

高等学校规划教材

工程应用型院校计算机系列教材

安徽省高等学校“十二五”规划教材
安徽省高等学校计算机教育研究会推荐用书

总主编 胡学钢



数据库技术与应用

SHUJUKU JISHU YU YINGYONG

(第2版)

主编 叶明全 伍长荣



北京师范大学出版集团
BEIJING NORMAL UNIVERSITY PUBLISHING GROUP
安徽大学出版社

数据库技术与应用

(第2版)

总主编 胡学钢
主 编 叶明全 伍长荣
副主编 汪 洋 潘洁珠 崔 琳 殷云霞 范庆春
编写人员 (以姓氏笔画为序)

马吴迪	蚌埠学院
王平水	安徽财经大学
叶明全	皖南医学院
伍长荣	安徽师范大学
朱金好	皖南医学院
张 冲	铜陵学院
杨 利	皖南医学院
汪 洋	安庆师范学院
范庆春	合肥师范学院
胡春玲	合肥学院
殷云霞	安徽中医药大学
崔 琳	宿州学院
潘洁珠	合肥师范学院



北京师范大学出版集团
BEIJING NORMAL UNIVERSITY PUBLISHING GROUP
安徽大学出版社

图书在版编目(CIP)数据

数据库技术与应用/叶明全,伍长荣主编.—2版.合肥:安徽大学出版社,2015.7

工程应用型院校计算机系列教材/胡学钢总主编

ISBN 978-7-5664-0981-2

I. ①数… II. ①叶… ②伍… III. ①数据库系统—高等学校—教材 IV. ①TP311.13

中国版本图书馆 CIP 数据核字(2015)第 160834 号

数据库技术与应用

胡学钢 总主编
叶明全 伍长荣 主 编

出版发行: 北京师范大学出版集团
安徽大学出版社
(安徽省合肥市肥西路3号 邮编 230039)
www.bnupg.com.cn
www.ahupress.com.cn

印 刷: 合肥市裕同印刷包装有限公司
经 销: 全国新华书店
开 本: 184mm×260mm
印 张: 21.25
字 数: 438 千字
版 次: 2015 年 7 月第 2 版
印 次: 2015 年 7 月第 1 次印刷
定 价: 42.50 元
ISBN 978-7-5664-0981-2

策划编辑:李 梅 蒋 芳
责任编辑:蒋 芳
责任校对:程中业

装帧设计:李 军
美术编辑:李 军
责任印制:赵明炎

版权所有 侵权必究

反盗版、侵权举报电话:0551—65106311

外埠邮购电话:0551—65107716

本书如有印装质量问题,请与印制管理部联系调换。

印制管理部电话:0551—65106311

编写说明

计算机科学与技术的迅速发展,促进了许多相关学科领域以及应用分支的发展,同时也带动了各种技术和方法、系统与环境、产品以及思维方式等的发展。由此而进一步激发了对各种不同类型人才的需求。按照教育部计算机科学与技术专业教学指导委员会的研究报告来分,学校培养的人才类型可以分为科学型、工程型和应用型三类,其中科学型人才重在基础理论、技术和方法等的创新;工程型人才以开发实现预定功能要求的系统为主要目标;应用型人才以系统集成为主要途径实现特定功能的需求。

虽然这些不同类型人才的培养有许多共同之处,但是因不同类型人才的就业岗位所需要的责任意识、专业知识能力与素质、人文素养、治学态度、国际化程度等方面存在一定的差异,因而培养目标、培养模式等方面也存在不同。对大多数高校来说,很难兼顾各类人才的培养。因此,合理定位培养目标是确保教学目标和人才培养质量的关键。

由于当前社会领域从事工程开发和应用的岗位数量远远超过从事科学人才的数量,结合当前绝大多数高校的办学现状,安徽省高等学校计算机教育研究会在和多所高校专业负责人以及来自企业的专家反复研究和论证的基础上,形成了以培养工程应用型人才为主的安徽省高等学校计算机类专业的培养目标,并组织研讨组共同探索相关问题,共同建设相关教学资源,共享研究和建设成果,为全面推动安徽省高等学校计算机教育教学水平做出积极的贡献。北京师范大学出版集团安徽大学出版社积极支持安徽省高等学校计算机教育研究会的工作,成立了编委会,组织策划并出版了该套工程应用型计算机系列教材。

为了做好教材的出版工作,编委会在许多方面都采取了积极的措施:

编委会组成的多元化:编委会不仅有来自高校的教育领域的资深教师和专家,而且还有从事工程开发、应用技术的资深专家,从而为教材内容的重组提供更为有力的支持。

教学资源建设的针对性:教材以及教学资源建设的目标就是要突出体现“学以致用”的原则,减少“学不好,用不上”的空泛内容,增加其应用案例,尤其是增设涵盖更多知识和应用能力的系统性、综合性的案例,以培养学生系统解决问题的能力,进而激发其学习兴趣。

建设过程的规范性:编委会对整体的框架建设、对每本教材和资源的建设都采取汇报、交流和研讨的方式,以听取多方意见和建议;每本书的编写组也都进行反复的讨论和修订,努力提高教材和教学资源的质量。

如果我们的工作能对安徽省高等学校计算机类专业人才的培养做出贡献,那将是我们的荣幸。真诚欢迎有共同志向的高校、企业专家提出宝贵意见和建议,更期待你们参与我们的工作。

胡学钢

2015年6月10日于合肥

编委会名单

主 任 胡学钢(合肥工业大学)

委 员 (以姓氏笔画为序)

王 浩(合肥工业大学)

王一宾(安庆师范学院)

叶明全(皖南医学院)

孙 力(安徽农业大学)

刘仁金(皖西学院)

朱昌杰(淮北师范大学)

沈 杰(合肥炜煌电子有限公司)

李 鸿(宿州学院)

陈 磊(淮南师范学院)

陈桂林(滁州学院)

张先宜(合肥工业大学)

张润梅(安徽建筑大学)

张燕平(安徽大学)

金庆江(合肥文康科技有限公司)

周国祥(合肥工业大学)

周鸣争(安徽工程大学)

宗 瑜(皖西学院)

郑尚志(巢湖学院)

钟志水(铜陵学院)

姚志峰(蓝盾信息安全技术股份有限公司)

郭有强(蚌埠学院)

黄 勇(安徽科技学院)

黄海生(池州学院)

潘地林(安徽理工大学)

前 言

随着计算机技术的发展,特别是在计算机硬件、软件和网络技术发展的前提下,人们对数据处理的要求不断提高,在此情况下,数据管理技术也不断改进。数据库技术是计算机科学与技术中发展最快的领域之一,也是应用最广的技术之一。数据库技术所研究的具体问题就是如何科学地组织和存储数据,以及如何高效地获取和处理数据。

本书针对培养工程应用型人才的要求,从开发数据库应用系统以及使用数据库管理系统的角度来讲解数据库的基础理论知识和基本技术。

本书由四部分组成,共分 10 章。

第一部分(第 1~6 章)介绍数据库系统的基础理论知识和基本应用方法。内容主要包括数据库系统的基本概念、关系模型和关系代数运算、关系数据库标准语言 SQL、关系模式规范化理论、数据库设计,以及数据库四种维护措施(安全性、完整性、并发控制和数据库恢复)的原理和实现技术。

第二部分(第 7~8 章)主要介绍 Microsoft SQL Server 2005 数据库管理系统的基本操作和应用开发技术。内容主要包括 SQL Server 2005 的安装和配置、在 SQL Server 2005 环境中数据库和表的创建和使用、数据库的备份和恢复以及高级应用技术,并分别以 VC++、C# 和 Java 前端开发工具为例,介绍数据库应用系统开发的主要技术。

第三部分(第 9~10 章)主要介绍数据库的新技术和新应用。内容主要包括数据仓库、联机分析处理和数据挖掘的基本知识,以及应用 SQL Server 2005 Analysis Services 实现数据仓库、联机分析处理和主要挖掘技术的解决方案,以及面向对象数据库、分布式数据库、多媒体数据库和空间数据库的基本知识。

第四部分(附录 A~B)主要介绍数据库实验指导和课程设计指导内容。

本书由多所高校作者合作完成,其中,叶明全、伍长荣担任主编,负责全书的策划与组织,承担全书统稿与校对工作;汪洋、潘洁珠、崔琳、殷云霞、范庆春担任副主编。具体编写分工:叶明全编写第 1 章、第 9 章和附录 A;朱金好、崔琳编写第 2 章;朱金好、张冲编写第 3 章;汪洋、范庆春编写第 4 章;殷云霞、王平水编写第 5 章;汪洋、潘洁珠编写第 6 章;伍长荣、杨利编写第 7 章;伍长荣、马吴迪编写第 8 章;殷云霞、胡春玲编写第 10 章;崔琳、范庆春编写附录 B。

数据库技术目前还处于迅速发展,由于编者水平和时间的限制,书中难免存在疏漏和不足之处,恳请广大师生和读者给予批评指正。

编者

2015 年 5 月

内容简介

本书侧重于数据库技术及其应用,重点介绍了关系数据库应用系统开发所必备的基础知识和基本技术,包括数据库系统的基本概念、基本技术,SQL Server 2005 数据库应用系统开发的主要步骤和相关技术,数据仓库、联机分析处理和数据挖掘新技术及其在 SQL Server 2005 中实施应用,以及数据库实验指导、课程设计指导等。全书内容新颖,理论与实践相结合,实例丰富,实用性强。

本书可作为高等学校计算机科学与技术、信息管理与信息系统、软件工程、网络工程等相关专业数据库课程的教材或教学参考书,也可以供从事数据库应用系统开发的工程技术人员以及其他相关人员参阅。

目 录

第 1 章 绪论	1
1.1 数据库技术概述	1
1.1.1 数据库技术常用术语	1
1.1.2 数据管理技术的发展	2
1.2 数据模型概述	5
1.3 数据库系统的内部体系结构	7
1.3.1 数据库的三级模式结构	7
1.3.2 数据库的两级独立性	9
1.4 数据库领域的图灵奖	9
1.5 数据库技术的演变与发展	10
1.6 小结	12
习题 1	12
第 2 章 关系数据库	15
2.1 关系数据模型	15
2.1.1 关系数据结构	15
2.1.2 关系操作	19
2.1.3 关系的完整性	20
2.2 关系代数	22
2.2.1 传统的集合运算	23
2.2.2 专门的关系运算	24
2.3 关系演算	29
2.3.1 元组关系演算	29
2.3.2 域关系演算	30
2.4 关系代数查询优化	31
2.4.1 关系代数等价变换规则	31
2.4.2 启发式优化规则	33
2.4.3 关系表达式的优化	33
2.5 小结	34
习题 2	34
第 3 章 关系数据库标准语言 SQL	39
3.1 SQL 概述	39
3.1.1 SQL 的产生和发展	39

3.1.2	SQL 数据库的体系结构	40
3.1.3	SQL 的组成和特点	40
3.2	数据定义	42
3.2.1	SQL 的基本数据类型	42
3.2.2	模式的定义与删除	43
3.2.3	基本表的定义、修改和删除	44
3.2.4	索引的定义和删除	47
3.3	基本表更新	48
3.3.1	插入数据	48
3.3.2	修改数据	49
3.3.3	删除数据	50
3.4	数据查询	50
3.4.1	单表查询	51
3.4.2	多表查询	57
3.4.3	嵌套查询	59
3.4.4	集合查询	63
3.4.5	带子查询的数据更新	64
3.5	视图	65
3.5.1	视图的建立和删除	65
3.5.2	更新视图	67
3.5.3	查询视图	68
3.6	小结	69
	习题 3	70
第 4 章	关系模式规范化设计	73
4.1	关系规范化问题的提出	73
4.2	函数依赖	75
4.2.1	函数依赖的概念	76
4.2.2	码	77
4.3	关系模式的范式	77
4.3.1	第 1 范式(1NF)	78
4.3.2	第 2 范式(2NF)	79
4.3.3	第 3 范式(3NF)	80
4.3.4	BC 范式(BCNF)	81
4.4	多值依赖与第 4 范式(4NF)	82
4.4.1	多值依赖	83
4.4.2	第 4 范式(4NF)	85
4.5	规范化	86
4.6	数据依赖公理系统	87
4.7	模式分解	90
4.7.1	分解的无损连接性	92

4.7.2	分解的保持函数依赖性	93
4.8	小结	94
	习题 4	94
第 5 章	关系数据库设计	98
5.1	数据库设计概述	98
5.1.1	数据库设计的定义、内容和特点	98
5.1.2	数据库系统生存期	100
5.1.3	数据库设计的基本步骤	101
5.2	实体—联系模型	103
5.2.1	E-R 模型的基本元素	103
5.2.2	E-R 模型设计	105
5.3	需求分析	106
5.3.1	需求分析的任务和方法	106
5.3.2	数据字典	108
5.4	概念结构设计	109
5.4.1	概念模型	109
5.4.2	数据抽象	110
5.4.3	概念结构设计的 E-R 方法	110
5.5	逻辑结构设计	114
5.5.1	E-R 模型向关系模型的转换	114
5.5.2	关系模型的优化	115
5.5.3	设计数据库外模式	116
5.6	数据库的物理设计	116
5.6.1	物理设计的内容	116
5.6.2	设计数据库的存取方法	117
5.6.3	设计数据库的存储结构	117
5.6.4	评价物理结构	118
5.7	数据库的实施和维护	118
5.7.1	数据库的实施	118
5.7.2	数据库运行与维护	119
5.8	小结	120
	习题 5	120
第 6 章	关系数据库维护	124
6.1	事务	124
6.1.1	事务和程序的区别	124
6.1.2	如何定义事务	125
6.1.3	事务的特性	125
6.2	数据库的恢复技术	126
6.2.1	故障类型	126

6.2.2	恢复的实现技术	127
6.2.3	恢复策略	130
6.2.4	具有检查点的恢复技术	132
6.3	数据库的并发控制	133
6.3.1	并发控制概述	134
6.3.2	封锁机制	136
6.3.3	活锁和死锁	139
6.3.4	并发调度的可串行化	141
6.3.5	两段锁机制	142
6.3.6	封锁的粒度	144
6.4	数据库的完整性	146
6.4.1	实体完整性	146
6.4.2	参照完整性	147
6.4.3	用户定义的完整性	149
6.4.4	SQL 中的触发器	151
6.5	数据库的安全性	153
6.5.1	用户标识和鉴别	153
6.5.2	存取控制	154
6.5.3	自主存取控制	155
6.5.4	授权与回收	155
6.5.5	数据库角色	158
6.5.6	强制存取控制	159
6.6	小结	160
	习题 6	160
第 7 章	SQL Server 2005 数据库管理系统	164
7.1	SQL Server 2005 概述	164
7.1.1	SQL Server 的发展过程	164
7.1.2	SQL Server 2005 的安装	164
7.1.3	SQL Server 2005 的常用工具	171
7.1.4	SQL Server 2005 的系统数据库	174
7.2	数据库的创建、修改和删除	175
7.2.1	数据库的创建	175
7.2.2	数据库的修改	176
7.2.3	数据库的删除	177
7.3	基本表的创建、修改和删除	178
7.3.1	基本表的创建	178
7.3.2	基本表的修改	180
7.3.3	基本表的删除	181
7.4	数据的添加、删除、修改和查询	182
7.5	数据库的备份与还原	184

7.5.1	数据库的备份	184
7.5.2	数据库的还原	185
7.6	SQL Server 程序设计	187
7.6.1	程序中的批处理、注释、变量	187
7.6.2	流程控制语句	189
7.7	高级应用技术	192
7.7.1	存储过程	192
7.7.2	触发器	195
7.8	小结	197
	习题 7	197
第 8 章	数据库编程	200
8.1	VC++ 数据库编程	200
8.1.1	ADO 概述	200
8.1.2	VC++ 中使用 ADO	201
8.1.3	ADO 的应用实例	202
8.2	C# 数据库编程	205
8.2.1	ADO.NET 概述	205
8.2.2	ADO.NET 用法	205
8.2.3	C# 数据库连接实例	210
8.3	Java 数据库编程	211
8.3.1	JDBC 概述	211
8.3.2	使用 JDBC 进行数据库开发	212
8.3.3	JDBC API 的应用实例	213
8.4	小结	216
	习题 8	217
第 9 章	数据仓库与数据挖掘	218
9.1	数据仓库	218
9.1.1	数据仓库概述	218
9.1.2	数据仓库的系统结构	220
9.1.3	数据仓库中的数据组织	221
9.2	联机分析处理	223
9.2.1	OLAP 概述	223
9.2.2	多维数据模型	224
9.2.3	多维分析操作	227
9.2.4	多维查询语言	228
9.2.5	多维数据的存储	229
9.3	SQL Server 2005 数据仓库和联机分析处理实例	230
9.4	数据挖掘	239
9.4.1	数据挖掘概述	239

9.4.2	关联规则	242
9.4.3	分类	245
9.4.4	聚类	249
9.5	SQL Server 2005 数据挖掘实例	253
9.5.1	Microsoft 数据挖掘解决方案概述	253
9.5.2	Microsoft 关联算法	257
9.5.3	Microsoft 决策树算法	260
9.5.4	Microsoft 聚类分析算法	263
9.6	小结	265
	习题 9	265
第 10 章	数据库新技术	267
10.1	数据库技术发展趋势	267
10.2	面向对象数据库	268
10.2.1	面向对象数据库概述	268
10.2.2	面向对象数据模型	268
10.2.3	面向对象数据库系统	270
10.2.4	ODMG 标准	270
10.2.5	面向对象的数据操作	272
10.3	分布式数据库	275
10.3.1	分布式数据库概述	275
10.3.2	分布式数据库的设计目标	276
10.3.3	分布式数据存储	276
10.3.4	分布式数据库系统	277
10.3.5	分布式查询处理	279
10.3.6	分布式事务处理	280
10.4	其他数据库新技术	282
10.4.1	多媒体数据库	282
10.4.2	空间数据库	283
10.4.3	NoSQL 数据库	285
10.5	小结	288
	习题 10	288
附录 A:	上机实验指导	291
附录 B:	课程设计指导	306
参考文献		328

第1章 绪论

本章内容

- 介绍数据管理技术的发展
- 定义数据库技术中的一些常用术语
- 描述数据模型及数据库的体系结构
- 简介数据库发展史上的重要人物
- 描述数据库技术的新发展

数据库技术产生于20世纪60年代中期,已经历了三代演变,先后造就了数据库技术的先驱查尔斯·巴赫曼(Charles W. Bachman)、关系数据库之父埃德加·考特(Edgar F. Codd)和数据库专家詹姆斯·格雷(James Gray)三位图灵奖得主。数据库技术是现代计算机信息系统和计算机应用系统的核心技术和重要基础,已经成为先进信息技术的重要组成部分,它的出现极大地促进了计算机应用向各行各业的渗透。数据库的建设规模、数据库信息量的大小和使用频度已成为衡量一个国家信息化程度的重要标志。

1.1 数据库技术概述

数据库技术是应数据管理任务需要而产生的。数据管理是利用计算机硬件和软件技术对数据进行有效的收集、存储、处理和应用的过程。数据管理的目的在于充分有效地发挥数据的作用,一方面借助计算机科学地保存和管理复杂的、大量的数据,以便人们能够方便而充分地利用这些数据资源;另一方面通过从大量原始的数据中获取对人们有价值的信息,然后利用信息作为管理决策的依据。

下面介绍一些数据库常用的术语和相关概念。

1.1.1 数据库技术常用术语

与数据库技术密切相关的4个基本概念:数据、数据库、数据库管理系统和数据库系统。

1. 数据(Data)

数据是描述事物的符号记录,既可以是数字,也可以是文字、图形、图像、音频、视频等。例如,可以这样描述一位同学的基本情况:张三,男,1990年10月出生,安徽省合肥市人。数据有多种形式,它们都是通过数字化后存入计算机中。数据的含义称为“数据的语义”,数据与语义是不可分的。在计算机中,可以把描述事物的相关数据组成一条记录来存储和处理。例如,在计算机中可用记录来描述一位同学,即:

(张三,男,199010,安徽省合肥市)

2. 数据库(DataBase,DB)

数据库,顾名思义,是存储数据的仓库,是长期存放在计算机内有组织的、可共享的大量

数据的集合。数据是数据库中存储的基本对象。

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、较高的独立性和易扩展性,并可为各种用户共享。

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

数据库管理系统是位于用户与操作系统之间的一种数据管理软件。和操作系统一样,数据库管理系统也是计算机的系统软件,是一个大型复杂的软件系统。

DBMS 的主用功能包括:

①数据定义功能。DBMS 提供数据定义语言(Data Definition Language,DDL),以便用户对数据库中的数据对象进行定义。

②数据组织、存储和管理。DBMS 能够提高存储空间利用率和存取效率。

③数据操纵功能。DBMS 提供数据操纵语言(Data Manipulation Language,DML),实现对数据库的基本操作。

④数据库的事务管理和运行管理。在建立、运用和维护数据库时,DBMS 对数据库进行统一管理和控制,以保证数据的安全性、完整性、多用户并发使用以及系统发生故障的恢复。

⑤数据库的建立和维护功能。DBMS 提供一些实用程序和管理工具来完成数据库的建立和维护。

4. 数据库系统(DataBase System,DBS)

数据库系统是指引入了数据库后的计算机系统,一般由数据库、数据库管理系统、应用系统和数据库管理员组成。数据库系统的组成如图 1-1 所示。

数据库管理人员(DataBase Administrator,DBA)是指建立、运行和维护数据库的专门人员。DBMS 是 DBS 的一个重要组成部分。通常情况下,可将数据库系统简称为“数据库”。

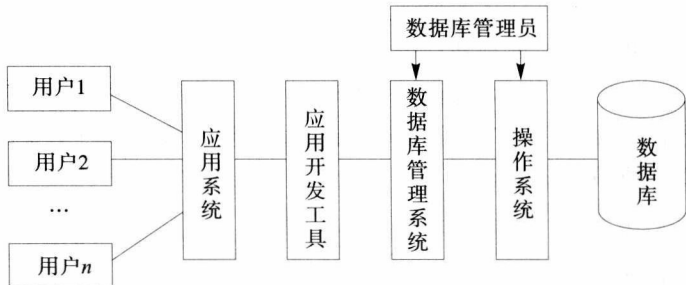


图 1-1 数据库系统的组成

关键概念:与数据库技术密切相关的 4 个基本概念是数据、数据库、数据库管理系统和数据库系统。

1.1.2 数据管理技术的发展

实现数据有效管理的关键是利用计算机对数据进行组织、存储、检索和维护等操作。在计算机技术发展和应用需求的推动下,数据管理技术经历了人工管理、文件系统和数据库系统 3 个阶段。

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。当时的计算机处于起步阶段,在硬件方面,外存只有磁带、卡片、纸带,没有磁盘等直接存取的存储设备,存储量非常小;在软件方面,没有操作系统,没有高级语言,数据处理采用批处理方式,即机器一次处理一批数据,直到运算完成为止,然后才能进行另外一批数据的处理,中间不能被打断,原因是此时的外存,如磁带、卡片等,只能顺序输入。

人工管理阶段的数据管理主要有以下特点:

①数据不保存。由于当时计算机主要用于科学计算,一般不需要长期保存数据,只是在计算某一个课题时,将数据随程序一起输入计算机,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。数据需要在应用程序中设计、说明和管理,没有相应的软件系统负责数据的管理工作。应用程序中要规定数据的逻辑结构,同时还要设计物理结构,包括数据的存储结构、存取方法、输入与输出方式等,这样必然导致程序员的负担很重。

③数据不共享。数据是面向应用程序的,一组数据仅仅对应一个应用程序。不同应用程序的数据之间是完全相互独立的,即使两个不同应用程序涉及相同的数据,也必须各自定义,无法相互利用、相互参照。因此,程序之间存在大量的冗余数据。

④数据不具备独立性。如果数据的逻辑结构或物理结构发生变化,则必须对应用程序做相应的修改,从而导致程序员负担过重。

图 1-2(a)所示为人工管理阶段程序与数据之间的一一对应关系。

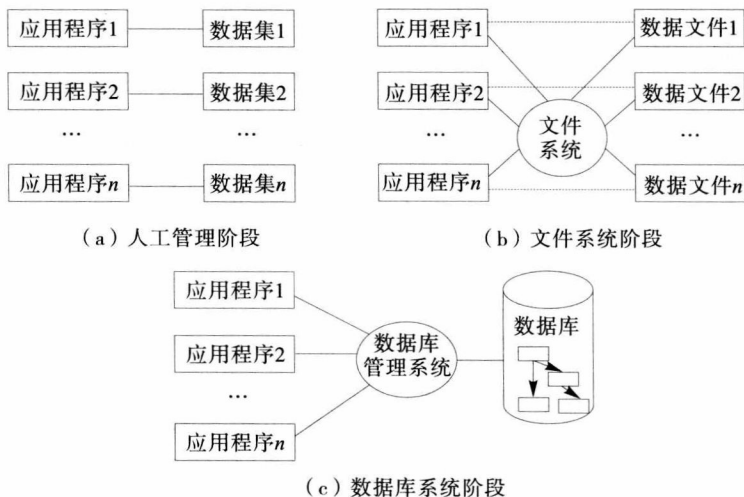


图 1-2 不同数据管理阶段程序与数据之间的对应关系

2. 文件系统阶段

20 世纪 50 年代后期到 60 年代中期,数据管理发展到文件系统阶段。此时的计算机有较大的发展,不仅用于科学计算,还大量用于管理。外存储器有了磁盘等直接存取的存储设备。在软件方面,操作系统中已有了专门的管理数据软件,称为“文件系统”。数据处理的方式不仅有批处理,也有联机实时处理。

文件系统阶段的数据管理主要有以下特点:

①数据以“文件”的形式长期保存。由于数据可以长期保存在外存上,因此可以经常对文件进行查询、修改和删除等操作。

②程序与数据之间具有“设备独立性”。利用操作系统中的文件系统进行专门的数据管理。文件系统把数据组织成相互独立的数据文件,采用“按文件名访问,按记录进行存取”的管理技术。因此,程序与数据之间有了一定的独立性。

③可以实时处理。由于有了直接存取设备,也有了索引文件、链接存取文件、直接存取文件等,所以既可以采用顺序批处理,也可以采用实时处理方式。

文件系统阶段是数据管理技术发展中的一个重要阶段,比人工管理阶段有了很大的改进。但这种方法仍有以下缺点:

①数据共享性差,冗余度大。当不同的应用程序所需的数据有部分相同时,仍需要建立各自独立的数据文件,而不能共享相同的数据。这样将会导致数据的冗余度大,严重浪费存储空间。另外由于相同的数据重复存储、各自管理,当需要修改相同部分的数据时,容易导致数据的不一致性,这就给数据的修改和维护带来了困难。

②数据和程序缺乏足够的独立性。文件中的数据面向特定的应用程序。当数据的逻辑结构改变时,应用程序必须重新修改,文件结构必须重新定义。同时,应用程序的改变也将引起文件结构的改变。因此,程序和数据之间仍缺乏独立性,文件之间是孤立的,不能反映现实世界事物之间的内在联系。

图 1-2(b)所示为文件系统管理阶段程序与数据之间的对应关系。

3. 数据库系统阶段

从 20 世纪 60 年代后期开始,数据管理进入数据库系统阶段。这一时期用计算机管理的数据规模日益庞大,应用越来越广泛,数据量急剧增长,数据要求共享的呼声越来越高。这种共享的含义是多种应用、多种语言互相覆盖地共享数据集合。此时的计算机有了大容量磁盘,计算能力也非常强,在这样的背景下,数据管理技术进入数据库系统阶段。

数据库系统的目标是解决数据冗余问题,实现数据独立性,实现数据共享,并解决由于数据共享而带来的数据完整性、安全性及并发控制等一系列问题。为实现这一目标,能够统一管理和共享数据的数据库管理系统应运而生。

数据库系统阶段数据管理具有以下特点:

①数据结构化。数据结构化是数据库系统与文件系统的根本区别。数据的结构化是数据库的主要特征之一,也是数据库系统与文件系统的根本区别。

②数据共享性高,冗余度小,系统易扩充。数据库从整体的观点来看待和描述数据,数据不再是面向某一应用,而是面向整个系统。这样就减小了数据的冗余,节约存储空间,缩短存取时间,避免数据之间的不一致性。当一些应用需求改变时,只要重新选择数据子集或者加上一部分数据,便可以满足更多的要求,从而保证了系统的易扩充性。

③数据独立性高。数据库提供数据的存储结构与逻辑结构之间的映象,使得当数据的物理存储结构改变时,数据的逻辑结构可以不变,程序也不用改变。这就是数据与程序的物理独立性。数据库还提供了数据的总体逻辑结构与某类应用所涉及的局部逻辑结构之间的映象,当数据的总体逻辑结构改变时,局部逻辑结构可以通过这种映象的转换保持不变,从而程序也不用改变,即数据与程序的逻辑独立性。

④数据由数据库管理系统进行统一管理和控制。通过数据库管理系统,多个用户可以