

第6-7章 关系数据理论&数据库设计

计科2201zzy友情分享😁😊

关系数据理论

- 数据依赖：一个关系内部属性与属性之间的一种约束关系。这种约束关系是通过属性间值的相等与否体现出来的数据间相关联系。
 - 函数依赖
 - X函数确定Y或Y函数依赖于X，记作 $X \rightarrow Y$
 - 语义范畴的概念，只能根据语义来确定一个函数依赖。
 - 若 $X \rightarrow Y$ ，但 $Y \not\subseteq X$ ，则称 $X \rightarrow Y$ 是非平凡的函数依赖。
 - 如： $Sno \rightarrow Sdept$
 - 反之若 $Y \subseteq X$ ，则称 $X \rightarrow Y$ 是平凡的函数依赖。如： $Sno \rightarrow Sname$
 - 平凡的函数依赖中，右侧的属性完全依赖于左侧的属性。换句话说，右侧属性的值对于左侧属性的任何特定值都是唯一确定的。
 - 如果 $X \rightarrow Y$ ，且对于 X 的任何一个真子集 X' ，都有 $X' \not\rightarrow Y$ ，则称 Y 对 X 完全函数依赖。记作 $X \xrightarrow{F} Y$
 - 否则是部分函数依赖，记作 $X \xrightarrow{P} Y$ 。
 - 如 $(Sno, Cno) \xrightarrow{F} Grade$ 是完全函数依赖； $(Sno, Cno) \xrightarrow{P} Sdept$ 是部分函数依赖。
 - 多值依赖
- 关系模式的设计：可能的问题
 - e.g. 把系主任的名字存在Student表里
 - 数据冗余：重复存储很多次。
 - 更新异常：更新数据代价很大，否则会不一致。
 - 插入异常：如果不存在某条数据，与之相关但独立的数据无法插入。系里没有学生时存不了系主任
 - 删除异常：删除某些数据的时候，其它与之相关但独立的数据也丢失了。这个系学生毕业了系主任也没了

规范化理论

- 关系数据库中的关系需要满足一定要求。满足不同程度要求的为不同范式。
- 第一范式(1NF)：每一个分量必须是不可分的数据项。
- 第二范式(2NF)
 - 首先必须满足1NF
 - 所有非主键字段必须完全函数依赖于主键。换句话说，表中的每个非主键字段必须完全取决于整个主键，而不仅仅取决于主键的一部分。

- 第三范式(3NF):
 - 首先必须满足2NF,
 - 消除非主键字段之间的传递依赖。换句话说, 如果一个非主键字段依赖于另一个非主键字段, 那么这个依赖关系应该通过创建一个新表来解决, 将依赖的字段抽取出来作为新表的主键。
- 一个低一级范式的关系模式通过模式分解可以转换为若干个高一级范式的关系模式的集合, 这个过程称为**规范化**。

数据库设计

- 数据库设计是指对于一个给定的应用环境, 构造(设计)优化的数据库逻辑模式和物理结构, 并据此建立数据库及其应用系统, 使之能够有效的存储和管理数据, 满足各种用户的应用需求, 包括信息管理要求和数据操作要求。
- 数据库设计的特点: “三分技术, 七分管理, 十二分基础数据”
- 数据库设计基本步骤
 - 需求分析
 - 数据字典: 需求分析-数据收集与分析后所获得的主要成果
 - 包括: 数据项, 数据结构, 数据流, 数据存储和处理过程
 - 数据项是不可再分的数据单位。
 - 一个数据结构可以由若干个数据项构成, 也可以由若干数据结构组成, 反映数据之间的组合关系
 - 概念结构设计
 - E-R模型: 描述现实世界的概念模型, 包括实体、属性、实体之间的联系
 - 三类联系: $1:1$, $1:n$, $m:n$
 - E-R图: 提供了表示实体型、属性和联系的方法。
 - 实体型用矩形表示
 - 属性用椭圆形表示
 - 联系用菱形表示
 - E-R图的集成
 - 合并。解决各部分E-R图之间的冲突将分E-R图合并起来生成初步的E-R图
 - 修改并重构。消除不必要的冗余, 生成基本的E-R图。
 - 逻辑结构设计
 - 将E-R图转换为关系数据模型
 - 一个实体型对应一个关系模式, 实体型间联系转换为关系模式, 详见教材 P232
 - 物理结构设计
 - 确定数据库的物理结构
 - 确定数据库的存储结构
 - 确定数据的存放位置
 - 确定系统配置
 - 对物理结构进行评价

- 数据库实施
- 数据库运行和维护
 - 数据库的转储与恢复
 - 数据库的安全性、完整性控制
 - 数据库性能的监督、分析和改造
 - 数据库的重组织和重构造