

高等学校计算机科学与技术系列教材
武汉大学“十一五”规划教材




数据库原理与技术

尹为民 主编

(第二版)

13
2.02

 科学出版社
www.sciencep.com

清华大学出版社
清华大学出版社



数据库原理与技术

清华大学出版社

清华大学出版社

清华大学出版社

学校计算机科学与技术系列教材
武汉大学“十一五”规划教材

数据库原理与技术

(第二版)

尹为民 主编

科学出版社

北京

版权所有,侵权必究

举报电话:010-64030229;010-64034315;13501151303

内 容 简 介

这是一部关于数据库系统的基本原理、技术和方法的教科书。分为两个部分,第一部分:基础篇。介绍数据库基础知识,描述数据库原理、关系数据库设计与实现方法。第二部分:系统篇。介绍数据库管理系统及其数据库保护与事务管理技术,描述数据库访问及应用系统建模方法,讨论现代数据库系统与高级论题。

本书是作者多年从事数据库原理课程教学与科研实践的结晶,注重核心理论的描述、注意基础理论和现行应用技术的结合。删减繁琐、过时的内容,同时又根据本学科教学科研发展的动态,介绍现代数据库系统的特点、技术及相关知识,本书配有方便教学的网站以及学习辅导书,以方便师生的教与学。

本书不仅可作为高等院校计算机及相关专业学生数据库课程的教材,还可供从事信息领域工作的科技人员及其他人员参阅。

图书在版编目(CIP)数据

数据库原理与技术/尹为民主编. —2版. —北京:科学出版社,2010.1

(高等学校计算机科学与技术系列教材)

武汉大学“十一五”规划教材

ISBN 978-7-03-026415-2

I. ①数… II. ①尹… 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(2010)第009124号

责任编辑:黄金文/责任校对:闫陶

责任印制:彭超/封面设计:苏波

科学出版社 出版

北京东黄城根北街16号

邮政编码:100717

<http://www.sciencep.com>

武汉市新华印刷有限责任公司印刷

科学出版社发行 各地新华书店经销

*

2010年1月第二版 开本:787×1092 1/16

2010年1月第一次印刷 印张:21 1/4

印数:1—2000 字数:480千字

定价:34.00元

(如有印装质量问题,我社负责调换)

《数据库原理与技术》(第二版)编委会

主 编 尹为民

副主编 金银秋

编 委 曾 慧 李石君 邓宏涛

朱三元 孙扬波

前 言

我国高等教育正进入大众化、高水平发展阶段,社会对计算机专业人才呈现出多样化的需求。同时,计算机学科的发展已突破了原有的学科系列框架,形成了在多个专业方向发展的新格局。在这种背景下,教育部高等学校计算机科学与技术教学指导委员会编制了《高等学校计算机科学与技术专业发展战略研究报告暨专业规范(试行)》。该专业规范按照“培养规格分类”的指导思想,提出了三大类、四方向的计算机专业发展建议,给出了《高等学校计算机科学与技术专业公共核心知识体系与课程》的描述。其中,三大类、四方向是:

科学型(计算机科学方向);

工程型(计算机工程方向、软件工程方向);

应用型(信息技术方向)。

上述分类体现了社会对不同人才类型的需求。但无论属于何种类型与方向,数据库系统是其公共核心知识体系中的课程。这些对于指导和规范我国计算机教学改革与课程建设具有重要的意义。

《数据库原理与技术》(第二版)是在前一版的基础上更新、精简而成的,其目的是使该版本更符合教学与应用的需要,并且通过总结本书第一版的使用经验来编著一本更好的数据库基础教材。

与第一版相比,本书各章均进行了适当的修订。主要有:为使本书能够与发展技术相适应,根据应用要求新增加了一章,即第二版的第九章 数据库访问与系统建模方法,其中描述了关于数据库访问的多种途径,介绍了数据库应用系统建模的新方法——UML方法;第六章增加了关于数据库调优的技术,介绍几种常用的方法;现代数据库部分添加了介绍XML数据库、数据仓库与数据挖掘的知识,讲授高年级数据库选修课程的教师可选讲这些相关主题的内容,以使课程内容更加丰富;此外,还扩充了几种数据模型的介绍、增加了SQL类型扩展等内容。

本书的组织结构分两部分,共计十章。第一部分:基础篇,从第一章至第五章,包括数据库基础知识和关系数据库设计;第二部分:系统篇,从第六章至第十章,包括数据库管理与保护、数据库访问与系统建模方法以及数据库系统新技术。

本书是现代数据库系统的基本原理、技术和方法的教科书,已列为武汉大学“十一五”规划教材。具有以下特点:

(1) 根据新出版的《高等学校计算机科学与技术专业发展战略研究报告暨专业规范(试行)》建议,围绕其《高等学校计算机科学与技术专业公共核心知识体系与课程》里有关数据库系统部分的教学大纲与知识单元增删改而成,符合发展要求的知识结构与体系,体现了教学内容改革的新成果。

(2) 既注重系统地介绍数据库的基本原理和方法,又补充新型数据库系统的主要技术及知识,还有理论方法的应用技术。本书缩减传统数据库系统的部分内容,突出数据库理论

与应用技术紧密结合的特征,结合现代的软、硬件环境及应用实例讲解,更适于作为高等学校本科生教材。

(3) 配有丰富的教学辅助资源,如有配套的辅导教材、课程网站(含配套的课件、实习指导和练习测试等)以方便教学。动态更新的课件可从武汉大学计算机学院的精品课程网站上获取,或者从武汉大学课程中心(<http://kczx.whu.edu.cn>)中找到。

本书力求用勤勉的工作、自己的思考诠释专业规范的思想,把新的课程体系和教学内容生动地表述出来,着眼于既帮助学生掌握数据库系统的基本原理、技术和方法,又了解现代数据库系统的特点及发展趋势。本书这次改版作了较多的修改和补充,全书内容丰富,叙述严谨清楚,每章后均有知识层次结构以及适量的思考题和习题,适于广大师生的教与学。本书的第一版,包括多媒体课件已多次用于本科生、双学位等“数据库原理”的教学。其知识结构的更新及开放的实习是数据库原理课程教改项目的内容,该书亦属精品课程建设项目的

一部分。
本书在编写过程中参考了一些书籍及文献资料,在此谨向被引用资料的作者表示感谢。本书存在的不足之处,恳请广大读者批评指正。

作 者

2009年9月 于武汉大学

目 录

前言

第一部分 基础篇	1
第一章 数据库系统概论	3
1.1 信息与数据库	3
1.1.1 基本概念	3
1.1.2 数据管理的发展	5
1.2 数据模型	7
1.2.1 三个世界及其关系	7
1.2.2 概念模型	10
1.2.3 数据模型的组成及分类	13
1.2.4 三种基本数据模型	14
1.2.5 面向对象数据模型	18
1.2.6 数据模型的新特征	19
1.3 数据库系统结构	20
1.3.1 三级模式结构	20
1.3.2 数据库系统的组成	24
1.3.3 数据库系统的特点	26
1.3.4 数据库系统的分类	27
1.4 本章小结	30
第二章 关系数据库	32
2.1 基本概念	32
2.1.1 关系及其性质	32
2.1.2 关系模式与关系数据库	35
2.2 关系模型及其描述	36
2.2.1 关系模型的特点及组成	36
2.2.2 关系数据操作	36
2.2.3 关系的完整性	37
2.3 关系代数	39
2.3.1 集合运算	40
2.3.2 关系运算	42
2.3.3 扩充的关系运算	47
2.4 关系演算	49
2.4.1 元组关系演算	49

2.4.2	域关系演算	50
2.4.3	域关系演算语言 QBE	51
2.5	本章小结	54
第三章 关系数据库语言 SQL		59
3.1	SQL 的特征	59
3.1.1	SQL 的主要标准	59
3.1.2	SQL 的功能特点	60
3.1.3	SQL 的基本组成	61
3.1.4	SQL 环境	62
3.2	SQL 的数据定义	63
3.2.1	模式的定义与撤消	63
3.2.2	表的建立和删除	64
3.2.3	表的扩充和修改	65
3.2.4	自定义的域类型	66
3.2.5	索引的建立与删除	67
3.3	SQL 的数据查询	68
3.3.1	单表查询	69
3.3.2	表达式与函数的使用	71
3.3.3	多表查询	73
3.3.4	自然连接与外连接	79
3.3.5	递归合并查询	81
3.3.6	空值处理	83
3.4	SQL 的数据更新	84
3.4.1	插入数据	84
3.4.2	修改数据	85
3.4.3	删除数据	86
3.5	SQL 中的视图	86
3.5.1	视图的概念及作用	86
3.5.2	创建与使用视图	87
3.5.3	更新视图	90
3.6	过程化 SQL	91
3.6.1	过程化 SQL 的概念	91
3.6.2	过程化 SQL 编程	92
3.6.3	存储过程	95
3.6.4	游标的设计	96
3.7	复杂数据类型	99
3.7.1	构造数据类型	100
3.7.2	大对象数据类型	101

3.8 本章小结	101
第四章 关系数据库理论	104
4.1 函数依赖	104
4.1.1 问题的提出	104
4.1.2 函数依赖的定义	106
4.1.3 函数依赖的分类	107
4.1.4 候选码的形式定义	108
4.2 关系模式的规范化	109
4.2.1 关系与范式	109
4.2.2 第1范式(1NF)	109
4.2.3 第2范式(2NF)	110
4.2.4 第3范式(3NF)	111
4.2.5 BC范式(BCNF)	113
4.2.6 多值依赖与第4范式	113
4.2.7 连接依赖与第5范式	116
4.3 数据依赖的公理系统	118
4.3.1 Armstrong公理	118
4.3.2 公理的完备性	119
4.3.3 属性闭包的计算	121
4.3.4 函数依赖集的等价	123
4.3.5 最小函数依赖集	124
4.4 关系模式的分解	126
4.4.1 等价模式分解的定义	126
4.4.2 无损连接性与依赖保持性	127
4.4.3 模式分解的算法	130
4.5 规范化的问题	133
4.6 本章小结	134
第五章 数据库设计	137
5.1 数据库设计概述	137
5.1.1 数据库设计的特点和方法	137
5.1.2 数据库设计的内容和步骤	139
5.2 需求分析	140
5.2.1 需求分析的任务	141
5.2.2 需求分析的方法	141
5.2.3 数据流图	142
5.2.4 数据字典	143
5.3 概念结构设计	147
5.3.1 概念结构设计的方法和步骤	148

5.3.2	E-R 图的扩展描述	149
5.3.3	局部视图设计	152
5.3.4	视图的集成	153
5.4	逻辑结构设计	156
5.4.1	逻辑结构设计的步骤	156
5.4.2	E-R 模型向关系模型的转换	156
5.4.3	模型的优化	160
5.4.4	子模式设计	161
5.5	数据库物理设计	161
5.5.1	物理设计的内容和要求	161
5.5.2	存取方法与存储结构	162
5.5.3	性能评价	164
5.6	数据库实施与维护	165
5.7	应用实例——录像出租公司数据库设计	165
5.7.1	应用的需求描述	165
5.7.2	需求收集和分析	166
5.7.3	概念结构设计	171
5.7.4	逻辑结构设计	175
5.7.5	物理结构设计	176
5.8	本章小结	179
第二部分 系统篇		181
第六章 数据库管理系统		183
6.1	DBMS 介绍	183
6.1.1	DBMS 的特点与功能	183
6.1.2	DBMS 的基本组成	184
6.1.3	DBMS 的工作与环境	185
6.2	关系数据库的查询优化	186
6.2.1	查询处理与查询优化	186
6.2.2	查询优化的一般策略	189
6.2.3	关系代数表达式的等价规则	190
6.2.4	关系代数表达式的优化算法	192
6.2.5	查询优化的一般步骤	195
6.3	数据库调优	195
6.3.1	数据库模式调优	196
6.3.2	SQL 语句调优	200
6.3.3	数据库调优工具	202
6.4	本章小结	203

第七章 数据库的安全性与完整性	205
7.1 安全性与完整性概述	205
7.1.1 基本概念	205
7.1.2 安全标准	207
7.2 数据库的安全性控制	210
7.2.1 DBMS 安全模型.....	210
7.2.2 存取控制方法	211
7.2.3 自主存取控制的实施	212
7.2.4 视图机制	217
7.2.5 其他安全控制方法	218
7.2.6 Web 数据安全策略.....	219
7.3 数据库的完整性控制	219
7.3.1 完整性控制的功能	219
7.3.2 完整性约束的设计	220
7.4 数据库触发器	223
7.4.1 触发器机制	223
7.4.2 SQL 触发器规则	224
7.4.3 SQL 触发器设计	227
7.5 安全性与完整性示例	230
7.6 本章小结	231
第八章 事务管理	233
8.1 事务的基本概念	233
8.1.1 事务及其生成	233
8.1.2 事务的状态与特性	234
8.2 数据库恢复的基本技术	235
8.2.1 基本概念	235
8.2.2 数据转储	236
8.2.3 登记日志文件	237
8.2.4 故障的种类及恢复策略	238
8.2.5 事务更新与恢复的分类	240
8.3 数据库恢复的扩展技术	241
8.3.1 具有检查点的恢复技术	241
8.3.2 数据库镜像	243
8.3.3 远程备份系统	244
8.4 事务的并发控制	245
8.4.1 并发控制的概念	245
8.4.2 封锁与封锁协议	247
8.4.3 并发调度的可串行性	249

8.4.4	基于时间戳的协议	252
8.4.5	封锁的粒度	254
8.5	数据库事务管理的示例	255
8.6	本章小结	257
第九章	数据库访问与系统建模	260
9.1	数据库访问概述	260
9.2	嵌入式 SQL	260
9.2.1	嵌入式 SQL 概述	260
9.2.2	嵌入式 SQL 的工作原理	261
9.2.3	嵌入式 SQL 的程序组成	263
9.2.4	嵌入式 SQL 的使用技术	265
9.2.5	动态 SQL 方法	268
9.3	JDBC 与 SQLJ	270
9.3.1	调用级接口——JDBC	270
9.3.2	语句级接口——SQLJ	274
9.4	数据库系统建模方法	275
9.4.1	统一建模语言(UML)	275
9.4.2	UML 的表示法	276
9.4.3	UML 的 E-R 建模	277
9.4.4	UML 的高级建模	280
9.4.5	E-R 模型与 UML 表示	282
9.5	数据库应用开发	286
9.5.1	数据库应用系统开发步骤	286
9.5.2	数据库平台及开发工具的选择	287
9.6	本章小结	291
第十章	现代数据库系统	293
10.1	现代数据库系统概述	293
10.1.1	数据库系统的发展	293
10.1.2	现代数据库系统简介	295
10.1.3	现代数据库系统的研究	296
10.2	基于对象的数据库	299
10.2.1	面向对象的数据模型	299
10.2.2	面向对象数据库	301
10.2.3	对象关系数据模型	302
10.2.4	对象关系数据库	303
10.3	基于网络的数据库	304
10.3.1	分布式数据库	304
10.3.2	网络数据库	307

10.3.3 网格环境下的数据库	310
10.4 XML 数据库	313
10.4.1 XML 数据库概念及分类	313
10.4.2 XML 数据模型	314
10.4.3 XML 查询语言	315
10.5 数据仓库与数据挖掘	318
10.5.1 数据仓库	318
10.5.2 联机分析处理	320
10.5.3 数据挖掘	321
10.6 本章小结	321
参考文献	324

第一部分 基础篇



第一章 数据库系统概论

本章主要描述数据库系统的有关基本概念、三级模式结构及系统的组成、特点。讲解几种典型的数据模型,介绍概念模型的表示以及数据库系统的分类。

1.1 信息与数据库

数据库技术产生于20世纪60年代末,当时的“软件危机”促使人们对软件设计的反思,于是数据库技术作为一门软件学科应运而生。它的出现使得计算机应用渗透到工农业生产、商业、行政管理、科学研究、工程技术以及国防军事等各个领域。20世纪80年代微型机的出现,现代网络的产生,使数据库技术得到了广泛的应用和普及。

数据库技术是计算机科学技术中发展最快的领域之一,已经成为计算机信息管理与应用系统的核心技术,它与网络技术构成计算机应用的两个重要平台。现在,数据库系统的建设规模、数据库描述信息量的大小以及网络应用的程度已成为衡量一个部门信息化程度的重要标志。

那么,什么是数据库?什么是数据库系统?数据库理论与技术是怎样发展而来的?这就是本节要介绍的内容。

1.1.1 基本概念

1. 信息与数据

现代社会是信息的社会,信息以惊人的速度增长。因此,如何有效地组织和利用它们成为急需解决的问题。数据库系统的目的就是为了高效地管理及共享大量的信息,而信息与数据是分不开的。

数据是描述事物的符号记录,也是数据库中存储、用户操纵的基本对象。数据不仅是数值,还可以是文字、图形、图像、声音、视频等。数据是信息的符号表示。例如描述新生信息,可用一组数据“2009年,武汉大学计算机学院研究生及本科新生,共18个班,800人”。这些符号被赋予了特定的语义,具体描述了信息,因此它们就具有了传递信息的功能。

信息是有一定含义的,经过加工处理的,对决策有价值的信息。它是关于现实世界事物存在方式或运动状态的反应,它是人类共享的一切知识及客观加工提炼出的各种消息的总和。

信息与数据的关系可以归纳为:数据是信息的载体,信息是数据的内涵。即数据是信息的符号表示,而信息通过数据描述,又是数据语义的解释。尽管两者在概念上不尽相同,但通常使用时并不严格去区分它们。