


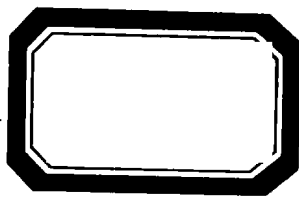
21世纪高等学校规划教材 | 计算机科学与技术



数据库系统原理 及应用

丁忠俊 主编
郭胜 王志 编著

清华大学出版社



等学校规划教材 | 计算机

术

数据库系统原理 及应用

丁忠俊 主编
郭胜 王志 编著

清华大学出版社
北京

内 容 简 介

本书系统全面地介绍了数据库系统的基本理论、实现技术和开发方法。内容包括数据库系统概论、关系数据库、关系数据库标准语言 SQL、关系模式设计理论、数据库设计、数据库的保护、数据库系统的新技术和数据库系统的应用及开发。另外,各章后面选编了许多练习题或实验题,以供复习或上机使用。

本书取材广泛,内容丰富,解析清楚,讲述明确,通俗易懂。本书可以作为高等学校计算机专业、信息和自动控制专业、经济和电子商务专业等学习数据库课程的教材,也可以供从事数据库开发应用的研究人员和工程技术人员参考。

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

数据库系统原理及应用/丁忠俊主编.--北京:清华大学出版社,2012.4

(21世纪高等学校规划教材·计算机科学与技术)

ISBN 978-7-302-28230-3

I. ①数… II. ①丁… III. ①数据库系统 IV. ①TP311.13

中国版本图书馆 CIP 数据核字(2012)第 038931 号

责任编辑:魏江江 薛 阳

封面设计:傅瑞学

责任校对:焦丽丽

责任印制:何 芊

出版发行:清华大学出版社

网 址: <http://www.tup.com.cn>, <http://www.wqbook.com>

地 址:北京清华大学学研大厦 A 座 邮 编:100084

社总机:010-62770175 邮 购:010-62786544

投稿与读者服务:010-62776969, c-service@tup.tsinghua.edu.cn

质量反馈:010-62772015, zhiliang@tup.tsinghua.edu.cn

课 件 下 载: <http://www.tup.com.cn>, 010-62795954

印 刷 者:北京富博印刷有限公司

装 订 者:北京市密云县京文制本装订厂

经 销:全国新华书店

开 本:185mm×260mm 印 张:16.25 字 数:394千字

版 次:2012年4月第1版 印 次:2012年4月第1次印刷

印 数:1~3000

定 价:28.00元

产品编号:042186-01

编审委员会成员

(按地区排序)

清华大学	周立柱	教授
	章 征	教授
	王建民	教授
	冯建华	教授
	刘 强	副教授
北京大学	杨冬青	教授
	陈 钟	教授
	陈立军	副教授
北京航空航天大学	马殿富	教授
	吴超英	副教授
	姚淑珍	教授
中国人民大学	王 珊	教授
	孟小峰	教授
	陈 红	教授
北京师范大学	周明全	教授
北京交通大学	阮秋琦	教授
	赵 宏	副教授
北京信息工程学院	孟庆昌	教授
北京科技大学	杨炳儒	教授
石油大学	陈 明	教授
天津大学	艾德才	教授
复旦大学	吴立德	教授
	吴百锋	教授
	杨卫东	副教授
同济大学	苗夺谦	教授
	徐 安	教授
华东理工大学	邵志清	教授
华东师范大学	杨宗源	教授
	应吉康	教授
东华大学	乐嘉锦	教授
	孙 莉	副教授

浙江大学	吴朝晖	教授
	李善平	教授
扬州大学	李 云	教授
南京大学	骆 斌	教授
	黄 强	副教授
南京航空航天大学	黄志球	教授
	秦小麟	教授
南京理工大学	张功萱	教授
南京邮电学院	朱秀昌	教授
苏州大学	王宜怀	教授
	陈建明	副教授
江苏大学	鲍可进	教授
中国矿业大学	张 艳	教授
武汉大学	何炎祥	教授
华中科技大学	刘乐善	教授
中南财经政法大学	刘腾红	教授
华中师范大学	叶俊民	教授
	郑世珏	教授
	陈 利	教授
江汉大学	颜 彬	教授
国防科技大学	赵克佳	教授
	邹北骥	教授
中南大学	刘卫国	教授
湖南大学	林亚平	教授
西安交通大学	沈钧毅	教授
	齐 勇	教授
长安大学	巨永锋	教授
哈尔滨工业大学	郭茂祖	教授
吉林大学	徐一平	教授
	毕 强	教授
山东大学	孟祥旭	教授
	郝兴伟	教授
中山大学	潘小轰	教授
厦门大学	冯少荣	教授
厦门大学嘉庚学院	张思民	教授
云南大学	刘惟一	教授
电子科技大学	刘乃琦	教授
	罗 蕾	教授
成都理工大学	蔡 淮	教授
	于 春	副教授
西南交通大学	曾华桑	教授

出版说明

随着我国改革开放的进一步深化,高等教育也得到了快速发展,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,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投入力度,通过教育改革合理调整和配置了教育资源,优化了传统学科专业,积极为地方经济建设输送人才,为我国经济社会的快速、健康和可持续发展以及高等教育自身的改革发展做出了巨大贡献。但是,高等教育质量还需要进一步提高以适应经济社会发展的需要,不少高校的专业设置和结构不尽合理,教师队伍整体素质亟待提高,人才培养模式、教学内容和方法需要进一步转变,学生的实践能力和创新精神亟待加强。

教育部一直十分重视高等教育质量工作。2007年1月,教育部下发了《关于实施高等学校本科教学质量与教学改革工程的意见》,计划实施“高等学校本科教学质量与教学改革工程”(简称“质量工程”),通过专业结构调整、课程教材建设、实践教学改革、教学团队建设等多项内容,进一步深化高等学校教学改革,提高人才培养的能力和水平,更好地满足经济社会发展对高素质人才的需要。在贯彻和落实教育部“质量工程”的过程中,各地高校发挥师资力量强、办学经验丰富、教学资源充裕等优势,对其特色专业及特色课程(群)加以规划、整理和总结,更新教学内容、改革课程体系,建设了一大批内容新、体系新、方法新、手段新的特色课程。在此基础上,经教育部相关教学指导委员会专家的指导和建议,清华大学出版社在多个领域精选各高校的特色课程,分别规划出版系列教材,以配合“质量工程”的实施,满足各高校教学质量和教学改革的需要。

为了深入贯彻落实教育部《关于加强高等学校本科教学工作,提高教学质量的若干意见》精神,紧密配合教育部已经启动的“高等学校教学质量与教学改革工程精品课程建设工作”,在有关专家、教授的倡议和有关部门的大力支持下,我们组织并成立了“清华大学出版社教材编审委员会”(以下简称“编委会”),旨在配合教育部制定精品课程教材的出版规划,讨论并实施精品课程教材的编写与出版工作。“编委会”成员皆来自全国各类高等学校教学与科研第一线的骨干教师,其中许多教师为各校相关院、系主管教学的院长或系主任。

按照教育部的要求,“编委会”一致认为,精品课程的建设工作从开始就要坚持高标准、严要求,处于一个比较高的起点上。精品课程教材应该能够反映各高校教学改革与课程建设的需要,要有特色风格、有创新性(新体系、新内容、新手段、新思路,教材的内容体系有较高的科学创新、技术创新和理念创新的含量)、先进性(对原有的学科体系有实质性的改革和发展,顺应并符合21世纪教学发展的规律,代表并引领课程发展的趋势和方向)、示范性(教材所体现的课程体系具有较广泛的辐射性和示范性)和一定的前瞻性。教材由个人申报或各校推荐(通过所在高校的“编委会”成员推荐),经“编委会”认真评审,最后由清华大学出版

社审定出版。

目前,针对计算机类和电子信息类相关专业成立了两个“编委会”,即“清华大学出版社计算机教材编审委员会”和“清华大学出版社电子信息教材编审委员会”。推出的特色精品教材包括:

(1) 21 世纪高等学校规划教材·计算机应用——高等学校各类专业,特别是非计算机专业的计算机应用类教材。

(2) 21 世纪高等学校规划教材·计算机科学与技术——高等学校计算机相关专业的教材。

(3) 21 世纪高等学校规划教材·电子信息——高等学校电子信息相关专业的教材。

(4) 21 世纪高等学校规划教材·软件工程——高等学校软件工程相关专业的教材。

(5) 21 世纪高等学校规划教材·信息管理与信息系统。

(6) 21 世纪高等学校规划教材·财经管理与应用。

(7) 21 世纪高等学校规划教材·电子商务。

(8) 21 世纪高等学校规划教材·物联网。

清华大学出版社经过三十多年的努力,在教材尤其是计算机和电子信息类专业教材出版方面树立了权威品牌,为我国的高等教育事业做出了重要贡献。清华版教材形成了技术准确、内容严谨的独特风格,这种风格将延续并反映在特色精品教材的建设中。

清华大学出版社教材编审委员会

联系人:魏江江

E-mail: weijj@tup.tsinghua.edu.cn

数据库技术是计算机应用领域中发展最快、应用最广的科学技术之一。数据库系统已成为信息系统和应用系统不可缺少的核心组成部分。数据库的应用已从数据处理、信息管理、事务处理扩大到计算机辅助设计、决策支持、人工智能和网络应用等领域。数据库系统的推广使用也使得计算机的应用迅速地渗透到各行各业和各个部门,如军事国防、航天航空、金融工商、交通能源、通信测控、文教卫生等,直至影响到人们的工作方式和生活方式。在高等学校中,数据库系统不仅是计算机专业重要的专业课程之一,而且也是信息、自控、经济、电子商务等相关专业必修的计算机应用课程。

本书是在参考了全日制高等学校本科数据库教学大纲的基础上,结合作者多年从事数据库课程的教学体会和科研实践成果编写而成的。本书旨在将数据库的基本理论、系统实现技术和应用开发方法紧密结合起来,以解析的观点,从应用的角度,站在开发与实现的立场来进行讨论。以求由浅入深,理论联系实际,通俗易懂地讨论数据库系统,重点是关系数据库系统的功能、结构、设计理论和实现方法,以及组织和开发过程。为大专院校学生和从事计算机应用的人员提供一本学能致用的教材和应用开发的参考书。

全书共分 8 章:

第 1 章 数据库系统概论。主要介绍数据库系统的基本概念、数据模型、数据库系统结构和功能等内容。

第 2 章 关系数据库。主要介绍关系数据库结构、关系的数学运算:关系代数运算和关系演算等内容。

第 3 章 关系数据库标准语言 SQL。主要介绍 SQL 语言结构、SQL 的数据定义、数据查询和数据更新、嵌入式 SQL 和动态 SQL 以及存储过程等内容。

第 4 章 关系模式设计理论。主要介绍函数依赖、模式分解的特性、范式和范式化等内容。

第 5 章 数据库设计。主要介绍数据库设计的步骤,概念结构设计,逻辑结构设计,物理设计,数据库的实施、运行和维护等内容。

第 6 章 数据库的保护。主要介绍事务的并发控制、数据库的完整性、数据库的安全性、数据库的恢复等内容。

第 7 章 数据库系统的新技术。主要介绍分布式数据库系统、对象关系数据库系统、多媒体数据库系统、数据仓库与数据挖掘等内容。

第 8 章 数据库系统的应用与开发。主要介绍 SQL Server 2005 集成环境和一个学生成绩管理系统的开发过程等内容。

本书第 1 章和第 4 章由丁忠俊编写,第 3 章、第 5 章和第 8 章由郭胜编写,第 2 章、第 6



章和第7章由王志编写。

丁忠俊负责全书的编审、总纂和统稿工作。

在成书过程中,得到了华中科技大学文华学院的大力支持,在此表示衷心的感谢。

由于作者水平有限,不足之处在所难免,恳请读者批评指正。

丁忠俊 郭胜 王志

2012年1月



目 录

第 1 章 数据库系统概论	1
1.1 数据管理技术	1
1.1.1 数据管理技术的发展.....	1
1.1.2 数据库管理技术.....	5
1.2 数据模型	8
1.2.1 数据模型概述.....	8
1.2.2 概念模型	11
1.2.3 层次模型	14
1.2.4 网状模型	17
1.2.5 关系模型	19
1.3 数据库系统结构.....	22
1.3.1 数据库系统的体系结构	22
1.3.2 数据库系统的三级模式结构	25
1.4 数据库管理系统.....	27
1.4.1 DBMS 的功能	27
1.4.2 DBMS 组成	28
1.4.3 DBMS 工作过程	29
1.5 数据库系统的组成.....	30
1.6 典型 RDBMS 产品介绍.....	32
1.6.1 Oracle	32
1.6.2 DB2	33
1.6.3 Sybase	33
1.6.4 SQL Server	34
1.6.5 MySQL	35
小结	35
习题 1	37
第 2 章 关系数据库	39
2.1 关系数据结构.....	39
2.1.1 关系	39
2.1.2 关系模式	41
2.1.3 关系数据库的概念	41



2.2	关系的完整性	41
2.2.1	实体完整性	41
2.2.2	参照完整性	42
2.2.3	用户定义的完整性	43
2.3	关系代数	43
2.3.1	关系代数的5种基本运算	44
2.3.2	关系代数的4种组合运算	48
2.3.3	关系代数表达式的优化	51
2.4	关系演算	56
2.4.1	元组关系演算	57
2.4.2	域关系演算	59
	小结	60
	习题2	61
第3章	关系数据库标准语言 SQL	63
3.1	SQL 概述	63
3.1.1	SQL 简介	63
3.1.2	SQL 数据库结构	63
3.1.3	SQL 的组成及特点	64
3.2	SQL 的数据定义	66
3.2.1	模式的创建与删除	68
3.2.2	SQL 的数据类型	69
3.2.3	基本表的创建、删除与修改	69
3.2.4	索引的创建与删除	72
3.3	SQL 的数据查询	74
3.3.1	SELECT 语句的结构	74
3.3.2	单表查询	75
3.3.3	关联查询	80
3.3.4	嵌套查询	82
3.4	SQL 的数据更新	85
3.4.1	数据的插入	85
3.4.2	数据的删除	87
3.4.3	数据的修改	87
3.5	视图	88
3.5.1	视图的创建与删除	88
3.5.2	视图的查询	90
3.5.3	视图的更新	91
3.5.4	视图的作用	91
3.6	嵌入式 SQL	92

3.6.1	嵌入式 SQL 的处理过程	92
3.6.2	嵌入式 SQL 的使用规定	92
3.6.3	嵌入式 SQL 的使用技术	93
3.7	动态 SQL 语句	97
3.7.1	使用 SQL 语句主变量	97
3.7.2	使用动态参数	97
3.8	存储过程	98
3.8.1	存储过程的概念	98
3.8.2	存储过程的操作	98
	小结	100
	习题 3	101
第 4 章	关系模式设计理论	104
4.1	关系模式中数据冗余和操作异常问题	104
4.2	函数依赖	106
4.2.1	函数依赖的定义	106
4.2.2	函数依赖的类型	107
4.2.3	关键字	108
4.2.4	FD 公理	108
4.2.5	属性集的闭包	110
4.2.6	FD 集的等价与最小依赖集	112
4.3	关系模式的分解	113
4.3.1	模式分解的两个特性	113
4.3.2	无损连接的分解	115
4.3.3	无损连接分解的判定	116
4.3.4	保持函数依赖的分解	118
4.4	关系模式的范式及规范化	120
4.4.1	第一范式	120
4.4.2	第二范式	121
4.4.3	第三范式	123
4.4.4	BC 范式	124
4.4.5	保持无损连接性的 BCNF 分解算法	126
4.4.6	保持无损连接和函数依赖的 3NF 合成算法	127
4.5	多值依赖与第四范式	128
4.5.1	多值依赖	128
4.5.2	FD 和 MVD 完备的公理系统	129
4.5.3	第四范式	130
	小结	130
	习题 4	131

第 5 章 数据库设计	135
5.1 数据库设计的步骤	135
5.2 需求分析	136
5.2.1 需求分析的任务.....	136
5.2.2 需求分析的方法.....	137
5.3 概念结构设计	138
5.3.1 概念结构设计的步骤.....	138
5.3.2 设计局部的 E-R 模型	139
5.3.3 设计全局的 E-R 模型	143
5.4 逻辑结构设计	146
5.4.1 E-R 模型向关系模型的转换.....	147
5.4.2 关系数据模型的优化.....	149
5.4.3 设计用户子模式.....	150
5.5 物理结构设计	151
5.5.1 物理设计的主要内容.....	151
5.5.2 关系数据库的存取方法.....	151
5.6 数据库的实施	153
5.7 数据库的运行与维护	154
小结.....	155
习题 5	156
第 6 章 数据库的保护	158
6.1 事务	158
6.1.1 事务的定义.....	158
6.1.2 事务的特性.....	158
6.2 事务的并发控制	159
6.2.1 并发操作中的三个问题.....	159
6.2.2 封锁技术.....	160
6.2.3 并发调度与两段封锁协议.....	162
6.3 数据库的完整性	165
6.3.1 数据完整性概念.....	165
6.3.2 数据库完整性的实施定义.....	165
6.3.3 数据库完整性的实施约束.....	166
6.3.4 数据库完整性的实施规则.....	170
6.4 数据库的安全性	171
6.4.1 安全性问题.....	172
6.4.2 数据库安全控制.....	175
6.4.3 SQL Server 的安全机制	177

6.4.4	Oracle 的安全机制	178
6.4.5	安全数据库的研究方向	180
6.5	数据库的恢复	181
6.5.1	故障类型	181
6.5.2	数据库的备份	182
6.5.3	日志文件	183
6.5.4	故障恢复的方法	184
6.5.5	数据库镜像	186
小结	187
习题 6	187
第 7 章	数据库系统的新技术	188
7.1	概述	188
7.1.1	传统数据库系统的局限性	188
7.1.2	数据库技术与相关技术的结合	190
7.2	分布式数据库系统	191
7.2.1	分布式数据库系统的结构	191
7.2.2	分布式数据库系统的特点	193
7.3	对象关系数据库系统	196
7.3.1	面向对象模型	196
7.3.2	对象关系数据库	197
7.4	多媒体数据库系统	202
7.5	数据仓库与数据挖掘	204
7.5.1	数据仓库	204
7.5.2	数据挖掘	211
7.5.3	数据仓库与数据挖掘的关系	216
小结	216
习题 7	217
第 8 章	数据库系统的应用与开发	218
8.1	SQL Server 2005 集成环境	218
8.1.1	SQL Server 2005 系统简介	218
8.1.2	SQL Server 2005 的主要工具	218
8.1.3	SQL Server 2005 数据类型	221
8.2	学生成绩管理系统的开发过程	222
8.2.1	创建数据库和表结构	222
8.2.2	VB 6.0 连接 SQL Server 2005 数据库	225
8.2.3	系统登录及权限的实现	226
8.2.4	系统主界面设计过程	228

8.2.5	学生信息的录入	230
8.2.6	学生信息的更新	233
8.2.7	学生信息的查询	235
8.2.8	学生成绩的录入	236
8.2.9	学生成绩的查询	240
小结		243

第 1 章

数据库系统概论

自 20 世纪 60 年代以来,数据管理已成为计算机的主要应用领域。数据库技术作为数据管理中的核心技术,已成为计算机软件领域中的一个重要分支。它的出现极大地提升了计算机数据处理的能力和数据库管理的水平,不仅拓宽了计算机的应用领域,同时也使数据库管理的水平提高到了一个更高的层次。

本章主要从整体上介绍数据库系统的基本概念、结构及功能,使读者从中领悟到数据库系统管理数据的重要作用。

1.1 数据管理技术

数据管理是指对数据的分类、组织、存储、加工、检索、传递和维护等操作,这些操作是数据管理中的中心问题。数据量越大、数据结构越复杂,其管理数据的难度就越大,要求数据库管理的技术水平也就越高。数据管理技术是随着计算机应用范围的不断扩大、对数据管理特性及处理要求的不断提高,而逐步地产生和发展起来的。

1.1.1 数据管理技术的发展

随着计算机硬件和软件的发展,以及人们对计算机数据处理的要求,数据管理技术的发展经历了三个阶段:人工管理阶段、文件系统管理阶段和数据库系统管理阶段。

1. 人工管理阶段

在 20 世纪 50 年代中期以前,计算机主要用于科学计算。当时的计算机硬件没有磁盘等直接存取存储设备,只有磁带、卡片和纸带等外部存储器;而软件没有操作系统,也没有数据管理方面的软件。数据处理的方式是批处理。数据的组织和管理由人工完成。

人工数据管理有下列特点。

(1) 数据不保存在计算机内。计算机主要用于计算,一般不需要长期保存数据。在计算某一课题任务时,将原始数据随程序一起输入内存,运算处理并将结果数据输出后,数据和程序也同时被撤销。

(2) 没有统一的数据管理软件。主要通过应用程序管理数据,程序员既要规定数据的逻辑结构,又要设计数据的物理结构,包括存储结构、存取方法和输入方式等。程序员的负担很重。

(3) 数据面向应用程序,即一组数据对应一个程序。需要相同数据的多个程序,它们都要各自定义这些数据。因此数据不能共享,在程序之间存在大量的数据冗余。

(4) 应用程序依赖于数据,一旦数据的逻辑结构或物理结构发生了变化,应用程序就必须要做相应的修改,因此数据不具备独立性。

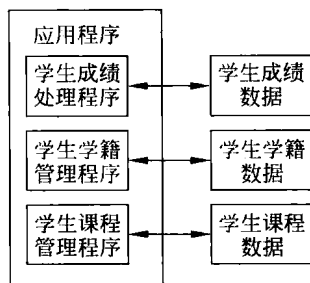


图 1.1 应用程序与数据
一一对应的关系

人工管理阶段应用程序与数据一一对应的关系如图 1.1 所示。

2. 文件系统管理阶段

20 世纪 50 年代后期至 60 年代中期,计算机的应用范围逐渐扩大,计算机不仅用于科学计算,还大量用于信息管理。此时,计算机硬件出现了磁盘、磁鼓等直接存取存储设备;软件出现了操作系统,操作系统中的文件系统是专门管理外存的数据管理软件。数据处理的方式有批处理,也有联机实时处理。

文件系统管理数据的特点:

(1) 数据以文件的形式可长期保存在磁盘上,供应用程序反复进行查询、修改、插入和删除等操作。

(2) 由文件系统管理数据,实现了“按文件名访问,按记录存取”的数据管理技术。程序与文件之间具有“设备独立性”,即当文件的物理存储位置发生改变时,用户不必关心数据的物理位置,也不必改变应用程序。

(3) 数据不再属于某个特定的程序,可以重复使用。但是程序是基于特定的文件结构和存取方法,只有符合文件的结构和存取方法的程序才能使用该文件。因此程序与文件之间的依赖关系并未根本改变。

(4) 文件之间相互独立,数据之间的联系是通过程序去构造的。

文件系统管理阶段文件与程序之间的对应关系如图 1.2 所示。

文件之间的相互独立性,以及文件与程序之间的依赖性,使得用文件系统组织和管理数据会呈现出如下的缺陷。

(1) 数据冗余度大、共享性差,数据之间的联系弱。由于文件之间缺乏联系,不同的文件中可能存在大量的相同数据,形成数据的冗余。数据冗余不仅浪费存储空间,更重要的是会造成对数据操作的不一致性。数据的不一致性极大地阻碍着信息的交流和共享。一个文件面向一个或几个应用程序,而不是面向整个系统的应用,因此文件的共享性比较差。数据之间联系弱是由于文件之间相互独立、缺乏联系造成的。

(2) 文件与程序之间缺乏“数据独立性”。文件的逻辑结构与对应的程序结构相互依赖。当文件的逻辑结构发生变化时,其程序结构也要做相应的修改,即文件与程序之间缺乏“数据独立性”。

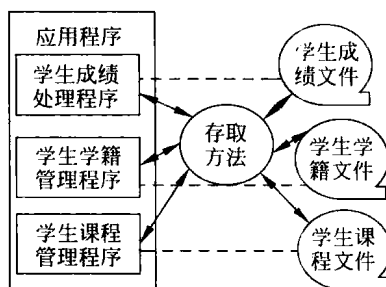


图 1.2 文件与程序之间的对应关系