

安徽省高等学校“十一五”省级规划教材

SHUJUKU XITONG JI YINGYONG

# 数据库系统及应用

主 编 戴小平  
副主编 王 丽 帅 兵 张润梅

中国科学技术大学出版社

安徽省高等学校“十一五”省级规划教材

# 数据库系统及应用

SHUJUKU XITONG JI  
YINGYONG

主 编 戴小平

副主编 王 丽 帅 兵 张润梅

中国科学技术大学出版社

## 内 容 简 介

本书是安徽省高等学校“十一五”省级规划教材。

全书较全面地介绍了数据库系统的基本原理、设计和应用技术。内容包括数据库基础知识、关系数据模型、关系数据库语言 SQL、关系数据库理论、关系数据库设计和应用系统开发、数据库的安全性与完整性、并发控制、数据库故障恢复技术和数据库新技术。

本书以学习数据库理论基础、培养数据库应用开发能力为目标,以大型数据库系统 Oracle 为实例贯穿全书。在重视学习数据库基本原理的基础上,突出了实用技术的学习,各章都备有适量的例题和习题。

本书既可以作为高等院校计算机、软件工程、信息管理与信息系统等工科类相关专业数据库课程的教材,也可供从事计算机软件以及数据库应用、管理和开发的工程技术人员阅读参考。

## 图书在版编目(CIP)数据

数据库系统及应用/戴小平主编. —合肥:中国科学技术大学出版社,2010.8

ISBN 978-7-312-02606-5

I. 数… II. 戴… III. 数据库系统 IV. TP311.13

中国版本图书馆 CIP 数据核字 (2010) 第 096397 号

**出版** 中国科学技术大学出版社

地址:安徽省合肥市金寨路 96 号,邮政编码:230026

网址: <http://press.ustc.edu.cn>

**印刷**

**发行** 中国科学技术大学出版社

**经销** 全国新华书店

**开本** 787 mm×1092 mm 1/16

**印张** 20.5

**字数** 512 千

**版次** 2010 年 8 月第 1 版

**印次** 2010 年 8 月第 1 次印刷

**印数** 1—4000 册

**定价** 35.00 元

# 前 言

数据库技术始于 20 世纪 60 年代,经过四十多年的发展,数据库已经与操作系统、通信网络、应用服务器一起成为 IT 基础设施的重要组成部分,工农业生产、银行、电信、商业、行政管理、科学研究、教育、国防军事等几乎每个行业都广泛应用数据库系统来管理和处理数据。可以说数据库技术和数据库系统已经成为计算机信息系统的核心技术和重要基础,围绕着数据库技术形成了一个巨大的软件产业。

目前,数据库技术已成为计算机领域内一个重要部分。关于数据库系统的课程已成为计算机科学与技术、信息管理与工程、软件工程等专业的核心课程,也是许多其他专业的重要选修课程。

本书共分 13 章。第 1 章主要介绍数据库基础知识,包括数据库概念、三层模式和数据库管理系统等内容;第 2 章介绍关系数据库,包括关系模型和关系代数;第 3 章主要介绍 Oracle 数据库基础及 Oracle 数据库体系结构;第 4 章与第 5 章分别介绍关系数据语言 SQL 和 Oracle 数据库的存储过程和触发器;第 6 章介绍关系数据理论,包括函数依赖、公理系统、规范化和模式分解等内容;第 7 章和第 8 章分别介绍数据库设计的基本方法和数据库应用系统开发的基本知识;第 9 章到第 12 章介绍数据库管理系统的统一数据控制功能的概念与知识,分别为数据库安全性、数据库故障与恢复技术、并发控制和数据库完整性;第 13 章介绍数据库的一些新的应用和研究领域,包括分布式数据库、面向对象数据库、数据仓库和数据挖掘技术等。

Oracle 作为一种大型数据库,是当前应用最广泛的数据库系统之一,因此本书选择 Oracle 9i 数据库作为实例贯穿其中。希望读者在学习数据库基础理论知识的同时,能够了解一种数据库产品,学会和掌握一种数据库环境的基本操作,并结合课程设计,能够进行数据库应用系统的开发。

本书由安徽工业大学、安徽理工大学、安徽建筑工业学院和安徽工程大学联合编写。第 1、9 章由安徽工业大学周兵编写,第 7、8 章由安徽工业大学陈业斌编写,第 10、11、12 章由安徽工业大学戴小平编写,第 2、6 章由安徽建筑工业学院张润梅编写,第 3、5 章由安徽理工大学王丽编写,第 4、13 章由安徽工程大学帅兵编写。全书由安徽工业大学戴小平统稿。

安徽工业大学方木云、张学锋,安徽理工大学吴观茂、朱晓娟、周花平,安徽工程大学伊芸芸、周文、汪婧,安徽建筑工业学院王坤等对本书提出了很多意见和建议,特此表示感谢。

在本书的编写过程中,所有参编者都力求跟踪数据库学科最新的技术水平和发展方向,引入新的技术和方法。但由于水平有限,书中疏漏之处在所难免,殷切希望同行专家和读者批评指正。

编 者

2010 年 3 月

## 目 录

前言	( I )
<b>第 1 章 数据库基础</b>	( 1 )
1.1 数据、信息与数据处理	( 1 )
1.2 数据管理技术的发展历史	( 2 )
1.2.1 手工管理阶段	( 2 )
1.2.2 文件系统阶段	( 3 )
1.2.3 数据库系统阶段	( 4 )
1.3 数据库概念	( 4 )
1.4 数据模型	( 7 )
1.4.1 3 个世界及其相互关系	( 7 )
1.4.2 概念模型	( 8 )
1.4.3 数据模型	( 12 )
1.4.4 层次模型	( 12 )
1.4.5 网状模型	( 14 )
1.4.6 关系模型	( 15 )
1.5 数据库体系结构	( 16 )
1.5.1 数据库体系结构中的三级模式	( 16 )
1.5.2 数据库体系结构中的二级映射与数据独立性	( 17 )
1.6 数据库管理系统(DBMS)	( 17 )
1.6.1 数据库管理系统的目标	( 17 )
1.6.2 数据库管理系统的基本功能	( 18 )
1.7 数据库系统(DBS)	( 19 )
1.7.1 数据库系统的组成	( 19 )
1.7.2 数据库系统的分类	( 21 )
本章小结	( 22 )
习题	( 22 )
<b>第 2 章 关系数据库</b>	( 23 )
2.1 关系模型的基本概念	( 23 )
2.1.1 基本术语	( 23 )
2.1.2 关系( Relation)	( 25 )

2.1.3	关系模式 .....	(25)
2.1.4	关系数据库 .....	(26)
2.2	关系的完整性 .....	(26)
2.2.1	实体完整性(Entity Integrity) .....	(26)
2.2.2	参照完整性(Referential Integrity) .....	(27)
2.2.3	用户定义完整性(User-defined Integrity) .....	(27)
2.3	关系数据语言概述 .....	(28)
2.3.1	关系操作的基本内容 .....	(28)
2.3.2	关系数据语言的特点 .....	(28)
2.3.3	关系数据语言的分类 .....	(28)
2.4	关系代数 .....	(29)
2.4.1	传统的集合操作 .....	(30)
2.4.2	扩充的关系操作 .....	(31)
2.4.3	关系代数运算的应用实例 .....	(34)
2.5	关系演算及其查询优化 .....	(36)
2.5.1	元组关系演算语言 ALPHA .....	(36)
2.5.2	元组关系演算 .....	(39)
2.5.3	域关系演算语言 QBE .....	(40)
2.5.4	关系系统及其查询优化 .....	(44)
	本章小结 .....	(48)
	习题 .....	(48)
<b>第3章</b>	<b>Oracle 数据库</b> .....	(51)
3.1	Oracle 数据库基础 .....	(51)
3.1.1	Oracle 简介 .....	(51)
3.1.2	Oracle 9i 产品结构及组成 .....	(51)
3.1.3	Oracle 9i 数据库特点 .....	(52)
3.2	Oracle 数据库的体系结构 .....	(52)
3.2.1	Oracle 数据库的逻辑结构 .....	(53)
3.2.2	Oracle 数据库的物理结构 .....	(55)
3.2.3	Oracle 实例 .....	(56)
3.2.4	Oracle 实例的内存结构 .....	(57)
3.2.5	Oracle 实例的进程结构 .....	(58)
3.3	Oracle 数据库的使用 .....	(59)
3.3.1	Oracle 9i 的安装 .....	(59)
3.3.2	数据库的启动与关闭 .....	(68)
3.3.3	数据库的创建与管理 .....	(69)
3.3.4	Oracle 的卸载 .....	(73)
3.4	SQL * Plus 初步操作 .....	(74)

3.4.1	SQL * Plus 的登录与退出	(74)
3.4.2	SQL * Plus 命令	(75)
	本章小结	(77)
	习题	(77)
<b>第4章</b>	<b>关系数据库标准语言 SQL</b>	<b>(78)</b>
4.1	SQL 语言概述	(78)
4.1.1	SQL 语言的特点	(79)
4.1.2	SQL 数据库的体系结构	(79)
4.1.3	SQL 语言的组成	(80)
4.2	SQL 的数据定义	(81)
4.2.1	SQL 的数据定义语句	(81)
4.2.2	SQL 语言的基本数据类型	(82)
4.2.3	基本表的创建、修改和撤销	(84)
4.2.4	索引的创建和撤销	(88)
4.3	SQL 的数据查询	(89)
4.3.1	SELECT 语句的基本格式	(90)
4.3.2	嵌套查询	(98)
4.3.3	多个 SELECT 语句的集合操作	(101)
4.4	SQL 的数据操作	(101)
4.4.1	插入数据	(102)
4.4.2	修改数据	(103)
4.4.3	删除数据	(103)
4.4.4	更新操作与数据库的一致性	(104)
4.5	视图	(104)
4.5.1	定义视图	(105)
4.5.2	撤销视图	(106)
4.5.3	视图的查询	(106)
4.5.4	视图的更新	(107)
4.5.5	视图的作用	(108)
4.6	SQL 的数据控制	(108)
4.6.1	授予权限语句 GRANT	(109)
4.6.2	撤销权限语句 REVOKE	(110)
4.7	SQL 的事务处理	(111)
4.7.1	事务的概念	(111)
4.7.2	事务的特性	(115)
4.7.3	SQL 对事务的支持	(116)
4.8	嵌入式 SQL 的应用	(116)
4.8.1	区分 SQL 语句与宿主语言语句	(117)

4.8.2 嵌入式 SQL 与宿主语言间的信息传递 .....	(117)
4.8.3 游标 .....	(119)
本章小结 .....	(121)
习题 .....	(122)
<b>第 5 章 Oracle 存储过程与触发器 .....</b>	<b>(124)</b>
5.1 基本概念 .....	(124)
5.1.1 PL/SQL 程序块 .....	(124)
5.1.2 PL/SQL 的变量、常量与字符集 .....	(126)
5.1.3 PL/SQL 的控制语句 .....	(128)
5.1.4 PL/SQL 中的异常 .....	(133)
5.2 Oracle 存储过程 .....	(135)
5.2.1 存储过程基本知识 .....	(135)
5.2.2 存储过程的相关操作 .....	(136)
5.2.3 存储过程示例 .....	(137)
5.2.4 包 .....	(139)
5.3 Oracle 触发器 .....	(141)
5.3.1 触发器基本知识 .....	(141)
5.3.2 触发器相关操作 .....	(142)
5.3.3 触发器实例 .....	(144)
本章小结 .....	(146)
习题 .....	(147)
<b>第 6 章 关系数据理论 .....</b>	<b>(148)</b>
6.1 基本概念 .....	(148)
6.1.1 函数依赖 .....	(148)
6.1.2 完全函数依赖 .....	(149)
6.1.3 传递函数依赖 .....	(149)
6.1.4 码 .....	(150)
6.2 函数依赖的公理系统 .....	(150)
6.2.1 函数依赖的逻辑蕴含 .....	(150)
6.2.2 Armstrong 公理系统 .....	(150)
6.2.3 函数依赖集闭包和属性依赖集闭包 .....	(151)
6.2.4 Armstrong 公理的有效性和完备性 .....	(152)
6.2.5 函数依赖集的等价和覆盖 .....	(153)
6.2.6 函数依赖集的最小化 .....	(153)
6.3 关系模式的规范化 .....	(155)
6.3.1 范式(Normal Form) .....	(155)
6.3.2 多值依赖与第四范式(4NF) .....	(159)



6.3.3 关系模式的规范化	(162)
6.4 模式分解	(163)
6.4.1 函数依赖集的投影	(163)
6.4.2 模式分解	(163)
6.4.3 无损连接分解	(164)
6.4.4 保持函数依赖的分解	(165)
6.4.5 模式分解算法	(165)
本章小结	(167)
习题	(167)
<b>第7章 数据库设计</b>	<b>(169)</b>
7.1 数据库设计概述	(169)
7.1.1 数据库设计的任务	(169)
7.1.2 数据库设计的内容	(170)
7.1.3 数据库设计的特点	(170)
7.1.4 数据库设计方法	(170)
7.1.5 数据库设计的步骤	(172)
7.2 需求分析	(174)
7.2.1 需求分析的任务	(174)
7.2.2 需求分析的方法	(175)
7.2.3 需求分析注意点	(178)
7.3 概念结构设计	(179)
7.3.1 概念结构设计的必要性	(179)
7.3.2 概念结构设计的方法与步骤	(180)
7.3.3 数据抽象与局部视图设计	(183)
7.4 逻辑结构设计	(190)
7.4.1 逻辑结构设计的任务和步骤	(190)
7.4.2 初始关系模式设计	(191)
7.4.3 关系模式规范化	(192)
7.4.4 模式评价与改进	(193)
7.5 数据库物理设计	(194)
7.5.1 确定物理结构	(195)
7.5.2 评价物理结构	(196)
7.6 数据库的实施	(197)
7.6.1 建立实际数据库结构	(197)
7.6.2 装入数据	(197)
7.6.3 应用程序编码与调试	(198)
7.6.4 数据库试运行	(198)
7.6.5 整理文档	(199)

7.7 数据库的运行和维护	(199)
7.7.1 维护数据库的安全性与完整性	(199)
7.7.2 监测并改善数据库性能	(200)
7.7.3 重新组织和构造数据库	(200)
本章小结	(200)
习题	(201)
<b>第8章 数据库应用系统开发</b>	<b>(203)</b>
8.1 数据库应用结构	(203)
8.1.1 基于客户机/服务器模式的数据库系统	(203)
8.1.2 基于浏览器/服务器模式的数据库系统	(206)
8.1.3 B/S模式与C/S模式比较	(207)
8.2 数据库访问接口	(208)
8.2.1 ODBC	(208)
8.2.2 OLE DB	(212)
8.2.3 ADO和ADO.NET	(213)
8.2.4 JDBC	(214)
8.3 数据库应用开发	(215)
8.3.1 数据库应用环境配置	(216)
8.3.2 数据库应用编程的步骤	(216)
8.3.3 建立连接	(216)
8.3.4 操作数据库	(218)
8.3.5 处理结果集	(225)
本章小结	(228)
习题	(228)
<b>第9章 数据库安全性</b>	<b>(230)</b>
9.1 数据库安全概述	(230)
9.2 计算机系统及数据库系统安全标准简介	(230)
9.2.1 TCSEC/TDI标准	(231)
9.2.2 CC标准	(233)
9.3 数据库安全控制技术	(234)
9.3.1 用户标识与鉴别	(234)
9.3.2 存取控制	(234)
9.3.3 数据库的视图机制	(235)
9.3.4 数据加密	(236)
9.3.5 数据库的审计	(236)
9.4 自主存取控制的SQL命令	(237)
9.4.1 权限	(237)

9.4.2 授权与收回授权	(238)
9.4.3 数据库角色	(238)
9.5 数据库安全性实例——Oracle 系统	(239)
9.5.1 用户管理	(239)
9.5.2 权限与角色管理	(240)
9.5.3 Oracle 审计	(241)
本章小结	(243)
习题	(243)
<b>第 10 章 数据库恢复技术</b>	<b>(245)</b>
10.1 数据库故障	(245)
10.2 恢复的实现技术	(247)
10.2.1 数据转储	(248)
10.2.2 日志文件	(249)
10.2.3 归档日志文件	(250)
10.3 恢复策略	(250)
10.3.1 事务故障的恢复	(250)
10.3.2 系统故障的恢复	(251)
10.3.3 介质故障的恢复	(251)
10.4 具有检查点的恢复技术	(252)
10.5 冗余磁盘阵列与数据库镜像	(253)
10.6 Oracle 备份与恢复技术	(254)
10.6.1 物理备份与恢复	(254)
10.6.2 逻辑备份与恢复	(255)
本章小结	(259)
习题	(260)
<b>第 11 章 并发控制</b>	<b>(261)</b>
11.1 并发控制概述	(261)
11.1.1 丢失修改(Lost Update)	(261)
11.1.2 不可重复读(Non-Repeatable Read)	(262)
11.1.3 读“脏”数据(Dirty Read)	(262)
11.2 封锁	(263)
11.2.1 封锁机制	(263)
11.2.2 封锁协议	(264)
11.3 死锁	(266)
11.3.1 产生死锁的原因	(266)
11.3.2 死锁的预防	(266)
11.3.3 死锁的诊断与解除	(267)
11.4 并发调度和可串行性	(268)

11.5 两段锁协议 .....	(269)
11.6 封锁的粒度 .....	(271)
11.6.1 封锁的粒度 .....	(271)
11.6.2 意向锁 .....	(271)
11.7 Oracle 的并发控制 .....	(272)
本章小结 .....	(273)
习题 .....	(274)
<b>第 12 章 数据库完整性</b> .....	(275)
12.1 完整性子系统 .....	(275)
12.2 完整性约束语义的定义与检查 .....	(276)
12.2.1 域完整性约束 .....	(277)
12.2.2 实体完整性约束 .....	(277)
12.2.3 参照完整性约束 .....	(277)
12.3 Oracle 系统的完整性 .....	(280)
12.3.1 实体完整性 .....	(280)
12.3.2 参照完整性 .....	(281)
12.3.3 用户定义的完整性 .....	(281)
12.3.4 通过触发器定义用户的完整性规则 .....	(283)
12.4 完整性约束命名 .....	(283)
12.4.1 CONSTRAINT 约束命名 .....	(283)
12.4.2 修改表中的完整性约束条件的定义 .....	(284)
本章小结 .....	(284)
习题 .....	(285)
<b>第 13 章 高级数据库技术</b> .....	(286)
13.1 分布式数据库系统 .....	(286)
13.1.1 分布式数据库系统概述 .....	(287)
13.1.2 分布式数据库设计 .....	(288)
13.1.3 分布式数据库系统的主要技术问题 .....	(291)
13.2 面向对象数据库系统 .....	(296)
13.2.1 面向对象基本概念 .....	(296)
13.2.2 面向对象技术与数据库技术相结合的途径 .....	(298)
13.2.3 对象—关系数据库系统 .....	(298)
13.3 数据库的新技术及新应用 .....	(300)
13.3.1 数据库系统的发展 .....	(300)
13.3.2 数据库技术与其他相关技术相结合 .....	(301)
本章小结 .....	(312)
习题 .....	(312)
<b>参考文献</b> .....	(313)

# 第 1 章 数据库基础

我们现在处在信息时代,信息量成爆炸趋势增长,处理信息的技术更是飞速发展,而处理信息的一项核心技术就是数据库技术。因此,对于一个国家来说,数据库的规模、数据库信息量的大小和使用频度已成为衡量这个国家信息化程度的重要标志之一。

## 1.1 数据、信息与数据处理

在我们的学习、生活和工作中,经常会接触到各式各样的信息。如文字、数字、声音、图形、图像等,可以说现实世界到处都充满了信息。

信息是现实世界中各种事物(包括有生命的和无生命的、有形的和无形的)的存在方式、运动形态以及它们之间的相互联系等诸多要素在人脑中的反映,是通过人脑的抽象后形成的概念。这些概念不仅被人们认识和理解,而且人们还可以对它们进行推理、加工和传播。

数据一般是指信息的一种符号化表示方法,就是说用一定的符号表示信息,而采用什么符号,完全是人为规定。例如,为了便于用计算机处理信息,就得把信息转换为计算机能够识别的符号。即采用 0 和 1 两个符号编码来表示各种各样的信息。所以数据的概念包括两方面的含义:一是数据的内容是信息,二是数据的表现形式是符号。

数据在数据处理领域中涵盖的内容非常广泛,这里的“符号”不仅仅是指数字、字母、文字等常见符号,还包括图形、图像、声音等多媒体数据。

信息和数据的关系是既有联系又有区别。数据是承载信息的物理符号或称之为载体,而信息是数据的内涵。二者的区别是:数据可以表示信息,但不是任何数据都能表示信息,同一数据也可以有不同的解释。正如人们常说的“如果给计算机输入的是垃圾,输出的也将是垃圾”。信息是抽象的,同一信息可以有不同的数据表示方式。例如,在足球世界杯期间,同一场比赛的新闻,可以分别在报纸上以文字形式、在电台以声音形式、在电视上以图像形式来表现。

数据处理是指将数据转换成信息的过程,这一过程主要是指对所输入的数据进行加工整理,包括对数据的收集、存储、加工、分类、检索和传播等一系列活动,其根本目的是从大量的、已知的数据出发,根据事物之间的固有联系和变化规律,采用分析、推理、归纳等手段,提取出对人们有价值、有意义的信息,作为某种决策的依据。

我们可以用如图 1.1 所示的过程简单地表示出信息与数据之间的关系。

图 1.1 中数据是输入,而信息是输出结果。人们所说的“信息处理”,其真正含义应该是为了产生信息而处理数据。例如,学生的“出生日期”是有生以来不可改变的基本特征之一,属于

原始数据,而“年龄”是当年与出生日期相减而得到的数字,具有相对性,可视为二次数据。同样道理,“参加工作时间”、产品的“购置日期”是职工和产品的原始数据,工龄、产品的报废日期则是经过简单计算所得到的结果。

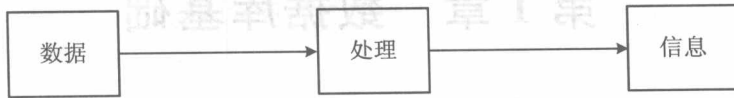


图 1.1 信息与数据之间的关系

在数据处理活动中,计算过程相对比较简单,很少涉及复杂的数学模型,但是却有数据量大,且数据之间有着复杂的逻辑联系的特点。因此数据处理任务的矛盾焦点不是计算,而是如何把数据管理好。数据管理是指数据的收集、整理、编目、组织、存储、查询、维护和传送等各种操作,是数据处理的基本环节,是任何数据处理任务必有共有的共性部分。因此,对数据管理应当加以突出,集中精力开发出通用而又方便实用的软件,把数据有效地管理起来,以便最大限度地减轻计算机软件用户的负担。数据库技术正是瞄准这一目标而逐渐完善起来的一门计算机软件技术。

## 1.2 数据管理技术的发展历史

利用计算机进行数据管理经历了从低级到高级的发展过程,这一过程大致分为 3 个阶段:

手工管理阶段 → 文件系统阶段 → 数据库系统阶段

### 1.2.1 手工管理阶段

20 世纪 50 年代中期以前,计算机主要用于科学计算。硬件存储设备主要有磁鼓、卡片机、纸带机等,还没有磁盘等直接存取的存储设备;软件上也处于初级阶段,没有操作系统和数据管理工具。数据的组织和管理完全靠程序员手工完成,因此称为手工管理阶段。这个阶段数据的管理效率很低。

在人工管理阶段,程序与数据之间的一一对应关系可用图 1.2 表示。

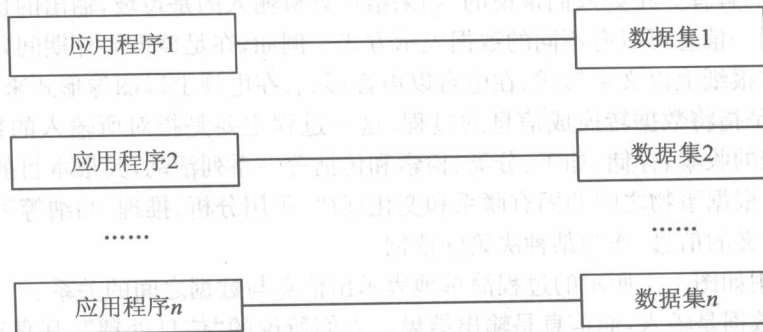


图 1.2 人工管理阶段应用程序与数据之间的对应关系

人工管理数据具有如下特点：

(1) 数据不保存

由于当时计算机主要用于科学计算，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，随着计算任务的完成，用户作业退出计算机系统，数据空间随着程序空间一起被释放。

(2) 应用程序管理数据

数据需要由应用程序自己管理，没有相应的软件系统负责数据的管理工作。应用程序负责所有数据的存取，因此程序员负担很重。

(3) 数据不共享

数据是面向应用的，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，由于必须各自定义，无法互相利用、互相参照，因此程序与程序之间有大量的冗余数据，容易产生数据的不一致性。

(4) 数据不具有独立性

数据的存储结构发生变化后，必须对应用程序做相应的修改，这就进一步加重了程序员的负担，导致软件维护成本增加。

## 1.2.2 文件系统阶段

20世纪50年代中期以后，计算机的硬件和软件得到快速发展，计算机不再仅用于科学计算的单一任务，还可以做一些非数值数据的处理。又由于大容量磁盘等辅助存储设备的出现，使得专门管理辅助存储设备上的数据的文件系统应运而生，它是操作系统中的一个子系统。在文件系统中，按一定的规则将数据组织成为一个文件，应用程序通过文件系统对文件中的数据进行存取和加工。文件系统对数据的管理，实际上是通过应用程序和数据之间的一种接口实现的。在文件系统阶段，程序与数据之间的关系如图1.3所示。

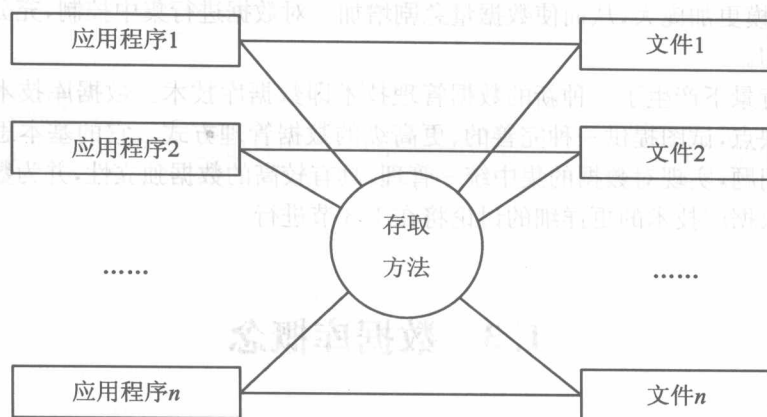


图 1.3 文件系统阶段应用程序与数据之间的对应关系

文件系统的最大特点是解决了应用程序和数据之间的一个公共接口问题，使得应用程序采用统一的存取方法来操作数据。在文件系统阶段中数据管理的特点如下：

① 数据可以长期保留,数据的逻辑结构和物理结构有了区别,程序可以按名访问,不必关心数据的物理位置,由文件系统提供存取方法。

② 数据不属于某个特定的应用程序,即应用程序和数据之间不再是直接的对应关系,可以重复使用。但是文件系统只是简单地存取数据,相互之间并没有有机的联系,即数据存取依赖于应用程序的使用方法,不同的应用程序仍然很难有效地共享同一数据文件。

③ 文件组织形式的多样化。有索引文件、链接文件和 Hash 文件等。但文件之间没有联系,相互独立,数据间的联系要通过程序去构造。

文件系统具有如下缺点:

① 数据冗余度大。文件与应用程序密切相关。相同的数据集合在不同的应用程序中使用,经常需要重复定义、重复存储。例如:学校中学生学籍管理系统中的学生情况,学生成绩管理系统的学生选课,教师教学管理的任课情况,所用到的数据很多都是重复的。这样相同的数据不能被共享,必然导致数据的冗余。

② 数据不一致性。由于相同数据的重复存储,单独管理,给数据的修改和维护带来难度,容易造成数据的不一致。例如,学校在学生学籍管理系统中修改了某学生的情况,但在学生成绩管理系统中该学生相应的信息没有被修改,造成同一个学生的信息在不同的管理部门结果不一样。

③ 数据联系弱。文件系统中数据组织成记录,记录由字段组成,记录内部有了一定的结构。但是文件之间是孤立的,从整体上看没有反映现实世界事物之间的内在联系,因此很难对数据进行合理地组织以适应不同应用的需要。

### 1.2.3 数据库系统阶段

自 20 世纪 60 年代末开始,磁盘技术已经成熟,并作为主要外存而广泛使用。计算机硬件的价格大幅度下降,可靠性增强,为数据管理技术的发展奠定了物质基础。另外,因为计算机被用于管理,其规模更加庞大,从而使数据量急剧增加。对数据进行集中控制,充分提供数据共享的要求日益迫切。

在这样的背景下产生了一种新的数据管理技术即数据库技术。数据库技术克服了以前所有管理方式的缺点,试图提供一种完善的、更高级的数据管理方式。它的基本思想是解决多用户数据共享的问题,实现对数据的集中统一管理,具有较高的数据独立性,并为数据提供各种保护措施。关于数据库技术的更详细的讨论将在 1.3 节进行。

## 1.3 数据库概念

数据库是一个很复杂的系统,涉及面很广,难以用简练的语言准确地概括其全部特征。因此,本节先从简单分析入手,逐步认识什么是数据库。

顾名思义,数据库就是存储数据的“仓库”。但它和普通的仓库是有所不同的。首先,数据



不是存放在容器或空间中,而是存放在计算机的外存储器中(如磁盘),并且是有组织地存放的。数据的管理和利用是通过计算机的数据管理软件——数据库管理系统来完成的。因此,我们讲的数据库,不单是指存有数据的计算机外存,而是指存放在外存上的数据集合以及管理它们的计算机软件的总和。数据库系统的组成可以用图 1.4 表示。

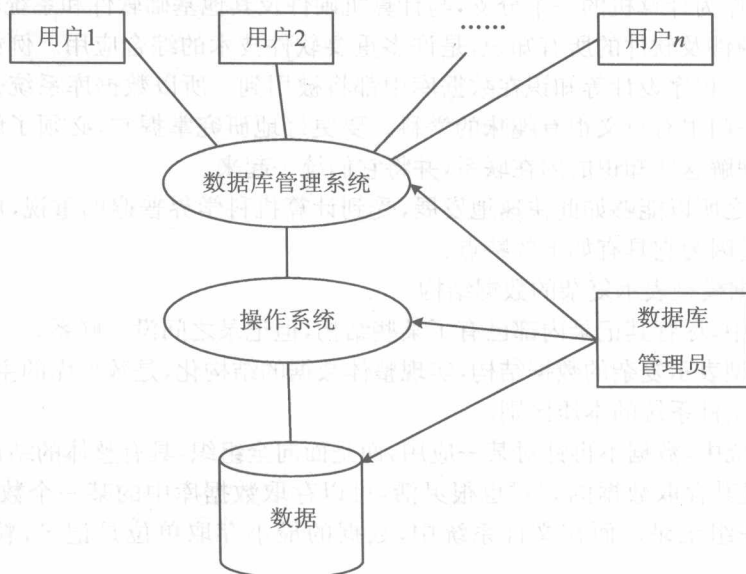


图 1.4 数据库系统

通常一个数据库系统包含下列内容:

(1) 有一个结构化的相关数据的集合

这一系统称为数据库(Database 或 DB)。在这个数据集中没有有害的或不必要的冗余,能够为多种应用服务,它独立于应用程序而存在。这种结构化的数据集合就是数据库本身,是数据库系统的核心和管理对象。

(2) 有一个负责数据库管理和维护的软件系统

这一系统称为数据库管理系统(Database Management System 或 DBMS)。

用户一般不直接加工或使用数据库中的数据,而必须通过数据库管理系统。DBMS 的主要功能是维持数据库系统的正常活动,接受并响应用户对数据库的一切访问要求,包括建立和删除数据文件、检索、统计、修改和组织数据库中的数据以及为用户提供对数据库的维护手段等。通过使用 DBMS,用户可逻辑、抽象地处理数据,不必关心数据在计算机中的具体存储方式以及计算机处理数据的过程细节。这样,把一切处理数据的具体而繁杂的工作交给 DBMS 去完成。就好像操作系统的出现,解放了用户,使用户不必关心数据的实际存放和读取,而只需给出文件名和路径一样。

(3) 有供数据库及数据库管理系统运行的运行环境

运行环境主要是指计算机硬件系统以及基础的软件支持系统(如:操作系统等)。

(4) 有一个(或一组)负责整个数据系统的建立、维护和协调工作的专门人员

这就是数据库管理员(Database Administrator 或 DBA)。在一个安全性较高的大型的数据