

21世纪高等学校计算机教育实用规划教材

周屹 李艳娟 主编

董莉 主审

数据库原理及 开发应用（第二版）

清华大学出版社

21世纪高等学校计算机教育实用规划教材

周屹 李艳娟 主编

崔琨 姜晓宏 副主编

数据库原理及 开发应用（第二版）

清华大学出版社
北 京

内 容 简 介

为了适应数据库技术的发展需要,满足广大读者和教师的建议和需求,本书是在《数据库原理及开发应用》的基础上修订而成,本版深入浅出地讲解了数据库原理及其应用。相对第一版,本版对原书各章节增加了大量的例题,本版更系统全面、例题更丰富、实用性更强、注重培养学生的应用技能和能力,既能更方便教师的讲授,又便于学生掌握和理解数据库知识。

本书的主要内容为数据库概述、数据模型、关系数据库的基本理论、Transact-SQL 语言、结构化查询语言 SQL、存储过程和触发器、关系模式的规范化设计、数据库设计、数据库的安全与保护、数据库应用系统设计实例、Web 数据库应用系统、数据库技术新进展。

本书注重教材内容循序渐进、由浅入深、理论与实践相结合,书中有大量的图形和文字说明,既便于教师教学又便于学生学习。本书可作为各类高等院校特别是应用型本科院校数据库技术课程教材,也可供计算机数据库爱好者参考使用。

本书封面贴有清华大学出版社防伪标签,无标签者不得销售。

版权所有,侵权必究。侵权举报电话:010-62782989 13701121933

图书在版编目(CIP)数据

数据库原理及开发应用/周屹等主编.--2 版.--北京:清华大学出版社,2013.2

21 世纪高等学校计算机教育实用规划教材

ISBN 978-7-302-30246-9

I. ①数… II. ①周… III. ①数据库系统 IV. ①TP311.13

中国版本图书馆 CIP 数据核字(2012)第 233231 号

责任编辑:闫红梅 李 晔

封面设计:常雪影

责任校对:时翠兰

责任印制:王静怡

出版发行:清华大学出版社

网 址: <http://www.tup.com.cn>, <http://www.wqbook.com>

地 址:北京清华大学学研大厦 A 座 邮 编:100084

社 总 机:010-62770175 邮 购:010-62786544

投稿与读者服务:010-62776969, c-service@tup.tsinghua.edu.cn

质 量 反 馈:010-62772015, zhiliang@tup.tsinghua.edu.cn

课 件 下 载: <http://www.tup.com.cn>, 010-62795954

印 刷 者:北京富博印刷有限公司

装 订 者:北京市密云县京文制本装订厂

经 销:全国新华书店

开 本:185mm×260mm 印 张:21.5 字 数:524 千字

版 次:2007 年 8 月第 1 版 2013 年 2 月第 2 版 印 次:2013 年 2 月第 1 次印刷

印 数:1~3000

定 价:35.00 元

产品编号:048603-01

出版说明

随着我国高等教育规模的扩大以及产业结构调整的不断深入,社会对高层次应用型人才的需求将更加迫切。各地高校紧密结合地方经济建设发展需要,科学运用市场调节机制,合理调整和配置教育资源,在改革和改造传统学科专业的基础上,加强工程型和应用型学科专业建设,积极设置主要面向地方支柱产业、高新技术产业、服务业的工程型和应用型学科专业,积极为地方经济建设输送各类应用型人才。各高校加大了使用信息科学等现代科学技术提升、改造传统学科专业的力度,从而实现传统学科专业向工程型和应用型学科专业的发展与转变。在发挥传统学科专业师资力量强、办学经验丰富、教学资源充裕等优势的同时,不断更新教学内容、改革课程体系,使工程型和应用型学科专业教育与经济建设相适应。计算机课程教学在从传统学科向工程型和应用型学科转变中起着至关重要的作用,工程型和应用型学科专业中的计算机课程设置、内容体系和教学手段及方法等也具有不同于传统学科的鲜明特点。

为了配合高校工程型和应用型学科专业的建设和发展,急需出版一批内容新、体系新、方法新、手段新的高水平计算机课程教材。目前,工程型和应用型学科专业计算机课程教材的建设工作仍滞后于教学改革的实践,如现有的计算机教材中有不少内容陈旧(依然用传统专业计算机教材代替工程型和应用型学科专业教材),重理论、轻实践,不能满足新的教学计划、课程设置的需要;一些课程的教材可供选择的品种太少;一些基础课的教材虽然品种较多,但低水平重复严重;有些教材内容庞杂,书越编越厚;专业课教材、教学辅助教材及教学参考书短缺,等等,都不利于学生能力的提高和素质的培养。为此,在教育部相关教学指导委员会专家的指导和帮助下,清华大学出版社组织出版本系列教材,以满足工程型和应用型学科专业计算机课程教学的需要。本系列教材在规划过程中体现了如下一些基本原则和特点。

(1) 面向工程型与应用型学科专业,强调计算机在各专业中的应用。教材内容坚持基本理论适度,反映基本理论和原理的综合应用,强调实践和应用环节。

(2) 反映教学需要,促进教学发展。教材规划以新的工程型和应用型专业目录为依据。教材要适应多样化的教学需要,正确把握教学内容和课程体系的改革方向,在选择教材内容和编写体系时注意体现素质教育、创新能力与实践能力的培养,为学生知识、能力、素质协调发展创造条件。

(3) 实施精品战略,突出重点,保证质量。规划教材建设仍然把重点放在公共基础课和专业基础课的教材建设上;特别注意选择并安排一部分原来基础比较好的优秀教材或讲义修订再版,逐步形成精品教材;提倡并鼓励编写体现工程型和应用型专业教学内容和课程体系改革成果的教材。

(4) 主张一纲多本,合理配套。基础课和专业基础课教材要配套,同一门课程可以有多种具有不同内容特点的教材。处理好教材统一性与多样化,基本教材与辅助教材,教学参考书,文字教材与软件教材的关系,实现教材系列资源配套。

(5) 依靠专家,择优选用。在制订教材规划时要依靠各课程专家在调查研究本课程教材建设现状的基础上提出规划选题。在落实主编人选时,要引入竞争机制,通过申报、评审确定主编。书稿完成后要认真实行审稿程序,确保出书质量。

繁荣教材出版事业,提高教材质量的关键是教师。建立一支高水平的以老带新的教材编写队伍才能保证教材的编写质量和建设力度,希望有志于教材建设的教师能够加入到我们的编写队伍中来。

21 世纪高等学校计算机教育实用规划教材编委会

联系人:魏江江 weijj@tup.tsinghua.edu.cn

前 言

《数据库原理及开发应用》第一版于 2007 年 8 月由清华大学出版社出版。近几年由于计算机的广泛普及与发展,特别是数据库技术的迅猛发展,数据库技术已经成为应用最广的技术之一。为了适应发展需要,并结合近年来的教学实践以及广大读者和教师提出的建议和需求,我们调整了第一版中部分内容,推出了第二版,对原书从结构到内容做了较大的增删和修改,特别是本书各章节增加了更多的例题,相对于第一版,教师和学生更容易利用这些实例进行学习,更好地掌握和理解数据库知识,本书更加注重全面培养学生的应用技能和能力。

全书共分为 12 章,第 1 章数据库概述、第 2 章数据模型、第 3 章关系数据库的基本理论、第 4 章 Transact-SQL 语言、第 5 章结构化查询语言 SQL、第 6 章存储过程和触发器、第 7 章关系模式的规范化设计、第 8 章数据库设计、第 9 章数据库的安全与保护、第 10 章数据库应用系统设计实例、第 11 章 Web 数据库应用系统、第 12 章数据库技术新进展。

本书第二版与第一版相比,主要有以下几方面的调整:

第一版第 1 章中增加了 SQL Server 2005 数据库的建立内容,将 SQL Server 2005 的系统数据类型部分的介绍提出放到第 4 章,对原第 3 章中各章节的顺序进行了调整,增加或删减了部分章节的内容。将第一版第 4 章的 Transact-SQL 语言部分的内容提出在第二版中扩充为单独一章,即第 4 章 Transact-SQL 语言,将第一版的第 4 章结构化查询语言 SQL 改为第 5 章,并对内容进行了调整,增加了大量的例题,并把系统存储过程提出放到第 6 章,第二版的第 6 章存储过程和触发器为新增加的内容,将第一版第 5 章关系模式的规范化设计改为第 7 章,后面的章节依次往后顺延,最后删去了第一版中第 11 章 Oracle 数据库管理系统的内容。

本书内容紧扣教学大纲,密切结合实际,以应用为目的,结构清晰,内容系统翔实,布局合理,加强基本概念和基本分析方法的介绍,尽量用实例解释概念,用例题来演绎方法和原理,叙述采用启发式,逻辑性强、文字流畅、通俗易懂。

本书由长期承担数据库技术基础课教学、具有丰富实践经验的一线教师编写,根据作者多年从事数据库技术课程教学活动以及数据库系统开发的经验,针对多数学生的认知规律编写,注重教材内容循序渐进、由浅入深、理论与实践相结合,书中有大量的图形和文字说明,既便于教师教授,又便于学生学习。

本书由周屹、李艳娟主编,董莉教授主审,崔琨、姜晓宏副主编,其中第 1 章和第 3 章由

崔琨编写,第2章、第5章和第12章由周屹编写,第4章和第11章由姜晓宏编写,第6章和第9章由李艳娟编写,第7章和第10章由李向宏编写,第8章由张谢群编写,全书由周屹统稿。

本书在编写过程中,得到了各方面有关专家的大力支持和帮助,在此对所有人的工作与支持表示衷心的感谢。由于编者水平有限,书中难免存在不足,敬请广大读者批评指正。编者的E-mail地址是zhouyi_1@163.com。

编 者

2012年7月

目 录

第 1 章 数据库概述	1
1.1 数据管理技术的发展	1
1.1.1 信息与数据	1
1.1.2 数据管理技术的发展	2
1.2 数据库系统的结构	5
1.2.1 数据库三级模式结构	5
1.2.2 数据库系统的体系结构	6
1.3 数据库、数据库管理系统与数据库系统	7
1.3.1 数据库	7
1.3.2 数据库管理系统	8
1.3.3 数据库系统	8
1.4 数据库技术的发展	9
1.4.1 数据库的发展历程	9
1.4.2 数据库研究的现状和未来的发展趋势	10
1.5 SQL Server 2005 简介	11
1.5.1 SQL Server 2005 概述	11
1.5.2 SQL Server 2005 常见版本	13
1.5.3 SQL Server 2005 工具和实用程序简介	14
1.5.4 SQL Server 2005 系统数据库简介	18
1.5.5 SQL Server 2005 数据库的建立	20
1.6 本章小结	24
1.7 习题	25
第 2 章 数据模型	26
2.1 信息的三个世界	26
2.1.1 现实世界	27
2.1.2 信息世界	27
2.1.3 数据世界	28
2.1.4 三个世界之间的关系	28
2.2 概念模型	29

2.2.1	实体、属性及联系	29
2.2.2	实体间的联系	30
2.2.3	实体联系方法	31
2.3	逻辑模型	31
2.3.1	逻辑模型的三要素	32
2.3.2	层次模型	32
2.3.3	网状模型	34
2.3.4	关系模型	35
2.3.5	面向对象模型	36
2.3.6	数据模型的比较	38
2.4	本章小结	39
2.5	习题	39
第3章	关系数据库的基本理论	42
3.1	关系模型概述	42
3.2	关系数据模型	44
3.2.1	关系数据结构和基本术语	45
3.2.2	关系的数学定义	46
3.2.3	关系的性质	47
3.3	关系模型的三类完整性	49
3.3.1	实体完整性(Entity Integrity)	49
3.3.2	参照完整性(Referential Integrity)	49
3.3.3	用户定义完整性	51
3.4	关系代数的基本运算	51
3.4.1	传统的集合运算	52
3.4.2	专门的关系运算	54
3.4.3	基本运算及其变换运算	59
3.5	关系演算	59
3.5.1	元组关系演算语言——ALPHA	59
3.5.2	域关系演算	63
3.6	本章小结	67
3.7	习题	67
第4章	Transact-SQL 语言	71
4.1	Transact-SQL 语言基本概念	71
4.1.1	数据类型	71
4.1.2	变量和运算符	74
4.2	流程控制语句	77
4.3	函数	80

4.3.1	系统函数	80
4.3.2	用户自定义函数	83
4.4	本章小结	85
4.5	习题	85
第 5 章	结构化查询语言 SQL	86
5.1	概述	86
5.1.1	SQL 的产生和发展	86
5.1.2	SQL 的特点	86
5.2	数据定义功能	87
5.2.1	基本表的定义	87
5.2.2	表结构的修改	89
5.2.3	基本表的删除	90
5.2.4	索引的建立	90
5.2.5	索引的删除	91
5.3	数据查询功能	91
5.3.1	投影查询	92
5.3.2	条件查询	94
5.3.3	连接查询	97
5.3.4	排序	100
5.3.5	分组及计算查询	100
5.3.6	子查询	102
5.3.7	集合操作	104
5.4	数据操作功能	105
5.4.1	插入数据语句	106
5.4.2	更新数据语句	107
5.4.3	删除数据语句	107
5.4.4	保持数据的一致性	108
5.5	视图	108
5.5.1	视图的建立	109
5.5.2	视图的修改	111
5.5.3	视图的删除	111
5.5.4	视图更新操作	111
5.6	数据控制功能	112
5.6.1	授予权限	112
5.6.2	回收权限	113
5.6.3	拒绝权限	114
5.7	嵌入式 SQL	114
5.7.1	嵌入式 SQL 简介	114

5.7.2	游标	115
5.7.3	动态 SQL 简介	118
5.8	本章小结	119
5.9	习题	120
第 6 章	存储过程和触发器	122
6.1	存储过程	122
6.1.1	存储过程的类型	122
6.1.2	存储过程的创建和执行	123
6.1.3	存储过程的修改和删除	127
6.1.4	存储过程的优点	128
6.2	触发器	128
6.2.1	触发器的类型	129
6.2.2	触发器的工作原理	130
6.2.3	触发器的创建和查看	130
6.2.4	触发器的修改和删除	134
6.2.5	触发器的优缺点	134
6.3	本章小结	135
6.4	习题	135
第 7 章	关系模式的规范化设计	137
7.1	问题提出	137
7.2	数据依赖	137
7.2.1	属性间联系	138
7.2.2	函数依赖	139
7.2.3	多值依赖	140
7.3	规范化的关系模式	141
7.3.1	第一范式	141
7.3.2	第二范式	141
7.3.3	第三范式	142
7.3.4	BC 范式	143
7.3.5	第四范式	144
7.3.6	规范化小结	145
7.4	数据依赖的公理系统	147
7.5	模式分解	150
7.6	本章小结	151
7.7	习题	152

第 8 章 数据库设计	155
8.1 概述	155
8.1.1 数据库设计的任务、内容和特点	155
8.1.2 数据库设计方法简述	156
8.1.3 数据库设计的步骤	158
8.2 需求分析	160
8.2.1 需求分析的任务	160
8.2.2 需求分析的步骤	161
8.3 概念结构设计	164
8.3.1 概念结构设计的必要性	164
8.3.2 概念模型的特点	165
8.3.3 概念结构设计的方法与步骤	165
8.4 逻辑结构设计	173
8.4.1 E-R 图向关系模型的转换	173
8.4.2 数据模型的优化	175
8.5 数据库物理设计	177
8.5.1 确定物理结构	177
8.5.2 评价物理结构	179
8.6 数据库实施	179
8.7 数据库运行与维护	181
8.8 数据库设计中的技巧	182
8.9 本章小结	186
8.10 习题	186
第 9 章 数据库的安全与保护	188
9.1 数据库的安全性	188
9.1.1 用户认证	188
9.1.2 存取控制	189
9.1.3 视图机制	191
9.1.4 审计	194
9.1.5 数据加密	196
9.1.6 SQL Server 2005 的安全性管理	198
9.2 数据库的完整性	209
9.2.1 数据库的完整性的含义	209
9.2.2 完整性约束条件	209
9.2.3 完整性控制	211
9.2.4 SQL Server 2005 的完整性保护	213
9.3 数据库的并发控制技术	218

9.3.1	事务概述	218
9.3.2	并发控制方法	219
9.3.3	并发调度的可串行性	224
9.3.4	SQL Server 2005 的并发控制	224
9.4	数据库的备份与恢复技术	225
9.4.1	数据库的故障种类	226
9.4.2	数据备份与恢复	227
9.4.3	SQL Server 2005 的数据库备份	232
9.4.4	SQL Server 2005 的数据库还原	237
9.5	本章小结	240
9.6	习题	240
第 10 章	数据库应用系统设计实例——学生管理系统	242
10.1	用户需求	242
10.1.1	系统功能分析	242
10.1.2	系统功能模块设计	243
10.2	数据库设计	243
10.2.1	数据库需求分析	243
10.2.2	数据库概念结构设计	244
10.2.3	数据库逻辑结构设计	244
10.3	数据库结构的实现	246
10.3.1	创建数据库和数据表	246
10.3.2	SQL Server 2005 的设置	248
10.3.3	前台客户端编程所使用的高级语言——Visual Basic 6.0 简介	249
10.3.4	创建数据库	257
10.3.5	创建 ODBC	258
10.4	学生管理系统主窗体的创建	258
10.4.1	创建工程项目——Student_MIS	259
10.4.2	创建学生管理系统的主窗体	259
10.4.3	创建主窗体的菜单	260
10.4.4	创建公用模块	260
10.5	系统用户管理模块的创建	263
10.5.1	用户登录窗体的创建	263
10.5.2	添加用户窗体的创建	265
10.6	本章小结	267
10.7	习题	267
第 11 章	Web 数据库应用系统	269
11.1	Web 概述	269

11.1.1	Web 的历史	269
11.1.2	Web 工作模式	270
11.1.3	与 Web 有关的一些概念	271
11.2	Web 数据库	271
11.2.1	Web 数据库的概念	272
11.2.2	Web 应用系统	272
11.2.3	常用 Web 数据库的比较	274
11.3	Web 数据库访问技术	276
11.3.1	ODBC	276
11.3.2	JDBC	282
11.3.3	JSP	284
11.4	Web 数据库应用系统	285
11.4.1	ASP 简介	285
11.4.2	ASP 访问数据库示例	292
11.5	本章小结	296
11.6	习题	296
第 12 章	数据库技术新进展	298
12.1	数据库技术发展概述	298
12.2	数据库技术与其他相关技术相结合	299
12.2.1	分布式数据库	299
12.2.2	并行数据库	307
12.2.3	主动数据库	309
12.2.4	面向对象数据库	313
12.2.5	多媒体数据库	316
12.3	面向应用领域的数据库新技术	317
12.3.1	数据仓库	317
12.3.2	工程数据库	318
12.3.3	空间数据库	320
12.3.4	模糊数据库	321
12.3.5	传感器数据库	322
12.3.6	网格数据管理	323
12.3.7	微小型数据库技术	324
12.3.8	移动数据管理	325
12.4	本章小结	327
12.5	习题	328
	参考文献	329

数据库技术从诞生到现在,在不到半个世纪的时间里,形成了坚实的理论基础、成熟的商业产品和广泛的应用领域,吸引了越来越多的研究者加入,使得数据库技术成为一个研究者众多且被广泛关注的研究领域。

数据库技术是信息系统的基础和核心,在计算机软件体系中占有极为重要的地位。在过去的数十年中,数据库技术经历了层次/网状数据库、关系数据库、对象关系数据库 3 代的发展,形成了一个巨大的产业。但随着信息管理内容的不断扩展和新技术的层出不穷,数据库技术面临着前所未有的挑战。面对新的数据形式,人们提出了各种各样的数据模型,同时也提出了众多新的数据库技术,如 XML 数据管理、数据流管理、Web 数据集成、数据挖掘等技术。

本章主要介绍数据库技术的基础知识,包括数据管理技术的发展、数据库系统的结构、概念描述、数据库技术的发展,最后简要介绍 SQL Server 数据库。

1.1 数据管理技术的发展

计算机的一个重要的功能就是对信息的处理。人类认识世界、改造世界的一切有意义的活动都越来越离不开信息资源的开发、加工和利用。信息资源开发越充分、加工越合理,信息利用的效率就越高、产生的价值就越大。在了解数据库概念之前,先来看看什么是信息与数据。

1.1.1 信息与数据

信息是人类社会最重要的战略资源之一,20 世纪 40 年代后期建立起来的信息科学,已经对科学的发展产生了广泛而深远的影响。信息一般通过数据形式来表示,而计算机能够实际处理的就是各种各样的数据。

1. 信息

信息作为技术术语被广泛使用是在计算机得到普遍应用之后。在计算机发展的早期,计算机处理的对象仍沿用过去的名词,如数据、记录、报表、文字等。但随着计算机的不断发展,计算机能够处理的对象越来越多,这样就希望用一个通用的名称把所有这些处理对象包含在内,信息这一名称恰好符合这一要求。在众多对信息的定义和解释中,下面 3 种比较典型:

- (1) Shannon 的定义——信息使事物的不确定性减少或消除。
- (2) Wiener 的定义——信息是系统与外界相互交换的内容。

(3)《人工智能辞典》的定义——信息是数据中所蕴涵的意义。

下面是信息的定义:

信息是现实世界在人们头脑中的反映,它以文字、数据、符号、声音、图像等形式记录下来,可以进行传递和处理,为人们的生产、建设、管理等提供依据。

2. 数据

数据是指对客观事物进行记录并可以鉴别的符号,是对客观事物的性质、状态以及相互关系等进行记载的物理符号或这些物理符号的组合。它是可识别的、抽象的符号。它不仅指狭义上的数字,还可以是文字、图形和声音等,也是客观事物的属性、数量、位置及其相互关系等的抽象表示。例如,“0、1、2…”、“阴、雨、下降、气温”等都是数据。

数据的概念有两方面的含义:数据的内容是信息,数据的表现方式是符号。数据的格式往往和具体的计算机系统有关,随载荷它的物理设备的形式而不同。

3. 信息与数据联系和区别

信息与数据既有联系,又有区别。数据是信息的载体,而信息是数据的内涵。信息是加载在数据之上,对数据作具有含义的解释。信息依赖数据来表达,数据则生动具体地表现出信息。数据是符号,是物理性的,信息是对数据进行加工处理之后所得到的并对决策产生影响的数据,是逻辑性(观念性)的;数据是信息的表现形式,信息是数据有意义的表示。数据是信息的表达、载体,信息是数据的内涵,是形与质的关系。只有数据对实体行为产生影响才成为信息,数据只有经过解释才有意义,才能成为信息。例如“1”、“0”及独立的 1、0 均无意义,而当它表示某实体在某个地域内存在与否,它就提供了“有”、“无”信息,当用它来标识某种实体的类别时,它就提供了特征码信息。

1.1.2 数据管理技术的发展

计算机的发展历史表明,应用领域的不断拓展与深化是推动计算机学科发展的内在动力,数据管理技术也是如此。随着计算机硬件、软件技术的发展和应用领域的改变,数据管理经历了由低级到高级的发展过程,数据管理技术的发展可以大体归为 3 个阶段:人工管理阶段、文件系统阶段和数据库系统阶段。

1. 人工管理阶段

20 世纪 50 年代中期以前,计算机主要应用于科学计算。由于计算机软、硬件技术发展水平的限制,外部存储器只有磁带、卡片和纸带等,而不是类似于磁盘等可直接存取的存储设备;软件只有汇编语言,尚无数据管理方面的软件;数据处理方式基本是批处理。数据管理任务,包括存储结构、存取方法、输入输出方式等都是针对每个具体应用,由编程人员单独设计解决的。数据被包含在程序之中,程序之中存在着大量的不必要的重复、冗余数据,数据也不被长期保存。

人工管理阶段的数据管理具有以下几个特点:

(1) 计算机系统不提供对用户数据的管理功能,应用程序中只包含自己要用到的全部数据。用户编制程序时,必须全面考虑好相关的数据,包括数据的定义、存储结构以及存取方法等。程序和数据是一个不可分割的整体。数据脱离了程序就无任何存在的价值,数据无独立性。

(2) 数据不能共享。不同的程序均有各自的数据,这些数据对不同的程序通常是不相

同的,不可共享;即使不同的程序使用了相同的一组数据,这些数据也不能共享,程序中仍然需要各自加入这组数据,哪个部分都不能省略。基于这种数据的不可共享性,必然导致程序与程序之间存在大量的重复数据,浪费存储空间。

(3) 不能单独保存数据。基于数据与程序是一个整体,数据只为本程序所使用,数据只有与相应的程序一起保存才有价值,否则就毫无用处。所以,所有程序的数据均不单独保存。

这一时期数据和程序之间的关系可以用如图 1-1 所示的结构来表示。

2. 文件系统阶段

20 世纪 50 年代后期至 60 年代中期,计算机开始大量地用于数据处理工作,大量的数据存储、检索和维护工作提上议事日程,计算机不仅用于科学计算,还用于信息管理方面。外部存储器已有磁盘、磁鼓等直接存取的存储设备;软件领域出现了操作系统和高级软件,操作系统中的文件系统是专门管理外存的数据管理软件,文件是操作系统管理的重要资源之一;数据处理方式有批处理,也有联机实时处理。



图 1-1 人工管理阶段数据和程序之间的关系

文件系统阶段的数据管理具有以下几个特点:

(1) 数据以“文件”形式可长期保存在外部存储器的磁盘上。由于计算机的应用转向信息管理,因此对文件要进行大量的查询、修改和插入等操作。

(2) 数据的逻辑结构与物理结构有了区别,但比较简单。数据的逻辑结构是指呈现在用户面前的数据结构形式。数据的物理结构是指数据在计算机存储设备上的实际存储结构。程序与数据之间具有“设备独立性”,即程序只需用文件名就可与数据打交道,不必关心数据的物理位置。由操作系统的文件系统提供存取方法(读/写)。

(3) 文件组织已多样化。有索引文件、链接文件和直接存取文件等。但文件之间相互独立、缺乏联系。数据之间的联系要通过程序去构造。

(4) 数据不再属于某个特定的程序,可以重复使用,即数据面向应用。但是文件结构的设计仍然是基于特定的用途,程序基于特定的物理结构和存取方法,因此程序与数据结构之间的依赖关系并未根本改变。

(5) 对数据的操作以记录为单位。这是由于文件中只存储数据,不存储文件记录的结构描述信息。文件的建立、存取、查询、插入、删除、修改等所有操作,都要用程序来实现。

这一时期数据和程序之间的关系可以用如图 1-2 所示的结构来表示。

文件系统对计算机数据管理能力的提高虽然起了很大的作用,但随着数据管理规模的扩大,数据量急剧增加,文件系统显露出一些缺陷。主要表现在:

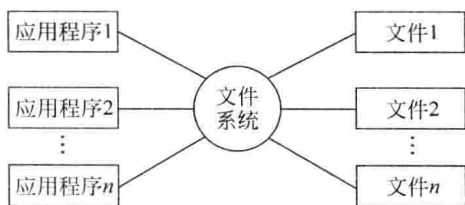


图 1-2 文件系统阶段数据和程序之间的关系

(1) 数据文件是为满足特定业务领域某一部门的专门需要而设计,数据和程序相互依赖,数据缺乏足够的独立性。

(2) 数据没有集中管理的机制,其安全性和完整性无法保障,数据维护业务仍然由应用程序