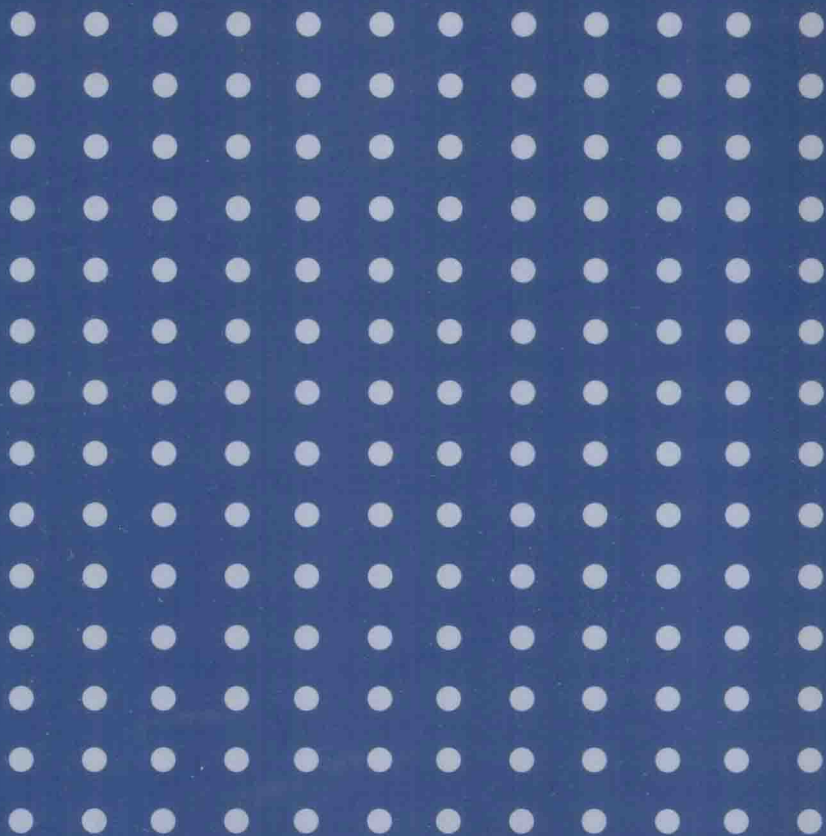


重点大学计算机专业系列教材

数据库技术及应用

王成良 柳玲 徐玲 编著



清华大学出版社

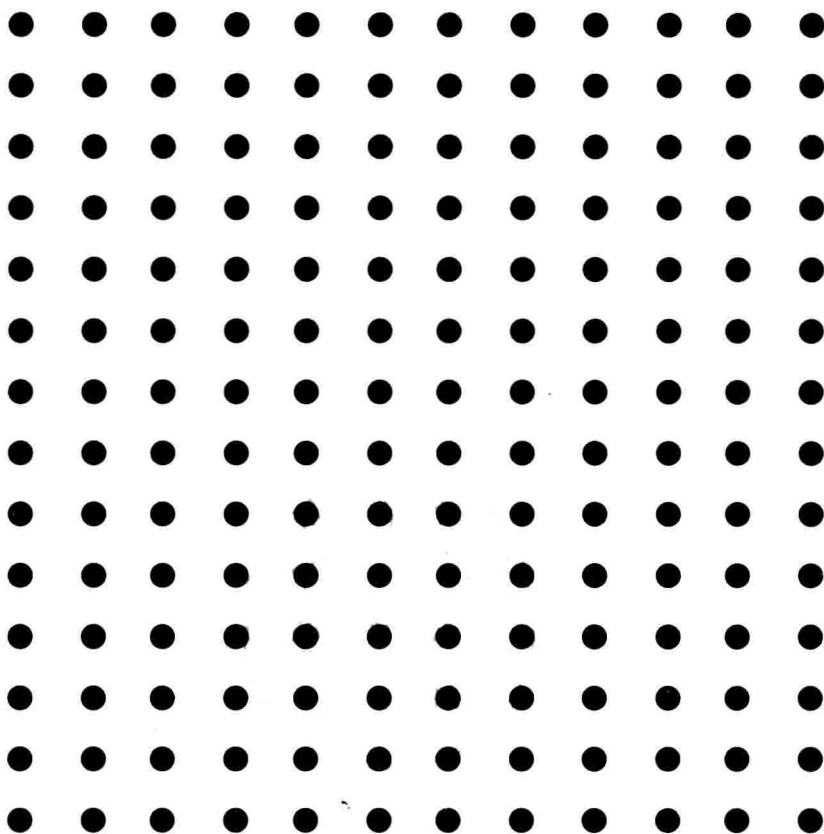




重点大学计算机专业系列教材

数据库技术及应用

王成良 柳玲 徐玲 编著



清华大学出版社

北京

内 容 简 介

本书以当前主流的关系数据库为主线,全面介绍了数据库技术的基本内容。全书共 10 章,分别为数据库基础知识,信息的三种世界与数据模型,关系模型,SQL Server 2008 关系数据库管理系统,关系数据库标准语言——SQL,数据库保护,关系数据库理论,数据库系统的设计,数据库高级应用技术,数据库技术的发展趋势。

本书以大型主流数据库管理系统 SQL Server 2008 作为丰富案例的演练平台,注重数据库技术的实际应用,强调理论与实践紧密结合。本书各章后均配有习题,具有较强的可读性。

本书是高等院校计算机、软件工程及相关专业本科生数据库课程教学的理想教材,也是从事数据库技术领域工作的科技人员的有价值的参考书。

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

数据库技术及应用/王成良,柳玲,徐玲编著.--北京:清华大学出版社,2011.11

(重点大学计算机专业系列教材)

ISBN 978-7-302-27145-1

I. ①数… II. ①王… ②柳… ③徐… 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 214062 号

责任编辑:付弘宇

责任校对:白 蕾

责任印制:李红英

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 装 者:北京鑫海金澳胶印有限公司

经 销:全国新华书店

开 本:185×260 印 张:23.75 字 数:589 千字

版 次:2011 年 11 月第 1 版 印 次:2011 年 11 月第 1 次印刷

印 数:1~3000

定 价:39.00 元

出版说明

随着国家信息化步伐的加快和高等教育规模的扩大,社会对计算机专业人才的需求不仅体现在数量的增加上,而且体现在质量要求的提高上,培养具有研究和实践能力的高层次的计算机专业人才已成为许多重点大学计算机专业教育的主要目标。目前,我国共有16个国家重点学科、20个博士点一级学科、28个博士点二级学科集中在教育部部属重点大学,这些高校在计算机教学和科研方面具有一定优势,并且大多以国际著名大学计算机教育为参照系,具有系统完善的教学课程体系、教学实验体系、教学质量保证体系和人才培养评估体系等综合体系,形成了培养一流人才的教学和科研环境。

重点大学计算机学科的教学与科研氛围是培养一流计算机人才的基础,其中专业教材的使用和建设则是这种氛围的重要组成部分,一批具有学科方向特色优势的计算机专业教材作为各重点大学的重点建设项目成果得到肯定。为了展示和发扬各重点大学在计算机专业教育上的优势,特别是专业教材建设上的优势,同时配合各重点大学的计算机学科建设和专业课程教学需要,在教育部相关教学指导委员会专家的建议和各重点大学的大力支持下,清华大学出版社规划并出版本系列教材。本系列教材的建设旨在“汇聚学科精英、引领学科建设、培育专业英才”,同时以教材示范各重点大学的优秀教学理念、教学方法、教学手段和教学内容等。

本系列教材在规划过程中体现了如下一些基本组织原则和特点。

1. 面向学科发展的前沿,适应当前社会对计算机专业高级人才的培养需求。教材内容以基本理论为基础,反映基本理论和原理的综合应用,重视实践和应用环节。

2. 反映教学需要,促进教学发展。教材要能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。规划教材建设的重点依然是专业基础课和专业主干课;特别注意选择并安排了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重点大学

前言

数据库技术自 20 世纪 60 年代产生至今已得到了迅猛的发展,目前已成为现代计算机信息系统与应用系统的核心技术。数据库的建设规模和数据库系统的应用水平是衡量一个国家信息化程度的重要标志之一,数据库应用技术已成为很多高校理工科学生应具备的重要技能之一。“数据库技术与应用”课程作为软件工程专业和计算机科学技术专业的一门重要专业基础课程,在整个专业课程体系起着承上启下、融会贯通的作用,是学生参加项目实践、毕业设计、软件开发和工作就业的重要的专业理论和实践课程,对提高本科学生的信息技术开发能力起着非常关键的作用。

本书融入作者从事数据库教学和数据库应用开发十多年来所积累的丰富经验,秉承拓宽基础、注重应用、提高能力的原则,以关系数据库原理、方法和技术为重点,以大型主流数据库管理系统 SQL Server 2008 作为丰富案例的演练平台,强调理论与实践紧密结合,注重数据库综合性知识和数据库技术应用能力的培养,通过实例讲解原理和方法,引导学生掌握理论方法的实际运用,不仅使学生由浅入深、循序渐进地完整掌握数据库技术的基本原理和基础知识,而且本教材中引入的许多数据库实用开发技术,可以培养学生具有较强的数据库综合应用开发能力,弥补了当前教材中存在的理论性强、实践性不够的缺陷。

为便于教师教学和读者学习,本书在每章的前面都列出了“本章学习目标”,指出本章的主要内容以及应该理解和掌握的知识点;在每章的最后附有习题,帮助读者巩固所学的知识。另外,本书的配套实验和课程设计教材《数据库技术及应用实验与课程设计教程》将在随后出版,通过验证性实验和综合性课程设计,使学生由浅入深、由点到面逐步提高,进一步巩固学生的数据库技术理论知识,并能结合实际问题熟练开发数据库应用系统,提高学生的综合实践与创新能力。

本书共分 10 章,第 1 章介绍数据库基础知识,包括数据库、数据库管理系统和数据库系统的概念、数据库的三级模式结构、数据库管理系统的主要功能等内容。第 2 章讲解数据模型的基础知识,包括常用的几种数据模型,并介绍了新一代数据模型——面向对象数据模型和半结构化数据模型。第 3

章讲解关系模型,包括关系的概念、关系代数的各种运算、关系演算等内容。第4章介绍 SQL Server 2008 数据库管理系统的应用,包括 SQL Server 2008 的体系结构、数据库文件管理及管理工具的使用方法等。第5章讲解关系数据库标准语言——SQL,包括 SQL 的数据定义功能、SQL 的数据查询功能、SQL 的数据操纵功能、SQL 的数据控制功能、视图、索引、存储过程和函数等内容。第6章讲解数据库保护,包括数据库完整性、安全性、并发控制、事务、数据库故障恢复等内容。在前面基础上,第7章为便于理解以大量实例讲解关系数据库理论,包括函数依赖、推理规则及逻辑蕴涵、关系模式分解以及关系模式的范式等。第8章讲解数据库系统的设计方法,采用案例分析的方式,让读者掌握数据库设计的全过程。第9章讲解数据库高级应用技术,包括数据库建模工具、存储过程、触发器和数据库事务处理的高级应用,以及数据库性能优化等内容。第10章介绍数据库技术的发展趋势,包括分布式数据库、面向对象数据库、数据仓库与数据挖掘、多媒体数据库、实时数据库、专家数据库、内存数据库、NoSQL 等内容。

本书由重庆大学软件学院王成良、柳玲、徐玲共同编写完成,其中柳玲负责第1、2、3、5章,徐玲负责第4、6、7章,王成良负责第8、9、10章,冉唯、焦晓军等参与了本书资料的搜集和整理工作,王成良对本书进行了编排和统稿。

本书编写过程中参考了许多相关书籍和资料,已在书后列出,在此对这些参考文献的作者表示感谢,同时感谢清华大学出版社对本书出版所给予的支持和帮助,也感谢重庆大学软件学院熊庆宇院长和陈蜀宇书记对编写本书所给予的大力支持。

由于编者水平有限,书中难免存在疏漏和不足,敬请读者批评指正,以利改进和提高。

本书的课件及源代码资源可以从清华大学出版社网站 www.tup.com.cn 下载,本书和课件的使用问题请联系 fuhy@tup.tsinghua.edu.cn。

编 者

2011年7月

目录

第 1 章 数据库基础知识	1
1.1 引言	1
1.2 数据库技术的相关概念	1
1.2.1 数据和信息	1
1.2.2 数据处理和数据管理	2
1.2.3 数据库	2
1.2.4 数据库管理系统	2
1.2.5 数据库系统	2
1.2.6 数据库管理员	3
1.2.7 数据库用户	3
1.3 数据管理的发展	4
1.3.1 人工管理	4
1.3.2 文件系统管理	5
1.3.3 数据库技术管理	6
1.3.4 高级数据库技术管理	7
1.3.5 XML 文件管理	9
1.3.6 数据管理技术的比较	9
1.4 数据库的体系结构	11
1.4.1 数据库系统的三级模式结构	11
1.4.2 数据库系统的应用构架	13
1.5 数据库管理系统	15
1.5.1 数据库管理系统的主要功能	15
1.5.2 数据库管理系统的组成	15
1.6 常用的数据库管理系统介绍	16
1.6.1 SQL Server	16
1.6.2 DB2	17
1.6.3 Oracle	17

1.6.4 Sybase ASA	18
1.6.5 Access	19
1.6.6 MySQL	19
习题 1	20
第 2 章 信息的三种世界与数据模型	21
2.1 信息的三种世界及其描述	21
2.1.1 现实世界	21
2.1.2 信息世界	21
2.1.3 机器世界	22
2.2 数据模型	22
2.2.1 数据模型的内容	22
2.2.2 数据模型的分类	22
2.2.3 实体联系模型	23
2.2.4 层次模型	25
2.2.5 网状模型	27
2.2.6 关系模型	28
2.2.7 面向对象数据模型	29
2.2.8 半结构化数据模型	30
习题 2	30
第 3 章 关系模型	32
3.1 关系模型的由来	32
3.2 关系数据库的结构	33
3.2.1 关系模型的基本术语	33
3.2.2 关系的键	34
3.2.3 基于集合论的关系定义	36
3.2.4 关系规则	37
3.2.5 关系操作	39
3.3 关系代数	40
3.3.1 传统的集合运算	41
3.3.2 自然关系运算	42
3.3.3 关系代数综合实例	50
3.3.4 扩展的关系代数运算	53
3.4 关系演算	55
3.4.1 元组关系演算	55
3.4.2 域关系演算	58
3.5 关系代数表达式的优化	59
习题 3	61

第 4 章 SQL Server 2008 关系数据库管理系统	64
4.1 SQL Server 2008 概述	64
4.1.1 SQL Server 2008 的各种版本	64
4.1.2 SQL Server 2008 的新特性	65
4.2 SQL Server 2008 体系结构	66
4.3 SQL Server 2008 系统数据库	67
4.3.1 SQL Server 2008 数据库的组成	67
4.3.2 SQL Server 2008 数据库	68
4.3.3 SQL Server 2008 数据库对象	70
4.4 SQL Server 2008 的管理工具	71
4.4.1 SQL Server 配置管理器	71
4.4.2 SQL Server Management Studio	75
4.4.3 SQL Server Profiler	77
4.4.4 数据库引擎优化顾问	78
4.4.5 实用工具	79
4.4.6 联机丛书	80
习题 4	81
第 5 章 关系数据库标准语言——SQL	82
5.1 SQL 概述及特点	82
5.1.1 SQL 的发展历程	82
5.1.2 SQL 的组成及特点	83
5.1.3 Transact-SQL 概述	84
5.1.4 SQL 语言的基本概念	85
5.2 SQL 的数据定义功能	90
5.2.1 数据库的创建和删除	90
5.2.2 基本表的创建、修改、删除	91
5.3 SQL 的数据查询功能	94
5.3.1 查询语句的基本结构	95
5.3.2 简单查询	96
5.3.3 连接查询	103
5.3.4 嵌套查询	106
5.3.5 集合查询	109
5.3.6 复杂查询	111
5.4 SQL 的数据操纵功能	112
5.4.1 插入数据	112
5.4.2 更新数据	113
5.4.3 删除数据	113

5.5	视图	114
5.5.1	视图的概念及特点	114
5.5.2	视图的创建和使用	114
5.5.3	视图的更新	115
5.5.4	视图的删除	117
5.6	索引	118
5.6.1	索引的概念及作用	118
5.6.2	索引的分类	118
5.6.3	索引的创建及删除	118
5.7	SQL 的数据控制功能	119
5.7.1	授予权限	119
5.7.2	收回权限	121
5.7.3	视图机制保证安全性	121
5.8	存储过程	121
5.8.1	存储过程简介	122
5.8.2	存储过程的创建与执行	123
5.8.3	存储过程的修改	124
5.8.4	重新编译存储过程	125
5.8.5	存储过程的删除	125
5.8.6	使用存储过程的注意事项	125
5.9	函数	126
5.9.1	函数的概念及优点	126
5.9.2	函数的创建与使用	126
5.9.3	函数的修改	127
5.9.4	函数的删除	127
5.9.5	SQL Server 2008 中的内置函数	128
5.10	Transact-SQL 的流程控制语句	134
5.10.1	begin...end 语句	134
5.10.2	if...else 语句	135
5.10.3	case 语句	135
5.10.4	while 语句	137
5.10.5	goto 语句	138
5.10.6	waitfor 语句	138
5.10.7	return 语句	139
5.10.8	try/catch 语句	139
5.11	SQL Server 2008 中 Transact-SQL 的扩展功能	139
5.12	嵌入式 SQL	141
5.12.1	嵌入式 SQL 的定义及实现	141
5.12.2	嵌入式 SQL 语句的使用	141

5.12.3	SQL 和宿主语言的接口	142
5.12.4	嵌入式 SQL 语句	144
5.12.5	动态 SQL 语句	146
习题 5		147
第 6 章	数据库保护	150
6.1	系统目录	150
6.1.1	系统目录简介	150
6.1.2	SQL Server 2008 的系统目录	153
6.2	数据库完整性	155
6.2.1	完整性规则	155
6.2.2	完整性约束	156
6.2.3	触发器	161
6.2.4	SQL Server 2008 的完整性控制	163
6.3	数据库的安全性	171
6.3.1	安全性概述	171
6.3.2	身份认证	172
6.3.3	存取控制	173
6.3.4	自主存取控制	174
6.3.5	强制存取控制	175
6.3.6	建立视图	176
6.3.7	数据加密	177
6.3.8	审计跟踪	178
6.3.9	SQL Server 2008 的安全机制	178
6.4	事务	180
6.4.1	事务的基本概念	180
6.4.2	事务的特性	181
6.4.3	SQL 事务处理模型	182
6.4.4	SQL Server 2008 的事务处理	183
6.5	并发控制	185
6.5.1	事务的并发执行	185
6.5.2	并发操作与数据的不一致性	186
6.5.3	封锁	187
6.5.4	事务调度与可串行化	193
6.5.5	两段锁协议	197
6.5.6	SQL Server 2008 的并发控制机制	197
6.6	数据库的恢复	199
6.6.1	故障的种类	199
6.6.2	故障恢复技术	200

6.6.3 检查点.....	202
6.6.4 事务故障恢复.....	203
6.6.5 系统与介质故障的恢复.....	204
6.6.6 SQL Server 2008 的备份和恢复	205
习题 6	211
第 7 章 关系数据库理论	213
7.1 关系模式规范化的必要性	213
7.2 函数依赖	215
7.2.1 函数依赖的定义.....	216
7.2.2 函数依赖的分类.....	217
7.2.3 函数依赖和键的联系.....	217
7.2.4 函数依赖的逻辑蕴涵.....	218
7.2.5 函数依赖的推理规则.....	219
7.2.6 函数依赖集的闭包和属性集的闭包.....	220
7.2.7 函数依赖集的最小依赖集.....	221
7.3 关系模式的分解	223
7.3.1 模式分解的规则.....	223
7.3.2 无损连接分解.....	224
7.3.4 保持函数依赖的分解.....	226
7.4 关系模式的范式	226
7.4.1 第一范式(1NF).....	227
7.4.2 第二范式(2NF).....	228
7.4.3 第三范式(3NF).....	229
7.4.4 BCNF 范式	230
7.4.5 多值依赖与第四范式(4NF)	232
7.4.6 规范化小结.....	234
习题 7	235
第 8 章 数据库系统的设计	237
8.1 数据库系统设计概述	237
8.2 系统需求分析	238
8.2.1 需求分析的必要性.....	238
8.2.2 需求分析的方法.....	239
8.2.3 数据流图和数据字典.....	239
8.3 概念结构的设计	242
8.3.1 概念模型的特点、设计方法和步骤	243
8.3.2 数据抽象与局部视图设计.....	244
8.3.3 视图的集成.....	245

8.4	数据库逻辑结构的设计	248
8.4.1	逻辑结构设计的过程	248
8.4.2	概念模型向关系模型的转换	248
8.4.3	设计用户子模式	252
8.5	数据库物理结构的设计	255
8.5.1	确定关系模式的存取方法	255
8.5.2	确定数据库的存储结构	257
8.5.3	评价物理结构	257
8.6	数据库的实施和维护	257
8.6.1	数据的载入和应用程序的调试	258
8.6.2	数据库的试运行	258
8.6.3	数据库的运行和维护	259
8.7	综合实例	259
8.7.1	库存管理的需求分析和相关文档	260
8.7.2	设计 E-R 图	262
8.7.3	将 E-R 图转换为关系模式	262
8.7.4	规范化处理	263
8.7.5	数据库实施	264
习题 8		266
第 9 章	数据库高级应用技术	269
9.1	数据库建模工具的应用	269
9.1.1	PowerDesigner 概述	270
9.1.2	PowerDesigner 15 的组成	270
9.1.3	基于 PowerDesigner 的数据库建模	272
9.2	存储过程的高级应用	293
9.2.1	存储过程应用实例	294
9.2.2	执行系统存储过程和扩展存储过程	295
9.3	函数的高级应用	298
9.3.1	函数的使用位置	298
9.3.2	日期函数的应用	299
9.3.3	isnull 函数的应用	301
9.3.4	复杂字段约束的实现	301
9.4	数据库连接技术	303
9.4.1	数据库应用开发接口	303
9.4.2	使用 ADO.NET 连接 SQL Server 2008	306
9.4.3	使用 JDBC 连接 SQL Server 2008	309

9.5 数据库性能优化技术	312
9.5.1 逻辑数据库规范化问题	312
9.5.2 改善物理数据库的存储	313
9.5.3 与 SQL Server 相关的硬件系统的优化	313
9.5.4 检索策略的优化	314
习题 9	318
第 10 章 数据库技术的发展趋势	319
10.1 分布式数据库系统	319
10.1.1 分布式数据库系统的概念	319
10.1.2 分布式数据库系统的特点	321
10.1.3 分布式数据库系统的分类	323
10.1.4 分布式数据库系统的结构	324
10.1.5 分布式数据库管理系统	326
10.1.6 分布式数据库的应用与发展	327
10.2 面向对象数据库系统	328
10.2.1 面向对象数据库系统的兴起	328
10.2.2 面向对象数据库模型的核心概念	330
10.2.3 面向对象数据库的模式演进	337
10.3 数据仓库与数据挖掘技术	339
10.3.1 数据仓库	339
10.3.2 数据挖掘技术	341
10.4 多媒体数据库	346
10.4.1 多媒体数据库的定义	346
10.4.2 多媒体数据的特点	346
10.4.3 多媒体数据库管理系统	347
10.5 实时数据库	348
10.5.1 实时数据库的定义	348
10.5.2 实时数据库的功能特征	349
10.5.3 实时数据库管理系统的功能特征	350
10.5.4 实时数据库系统的主要技术	350
10.5.5 RTDBMS 的体系结构	352
10.6 专家数据库	352
10.6.1 专家数据库的目标	352
10.6.2 专家数据库的系统结构	353
10.7 内存数据库	354
10.7.1 内存数据库的定义	354

10.7.2 常见的通用内存数据库	354
10.8 NoSQL 数据库	355
10.8.1 NoSQL 数据库的产生	355
10.8.2 NoSQL 数据库的概念	356
10.8.3 NoSQL 数据库的分类	357
习题 10	360
参考文献	361

数据库基础知识

第 1 章

本章学习目标

- 了解数据管理技术的发展历程。
- 掌握数据库、数据库管理系统和数据库系统的概念及相互关系。
- 掌握数据库系统的三级模式结构及它们之间的映像。
- 掌握数据库管理系统的主要功能。
- 了解常用的数据库管理系统。

1.1 引言

数据是信息的基础,如何实现数据的存储、操纵、管理和检索,进而从中获取有价值的信息,已成为当今计算机技术研究和应用的重要课题。数据库技术在 20 世纪 60 年代产生,主要研究如何存储、操纵、管理和检索数据。经过 40 多年的发展,数据库技术已经形成了坚实的理论基础,开发了成熟的商业产品,在众多领域得到了广泛应用。例如,在教育、商业、医疗保健、政府部门、图书馆、军事、工业控制等领域都有数据库技术的应用。数据库技术已成为信息管理、电子商务、电子政务、办公自动化、网络服务等应用系统的核心技术和重要基础,吸引了众多的数据库理论研究、系统研制和应用开发等不同层次的学者、专家和技术人才致力其研究和应用。

1.2 数据库技术的相关概念

1.2.1 数据和信息

1. 数据

数据(data)实际上就是描述事物的符号记录,可以是数字、文字、图形、图像、视频或音频等。数据本身并不能完全表达其内容,需要经过语义解释。数据与其语义是不可分的。在日常生活中,我们一般采用自然语言描述事

物,例如,张强出生于1980年,他的专业是软件工程。但是在计算机中,因为存储和处理的需要,应抽象出事物相关特征组成一个记录来描述这些事物。例如,描述一个学生,可以抽象出学生的相关特征,组成如下记录:

张强,男,1980,重庆,软件工程,2008

从这条记录中,了解其含义的人可以得到这些信息:张强是一个男生,出生于1980年,重庆人,是软件工程专业2008级学生。不了解其语义的人则不能理解其含义。

2. 信息

信息(information)通常定义为经过加工处理后的数据,所以,数据和信息都是客观事物的反映,反映了人们对事物的认识。信息是以数据的形式表示的,即数据是信息的载体,信息则是数据加工的结果,是对数据的解释。数据和信息的相互关系如图1.1所示。计算机系统的每项操作,均是对数据进行某种处理。数据输入计算机后,经存储、传输、排序、计算、转换、检索和仿真等操作,输出人们需要的结果,亦即产生信息。

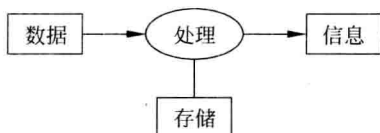


图 1.1 数据和信息的相互关系

1.2.2 数据处理和数据管理

数据处理是指将数据转换成信息的过程,如对数据的收集、存储、传播、检索、分类、加工或计算、打印各类报表或输出各种需要的文本和图形等。在数据处理的一系列活动中,数据收集、存储、传播、检索、分类等操作是基本环节,这些基本环节统称为数据管理。

1.2.3 数据库

数据库(DataBase,DB)是长期存储在计算机内的、有组织的、可共享的数据集合。数据库具有如下特点: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、较高的独立性和易扩展性,并且可以为各种用户共享等。数据库是数据库系统的组成部分。

1.2.4 数据库管理系统

数据库管理系统(DataBase Management System,DBMS)是位于用户与操作系统(Operation System,OS)之间,使人们能对数据库中的数据进行科学的组织、高效的存取和维护、管理的一种数据管理软件。比如我们常常听说的SQL Server、DB2、Oracle等就是商用的数据库管理系统。

1.2.5 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构成,一般由数据库、操作系统、数据库管理系统、应用开发工具、应用系统、数据库管理员和用户构成(见图1.2)。