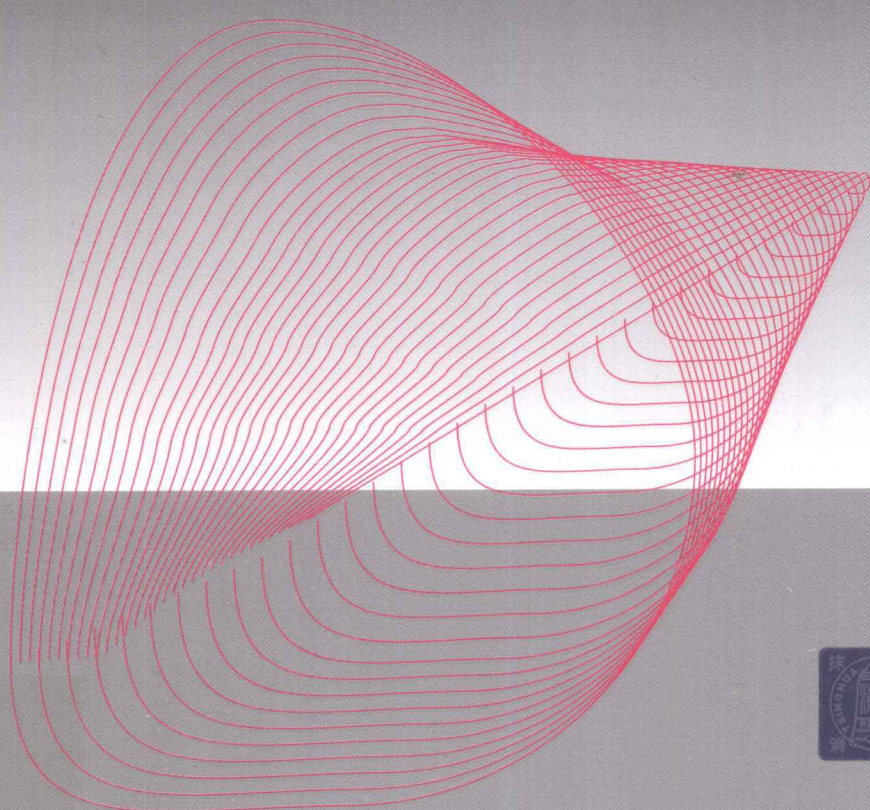


21 世纪高等学校计算机教育实用规划教材

数据库原理及应用

刘玉宝 李念峰 主编
祝海英 李纯莲 副主编



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

数据库原理及应用

刘玉宝 李念峰 主 编
祝海英 李纯莲 副主编

清华大学出版社
北京

内 容 简 介

本书是面向高等院校培养应用型本科人才的发展目标而编写的,全面介绍了数据库系统基本原理以及数据库应用程序开发技术。全书共 11 章,主要内容包括数据库系统概述、关系数据库系统理论基础、SQL Server 2005 的使用、数据库安全及维护、数据库设计、数据库访问技术、C 语言数据库应用程序开发技术、C# 和 ADO.NET 数据库应用程序开发技术、Java 数据库应用程序开发技术、数据库新技术概述、实验。本书在讲述理论的同时与 SQL Server 2005 有机结合,使理论与实践同步,同时介绍了使用 C、C# 和 Java 三种语言开发数据库应用程序的基本方法和技术,使具有不同语言基础的读者有选择性地学习,也为嵌入式系统相关专业的读者开发嵌入式软件打下良好的基础。

本书内容丰富、结构合理、实用性强。力求理论叙述严谨,应用能力培养目标明确,使读者学完本书后,能够具备数据库应用系统的独立开发能力。

本书可作为高等学校计算机及信息专业本科生的教科书,也可作为高职高专院校专科生的教科书,同时也适合具有同等文化程度的读者自学以及从事数据库应用程序的开发人员参考。为方便教师教学,本书配有教学课件和实例源代码,读者可到清华大学出版社网站(www.tup.com.cn)下载。

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

数据库原理及应用/刘玉宝,李念峰主编. —北京:清华大学出版社,2011.6

(21 世纪高等学校计算机教育实用规划教材)

ISBN 978-7-302-24742-5

I. ①数… II. ①刘… ②李… 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(2011)第 021428 号

责任编辑:闫红梅 张为民

责任校对:李建庄

责任印制:王秀菊

出版发行:清华大学出版社

地 址:北京清华大学学研大厦 A 座

<http://www.tup.com.cn>

邮 编:100084

社 总 机:010-62770175

邮 购:010-62786544

投稿与读者服务:010-62795954, jsjic@tup.tsinghua.edu.cn

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

印 刷 者:北京四季青印刷厂

装 订 者:三河市溧源装订厂

经 销:全国新华书店

开 本:185×260 印 张:22.5 字 数:563 千字

版 次:2011 年 6 月第 1 版 印 次:2011 年 6 月第 1 次印刷

印 数:1~3000

定 价:33.00 元

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

数据库技术是数据管理的最新技术,是计算机科学的重要分支。从 20 世纪 60 年代中期诞生到今天,它已经应用于社会生产和生活的各个领域,目前数据库技术已成为信息系统和应用软件系统的核心技术和重要基础,而且围绕数据库技术已形成了一个巨大的软件产业。

本书是多年讲授数据库原理与数据库应用技术的一线教师结合自己的教学经验和教学体会,整理和丰富了教学讲义而编写的。本书的特点在于能够把数据库系统原理和 SQL Server 2005 以及 C、C# 和 Java 语言有机结合起来,理论叙述严谨,应用能力培养目标明确,使读者在学习过程中做到理论与实践相结合;而且,叙述力求简单明了、深入浅出,在数据库技术应用相关章节尽量避免冗长的理论叙述,而侧重于技术的应用与程序开发能力的培养,使学生学完本课程后能具备数据库应用程序开发能力,快速适应实际工作。

全书共分 11 章和 1 个附录,参考学时为 56~72 学时,读者可以根据实际情况进行适当地取舍。

第 1 章数据库系统概述,主要介绍了数据库的基本概念、数据库技术的发展、数据库系统的组成与结构、数据模型的概念与分类。

第 2 章关系数据库系统理论基础,主要介绍了关系模型、关系数据结构及形式化定义、关系代数、关系数据库标准语言 SQL 以及关系规范化理论等。

第 3 章 SQL Server 2005 的使用,主要介绍了 SQL Server 2005 系统的组成及基本特性、SQL Server 2005 安装方法、Transact-SQL 语句的使用、SQL Server 2005 数据库管理、表的管理与使用、视图的创建与管理、索引的创建与管理、存储过程与触发器的使用。

第 4 章数据库安全及维护,主要介绍了数据库安全性控制原理、使用 SQL Server 2005 实现数据库安全性控制、数据库完整性控制原理、使用 SQL Server 2005 实现数据库完整性约束、数据库恢复技术、使用 SQL Server 2005 实现数据库的备份与恢复以及并发控制等。

第 5 章数据库设计主要介绍了数据库设计的内容和特点,数据库设计的步骤、需求分析、概念结构设计、逻辑结构设计、物理结构设计,以及数据库的实施和维护等。

第 6 章数据库访问技术,主要介绍了 ODBC 工作原理及使用方法、ADO 模型的层次结构、使用 ADO 技术访问数据库的方法、ADO.NET 的体系结构的组成及工作原理、JDBC 数据库访问技术。

第 7 章 C 语言数据库应用程序开发技术,主要介绍了 C 语言嵌入式 SQL 程序开发环境、嵌入式 SQL 语句中使用的 C 变量、数据库的连接、查询和更新、SQL 通信区、游标的使用、SQLDA。

第 8 章 C# 和 ADO.NET 数据库应用程序开发技术,主要介绍了数据提供程序的选

择,数据库的连接,数据的获取,DataReader、DataSet 和 DataAdapter 的使用等。

第 9 章 Java 数据库应用程序开发技术,主要介绍了 JDBC API、SQL 和 Java 数据类型的映射关系、Java 数据库操作的基本步骤、使用 JDBC 实现对数据库的操作、JDBC 连接其他类型的数据库。

第 10 章数据库新技术概述,主要介绍了分布式数据库的概念、特点和体系结构,面向对象数据库的理论和实现方法,数据仓库技术以及数据挖掘技术等。

第 11 章实验,针对本书内容介绍了 8 个实验,通过操作练习,加深对数据库原理的掌握。

全书内容丰富,结构合理,实用性强,其中第 7、8、9 章的内容可以根据读者已掌握的语言基础进行选择性的学习。

本书由刘玉宝、李念峰担任主编,祝海英、李纯莲担任副主编,参加编写的人员还有戴银飞、王士刚、高鹏、肖治国。其中,第 11 章及附录由刘玉宝编写,第 9 章由李念峰编写,第 3 章由祝海英编写,第 1 章由李纯莲编写,第 4 章由戴银飞编写,第 7、8 章由王士刚编写,第 2、5 章由高鹏编写,第 6、10 章由肖治国编写,全书由刘玉宝统一定稿。

在本书编写的过程中得到了单位领导和同仁的热情帮助和支持,在此向他们表示衷心的感谢!

本书的编写参考了广大同行专家的著作和成果,在此也对他们表示衷心的感谢!

计算机技术日新月异,数据库技术的发展更是十分迅速,由于时间仓促,加之作者的水平有限,书中难免有疏漏和不足之处,恳请同行专家和广大读者批评指正。

编 者

2011 年 4 月

目 录

第 1 章 数据库系统概述	1
1.1 基本概念	1
1.1.1 数据	1
1.1.2 数据库	2
1.1.3 数据库管理系统	2
1.1.4 数据库系统	3
1.2 数据库技术的发展	3
1.2.1 人工管理阶段	3
1.2.2 文件系统阶段	4
1.2.3 数据库系统阶段	4
1.3 数据库系统的组成与结构	5
1.3.1 数据库系统的组成	5
1.3.2 数据库系统结构	6
1.4 数据模型	7
1.4.1 数据模型的概念	8
1.4.2 概念模型	8
1.4.3 层次模型	10
1.4.4 网状模型	10
1.4.5 关系模型	10
小结	12
习题	12
第 2 章 关系数据库系统理论基础	14
2.1 关系模型概述	14
2.2 关系数据结构及形式化定义	15
2.2.1 关系的数学定义	15
2.2.2 关系的性质	16
2.2.3 关系模式	16
2.3 关系的完整性	17
2.3.1 实体完整性	17

2.3.2	参照完整性	17
2.3.3	用户定义完整性	18
2.4	关系代数	18
2.4.1	传统的集合运算	18
2.4.2	专门的关系运算	19
2.5	关系数据库标准语言 SQL	24
2.5.1	SQL 语言基本知识	25
2.5.2	数据定义命令	26
2.5.3	数据查询语言	28
2.5.4	数据更新语言	38
2.5.5	视图	40
2.5.6	数据控制	45
2.6	关系规范化理论	47
2.6.1	问题的提出	47
2.6.2	关系的规范化	48
小结	53
习题	54
第3章	SQL Server 2005 的使用	56
3.1	SQL Server 2005 系统概述	56
3.1.1	SQL Server 2005 版本说明	57
3.1.2	Microsoft SQL Server 2005 Express Edition 简介	58
3.1.3	SQL Server 2005 安装与配置	61
3.2	Transact-SQL 语言的使用	75
3.2.1	数据定义语言	75
3.2.2	数据查询操作	77
3.2.3	添加数据操作	84
3.3	SQL Server 数据库管理	85
3.3.1	创建数据库	85
3.3.2	查看数据库信息	87
3.3.3	管理数据库	88
3.3.4	删除数据库	89
3.4	表的管理与使用	89
3.4.1	数据库中表的创建	89
3.4.2	数据库中表的删除	90
3.4.3	数据库中表的修改	91
3.4.4	查看表的属性	91
3.4.5	表的重命名	91
3.4.6	编辑维护表格数据	91

3.5	视图及其应用	95
3.5.1	视图的创建	95
3.5.2	视图的修改和删除	96
3.6	索引的创建与管理	98
3.6.1	创建索引	99
3.6.2	查看、修改和删除索引	101
3.7	存储过程与触发器	101
3.7.1	存储过程概述	101
3.7.2	存储过程的使用与管理	101
3.7.3	触发器概述	110
3.7.4	触发器的使用与管理	110
	小结	114
	习题	114
第4章	数据库安全及维护	120
4.1	数据库安全性	120
4.1.1	安全性概述	120
4.1.2	数据库安全性控制	121
4.1.3	SQL Server 2005 实现数据库安全性	123
4.2	数据库完整性	133
4.2.1	完整性约束条件	133
4.2.2	完整性控制	134
4.2.3	SQL Server 2005 实现数据库完整性	135
4.3	数据库恢复技术	144
4.3.1	事务的基本概念	144
4.3.2	故障的种类	145
4.3.3	转储和恢复	146
4.3.4	日志文件	148
4.3.5	SQL Server 2005 实现数据库的备份与恢复	149
4.4	并发控制	153
4.4.1	并发控制概述	153
4.4.2	封锁	154
	小结	157
	习题	157
第5章	数据库设计	160
5.1	数据库设计概述	160
5.1.1	数据库设计内容	160
5.1.2	数据库设计特点	161

5.2	数据库设计步骤	161
5.2.1	需求分析	162
5.2.2	概念结构设计	166
5.2.3	逻辑结构设计	171
5.2.4	物理结构设计	175
5.2.5	数据库的实施和维护	177
	小结	180
	习题	180
第 6 章	数据库访问技术	182
6.1	ODBC 的使用	182
6.1.1	ODBC 概述	182
6.1.2	ODBC 数据源的配置	183
6.2	ADO 的使用	186
6.2.1	ADO 技术概述	186
6.2.2	使用 ADO 技术访问数据库举例	189
6.3	ADO.NET 简介	191
6.3.1	ADO.NET 技术的设计目标	191
6.3.2	ADO.NET 的体系结构	192
6.3.3	ADO.NET 数据对象介绍	198
6.4	JDBC 技术介绍	199
6.4.1	JDBC 概述	199
6.4.2	JDBC 驱动程序	201
6.4.3	JDBC 常用类	202
	小结	203
	习题	203
第 7 章	C 语言数据库应用程序开发技术	205
7.1	嵌入式 SQL 语句	205
7.1.1	C 语言嵌入式 SQL 程序开发环境	206
7.1.2	第一个 C 语言嵌入式 SQL 程序	207
7.2	静态 SQL 语句	210
7.2.1	声明嵌入式 SQL 语句中使用的 C 变量	210
7.2.2	连接数据库	212
7.2.3	数据的查询和更新	213
7.2.4	SQL 通信区	215
7.3	动态 SQL 语句	217
7.3.1	动态修改	217
7.3.2	动态游标	218

7.3.3	SQLDA	220
7.4	应用实例	222
	小结	228
	习题	228
第 8 章	C# 和 ADO.NET 数据库应用程序开发技术	230
8.1	数据库的连接	230
8.1.1	SqlConnection 的使用	232
8.1.2	OleDbConnection 的使用	234
8.1.3	OdbcConnection 的使用	234
8.1.4	OracleConnection 的使用	235
8.2	数据的获取	238
8.2.1	创建 Command 对象	238
8.2.2	执行命令	238
8.2.3	参数化查询	243
8.2.4	执行存储过程	245
8.3	DataReader 的使用	247
8.3.1	DataReader 简介	247
8.3.2	使用 DataReader 读取数据	248
8.3.3	在 DataReader 中使用多个结果集	250
8.4	DataSet 和 DataAdapter 的使用	251
8.4.1	DataSet 简介	251
8.4.2	DataAdapter 简介	252
8.4.3	利用 DataSet 和 DataAdapter 访问数据	253
8.4.4	类型和无类型 DataSet	255
8.5	性能	256
	小结	256
	习题	256
第 9 章	Java 数据库应用程序开发技术	258
9.1	JDBC API 介绍	258
9.2	SQL 和 Java 之间的映射关系	259
9.3	JDBC 编程	260
9.3.1	数据库操作基本步骤	261
9.3.2	JDBC 数据库操作实现	262
9.4	连接其他类型数据库	284
9.4.1	连接 Oracle 数据库	284
9.4.2	连接 Mysql 数据库	285
9.4.3	连接 SQL Server 数据库	286

9.4.4 连接 Access 数据库	287
小结	290
习题	290
第 10 章 数据库新技术概述	293
10.1 分布式数据库	293
10.1.1 分布式数据库系统的概念	294
10.1.2 分布式数据库系统的特点	295
10.1.3 分布式数据库系统的体系结构	296
10.1.4 分布式数据库系统的发展前景	297
10.2 面向对象数据库	297
10.2.1 面向对象数据模型	297
10.2.2 面向对象数据库建模	301
10.2.3 对象-关系数据库	304
10.3 数据仓库	305
10.3.1 数据仓库的定义与特征	306
10.3.2 数据仓库系统的体系结构	307
10.3.3 数据仓库的数据库模式	309
10.3.4 数据仓库工具	310
10.4 数据挖掘技术	311
10.4.1 数据挖掘的主要功能	312
10.4.2 数据挖掘的方法及工具	312
10.4.3 数据挖掘的实施步骤	313
10.4.4 数据挖掘应用现状	314
10.4.5 数据挖掘中存在的问题	314
小结	314
习题	315
第 11 章 实验	317
实验 1 SQL Server 2005 的安装及其管理工具的使用	317
实验 2 创建数据库和表	320
实验 3 表数据的插入、修改和删除	324
实验 4 数据查询	326
实验 5 存储过程和触发器的使用	328
实验 6 安全性控制	330
实验 7 数据完整性	333
实验 8 数据库备份与还原	335
附录 A 课程设计指导书	338

内容提要

- 数据库的基本概念;
- 数据库技术的发展;
- 数据库系统的组成与结构;
- 数据模型的概念与分类。

数据库是数据管理的最新技术,是计算机科学的重要分支。今天,信息资源已成为各个部门的重要财富和资源。建立一个满足各级部门信息处理要求的行之有效的信息系统也成为企业或组织生存和发展的重要条件。因此,作为信息系统核心和基础的数据库技术得到越来越广泛的应用,从小型单项事务处理系统到大型信息系统,从联机事务处理到联机分析处理,从一般企业管理到计算机辅助设计与制造(CAD/CAM)、计算机集成制造系统(CIMS)、办公信息系统(OIS)、地理信息系统(GIS)等,越来越多新的用户采用数据库存储和处理信息资源。对于一个国家来说,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其信息化程度的重要标志。

1.1 基本概念

数据库技术涉及许多基本概念,主要包括数据、数据处理、数据库、数据库管理系统以及数据库系统等。

1.1.1 数据

数据是数据库中存储的基本对象。数据在大多数人头脑中的第一个反应就是数字。其实数字只是最简单的一种数据,是数据的一种传统和狭义的理解。从广义上理解,数据的种类有很多,包括文字、图形、图像、声音、学生的档案记录、货物的运输情况等,这些都是数据。

数据就是描述事物的符号记录。描述事物的符号可以是数字,也可以是文字、图形、图像、声音、语言等,数据有多种表现形式,都可以经过数字化后存入计算机。

为了了解世界、交流信息,人们需要描述这些事物。在日常生活中,直接用自然语言(如汉语)描述。在计算机中,为了存储和处理这些事物,就要抽出这些事物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、出生日期、籍贯、所在系别、入学时间,那么可以这样描述:

(王晓楠,女,22,1988,北京,英语系,2007)

因此这里的学生记录就是数据。对于上面这条学生记录,了解其含义的人会得到如下

信息：王晓楠是个大学生，1988年出生，女，北京人，2007年考入英语系；而不了解其语义的人则无法理解其含义。可见，数据的形式还不能完全表达其内容，需要经过解释。所以数据和关于数据的解释是不可分的，数据的解释是指对数据含义的说明，数据的含义称为数据的语义，数据与其语义是不可分的。

1.1.2 数据库

数据库，顾名思义，是存放数据的仓库。只不过这个仓库是在计算机存储设备上，而且数据是按一定的格式存放的。人们收集并抽取应用所需要的大量数据之后，应将其保存起来以供进一步加工处理，进一步抽取有用信息。在科学技术飞速发展的今天，人们的视野越来越广，数据量急剧增加。过去人们把数据存放在文件柜里，现在人们借助计算机和数据库技术科学地保存和管理大量的复杂的数据，以便能方便而充分地利用这些宝贵的信息资源。

所谓数据库是指长期储存在计算机内的、有组织的、可共享的数据集合。数据库中的数据按一定的数据模型组织、描述和存储，具有较小的冗余度、较高的数据独立性和易扩展性，并可以为各种用户共享。

数据库中的数据具有如下特点：

- (1) 按一定的数据模型组织、描述和存储；
- (2) 具有较小的冗余度；
- (3) 具有较高的数据独立性和易扩展性；
- (4) 可为各种用户共享。

更加直观地，可以把数据库理解为存放数据的仓库。只不过这个仓库是在计算机的大容量存储器上，比如硬盘。数据库中的数据按一定的格式存放，便于查找。

1.1.3 数据库管理系统

数据库管理系统是数据库系统的一个重要组成部分。它是位于用户与操作系统之间的一层数据管理软件。数据库管理系统的主要任务是科学有效地组织和存储数据、高效地获取和管理数据，接受和完成用户提出的访问数据的各种请求。主要包括以下几方面的功能。

数据库管理系统的主要功能包括以下几个方面。

1. 数据定义功能

该功能提供数据定义语言(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。例如，对数据库、表、索引进行定义。

2. 数据操纵功能

该功能提供数据操纵语言(Data Manipulation Language, DML)，用户通过它可以实现对数据库的基本操作。例如，对表中数据的查询、插入、删除和修改等。

3. 数据库运行控制功能

该功能包括并发控制(即处理多个用户同时使用某些数据时可能产生的问题)、安全性检查、完整性约束条件的检查和执行、数据库的内部维护(例如，索引的自动维护)等。这是数据库管理系统的核心部分。数据库在建立、运用和维护时所有操作都要由这些控制程序统一管理、统一控制，以保证数据的安全性、完整性、多用户对数据的并发使用及发生故障后

的系统恢复。

4. 数据库的建立和维护功能

该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,数据库的重新组织功能和性能监视、分析功能等。这些功能通常是由一些实用程序完成的。它是数据库管理系统的一个重要组成部分。

1.1.4 数据库系统

数据库系统是应用数据库技术进行数据管理的计算机系统,它由计算机硬件系统、软件系统、数据和用户组成,其中软件系统包括操作系统、数据库管理系统和应用程序系统,属于应用平台。

与文件系统相比,数据库系统具有以下特点:

(1) 数据的结构化。同一数据库中的数据整体上按照一定的结构形式存储,但数据文件之间相互有联系。

(2) 最小的冗余度。数据库系统中的数据是集成化的,只是在逻辑存储上存在重复,而在物理存储上不存在。实现数据共享后,将消除不必要的重复。但有时可保留少量冗余以提高查询效率。

(3) 数据的共享。不同的用户可以使用同一数据库资源,这样节约了存储空间。

(4) 数据与程序独立。数据按照某种规则,以能反映数据之间内在联系的形式组织在库文件中,数据不会受到应用程序变化的影响,数据的变动也不会影响应用程序。

(5) 数据的安全性和完整性。数据安全指数据的保密,数据的完整性指数据的正确性、有效性和相容性。数据库系统提供了管理和控制数据的各种操作命令及程序设计语言,使用户控制数据库,例如设置口令来保证数据安全,防止数据被破坏或窃取。

1.2 数据库技术的发展

数据库技术是数据处理中的一门技术,数据处理是指对各种类型的数据进行收集、存储、分类、计算或加工、检索、传输、维护的一系列操作。这样的操作可以达到两个目的,第一是从大量的、原始的数据中抽取、推导出对人们有价值的信息以作为行动和决策的依据;第二是为了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宝贵的信息资源。

随着计算机硬件、软件技术和计算机应用范围的发展,数据处理的主要工作已不再是计算,而是进行管理。多年来数据库管理技术大致经历了人工管理阶段、文件系统阶段和数据库系统阶段。

1.2.1 人工管理阶段

20世纪50年代以前,计算机主要用于数值计算。这一时期的数据,数据量小,无结构,由用户直接管理,且数据间缺乏逻辑组织,由于是面向应用程序的,数据缺乏独立性,应用程序与其处理的数据结合成一个整体。程序与数据的关系如图1-1所示。

在人工管理阶段,外存只有纸带、卡片、磁带,并没有磁盘等直接存取的存储设备。这一

时期,还未形成软件的整体概念,没有操作系统,没有管理数据的软件。

特点:

(1) 数据不保存。

(2) 应用程序管理数据。应用程序承担设计数据的逻辑结构和物理结构任务。

(3) 数据不能共享。一组数据只能对应一个程序。

(4) 数据不具有独立性。数据的逻辑或物理结构改变,应用程序随之改变。

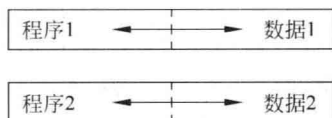


图 1-1 人工管理阶段程序与数据的关系

1.2.2 文件系统阶段

在 20 世纪 50 年代后期到 60 年代中期,这一时期程序与数据的关系如图 1-2 所示。

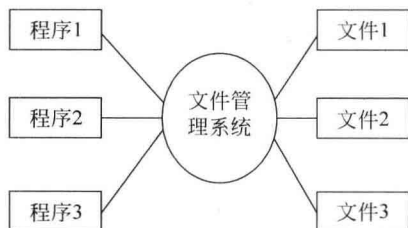


图 1-2 文件系统阶段程序与数据的关系

这一时期,有了磁盘、磁鼓等直接存取存储设备,操作系统中有了专门的数据管理软件——文件系统。

特点:

(1) 数据可以长期保存。

(2) 由文件系统进行数据管理。数据按文件名访问,按记录进行存取,可以对文件进行修改、插入和删除操作。

(3) 数据共享性差,冗余度大。

① 一个文件对应一个应用程序。

② 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。

(4) 数据独立性差。

在数据文件中常涉及下列术语:

(1) 数据项:描述事物性质的最小单位。

(2) 记录:若干数据项的集合,一个记录表达一个具体事物。

(3) 文件:若干记录的集合。

1.2.3 数据库系统阶段

20 世纪 60 年代后期至今,程序与数据的关系如图 1-3 所示。

这一时期,出现了大容量磁盘,且价格下降。同时,软件价格上升,编制、维护系统软件及应用程序的成本相对增加,因此出现了统一管理数据的专门软件——数据库管理系统。

特点:

(1) 数据结构化。数据库系统与文件系统是有根本区别的。对于文件系统来讲,相互独立的文件的记录内部是有结构的,而数据库系统主要实现整体数据的结构化。

(2) 数据的共享性高,冗余度低,易扩充。

① 数据可以被多个用户、多个应用共享

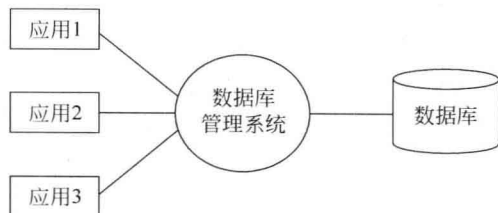


图 1-3 数据库系统阶段程序与数据的关系