

新编高等院校计算机科学与技术规划教材

数据库原理及应用（第4版）

SHU JU KU YUAN LI JI YING YONG

主 编 钱雪忠 王月海

副主编 陈国俊 周頔 徐华 钱瑛



北京邮电大学出版社
www.buptpress.com

新编高等院校计算机科学与技术规划教材

数据库原理及应用

(第4版)

钱雪忠 王月海 主编

陈国俊 周 颀
徐 华 钱 瑛 副主编

北京邮电大学出版社

·北京·

内 容 简 介

本书重点地介绍了数据库系统的基本概念、基本原理和基本设计方法,同时以 SQL Server 为背景介绍了数据库原理的应用。本书力求对传统的数据库理论和应用进行精炼,保留实用的部分,使其更为通俗易懂,更为简明与实用。

全书共有 8 章,主要内容包括:数据库系统概述、数据模型、数据库系统结构、关系数据理论、SQL 语言、关系数据库设计理论、数据库安全保护、数据库设计、SQL Server 数据库管理系统和 XML 应用基础等。

本书内容循序渐进、深入浅出,每章都给出了较多的实例,各章后有适量的习题以便于读者练习与巩固所学知识。

本书可作为计算机各专业及信息类、电子类等相关专业的本科、专科“数据库原理及应用”类课程的教材,也可以供参加自学考试人员、数据库应用系统开发设计人员、工程技术人员及其他相关人员参阅。

图书在版编目(CIP)数据

数据库原理及应用 / 钱雪忠, 王月海主编. -- 4 版. -- 北京: 北京邮电大学出版社, 2015. 4
ISBN 978-7-5635-4291-8

I. ①数… II. ①钱…②王… III. ①数据库系统 IV. ①TP311.13

中国版本图书馆 CIP 数据核字(2015)第 021598 号

书 名: 数据库原理及应用(第 4 版)

著作责任者: 钱雪忠 王月海 主编

责任编辑: 徐振华 孙宏颖

出版发行: 北京邮电大学出版社

社 址: 北京市海淀区西土城路 10 号(邮编:100876)

发行部: 电话:010-62282185 传真:010-62283578

E-mail: publish@bupt.edu.cn

经 销: 各地新华书店

印 刷: 北京鑫丰华彩印有限公司

开 本: 787 mm×1 092 mm 1/16

印 张: 18.5

字 数: 481 千字

版 次: 2005 年 9 月第 1 版 2007 年 8 月第 2 版 2010 年 8 月第 3 版 2015 年 4 月第 4 版

2015 年 4 月第 1 次印刷

ISBN 978-7-5635-4291-8

定 价: 38.00 元

· 如有印装质量问题,请与北京邮电大学出版社发行部联系 ·

前言

数据库技术是计算机科学技术中发展最快的领域之一,也是应用最广的技术之一,它已成为计算机信息系统与应用系统的核心技术和重要基础。

随着计算机技术飞速发展及其应用领域的扩大,特别是计算机网络和 Internet 的发展,基于计算机网络和数据库技术的信息管理系统、应用系统得到了飞速的发展。当前,计算机的计算模式已由单用户→主从式或主机/终端式结构→C/S 结构→B/S 结构,而发展到了 Web 服务、网格与云计算时代,然而数据库及其技术一直是它们的后台与基础,并在发展中得到不断的完善改进与功能增强。目前,数据库技术已成为社会各行各业进行数据管理的必备技能。数据库技术相关的基本知识和基本技能是计算机及相关专业的必学内容。

“数据库原理及应用”课程就是为使学生全面掌握数据库技术而开设的专业基础课程。它现已是计算机各专业及信息类、电子类专业等的必修课程。该课程的主要目的是使学生在较好掌握数据库系统原理的基础上,能理论联系实际,使学生较全面透彻地掌握数据库应用技术。本书追求的目标也正是如此。

本书内容循序渐进、深入浅出、抓住要点、内容精选。

本书围绕数据库系统的基本原理与应用技术两个核心点展开。全书内容编排成 8 章,第 1 章集中介绍了数据库系统的基本概念、基本知识与基本原理,内容包括数据库系统概述、数据模型、数据库系统结构、数据库系统的组成、数据库技术的研究领域及其发展等;第 2 章借助数学的方法,较深刻透彻地介绍了关系数据库理论,内容包括关系模型、关系数据结构及形式化定义、关系的完整性、关系代数、关系演算等;第 3 章介绍了实用的关系数据库操作语言 SQL,内容包括 SQL 语言的基本概念与特点,SQL 语言的数据定义、数据查询、数据更新、数据控制、视图等功能介绍及嵌入式 SQL 语言应用初步;第 4 章是关于数据库设计理论方面的内容,主要介绍了规范化问题的提出、规范化、数据依赖的公理系统等;第 5 章是关于数据库统一管理与控制方面的内容,主要介绍了数据库的安全性、完整性控制、并发控制与封锁、数据库的恢复等;第 6 章介绍数据库设计方面的概念与开发设计过程,包括数据库设计概述及规范化数据库开发设计六步骤等;第 7 章以 SQL Server 数据库系统为背景,介绍了 SQL Server 数据库系统的变迁、新版 SQL Server 的核心 Transact-SQL 的全面简要介绍等;第 8 章简要介绍 XML 的基本知识、XML 数据库及 XML 数据的 XQuery 查询操作等。可见本书精炼、核心与实用,内容适合数据库原理及应用类课程教学需要。

本书在第 3 版的基础上,在内容编排、叙述、图示释义等方面继续做改进,力求使书本理论知识深入浅出,更适合于教与学;增补了数据库行业发展趋势(NoSQL 等)、候选码的求解、XML 应用基础(新增一章)等新内容,使数据库原理知识更系统更完整;数据库应用技术方面,基于功能更强大的 SQL Server 新版本做了更新;注重更正了第 3 版中存在的不足。

本书每章除基本知识外,还有章节要点、小结、适量的练习题等,以配合对知识点的掌握。

讲授时可根据学生、专业、课时等情况对内容适当取舍,带有“*”的章节内容是取舍的首选对象。本套教材提供了PPT演示稿,可从出版社网站下载。实验教学请使用本书配套最新版实验指导书——《数据库原理及应用实验指导(第3版)》(钱雪忠主编,北京邮电大学出版社出版)。

本书可作为计算机各专业及信息类、电子类专业等的数据库相关课程教材,也可以供参加自学考试人员、数据库应用系统开发设计人员、工程技术人员及其他相关人员参阅。

本书由钱雪忠、王月海等主编,参编人员有钱雪忠、陈国俊、周頔、徐华、钱瑛、马晓梅、程建敏、李京、黄学光、王萌等。另外,钱恒、邓杰、任看看、施亮、孙志鹏、赵娜娜、冯振华等参与了校稿、技术支持等相关工作。编写中得到江南大学物联网工程学院数据库课程组教师们的大力协助与支持,使编者获益良多,谨此表示衷心的感谢。

由于时间仓促,编者水平有限,书中难免有错误、疏漏和欠妥之处,敬请广大读者与同行专家批评指正。

编者

目 录

第 1 章 绪论	1
1.1 数据库系统概述	1
1.1.1 数据、数据库、数据库管理系统、数据库系统	1
1.1.2 数据管理技术的产生和发展	3
1.1.3 数据库系统的特点	6
1.2 数据模型	9
1.2.1 数据模型的组成要素	10
1.2.2 概念模型	12
1.2.3 基本 E-R 模型的扩展*	16
1.2.4 层次模型概述	22
1.2.5 网状模型	26
1.2.6 关系模型	27
1.2.7 面向对象模型*	29
1.3 数据库系统结构	33
1.3.1 数据库管理系统的三级模式结构	33
1.3.2 数据库的二级映像功能与数据独立性	35
1.3.3 数据库管理系统的工作过程	36
1.4 数据库系统的组成	37
1.5 数据库技术的研究领域及其发展*	38
1.5.1 数据库技术的研究领域	38
1.5.2 数据库技术的发展	39
1.5.3 数据库行业发展趋势	41
1.6 小结	46
习题	46
第 2 章 关系数据库	49
2.1 关系模型	49
2.2 关系数据结构及形式化定义	50
2.2.1 关系	50
2.2.2 关系模式	53
2.2.3 关系数据库	54

2.3 关系的完整性	54
2.4 关系代数	56
2.4.1 传统的集合运算	57
2.4.2 专门的关系运算	59
2.5 关系演算	65
2.5.1 抽象的元组关系演算*	65
2.5.2 元组关系演算语言 ALPHA	67
2.5.3 域关系演算语言 QBE*	74
2.6 小结	78
习题	79
第3章 关系数据库标准语言 SQL	82
3.1 SQL语言的基本概念与特点	82
3.1.1 语言的发展及标准化	82
3.1.2 SQL语言的基本概念	83
3.1.3 SQL语言的主要特点	84
3.2 SQL数据定义	85
3.2.1 字段数据类型	85
3.2.2 创建、修改和删除数据表	87
3.2.3 设计、创建和维护索引	90
3.3 SQL数据查询	91
3.3.1 SELECT命令的格式及其含义	91
3.3.2 SELECT子句的基本使用	94
3.3.3 WHERE子句的基本使用	95
3.3.4 常用集函数及统计汇总查询	98
3.3.5 分组查询	98
3.3.6 查询的排序	99
3.3.7 连接查询	100
3.3.8 合并查询	102
3.3.9 嵌套查询	102
3.3.10 子查询别名表达式的使用*	107
3.3.11 存储查询结果到表中	107
3.4 SQL数据更新	108
3.4.1 插入数据	108
3.4.2 修改数据	109
3.4.3 删除数据	109
3.5 视图	110
3.5.1 定义和删除视图	110
3.5.2 查询视图	111
3.5.3 更新视图	111

3.5.4 视图的作用	112
3.6 SQL 数据控制	113
3.6.1 权限与角色	113
3.6.2 系统权限和角色的授予与收回	113
3.6.3 对象权限和角色的授予与收回	114
3.7 嵌入式 SQL 语言*	115
3.7.1 嵌入式 SQL 简介	115
3.7.2 嵌入式 SQL 要解决的 3 个问题	115
3.7.3 第四代数据库应用开发工具或高级语言中 SQL 的使用	119
3.8 小结	122
习题	123
第 4 章 关系数据库设计理论	126
4.1 问题的提出	126
4.1.1 规范化理论概述	126
4.1.2 不合理的关系模式存在的问题	126
4.2 规范化	129
4.2.1 函数依赖	129
4.2.2 码	131
4.2.3 范式	135
4.2.4 第一范式	135
4.2.5 第二范式	136
4.2.6 第三范式	137
4.2.7 BC 范式	139
4.2.8 多值依赖与 4NF	140
4.2.9 连接依赖与 5NF*	143
4.2.10 规范化小结	144
4.3 数据依赖的公理系统*	145
4.4 关系分解保持性*	148
4.4.1 关系模式的分解	148
4.4.2 模式分解的无损连接性	149
4.4.3 模式分解的函数依赖保持性	150
4.4.4 模式分解算法	150
4.5 小结	152
习题	153
第 5 章 数据库安全保护*	156
5.1 数据库的安全性	156
5.1.1 数据库安全性概述	156
5.1.2 安全性控制的一般方法	156

5.1.3	安全性控制的其他方法	161
5.1.4	SQL Server 安全性概述	162
5.2	完整性控制	163
5.2.1	数据库完整性概述	163
5.2.2	完整性规则的组成	163
5.2.3	完整性约束条件的分类	164
5.2.4	SQL Server 完整性概述	165
5.3	并发控制与封锁	166
5.3.1	数据库并发性概述	166
5.3.2	事务的基本概念	167
5.3.3	并发操作与数据的不一致性	168
5.3.4	封锁及其产生问题的解决	170
5.3.5	SQL Server 的并发控制机制	175
5.4	数据库的恢复	178
5.4.1	数据库恢复概述	178
5.4.2	数据库恢复的基本原理及其实现技术	178
5.4.3	数据库的故障及其恢复策略	179
5.4.4	SQL Server 的备份和还原机制	181
5.5	小结	182
	习题	183
第6章	数据库设计	185
6.1	数据库设计概述	185
6.1.1	数据库设计的任务、内容和特点	185
6.1.2	数据库设计方法简述	186
6.1.3	数据库设计的步骤	187
6.2	系统需求分析	190
6.2.1	需求分析的任务	190
6.2.2	需求分析的方法	191
6.3	概念结构设计	193
6.3.1	概念结构设计的必要性	193
6.3.2	概念模型设计的特点	193
6.3.3	概念结构的设计方法和步骤	194
6.4	逻辑结构设计	201
6.4.1	逻辑结构设计的任务和步骤	201
6.4.2	初始化关系模式设计	202
6.4.3	关系模式的规范化	204
6.4.4	关系模式的评价与改进	205
6.5	物理结构设计	205
6.5.1	确定物理结构	205

6.5.2 评价物理结构	206
6.6 数据库实施	206
6.6.1 建立实际数据库结构	207
6.6.2 装入数据	207
6.6.3 编制与调试应用程序	207
6.6.4 数据库试运行	207
6.6.5 整理文档	208
6.7 数据库运行和维护	208
6.7.1 数据库的安全性和完整性	208
6.7.2 监视并改善数据库性能	208
6.7.3 数据库的重组和重构造	209
6.8 UML 简介*	209
6.8.1 概述	209
6.8.2 UML 的构成	210
6.8.3 UML 的系统开发思路及开发阶段	212
6.8.4 用例图	213
6.8.5 类图	215
6.8.6 交互图	221
6.8.7 状态图	222
6.8.8 活动图	223
6.8.9 构件图	223
6.8.10 配置图	224
6.8.11 包图	224
6.9 小结	225
习题	225
第 7 章 SQL Server 数据库管理系统*	229
7.1 微软数据平台的进化	229
7.1.1 SQL Server 2014 新特色	229
7.1.2 SQL Server 2012 简介	231
7.2 Transact-SQL 语言	232
7.2.1 语法规则	233
7.2.2 运算符	235
7.2.3 数据类型	238
7.2.4 函数	240
7.2.5 变量	246
7.2.6 表达式	248
7.2.7 控制流	249
7.2.8 事务语句	254
7.2.9 批处理	254

7.2.10 游标	255
7.3 小结	256
习题	256
第8章 XML应用基础	258
8.1 XML的基本知识	258
8.1.1 XML简介	258
8.1.2 XML的用途	259
8.1.3 XML树结构	261
8.1.4 XML语法规则	262
8.1.5 XML元素	264
8.1.6 XML属性	265
8.2 XML文档基本操作	267
8.2.1 创建XML文件	268
8.2.2 查看XML文件	269
8.2.3 使用XSLT显示XML	270
8.2.4 使用CSS显示XML	270
8.3 XML数据库简介	271
8.3.1 XML数据模型	272
8.3.2 XML数据库系统介绍	273
8.4 XML数据查询	275
8.4.1 XQuery简介	275
8.4.2 XQuery实例	277
8.4.3 XQuery FLWOR表达式	278
8.4.4 XQuery FLWOR+HTML	278
8.4.5 XQuery语法	279
8.4.6 XQuery添加元素和属性	280
8.4.7 XQuery函数	281
习题	282
参考文献	283

第1章 绪 论

本章要点

本章从数据库基本概念与知识出发,依次介绍了数据库系统的特点、数据模型的三要素及其常见数据模型、数据库系统的内部体系结构等重要概念与知识。本章的另一重点是围绕 DBMS 介绍其功能、组成与操作,还介绍了数据库技术的研究点及其发展变化情况。

1.1 数据库系统概述

数据库技术自从 20 世纪 60 年代中期产生以来,无论是理论还是应用方面都已变得相当的重要和成熟,成为计算机科学的重要分支。数据库技术是计算机领域发展最快的学科之一,也是应用很广、实用性很强的一门技术。目前,数据库技术已从第一代的网状、层次数据库系统,第二代的关系数据库系统,发展到以面向对象模型为主要特征的第三代数据库系统。

随着计算机技术飞速发展及其应用领域的扩大,特别是计算机网络和 Internet 的发展,基于计算机网络和数据库技术的信息管理系统,各类应用系统得到了突飞猛进地发展,如事务处理系统(TPS)、地理信息系统(GIS)、联机分析系统(OLAP)、决策支持系统(DSS)、企业资源规划(ERP)、客户关系管理(CRM)、数据仓库(DW)和数据挖掘(DM)等系统都是以数据库技术作为其重要的支撑。可以说,只要有计算机的地方,就在使用着数据库技术。因此,数据库技术的基本知识和基本技能正在成为信息社会人们的必备知识。

1.1.1 数据、数据库、数据库管理系统、数据库系统

数据、数据库、数据库管理系统、数据库系统是数据库技术密切相关的 4 个基本概念,它们是我们首先要认识的。

1. 数据(Data)

(1) 数据的定义

数据是用来记录信息的可识别的符号,是信息的具体表现形式。

(2) 数据的表现形式

数据是数据库中存储的基本对象。数据在大多数人的第一印象中就是数字,其实数字只是其中一种最简单的表现形式,是数据的一种传统和狭义的理解。按广义的理解来说,数据的种类有很多,如文字、图形、图像、音频、视频、语言以及学校学生的档案管理等,这些都是数据,都可以转化为计算机可以识别的标识,并以数字化后的二进制形式存入计算机。

为了了解世界,交流信息,人们需要描述各种事物。在日常生活中直接用自然语言描述。在计算机中,为了存储和处理这些事物,就要抽出对这些事物感兴趣的特征组成一个记录来描述。例如,在学生档案中,如果人们最感兴趣的是学生的姓名、性别、年龄、出生年月,那么可以这样来描述某一学生:(赵一,女,23,1982.05)。目前数据库中的数据主要是以这样的结构化记录形式存在着。

(3) 数据与信息的联系

上面表示的学生记录就是一个数据。对于此记录来说,要表示特定的含义,必须对它给予解释说明,数据解释的含义称为数据的语义(即信息),数据与其语义是不可分的。可以这样认为:数据是信息的符号表示或载体,信息则是数据的内涵,是对数据的语义解释。

再如:“小明今年12岁了”,数据“12”被赋予了特定的语义“岁”,它才具有表达年龄信息的功能。

2. 数据库(DataBase,DB)

数据库,从字面意思来说就是存放数据的仓库。具体而言就是长期存放在计算机内的有组织的可共享的数据集合,可供多用户共享,数据库中的数据按一定的数据模型组织、描述和储存,具有尽可能小的冗余度和较高的数据独立性以及易扩展性。

数据库具有两个比较突出的特点。

① 把在特定的环境中与某应用程序相关的数据及其联系集中在一块并按照一定的结构形式进行存储,即集成性。

② 数据库中的数据能被多个应用程序的用户所使用,即共享性。

3. 数据库管理系统(DataBase Management System,DBMS)

数据库管理系统是数据库系统的核心组成部分,是对数据进行管理的大型系统软件,用户在数据库系统中的一些操作,例如,数据定义、数据操作、数据查询和数据控制,这些操作都是由数据库管理系统来实现的。

数据库管理系统主要包括以下几个功能。

(1) 数据定义

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(包括表、视图、索引、存储过程等)进行定义,定义相关的数据库系统的结构和有关的约束条件。

(2) 数据操作

DBMS提供数据操作语言(Data Manipulation Language,DML),通过DML操作数据实现对数据库的一些基本操作,如查询、插入、删除和修改等。其中,国际标准数据库操作语言——SQL——就是DML的一种。

(3) 数据库的运行管理

这一功能是数据库管理系统的核心所在。DBMS通过对数据库在建立、运用和维护时提供统一管理和控制,以保证数据安全、正确、有效地正常运行。DBMS主要通过数据的安全性控制、完整性控制、多用户应用环境的并发性控制和数据库数据的系统备份与恢复4个方面来实现对数据库的统一控制功能的。

(4) 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、重组功能和性能监视、分析功能等。

(5) 其他功能

包括:DBMS 与网络中其他软件系统的通信;多个 DBMS 系统间的数据转换;异构数据库之间的互访和互操作;DBMS 开发工具的支持功能;DBMS Internet 网络功能。

常用的数据库管理系统有:Oracle、MS SQL Server、DB2、MySQL、PostgreSQL、Sybase、Informix、Ingres、ACCESS、VFP 系列、Kingbase ES、PBASE、EASYBASE、Openbase、Ipedo、Tamino 等。

4. 数据库系统(DataBase System, DBS)

数据库系统是指在计算机系统中引入数据库后的系统,其构成主要有数据库(及相关硬件)、数据库管理系统及其开发工具、应用系统、数据库管理员和各类用户这几部分。其中,在数据库的建立、使用和维护的过程中要有专门的人员来完成,这些人就被称为数据库管理员(DataBase Administrator, DBA)。

常用开发工具有:Java 语言、.NET 平台及语言如 C# 等、C 语言、PHP、VC++ 等。

数据库系统可以用图 1.1 表示。数据库系统在整个计算机系统中的地位如图 1.2 所示。

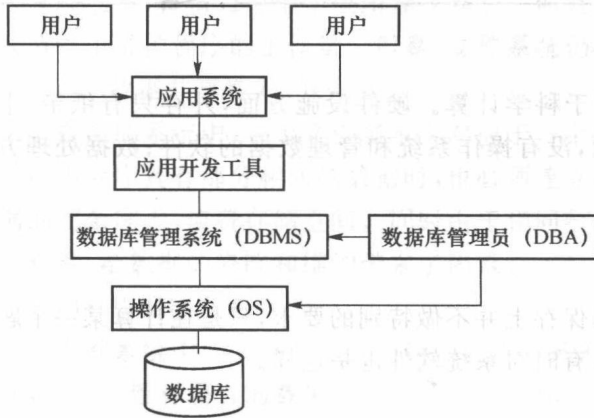


图 1.1 数据库系统

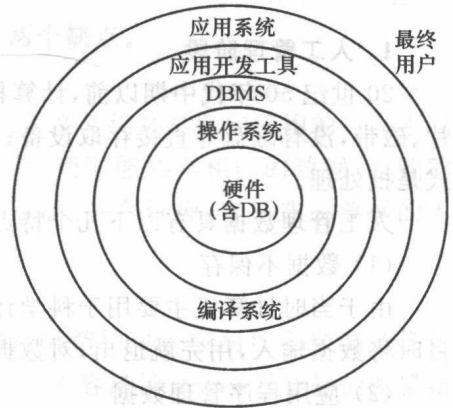


图 1.2 数据库在计算机系统中的地位

1.1.2 数据管理技术的产生和发展

谈数据管理技术,先要讲到数据处理,所谓数据处理是指对各种数据进行收集、存储、加工和传播的一系列活动的总和。数据管理则是数据处理的中心问题,为此,数据管理是指对数据进行分类、组织、编码、存储、检索和维护的管理活动总称。就用计算机来管理数据而言,数据管理是指数据在计算机内的一系列活动的总和。

随着计算机技术的发展,特别是在计算机硬件、软件与网络技术发展的前提下,人们对数据处理要求不断提高,在此情况下,数据管理技术也随之不断改进。人们借助计算机来进行数据管理虽是近五十多年的事,然而数据管理技术已经历了人工管理、文件系统及数据库系统 3 个发展阶段。这 3 个阶段的特点及其比较如表 1.1 所示。

表 1.1 数据管理 3 个阶段的比较

比较项目		人工管理阶段	文件系统阶段	数据库系统阶段
背景	应用背景	科学计算	科学计算、管理	大规模管理
	硬件背景	无直接存取存储设备	磁盘、磁鼓	大容量磁盘
	软件背景	没有操作系统	有文件系统	有数据库管理系统
	处理方式	批处理	联机实时处理、批处理	联机实时处理、分布处理、批处理
特点	数据的管理者	用户(程序员)	文件系统	数据库管理系统
	数据面向的对象	某一应用程序	某一应用	现实世界
	数据的共享程度	无共享,冗余度极大	共享性差,冗余度大	共享性好,冗余度小
	数据的独立性	不独立,完全依赖于程序	独立性差	具有高度的物理独立性和一定的逻辑独立性
	数据的结构化	无结构	记录内有结构、整体无结构	整体结构化,用数据模型描述
	数据控制能力	应用程序自己控制	应用程序自己控制	由数据库管理系统提供数据安全性、完整性、并发控制和恢复能力

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。硬件设施方面,外存只有纸带、卡片、磁带,没有磁盘等直接存取设备;软件方面,没有操作系统和管理数据的软件;数据处理方式是批处理。

人工管理数据具有以下几个特点。

(1) 数据不保存

由于当时计算机主要用于科学计算,数据保存上并不做特别的要求,只是在计算某一个题目时将数据输入,用完就退出,对数据不保存,有时对系统软件也是这样。

(2) 应用程序管理数据

数据没有专门的软件进行管理,需要应用程序自己进行管理,应用程序中要规定数据的逻辑结构和设计物理结构(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。如果多个应用程序涉及某些相同的数据,则由于必须各自进行定义,无法进行数据的参照,因此程序间有大量的冗余数据。

(4) 数据不具有独立性

数据的独立性包括了数据的逻辑独立性和数据的物理独立性。当数据的逻辑结构或物理结构发生变化时,必须对应用程序做相应的修改。

在人工管理阶段,程序与数据之间的对应关系可用图 1.3 表示,可见两者间是一对一的紧密依赖关系。

2. 文件系统阶段

20世纪50年代后期到60年代中期,这时计算机已大量用于数据的管理。硬件方面,有了磁盘、磁鼓等直接存取存储设备;软件方面,操作系统中已经有了专门的管理软件,一般称为

文件系统;处理方式有批处理、联机实时处理。

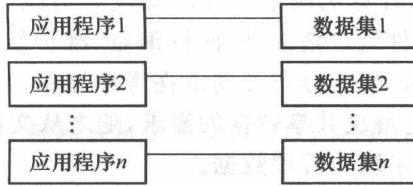


图 1.3 人工管理阶段应用程序与数据之间的对应关系

文件系统管理数据特点如下。

(1) 数据长期保存

由于计算机大量用于数据处理,数据需要长期保留在外存上反复进行查询、修改、插入和删除等操作。

(2) 文件系统管理数据

由专门的软件(即文件系统)进行数据管理,文件系统把数据组织成相互独立的数据文件,利用“按文件名访问,按记录进行存取”的管理技术,可以对文件进行修改、插入和删除等操作。文件系统实现了记录内的结构性,但大量文件之间整体无结构。程序和数据之间由文件系统提供存取方法进行转换,使应用程序与数据之间有了一定的独立性,程序员可以不必过多地考虑物理细节,将精力集中于应用程序算法。而且数据在存储上的改变不一定反映在程序上,大大节省了维护程序的工作量。但是,文件系统仍存在以下两个缺点。

(3) 数据共享性差,冗余度大

在文件系统中,一个文件基本上对应于一个应用程序,即文件仍然是面向应用的。当不同的应用程序具有部分相同的数据时,也必须建立各自的文件,而不能共享相同的数据,因此数据的冗余度大,浪费存储空间。同时由于相同数据的重复存储、各自管理,容易造成数据的不一致性,给数据的修改和维护带来了困难。

(4) 数据独立性差

文件系统中的文件是为某一特定应用服务的,文件的逻辑结构对该应用程序来说是优化的,因此要想对现有的数据增加一些新的应用会很困难,系统不容易扩充。一旦数据的逻辑结构改变,必须修改应用程序,修改文件结构的定义。应用程序的改变,例如,应用程序改用不同的高级语言等,也将引起文件的数据结构的改变。因此数据与程序之间仍缺乏独立性。可见,文件系统仍然是一个不具有弹性的整体无结构的数据集合,即文件之间是孤立的,不能反映现实世界事物之间的内存联系。在文件系统阶段,程序与数据之间的关系如图 1.4 所示,可见两者间仍有固定的对应关系。

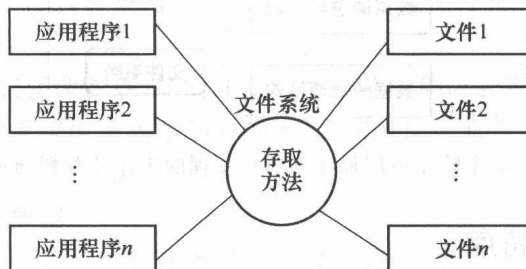


图 1.4 文件系统阶段应用程序与数据之间的对应关系

3. 数据库系统阶段

20世纪60年代后期以来,计算机用于管理的规模更为庞大,数据量急剧增长,硬件已有大容量磁盘,硬件价格下降;软件则价格上升,使得编制、维护软件及应用程序成本相对增加;处理方式上,联机实时处理要求更多,分布处理也在考虑之中。介于这种情况,文件系统的数据库管理满足不了应用的需求,为解决共享数据的需求,随之从文件系统中分离出了专门软件系统——数据库管理系统——用来统一管理数据。

数据库系统阶段应用程序与数据之间的对应关系可用图 1.5 表示,可见两者间已没有固定的对应关系。

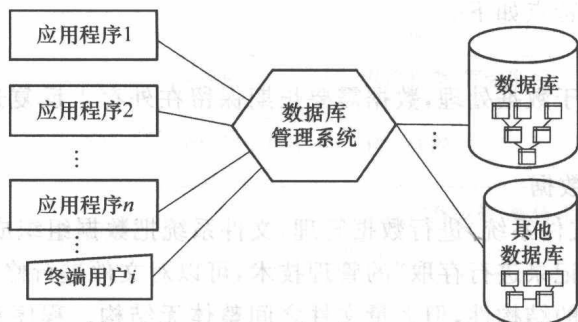


图 1.5 数据库系统阶段应用程序与数据之间的对应关系

综上所述,3个阶段应用程序与数据管理的工作任务如图 1.6 所示,随着数据管理技术的不断发展,应用程序不断从底层的、低级的、物理的数据管理工作中解脱出来,能独立地、较高逻辑级别地轻松处理数据库数据。从而能极大地提高应用软件的生产力。

数据库技术从20世纪60年代中期产生到现在仅仅50余年的历史,但其发展速度之快、使用范围之广是其他技术所不及的。60年代末出现了第一代数据库——层次数据库、网状数据库,70年代出现了第二代数据库——关系数据库。目前关系数据库系统已逐渐淘汰了层次数据库和网状数据库,成为当今最流行的商用数据库系统。

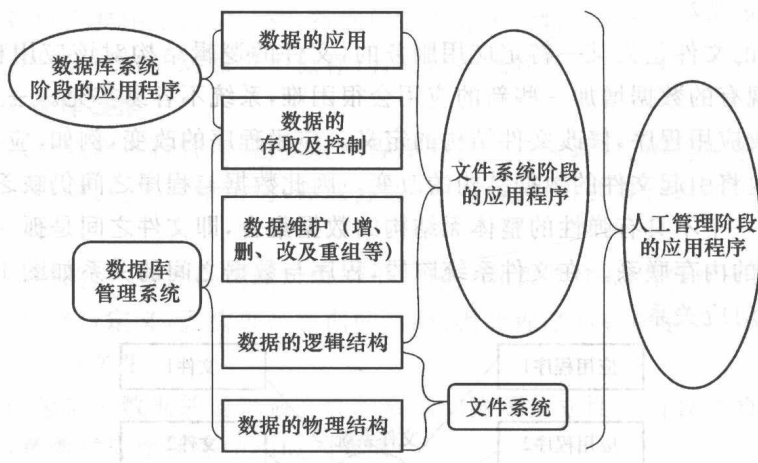


图 1.6 3个阶段应用程序与数据管理的工作任务划分示意图

1.1.3 数据库系统的特点

与其他两个数据管理阶段相比,数据库系统阶段数据管理有其自己的特点,主要体现在以