

数据库系统原理

复习

课程内容

第1章 绪论

历史、发展、基本特征、模型与模式

第2章 关系数据库

关系代数基础

第3章 关系数据库标准语言SQL

建表、改表、插入、查询、删除、更新

索引的概念

查询：连接、分组、嵌套

第4章 数据库安全性

基本安全机制、方法、原理

第5章 数据库完整性

完整性机制、关系模型的基本完整性

第6章 关系数据理论

关系规范化分析与理论

函数依赖、多值依赖、操作异常、范式（主属性/非主属性）

函数依赖的公理系统

算法：属性闭包、最小依赖集、关系分解
分解的正确性

第7章 数据库设计

主要阶段及其基本手段、概念、逻辑、物理

ER图、关系模型

第8章 数据库编程

第8章 关系数据库引擎基础

存储的内部数据结构、进程、缓冲区、文件、页、元组、执行引擎

第9章 关系查询处理和查询优化

问题的分析 (I/O开销分析)

代数和物理优化

代数优化的启发式规则：一种树结构的变形

物理优化的启发式规则：代数操作实现算法的选择

第10章 数据库恢复技术

故障类型，日志的恢复技术，检查点，具体恢复的案例

第11章 并发控制

并发错误、锁的使用、可串行化（结果的正确性判断依据）、冲突可串行化（一种可行的判断依据，针对一个调度序列展开验证）、3级封锁协议、两段锁协议（2PL）、死锁、多粒度锁调度序列的表达与分析（表格、操作序列字符串）

考试题型

没有单独的选择、填空题，主要题目涉及：分析、计算、推理、综合应用、设计等

考试时间：**150**分钟 考试形式：闭卷

- 原则上各章都有，应熟练掌握关系代数、SQL语言、ER图等基本应用技能和恢复、并发的策略与动作。
- 各章可结合实际案例的方法、原理都是考察的范围。
- 注意题量和答卷的书写。

第1章 绪论

产生与发展、基本特征、和数据管理有关的两种基本方法

- 数据库系统的产生与发展
人工方法、文件系统方法、数据库系统方法在应用特征、存储设备、存储需求、管理方式等方面的区别。

	人工方法	文件系统	数据库系统
应用背景	科学计算	科学计算、管理	大规模管理
硬件背景	无直接存取存储设备	磁盘、磁鼓	大容量磁盘
软件背景	无操作系统	有文件系统	有数据库管理系统
处理方式	批处理	批处理、联机实时处理	联机实时处理、分布处理、批处理
数据管理者	用户（程序员）	文件系统	数据库管理系统
数据面向的对象	某一应用程序	某一应用	现实世界
数据的共享程度	无共享，冗余极大	共享性差，冗余度大	共享性高，冗余小
数据独立性	不独立，完全依赖于程序	独立性差	物理独立性和逻辑独立性
数据的结构化	无结构	记录内有结构，整体无结构	整体结构化、使用数据模型描述
数据控制能力	应用程序控制	应用程序控制	Dbms 控制，提供安全、完整、并发、恢复能力。

● 数据模型

模型→数据模型

数据模型的功能、概念、基本要求、三个要素，现实世界到计算机世界的过程，对应的一些典型的数据模型。

概念模型的概念、表示方法。（会画概念模型）

实体、属性、码、实体型、实体集、联系（**3种**）。

层次模型、网状模型、关系模型的基本数据结构、操作原理、基本约束、优缺点。

- 数据库系统的**三层模式两级映像**结构。
作用、逻辑和物理两种独立性。
- 数据库系统的主从式、分布式、客户机/服务器方式的体系结构及其对应特征。
- 数据库系统（DBS）的组成
（DBS——数据库系统不等于DBMS——数据库管理系统）
数据库系统的主要功能、由哪些软件构成、应用程序的特征、有哪些用户。

功能:

- 1.数据定义、2.数据操纵、3.数据库运行管理（安全性、完整性检查，DD、索引维护、并发控制）、4.数据组织、存储和管理、5.数据库的建立（初始数据输入，数据转换）、6.维护（转储与恢复、重组、重构、性能监视与分析）、7.数据通讯接口。

包含的程序:

- 1.数据定义语言及其翻译处理程序、数据操纵语言及其编译（或解释）程序
- 2.数据库运行控制程序
初启程序、I/O，存取路径管理、缓冲区管理、安全控制、完整性控制、并发控制、事务管理、日志管理。
- 3.实用程序
初装、转储、恢复、监测、转换、重组、重构、通讯。

- 数据库系统的工作过程
- 数据库系统的主要特点：
 - 1) 数据结构化（整体结构化）
 - 2) 数据的共享性高、冗余度低、易扩充
 - 3) 数据独立性高
 - 4) 数据由**DBMS**统一管理和控制（安全、完整、并发、恢复）
- 数据库系统和手工、文件方法的区别
- 应用程序特征：

主语言+**SQL**语句（或者经过主语言包装的数据操纵控件）

DBS用户包括：

1、DBA (Database Adminitor)

DBMS、DB其它软件管理与维护

（安全授权、监测和改进性能）

2、系统分析员

分析用户需求，确定数据库事务

3、应用程序员

应用软件编码、调试和维护

4、终端用户

使用DB

第1章作为绪论，初学时只有概念层面的印象，学完所有章节后，应该能够结合各章原理、方法做到全面透彻的领悟——融会贯通。

第2章 关系数据库

- 关系数据模型

- 关系数据结构及形式化定义

- 域、笛卡儿积、元组、分量、基数、关系、关系的目、属性、候选码、主码、主属性、非码属性、全码、关系模式

- 。 。 。

- 关系数据操作

- 关系完整性：关系的两个不变性——实体完整性、参照完整性——应该由系统自动支持，用户定义完整性——体现具体应用领域中的语义约束。

- 关系代数

8种运算符（并、交、差、笛卡尔积、投影、选择、连接、除），其中5种基本运算符（并、差、笛卡尔积、投影、选择）。

理解关系代数，掌握关系代数表达式。

- *关系演算

元组关系演算、域关系演算。

- *关系代数、元组关系演算、域关系演算的等价性。

- Sql语言和上述语言的等价性。

- 关系演算的具体语言不考

第3章 关系数据库标准语言SQL、

- SQL的几个动词：
select、create、drop、alter、insert、
update、delete、grant、revoke。

- SQL语言的基本功能
定义、查询、更新、控制（安全、完整、一致性）
- SQL支持三级模式
- 数据的定义功能
表的创建、修改、删除。
索引的创建、删除。
- 数据的查询功能
一般格式（语法结构）
单表查询：
 - 简单查询**（别名）、
 - 条件查询：**比较查询、范围查询、集合查询、匹配查询（通配字符、转义字符）、空值查询、复合条件查询、
 - 排序查询、集函数查询、分组查询**

连接查询

等值连接、不等值连接、自身连接、外连接、复合条件连接。

嵌套查询

简单嵌套查询（例如使用**IN**谓词）、多层嵌套查询、同表嵌套查询、相关嵌套查询、比较嵌套查询、使用**ANY**谓词的嵌套查询、使用**ALL**谓词的嵌套查询、使用**EXISTS**谓词的嵌套查询、全称谓词的**EXISTS**表示方法。

- 数据更新功能

插入：插入单个元组、插入子查询结果。

修改：修改语句的条件、带子查询的修改。一致性修改

。

删除：删除语句的条件、带子查询的删除。

- 视图功能

- 视图的概念

- 视图的定义语句（视图列名定义的**3**个要求：聚集函数或表达式、多表连接选出公共属性、需要改名）

- 单表视图、多表视图、基于视图的视图、表达式视图、集函数视图。。。

- 视图结构的删除

- 视图数据的更新：插入、删除、修改。（with check option）

- 视图数据的查询

第4章 数据库安全性

- 安全性控制的概念与行业标准
- DBS安全控制的一般方法
 - 用户鉴别、访问控制（自主、强制）、视图、审计、加密。
- 自主存取控制：grant、revoke、role
- 强制存取控制：客体（敏感度标记）、主体（许可证级别）、读写规则
- 其他安全性保护：推理控制、隐蔽信道、数据隐私保护

第5章 数据库完整性

- 定义、功能（三个部分：定义、检查、处理）
- 完整性约束条件
完整性控制
实体完整性和参照完整性（能否置空、删除、修改、插入）。
- 触发器

第6章 关系数据理论

- 关系模式定义回顾
- 数据依赖
 函数依赖、多值依赖
- 关系中存在的操作异常：
 插入异常、删除异常、冗余太大导致的修改麻烦。
- 分析上述异常后得出的结论——规范化
- 函数依赖的定义

- 完全函数依赖、部分函数依赖。
- 传递函数依赖（不等同于函数依赖的传递性）
- 候选码、主码、主属性、非主属性（非码属性）、全码、外码。
- 范式的含义
- 1NF——2NF——3NF——BCNF——4NF、相关的结论。
 -
- 多值依赖的概念、性质、4NF。

- 模式的分解
 - 分解的定义
 - 分解的多样性
 - 分解的正确性——无损连接性、依赖保持性——“等价”的三个定义。
- Armstrong公理系统
 - 公理本身、基本推导规则
 - 逻辑蕴涵和Armstrong推导的等价
 - 最小依赖集、属性闭包
 - 分解的正确性判定（算法）、分解的算法

第7章 数据库设计

- 数据库设计的一般步骤：

需求分析

目标、策略、步骤、需求分析说明书。

概念设计

任务、工具、方法、步骤、局部ER图设计（确定实体、联系、属性、码）、视图集成（控制冗余、消除冲突、实施集成）、冲突的类型及其处理、集成的方法。

命名冲突——同名异义和同义异名、特征（属性）冲突——值、类型、范围、单位等等、结构冲突——联系方式、属性个数、属性次序、抽象级别。

逻辑设计

ER模型→一般的关系模型

联系的消化

一般关系模型→特定的关系系统模型

优化：关系的水平分解、垂直分解（规范化）。

抽取子模式。

物理设计

存储结构（存取路径——索引、聚簇、存放位置、系统配置）、存取方法。

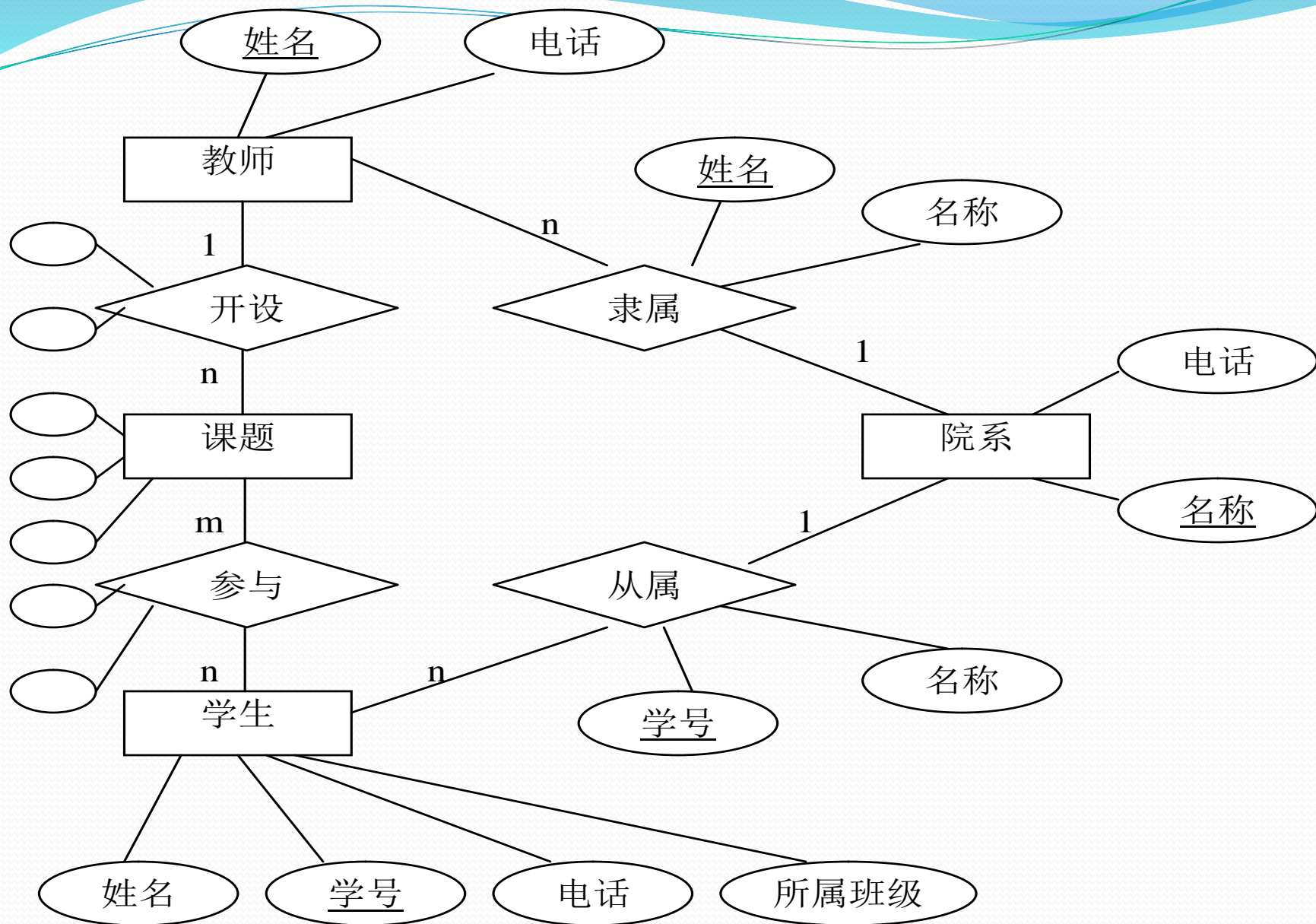
实现设计

DB定义、装载、编程、**DBS**试运行。

运行与维护

安全性和完整性控制、转储与恢复、性能监测、分析与改进、重组、重构。

课题一：	数据库实验	地点	南一楼		
	指导教师： 王宏		电话： 87541234		所在院系： 系办电话： 计算机 87544321
	学生姓名	学号	所属院系	班级	电话
	张三	01	计算机	0601	13900000001
	李四	02	计算机	0602	87546789
课题二：	电子器件制作	地点	西九楼		
	指导教师： 张悦		电话： 87542345		所在院系： 系办电话： 电子系 87545432
	学生姓名	学号	所属院系	班级	电话
	张三	01	计算机	0601	13900000001
	李四	02	计算机	0602	87546789
	赵六	05	电子	0603	87549876



第9章 关系查询处理和查询优化

- 查询处理的步骤
- 查询优化的实例分析
- 选择、连接操作算法的不同实现方式
- 代数优化（关系代数语法树的启发式优化）
- 物理优化的启发式规则

选择操作、连接操作、基于代价的优化（统计信息、代价估算模型）

第10章 数据库恢复技术

- 恢复的含义、目标。
 - **DBS**故障分类——事务、系统（硬件、软件、操作、停掉电）、介质。
 - 恢复的基本思想。
 - 备份的方式——静态、动态。
 - 备份的策略——海量、增量。
 - 日志的概念、日志记录、日志的管理。
 - 恢复的策略——事务故障、系统故障、介质故障。
 - 检查点对恢复的影响
 - 镜像技术。
- 从原理上掌握对某种故障及日志应该如何恢复保证数据库能恢复到某个正确状态，恢复的过程。

第11章 并发控制

- 事务的概念、构成方式、**ACID**特性。
 - 事务并发可能存在的冲突（写写、读写、写读）、幻象。
 - 并发调度、并发调度的正确性标准。
 - 并发控制的几种方法：锁、乐观、时标。
 - 锁的概念、封锁规则、锁的类型、申请的时机（事务、**sql**语句）、申请的方式（隐式、显示）、封锁的粒度。
 - 封锁协议的概念、目的、类别（三级封锁协议、两阶段锁协议、避免死锁协议）。
 - 死锁的含义、产生死锁的调度、死锁的处理机制。
 - 活锁的含义、解决办法。
 - 冲突可串行化（冲突操作、调度的变形规则、结果的判断、可串行化的充分而非必要条件）
-
- 理解调度的多种表现形式，理解锁的申请与释放对调度的影响，理解并发的正确性判断。