

数据库系统概论（第四版）

1~3 数据库基本概念，4~8 三种重要数据库模型，9 数据库设计控制，10 数据库设计

目录

=====基础=====

一 绪论

数据模型的组成要素，最主要的数据模型，数据库系统三级模式结构，数据库系统的主要组成部分

二 关系数据库

关系数据库概念：关系模型、关系代数

三 数据库标准语言

介绍sql语句，定义、查询、更新三部分

四 数据库安全性

全面讲解数据库系统安全性的技术和方法，着重讲解SQL存取控制功能

五 数据库完整性

实体完整性、参照完整性、用户定义完整性。完整性约束的定义方法、完整性检查机制、违约处理。

=====设计应用开发=====

六 关系数据理论

七 数据库设计

八 数据库编程

=====系统=====

九 关系查询与查询优化

十 数据库恢复技术

十一 并发控制

十二 数据库管理系统

=====新技术=====（选择看）

十三 数据库技术新发展

十四 分布式数据库系统

十五 对象关系数据库系统

十六 XML数据库

十七 数据仓库和练级分析处理

绪论

数据、数据库、数据库管理系统、数据库系统是4个基本概念

数据

数据不仅是数字，种类很多：文本、图形、图像、音频、视屏....。

具体定义为：描述事物的符号成为数据。描述事物的符号可以是文字、图形..。他们可以经过数字化存入计算机。

数据的表现形式不能表达其内容，需要经过解释，数据和关于数据的解释是不可分的。例如90，可以表示成绩，体重，长度...

数据库

是存放数据的仓库，只是这个仓库在计算机的存储设备上，数据按照一定格式存放。

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 **按照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**。

数据库管理系统

如何科学的组织和存储数据，如何高效的获取和维护数据就是数据库管理系统的任务。

数据库管理系统位于用户和操作系统之间的一层数据管理软件。主要功能如下：

数据定义功能

数据定义语言（Data Definition Language,DDL），用户通过他可以方便的对数据库中的数据对象进行定义。

数据组织、存储、管理

DBMS要 **分类组织、存储、管理各种数据**，包括数据字典、用户数据、数据的存取路径等。要确定以何种 **文件结构** 和 **存取方式** 在存储级上组织这些数据,如何 **实现数据之间的联系**。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储，提供多种存取方法（如索引查找、Hash查找、顺序查找）等来提高存取效率。

数据操纵功能

DBMS提供数据操作语言DML。用户可以使用DML操作数据，实现对数据的基本操作，如查询，插入，删除，修改等

数据库事务管理和运行管理

数据库在建立、运用和维护时有DBMS统一管理、统一控制，保证数据的完整性、安全性、多用户对数据的并发使用及

发生故障后的恢复

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，这些功能通常由一些使用程序或管理工具完成

其他功能

包括：DBMS与网络中其他软件系统的通信功能，一个DBMS与另一个DBMS或文件系统的数据库转换功能，异构数据库之间的互访和互操作功能等

数据库系统

数据库系统指在计算机系统中引入数据库后的系统，一般指数据库、数据库管理系统（以及相关工具）、应用系统、数据库管理员构成。数据库的建立和维护还需专门的人员来完成，这些人是数据库管理员（DBA）

数据模型

模型，例如一张地图、飞机。模型是对现实世界中某个对象特征的模拟和抽象。

现实世界的人、物、活动、概念等用数据模型这个工具来抽象、表示和处理。

数据模型也是一种模型，他是现实世界数据特征的抽象。通俗的讲数据模型就是现实世界的模拟。

现有的数据库系统均是基于某种数据模型，数据模型是数据库系统的核心和基础。

数据模型：能比较真实的模拟现实世界、容易为人所理解、便于在计算机上实现。

如同在建筑设计和施工的不同阶段需要不同的图纸一样，开始实施数据库应用系统中也需要使用不同的数据模型：概念模型、逻辑模型、物理模型

概念模型

也称做信息模型、他是按用户的观点来对数据和信息建模，主要用于数据库设计

逻辑模型

包括：层次模型、网状模型、关系模型、面向对象模型、对象关系模型。它是按计算机系统的观点对数据建模，主要用于DBMS实现

物理模型

对数据最底层的抽象，描述数据在系统内部的表示方式和存取方法，在磁盘或磁带上磁带上的存储方式和存取方法。物理模型的具体实现是DBMS的任务，数据库设计人员要了解和选择物理模型，一般用户则不必考虑物理级的细节

各种机器上实现的DBMS软件都是基于某种数据模型或者说支持某种数据模型。

人们首先将现实世界抽象为信息世界--》在把信息世界转换为机器世界。信息世界不依赖于具体的计算机系统，也不是某一个DBMS支持的数据模型，而是概念级别的模型。然后再把概念模型转换为计算机上某一DBMS支持的数据模型。

数据模型组成要素

数据结构

数据结构描述数据库组成的对象以及对象的联系。数据结构描述的内容有两类：

- 对象的类型、内容、性质，例如网状模型中的数据项、记录，关系模型中的域、属性、关系。

- **数据之间联系有关的对象**，例如网状模型中的系型

数据结构是刻画数据模型性质最重要的方面，所以通常用数据结构来命名数据模型，例如关系模型、网状模型、层次模型。

数据操作

数据操作值对数据库中各种对象（型）的实例（值）允许执行的操作的集合，包括操作及操作规则。数据库主要有查询和更行两类操作

数据的完整性约束条件

数据的完整性约束是一组完整性规则。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，以保证数据的正确、有效、相容。

概念模型

概念模型应当有较强的语义表达能力，能够方便、直接的表达应用中的各种语义只是，还应当简单、清晰、易于用户理解。

信息世界的两个基本概念：

- **实体**：客观存在并可相互区别的事物称为实体，例如：一个学生
- **属性**：实体具有的某一特性称为属性，例如：（学号，姓名，性别，入学时间）
- **码(key)**：唯一标识实体的属性集称为码，例如，学生的学号
- **域**：属性的取值方位称为该属性的域，例如：性别的域(男，女)
- **实体型(Entity Type)**：具有相同属性的实体必然具有相同的特征和性质。用实体名机器属性名集合来抽象和刻画同类实体，称为实体型。例如：学生(学号，姓名，性别，入学时间)

实体之间的联系 1:1、1: n、n: n。

概念模型的一种表示方法：实体-联系

常用数据模型

- 层次模型
- 网状模型
- 关系模型
- 面向对象模型
- 对象关系模型 层次模型和网状模型统称为非关系模型，现在基本不用非关系模型。

层次模型

层次模型以树形结构表示实体以及实体之间的联系。

关系表示方法 邻接法

链接法，有两种：兄弟链接、层次链接

关系模型

数据结构

- **关系**：一个关系通常对应一张表
- **元组**：表中的一行为一个元组
- **属性**：表中的一列为一个属性
- **码**：码键。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，例如：关系名（属性1，属性2，属性3， ...,属性n）

数据操作和完整性约束

主要操作为查询、插入、删除、更新。完整性约束三大类：实体完整性、参照完整性、用户定义完整性。数据操作为集合操作，操作对象和操作结果都是关系，即若干元组的集合。

关系模型的存储结构

关系模型中实体以及实体的联系都用表来表示。有的DBMS一个表对应一个文件，有的对应多个文件，自己设计表、索引等存储结构

数据库系统结构

数据库系统的结构可以有不同层次和不同角度

从数据库管理系统角度看，数据库系统通常采用三级模式结构

从用户角度看，数据库系统结构分为，单用户、主从式、分布式、客户/服务、浏览器/应用服务器/数据库服务器。这是
数据库系统外部体系结构

数据库系统模式概念

数据模型中有型(type)和值(value)概念。型指一类数据的结构和属性的说明，值是型的一个具体赋值。例如型（学号，姓名，性别） 值（001，张三，男）

模式(Schema) 是数据库中全体数据的逻辑结构和特征的描述，她仅仅涉及到型的描述，不涉及具体的值。模式的一个具体值称为模式的一个实例(Instance)，同一个模式可以有多个实例。

例如：学生选课数据库模式中，2010年为一个实例，2011年一个实例

数据库系统的三级模式：外模式、模式、内模式

多个外模式，一个模式，一个内模式

![image](./src/数据库系统概论/Screen Shot 2017-02-21 at 8.55.49 PM.png)

外模式 也称子模式或者用户模式，外模式是模式的子集，
他是数据库用户能够看见和使用的局部数据的逻辑结构和特征的描述。他是保证数据库安全的一个有力措施

模式 也称逻辑模式，是数据库中全体数据的逻辑结构和特征的描述，**是所有用户的公共视图**。他是数据库模式的中间层，既不涉及数据的物理存储细节和硬件环境，也与具体的应用程序，开发工具和高级程序设计语言（C，COBOL，FORTRAN）无关

内模式 也称存储模式。一个数据库只有一个内模式，他是 **数据的物理结构和存储方式的描述**，是数据在数据库内部的表示方式，例如记录的存储方式是堆存储、按照某个属性值升降序存储、按照属性值聚簇存储；索引按照什么方式组织，B+树还是Hash；数据是否压缩，是否加密；数据的存储记录结构有何规定，如定长结构或变长结构，一个记录能不能跨物理页存储等等。

