题型：

选择题

填空

简答

综合应用

综合设计

# 第一章 绪论

**一、 数据库系统概述**

1. **数据**是数据库中存储的基本对象

**记录**是计算机中表示和存储数据的一种格式或一种方法

1. **数据库**是长期存储在计算机内、有组织的、可共享的大量数据的集合。数据库中的数据按一定的**数据模型**组织、描述和存储，具有较小的冗余度、较高的数据独立性和易扩展性，并可为各种用户共享。
2. **数据库管理系统（DBMS）**是位于用户与操作系统之间的一层数据管理软件，和操作系统一样是计算机的**基础软件，**也是一个大型复杂的软件系统。
3. **数据库系统（DBS）**是由**数据库、数据库管理系统、应用程序和数据库管理员（DBA）**组成的存储、管理、处理和维护数据的系统。
4. 数据管理技术经历了**人工管理、文件系统、数据库系统**三个阶段（简单理解各自特点）

数据库系统的特点：

1. 数据结构化
2. 数据的共享性高、冗余度低且易扩展
3. 数据独立性高（**物理独立性和逻辑独立性**）
4. 数据由数据库管理系统统一管理和控制

数据的安全性保护

数据的完整性检查

并发控制

数据库恢复

**二、 数据模型**

1. **数据模型**也是一种模型，**它是对现实时间数据特征的抽象**。也就是数据模型是用来**描述数据、组织数据和对数据进行操作**的。
2. 两类数据模型：**概念模型；逻辑模型和物理模型**
3. 概念模型实际上是**现实世界**到**机器世界**的一个中间层次

信息世界的基本概念：实体（客观存在并可相互区别的事物称为实体）、属性（实体所具有的某一特性称为属性）、码（唯一标识实体的属性）、实体型（用实体名及其属性名集合来抽象和刻画同类实体，称为实体型）、联系（实体之间的联系有**一对一、一对多、多对多**）、域（属性的取值范围）

概念模型的一种表示方法：**实体-联系方法**（E-R图）

1. 数据模型通常由**数据结构**、**数据操作**和**数据的完整性约束条件**三部分组成
2. 逻辑数据模型：**层次模型、网状模型、关系模型**

层次模型是数据库系统中**最早**出现的数据模型（用树形结构来表示）

关系模型是**最重要**的一种数据模型。关系数据库系统采用**关系模型**作为数据的组织方式。

关系模型由一组关系组成，每个关系的数据结构是一张规范化的**二维表**

关系模型的一些术语：关系（一个关系对应通常说的一张表）、元组（表中的一行即为一个元组）、属性（表中的一列即为一个属性）、码（表中的某个属性组，可以唯一确定一个元组）、域（属性的取值范围）、分量（元组中的一个属性值）、关系模式（对关系的描述，关系模式要求关系必须是**规范化的**，这些规范条件中最基本的一条就是，关**系的每一个分量必须是一个不可分的数据项**）

关系模型要求关系必须是规范化的

关系的完整性约束条件包括三大类：**实体完整性、参照完整和用户定义的完整性**

**三、 数据库系统的结构、数据库系统的组成**

1. 数据库系统的**三级模式结构**是指数据库系统是由**外模式（用户模式）、模式（逻辑模式）和内模式（存储模式）**三级构成
2. **数据库系统（DBS）**是由**数据库、数据库管理系统、应用程序和数据库管理员（DBA）**组成。

# 第二章 关系数据库

**一、关系数据结构及形式化定义**

现实世界的实体以及实体间的各种联系均用单一的结构类型，即**关系**来表示。

【**候选码：**某一属性组的值能唯一地标识一个元组，而子集不能，则称该属性组为**候选码**；

**主码**：若一个关系有多个候选码，则选定其中一个为**主码；**

**外码：**一个关系中的属性在另一个关系中当主码；】

（候选码的诸属性称为**主属性**，不包含在任何候选码中的属性称为**非主属性或非码属性**。

在最极端的情况下，关系模式的所有属性是这个关系模式的候选码，称为**全码**；）

关系可以有三种类型：**基本关系、查询表和视图表**；

关系模式要求关系必须是**规范化**；

【**关系模式**：关系的描述称为**关系模式，**可以形式化的表示为R（U，D，DOM,F）；

**关系**：现实世界的实体以及实体间的各种联系均用单一的结构类型，即**关系**来表示。是若干组的集合，指数据库表

**关系数据库**：在一个给定的应用领域中，所有关系的集合构成一个关系数据库。关系数据库也有型和值之分。关系数据库的型也称为关系数据库模式，是对关系数据库的描述。关系数据库的值是这些关系模式在某一时刻对应关系的集合，通常就称为关系数据库。】

**二、关系操作**

关系模型中常用的关系操作包括**查询、插入、删除、修改（增删改查）**

关系的**查询**表达能力很强，是关系操作中**最主要**的部分。查询操作又可以分为**选择、投影、连接、除、并、差、交、笛卡儿积**等，其中选择、投影、并、差、笛卡儿积是5种基本操作。

关系操作的特点是**集合**操作方式，即**操作的对象和结果都是集合**。

早期的关系操作能力通常用代数方式或逻辑方式来表示，分别称为**关系代数和关系演算**。

还有一种介于关系代数和关系演算之间的**结构化查询语言（Structured Query Language）。SQL**是集**查询、数据定义语言、数据操纵语言和数据控制**语言于一体的关系数据语言。

关系数据语言分为三类（关系代数语言、关系演算语言、具有关系代数和关系演算双重特点的语言）

**三、关系的完整性**

关系模型中有三类完整性约束：**实体完整性、参照完整性和用户定义的完整性**。

实体完整性：主码不能为空；

参照完整性：外码为空（即每个属性的值均为空值）、或者只能取相应被参照关系中已经存在的主码值；

用户定义完整性：针对某一具体关系数据库的约束条件

**四、关系代数**

关系代数是一种抽象的查询语言，它用对**关系的运算**来表达查询。

集合运算：

1. 并
2. 差
3. 交
4. 笛卡儿积

专门的关系运算

1. 选择
2. 投影
3. 连接
4. 除

# 第三章 关系数据库标准语言（SQL）

**一、SQL概述**

1. 视图是从一个或几个基本表导出的表，是一个虚表，用户可以在视图上再定义视图

2.

|  |  |
| --- | --- |
| SQL功能 | 动词 |
| 数据查询 | SELECT |
| 数据定义 | CRATE DROP ALTER |
| 数据操纵 | INSERT UPDATE DELETE |
| 数据控制 | GRANT REVOKE |

**各种例题⭐**

# 第四章 数据库安全性

**一、数据库安全性概述**

1. 数据库的安全性是指保护数据库以防不合法使用所造成的数据泄露、更改或破坏

2. 自主存取控制：用户对于不同的数据库对象有着不同的存取权限，不同的用户对同一对象也有不同的权限，而且用户还可将其拥有的权限转授给其他用户

强制存取控制：每一个数据库对象被标以一定的密级，每一个用户也被授予某一个级别的许可证。

1. 自主存取控制方法：

Grant（授予）、revoke（收回）

例题⭐

数据库角色是被命名的一组于数据库操作相关的权限，**角色**是权限的集合

1. 强制存取控制方法：

强制存取控制是对数据本身进行密级标记，无论数据如何复制，标记与数据是一个不可分的整体，只有符合密级标记要求的用户才可以操纵数据，从而提供了更高级别的安全性。

即自主存取控制和强制存取控制共同构成数据库管理系统的安全机制

# 第五章 数据库完整性

1. 数据库的完整性是指数据的**正确性的相容性**

2. 定义实体完整性：primary key

定义参照完整性：foreign key （Sno）references student（sno）

参照完整性检查和违约处理：拒绝执行（No action）、级联操作（cascade）、设置为空值

用户定义的完整性：列值非空（Not null）、列值唯一（unique）、检查是否满足一个条件表达式（check）例题⭐

3. 完整性约束命名子句 constraint

域中的完整性限制 domain

断言 assertion

触发器 trigger

# 第六章 关系数据理论

1. 数据依赖：**函数依赖和多值依赖（**填空题），函数依赖是多值依赖的一种特殊情况。例6.1 P178

2. 非平凡函数的依赖

平凡函数的依赖

部分函数依赖

完全函数依赖

传递函数依赖

3. 规范化：前3NF判断

1NF：每一个分量不可分

2NF：每一个非主属性完全函数依赖于任何一个候选码

3NF：每一个非主属性既不传递依赖于码，也不部分依赖于码

# 第七章 数据库设计

1. 结构化分析方法是一种简单实用的方法，采用自顶向下、逐层分解的方式

2. **数据字典**通常包括**数据项、数据结构、数据流、数据存储和处理过程**几部分。其中数据项是数据的最小组成单位，若干个数据项可以组成一个数据结构。数据字典通过对数据项和数据结构的定义来描述数据流、数据存储的逻辑内容。

3. 数据项是不可再分的数据单位

4. 各子系统的E-R图之间的冲突主要有三类：属性冲突、命名冲突和结构冲突

# 剩余章节

1. 使用游标的SQL语句

说明游标（declare）、打开游标（open）、推进游标指针并取当前记录（fetch）、关闭游标（close）

1. 查询优化分为**代数优化**（逻辑优化）和**物理优化**（非代数优化）
2. **数据库恢复机制**和**并发控制机制**是数据库管理系统的重要组成部分
3. **事务**：所谓事务是用户定义的一个数据库操作序列，这些操作要么全做，要么全不做，是一个不可分割的工作单位

事务具有4个特性：**原子性、一致性、隔离性和持续性**

事务是恢复和并发控制的基本单位

5．故障的种类

事物内部的故障、系统故障、介质故障、计算机病毒

6. 恢复机制涉及两个关键问题：**如何建立冗余数据、以及如何利用这些冗余数据实施数据库恢复**。建立冗余数据最常用的技术是**数据转储和登记日志文件**。

7. 转储可以分为**静态转储和动态转储、海量转储和增量转储**

8. **日志文件**是用来记录事务对数据库的更新操作的文件，可以用来进行事务故障恢复和系统故障恢复，并协助后备副本进行介质故障恢复。

9. 数据库管理系统自动把更新后的数据复制过去，由数据库管理系统自动保证**镜像**数据与主数据库的一致性。

10. 事务是并发控制的基本单位

11. 并发操作带来的数据不一致性包括**丢失修改、不可重复读、读“脏”数据**

12．**封锁**是实现并发控制的一个非常重要的技术。基本的封锁有两种类型：**排他锁**和**共享锁**

13. 封锁的方法可能引起活锁和死锁等问题，避免活锁的简单方法是采用先来先服务的策略。封锁对象的大小称为封锁粒度。多粒度封锁中一个数据对象可能以两种方式封锁，显式封锁和隐式封锁。显式封锁是应事务的要求直接加到数据对象上的锁；隐式封锁是该数据对象没有被独立加锁，是由于其上级结点加锁而使该数据对象加了锁。

# 简答题

1. **（DBMS）数据库管理系统的主要功能：**
2. 数据定义功能

数据库管理系统提供数据定义语言，用户通过它可以方便地对数据库中的数据对象的组成与结构进行定义。

1. 数据组织、存储和管理

数据组织和存储的基本目标是提高存储空间利用率和方便存取，提供多种存取方法来提高存取效率

1. 数据操纵功能

数据库管理系统还提供数据操纵语言，用户可以使用它操纵数据，实现对数据库的基本操作，如查询、插入、删除和修改等

1. 数据库的事物管理和运行管理

数据库在建立、运用和维护时由数据库管理系统统一管理和控制，以保证事物的正确运行，保证数据的安全性、完整性、多用户对数据的并发使用及发生故障后的系统恢复

1. 数据库的建立和维护功能

数据库的建立和维护功能包括数据库初始数据的输入、转换功能，数据库的转储、恢复功能，数据库的重组织功能和性能监视、分析功能。

1. 其他功能

其他功能包括数据库管理系统与网络中其他软件系统的通信功能，一个数据库管理系统与另一个数据库管理系统或文件系统的数据转换功能，异构数据库之间的互访和互操作功能

1. **SQL的特点：**
2. 综合统一。SQL集数据定义语言、数据操纵语言、数据控制语言的功能于一体，语言风格统一，可以独立完成数据库生命周期中的全部活动。
3. 高度非过程化

只要提出“做什么”，而无须指明“怎么做”，因此无须了解存取路径。存取路径的选择以及SQL的操作过程由系统自动完成。

1. 面向集合的操作方式

SQL采用集合操作方式，不仅操作对象，查找结果可以是元组的集合，而且一次插入、删除、更新操作的对象也可以是元组的集合。

1. 以同一种语法结构提供多种使用方式

SQL既是独立的语言，又是嵌入式语言

1. 语言简洁，易学易用

非常接近英语口语

1. **视图的作用：**
2. 视图能够简化用户的操作

试图机制使用户可以将注意力集中在所关心地数据上，用户所做的只是对一个虚表的简单查询，而这个虚表是怎样得来的，用户无须了解。

1. 视图使用户能以多种角度看待同一数据

视图机制能使不同的用户以不同的方式看待同一数据

1. 视图对重构数据库提供了一定程度的逻辑独立性

尽管数据库的逻辑结构改变了，但应用程序不必修改，因为新建立的视图定义为用户原来的关系，使用户的外模式保持不变，用户的应用程序通过视图仍然能够查找数据。

1. 视图能够对机密数据提供安全保护

对不同的用户定义不同的视图，使机密数据不出现在不应看到的这些数据的用户视图上

1. 适当利用视图可以更清晰地表达查询
2. **数据库安全性控制**
3. 用户标识与鉴定

包括静态口令鉴别（最常用，密码）、动态口令鉴别（如短信验证码）、生物特征鉴别（指纹）、智能卡鉴别（比如校园卡的门禁）

1. 存取控制

存取控制机制主要包括定义用户权限和合法权限检查；

存储控制支持自主存取控制和强制存取控制。

1. 视图机制

可以为不同的用户定义不同的视图，把数据对象限制在一定的范围内

1. 审计

审计功能把用户对数据库的所有操作自动记录下来放入审计日志中。审计员可以利用审计日志监控数据库中的各种行为，重现导致数据库现有状况的一系列事件，找出非法存取数据的人

1. 加密

加密的基本思想是根据一定的算法将原始数据——明文变换为不可直接识别的格式——密文

1. 其他安全性保护

在自主存取控制和强制存取控制之外，还有推理控制以及数据库应用中隐蔽信道和数据隐私保护等技术。

1. **什么是数据库的完整性？**
2. 实体完整性。

主码不能为空，主码不能重复使用，用primary key定义，有两种说明，一种是定义为列级约束条件，另一种是定义为表级约束条件。

1. 参照完整性

即在create table中用foreign key短语定义哪些列是外码，用references 短语指明这些外码参照哪些表的主码

1. 用户定义完整性

在属性上的约束条件，有列值非空（not null），列值唯一（unique），检查列值是否满足一个条件表达式（check）

1. **数据库设计的基本步骤**
2. 需求分析阶段

分析用户要求，可以通过详细调查现实世界要处理的对象，明确用户的各种需求，在此基础上确定新系统的功能

1. 概念结构设计

即E-R图

1. 逻辑结构设计

**逻辑结构设计**的任务就是把概念结构设计阶段设计好的基本E-R图转换为与选用数据库管理系统产品所支持的数据模型相符合的逻辑结构

实体型转换为一个关系模式P232

1. 物理结构设计

确定数据库的物理结构

对物理结构进行评价

1. 数据库实施
2. 数据库运行和维护