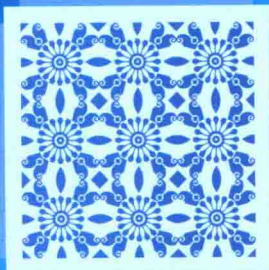


DATABASE PRINCIPLES
AND APPLICATIONS 3RD EDITION

数据库原理与应用

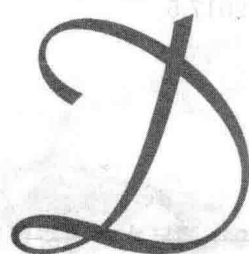
第3版

何玉洁 编著



机械工业出版社
China Machine Press

高等院校计算机教材系列



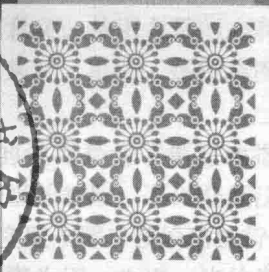
DATA

AND APPLICATIONS 3RD EDITION

数据库原理与应用

第3版

何玉洁 编著



机械工业出版社
China Machine Press

图书在版编目 (CIP) 数据

数据库原理与应用 / 何玉洁编著. —3 版. —北京: 机械工业出版社, 2017.5
(高等院校计算机教材系列)

ISBN 978-7-111-56827-8

I. 数… II. 何… III. 数据库系统—高等学校—教材 IV. TP311.13

中国版本图书馆 CIP 数据核字 (2017) 第 103094 号

本书主要介绍以下内容: 数据库理论, 包括数据模型、关系代数、关系数据库规范化理论、数据库设计、事务与并发控制; 数据库相关的一些内容, 包括创建数据库、创建数据库对象、数据查询、数据修改等; 数据库的维护性工作, 包括安全管理、备份和恢复数据库; 处理非结构化数据的新型数据库 NoSQL 的一些基本概念。

本书采用的实践平台为 SQL Server 2012, 该软件具有界面友好、使用方便、功能全面的特点, 非常适合学生作为数据库实践平台使用。

本书内容全面、实例丰富, 并为教师配备了电子教案, 方便教师开展教学工作。本书可作为高等院校计算机专业以及信息管理等相关专业本科生的数据库教材, 也可作为相关人员学习数据库知识的参考书。

出版发行: 机械工业出版社 (北京市西城区百万庄大街 22 号 邮政编码: 100037)

责任编辑: 余 洁

责任校对: 殷 虹

印 刷: 北京诚信伟业印刷有限公司

版 次: 2017 年 6 月第 3 版第 1 次印刷

开 本: 185mm×260mm 1/16

印 张: 19.25

书 号: ISBN 978-7-111-56827-8

定 价: 39.00 元

凡购本书, 如有缺页、倒页、脱页, 由本社发行部调换

客服热线: (010) 88378991 88361066

投稿热线: (010) 88379604

购书热线: (010) 68326294 88379649 68995259

读者信箱: hzsj@hzbook.com

版权所有·侵权必究

封底无防伪标均为盗版

本书法律顾问: 北京大成律师事务所 韩光 / 邹晓东

前言

数据库技术起源于 20 世纪 60 年代末, 经过几十余年的迅速发展, 已经形成一套较完整的理论体系, 产生了一大批商用软件产品。随着数据库技术的推广使用, 计算机应用已深入到国民经济和社会生活的各个领域, 这些应用一般都以数据库技术及其应用为基础和核心。因此, 数据库技术与操作系统一起构成信息处理的平台已成为业界的共识。在计算机应用中, 数据存储和数据处理是计算机最基本的功能, 数据库技术为人们提供了科学和高效地管理数据的方法。从某种意义上讲, 数据库技术的教学成为计算机专业教学的重中之重, 数据库课程也成为很多高校计算机专业的重点核心课程。目前市场上数据库类的教科书非常之多, 每本书各有其特色, 本书博采众家之所长, 在完整包括数据库基础理论知识的同时, 加入了将数据库知识与具体数据库管理系统结合的内容, 以方便学生在实践中更好地掌握所学知识。

本书具有如下特色。

- 内容安排求全、求新。本教材从数据库基础理论、数据库设计、数据库发展、数据库实践几个方面全面阐述了数据库技术的应用体系。在选择实践平台时, 充分考虑软件的流行性和易获得性, 后台数据库管理系统选用的是 SQL Server 2012, 它是目前应用范围广泛且功能完善、操作界面友好的数据库管理系统。
- 理论阐述求精、求易。数据库基础理论较为抽象, 但又是实践的基础, 没有扎实的基本功是无法灵活运用并付诸实践的。因而基础理论的教学历来是重点和难点。在理论阐释方面, 本书力求深入浅出, 突出概念和技术的直观意义, 并用大量图表和示例帮助理解, 启发思维, 使读者不仅能深刻理解相关理论的来源、思路、适用范围和条件, 并能灵活运用, 举一反三。
- 理论实践丝丝相扣。知之明也, 因知进行, 理论和技术的学习是为了更好地指导实践。本书的每部分内容根据相关理论 and 应用需求进行了精当的选取, 不以全面泛泛取胜, 但求精而实用。本书不但以图例的形式细致地描述了实践步骤, 还给出执行结果, 使学生能够以行验知, 以行证知, 最后达到知行并进, 相资为用, 为进一步的学习和实践打下良好的基础。同时, 各章后都有大量的习题, 供读者验证自己对知识的掌握程度。在实践部分除概念题之外, 还附有上机练习题, 以方便读者上机实践。

相对于第 2 版, 第 3 版主要修订的内容如下。

1) 删去了第 2 版中客户端编程 (ASP.NET) 部分的全部内容。主要是基于这样的考虑: 一方面, 学时的减少, 使得有些内容不得不放弃; 另一方面, 客户端访问数据库的技术不断发展, 一般高校都开设有专门介绍客户端应用编程的课程, 学生在这些课程中学习数据库应用编程的新知识更加合适。

2) 将实践平台从 SQL Server 2005 升级到 SQL Server 2012。

3) 将数据类型全部更新为 SQL Server 2012 所支持的。

4) 将第 2 版的第 10 章“SQL Server 基础”移至第 4 章, 以更利于讲解和实践。

5) 将“主码”改为“主键”, “外码”改为“外键”, “候选码”改为“候选键”, 更符合

当下普遍使用的术语。

6) 将第2版的第11章“创建数据库”内容移至第4章,使得知识内容更加合理(先建库,再建表)。

7) 交换了“数据库设计”与“事务与并发控制”两章顺序,更利于对数据库设计中事务设计的理解。

8) 在“数据操作语句”部分增加了两项内容:将查询结果保存到新表中;查询结果的并、交、差运算。

9) 增加了对NoSQL数据库的介绍。

作者在修订本书过程中得到了机械工业出版社华章公司姚蕾等人的大力支持和鼓励,是他们认真的工作态度以及一直以来的热情帮助,鼓励着我坚持完成此教材的修订工作。在此,对机械工业出版社的全体人员表示诚挚的感谢。同时非常感谢我们数据库课程组的全体同仁:殷旭、谷葆春、李宝安、岳清、张良、刘京志、张鸿斌、梁琦、韩麦燕老师。最后感谢我的学生们,是他们对知识的渴求,对教师的尊重让我感受到了自己的责任和价值。师者之尊,缘自“用心”。

真诚地希望读者和同行们对本书提出宝贵的意见。我深知教学探索的道路没有止境,教师是我的职业,但是在人生的道路上我永远是一名学生。

何玉洁

2017年3月

教学建议

教学章节	教学要求	课时
第 1 章 数据库概述	理解数据库管理与文件管理的区别 理解数据独立性的概念 掌握数据库管理系统的组成	2
第 2 章 数据模型与数据库系统结构	理解数据模型的概念 掌握概念层数据模型 理解关系数据模型及其特点 掌握数据库三级模式结构及两个数据独立性	2
第 3 章 关系数据库	掌握关系数据模型、关系的含义、关系操作 了解数据完整性约束在数据库系统中的作用	2
	掌握传统的集合运算	1
	掌握专门的关系运算	2
第 4 章 SQL Server 2012 基础	了解 SQL Server 2012 的主要组件和工具 了解 SQL Server 2012 的安装过程	1
	了解 SQL Server 2012 基本工具的使用 掌握数据库的组成和创建方法	2
第 5 章 数据类型及关系表创建	了解 SQL 的发展 理解 SQL 的特点 理解主要的数据类型	2
	掌握基本表的定义、删除和修改语句	2
	理解数据完整性约束的含义 掌握数据完整性约束的定义语句	2
第 6 章 数据操作语句	掌握单表查询语句 掌握多表连接查询	4
	掌握子查询语句及其作用	2
	理解将查询结果保存到新表中 理解查询结果的并、交、差运算	2
	掌握数据插入语句 掌握数据更新语句 掌握数据删除语句	2
第 7 章 索引和视图	掌握索引的概念及优点 掌握索引的存储结构和分类 掌握索引的创建、修改和删除	2
	了解视图的概念及优点 掌握视图的创建、修改和删除 掌握通过视图修改数据表的数据	2
第 8 章 关系数据库规范化理论	掌握函数依赖的概念 掌握第一范式、第二范式和第三范式的概念 掌握模式分解的方法	4

(续)

教学章节	教学要求	课时
第9章 事务与并发控制	掌握事务的4个基本特征 理解并发控制的基本概念 理解事务相互干扰的几种情况 掌握共享锁、排他锁的含义 理解三个封锁协议 理解两段锁的含义及作用	4
第10章 数据库设计	理解数据库设计步骤 了解需求分析的主要方法 掌握概念结构设计和逻辑结构设计方法 了解物理结构设计方法 了解数据库的行为设计 了解数据库的实施和运行、维护	6 ~ 8
第11章 存储过程和触发器	理解 SQL 中的变量和流程控制语句	2
	理解存储过程的作用 掌握存储过程的创建、修改、删除方法 理解触发器的作用和分类	2
	掌握创建后触发型触发器的方法 理解创建前触发型触发器的方法 掌握删除、修改触发器的方法	2
第12章 函数和游标	理解系统内置的主要函数	2
	理解三种用户自定义函数的定义方法 理解函数的修改、删除方法	2
	理解游标的概念和作用 理解游标的定义步骤和方法	2
第13章 安全管理	理解 SQL Server 安全认证过程 了解在 SQL Server 中创建登录名、数据库用户及管理权限的方法 掌握授权、收权语句 理解角色的作用 了解在 SQL Server 中创建角色及为角色授权的方法 了解在 SQL Server 中为角色添加、删除成员的方法	4
第14章 备份和恢复数据库	掌握数据库的备份策略 了解备份设备的概念 理解备份和恢复的方法	2
第15章(选讲) NoSQL 数据库	了解 NoSQL 数据库的基本概念 了解 NoSQL 数据库的常见分类 了解目前的一些 NoSQL 数据库 了解 NoSQL 数据库发展现状及挑战	2
附录(选讲) 数据库分析与设计示例		2
总课时		68 ~ 70

说明:

- 1) 建议课堂教学全部在多媒体机房内完成, 实现“讲-练”结合。
- 2) 建议教学分为核心知识技能模块(前14章的内容)和技能提高模块(第15章、附录内容), 其中核心知识技能模块建议教学学时为64~68, 技能提高模块建议学时为4, 不同学校可以根据各自的教学要求和计划学时数对教学内容进行取舍。

推荐阅读



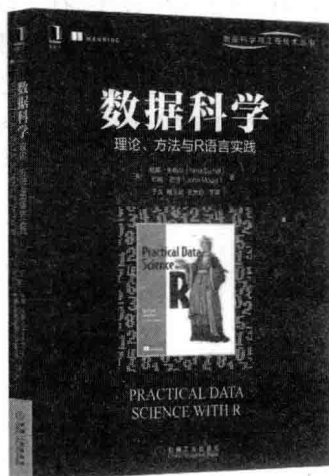
数据挖掘与商务分析：R语言

作者：约翰尼斯·莱道尔特 ISBN: 978-7-111-54940-6 定价：69.00元



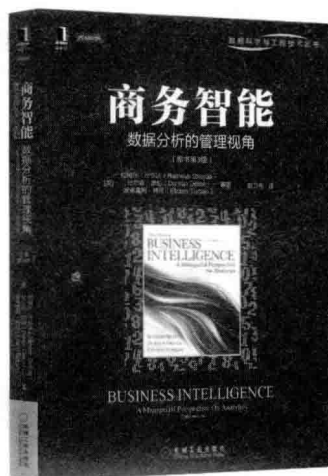
统计学习导论——基于R应用

作者：加雷斯·詹姆斯等 ISBN: 978-7-111-49771-4 定价：79.00元



数据科学：理论、方法与R语言实践

作者：尼娜·朱梅尔等 ISBN: 978-7-111-52926-2 定价：69.00元



商务智能：数据分析的管理视角（原书第3版）

作者：拉姆什·沙尔达等 ISBN: 978-7-111-49439-3 定价：69.00元

目 录

前言	
教学建议	
第1章 数据库概述	1
1.1 一些基本概念	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 数据管理技术的发展	3
1.2.1 文件管理	3
1.2.2 数据库管理	6
1.3 数据独立性	9
1.4 数据库系统的组成	10
小结	11
习题	11
第2章 数据模型与数据库系统结构	12
2.1 数据和数据模型	12
2.1.1 数据与信息	12
2.1.2 数据模型	13
2.2 概念层数据模型	14
2.2.1 基本概念	14
2.2.2 实体-联系模型	15
2.3 组织层数据模型	17
2.3.1 层次数据模型	17
2.3.2 网状数据模型	19
2.3.3 关系数据模型	20
2.4 数据库系统结构	21
2.4.1 模式的基本概念	21
2.4.2 三级模式结构	22
2.4.3 模式映像与数据独立性	23
2.5 数据库管理系统	24
小结	26
习题	26
第3章 关系数据库	27
3.1 关系数据模型的组成	27
3.1.1 关系数据结构	27
3.1.2 关系操作	27
3.1.3 数据完整性约束	29
3.2 关系模型的基本术语	30
3.3 关系模型的形式化定义	31
3.3.1 形式化定义	31
3.3.2 对关系的限定	33
3.4 关系模型的完整性约束	33
3.4.1 实体完整性	33
3.4.2 参照完整性	35
3.4.3 用户定义的完整性	36
3.5 关系代数	36
3.5.1 传统的集合运算	37
3.5.2 专门的关系运算	39
3.5.3 关系代数操作总结	45
小结	46
习题	46
第4章 SQL Server 2012 基础	48
4.1 SQL Server 2012 预备知识	48
4.1.1 主要服务器组件	48
4.1.2 管理工具	49
4.1.3 主要版本	49

4.1.4 软 / 硬件要求	50	第 6 章 数据操作语句	83
4.1.5 实例	51	6.1 数据查询	83
4.2 安装 SQL Server 2012	52	6.1.1 查询语句的基本结构	84
4.3 管理工具	60	6.1.2 单表查询	84
4.3.1 SQL Server 配置管理器	60	6.1.3 多表连接查询	101
4.3.2 SQL Server Management	62	6.1.4 使用 TOP 限制结果集	107
4.4 创建数据库	65	6.1.5 子查询	109
4.4.1 SQL Server 数据库的组成	65	6.1.6 将查询结果保存到新表中	118
4.4.2 数据库文件的属性	66	6.1.7 查询结果的并、交、差	119
4.4.3 创建数据库的图形化方法	66	6.2 数据更改	122
小结	69	6.2.1 插入数据	122
习题	70	6.2.2 更新数据	122
上机练习	70	6.2.3 删除数据	124
第 5 章 数据类型及关系表创建	71	小结	125
5.1 基本概念	71	习题	126
5.1.1 SQL 的发展	71	上机练习	126
5.1.2 SQL 的特点	71	第 7 章 索引和视图	128
5.1.3 SQL 功能概述	72	7.1 索引	128
5.2 SQL Server 提供的主要数据	72	7.1.1 基本概念	128
类型	72	7.1.2 索引的存储结构及分类	129
5.2.1 数字类型	73	7.1.3 创建和删除索引	134
5.2.2 字符串类型	73	7.2 视图	135
5.2.3 日期和时间类型	75	7.2.1 基本概念	135
5.3 关系表的创建与维护	75	7.2.2 定义视图	135
5.3.1 创建关系表	76	7.2.3 通过视图查询数据	138
5.3.2 删除关系表	78	7.2.4 修改和删除视图	139
5.3.3 修改关系表	78	7.2.5 视图的作用	140
5.4 数据完整性	79	小结	141
5.4.1 完整性约束条件的作用对象	79	习题	141
5.4.2 实现数据完整性	79	第 8 章 关系数据库规范化理论	143
小结	81	8.1 函数依赖	143
习题	81	8.1.1 基本概念	143
上机练习	82	8.1.2 一些术语和符号	144

8.1.3 为什么要讨论函数依赖	144
8.2 关系规范化	146
8.2.1 关系模式中的键	146
8.2.2 范式	147
8.3 关系模式的分解准则	151
小结	153
习题	154

第9章 事务与并发控制

9.1 事务	155
9.1.1 基本概念	155
9.1.2 事务的特征	156
9.1.3 事务处理模型	156
9.2 并发控制	157
9.2.1 并发控制概述	157
9.2.2 并发控制措施	159
9.2.3 封锁协议	160
9.2.4 活锁和死锁	162
9.2.5 并发调度的可串行性	164
9.2.6 两段锁协议	165
小结	166
习题	166

第10章 数据库设计

10.1 数据库设计概述	168
10.1.1 数据库设计的特点	169
10.1.2 数据库设计方法概述	169
10.1.3 数据库设计的全过程	170
10.2 数据库需求分析	171
10.2.1 需求分析的任务	171
10.2.2 需求分析的方法	172
10.3 数据库结构设计	173
10.3.1 概念结构设计	173
10.3.2 逻辑结构设计	176
10.3.3 物理结构设计	180
10.4 数据库行为设计	182

10.4.1 功能分析	183
10.4.2 功能设计	183
10.4.3 事务设计	184
10.5 数据库实施	184
10.6 数据库的运行和维护	185
小结	186
习题	186

第11章 存储过程和触发器

11.1 变量及流程控制语句	188
11.1.1 变量	188
11.1.2 流程控制语句	189
11.2 存储过程	190
11.2.1 基本概念	190
11.2.2 创建和执行存储过程	191
11.2.3 查看和修改存储过程	196
11.2.4 删除存储过程	196
11.3 触发器	197
11.3.1 创建触发器	198
11.3.2 后触发型触发器	198
11.3.3 前触发型触发器	200
11.3.4 查看和修改触发器	200
11.3.5 删除触发器	201
小结	201
习题	201

上机练习

第12章 函数和游标

12.1 系统提供的内置函数	203
12.1.1 日期和时间函数	203
12.1.2 字符串函数	207
12.1.3 类型转换函数	209
12.2 用户自定义函数	210
12.2.1 基本概念	210
12.2.2 标量函数	211
12.2.3 内联表值函数	213

12.2.4 多语句表值函数	214	上机练习	253
12.2.5 查看和修改用户自定义函数	215	第 14 章 备份和恢复数据库	254
12.2.6 删除用户自定义函数	216	14.1 备份数据库	254
12.3 游标	217	14.1.1 备份内容及备份时间	254
12.3.1 基本概念	217	14.1.2 备份设备	255
12.3.2 使用游标	217	14.1.3 SQL Server 支持的备份类型	256
12.3.3 游标示例	220	14.1.4 备份策略	258
小结	223	14.1.5 实现备份	259
习题	223	14.2 恢复数据库	264
上机练习	223	14.2.1 恢复的顺序	264
第 13 章 安全管理	225	14.2.2 实现恢复	265
13.1 安全控制概述	225	小结	271
13.1.1 安全控制模型	225	习题	271
13.1.2 SQL Server 安全控制过程	226	上机练习	271
13.2 登录名	226	第 15 章 NoSQL 数据库	274
13.2.1 身份验证模式	226	15.1 NoSQL 简介	274
13.2.2 建立登录名	227	15.2 NoSQL 数据库常见分类	275
13.2.3 删除登录名	231	15.3 NoSQL 的优缺点	276
13.3 数据库用户	232	15.3.1 优点	276
13.3.1 建立数据库用户	233	15.3.2 缺点	276
13.3.2 删除数据库用户	235	15.4 目前一些常见的 NoSQL 数据库	276
13.4 权限的种类和管理	236	15.4.1 Hypertable	276
13.4.1 权限种类及用户分类	236	15.4.2 MongoDB	277
13.4.2 权限管理	237	15.4.3 HBase	279
13.5 角色	247	15.4.4 MemcacheDB	279
13.5.1 建立用户定义的角色	247	15.4.5 Redis	280
13.5.2 为用户定义的角色授权	248	15.4.6 Tokyo Cabinet/Tokyo Tyant	280
13.5.3 为用户定义的角色添加成员	248	15.4.7 db4o	280
13.5.4 删除用户定义角色中的成员	251	15.4.8 Versant	281
小结	252	15.4.9 CouchDB	281
习题	252	15.4.10 DynamoDB	282
		15.4.11 Cassandra	282

15.4.12	Flare	282	15.5	NoSQL 数据库发展现状及挑战	284
15.4.13	Berkeley DB	283		小结	285
15.4.14	Memlink	283		习题	285
15.4.15	BaseX	283	附录	数据库分析与设计示例	286
15.4.16	Neo4j	283	参考文献		295
15.4.17	NoSQL Database	284			

第1章 数据库概述

随着管理水平的不断提高和应用范围的日益扩大,信息已成为企业的重要财富和资源。同时,作为管理信息的数据库技术也得到了很大的发展,应用领域越来越广泛。人们在不知不觉中扩展着对数据库的使用,信用卡购物,飞机、火车订票系统,商场的进货与销售、图书馆对书籍及借阅的管理等,无一不使用了数据库技术。从小型事务处理到大型信息系统,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)、地理信息系统等,数据库系统已经渗透到日常生活的方方面面,数据库中信息量的大小以及使用的程度已经成为衡量企业的信息化程度的重要标志。

数据库技术是数据管理的最新技术,其主要研究内容是如何对数据进行科学的管理,以提供可共享、安全、可靠的数据。数据库技术一般包含数据管理和数据处理两部分内容。

数据库系统本质上是一个用计算机存储数据的系统,数据库本身可以看作一个电子文件柜,也就是说,数据库是收集数据文件的仓库或容器。

本章介绍数据库的基本概念,包括数据管理的发展过程、数据库系统的组成以及使用数据库技术的一些考虑等。读者可从本章了解为什么要学习数据库技术,并为后续章节的学习做好准备。

1.1 一些基本概念

在系统地介绍数据库技术之前,首先介绍数据库中常用的一些术语和基本概念。

1.1.1 数据

数据(data)是数据库中存储的基本对象。早期的计算机系统主要用在科学计算领域,处理的数据基本是数值型数据,因此数据在人们头脑中的直觉反应就是数字。但其实数字只是数据的一种最简单的形式,是对数据的传统和狭义的理解。目前计算机的应用范围已十分广泛,因此数据种类也更加丰富,如文本、图形、图像、音频、视频、商品销售情况等都是数据。

可以将数据定义为:数据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,它们都可以经过数字化后保存在计算机中。

数据的表现形式并不一定能完全表达其内容,有些还需要经过解释才能明确其表达的含义。比如20,当解释其代表人的年龄时是20岁,当解释其代表商品的价格时,就是20元。因此,数据和数据的解释是不可分的。数据的解释是对数据演绎的说明,数据的含义称为数据的语义。

在日常生活中,人们一般直接用自然语言来描述事物,如一门课程的信息可以描述为:数据库系统基础,4个学分,第5学期开设。但在计算机中经常按如下形式描述:

(数据库系统基础, 4, 5)

以上形式是把课程名、学分、开课学期信息组织在一起, 形成一个记录, 这个记录就是描述课程的数据。这样的数据是有结构的。记录是计算机表示和存储数据的一种格式或方法。

1.1.2 数据库

数据库 (DataBase, 简称 DB), 顾名思义, 就是存放数据的仓库, 只是这个仓库是存储在计算机存储设备上的, 而且是按一定的格式存储的。

人们在收集并抽取出一个应用所需要的大量数据之后, 就希望将这些数据保存起来, 以供进一步从中得到有价值的信息, 并进行相应的加工和处理。在科学技术飞速发展的今天, 人们对数据的需求越来越多, 数据量也越来越大。最早人们把数据存放在文件柜里, 现在人们可以借助计算机和数据库技术来科学地保存和管理大量的复杂数据, 以便能方便而充分地利用宝贵的数据资源。

严格地讲, 数据库是长期存储在计算机中的、有组织的、可共享的大量数据的集合。数据库中的数据按一定的数据模型组织、描述和存储, 具有较小的数据冗余、较高的数据独立性和易扩展性, 并可为多个用户共享。

概括起来, 数据库数据具有永久存储、有组织和可共享三个基本特点。

1.1.3 数据库管理系统

在了解了数据和数据库的基本概念之后, 下一个需要了解的就是如何科学有效地组织和存储数据, 如何从大量的数据中快速地获得所需的数据以及如何对数据进行维护, 这些都是数据库管理系统 (Database Management System, 简称 DBMS) 要完成的任务。数据库管理系统是一个专门用于对数据进行管理和维护的系统软件。

数据库管理系统位于用户应用程序与操作系统软件之间, 如图 1-1 所示。数据库管理系统与操作系统一样都是计算机的基础软件, 同时也是一个非常复杂的大型系统软件, 其主要功能包括如下几个方面。



图 1-1 数据库管理系统在计算机系统的位置

1. 数据库的建立与维护功能

该功能包括创建数据库及对数据库空间的维护、数据库的备份与恢复、数据库的重组、数据库的性能监视与调整功能等。这些功能一般是通过数据库管理系统中提供的一些实用工具实现的。

2. 数据定义功能

该功能包括定义数据库中的对象, 比如表、视图、存储过程等。这些功能的实现一般是通过数据库管理系统提供的数据库定义语言 (Data Definition Language, DDL) 实现的。

3. 数据组织、存储和管理功能

为提高数据的存取效率, 数据库管理系统需要对数据进行分类存储和管理。数据库中的数据包括数据字典、用户数据和存取路径数据等。数据库管理系统要确定这些数据的存储结构、存取方法以及存储位置, 以及如何实现数据之间的关联。确定数据的组织和存储的主要目的是提高存储空间利用率和存取效率。一般的数据库管理系统都会根据数据的具体组织和存储方式提供多种数据存取方法, 比如索引查找、Hash 查找、顺序查找等。

4. 数据操作功能

数据操作功能包括对数据库数据的查询、插入、删除和更改操作。这些操作一般是通过数据库管理系统提供的数据库操作语言 (Data Manipulation Language, DML) 实现的。

5. 事务的管理和运行功能

数据库中的数据是可供多个用户同时使用的共享数据。为保证数据能够安全、可靠地运行, 数据库管理系统提供了事务管理功能。这些功能保证数据能够并发使用并且不会产生相互干扰的情况, 而且在发生故障时 (包括硬件故障和操作故障等) 能够对数据库进行正确的恢复。

6. 其他功能

其他功能包括与其他软件的网络通信功能、不同数据库管理系统间的数据传输以及互访问功能等。

1.1.4 数据库系统

数据库系统 (DataBase System, DBS) 是指在计算机中引入数据库后的系统, 一般由数据库、数据库管理系统 (及相关的实用工具)、应用程序、数据库管理员组成。为保证数据库中的数据能够正常、高效地运行, 除了数据库管理系统之外, 还需要一个 (或一些) 专门人员来对数据库进行维护, 这个专门人员就称为数据库管理员 (Database Administrator, DBA)。

一般在不引起混淆的情况下, 常常把数据库系统简称为数据库。

1.2 数据管理技术的发展

数据库技术是应数据管理任务的需要而产生和发展的。数据管理包括对数据进行分类、组织、编码、存储、检索和维护, 是数据处理的核心, 而数据处理则是对各种数据进行收集、存储、加工和传播等一系列活动的总和。

自计算机产生之后, 人们就希望用它来帮助我们对数据进行存储和管理。最初对数据的管理是以文件方式进行的, 也就是通过编写应用程序来实现对数据的存储和管理。后来, 随着数据量越来越大, 人们对数据的要求越来越多, 希望达到的目的也越来越复杂, 文件管理方式已经很难满足人们对数据的需求, 由此产生了数据库技术, 也就是用数据库来存储和管理数据。数据管理技术的发展因此也就经历了文件管理和数据库管理两个阶段。

本节介绍文件管理方式和数据库管理方式在数据管理上的主要差别。

1.2.1 文件管理

理解今日数据库特征的最好办法是了解在数据库技术产生之前, 人们是如何通过文件的方式对数据进行管理的。

20 世纪 50 年代后期到 60 年代中期, 计算机的硬件方面已经有了磁盘等直接存取的存储设备, 在软件方面, 操作系统中已经有了专门的数据管理软件, 一般称为文件管理系统。文件管理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, 可以对文件中的数据进行修改、插入和删除等操作。

在出现程序设计语言之后, 开发人员不但可以创建自己的文件并将数据保存在自己定义的文件中, 而且还可以通过编写应用程序的方式来处理文件中的数据, 即编写应用程序来定义文件的结构, 实现对文件内容的插入、删除、修改和查询操作。当然, 真正实现磁盘文件

的物理存取操作的还是操作系统中的文件管理系统，应用程序只是告诉文件管理系统对哪个文件的哪些数据进行哪些操作。我们将由开发人员定义存储数据的文件及文件结构，并借助文件管理系统的功能编写访问这些文件的应用程序，以实现用户对数据进行处理的方式称为**文件管理**。为叙述简单，在本章后面的讨论中将忽略文件管理系统，假定应用程序是直接对磁盘文件进行操作的。

如果用文件管理数据，用户必须编写应用程序来管理存储在文件中的数据，其操作模式如图 1-2 所示。

假设某学校要用文件的方式保存学生及其选课的数据，并在这些数据文件基础之上构建对学生进行管理。此系统主要实现两部分功能：学生基本信息管理和学生选课情况管理。假设教务部门管理学生选课情况，各系部管理学生基本信息。学生基本信息管理中涉及学生的基本信息数据，假设这些数据保存在 F1 文件中；学生选课情况管理涉及学生的部分基本信息、课程基本信息和学生选课信息。假设用 F2 和 F3 文件分别保存课程基本信息和学生选课信息的数据。

设 A1 为实现“学生基本信息管理”功能的应用程序，A2 为实现“学生选课管理”功能的应用程序。由于学生选课管理中要用到 F1 文件中的一些数据，为减少冗余，它将使用“学生基本信息管理”（即 F1 文件）中的数据，如图 1-3 所示（图中省略了操作系统部分）。

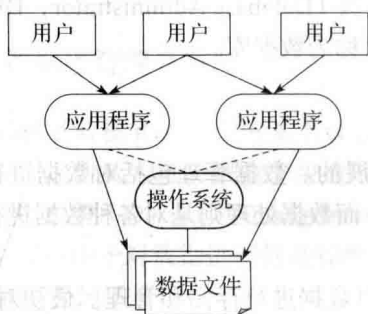


图 1-2 文件管理的操作模式



图 1-3 文件管理实现示例

假设 F1、F2 和 F3 文件分别包含如下信息：

F1 文件——学号、姓名、性别、出生日期、联系电话、所在系、专业、班号。

F2 文件——课程号、课程名、授课学期、学分、课程性质。

F3 文件——学号、姓名、所在系、专业、课程号、课程名、选课类型、选课时间、考试成绩。

我们将文件中所包含的每一个子项称为文件结构中的“字段”或“列”，将每一行数据称为一个“记录”。

“学生选课管理”的处理过程大致为：在学生选课管理中，若有学生选课，则先查 F1 文件，判断有无此学生；若有，则再访问 F2 文件，判断其所选的课程是否存在；若一切符合规则，就将学生选课信息写到 F3 文件中。

这看似很好，但仔细分析一下，就会发现直接用文件管理数据有如下缺点。

(1) 编写应用程序不方便

应用程序编写者必须清楚地了解所用文件的逻辑及物理结构，如文件中包含多少字段，每个字段的数据类型，采用何种逻辑结构和物理存储结构。操作系统只提供了打开、关闭、