

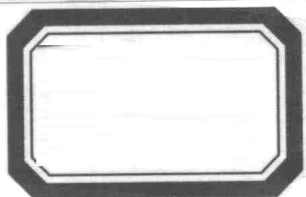
P RINCIPLES AND APPLICATIONS
OF DATABASE

数据库原理及应用

王丽艳 郑先锋 刘亮 编著



机械工业出版社
China Machine Press



计算机基础课程系列教材

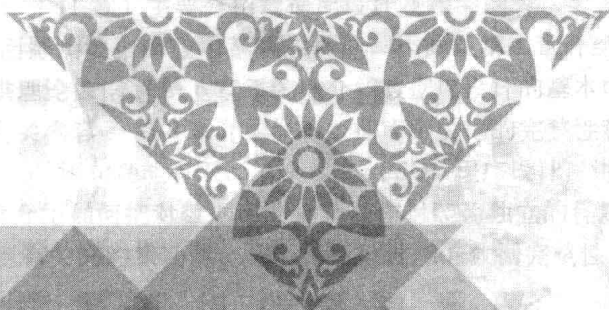
P

PRINCIPLES AND APPLICATIONS
OF DATABASE

数据库原理及应用

王丽艳 郑先锋 刘亮 编著

陈昌志 主审



机械工业出版社
China Machine Press

图书在版编目 (CIP) 数据

数据库原理及应用 / 王丽艳等编著. —北京: 机械工业出版社, 2013.2

(计算机基础课程系列教材)

ISBN 978-7-111-40997-7

I. 数… II. 王… III. 数据库系统—教材 IV. TP311.13

中国版本图书馆 CIP 数据核字 (2012) 第 315019 号

版权所有·侵权必究

封底无防伪标均为盗版

本书法律顾问 北京市展达律师事务所

本书讲授数据库基本原理, 并以 SQL Server 2008 为平台介绍数据库管理系统的应用。全书分为 4 部分: 第 1 ~ 第 4 章为第一部分, 介绍数据库的基本原理; 第 5 ~ 11 章为第二部分, 介绍数据库的设计; 第 12 ~ 15 章为第三部分, 介绍 SQL Server 2008 的使用; 第四部分给出 10 个实验和一个课程设计作为实践内容。

本书由浅入深、循序渐进地介绍各个知识点。书中提供了大量例题, 有助于读者理解概念和巩固知识, 各章还提供了一定数量的习题, 便于学生巩固所学内容。

本书可以作为各类院校相关专业及培训班的“数据库原理与应用”课程的教学用书, 也可作为计算机应用人员和计算机爱好者的自学参考书。

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: 朱秀英

北京京师印务有限公司印刷

2013 年 2 月第 1 版第 1 次印刷

185mm × 260mm · 19.25 印张

标准书号: ISBN 978-7-111-40997-7

定 价: 33.00 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: hzjsj@hzbook.com

前 言

在信息时代, 各行各业都需要使用计算机来进行信息的管理, 信息量呈现级数增长。管理数量庞大的数据, 需要数据库技术的支持, 因此数据库技术成为信息社会的重要支撑。随着数据库应用领域的日益广泛, 数据库技术本身也得到了不断的发展和完善, 对于数据库有关内容的教学也得到了各大院校相关专业的重视。

本书着重介绍数据库的基本概念、基本原理和基本应用, 力图使读者获得对数据库原理和设计的全面、深入的了解。每章后面还提供了丰富的习题, 以加深读者对各章节内容的理解, 巩固相应的知识和基本操作。具体来说, 本书将内容分成四部分: 基础篇、设计篇、应用篇和实验篇。

第一部分基础篇(第1~4章)阐述了数据库的基本概念、关系模型和SQL语句。其中, 第1章介绍数据库的基本概念、发展等内容; 第2章介绍数据模型及其要素, 并阐述了数据库的三级模式结构和数据独立性; 第3章介绍关系数据的基本概念, 并在关系代数的基础上介绍了对关系的并、交、差、投影、选择和连接等操作; 第4章对关系数据库的标准语言SQL做了详细的介绍, 使读者进一步理解数据库和数据操作本身。

第二部分设计篇(第5~11章)主要介绍数据库设计及相关内容。关系模式结构的好坏将直接影响数据库的质量, 进而影响数据库应用系统的性能。因此, 第5章介绍函数依赖和规范化理论, 为关系模型的设计提供理论依据; 第6章系统地介绍数据库设计的基本内容和基本步骤; 第7章介绍事务及并发控制的有关内容; 第8章介绍保护数据库中数据的完整性和安全性的必要性以及有关方法和措施; 单个的SQL语句功能有限, 第9章介绍SQL程序设计, 讲解如何利用SQL语句实现复杂的操作; 第10章介绍如何把对数据库的复杂操作封装为独立的程序模块, 以及如何自动调用执行的程序模块, 实现有关的约束功能, 从而保证数据库的数据完整性; 第11章介绍数据库的发展趋势和有关研究方向。

第三部分应用篇(12~15章)介绍SQL Server 2008的基础知识和应用。SQL Server 2008是Microsoft公司推出的数据库管理系统, 它是一个全面的数据库平台, 使用集成的商业智能(Business Intelligence, BI)工具提供了企业级的数据管理。第12章介绍SQL Server 2008的安装和配置; 第13章介绍如何创建、管理数据库和表; 第14章介绍SQL Server 2008中的安全管理和控制; 第15章介绍如何进行数据库的备份和恢复。

第四部分实验篇给出10个实验和一个课程设计。通过有关的实验和课程设计, 使读者把理论与实践结合起来, 提高动手能力。

本书第1~8章由王丽艳编写, 第9~12章及实验1~10由郑先锋编写, 第13~15章及课程设计由刘亮编写, 由王丽艳和郑先锋组织设计及统稿。在编写过程中, 教研室的同仁给予了编者许多帮助, 在此表示感谢。另外, 在本书的编写过程中参阅了一些著作和资料, 在此对其作者表示感谢。

由于水平有限, 书中难免有疏漏和不足之处, 恳请读者批评指正。

编者

2012年11月

教学建议

教学章节	教学要求	课时
第 1 章 数据库系统概述	了解数据库的发展历史、特点 掌握数据库的三级模式和两级映像	2
第 2 章 数据模型	了解数据的三个范畴, 概念模型和组织层数据模型的特点 掌握数据模型的组成要素, E-R 模型的特点	2
第 3 章 关系模型	了解关系模型的数据结构 掌握关系的完整性含义, 关系代数的有关操作	2
第 4 章 关系数据库标准语言 SQL	了解 SQL 数据库的结构、功能和特点 掌握 SQL 的数据类型, 模式的定义与删除	2
	掌握基本表的定义、修改与删除 掌握索引的建立与删除	2
	掌握数据库查询语句	7
	了解视图的作用 掌握数据更新和数据控制语句的使用 掌握视图的定义、查询和更新	2
第 5 章 数据库规范化理论	了解规范化理论的意义, 多值依赖和 4NF 的含义, 数据依赖的公理系统 掌握函数依赖的定义和分类 掌握 1NF、2NF、3NF 和 BCNF 的判别和模式分解	2
第 6 章 数据库设计	了解数据库设计的特点 掌握数据库设计的步骤和内容 掌握 E-R 模型转换的规则	2
第 7 章 事务处理技术	了解事务和并发控制的特点 掌握事务的状态变迁和处理模型 掌握并发控制的措施 掌握数据库的备份和恢复	4
第 8 章 数据库完整性和安全性	了解数据库完整性和安全性的含义 掌握定义约束、默认和规则的使用, 数据安全性控制方法	2
第 9 章 SQL 程序设计	掌握全局变量与局部变量的定义和使用 掌握临时表和表变量的区别和使用 掌握流程控制语句、表达式和函数以及游标的使用	4
第 10 章 存储过程和触发器	了解存储过程和触发器的概念 掌握存储过程的创建、修改和删除 掌握触发器的创建、查看、修改和删除	4
第 11 章 数据库技术的发展	了解数据库技术的发展历程、特点和发展趋势 了解 XML 数据库的数据和数据模型等 掌握数据仓库和数据挖掘的定义 掌握联机分析处理的定义、特征、数据模型和衡量标准等	2
第 12 章 安装与配置 SQL Server 2008	了解 SQL Server 的发展史、版本和安装要求 掌握 SQL Server 2008 的安装和配置	1
第 13 章 创建与管理数据库及表	了解 SQL Server 数据库类型 掌握数据库的创建、管理 掌握数据库快照的创建和管理 掌握数据库中基本表的创建与管理	2

(续)

教学章节	教学要求	课时
第 14 章 SQL Server 2008 安全性控制	了解 SQL Server 2008 身份验证模式 掌握两种登录账户的管理 掌握数据库用户管理 掌握角色、权限管理	4
第 15 章 SQL Server 2008 的备份和恢复	了解数据的导入和导出方法 了解备份设备的概念、恢复数据方法 掌握数据库备份和恢复方法	2
实验 1 ~ 10 课程设计	掌握 SQL Server 2008 的安装和使用 掌握数据库系统的设计	16
总课时	第 1 ~ 15 章建议课时	48
	实验篇建议课时	16

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) 建议教学分为基础知识模块(前 4 章)、数据库设计模块(第 5 ~ 11 章)、SQL Server 2008 使用模块(第 12 ~ 15 章)和实验模块, 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。

3) 第 12 ~ 15 章可以适当调整到实验环节讲解。

目 录

前 言
教学建议

第一部分 基础篇

第 1 章 数据库系统概述	1	第 3 章 关系模型	25
1.1 数据库系统基本概念	1	3.1 关系模型的数据结构	25
1.1.1 信息和数据	1	3.2 关系操作	28
1.1.2 数据管理	2	3.3 关系的完整性	28
1.1.3 数据库	2	3.3.1 实体完整性	28
1.1.4 数据库管理系统	3	3.3.2 参照完整性	29
1.1.5 数据库系统	3	3.3.3 用户定义的完整性	29
1.2 数据库分类	5	3.4 关系代数	29
1.3 数据管理技术的产生和发展	5	3.4.1 传统的集合运算	30
1.4 数据库系统结构	6	3.4.2 专门的关系运算	32
1.4.1 数据库系统模式的概念	6	3.4.3 关系代数表达式的优化策略	35
1.4.2 三级模式结构	7	3.5 关系代数的应用	36
1.4.3 映像功能与数据独立性	8	3.6 小结	36
1.4.4 应用程序的体系结构	9	习题 3	37
1.5 小结	11	第 4 章 关系数据库标准语言 SQL	38
习题 1	12	4.1 SQL 概述	38
第 2 章 数据模型	13	4.1.1 SQL 简介	38
2.1 数据的三个范畴	13	4.1.2 SQL 数据库结构	38
2.2 数据模型概念	14	4.1.3 SQL 的功能	39
2.3 数据模型的组成要素	15	4.1.4 SQL 的特点	40
2.4 概念模型	15	4.1.5 Transact-SQL	41
2.4.1 基本概念	15	4.1.6 SQL 的数据类型	41
2.4.2 E-R 模型	17	4.2 SQL 数据定义	42
2.4.3 UML	18	4.2.1 模式的定义与删除	43
2.5 组织层数据模型	19	4.2.2 基本表的定义、删除与修改	44
2.5.1 层次模型	19	4.2.3 索引的建立与删除	46
2.5.2 网状模型	20	4.3 数据查询	47
2.5.3 关系模型	21	4.3.1 单表查询	48
2.5.4 面向对象模型	22	4.3.2 连接查询	56
2.6 小结	24	4.3.3 嵌套查询	59
习题 2	24	4.3.4 集合查询	63
		4.3.5 查询语句的性能优化	64
		4.4 数据更新	65
		4.4.1 插入数据	65
		4.4.2 修改数据	67

4.4.3 删除数据	67	6.2.4 数据字典	94
4.5 SQL 数据控制	68	6.3 概念结构设计	96
4.5.1 授权语句	68	6.3.1 概念结构设计的必要性及要求	96
4.5.2 收权语句	69	6.3.2 概念结构设计的方法与步骤	97
4.6 视图	70	6.3.3 采用 E-R 模型设计概念结构的方法	99
4.6.1 视图的作用	70	6.3.4 E-R 图的高级技术	102
4.6.2 定义视图	70	6.4 逻辑结构设计	103
4.6.3 查询视图	72	6.4.1 E-R 图向关系模型的转换	104
4.6.4 更新视图	73	6.4.2 数据模型的优化	105
4.7 小结	74	6.5 物理结构设计	106
习题 4	74	6.5.1 物理结构设计的内容和方法	107
		6.5.2 评价物理结构	108
		6.6 数据库行为设计	108
		6.7 数据库实施	109
		6.7.1 加载数据	109
		6.7.2 调试和运行应用程序	110
		6.8 数据库的运行与维护	110
		6.9 小结	111
		习题 6	111
		第 7 章 事务处理技术	112
		7.1 事务	112
		7.1.1 事务的基本概念	112
		7.1.2 事务的特性	112
		7.1.3 事务的状态变迁	113
		7.1.4 SQL 事务处理模型	114
		7.2 并发控制	115
		7.2.1 并发控制概述	115
		7.2.2 事务一致性级别	116
		7.2.3 并发控制措施	117
		7.2.4 封锁协议	118
		7.2.5 活锁和死锁	119
		7.2.6 并发调度的可串行性	120
		7.2.7 两段锁协议	121
		7.2.8 封锁粒度	122
		7.3 数据库的备份与恢复	122
		7.3.1 数据库故障	123
		7.3.2 数据库备份	124
		7.3.3 数据库恢复	125
		7.4 小结	126
		习题 7	127
第二部分 设计篇			
第 5 章 数据库规范化理论	75		
5.1 概述	75		
5.2 函数依赖	76		
5.2.1 函数依赖的定义	76		
5.2.2 函数依赖的分类	76		
5.3 关系规范化	77		
5.3.1 关系模式中的码	77		
5.3.2 范式概念	78		
5.3.3 1NF	78		
5.3.4 2NF	79		
5.3.5 3NF	80		
5.3.6 BCNF	81		
5.3.7 多值依赖	81		
5.3.8 4NF	84		
5.4 数据依赖的公理系统	84		
5.5 小结	85		
习题 5	85		
第 6 章 数据库设计	87		
6.1 数据库设计概述	87		
6.1.1 数据库设计的特点	87		
6.1.2 数据库设计方法	88		
6.1.3 数据库设计的阶段	89		
6.2 需求分析	91		
6.2.1 需求描述与分析	91		
6.2.2 需求分析分类	92		
6.2.3 需求分析的内容、方法和步骤	92		

第 8 章 数据库完整性和安全性	128	习题 9	164
8.1 数据库完整性概述	128	第 10 章 存储过程和触发器	165
8.1.1 实体完整性	128	10.1 存储过程	165
8.1.2 域完整性	129	10.1.1 存储过程的优点	165
8.1.3 参照完整性	129	10.1.2 存储过程的类型	166
8.1.4 用户定义的完整性	130	10.2 用户自定义存储过程的创建与执行	166
8.2 定义约束	130	10.2.1 创建存储过程	167
8.2.1 约束的类型	130	10.2.2 调用存储过程	169
8.2.2 约束的创建	131	10.2.3 存储过程的错误处理	170
8.2.3 查看约束的定义信息	132	10.2.4 管理存储过程	171
8.2.4 删除约束	133	10.3 触发器	173
8.3 使用默认	133	10.3.1 触发器的概念	174
8.4 使用规则	134	10.3.2 触发器的作用	174
8.5 数据完整性方法选用	135	10.3.3 触发器的类型	174
8.6 数据库安全性控制	135	10.4 DML 触发器	174
8.6.1 用户标识与鉴别	136	10.4.1 DML 触发器的优点	175
8.6.2 存取控制	137	10.4.2 DML 触发器的创建	175
8.6.3 审计	138	10.4.3 DML 触发器的应用	176
8.6.4 数据加密	139	10.5 DDL 触发器的创建与应用	177
8.6.5 权限控制	139	10.6 激活触发器	178
8.7 小结	141	10.7 管理触发器	178
习题 8	141	10.8 小结	180
第 9 章 SQL Server 2008 程序设计	142	习题 10	180
9.1 SQL 程序基础	142	第 11 章 数据库技术的发展	181
9.1.1 批处理	142	11.1 数据库技术的发展历程	181
9.1.2 数据类型	143	11.2 数据库系统发展的特点	181
9.1.3 常量	147	11.3 XML 数据库	182
9.1.4 变量	148	11.3.1 XML 简介	182
9.2 流程控制语句	152	11.3.2 XML 数据	183
9.3 表达式和函数	154	11.3.3 XML 数据模型	184
9.3.1 表达式	154	11.3.4 XML 数据库	185
9.3.2 函数	156	11.4 数据仓库	185
9.4 临时表和表变量	158	11.4.1 数据仓库概述	185
9.4.1 临时表	158	11.4.2 数据仓库中的数据组织	186
9.4.2 表变量	159	11.4.3 数据仓库系统的体系结构	186
9.5 游标	159	11.4.4 数据仓库的开发方法	188
9.5.1 游标的作用及特点	160	11.5 联机分析处理	188
9.5.2 游标的类型	160	11.5.1 OLAP 的一些基本概念	189
9.5.3 游标的使用	161	11.5.2 OLAP 的定义与特征	189
9.6 小结	164	11.5.3 多维分析操作	190

11.5.4 OLAP 的基本数据模型	190	13.4.1 创建数据库快照	219
11.5.5 OLAP 的衡量标准	192	13.4.2 管理数据库快照	220
11.6 数据挖掘	193	13.5 数据库中基本表的创建与管理	220
11.6.1 数据挖掘的定义	193	13.5.1 创建表	220
11.6.2 数据挖掘的数据源	193	13.5.2 定义完整性约束	222
11.6.3 数据挖掘的功能	194	13.5.3 修改表	227
11.6.4 数据挖掘工具	194	13.5.4 删除表	227
11.7 数据库技术的发展趋势	195	13.6 小结	228
11.8 小结	195	习题 13	228
习题 11	195	第 14 章 SQL Server 2008 安全性控制	229
第三部分 应用篇			
第 12 章 安装与配置 SQL Server 2008	197	14.1 安全管理概述	229
12.1 初识 SQL Server 2008	197	14.2 SQL Server 2008 的身份验证模式	229
12.1.1 SQL Server 发展简史	197	14.3 管理服务器登录	230
12.1.2 SQL Server 2008 版本和安装要求	197	14.3.1 查看和编辑现有的登录	230
12.1.3 SQL Sever 2008 的体系结构	198	14.3.2 创建登录	233
12.2 SQL Server 2008 的安装	199	14.3.3 启用、禁止和解锁登录	237
12.3 SQL Server 2008 的配置	204	14.3.4 修改登录	238
12.4 SQL Server 2008 管理工具	206	14.3.5 删除登录	239
12.4.1 SQL Server 管理平台	206	14.4 数据库用户	239
12.4.2 商业智能开发平台	207	14.4.1 建立数据库用户	240
12.4.3 SQL Server 配置管理器	208	14.4.2 修改数据库用户	241
12.4.4 SQL Server 分析器	209	14.4.3 删除数据库用户	241
12.4.5 数据库引擎优化顾问	209	14.5 权限管理	242
12.4.6 报表服务配置管理器	210	14.5.1 权限种类及用户分类	242
12.4.7 文档和教程	210	14.5.2 权限的管理	242
12.5 小结	211	14.6 角色	243
习题 12	211	14.6.1 服务器角色	243
第 13 章 创建与管理数据库及表	212	14.6.2 数据库角色	245
13.1 SQL Server 2008 数据库	212	14.6.3 新建数据库角色	246
13.1.1 系统数据库	212	14.7 小结	248
13.1.2 文件和文件组	213	习题 14	248
13.2 创建数据库	214	第 15 章 SQL Server 2008 的备份和恢复	249
13.3 数据库管理	217	15.1 数据的导入和导出	249
13.3.1 查看数据库	217	15.1.1 数据导出	249
13.3.2 修改数据库	218	15.1.2 数据导入	251
13.3.3 删除数据库	218	15.2 备份数据库	254
13.4 数据库快照	219	15.2.1 备份数据库概述	254
		15.2.2 备份方式	254
		15.2.3 备份时机	255

15.2.4 备份设备	255	实验 2 SQL Server 2008 表的操作	270
15.2.5 实现备份	258	实验 3 SQL Server 2008 数据库表 中记录的操作	271
15.3 恢复数据库	260	实验 4 SQL Server 2008 简单查询	274
15.3.1 恢复的顺序	260	实验 5 SQL Server 2008 高级查询	275
15.3.2 使用 T-SQL 语句恢复数据库	261	实验 6 SQL Server 2008 连接查询	277
15.3.3 使用 SQL Server Management Studio 恢复数据库	262	实验 7 SQL Server 2008 视图操作	279
15.4 分离和附加数据库	263	实验 8 SQL Server 2008 的存储 过程	281
15.4.1 分离数据库	263	实验 9 SQL Server 2008 的触发器	284
15.4.2 附加数据库	265	实验 10 SQL Server 2008 数据备份 和恢复	286
15.5 小结	266	课程设计 教学管理信息的分析 与设计	291
习题 15	266	参考文献	296
第四部分 实 验 篇			
实验 1 SQL Server 2008 的安装及 数据库操作	267		

第一部分 基础篇

第1章 数据库系统概述

我们所处的时代是一个信息化的时代。快速增长的信息资源是各个企业的重要财富。如何有效地管理和使用这些海量信息成为企业生存与发展的重要条件。数据库技术是20世纪60年代后期发展起来的一项管理数据的重要技术，其发展极为迅速，促进了计算机在信息处理方面的应用。它已成为人们存储数据、管理信息的最常用和最先进的技术。

本章主要介绍数据库系统的基本概念、数据管理的发展过程及数据库系统的组成等内容。

1.1 数据库系统基本概念

1.1.1 信息和数据

信息是针对某一特定目的的事实或事物，它是现实世界在人脑中的反应，是现实世界事物存在方式或运动状态的反映。它以文字、数据、符号、声音等形式记录下来，可以进行传递和处理，为人们进行判断和决策提供依据。

早期的计算机系统主要用于科学计算，处理的数据是数值型数据，例如整数、实数等。在信息时代，计算机的应用领域越来越广泛，需要计算机进行存储和处理的对象多种多样，表示这些对象的数据也变得越来越复杂，可以是文本、图像、音频、视频等。

数据是描述客观事物的一组符号集合，是计算机可以识别和处理的符号集。

数据包含两方面含义：第一，数据的内容是信息；第二，其表现方式是符号。例如数字、文本、声音、图像等都是数据。数据是数据库存储和处理的基本对象。

数据的格式与具体的计算机系统有关，随承载它的物理设备的形式而不同。现实世界中存在的事物可以用数据进行描述。例如，一个学生的情况可以用学号、姓名、性别、年龄、入学时间等进行描述。例如，(091102 李莉 女 18 2010-09-01)是一条记录，从中我们可以知道，有一个学生，学号为091102，名字叫李莉，女性，18岁，2010年9月1日入学。可见，数据是信息的载体，信息是数据的内涵，二者是形和质的关系。然而，如果事先不知道这个记录中各个字段的含义，就无法得到上述内容。也就是说，数据的表示形式并不一定能完全表达其实际的内容，需要经过解释后才能明确其具体的含义，才有意义，才能成为信息。因此数据和关于数据的解释是不可分割的。数据的解释是对数据含义的说明，数据的含义称为数据的语义，数据与其语义是不可分的。

简单来说，数据有如下特征：

1) 数据有型和值之分。数据的型指数据的结构，而数据的值指数据的具体值。例如，课程的数据由“课程编号”、“课程名”、“学分”等属性组成，它体现的是数据的结构。而“007”、“数据库原理”、“2”是具体一门课程的信息，它就是数据的值。

2) 数据受数据类型和取值范围的约束。数据类型是对数据的约束。给出一个数据的数据类型，其在内存中的存储单元数、取值范围和可以进行的有关操作就确定了。

3) 数据有多种表示形式。数据在表示或描述时, 可以采用多种形式, 如文字、图表、声音等。

信息通过数据的形式存储, 便于信息在计算机中存储、加工和传递。数据和信息是两个相互联系但又相互区别的概念。数据是信息的具体表现形式或称载体; 信息是经过加工处理的数据, 是人们消化理解了的数据, 是数据的内涵, 是数据的语义解释。

1.1.2 数据管理

如何从海量数据中找出人们感兴趣的、有意义有价值的信息, 这就需要对各种数据进行收集、保存、加工和传播, 上述一系列过程称为数据处理。数据处理分为数据管理、数据加工和数据传播。数据管理的核心是数据管理。数据管理是对数据进行分类、组织、编码、保存和检索。

1.1.3 数据库

数据库 (database, DB), 简单理解就是“存放数据的仓库”。只不过这个仓库是计算机存储设备。人们针对某一具体应用, 采集并整理出所需的大量数据, 为了方便对数据进行管理, 数据在存放时是按一定格式或组织结构进行的。例如, 将全校学生的基本信息存放在表中, 这张表就可以看做是一个数据库。有了这个数据库, 我们可以随时查询学生的基本信息, 对基本信息进行相应操作。

对数据库一个比较完整的定义是: 数据库是长期存储在计算机内, 有组织的、可共享的大量数据的集合。数据库中的数据能反映出数据之间的联系。

数据库具有如下的特点:

1) 实现数据集中管理。文件管理方式中, 数据处于一种分散的状态, 不同的用户或同一用户在不同的处理中其文件之间毫无关系。利用数据库可以对数据进行集中管理。数据库中的数据是按照一定的结构存储在一起的, 数据之间存在一定的联系。

2) 实现数据共享。数据库中的数据可以被多个不同的用户共享, 用户可以并发地存取数据。多个用户可以用各种方式通过接口使用数据库, 并共享数据。

3) 降低数据的冗余度。数据库中的数据面向多个应用, 可以被多个用户、多个应用程序共享和使用, 大大降低了数据的重复存储, 降低了数据的冗余。

4) 数据独立性高。数据独立性是指应用程序不会因数据的物理表示方式和访问技术的变化而改变, 即应用程序不依赖于任何特定的物理表示方式和访问技术。数据独立性包括两个方面: 物理独立性和逻辑独立性。

物理独立性是指当数据的存储位置或存储结构发生变化时, 不影响应用程序的特性; 逻辑独立性是指当表示现实世界的信息内容发生变化时, 不影响应用程序的特性。数据独立性是由数据库管理系统的二级映像功能来保证的。

数据与程序的独立, 具有以下优点:

1) 把数据的定义从程序中分离出去, 简化了应用程序的编写, 减少了应用程序的维护和修改。

2) 易于使用和扩充。

3) 便于统一的数据控制, 包括数据安全性控制、完整性控制和并发控制。

1.1.4 数据库管理系统

数据库管理系统 (Database Management System, DBMS) 是一种操纵和管理数据库的大型软件, 用于建立、使用和维护数据库。它对数据库进行统一的管理和控制, 以保证数据库的安全性和完整性。用户通过 DBMS 访问数据库中的数据, 数据库管理员也可以通过 DBMS 对数据库进行维护。它提供了多种功能, 可使多个应用程序的用户用不同的方法同时或不同时去建立、修改和访问数据库。DBMS 与数据库和应用程序的关系如图 1-1 所示。

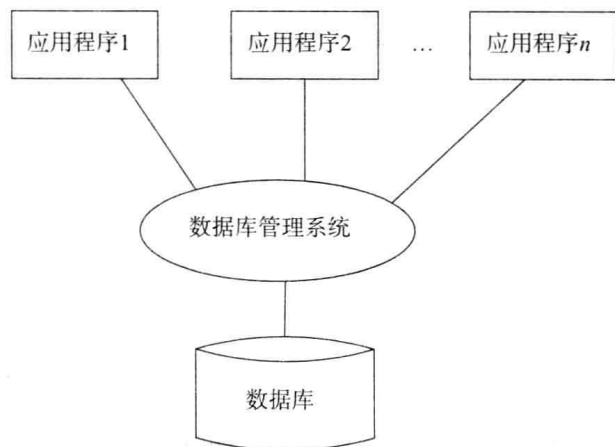


图 1-1 数据库管理系统与数据库和应用程序的关系

数据库管理系统是数据库系统的一个重要组成部分。数据库管理系统的任务是如何科学地组织和存储海量的数据, 如何高效地获取和维护它们, 并保证它们的安全性和完整性。

DBMS 的主要功能包括:

1) 数据定义功能。借助 DBMS 提供的数据定义语言 (Data Definition Language, DDL) 来对数据库中的数据对象进行定义。数据定义包括定义模式、存储模式和外模式, 定义外模式与模式之间的映射, 定义存储模式与模式之间的映射, 定义有关的约束条件。

2) 数据组织、存储和管理。为了提高数据的存取效率, 数据库管理系统要分类组织、存储和管理各种数据。数据库中的数据包括数据字典、用户数据和存取路径等。数据库管理系统要确定组织数据的文件结构和存取方式, 以及实现数据之间的联系。数据存储和组织的基本目标是提高存储空间利用率和方便存取, 提供多种存取方法, 提高存取效率。

3) 数据操纵功能。借助 DBMS 提供的数据操纵语言 (Data Manipulation Language, DML) 来完成用户对数据库提出的各种要求, 实现数据的查询、插入、删除和修改等操作。

4) 数据库的事务管理和运行管理。数据库中的数据可以供多个用户同时使用, 具有共享性。为了保证数据能安全、可靠的运行, 数据库提供了事务管理功能。数据库在建立、运行和维护时由 DBMS 统一管理和控制, 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。

5) 数据库的建立和维护功能。它包括数据库初始数据装载转换、数据库转储、介质故障恢复、数据库的重组织、性能监视分析等。为了提高处理的效率, 数据库管理系统提供了多种维护工具软件。

6) 其他功能。DBMS 的其他功能包括: DBMS 与网络中其他软件系统的通信、两个 DBMS 系统的数据转换、异构数据库之间的互访和互操作等。

1.1.5 数据库系统

数据库系统 (Database System, DBS) 是指在计算机系统中引入数据库后的系统, 一般由数据库、数据库管理系统、应用开发工具、应用系统和用户构成, 核心是数据库管理系统。数据库系统各组成部分的关系如图 1-2 所示。

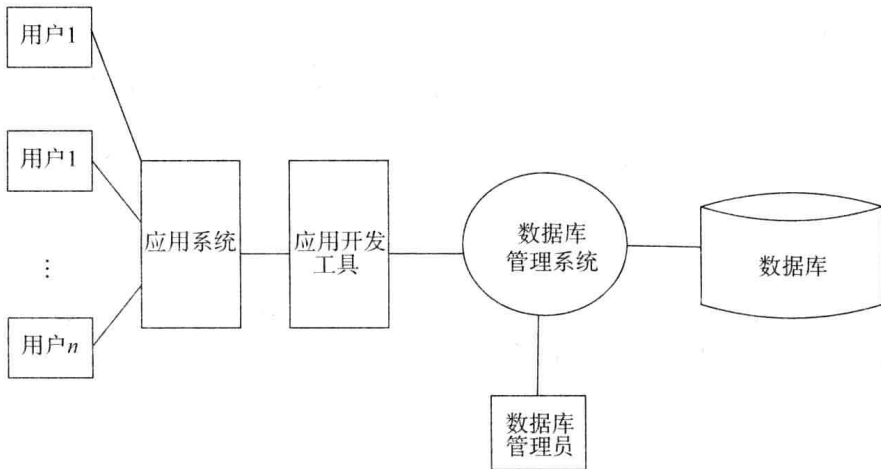


图 1-2 数据库系统

数据库系统各部分具体说明如下：

1) 数据库：它是数据库系统的处理对象，是以一定的组织方式存储在一起的、大量的、有组织的、能够被多个用户共享的数据的集合。不同厂家的数据库系统中，数据库的物理存储形式是不同的。例如，xBASE 微机数据库中，并没有数据库实体，可见的是组成数据库的一个个数据文件和索引文件；在 Access 中，数据库是以 *.mdb 文件的形式存在，没有独立的数据文件和索引文件，数据库中有的是一张张的表。其他大型的数据库还可能以其他的物理形式存在。

2) 数据库管理系统：它是数据库系统的核心，主要负责对数据库的管理和维护工作。用户一般不直接对数据库中的数据进行处理，必须通过数据库管理系统来实现。通过 DBMS，用户可以逻辑、抽象地处理数据，不用关心数据在计算机中的具体存储方式以及计算机处理数据的过程细节。一切具体而繁琐的工作由数据库管理系统完成。

3) 应用系统：它介于用户和数据库管理系统之间，是计算机软件开发人员开发的面向最终用户的软件，它是在数据库管理系统的基础上实现的。数据库应用系统一般具有友好的用户界面，便于用户实现自己的需求。

4) 用户：它包括数据库的开发、管理和使用者。按照其职责分为数据库管理员（Database Administrator, DBA）、数据库分析员、数据库设计员、应用程序员和终端用户。

- 数据库管理员是数据库系统中的重要人员，是专门负责数据库的建立、配置、管理和维护工作的人员。其主要职责体现在建库、用库和改进三个方面。他决定数据中的信息的内容和结构，决定数据库的存储结构和存取策略，同时，他要确定各个用户对数据库的存取权限、数据的保密级别和完整性约束条件，保证数据库的安全性和完整性。并且，数据库管理员还要监视数据库系统的运行情况，及时处理运行过程中出现的问题，进行数据库的备份、恢复和改进等工作。
- 数据库分析员负责应用系统的需求分析和规范说明，和终端用户及数据库管理员相配合，分析确定系统的软硬件配置，并参与数据库系统的设计。
- 数据库设计员负责数据库中数据的确定、数据库各级模式的设计。数据库设计员必须参加用户需求调查和系统分析，然后进行数据库设计。
- 应用程序员负责设计和编写应用系统的程序模块，并进行调试和安装。
- 终端用户是数据库系统的使用者，通过应用系统的用户接口实现各种操作，终端用户可以分为偶然用户、简单用户和复杂用户三类。

在不引起混淆的情况下，常常把数据库系统称为数据库。

1.2 数据库分类

随着数据库技术的发展和用户对数据需求的多元化，目前有各种各样的特色数据库系统，其目的是解决各种工程、应用问题。数据库的种类很多，可从数据库结构、应用对象、应用范围等几个方面分类。

1) 按照数据库结构分类：关系数据库、层次数据库、网状数据库和面向对象数据库。

2) 按照应用对象分类：以管理字符数据为主的传统数据库、以管理多媒体数据为主的多媒体数据库、以提供信息共享为主的浏览器/服务器网络数据库、以解决网络数据共享为主的分布式数据库等。

3) 按应用范围分类：主要有空间数据库、工程数据库等。

4) 按工作方式分类：主要有主动数据库、并行数据库、实时数据库等。

1.3 数据管理技术的产生和发展

数据管理是数据处理的核心问题，是指对数据进行分类、组织、编码、存储、检索和维护。数据管理技术的发展大致经历了人工管理阶段、文件管理阶段和数据库管理阶段三个阶段。

1. 人工管理阶段

20世纪50年代中期以前，计算机主要应用是进行科学计算。当时硬件方面，外存只有纸带、卡片和磁带，没有直接进行存取的设备，在计算时将数据输入，计算完数据不保存。对用户数据如此处理，对系统软件有时也是这样。软件方面，没有操作系统，没有管理数据的专门软件。其主要的处理方式是批处理。

在该阶段，每个应用程序需要自己设计、说明和管理数据。数据的共享性和独立性比较差，数据的逻辑结构或物理结构发生变化后，必须对应用程序做相应的修改，这加重了程序员的负担。

在人工管理阶段，程序与数据之间一一对应的关系如图1-3所示。

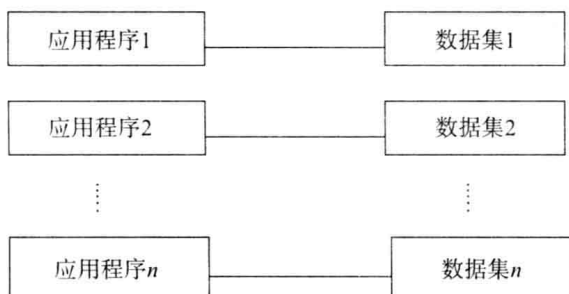


图1-3 人工管理阶段程序与数据的关系

2. 文件系统阶段

20世纪50年代后期到60年代中期，计算机在硬件方面已经有了磁盘、磁鼓等直接存储设备；数据以文件的形式长期保存。软件方面，操作系统中已经有了专门的数据管理软件，一般称为文件系统。处理方式上不仅有了批处理，而且能够联机实时处理。这时，计算机不仅用于科学计算，也已大量用于数据处理。

在文件系统阶段，利用“按文件名访问，按记录进行存取”的管理技术，对文件中的数据进行修改、插入和删除操作。应用程序和数据之间有了一定的独立性，程序员不必过多地考虑物理细节，大大节省了维护程序的工作量。但是文件系统中，文件仍然是面向应用的。当不同的应用程序具有部分相同数据时，不能共享相同的数据。因此数据的冗余度大，浪费存储空间，而且容易造成数据的不一致性，给数据的修改和维护带来困难。另外，数据独立性差，文件之间是孤立的，系统不容易扩充。

文件系统阶段，程序与数据之间的关系如图1-4所示。

3. 数据库系统阶段

20 世纪 60 年代末以来, 由于计算机处理数据量的急剧增长, 对数据的管理和分析力度需求也随之加大, 同时, 对多种应用、多种语言的共享数据的要求越来越强烈。

随着大容量磁盘的出现, 硬件价格不断下降, 软件价格不断上升, 编写和维护系统软件 and 应用程序的成本不断增加。同时, 对实时处理的要求越来越多, 以文件系统作为数据管理手段已经不能满足应用的需求, 于是, 出现了新的数据管理技术——数据库技术, 出现了统一管理数据的专门软件系统——数据库管理系统。

在数据库系统中, 所有相关的数据都存储在一个称为数据库的集合中, 它们作为一个整体定义。由于数据是统一管理的, 因此可以从全局出发, 合理组织数据, 避免了数据冗余。另外, 在数据库中, 程序与数据相互独立, 数据通过 DBMS 而不是应用程序来操作和管理, 应用程序不再处理文件和记录的格式。

数据库系统阶段, 应用程序与数据之间的对应关系如图 1-5 所示。

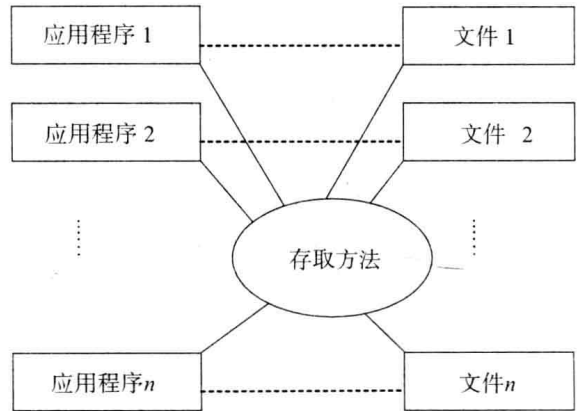


图 1-4 文件系统阶段程序与数据的关系

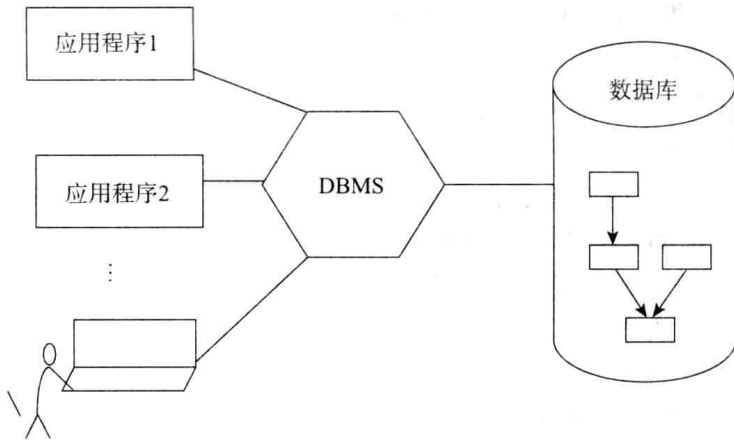


图 1-5 数据库系统阶段程序与数据的关系

1.4 数据库系统结构

数据库系统的结构是数据库系统的一个总框架。这个框架结构用于描述一般的数据库概念, 并可以解释特定的数据库结构。可以从不同的角度来考察数据库系统的结构, 从应用程序(用户)的角度, 数据库系统可以分为集中式、客户/服务器(C/S)等结构; 从数据库管理系统的角度, 数据库领域公认的标准结构是三层模式结构——外模式、概念模式和内模式。

1.4.1 数据库系统模式的概念

在数据模型中有“型”(type)和“值”(value)的概念。型是对某一类数据的结构和属性的说明, 值是型的一个具体赋值。例如, (学号, 姓名, 性别, 年龄, 所在系)是学生记录的型, 而(09111101, 黄维, 男, 18, 计算机系)是一个记录值。