

普通高等教育
软件工程

“十三五”规划教材



工业和信息化普通高等教育
“十三五”规划教材

13th Five-Year Plan Textbooks
of Software Engineering

数据库原理 与系统开发教程

张克君 ◎ 主编

章小莉 刘瑾 姜湘岗 傅仕净 谢婷婷 ◎ 编著

*Database Principle and
System Development Tutorial*

非
外
借

中国工信出版集团

人民邮电出版社
POSTS & TELECOM PRESS



普通高等教育

软件工程

“十三五”规划教材



工业和信息化普通高等教育

“十三五”规划教材

13th Five-Year Plan Textbooks
of Software Engineering

数据库原理 与系统开发教程

张克君 ○ 主编

章小莉 刘瑾 姜湘岗 傅仕净 谢婷婷 ○ 编著

*Database Principle and
System Development Tutorial*



人民邮电出版社

北京

图书在版编目(CIP)数据

数据库原理与系统开发教程 / 张克君主编 ; 章小莉
等编著. — 北京 : 人民邮电出版社, 2018.1
普通高等教育软件工程“十三五”规划教材
ISBN 978-7-115-47754-5

I. ①数… II. ①张… ②章… III. ①关系数据库系
统一系统开发—高等学校—教材 IV. ①TP311.138

中国版本图书馆CIP数据核字(2018)第010332号

内 容 提 要

本书简洁明了地讲述了数据库技术的基本原理,并以一个具体的数据库系统为例,完整地介绍了数据库系统项目开发的整个过程。全书分为三篇,第一篇为数据库原理篇,主要内容包括数据库系统结构、数据模型、关系数据库、关系数据库标准语言 SQL、关系数据库设计规范化理论、数据库的安全性、数据库的完整性、数据处理新技术,共7章;第二篇为数据库系统开发篇,主要内容包括数据库系统软件开发模型、以一个仓库管理系统为例的需求分析、业务设计、编码实现等数据库系统开发步骤的详细过程,以及所用到的工程开发工具的使用方法,共5章;第三篇为数据库系统开发任务集篇,主要内容包括对管理信息系统、电子政务和电子商务平台的简介,以及三类开发项目的待开发任务集的内容描述,为学生开展数据库系统项目开发实践提供参考选题,共1章。

本书从实际问题出发,完整地剖析了数据库系统项目开发的分析、设计和实现的全过程,力求达到提高读者设计能力以及激发创新的目的。

本书可作为计算机科学与技术、信息管理与信息系统等专业本科生和研究生相关课程的教材,也可作为IT行业从业人员的岗前培训参考书。

-
- ◆ 主 编 张克君
编 著 章小莉 刘 瑾 姜湘岗 傅仕净 谢婷婷
责任编辑 邹文波
责任印制 彭志环
 - ◆ 人民邮电出版社出版发行 北京市丰台区成寿寺路11号
邮编 100164 电子邮件 315@ptpress.com.cn
网址 <http://www.ptpress.com.cn>
固安县铭成印刷有限公司印刷
 - ◆ 开本: 787×1092 1/16
印张: 13.75 2018年1月第1版
字数: 334千字 2018年1月河北第1次印刷
-

定价: 45.00 元

读者服务热线: (010)81055256 印装质量热线: (010)81055316

反盗版热线: (010)81055315

随着计算机信息处理技术、网络技术的快速发展,电子商务、电子政务等社会信息化应用的不断深入,云计算、大数据等数据智能处理技术不断涌现,信息化正深刻地影响和改变着人类社会的管理及生活方式。信息化水平已成为衡量一个国家和地区现代化水平的重要标志。数据库技术为信息系统的构建提供信息存储和处理平台,是信息化的核心技术。培养掌握数据库技术的专业人才对国家发展具有重要意义。高等学校主要通过开设数据库相关课程来培养学生的数据库系统开发技能。

目前,我国高等教育正从粗放式的规模发展转向内涵式发展,教学质量已经成为关注重点,人才培养模式也更加注重理论联系实际的人才培养理念。高等教育“质量工程”的重要任务之一是强化各实践教学环节,推进实验内容和模式的改革与创新,培养学生分析和解决问题的能力。数据库系统开发建设具有很强的工程性,数据库教学必须为学生建立起完整的理论体系和培养扎实的实践功底。本书系统地梳理了数据库的相关原理,然后以一个实际工程建设案例为主线,引入系统建设工程方法,经过类瀑布模型的需求分析、设计、编码、测试等建设环节,运用当前流行的业务建模、存储建模开发工具完成数据库信息系统的设计开发。让读者在感性理解数据库原理的同时,全面地掌握数据库系统的工程建设方法,并能够独立完成从需求分析到应用系统开发的全过程。本书旨在使读者得到全面的训练,进而提升分析问题和解决问题的能力。

本教材分为三篇。第一篇为数据库原理篇,表述简洁,逻辑性强,通俗易懂,符合初学者学习的认知规律;第二篇为数据库系统开发篇,提供了完整的数据库系统开发案例,引入了软件工程建设方法,运用了多种新版本的设计和开发工具,如关系型数据库管理系统 MySQL、UML 业务建模工具 RSA、数据模型建模工具 PowerDesigner、集成开发环境 Eclipse 等,有利于培养学生分析问题和系统开发的能力;第三篇为数据库系统开发任务集篇,提供了多个不同表达形式的待开发系统业务描述,供学生开发训练使用,以达到通过实践深入理解原理知识的目的。

本教材为读者提供了教学视频和互动平台。读者可以通过扫描书中相应位置的二维码,观看相关教学视频;读者还可以通过关注“北电院大数据分析 BestiDA”微信公众号,及时获取最新知识,反馈意见。

在本书的编写过程中,编者参考了很多优秀的同类教材和资料,并有选择地将它们纳入到教材之中,在此一并致谢。由于本学科发展日新月异,加之编者水平所限,书中难免存在欠妥之处,敬请读者批评指正。

编者

2018年1月

目录 CONTENTS

第一篇 数据库原理篇

第1章 数据库系统概述..... 2

- 1.1 数据库技术的发展历程..... 2
- 1.2 数据库有关基本概念..... 4
- 1.3 数据库系统的组成..... 5
- 1.4 数据库系统的结构..... 8
 - 1.4.1 数据库系统的内部结构..... 8
 - 1.4.2 数据库系统的外部结构..... 10
- 1.5 数据模型..... 13
 - 1.5.1 数据模型的概念与分类..... 13
 - 1.5.2 数据模型的组成要素..... 14
 - 1.5.3 概念模型及其 E-R 表示方法..... 15
 - 1.5.4 逻辑模型..... 17
- 小结..... 21
- 习题..... 21

第2章 关系数据库..... 22

- 2.1 关系模型的数据结构及形式化定义..... 22
 - 2.1.1 关系的基本术语..... 22
 - 2.1.2 关系的性质..... 26
 - 2.1.3 关系模式..... 27
 - 2.1.4 关系数据库与关系数据库模式..... 27
- 2.2 关系的完整性..... 28
 - 2.2.1 主码与实体完整性..... 28
 - 2.2.2 外码与参照完整性..... 28
 - 2.2.3 用户自定义完整性..... 29
- 2.3 关系代数..... 30

2.3.1 传统的集合运算..... 30

2.3.2 专门的关系运算..... 31

小结..... 35

习题..... 35

第3章 关系数据库标准语言 SQL..... 36

- 3.1 SQL 概述..... 36
 - 3.1.1 SQL 的产生与发展..... 37
 - 3.1.2 SQL 的基本概念..... 37
 - 3.1.3 SQL 的特点..... 38
- 3.2 MySQL 数据库简介..... 39
 - 3.2.1 MySQL 常用的语句..... 39
 - 3.2.2 MySQL 的数据类型..... 40
- 3.3 定义数据..... 42
 - 3.3.1 创建与使用模式..... 42
 - 3.3.2 创建与使用基本表..... 43
 - 3.3.3 创建与使用索引..... 45
- 3.4 查询数据..... 45
 - 3.4.1 Select 命令的一般格式..... 46
 - 3.4.2 单表查询..... 46
 - 3.4.3 连接查询..... 50
 - 3.4.4 嵌套查询..... 52
 - 3.4.5 集合查询..... 54
- 3.5 更新数据..... 55
 - 3.5.1 插入数据..... 55
 - 3.5.2 修改数据..... 57
 - 3.5.3 删除数据..... 58
- 3.6 关于视图..... 59
- 小结..... 60
- 习题..... 62

第4章 关系数据库设计**规范化理论** 63

4.1 规范化问题的提出 63

4.2 函数依赖 64

4.2.1 函数依赖的定义 64

4.2.2 平凡函数依赖与非平凡
函数依赖 644.2.3 完全函数依赖与部分
函数依赖 65

4.2.4 传递函数依赖 65

4.3 码(键)的概念 65

4.4 关系模式的范式 65

4.4.1 第一范式(1NF) 66

4.4.2 第二范式(2NF) 67

4.4.3 第三范式(3NF) 67

4.4.4 BC范式(BCNF) 67

小结 68

习题 69

第5章 数据库的安全性 70

5.1 数据库安全性概述 70

5.1.1 数据库的安全性问题 71

5.1.2 数据库系统相关
安全标准 725.2 数据库安全性控制的
常用方法 72

5.2.1 用户身份标识与鉴别 73

5.2.2 存取权限控制 73

5.2.3 视图机制 73

5.2.4 数据加密 74

5.2.5 审计机制 75

5.2.6 统计数据库 75

5.3 MySQL 数据库的数据
安全性机制 75

5.3.1 创建数据库角色和用户 76

5.3.2 访问权限的授予和回收 78

小结 81

习题 81

第6章 数据库的完整性 82

6.1 数据库完整性的含义 82

6.2 DBMS 完整性控制机制 85

6.2.1 完整性约束条件定义功能 85

6.2.2 检查功能 85

6.2.3 违约方反应 86

6.3 完整性约束条件的分类 87

6.4 MySQL 中的完整性机制 88

小结 90

习题 91

第7章 数据处理新技术 92

7.1 数据仓库与数据挖掘 92

7.1.1 数据仓库 92

7.1.2 数据挖掘 93

7.1.3 联机分析处理 94

7.2 大数据技术 95

7.2.1 大数据的概念与特征 95

7.2.2 大数据管理系统 96

7.2.3 大数据应用 97

小结 98

习题 98

第二篇 数据库系统开发篇**第8章 数据库系统开发****概述** 100

8.1 数据库系统开发的基本流程 100

第 10 章 仓库管理系统的业务设计 129

10.1 功能设计 129

10.1.1 商品清单 129

10.1.2 商品出货 130

10.1.3 出货记录 131

10.1.4 进货记录 132

10.1.5 基础数据 133

10.1.6 库存盘点 139

10.2 页面设计 139

10.2.1 商品清单页面 140

10.2.2 商品出货页面 140

10.2.3 出货记录页面 140

10.2.4 进货记录页面 141

10.2.5 库存盘点页面 141

10.2.6 基础数据——商品信息
页面 142

10.2.7 基础数据——规格信息
页面 144

10.2.8 基础数据——联系方式
页面 145

10.2.9 基础数据——收款账号
页面 146

10.2.10 基础数据——客户信息
页面 147

小结 149

习题 149

第 11 章 仓库管理系统的数据库设计 150

11.1 概念模型的设计 150

11.2 逻辑模型的设计 153

11.3 物理模型的设计 154

小结 154

8.1.1 系统分析 100

8.1.2 业务设计 101

8.1.3 数据库设计 101

8.1.4 编码与测试 103

8.1.5 部署与运维 103

8.2 数据库系统开发常用的
建模工具 103

8.3 Rational Software Architect
建模工具的使用 104

8.3.1 用 RSA 创建项目 104

8.3.2 用 RSA 进行
UML 建模 106

8.4 PowerDesigner 数据库
设计建模工具的使用 109

8.4.1 PowerDesigner 基础 109

8.4.2 用 PowerDesigner 进行
数据建模 110

8.5 仓库管理系统 (WMS)
开发案例概述 114

小结 115

习题 115

第 9 章 仓库管理系统的需求分析 116

9.1 问题描述 116

9.2 问题分析 116

9.3 功能描述 117

9.4 系统涉众 119

9.5 概要结构 119

9.6 用例解析 119

9.7 用例规约 121

9.8 活动解析 125

小结 128

习题 128

习题	155
----	-----

第 12 章 仓库管理系统

编码实现	156
------	-----

12.1 实体类的编码实现	156
---------------	-----

12.2 业务功能的编码实现	169
----------------	-----

12.2.1 商品清单	170
-------------	-----

12.2.2 商品出货	171
-------------	-----

12.2.3 出货记录	178
-------------	-----

12.2.4 进货记录	181
-------------	-----

12.2.5 库存盘点	183
-------------	-----

12.2.6 基础数据	184
-------------	-----

小结	199
----	-----

习题	199
----	-----

第三篇 数据库系统 开发任务集篇

第 13 章 各类待开发

应用系统	202
------	-----

13.1 管理信息系统	202
-------------	-----

13.2 电子政务	205
-----------	-----

13.2.1 政府间的电子政务	206
-----------------	-----

13.2.2 政府对企业的电子政务	206
-------------------	-----

13.2.3 政府对公民的电子政务	207
-------------------	-----

13.3 电子商务	208
-----------	-----

小结	210
----	-----

习题	210
----	-----

第一篇 数据库原理篇

- 第 1 章 数据库系统概述
- 第 2 章 关系数据库
- 第 3 章 关系数据库标准语言 SQL
- 第 4 章 关系数据库设计规范化理论
- 第 5 章 数据库的安全性
- 第 6 章 数据库的完整性
- 第 7 章 数据处理新技术

众所周知,世界上第一台真正意义上的计算机是为了满足军事领域中的数学运算而产生的,运算是计算机的第一任务。然而随着计算机技术的不断成熟和发展,越来越多的用途被提出并得到广泛的应用,其中,数据处理是迄今为止最主要的计算机应用之一。在计算机领域中数据处理是指借助计算机对数据进行收集、管理、加工利用乃至信息输出的演变与推导,最终获取信息的过程。数据管理是处理的核心内容。数据库作为一门研究数据管理的技术,始于20世纪60年代,经过50多年的发展已经形成非常完善的理论体系,成为计算机应用领域中非常重要的一个分支。

在信息化社会,充分有效地管理和利用各类信息资源,是进行科学研究和决策管理的前提条件。作为管理信息系统、办公自动化系统、决策支持系统等各类信息系统的核心部分,数据库是进行科学研究和决策管理的重要技术手段。对数据库的学习应当成为计算机相关领域非常重要的环节。

1.1 数据库技术的发展历程

数据库是数据管理的产物。数据管理是数据库的核心任务,内容包括对数据的分类、组织、编码、储存、检索和维护。随着计算机硬件和软件的发展,数据库技术也不断发展。数据管理经过了人工管理、文件系统和数据库系统三大阶段。

1. 人工管理阶段

人工管理阶段是指现代计算机初次出现的时期(20世纪50年代)。这一时期的计算机主要用于科学计算。从硬件看,没有磁盘等直接存取的存储设备;从软件看,没有操作系统和管理数据的软件,数据处理方式是批处理。

这个时期数据管理的特点如下。

(1) 数据不保存

该时期的计算机主要应用于科学计算,一般不需要将数据长期保存,只是在计算某一课题时将数据输入,用完后既不保存原始数据,也不保存计算结果。

(2) 没有对数据进行管理的软件系统

程序员不仅要规定数据的逻辑结构,而且要在程序中设计物理结构,包括存储结构、存取方法、输入/输出方式等。因此程序中存取数据的子程序随着存储的改变而改变,数据与程序不具有 consistency。

(3) 没有文件的概念

数据的组织方式必须由程序员自行设计。

(4) 一组数据对应一个程序，数据是面向应用的

即使两个程序用到相同的数据，也必须各自定义、各自组织，数据无法共享、无法相互利用和互相参照，从而导致程序和程序之间存在大量重复的数据。

2. 文件系统阶段

随着计算机处理数据的应用规模越来越大，需要处理的数据量越来越多，人工管理阶段对数据的管理方式已经越来越跟不上数据管理的发展，文件系统应运而生。这一时期从 20 世纪 50 年代后期开始直到 20 世纪 60 年代中期。文件系统阶段的出现是计算机技术及数据管理技术发展 to 一定程度的产物。在这一时期计算机不仅用于科学计算，还被用于大量数据的管理，而且计算机软硬件有了长足的发展。在硬件方面，外存储器有了磁盘、磁鼓等直接存取的存储设备。在软件方面，操作系统中已经有了专门用于管理数据的软件，称为文件系统。

这个时期数据管理的特点如下。

(1) 数据需要长期保存在外存上供反复使用。由于计算机大量用于数据处理，经常对文件进行查询、修改、插入和删除等操作，所以数据需要长期保留，以便于反复操作。

(2) 程序之间有了一定的独立性。操作系统提供了文件管理功能和访问文件的存取方法，程序和数据之间有了数据存取的接口，程序可以通过文件名和数据打交道，不必再寻找数据的物理存放位置，至此，数据有了物理结构和逻辑结构的区别，但此时程序和数据之间的独立性还不充分。

(3) 文件的形式已经多样化。由于已经有了直接存取的存储设备，文件也就不再局限于顺序文件，还有了索引文件、链表文件等，因而，对文件的访问可以是顺序访问，也可以是直接访问。

(4) 数据的存取基本上以记录为单位。

3. 数据库系统阶段

数据库系统阶段是从 20 世纪 60 年代后期开始的。在这一阶段，数据库中的数据不再是面向某个应用或某个程序，而是面向整个企业（组织）或整个应用的。数据库系统阶段，数据管理的特点如下。

(1) 采用复杂的结构化的数据模型

数据库系统不仅要描述数据本身，还要描述数据之间的联系。这种联系是通过存取路径来实现的。

(2) 较高的数据独立性

数据和程序彼此独立，数据存储结构的变化尽量不影响用户程序的使用。

(3) 最低的冗余度

数据库系统中的重复数据被减少到最低程度，这样，在有限的存储空间内可以存放更多的数据并减少存取时间。

(4) 统一的数据控制功能

数据库系统中的数据由数据库管理系统统一管理和控制。

数据库系统具有数据的安全性，以防止数据丢失和被非法使用；具有数据的完整性，以保护数据的正确、有效和相容；具有数据的并发控制，避免并发程序之间相互干扰；具有数据的恢复功能，在数据库被破坏或数据不可靠时，系统有能力把数据库恢复到最近某个时刻的正确状态。

1.2 数据库有关基本概念

作为一套相对较为成熟的理论体系，数据库有其主要的基本术语和常用的概念，其中数据、数据库、数据库管理系统和数据库系统是最为基础的，也是理解数据库的入门知识。

1. 数据

数据 (Data) 是数据库中存储的基本对象。生活中，例如文本 (Text)、图形 (Graph)、图像 (Image)、音频 (Audio)、学生记录档案、货物的运输情况等，都可以称为数据。

我们通常把描述事物的符号记录称为数据。数据有多种表现形式，如文字、图形、图像、音频、视频等，它们都可以经过数字化后存入计算机。

然而，仅仅一个数据还不能表达完整的意思，例如，93 是一个数据，它可以是一门课的成绩，也可以是一个班级的人数，还可以是桌子的高度，所以需要说明数据的含义。数据的含义称为数据的语义，数据与其语义是不可分的。

2. 数据库

数据库 (Database, DB) 从字面上理解，就是存放数据的仓库。大量的数据按照一定的组织形式存放在计算机的存储设备上就形成了数据库。

严格地说，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 (Redundancy)、较高的数据独立性 (Data Independency) 和易扩展性 (Scalability)，并可为各种用户共享。

数据库具有永久存储、有组织和可共享三个基本特点。

3. 数据库管理系统

了解了数据和数据库的概念，那么如何科学地组织和存储数据，如何高效地获取和维护数据？完成这个任务的是一个系统软件——数据库管理系统。

数据库管理系统 (Database Management System, DBMS) 是一种操纵和管理数据库的大型软件，用于建立、使用和维护数据库。它对数据库进行统一管理和控制，以保证数据库的安全性和完整性。用户通过 DBMS 访问数据库中的数据，数据库管理员也通过 DBMS 进行数据库的维护工作。它可使多个应用程序和用户用不同的方法在同时或不同时刻建立、修改和询问数据库。大部分 DBMS 提供数据定义语言 (Data Definition Language, DDL) 和数据操作语言 (Data Manipulation Language, DML)，供用户定义数据库的模式结构与权限约束，实现对数据的追加、删除等操作。数据库管理系统的主要功能包括以下几个方面。

(1) 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定义数据库中的数据对象的组成与结构。

(2) 数据组织、存储和管理

数据库管理系统要分类组织、存储和管理各种数据，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取，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。

(3) 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。

(4) 数据库的事务管理和运行管理

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,以保证事务的正确运行,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。

(5) 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,数据库的重组功能和性能监视、分析功能等。这些功能通常是由一些实用程序或管理工具完成的。

(6) 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库转换功能、异构数据库之间的互访和互操作功能等。

4. 数据库系统

数据库系统(Database System, DBS)是由数据库、数据库管理系统(及其应用开发工具)、应用程序和数据库管理员(Database Administrator, DBA)组成的存储、管理、处理和维持数据的系统。数据库系统是为适应数据处理的需要而发展起来的一种较为理想的数据处理的核心机构。计算机的高速处理能力和大容量存储器提供了实现数据管理自动化的条件。

1.3 数据库系统的组成

数据库系统是引入了数据库的应用系统。从这一概念出发,可以得出数据库系统的基本组成。数据库系统通常情况下可以分成3个部分:硬件平台及数据库、软件体系、人员体系,如图1.1所示。

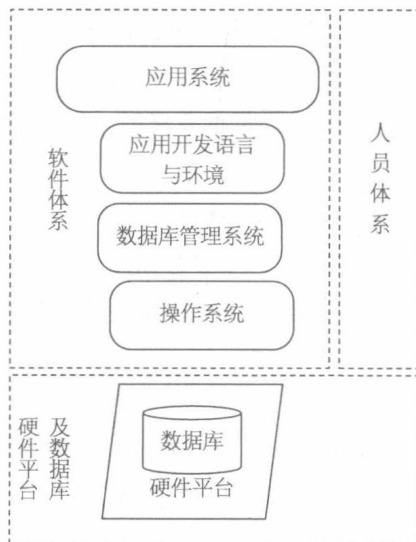


图1.1 数据库系统的组成

1. 硬件平台及数据库

数据库作为数据的集合是存储在计算机硬件设备上的，同时，构成数据库系统的软件体系也需要由硬件平台来支撑，所以硬件平台是整个数据库系统的基础设施层。对硬件平台的要求主要从存储和数据交换两个方面来衡量。

首先在存储方面，包含存放数据的外存储器和协助运行程序的内存存储器。外存储器主要用于存放数据库本身，数据库本身由大量静态的数据和对数据的定义组成，除了存放原始的数据之外，还需要提供足够的备份空间。内存存储器主要用于存放操作系统、数据库管理系统和应用程序等需要运行的程序以及提供运行期间的数据缓冲区。

其次，由于在硬件平台与软件体系之间需要进行大量的数据传输，所以对于硬件平台来说，提高系统通道能力，提供较高的数据传送率是必要的需求。

2. 软件体系

数据库系统的软件体系由操作系统、数据库管理系统、应用开发语言与环境（应用开发工具）和应用系统 4 部分构成。

（1）操作系统。作为管理和控制计算机硬件与软件资源的系统软件，操作系统为数据库系统提供了软件支持，所有其他软件都只有在操作系统的支持下才能运行。

（2）数据库管理系统。数据库管理系统用于操纵和管理数据库，它借助于操作系统建立、使用和维护配置数据库，通过 DBMS 对数据库进行统一管理和控制，以保证数据库的完整性和安全性。

（3）应用开发语言与环境（应用开发工具）。DBMS 提供了对数据库本身的定义及访问的功能接口。应用开发语言与环境为应用开发人员提供了构建用户接口的主要应用层。这一层由应用开发高级语言及其编译系统和应用开发环境构成，为数据库应用系统的开发提供了软件支持。

（4）应用系统。在整个数据库系统的最上层，呈现给用户的是借助于应用开发语言基于数据库管理系统，为特定应用环境开发的数据库应用系统。

3. 人员体系

数据库系统是与实际应用密不可分的系统，除了软硬件环境之外，还需要各种人员的密切参与。开发、管理和使用数据库系统的人员主要有：数据库管理员、系统分析员和数据库设计人员、应用程序员和最终用户。不同的人员涉及不同的数据抽象级别，具有不同的数据视图，其各自的职责分别如下。

（1）数据库管理员（DataBase Administrator, DBA）

在数据库系统环境下，有两类共享资源，一类是数据库，另一类是数据库管理系统软件。因此需要有专门的管理机构来监督和管理数据库系统。DBA 则是这个机构的一个（组）人员，负责全面管理和控制数据库系统。具体职责如下。

① 决定数据库中的信息内容和结构。

DBA 要参与决策数据库中要存放哪些信息。因此 DBA 必须参加数据库设计的全过程，并与用户、应用程序员、系统分析员密切合作共同协商，设计好数据库。

② 决定数据库的存储结构和存取策略。DBA 要综合各用户的应用要求，和数据库设计人员共同决定数据的存储结构和存取策略，以求获得较高的存取效率和存储空间利用率。

③ 定义数据的安全性要求和完整性约束条件。

DBA 的重要职责是保证数据库的安全性和完整性。因此 DBA 负责确定各个用户对数据库的存取权限、数据的保密级别和完整性约束条件。

④ 监控数据库的使用和运行。

DBA 还有一个重要职责就是监视数据库系统的运行情况，及时处理运行过程中出现的问题。比如系统发生各种故障时，数据库会因此遭到不同程度的破坏，DBA 必须在最短时间内将数据库恢复到正确状态，并尽可能不影响或少影响计算机系统其他部分的正常运行。为此，DBA 要定义和实施适当的后备和恢复策略，如周期性地转储数据、维护日志文件等。有关这方面的内容将在下面进一步讨论。

⑤ 数据库的改进和重组、重构

DBA 还负责在系统运行期间监视系统的空间利用率、处理效率等性能指标，对运行情况进行记录、统计分析，依靠工作实践并根据实际应用环境，不断改进数据库设计。不少数据库产品都提供了对数据库运行状况进行监视和分析的实用程序，DBA 可以使用这些实用程序完成这项工作。

另外，在数据运行过程中，大量数据不断插入、删除、修改，时间一长，会影响系统的性能。因此，DBA 要定期对数据库进行重组，以提高系统的性能。

当用户的需求增加和改变时，DBA 还要对数据库进行较大的改造，包括修改部分设计，即数据库的重构造。

(2) 系统分析员和数据库设计人员

系统分析员负责应用系统的需求分析和规范说明，要和用户及 DBA 结合，确定系统的硬件、软件配置，并参与数据库系统的概要设计。

数据库设计人员负责确定数据库中的数据、设计数据库各级模式。数据库设计人员必须参加用户需求调查和系统分析，然后设计数据库。在很多情况下，数据库设计人员就由数据库管理员担任。

(3) 应用程序员

应用程序员负责设计和编写应用系统的程序模块，并进行调试和安装。

(4) 用户

这里用户是指最终用户 (End User)。最终用户通过应用系统的用户接口使用数据库。常用的接口方式有浏览器、菜单驱动、表格操作、图形显示、报表书写等，给用户提供简明直观的数据表示。

最终用户可以分为如下 3 类。

① 偶然用户。这类用户不经常访问数据库，但每次访问数据库时往往需要不同的数据库信息，这类用户一般是企业或组织机构的中高级管理人员。

② 简单用户。数据库的多数最终用户都是简单用户。其主要工作是查询和修改数据库，一般都是通过应用程序员精心设计并具有友好界面的应用程序存取数据库。银行的职员、航空公司的机票预定工作人员、旅馆总台服务员等都属于这类用户。

③ 复杂用户。复杂用户包括工程师、科学家、经济学家、科学技术工作者等具有较高科学技术背景的人员。这类用户一般都比较熟悉数据库管理系统的各种功能，能够直接使用数据库语言访问数据库，甚至能够基于数据库管理系统的 API 编制自己的应用程序。

1.4 数据库系统的结构

数据库系统的组成阐述了一个完整的数据库系统应该包含的软硬件层次和人员的配备。对于包含了多种软件体系和人员体系的数据库系统来说，它的体系结构也是需要重点掌握的内容。从不同的用户体系出发，可以从两个角度来讨论数据库系统结构，一个是关注于数据库管理系统的内部结构，另一个是关注于应用层面的外部结构。

1.4.1 数据库系统的内部结构

数据库系统的内部结构是从数据库管理系统的角度来理解的数据库系统的结构，它围绕数据库管理系统对数据的描述展开讨论，更多的着眼点在数据库系统本身，所以往往被称为数据库系统的内部结构。

数据库系统的内部结构通常采用三级模式和两级映像来描述，如图 1.2 所示。

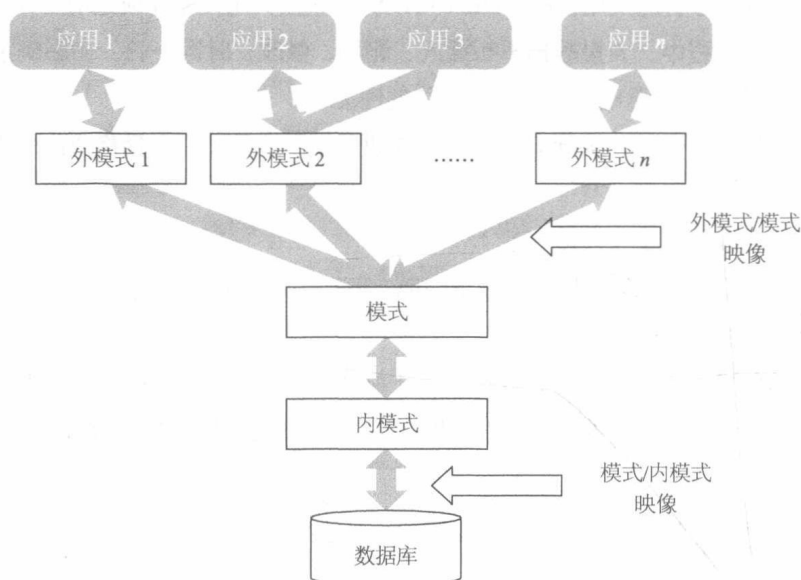


图1.2 数据库系统的三级模式和两级映像

1. 模式

模式 (Schema) 是数据库中全体数据的逻辑结构和特征的描述。它仅涉及结构的描述，不涉及数据库本身具体的值。模式的一个具体值称为模式的一个实例 (Instance)，同一个模式可以有很多实例。模式是相对稳定的，而实例是相对变动的，因为数据库中的数据在不断更新。模式反映的是数据的结构及其联系，而实例反映的是数据库某一时刻的状态。例如，学生记录定义为 (学号，姓名，性别，系别，年龄，籍贯)，这是记录结构，而 (111987，张明，男，计算机，22，郑州) 则是该记录结构的一个记录值。

虽然实际的数据库管理系统产品种类很多，它们支持不同的数据模型，使用不同的数据库语言，建立在不同的操作系统之上，数据的存储结构也各不相同，但在体系结构上通常都具有相同的特征，即采用三级模式结构并提供两级映像功能。

2. 三级模式结构

(1) 概念模式

概念模式简称模式，又称为数据库模式、逻辑模式，是数据库中全体数据的逻辑结构和特征的描述，是全体用户的公共数据视图。它是数据库系统三级模式结构的中间层，既不涉及数据的物理存储细节和硬件环境，又与具体的应用程序、所使用的应用开发工具及高级程序设计语言无关。

模式实际上是数据库数据在概念级上的视图。一个数据库只有一个概念模式，数据库概念模式以某一种数据模型为基础，统一综合地考虑所有用户的需求，并将这些需求有机地结合成一个逻辑整体。定义模式时不仅要定义数据的逻辑结构，如数据记录由哪些数据项构成，数据项的名称、类型、取值范围等，而且要定义数据之间的联系，以及与数据有关的安全性、完整性要求。

DBMS 提供模式描述语言（模式 DDL）来严格定义模式。

(2) 外模式

外模式是概念模式的子集，也称子模式或用户模式。外模式是与某一具体应用有关的数据的逻辑结构和特征的描述，是数据库用户（包括应用程序员和最终用户）所看到的数据视图。

一个数据库可以有多个外模式。由于不同的用户在应用需求、看待数据的方式、对数据保密的要求等方面存在差异，其外模式描述也有所差异。即使模式中的同一数据，在外模式中的结构、类型、长度、保密级别等方面都可以不同。另一方面，同一外模式也可以为某一用户的多个应用系统使用，但一个应用程序只能使用一个外模式。

外模式是保证数据库安全性的有力措施，每个用户只能看见和访问对应的外模式中的数据，数据库中的其余数据不可见。

DBMS 提供子模式描述语言（子模式 DDL）来严格定义子模式。

(3) 内模式

内模式也称存储模式或物理模式，它是数据库的物理存储结构和存储方式的描述，是数据在数据库内部的表示方式。一个数据库只有一个内模式，在内模式中规定了数据项、记录、键、索引和存取路径等所有数据的物理组织以及优化性能、响应时间和存储空间需求等信息，还规定了记录的位置、块的大小和溢出区等。例如，记录的存储方式是顺序存储；是按照 B 树结构存储，还是按 hash 方法存储；索引按照什么方式组织；数据是否压缩存储，是否加密；数据的存储记录结构有何规定等。

内模式是 DBMS 管理的最低层。虽然称其为物理模式，但它不涉及物理记录的形式，如物理块或页、具体设备的柱面与磁道大小等，内部视图仍然不是物理层，是最接近物理存储的数据存储方式，是物理存储设备上存储数据时的物理抽象。DBMS 提供内模式描述语言（内模式 DDL，或者存储模式 DDL）来严格定义内模式。

3. 三级模式和两级映像

数据库系统的三级模式是对数据 3 个级别的抽象视图的描述。使用户能逻辑地、抽象地处理数据，而不必关心数据在计算机中的具体表示方式与存储方式。为了能够在内部实现这 3 个抽象层次的联系和转换，数据库管理系统在这三级模式之间提供了两级映像：外模式/模式映像和模式/内模式映像，正是这两层映像保证了数据库系统中的数据能够具有较高的逻辑独立性和物理独立性。

(1) 外模式/模式映像

模式描述的是数据的全局逻辑结构，外模式描述的是数据的局部逻辑结构，同一个模式可以有任