

## 第一章 绪论

### 1.1 数据库系统概述

#### 1.1.1 数据库的 4 个基本概念

Data、DB、DBMS、DBS

##### 1、数据

描述事物的符号成为数据

数据的含义称为语义

记录是计算机表示和存储数据的一种格式或方法

##### 2、数据库

定义：数据库是长期存储在计算机内的、有组织的、可共享的大量数据的集合

特点：永久存储、有组织和可共享、较小冗余度、较高的数据独立性和易扩展性

##### 3、数据库管理系统

用途：可以科学的组织和存储数据，高效地获取和维护数据

主要功能：

数据定义：提供 DDL，定义数据对象

数据组织、存储和管理

数据操纵功能：提供 DML 语言，实现对数据库的增删改查

数据库的事务管理和运行管理：保证数据的安全性、完整性多用户对数据的并发使用

数据库的建立和维护功能

其他功能：DBMS 与其他软件系统的通信、DBMS 之间数据转换等

##### 4、数据库系统

数据库系统构成：数据库、数据库管理系统、应用程序、数据库管理员

数据库系统简称为数据库

#### 1.1.2 数据管理技术的产生和发展

数据管理定义：对数据进行分类、组织、编码、存储、检索和维护

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

##### 1、人工管理阶段

数据不独立，无结构，无共享，冗余度极大，应用程序自己控制数据

## 2、文件系统阶段

数据共享性差、冗余度大，记录内有结构整体无结构，独立性差，程序控制

### 1.1.3 数据库系统的特点

#### 1、数据结构化

数据库系统实现整体数据的结构化，是数据库系统与文件系统的本质区别

数据用数据模型描述，无需应用程序定义

#### 2、数据的共享性高，冗余度低且易扩充

数据面向整个系统

数据共享的好处：减少冗余节约空间、系统易于扩充、避免数据不相容与不一致

不一致是指在文件中数据重复存储，修改某一数据其其他副本数据没有被修改，起冲突

#### 3、数据独立性高

指物理独立性和逻辑独立性

物理独立性：应用程序中的数据与数据库中的数据存储是相互独立的

逻辑独立性：应用程序与数据库的逻辑结构是相互独立的

数据独立性由数据库管理系统的二级映像功能来保证

#### 4、数据由数据库管理系统统一管理和控制

(1) 数据的安全性保护：防止不合法使用导致数据泄露或被破坏

(2) 数据的完整性检查：保证数据正确性、有效性、相容性

(3) 并发控制：对并发操作加以协调

(4) 数据库恢复：将数据库从错误状态恢复到某一已知的正确状态

### 1.2 数据模型

数据模型是对现实世界数据特征的抽象，也就是对现实世界的模拟

数据模型是用来描述数据、组织数据和对数据进行操作的

数据模型是数据库系统的核心和基础

数据模型要求：真实地模仿现实世界、容易被人理解、便于在计算机上实现

#### 1.2.1 两类数据模型

分为两大类 概念模型：物理模型和逻辑模型

(1) 概念模型：也称信息模型，按用户观点对数据进行建模

## （2）逻辑模型和物理模型

逻辑模型：包括网状模型、层次模型、关系模型、面向对象数据模型、对象关系数据模型、半结构化数据模型。是按计算机观点对数据建模

物理模型：是对数据最底层的抽象，描述数据在系统内部或在磁盘磁带上的表示方式和存取方法

两步抽象法：先将现实抽象为概念模型，再将概念模型转换为数据库支持的数据模型。即现实世界->概念模型->逻辑模型->物理模型

### 1.2.2 概念模型

#### 1、信息世界中的基本概念

（1）实体 （2）属性 （3）码：唯一标识实体的属性集称为码

（4）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

（5）实体集：同一类型实体的集合

（6）联系：实体内部：指各属性

实体之间：不同实体集之间

有一对一、一对多、多对多等多种类型

#### 2、概念模型的一种表示方法：ER 图（ER 模型）

### 1.2.3 数据模型的组成要素

数据模型通常由数据结构、数据操作、数据的完整性约束条件组成

#### 1、数据结构

数据结构描述数据库的组成对象以及对象之间的联系

数据结构是对系统静态特性的描述

#### 2、数据操作

数据操作是指对数据库中各种对象（型）的实例（值）允许执行的操作的集合，包括操作及有关的操作规则

类型： 查询和更新（增删改）

数据结构是对系统动态特性的描述

#### 3、数据的完整性约束条件

数据的完整性约束条件是一组完整性规则

完整性规则是给定的数据模型中数据及其联系所具有的制约和依存规则，用以限

定符合数据模型的数据库状态以及状态的变化

#### 1.2.4 常用的数据模型

层次模型、网状模型、关系模型、面向对象数据模型、对象关系数据模型、半结构化数据模型

#### 1.2.5 层次模型

##### 1、特点：

结点双亲唯一；只有一对多；每个记录类型可以定义一个码字段；子女不能脱离双亲记录值而独立存在；任何记录值只能按其路径查看

##### 2、数据操纵：增删改查

##### 3、完整性约束条件

无双亲就不能插子女；删掉双亲子女也被删；更新时要更新所有相应记录，保证数据一致性

##### 4、优缺点

优点：数据结构简单清晰；查询效率高；有良好完整性支持

缺点：多对多表示不自然；插入删除限制多；查询子女必须通过双亲；层次命令趋于程序化

#### 1.2.6 网状模型

##### 1、特点：一个节点可有零个或多个双亲

##### 2、表示方法：与层次模型相同

3、与层次模型区别：结点问题；网状模型可以有复合联系；网状模型可以更直接地描述世界；层次模型是网状模型的特例

##### 4、数据操纵与完整性约束

记录码；双亲与子女记录是一对多；支持双亲记录和子女记录间的某些约束条件

##### 5、优缺点

优点：更直观；良好性能

缺点：结构复杂，越来越复杂；DDL/DML 语言复杂；用户必须了解系统结构细节

#### 1.2.7 关系模型

1、数据结构：用户来看，关系模式的逻辑结构是一张二维表；关系模型要求关系必须规范化；关系的每一个分量都不可分割

## 2、数据操纵与完整性约束

数据操纵：增删改查

完整性约束：实体完整性、参照完整性、用户定义的完整性

## 3、优缺点

优点：建立在严格的数学概念的基础上；概念单一；关系模型的存取路径对用户透明，数据更独立，更安全

缺点：查询效率低；需要优化查询，增加难度

### 1.3 数据库系统的结构

从内看，数据库采用三级模式结构

从外看，数据库系统的结构分为：单用户结构，主从式结构，分布式结构，客户-服务器结构，浏览器-应用服务器、数据库服务器多层结构等

#### 1.3.1 数据库系统模式的概念

“型”和“值”：型是某一类数据结构和属性的说明；值是型的一个具体复赋值  
模式是数据库中全体数据的逻辑结构和特征的描述，只有型没有值

#### 1.3.2 数据库系统的三级模式结构

内模式、模式、外模式

##### 1、内模式（称存储模式）

是数据物理结构和存储方式的描述

一个数据库只有一个内模式

##### 2、模式（称逻辑模式）

数据库中全体数据的逻辑结构和特征的描述，是所有用户的公共数据视图

一个数据库只有一个模式

模式与数据库的物理存储细节和硬件环境无关，与具体应用程序无关

##### 3、外模式（用户模式或子模式）

是数据库用户能够看见啊和使用的局部数据的逻辑结构和特征的描述

外模式通常是模式的子集

一个数据库系统可以有多个外模式，一个外模式可以对应多个应用系统

是保证数据库安全性的一个有力措施，用户只能看见和访问外模式中的数据

#### 1.3.3 数据库的二级映像功能与数据独立性

三级模式是对数据的三个抽象级别

三级模式有两层映像

#### 1、外模式/模式映像

该映像定义了外模式与模式之间的对应关系，定义一般包含在外模式描述中  
当模式改变时，该映像发生改变，可以使外模式薄啊吃不变，所以应用程序可以不变，这是逻辑独立性

#### 2、模式/内模式映像

模式/内模式映像是唯一的

定义了数据全局逻辑结构与存储结构之间的对应关系，通常包含在模式描述中  
当数据库存储结构变了，修改该映像，模式可以保持不变，应用程序也不变，这是数据的物理独立性

数据库模式即全局逻辑结构是数据库的中心与关键

设计数据库模式时应首先确定数据库的逻辑模式

数据库的二级映像保证了数据库的稳定性，除非需求发生变化，否则应用程序不需要修改

数据与程序之间的独立性，使得数据的定义和描述可以从应用程序中分离

### 1.4 数据库系统的组成 P31