



普通高等教育“十三五”计算机类规划教材

数据库原理及应用

◎ 赵鲁涛 张志刚 主编



机械工业出版社
CHINA MACHINE PRESS

普通高等学

类规划教材

数据库原理及应用

主 编 赵鲁涛 张志刚
参 编 李 晔 安 璐 吕国才

机械工业出版社

本书涵盖了数据库基本理论、应用和实际操作三部分内容。基础理论部分着重讲述了数据库的概念、数据模型、关系数据库、结构化查询语句 SQL、关系数据库的规范化理论等数据库基本理论知识；应用部分结合 ADO. NET 语言，重点讲述了数据库应用系统的详细开发范例；实际操作部分给出了 8 个数据库上机实验，供读者进行练习。

本书从理论到案例展示，再到读者实践，知识体系设计严密，概念表述准确，内容紧跟数据库发展，讲解深入浅出，并配有习题解答、程序代码和教学 PPT，方便教师教学和读者自学。

本书可作为计算机类专业、通信类专业和数学类专业的本科或专科教材，也可供从事计算机工程与应用工作的科技工作者参考。

图书在版编目 (CIP) 数据

数据库原理及应用/赵鲁涛, 张志刚主编. —北京: 机械工业出版社, 2017. 5
普通高等教育“十三五”计算机类规划教材
ISBN 978-7-111-56834-6

I. ①数… II. ①赵… ②张… III. ①关系数据库系统—高等学校—教材
IV. ①TP311. 138

中国版本图书馆 CIP 数据核字 (2017) 第 104748 号

机械工业出版社 (北京市百万庄大街 22 号 邮政编码 100037)

策划编辑: 王 康 责任编辑: 王 康 王小东

责任校对: 杜雨霏 封面设计: 张 静

责任印制: 李 昂

三河市宏达印刷有限公司印刷

2017 年 5 月 1 版第 1 次印刷

184mm × 260mm · 16.25 印张 · 403 千字

标准书号: ISBN 978-7-111-56834-6

定价: 36.00 元

凡购本书, 如有缺页、倒页、脱页, 由本社发行部调换

电话服务

网络服务

服务咨询热线: 010-88379833 机工官网: www.cmpbook.com

读者购书热线: 010-88379649 机工官博: weibo.com/cmp1952

教育服务网: www.cmpedu.com

封面无防伪标均为盗版

金书网: www.golden-book.com

前言

数据库技术自 20 世纪 60 年代产生以来,经过 50 多年的快速发展,已经成为数据管理最重要的方法。当前,绝大部分计算机应用均以数据库技术作为数据管理方式,数据库技术已经成为计算机信息系统和应用系统的核心技术和重要基础。目前,基于数据库基本技术,又衍生出数据仓库、数据集市、数据湖等先进的大数据技术,极大地促进了大数据的发展。

本书由浅入深地对数据库原理及应用进行了全面讲解,在讨论理论的同时更侧重于应用和实践,可作为高等院校计算机类、通信类、数学类专业数据库相关课程的教材,既适用于数据库技术的初学者,也适用于想要完整学习数据库理论的开发者。

本书从理论、操作和实际应用三个层面,系统地阐述了数据库技术的基本理论,数据库的设计、建立、管理,数据库技术在实际中的应用。全书共分为 14 章,各章内容为:第 1 章为绪论,概述数据库系统的基本概念、发展历程、数据库系统组成和结构以及常用的数据库管理系统;第 2 章为数据模型,主要讲解数据模型的相关知识,包括数据模型的概念、组成、类型以及概念模型和三大数据模型;第 3 章为关系数据库,阐述关系数据库的结构,即关系数据模型的关系数据结构、关系完整性约束和关系操作集合三大要素,重点介绍了关系的数学定义和关系代数等;第 4 章为结构化查询语言 SQL,主要讲述 SQL 语言的概念、语法及其使用方法;第 5 章为关系数据库的规范化理论,主要介绍数据库规范化理论,讲解函数依赖、多值依赖等重要概念,着重讨论用规范化理论分析关系模式设计的合理性;第 6 章为数据库恢复技术,主要讲解数据库恢复的概念和常用技术;第 7 章为并发控制,主要介绍数据库管理系统中的并发控制操作,通过并发操作,可以满足不同用户同时访问数据库的需要;第 8 章为数据库结构设计,主要阐述数据库结构设计的基本方法;第 9 章为数据库安全,主要从计算机、数据库、软件三个方面讲解数据库安全问题以及解决手段;第 10 章为现代数据库技术及其发展,主要介绍几种当前常用的现代数据库技术;第 11 章为 ADO.NET 入门教程,主要讲解 ADO.NET 技术的基本概念、组件以及执行过程,使读者对 ADO.NET 技术有基本的了解;第 12 章讲解在 Web 开发中的数据库应用系统访问技术;第 13 章以在线购物平台为例,用 B/S 和 C/S 两种架构来实现数据库应用系统开发;第 14 章为上机实验,为了让读者更好地理解和应用数据库技术,设计了 SQL Server 认识实验、数据库管理、数据库应用系统开发等 8 个实验供读者操作学习。

本书侧重于数据库系统的设计、建立和应用,以关系数据库为重点,介绍和演示数据库的使用方式,兼顾基础理论和实践拓展,包含严格的数据库理论以及详细的实践指南,使读者在理解数据库理论知识的同时能很快将技术运用到实际开发中;以在线购物平台为例,展示数据库的应用方式,全书主要运用多个案例贯通全文,演示关系代数、SQL 等数据库的重



要应用方法，思路连贯，也便于读者理解；除提供完整的案例之外，还有8个上机实验，同时提供相关程序代码，方便读者自行学习和实践；为了紧跟数据库发展前沿，穿插讲解了数据库目前的前沿技术，为读者进一步学习提供建议。每一章后面都配有习题，让读者在实践中巩固所学理论知识。本书内容可按96学时、64学时（重点讲授1~5章以及11~14章部分内容）和48学时（重点讲授1~5章及14章部分内容）讲授。

本书第1、2、3、5章由李晔、赵鲁涛编写，第4、9、10章由吕国才、赵鲁涛编写，第6、7、8章由安璐、李晔编写，第11、12、13、14章由赵鲁涛、张志刚、安璐编写。研究生曾冠荣、汪兆洋也参与了部分章节的编写。赵鲁涛、张志刚负责本书的构思和统稿。

本书的编写得到了“十二五”期间北京科技大学教材建设经费的资助和北京高等学校教育教学改革立项项目（2015-ms028）的资助。在本书写作过程中，还得到了北京科技大学有关老师的热情帮助和大力支持，在此表示衷心的感谢！同时，还要感谢北京科技大学李安贵教授和中国科学院软件研究所吴文玲研究员对本书提出的宝贵建议。

由于作者的水平有限，书中错漏之处在所难免，敬请广大读者批评指正。

编 者

目 录

前 言

第 1 章 绪论 1

1.1 信息、数据与数据处理 1

1.2 数据管理技术的产生和发展 3

1.2.1 人工管理阶段 3

1.2.2 文件系统阶段 3

1.2.3 数据库系统阶段 4

1.3 数据库系统的组成 5

1.3.1 数据库 6

1.3.2 数据库管理系统 6

1.3.3 应用系统 7

1.3.4 数据库管理员 7

1.4 数据库系统结构 7

1.4.1 数据库系统模式的概念 8

1.4.2 数据库系统的三级模式结构 8

1.4.3 数据库的二级映像与数据 独立性 9

1.5 常见数据库管理系统 10

1.5.1 Oracle 10

1.5.2 SQL Server 10

1.5.3 Access 10

1.5.4 MySQL 11

1.6 小结 11

习题 12

第 2 章 数据模型 14

2.1 数据模型的概念、组成及类型 14

2.1.1 数据模型的概念 14

2.1.2 数据模型的组成要素 14

2.1.3 数据模型的类型 15

2.2 概念模型 16

2.2.1 概念模型概述 16

2.2.2 信息世界中的基本概念 16

2.2.3 实体联系方法 18

2.3 三大逻辑数据模型 20

2.3.1 层次模型 20

2.3.2 网状模型 21

2.3.3 关系模型 22

2.4 小结 23

习题 24

第 3 章 关系数据库 26

3.1 关系数据库系统概述 26

3.1.1 关系数据库系统的产生 26

3.1.2 关系数据模型概述 26

3.2 关系数据模型 27

3.2.1 关系数据模型的数据结构和基本 术语 27

3.2.2 关系的形式定义 29

3.2.3 关系的性质 30

3.3 关系的完整性 31

3.3.1 实体完整性 32

3.3.2 参照完整性 32

3.3.3 用户定义完整性 33

3.4 关系代数 34

3.4.1 关系代数概述 34

3.4.2 传统的集合运算 35

3.4.3 专门的关系运算 37

3.4.4 基本运算及变换 43

3.4.5 关系代数的一般书写方法 43

3.5 小结 44

习题 45

第 4 章 结构化查询语言 SQL 49

4.1 SQL 语言概述 49

4.2 电子商务客户端管理数据库 50

4.3 SQL 的数据定义功能 51

4.3.1 基本表的定义 51

4.3.2 修改表结构 54



4.3.3 删除表	54	5.4.3 第三范式 (3NF)	93
4.3.4 索引的定义	54	5.4.4 BC 范式 (BCNF)	93
4.3.5 索引的删除	55	5.5 多值依赖与 4NF	94
4.4 SQL 的数据查询功能	56	5.5.1 多值依赖	94
4.4.1 简单查询	57	5.5.2 第四范式——4NF	96
4.4.2 连接查询	61	5.6 关系模式的规范化	97
4.4.3 嵌套查询	62	5.6.1 规范化的含义	97
4.4.4 集合查询	65	5.6.2 关系规范化的过程	97
4.5 SQL 的数据操纵功能	65	5.6.3 关系规范化的要求	98
4.5.1 插入数据	65	5.7 小结	98
4.5.2 修改数据	67	习题	99
4.5.3 删除数据	67	第 6 章 数据库恢复技术	103
4.6 视图	68	6.1 事务	103
4.6.1 定义视图	68	6.1.1 事务的定义	103
4.6.2 查询视图	70	6.1.2 实例	103
4.6.3 更新视图	71	6.1.3 事务的特性	103
4.6.4 视图的作用	73	6.2 数据库恢复概述	104
4.7 SQL 的数据控制功能	74	6.3 故障的种类	105
4.7.1 授权	74	6.3.1 事务故障 (UNDO)	105
4.7.2 收回权限	76	6.3.2 系统故障 (UNDO + REDO)	105
4.8 嵌入式 SQL	76	6.3.3 介质故障	105
4.8.1 嵌入式 SQL 一般形式及相关规定	76	6.3.4 计算机病毒或恶意攻击	106
4.8.2 游标	77	6.4 恢复的实现技术	106
4.8.3 不用游标的 SQL 语句	78	6.4.1 数据转储	106
4.8.4 使用游标的 SQL 语句	78	6.4.2 登记日志文件 (Logging)	108
4.8.5 动态 SQL 简介	79	6.5 恢复策略	109
4.9 小结	80	6.5.1 事务故障的恢复	109
习题	81	6.5.2 系统故障的恢复	110
第 5 章 关系数据库的规范化理论	84	6.5.3 介质故障的恢复	111
5.1 规范化问题的提出	84	6.6 小结	111
5.1.1 关系模式的表示	84	习题	112
5.1.2 设计中常见的问题	84	第 7 章 并发控制	114
5.2 函数依赖	86	7.1 并发控制概述	115
5.2.1 函数依赖的定义	86	7.2 封锁	116
5.2.2 函数依赖的三种基本情形	87	7.3 封锁协议	117
5.2.3 码 (键)	87	7.4 活锁和死锁	119
5.3 函数依赖的公理系统	88	7.4.1 活锁	119
5.3.1 函数依赖集的完备性	88	7.4.2 死锁	119
5.3.2 函数依赖的推理规则	89	7.5 并发调度的可串行性	121
5.4 范式	90	7.6 两段锁协议	122
5.4.1 第一范式 (1NF)	91	7.7 封锁的粒度	124
5.4.2 第二范式 (2NF)	92	7.8 小结	124
		习题	126

第 8 章 数据库结构设计	129	10.5.1 云计算	155
8.1 数据库概念设计	129	10.5.2 云数据库	156
8.1.1 概念设计的任务	129	10.5.3 云数据库的发展趋势	157
8.1.2 概念设计的依据及过程	129	10.6 大数据下数据库的发展	157
8.1.3 概念设计实例	131	10.6.1 大数据时代的外部环境的 变化	157
8.2 数据库逻辑设计	133	10.6.2 大数据时代下数据库的发展趋势 与影响	158
8.2.1 概述	133	10.6.3 结语	159
8.2.2 逻辑设计实例	133	10.7 小结	160
8.3 小结	134	习题	161
习题	135	第 11 章 ADO.NET 入门教程	162
第 9 章 数据库安全	137	11.1 ADO.NET 的基本概念	162
9.1 计算机系统安全性概述	137	11.1.1 .NET	162
9.2 数据库安全性控制	137	11.1.2 ASP.NET	163
9.2.1 用户标识与鉴别	138	11.1.3 ADO.NET	164
9.2.2 存取控制	138	11.2 ADO.NET 的组件	165
9.3 SQL Server 的安全性	139	11.2.1 Connection	166
9.3.1 SQL Server 的安全机制	139	11.2.2 Command	167
9.3.2 SQL Server 的身份认证模式	139	11.2.3 DataReader	169
9.4 数据库存取控制方法	140	11.2.4 DataAdapter	170
9.4.1 用户权限	140	11.2.5 DataSet	171
9.4.2 授权与回收	140	11.3 ADO.NET 的执行过程	172
9.4.3 数据库角色	143	11.3.1 ADO.NET 的两种连接方式	172
9.5 其他数据库的安全性手段	144	11.3.2 ADO.NET 在连接环境中的执行 过程	173
9.5.1 视图机制	144	11.3.3 ADO.NET 在非连接环境中的执行 过程	174
9.5.2 审计	145	11.4 小结	175
9.6 小结	145	习题	176
习题	146	第 12 章 数据库应用系统访问技术	177
第 10 章 现代数据库技术及其发展	148	12.1 数据库访问技术	177
10.1 面向对象数据库	148	12.1.1 ODBC	177
10.1.1 面向对象数据库概述	148	12.1.2 OLE DB	177
10.1.2 面向对象数据模型	149	12.1.3 DAO	177
10.2 实时数据库	151	12.1.4 JDBC	178
10.2.1 实时数据库概述	151	12.2 Visual Studio 中的 Windows 窗体 程序	178
10.2.2 实时数据库的作用	151	12.2.1 在 Visual Studio 中建立 Windows 窗体	178
10.3 时空数据库	152	12.2.2 Windows 窗体中的数据控件及其 应用	180
10.3.1 时空数据库概述	152	12.3 Visual Studio 中的 WebForm 网页	
10.3.2 时空数据模型	153		
10.4 多媒体数据库	154		
10.4.1 多媒体数据库概述	154		
10.4.2 多媒体数据模型	154		
10.4.3 多媒体数据库管理系统	155		
10.5 云计算下数据库的发展	155		



程序	184	13.5 C/S 架构举例	197
12.3.1 在 Visual Studio 中建立 WebForm 应用程序	184	13.5.1 连接数据库的控件的配置与数据库 连接字的设置	198
12.3.2 ASP.NET 服务器控件及其 应用	185	13.5.2 商品管理模块	199
12.4 主流 Web 开发技术	189	13.6 B/S 架构举例	207
12.4.1 ASP.NET	189	13.6.1 页面架构	207
12.4.2 JSP	189	13.6.2 IIS 部署	214
12.4.3 SSH	189	13.7 小结	220
12.4.4 Spring MVC	189	习题	221
12.4.5 ASP.NET MVC	190	第 14 章 上机实验	222
12.5 小结	190	14.1 SQL Server2016 安装	222
习题	191	14.2 SQL Server 认识实验	234
第 13 章 数据库应用系统开发实例	192	14.3 DBMS 管理表数据	240
13.1 系统概述	192	14.4 完整性约束	242
13.2 需求分析	192	14.5 结构化查询	244
13.2.1 用户模块	193	14.6 数据备份和恢复	245
13.2.2 商品-购物模块	193	14.7 数据库应用系统开发（概念设计和 逻辑设计）	246
13.2.3 商品-管理模块	193	14.8 数据库应用系统开发（实际开发与 编码实现）	247
13.2.4 购物车模块	193	参考文献	249
13.2.5 订单模块	194		
13.3 概念模型	194		
13.4 数据表	195		



第1章

绪 论

数据库技术是计算机技术的重要分支，是数据管理的实用技术。本章介绍数据库系统的基本概念，包括数据管理的发展过程、数据库系统的组成等。读者从中可以学到为什么要使用数据库技术，以及了解数据库技术的重要性。本章是后面各章节的准备和基础。

数据库技术是一门研究、存储、管理和应用数据库的科学，是现代信息科学与技术的重要组成部分，也是计算机数据处理与信息管理的核心内容。在数据库技术的研究中，研究人员逐步解决了计算机信息处理过程中大量数据组织和存储的有效性问题，并在数据库系统中减少了数据存储冗余的出现，实现了数据的共享，保障了数据的安全性以及高效地检索和处理数据。在这里，数据库技术的根本目标是要解决数据的高效存储与高可靠性的数据共享问题。

数据库技术研究和管理的对象是数据，具体来说就是要建立一套可靠、高效、安全的数据存储机制来对数据进行高效存储以及快速查询与变更。数据库技术所涉及的具体内容主要包括：通过对数据的统一组织和管理，按照指定的结构建立相应的数据库和数据仓库；利用数据库管理系统和数据挖掘系统设计出能够实现对数据库中的数据进行添加、修改、删除、处理、分析、理解、报表和打印等多种功能的数据管理和数据挖掘应用系统；利用应用管理系统最终实现对数据的处理、分析和理解。

1.1 信息、数据与数据处理

在数据处理中，人们最常用到的基本概念就是数据和信息，信息与数据是两个不同的概念，不能一概而论，但是它们之间是互相联系、密不可分的。

1. 数据

数据是用来记录信息的可识别符号，是信息的具体表现形式。数据有多种表示形式，最常用的有两种：一种是用数量形式表示定量的数据，例如，张明的身高是1.75米，张明今年18岁；另一种是用文字形式表示定性的数据，例如，张明的性别是男。两种形式可以转换，即用数量形式表示定性的数据，用文字形式表示定量的数据。例如，用1表示男，0表示女；用高、中、低定义身高；用老、中、青定义年龄。

数据的概念在数据处理领域中已大大地拓宽了，其表现形式不仅包括数字和文字，还包括图形、图像、声音等。可用多种不同的数据形式表示同一信息，但其所携带的信息不随数据形式的不同而改变。

2. 信息

信息是关于现实世界中事物的存在方式和运动状态的综合反映，具体来说是一种被加工



为特定形式的数据，但这种数据形式对接收者来说是有意义的，而且对当前和将来的决策具有明显的实际价值。也就是说信息是具有一定含义的数据。如“2017 年汽车购置税率将调整为 7.5%”，对接收者有意义，他们可以据此做出决策。

信息源于物质和能量，它不可能脱离物质而单独存在，信息的传递需要物质载体，信息的获取和传递要消耗能量。如信息可以通过报纸、电台、电视、计算机网络进行传递。信息是可以感知的，人类对客观事物的感知，可以通过感觉器官，也可以通过各种仪器仪表和传感器等，不同的信息源有不同的感知形式。如报纸上刊登的信息通过视觉器官感知，电台中广播的信息通过听觉器官感知。信息是可存储、加工、传递和再生的。动物用大脑存储信息，叫作记忆。计算机存储器、录音、录像等技术的发展，进一步扩大了信息存储的范围。借助计算机，还可对收集到的信息进行整理。

3. 数据与信息的联系

数据是信息的符号表示，或称载体；信息是数据的内涵，是数据的语义解释。数据是现象，而信息更反映实质。信息只有借助数据符号的表示，才能被人们感知、理解和接受。例如有一段话“2016 年全国高考报名人数为 940 万，同比 2015 年的考生人数下降了 2 万人”。从这段话中，可以知道：“940 万”和“2 万”是数据，而 2016 年考生人数低于 2015 年的人数是其所携带的信息。

4. 数据处理

数据处理是将数据转换成信息的过程，包括对数据的收集、存储、加工、检索、传输等一系列活动。其目的是从大量的原始数据出发，根据事物之间的固有联系和运动规律，通过分析归纳、演绎推导等手段，萃取出对人们有价值、有意义的信息，作为决策的依据。可用下式简单的表示信息、数据与数据处理的关系：

$$\text{信息} = \text{数据} + \text{数据处理}$$

数据是原料，是输入，而信息是产出，是输出结果。“信息处理”的真正含义应该是为了产生信息而处理数据。

数据→加工处理→信息

人们将原始信息表示成数据，称为源数据，然后经过汇集、存储、综合、推导得到新的数据，这些新的数据称为结果数据，结果数据对某些特定的人们来说是有价值的、有意义的，表示了新的信息，可以作为某种决策的依据或用于新的推导。这一过程通常称为数据处理或信息处理。其处理活动的基本环节如图 1-1 所示。

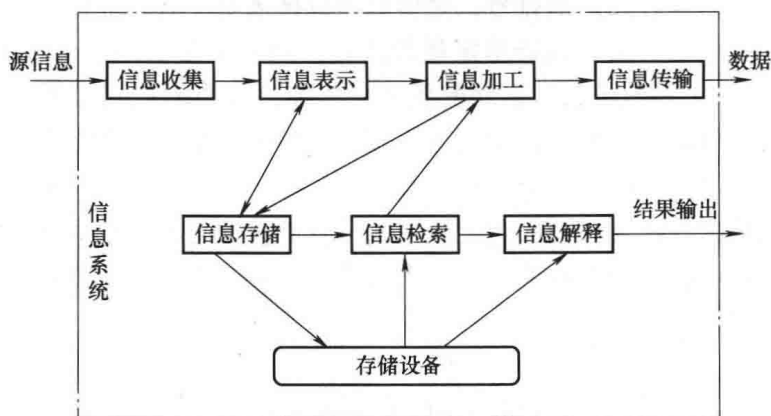


图 1-1 信息处理的基本环节

1.2 数据管理技术的产生和发展

数据处理的中心问题是数据管理。数据管理是指对数据的组织、分类、编码、存储、检索和维护等操作,这些操作是数据处理的基本环节,是任何数据处理业务中不可缺少的部分。数据管理的基本目的主要是为了提高数据的独立性、降低数据的冗余度、提高数据共享性、提高数据的安全性和完整性,从而能更加有效地管理和使用数据资源。

随着计算机硬件、软件技术的不断发展,数据管理技术经历了人工管理、文件系统和数据库系统三个发展阶段。

1.2.1 人工管理阶段

20世纪50年代中期以前,计算机主要用于科学计算。当时在硬件方面,外存储器只有卡片、纸带、磁带,没有像磁盘这样可以随机访问、直接存取的外部存储器设备。软件方面,只有汇编语言,没有操作系统和管理数据的软件。数据处理方式基本是批处理。

这一时期数据管理的特点是:

1. 数据不保存

因为当时计算机主要用于科学计算,对于数据保存的需求尚不迫切。

2. 系统没有专用的软件对数据进行管理

每个应用程序都要包括数据的存储结构、存取方法、输入方式等,程序员编写应用程序时,还要安排数据的物理存储,因此程序员负担很重。

3. 数据不共享

数据是面向程序的,一组数据只能对应一个程序。当多个应用程序涉及某些相同的数据时,也必须各自定义,无法互相利用、互相参照,因此程序之间有大量的冗余数据。

4. 数据不具有独立性

程序依赖于数据,如果数据的类型、格式或输入输出方式等逻辑结构或物理结构发生变化,必须对应用程序做出相应的修改。

在人工管理阶段,程序与数据之间的关系可用图1-2表示。

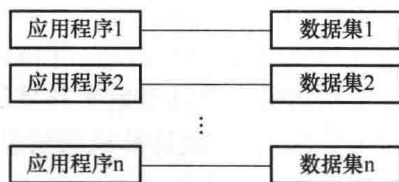


图 1-2 人工管理阶段

1.2.2 文件系统阶段

20世纪50年代后期至60年代中后期,计算机开始大量地用于管理中的数据处理工作。大量的数据存储、检索和维护成为紧迫的需求。硬件方面有了磁盘、磁鼓等直接存储设备。软件方面,出现了高级语言和操作系统。操作系统中有了专门管理数据的软件,一般称为文件系统。数据处理方式有批处理,也有联机实时处理。这一阶段有如下几个特点:

1. 数据以文件形式可长期保存下来

数据文件可以长期保存在外存储器上多次存取,用户可随时对文件进行查询、修改和增删等处理。

2. 文件系统可对数据的存取进行管理

文件系统使应用程序与数据之间有了初步的独立性，程序员只与文件名打交道，不必明确数据的物理存储，大大减轻了程序员的负担。

3. 文件形式多样化

有顺序文件、倒排文件、索引文件等，因而对文件的记录可顺序访问，也可随机访问，更便于存储和查找数据。

4. 程序与数据间有一定独立性

由专门的软件即文件系统进行数据管理，程序和数据间由软件提供的存取方法进行转换，数据存储发生变化不一定影响程序的运行。

在文件系统阶段，程序与数据之间的关系可用图 1-3 表示。

与人工管理阶段相比，文件系统阶段对数据的管理有了很大的进步，但一些根本性问题仍没有彻底解决，主要表现在以下三方面：

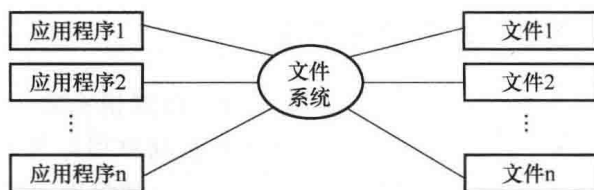


图 1-3 文件系统阶段

1. 数据冗余度大

各数据文件之间没有有机的联系，一个文件基本上对应于一个应用程序，数据不能共享。

2. 数据独立性低

数据和程序相互依赖，一旦改变数据的逻辑结构，必须修改相应的应用程序。而应用程序发生变化，如改用另一种程序设计语言来编写程序，也需修改数据结构。

3. 数据一致性差

由于相同数据的重复存储、各自管理，所以在进行更新操作时容易造成数据的不一致性。

1.2.3 数据库系统阶段

从 20 世纪 60 年代末开始，计算机应用于管理的规模更加庞大，需要计算机管理的数据量急剧增长，并且对数据共享的需求日益增强。这时，硬件方面出现了大容量磁盘（数百兆字节以上），使计算机联机存取大量数据成为可能；软件价格上升，硬件价格下降，使开发和维护系统软件的成本增加。文件系统的管理方法已无法适应开发应用系统的需要。为解决多用户、多个应用程序共享数据的需求，出现了统一管理数据的专门软件系统，即数据库管理系统（Database Management System, DBMS）。

数据库系统的主要特点是：

1. 数据共享性高、冗余少

这是数据库系统阶段的重大改进，数据不再面向某个应用程序而是面向整个系统，当前所有用户可同时存取库中的数据。这样便减少了不必要的数据冗余，节约存储空间，同时也避免了数据之间的不相容性与不一致性。

2. 数据结构化

按照某种数据模型，将全组织的各种数据组织到一个结构化的数据库中，整个组织的数据不是一盘散沙，可表示出数据之间的有机关联。例如，要建立学生成绩管理系统，系统包含学生（学号、姓名、性别、系别、年龄）、课程（课程号、课程名）、成绩（学号、课程

号、成绩)等数据,分别对应三个文件。若采用文件处理方式,因为文件系统只表示记录内部的联系,而不涉及不同文件记录之间的联系,要想查找某个学生的学号、姓名、所选课程的名称和成绩,必须编写一段程序来实现。而采用数据库方式,数据库系统不仅描述数据本身,还描述数据之间的联系,上述查询可以非常容易地联机查到。

3. 数据独立性高

数据的独立性是指逻辑独立性和物理独立性。

数据的逻辑独立性是指当数据的总体逻辑结构改变时,数据的局部逻辑结构不变,由于应用程序是依据数据的局部逻辑结构编写的,所以应用程序不必修改,从而保证了数据与程序之间的逻辑独立性。例如,在原有的记录类型之间增加新的联系,或在某些记录类型中增加新的数据项,均可确保数据的逻辑独立性。

数据的物理独立性是指当数据的存储结构改变时,数据的逻辑结构不变,从而应用程序也不必改变。例如,改变存储设备、增加新的存储设备或改变数据的存储组织方式,均可确保数据的物理独立性。

4. 有统一的数据控制功能

数据库为多个用户和应用程序所共享,对数据的存取往往是并发的,即多个用户可以同时存取数据库中的数据,甚至可以同时存取数据库中的同一个数据,为确保数据库数据的正确有效和数据库系统的有效运行,数据库管理系统提供数据控制功能。

数据库系统阶段,程序与数据之间的关系可用图 1-4 表示。

从文件系统管理发展到数据库系统管理是信息处理领域的一个重大变化。在文件系统阶段,人们关注的是系统功能的设计,因此程序设计处于主导地位,数据服从于程序设计;而在数据库系统阶段,数据的结构设计成为信息系统首先关心的问题。数据库技术经历了以上三个阶段的发展,已有了比较成熟的数据库技术,但随着计算机软硬件的发展,数据库技术仍需不断向前发展。

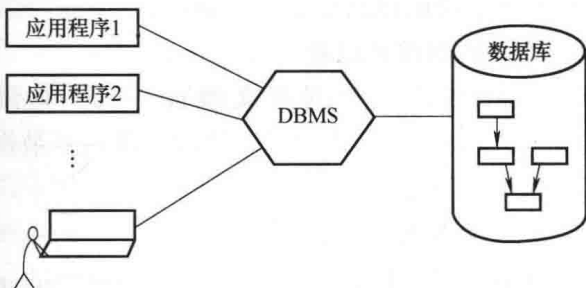


图 1-4 数据库系统阶段

1.3 数据库系统的组成

数据库系统是指在计算机系统中引入数据库后形成的集存储、管理和应用为一体的系统构成。数据库系统由五部分组成:以数据为主体的数据库;管理数据库的系统软件 DBMS;支持数据库系统的计算机硬件环境和操作系统环境;管理和使用数据库系统的人,特别是负责设计、维护数据库的技术人员——数据库管理员;方便使用和管理系统的各种技术说明书和使用说明书。简单来说,数据库是数据的存储仓库,和生活中用到的仓库类似,指在计算机中用于存储和管理数据的结构。例如在计算机中录入人的身高体重、学生的成绩、公司的业绩报表等,这些都可以存储在数据库中,并由数据库管理员统一管理。

数据库系统有大小之分,大型数据库系统有 SQL Server、Oracle、DB2 等,中小型数据库系统有 Foxpro、Access、MySQL。但是不论是大型数据库还是小型数据库,一套完整的数

数据库系统包括：数据库、数据库管理系统、应用系统、数据库管理员和说明文档 5 部分。

1.3.1 数据库

数据库 (DataBase, DB), 顾名思义, 是存放数据的仓库, 这个仓库不是一般意义上的库房, 而是存储在计算机存储设备上的相互关联的数据集合。

数据库是相互关联的数据的集合, 它用综合的方法组织数据, 具有较小的数据冗余, 可供多个用户共享, 具有较高的数据独立性, 具有安全控制机制, 能够保证数据的安全、可靠, 允许并发地使用数据库, 能有效、及时地处理数据, 并能保证数据的一致性和完整性。它是长期存储在计算机内有组织的、大量的、共享的数据集合。

数据库数据具有永久存储、有组织和可共享 3 个基本特点。

1.3.2 数据库管理系统

数据库管理系统 (DBMS) 是建立、管理和维护数据库的软件系统, 是位于用户与操作系统之间, 实现数据库管理功能的系统软件。

DBMS 在计算机体系结构中的层次如图 1-5 所示, 由图 1-5 可以看出, DBMS 在操作系统 (OS) 的支持下工作, 应用程序在 DBMS 支持下才能使用数据库。它的基本功能包括以下几个方面:

1. 数据定义功能

DBMS 提供数据定义语言 (Data Definition Language, DDL), 用户通过它可以方便地对数据库中的数据对象进行定义。

2. 数据操纵功能

DBMS 提供数据操纵语言 (Data Manipulation Language, DML), 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、删除和更新。

3. 数据库的建立和维护功能

它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。

4. 数据库的运行管理和控制功能

数据库为多个用户和应用程序所共享, 对数据的存取往往是并发的, 即多个用户可以同时存取数据库中的数据, 甚至可以同时存取数据库中的同一个数据, 为确保数据库数据的正确有效和数据库系统的有效运行, 数据库管理系统提供下述五方面的数据控制功能。

1) 数据的安全性 (Security) 控制: 防止不合法使用数据造成数据的泄露和破坏, 保证数据的安全和机密 (如系统提供口令检查或其他手段来验证用户身份, 防止非法用户使用系统); 也可以对数据的存取权限进行限制, 只有通过检查后才能执行相应的操作。

2) 数据的完整性 (Integrity) 控制: 系统通过设置一些完整性规则以确保数据的正确性、有效性和相容性。正确性是指数据的合法性, 如年龄属于数值型数据, 只能含有 0, 1, ..., 9, 不能含有字母或特殊符号; 有效性是指数据是否在其定义的有效范围, 如月份只

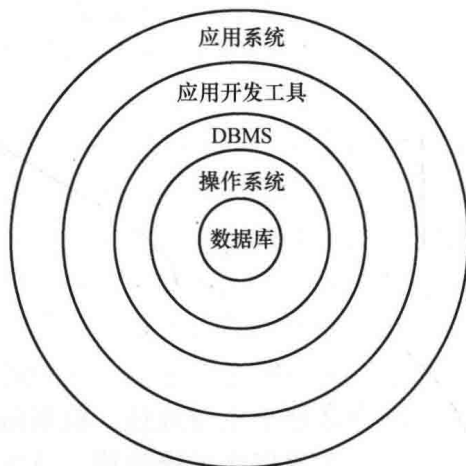


图 1-5 应用程序、DBMS、OS 及 DB 之间的关系

能用1~12之间的正整数表示；相容性是指表示同一事实的两个数据应相同，否则就不相容，如一个人不能有两个性别。

3) 并发 (Concurrency) 控制：多用户同时存取或修改数据库时，防止因相互干扰而提供给用户不正确的数据，并使数据库受到破坏。例如，网上并发订票操作、并发选课操作等。

4) 数据恢复 (Recovery)：当数据库被破坏或数据不可靠时，系统有能力将数据库从错误状态恢复到最近某一时刻的正确状态。

5) 事务 (Transation) 支持：事务是一个逻辑单元，它包括对数据库的一系列操作（特别是一些更新操作）。

1.3.3 应用系统

应用系统一般由计算机硬件系统、系统软件和应用软件组成。计算机基本硬件系统由运算器和控制器、存储器、外围接口和外围设备组成。硬件的配置应满足整个数据库系统的需要。系统软件包括操作系统、编译程序、数据库管理系统、各种高级语言等。应用软件由用户程序和各種应用软件包组成。

1.3.4 数据库管理员

数据库管理员 (Database Administrator, DBA) 是数据库的核心角色，负责设计、建立、管理和维护整个数据库，使数据能够被任何有权使用的人有效使用。

DBA 应始终参加整个数据库系统的研发工作，之后全面负责数据库系统的管理、维护和正常使用。DBA 对大型数据库系统应具有较强的管理能力和丰富的管理经验，其主要职责为：

- 1) 参与数据库设计的全过程，决定数据库的结构和内容。
- 2) 决定数据库的存储结构和存取策略，以获得较高的存取效率和存储空间利用率。
- 3) 帮助终端用户使用数据库系统，如培训终端用户、解答终端用户日常使用数据库系统时遇到的问题等。
- 4) 控制和监控用户对数据库的存取访问，维护数据库的安全性。
- 5) 监督控制数据库的使用和运行，负责定义和实施适当的数据库备份和恢复策略。当数据库受到破坏时，在最短时间内将数据库恢复到正确状态，并尽可能地不影响或少影响计算机系统其他部分的正常运行。

6) 改进和重组重构数据库，负责监视和分析系统的性能，使系统的空间利用率和处理效率总是处于较高水平。当系统性能降低时，根据实际情况不断改进数据库的设计，不断提高系统的性能；当用户需求情况发生变化时，对数据库进行重新构造。

1.4 数据库系统结构

考察数据库系统的结构可以有多种不同的层次或不同的角度。

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：外模式、模式和内模式，这是数据库管理系统内部的体系结构。

从数据库最终用户角度看，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（包括二层、三层和多层结构）等。这是数据库系统外部的体系结构。

本节介绍数据库系统的三级模式结构。

1.4.1 数据库系统模式的概念

型是指对某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，学生记录定义为（学号，姓名，性别，系别，年龄，籍贯）称为记录型，而（900201，李明，男，计算机，22，江苏）是该记录型的一个记录值。

模式是数据库中所有数据的逻辑结构和特征的描述，它仅仅涉及模型的描述，不涉及具体的值。模式的一个具体值称为一个实例。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。

1.4.2 数据库系统的三级模式结构

为了有效地组织、管理数据，提高数据库的逻辑独立性和物理独立性，人们为数据库设计了一个严谨的体系结构，数据库领域公认的标准结构是三级模式结构，即外模式、模式、内模式结构。用户级对应外模式，概念级对应模式，物理级对应内模式，使不同级别的用户对数据库形成不同的视图。所谓视图，就是指观察、认识和理解数据的范围、角度和方法，是数据库在用户“眼中”的反映，很显然，不同层次（级别）用户所“看到”的数据库是不相同的。数据库的三级模式结构如图 1-6 所示。

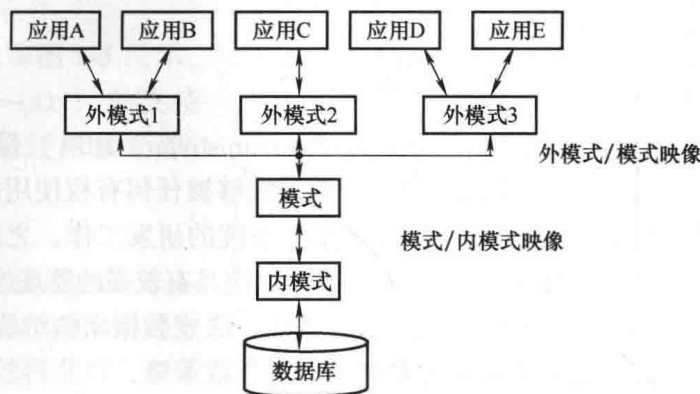


图 1-6 数据库的三级模式结构

1. 模式

模式又称概念模式或逻辑模式，对应于概念级。它是由数据库设计者综合所有用户的数据，按照统一的观点构造的全局逻辑结构，是对数据库中全部数据的逻辑结构和特征的总体描述，是所有用户的公共数据视图（全局视图）。它是由数据库管理系统提供的数据模式描述语言（Data Description Language，DDL）来描述和定义的，体现、反映了数据库系统的整体观。

2. 外模式

外模式又称子模式或用户模式，对应于用户级。它是数据库用户（包括应用程序员和最终用户）使用的局部数据的逻辑结构和特征的描述，是与某一应用有关的数据的逻辑表示。

外模式通常是模式的子集，一个数据库可以有多个外部模式，反映了不同用户的应用需求、看待数据的方式、对数据保密的要求。对模式中同一数据，在外部模式中的结构、类型、长度、保密级别等都可以不同。同一外部模式也可以为某一用户的多个应用系统所使用，但一个应用程序只能使用一个外部模式。

用户可以通过外模式描述语言来描述、定义对应于用户的数据记录（外模式），也可以