow \rightarrow Y$,则Z直接依赖于X,而不是传递函数依赖。

[例] 在关系Std(Sno, Sdept, Mname)中,有:

 $Sno \to Sdept, Sdept \to Mname,$

Mname传递函数依赖于Sno



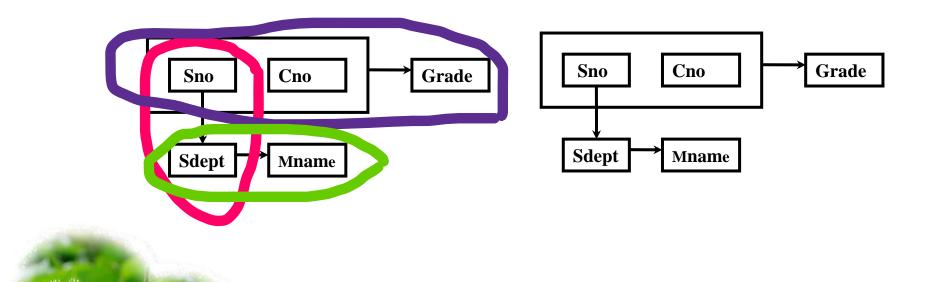






范式一2NF

❖ 定义6.6 若关系模式R∈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∈2NF





范式一2NF

❖ 定义6.6 若关系模式R∈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∈2NF

[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。

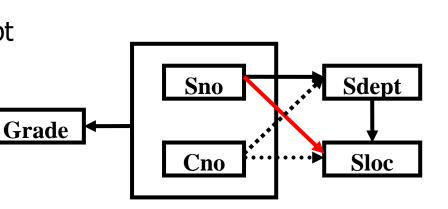
S-L-C的码为(Sno,Cno)。函数依赖有:

 $(Sno,Cno) \xrightarrow{\mathbf{F}} Grade$

Sno→Sdept, (Sno,Cno)→Sdept

Sno \rightarrow Sloc, (Sno,Cno) $\xrightarrow{\mathbf{P}}$ Sloc

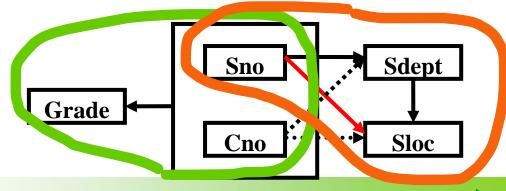
Sdept→Sloc





范式—2NF

- * 出现这种问题的原因
 - 例子中有两类非主属性:
 - 一类如Grade,它对码完全函数依赖
 - 另一类如Sdept、Sloc,它们对码不是完全函数依赖
 - •增删改会出现哪些异常?
- * 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)





华中科技大学网络空间安全学院





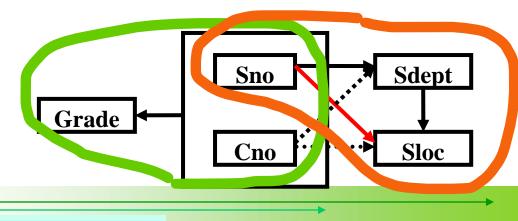
范式—2NF

- * 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)



●还存在什么问题吗?哪些增删改会出现异常?

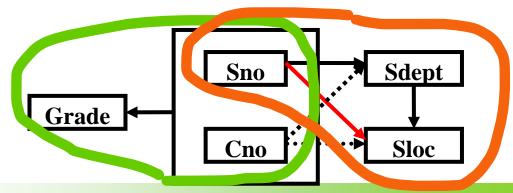






范式一3NF

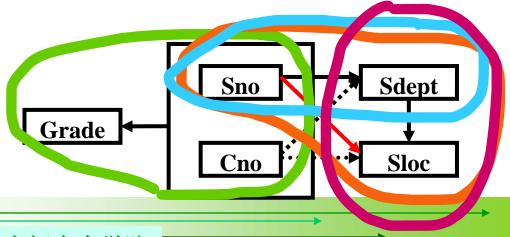
- ② 定义6.7 设关系模式 $R < U, F > \in 1$ NF,若R中不存在这样的码 X 属性组Y及非主属性Z ($Z \supseteq Y$),使得 $X \rightarrow Y$, $Y \rightarrow Z$ 成 立, $Y \rightarrow X$ 不成立,则称 $R < U, F > \in 3$ NF。
 - SC没有传递依赖,因此SC ∈ 3NF
 - S-L中Sno →Sdept(Sdept → Sno), Sdept→Sloc, 可得Sno → Sloc。
 - 解决的办法是将S-L分解成
 - S-D(Sno,Sdept)∈ 3NF
 - D-L(Sdept,Sloc) \in 3NF





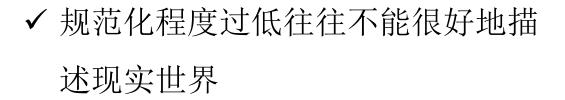
范式一3NF

- ❖ 定义6.7 设关系模式 $R < U, F > \in 1$ NF,若R中不存在这样的码 X 属性组Y及非主属性Z ($Z \supseteq Y$),使得 $X \rightarrow Y$, $Y \rightarrow Z$ 成立, $Y \nrightarrow X$ 不成立,则称 $R < U, F > \in 3$ NF。
 - SC没有传递依赖,因此SC ∈ 3NF
 - S-L中Sno →Sdept(Sdept → Sno), Sdept→Sloc, 可得Sno ^{传递} Sloc。
 - 解决的办法是将S-L分解成
 - S-D(Sno,Sdept)∈ 3NF
 - D-L(Sdept,Sloc)∈ 3NF





小结





- ✓ 可能存在插入异常、删除异常、修 改复杂等问题
- ✔ 解决方法就是
 - > 减少数据保存的冗余
 - ▶ 进行规范化,转换成高级范式。





