

数据库系统原理

第6章 关系数据理论2



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







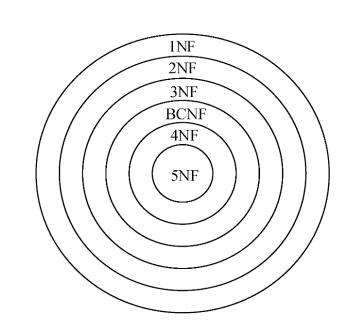
全民

阅读

范式



- 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)

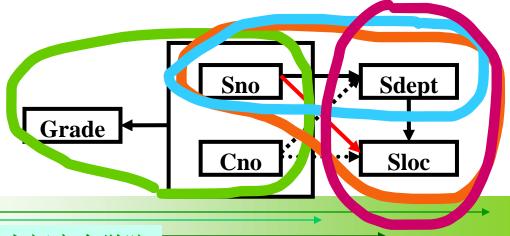






范式一3NF

- * 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($Z \not\supseteq Y$),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





范式一BCNF

* 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 $R < U, F > \in 1$ NF,若 $X \rightarrow Y$ 且 $Y \not\subseteq X$ 时X必含有码,则 $R < U, F > \in BCNF$ 。

◆ 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∈BCNF。







范式一BCNF

[例6.5]考察关系模式C(Cno,Cname,Pcno)

- 它只有<mark>一个码Cno</mark>,没有任何属性对Cno部分依赖或传递依赖,所以C∈3NF。
- 同时C中Cno是唯一的决定因素,所以C∈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∈3NF。
 - 同时S中除Sno, Sname外没有其他决定因素, 所以S也属于BCNF



码

- ② 定义6.4 设K为R<U,F>中的属性或属性组合。若K $\stackrel{F}{\rightarrow}$ U,则K称为R的一个候选码(Candidate Key)。
 - 如果U部分函数依赖于K,即 $K \xrightarrow{P} U$,则K称为超码(Surpkey)。候 选码是最小的超码。(K的任意一个真子集都不是候选码)
 - R有多个候选码,则可以选定一个做为主码(Primary key)。
 - ❖ 包含在任何一个候选码中的属性 , 称为主属性 (Prime attribute)
 - ❖ 全码:整个属性组是码,称为全码(All-key)

[例6.3] R(P,W,A) P: 演奏者 W: 作品 A: 听众

一个演奏者可以演奏多个作品,某一作品可被多个演奏者演奏 听众可以欣赏不同演奏者的不同作品

码为(P,W,A), 即All-Key

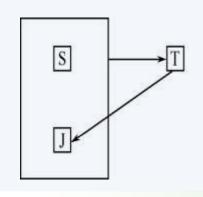


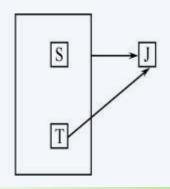


范式—BCNF

- ❖ [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)→T; (S,T)→J; T→J
 因为没有任何非主属性对码传递依赖或部分依赖, STJ ∈ 3NF。
 因为T是决定因素,而T不包含码,所以STJ ∉ BCNF关系。









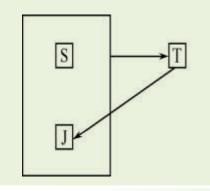


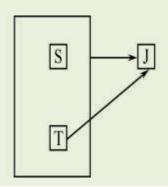
范式一BCNF

- ❖ 对于不是BCNF的关系模式,仍然存在不合适的地方。
- ❖ 非BCNF的关系模式也可以通过分解成为BCNF。例如STJ可分解为ST(S,T)与TJ(T,J),它们都是BCNF。
- * 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。
- ◆ 因为T是决定因素,而T不包含码,所以STJ

 BCNF关系。











范式一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。





范式

例[6.9]设学校中某一门课程由多个教师讲授,他们使用相同的一套参考书。每个教员可以讲授多门课程,每种参考书可以供多门课程使用

■ 用关系模式Teaching(C,T,B)来表示课程C、教师T和参考书B之间

的关系。「

课程 C	教员 T	参考书 B
物理	∫ 李勇	
数学	~ 李 勇 (张 平	物 堡 刁 越 果 数 学 分 析 微 分 方 程 高 等 代 数
计算数学	{ 张 平 } { 周 峰 }	大型 《数学分析》 《数学分析》 ···





范式

规范化的 二维表 **Teaching**

- ❖ Teaching具有
- 唯一候选码(C,T,B)
- ▶ 全码
- ➤ Teaching ∈ BCNF



课程C	教员 T	参考书 B
物理	李 勇	普通物理学
物理	李 勇	光学原理
物理	李勇	物理习题集
物理	王军	普通物理学
物理	王军	光学原理
物理	王军	物理习题集
数学	李 勇	普通物理学
数学	李 勇	光学原理
数学	李勇	物理习题集
数学	张平	普通物理学
数学	张平	光学原理
数学	张平	物理习题集
•••		

华中科技大学网络空间安全学院



❖ 定义6.9 设*R(U)*是属性集*U*上的一个关系模式。

新概念 NEW CONCEPT ENGLISH

X,Y,Z是**U**的子集,并且**Z=U-X-Y**。

关系模式R(U)中多值依赖 $X \rightarrow Y$ 成立,当且仅当对R(U)的任一关系r,给定的一对(x,z)值,有一组Y的值,这组值仅仅决定于x值而与z值无关。

※ 例 Teaching (C, T, B)
 对于C的每一个值, T有
 一组值与之对应, 而不论B
 取何值。因此T多值依赖
 于C, 即C→→T。

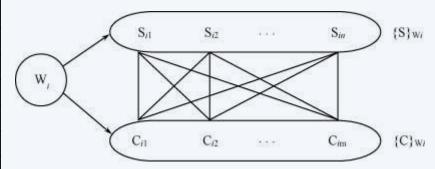
课程 C	教员 T	参考书 B
物理	李 勇	普通物理学
物理	李 勇	光学原理
物理	李 勇	物理习题集
物理	王军	普通物理学
物理	王军	光学原理
•••		

华中科技大学网络空间安全学院



[例6.10]关系模式WSC(W,S,C)中,W表示仓库,S表示保管员,C表示商品。假设每个仓库有若干个保管员,有若干种商品。每个保管员保管所在仓库的所有商品,每种商品被所有保管员保管。

W	S	C
W1	S1	C1
W1	S1	C2
W1	S1	C3
W1	S2	C1
W1	S2	C2
W1	S2	C3
W2	S3	C4
W2	S3	C5
W2	S4	C4
W2	S4	C5



❖ 按照语义对于W的每一个值 W_i , S有一个完整的集合与之对应而 不问C取何值。所以 $W\to\to S$ 。



◆ 由于C与S的完全对称性,必然有W→→C成立。



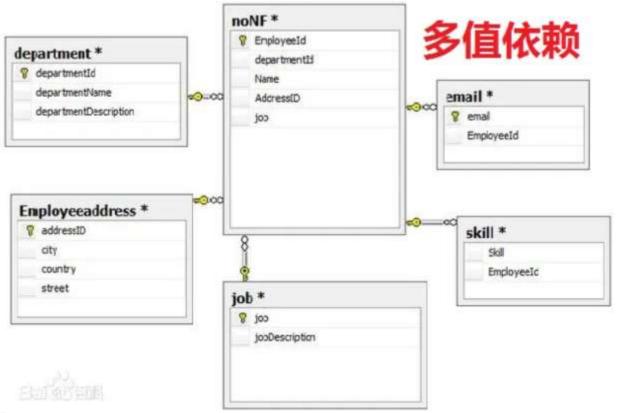




- (1)多值依赖的有效性与属性集的范围有关
 - 若 $X \rightarrow Y$ 在U上成立,则在W($XY \subseteq W \subseteq U$)上一定成立,反之则不然,即 $X \rightarrow Y$ 在W($W \subset U$)上成立,在U上并不一定成立。
 - 原因:多值依赖的定义中不仅涉及属性组X和Y,而且涉及U中其余属性Z。
- (2) 若函数依赖 $X \rightarrow Y$ 在R(U)上成立,则对于任何 $Y \subset Y$ 均有 $X \rightarrow Y'$ 成立。多值依赖 $X \rightarrow \rightarrow Y$ 若在R(U)上成立,不能断言对于任何 $Y' \subset Y$ 有 $X \rightarrow \rightarrow Y'$ 成立。















范式—4NF

定义6.10 关系模式 $R < U, F > \in 1$ NF,如果对于R的每个非平凡多值依赖 $X \rightarrow \rightarrow Y$ ($Y \nsubseteq X$),X都含有码,则 $R < U, F > \in 4$ NF。

- * 4NF就是限制关系模式的属性之间不允许有非平凡且非 函数依赖的多值依赖。4NF所允许的非平凡多值依赖实际上是 函数依赖。
- ❖ 如果一个关系模式是4NF, 则必为BCNF。

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 $R < U, F > \in BCNF$ 。





- ◆ 平凡多值依赖和非平凡的多值依赖
 - $X \to Y$,而 $Z = \Phi$,即Z为空,则称 $X \to Y$ 为平凡的多值依赖。
 - 否则称X→→Y为非平凡的多值依赖。
- ◆ 函数依赖是多值依赖的特殊情况。即若 $X \rightarrow Y$,则 $X \rightarrow Y$ 。

在[例**6.10**]的WSC中,W $\rightarrow\rightarrow$ S, W $\rightarrow\rightarrow$ C,他们都是非平凡多值依赖。

- ▶ 可以把WSC分解成WS(W,S),WC(W,C), WS∈4NF, WC∈4NF。





范式—5NF

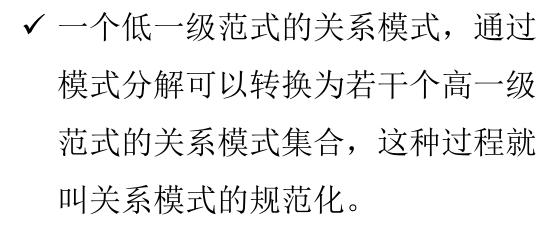
- 如果关系模式R中的每一个连接依赖均由R的候选码所隐含,则称此关系模式符合第五范式。
 - ▶ 在4NF的基础上,消除不是由候选码所蕴含的连接依赖。
 - 必须满足第四范式;表必须可以分解为较小的表,除非那些表在 逻辑上拥有与原始表相同的主键。

关系模式 $R < U, F > \in 1$ NF,如果对于R的每个非平凡多值依赖 $X \rightarrow \rightarrow Y (Y \not\subseteq X)$,X都含有码,则 $R < U, F > \in 4$ NF。











✓ 关系数据库的规范化理论是数据库 逻辑设计的工具。





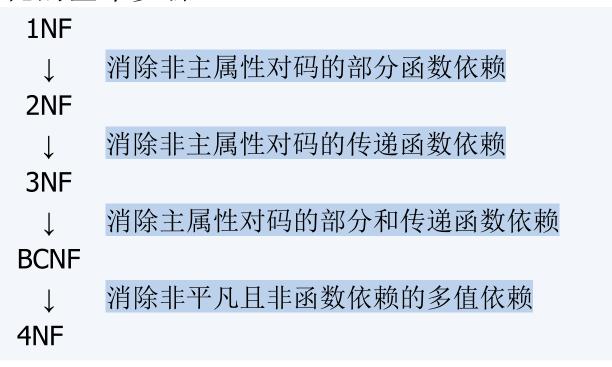
- ✓ 规范化的基本思想——概念的单一化
 - ➤ 是逐步消除数据依赖中不合适的部分, 使模式中的各关系模式达到某种程度的 "分离"。
 - ➤ 让一个关系描述<mark>一个</mark>概念、一个实体或 者实体间的一种联系。
 - ▶ 若多于一个概念就把它"分离"出去。





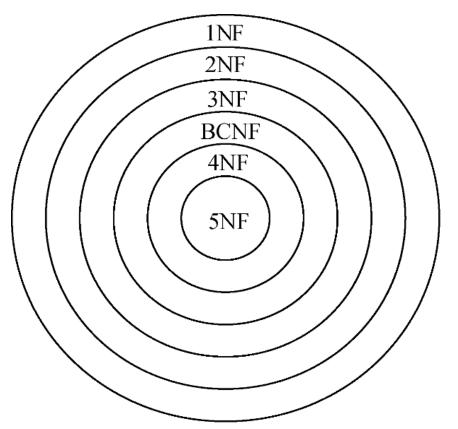


关系模式规范化的基本步骤















- ✓ 若要求分解具有无损连接性,那么模式分解一定能够达 到4NF。
- ✓ 若要求分解保持函数依赖,那么模式分解一定能够达到 3NF,但不一定能够达到BCNF。
- ✓ 若分解既具有无损连接性,又保持函数依赖,则模式分解一定能够达到3NF,但不一定能够达到BCNF。
- 并不是规范化程度越高,模式就越好。<mark>必须结合应用环</mark>境和现实世界的具体情况合理地选择数据库模式

