

普通高校本科计算机专业

特色

教材精选

数据库原理及设计

陶宏才 主编

2

<http://www.tup.com.cn>

清华大学出版社



普通高校本科计算机专业 **特色** 教材精选

数据库原理及设计

陶宏才 主编

清华大学出版社
北 京

内 容 简 介

本书是作者在长期数据库教学和项目开发基础上,为高等学校计算机及相关专业编写的数据库教材。本书对数据库的原理、应用与设计3个方面的内容进行了深入浅出和全新的诠释。主要包括:数据库系统概述,实体联系数据模型,关系数据模型,SQL语言及其操作,数据库的保护,关系数据库设计理论,数据库应用设计,数据库应用系统设计实例,主流数据库产品、工具及比较,数据仓库与数据挖掘及数据库新进展,数据库上机实验及指导。

本书以数据库系统的核心——DBMS的出现背景为线索,引出了数据库的相关概念及数据库的整个框架体系,理顺了数据库原理、应用与设计之间的有机联系。本书突出理论产生的背景和根源,强化理论与商用RDBMS产品,以及理论与应用开发的结合,重视知识的实用,跟踪数据库技术发展前沿,反映最新的主流数据库产品,并免费提供配套的电子课件。

本书逻辑性、系统性、实践性和实用性强,可作为计算机及相关专业本科生及研究生教材,也可作为从事数据库开发的专业人员的参考书。

版权所有,翻印必究。

本书封面贴有清华大学出版社激光防伪标签,无标签者不得销售。

图书在版编目(CIP)数据

数据库原理及设计/陶宏才主编. —北京:清华大学出版社,2004.2

(普通高校本科计算机专业特色教材精选)

ISBN 7-302-07933-1

I. 数… II. 陶… 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(2004)第001129号

出 版 者: 清华大学出版社

<http://www.tup.com.cn>

社 总 机: 010-62770175

地 址: 北京清华大学学研大厦

邮 编: 100084

客户服务: 010-62776969

组稿编辑: 王听讲

文稿编辑: 王冰飞

印 装 者: 清华大学印刷厂

发 行 者: 新华书店总店北京发行所

开 本: 185×260 印张: 25.25 字数: 572千字

版 次: 2004年2月第1版 2004年2月第1次印刷

书 号: ISBN 7-302-07933-1/TP·5758

印 数: 1~6000

定 价: 32.00元

本书如存在文字不清、漏印以及缺页、倒页、脱页等印装质量问题,请与清华大学出版社出版部联系调换。联系电话:(010)62770175-3103或(010)62795704

编审委员会

主任：蒋宗礼

副主任：李仲麟 何炎祥

委员：（排名不分先后）

王向东 宁 洪 朱庆生 吴功宜 吴 跃

张 虹 张 钢 张为群 余雪丽 陈志国

武 波 孟祥旭 孟小峰 胡金初 姚放吾

原福永 黄刘生 廖明宏 薛永生

秘书长：王听讲

AJS-256/06

出版说明

INTRODUCTION

在我国高等教育逐步实现大众化后，越来越多的高等学校将会面向国民经济发展的第一线，为行业、企业培养各级各类高级应用型专门人才。为此，教育部已经启动了“高等学校教学质量和教学改革工程”，强调要以信息技术为手段，深化教学改革和人才培养模式改革。如何根据社会的实际需要，根据各行各业的具体人才需求，培养具有特色显著的人才，是我们共同面临的重大问题。具体地，培养具有一定专业特色的和特定能力强的计算机专业应用型人才则是计算机教育要解决的问题。

为了适应 21 世纪人才培养的需要，培养具有特色的计算机人才，急需一批适合各种人才培养特点的计算机专业教材。目前，一些高校在计算机专业教学和教材改革方面已经做了大量工作，许多教师在计算机专业教学和科研方面已经积累了许多宝贵经验。将他们的教研成果转化为教材的形式，向全国其他学校推广，对于深化我国高等学校的教学改革是一件十分有意义的事。

清华大学出版社在经过大量调查研究的基础上，决定编写出版一套“普通高校本科计算机专业特色教材精选”。本套教材是针对当前高等教育改革的新形势，以社会对人才的需求为导向，主要以培养应用型计算机人才为目标，立足课程改革和教材创新，广泛吸纳全国各地的高等院校计算机优秀教师参与编写，从中精选出版确实反映计算机专业教学方向的特色教材，供普通高等院校计算机专业学生使用。

本套教材具有以下特点：

1. 编写目的明确

本套教材是深入研究各地各学校办学特色的基础上，面向普通高校的计算机专业学生编写的。学生通过本套教材，主要学习计算机科学与技术专业的基本理论和基本知识，接受利用计算机解决实际问题的基本训练，培养研究和开发计算机系统，特别是应用系统的基本能力。

2. 理论知识与实践训练相结合

根据计算学科的三个学科形态及其关系,本套教材力求突出学科的理论与实践紧密结合的特征,结合实例讲解理论,使理论来源于实践,又进一步指导实践,学生通过实践深化对理论的理解,更重要的是使学生学会理论方法的实际运用。在编写教材时突出实用性,并做到通俗易懂,易教易学,使学生不仅知其然,知其所以然,还要会其如何然。

3. 注意培养学生的动手能力

每种教材都增加了能力训练部分的内容,学生通过学习和练习,能比较熟练地应用计算机知识解决实际问题。既注重培养学生分析问题的能力,也注重培养学生解决问题的能力,以适应新经济时代对人才的需要,满足就业要求。

4. 注重教材的立体化配套

大多数教材都将陆续配套教师用课件、习题及其解答提示,学生上机实验指导等辅助教学资源,有些教材还提供能用于网上下载的文件,以方便教学。

由于各地区各学校的培养目标、教学要求和办学特色均有所不同,所以对特色教学的理解也不尽一致,我们恳切希望大家在使用教材的过程中,及时地给我们提出批评和改进意见,以便我们做好教材的修订改版工作,使其日趋完善。

我们相信经过大家的共同努力,这套教材一定能成为特色鲜明、质量上乘的优秀教材,同时,我们也希望通过本套教材的编写出版,为“高等学校教学质量和教学改革工程”作出贡献。

清华大学出版社

前言

PREFACE

本教材是为跟上目前数据库发展新形势、反映数据库开发新技术、体现数据库教学新思维而编写的。为此,本教材对数据库的教学内容、结构进行了调整、取舍和更新,赋予内容以新的视角和主线,力图使学生通过背景知识的了解、理论与实际产品和实践相结合,来掌握、运用和开发数据库,同时了解数据库发展的新方向。

本书以关系数据库为主,以 RDBMS 为核心,将数据库的主体内容划分为原理、应用和设计三大部分,并通过数据库系统总体结构,将三部分内容有机地统一于一个主体框架内,全书的内容均围绕着这个主体框架来组织。各章内容如下:

第 1 章为数据库系统概述。内容包括:DBMS 出现的背景、数据库系统的抽象层次、数据库语言与 SQL/ODBC/JDBC、数据库系统的总体结构、数据库应用系统的 C/S 与 B/S 模式和最新开发技术、数据库系统中的关键技术语与概念、数据库系统的用户、由应用需求看数据库技术的发展。

第 2 章是实体联系数据模型。内容包括:关于数据模型的几个重要问题、数据库设计概述、实体联系模型、扩展实体联系模型、利用 ER 模型的概念数据库设计。

第 3 章对目前占主导地位的关系模型,以及对其理论基础之一的关系代数和关系运算做了较全面的描述。其中,详细介绍了在数据库应用和设计中,用得相当多的关系模型的完整性约束以及在 Sybase 和 MS SQL Server 中的具体体现和支持;对关系代数及运算用大量例子进行讲解,同时与 SQL 语言查询进行对比。

第 4 章着重从实用角度,通过列举大量示例并结合 Sybase 和 MS SQL Server 的 T-SQL,对结构化查询语言 SQL 进行了比较详细的介绍。

第 5 章主要从原理和应用的角度,分别介绍了数据库的安全性、完整性、故障恢复和并发控制,是数据库课程的必学内容之一。

第 6 章涉及的是关系模型的理论基础之二,该理论是指导数据库设计的重要依据,揭示了关系数据中的一些特性——函数依赖、多值依赖和

连接依赖,以及由此引起的诸多问题。

第7章系统地介绍如何通过数据库的需求分析、概念设计、逻辑设计和物理设计等若干步骤一步一步地将企业的管理业务、数据等转变成数据库管理系统所能接受的形式,从而达到利用计算机管理信息的目的。

第8章用一个实际的应用系统开发的实例,详细展示其中的精髓。通过遵从本章的设计、构建和开发步骤,完成从理论到实践的跨越。

第9章对目前数据库市场上比较活跃的主流数据库厂商、最新产品及工具,作了较全面的介绍。

第10章是数据仓库与数据挖掘及数据库新进展。内容包括:数据仓库技术、数据挖掘技术、数据库技术的研究与发展。

第11章为数据库上机实验及指导。本章为配合教学,同时也使学生能更好地掌握和运用数据库,有针对性地罗列出一批上机实验,并给出相应的实验指导。

本书由陶宏才主编。本书第1章、第2章、第3章、第4章的4.1~4.4节、第5章、第6章、第9章、附录A、附录C由陶宏才编写;第4章的4.5节、第8章、第10章由陈安龙编写;第7章由张跃编写;第11章、附录B由梁斌梅编写。陶宏才制定了编写大纲,并最后对全书文稿进行了统纂和定稿。

在本书编写过程中,西南交通大学计算机与通信工程学院的周荣辉教授与何大可教授、西南交通大学软件学院的尹治本教授对本书的编写提出了宝贵的建议,作者在此对他们表示衷心的感谢。

由于作者水平有限,书中难免会存在缺点和错误,敬请读者及各位专家指教。

作 者
2003.10 于成都

目 录

CONTENTS

第 1 章 数据库系统概述	1
1.1 数据库系统及其总体结构	2
1.1.1 数据库管理系统出现的背景	2
1.1.2 数据库系统的抽象层次	9
1.1.3 数据库语言与 SQL/ODBC/JDBC	11
1.1.4 数据库系统总体结构	14
1.2 数据库系统中的关键术语与概念	15
1.2.1 数据库及其相关概念	15
1.2.2 视图及其相关概念	17
1.2.3 数据抽象、数据模型、数据模式及其相互关系	20
1.2.4 传统数据模型回顾	22
1.3 数据库系统的用户	24
1.4 数据库应用系统开发概述	25
1.4.1 C/S 模式	25
1.4.2 B/S 模式	26
1.4.3 组件与分布式对象开发技术	27
1.4.4 CGI/ISAPI/NSAPI	31
1.4.5 ASP/JSP/PHP	32
1.5 由应用需求看数据库技术的发展	34
1.5.1 基于文件系统的数据管理	34
1.5.2 第一代数据库系统	35
1.5.3 第二代数据库系统	35
1.5.4 OLTP 及 OLAP	36
1.5.5 数据仓库与数据挖掘	37
1.5.6 并行与分布式数据库系统	38
1.5.7 Internet/Web 数据库	39
1.5.8 面向对象的数据库系统	40

本章小结	41
习题	42
第2章 实体联系数据模型	43
2.1 关于数据模型的几个重要问题	43
2.2 数据库设计概述	45
2.3 实体联系模型	47
2.3.1 实体、实体型及属性	47
2.3.2 联系及联系型	49
2.3.3 ERM 中的完整性约束	50
2.3.4 联系型属性的移动处理	56
2.3.5 弱实体	56
2.3.6 ERM 各元素图示	57
2.4 扩展实体联系模型	58
2.4.1 类层次	58
2.4.2 演绎与归纳	59
2.4.3 演绎的原则	59
2.4.4 聚集	60
2.5 利用 ER 模型的概念数据库设计	61
2.5.1 实体与属性的取舍	61
2.5.2 属性在实体与联系间的取舍	61
2.5.3 二元联系与三元联系的取舍	62
2.5.4 三元联系与聚集的取舍	63
2.5.5 大型系统的概念数据库设计方法	63
2.5.6 应用实例	64
本章小结	65
习题	66
第3章 关系数据模型	67
3.1 关系数据模型的基本概念	67
3.1.1 SQL 语言简介	68
3.1.2 关系模型	68
3.1.3 SQL Server 和 Sybase 支持的完整性约束及其设定	75
3.1.4 视图及其操作	86
3.1.5 实体联系模型向关系模型的转换	89
3.2 关系代数及关系运算	91
3.2.1 关系代数	91
3.2.2 关系运算	108

本章小结.....	112
习题.....	113
第 4 章 SQL 语言及其操作	115
4.1 SQL 语言概述	115
4.1.1 SQL 语言及其标准	115
4.1.2 SQL 语言的特点	116
4.1.3 SQL-92 标准的分级	117
4.1.4 标准 SQL 语言与实际数据库产品中的 SQL 语言	117
4.2 数据定义子语言及其操作	118
4.2.1 定义子语言的 3 个命令关键字.....	118
4.2.2 定义数据库.....	119
4.2.3 定义表.....	122
4.2.4 定义视图.....	125
4.2.5 定义索引.....	127
4.3 数据操纵子语言及其操作	129
4.3.1 数据操纵子语言的 4 个命令关键字.....	129
4.3.2 数据插入.....	129
4.3.3 数据修改.....	130
4.3.4 数据删除.....	132
4.3.5 数据查询.....	133
4.4 Sybase 和 MS SQL Server 中的 T-SQL 语言	138
4.4.1 T-SQL 语言简介	138
4.4.2 T-SQL 的数据类型	139
4.4.3 T-SQL 编程	141
4.4.4 T-SQL 提供的函数	145
4.4.5 T-SQL 游标	149
4.4.6 T-SQL 存储过程	152
4.5 嵌入式 SQL	156
4.5.1 嵌入式 SQL 语句的基本形式	156
4.5.2 嵌入式 SQL 与宿主语言的通信	159
4.5.3 SQL 通信区	159
4.5.4 主变量的定义与使用.....	160
4.5.5 嵌入 SQL 中的游标定义与使用	164
4.5.6 C 语言中的嵌入式 SQL 实例	165
本章小结.....	170
习题.....	171

第 5 章 数据库的保护	173
5.1 数据库的保护概述	174
5.1.1 破坏数据库的类型	174
5.1.2 DBMS 对数据库的保护措施	174
5.2 数据库安全性	175
5.2.1 数据库安全性的概念	175
5.2.2 自主式访问控制的授权与撤权	178
5.2.3 Sybase 及 MS SQL Server 的安全体系及其设置	179
5.2.4 数据库安全性的其他相关内容	182
5.3 数据库完整性	183
5.3.1 数据库完整性的概念	184
5.3.2 完整性约束的类型	184
5.3.3 完整性约束的定义	185
5.3.4 完整性约束的验证	188
5.3.5 SQL-92 和数据库产品对完整性的支持	189
5.4 故障恢复技术	189
5.4.1 事务管理概述	189
5.4.2 故障恢复导论	193
5.4.3 日志结构	198
5.4.4 DBMS 围绕更新事务的工作	198
5.4.5 事务内消息的处理	202
5.4.6 故障类型及恢复对策	203
5.4.7 MS SQL Server 中的事务及故障恢复	205
5.5 并发控制	208
5.5.1 并发控制导论	208
5.5.2 并发执行可能引起的问题	210
5.5.3 并发控制的正确性准则	212
5.5.4 基于锁的并发控制协议	215
5.5.5 多粒度加锁协议	220
5.5.6 死锁及其预防、检测与处理	224
5.5.7 MS SQL Server 中的并发控制技术	227
本章小结	230
习题	230
第 6 章 关系数据库设计理论	233
6.1 关系模式中可能存在的异常	233
6.1.1 存在异常的关系模式示例	233
6.1.2 可能存在的异常	235

6.2	关系模式中存在异常的原因	235
6.3	函数依赖	236
6.3.1	函数依赖定义	236
6.3.2	函数依赖分类及其定义	237
6.3.3	其他相关定义	238
6.3.4	函数依赖示例	238
6.3.5	Armstrong 公理系统	239
6.4	关系模式的规范形式	240
6.4.1	范式	240
6.4.2	规范化	240
6.4.3	以函数依赖为基础的范式	241
6.4.4	多值依赖与第四范式	245
6.4.5	连接依赖与第五范式	247
6.5	关系模式的规范化	249
6.5.1	规范化步骤	249
6.5.2	关系模式的分解及其指标	250
	本章小结	251
	习题	252
第7章	数据库应用设计	253
7.1	数据库应用设计的步骤	254
7.2	数据库设计工具简介	255
7.2.1	Oracle Case	256
7.2.2	ERwin	256
7.2.3	SmartER	257
7.2.4	InfoModeler	257
7.3	用户需求描述与分析	257
7.3.1	需求分析的内容	257
7.3.2	用户需求调研的方法	258
7.3.3	用户需求描述与分析	259
7.3.4	用户需求描述与分析实例	259
7.4	概念设计	263
7.4.1	概念设计的方法	263
7.4.2	视图设计	264
7.4.3	视图集成	264
7.5	逻辑设计	265
7.5.1	ER 图向关系模型的转换	265
7.5.2	数据模式的优化	268

7.5.3 设计用户外模式	270
7.6 物理设计	270
7.6.1 索引的有关概念	270
7.6.2 聚簇索引与非聚簇索引	271
7.6.3 建立索引	272
7.7 数据库实施	273
7.7.1 数据载入	273
7.7.2 编写、调试应用程序	273
7.7.3 数据库试运行	274
7.8 数据库使用与维护	275
本章小结	276
习题	276
第8章 数据库应用系统设计实例	279
8.1 系统总体需求简介	279
8.1.1 用户总体业务结构	279
8.1.2 总体安全要求	280
8.2 系统总体设计	280
8.2.1 系统设计考虑的主要内容	280
8.2.2 系统的体系结构	280
8.2.3 系统软件开发平台	282
8.2.4 系统的总体功能模块	285
8.3 系统需求描述	285
8.3.1 系统全局数据流图	285
8.3.2 系统局部数据流图	286
8.3.3 系统数据字典	287
8.4 系统概念模型描述	289
8.4.1 构成系统的实体型	290
8.4.2 系统局部 ER 图	290
8.4.3 合成全局 ER 图	290
8.4.4 优化全局 ER 图	292
8.5 系统的逻辑设计	293
8.5.1 转化为关系数据模型	294
8.5.2 关系数据模型的优化与调整	294
8.5.3 数据库表的结构	295
8.6 数据库的物理设计	299
8.6.1 存储介质类型的选择	299
8.6.2 定义数据库	300

8.6.3 创建表及视图	301
8.6.4 创建索引	304
8.6.5 数据库服务器性能优化	304
本章小结	305
习题	306
第9章 主流数据库产品、工具及比较	307
9.1 主流数据库厂商、产品及工具	307
9.1.1 Oracle 公司的 Oracle	307
9.1.2 IBM 公司的 DB2 及 Informix	309
9.1.3 Sybase 公司的 ASE	311
9.1.4 CA 公司及 Ingres	314
9.2 信息系统对数据库产品的需求	316
9.3 主流数据库产品功能的粗略比较	317
本章小结	319
习题	319
第10章 数据仓库与数据挖掘及数据库新进展	321
10.1 数据仓库技术	321
10.1.1 数据仓库概述	321
10.1.2 数据仓库的多维数据模型	323
10.1.3 数据仓库的相关概念	325
10.1.4 数据仓库的系统结构	327
10.1.5 数据仓库系统开发与工作过程	328
10.2 数据挖掘技术	330
10.2.1 数据挖掘概述	330
10.2.2 数据挖掘的主要技术	332
10.2.3 数据挖掘与数据仓库	333
10.2.4 数据挖掘在各行业的应用	335
10.3 数据库技术的研究与发展	336
10.3.1 数据库技术研究的新特点	336
10.3.2 数据库技术的研究热点	338
10.3.3 国内数据库技术的发展状况	342
10.3.4 数据库技术的发展方向	343
本章小结	344
习题	344

第 11 章 数据库上机实验及指导	345
11.1 SQL 数据库语言操作实验	345
11.1.1 SQL 定义子语言实验	345
11.1.2 SQL 操纵子语言实验	347
11.1.3 T-SQL 编程实验	352
11.2 数据库完整性实验	353
11.2.1 表本身的完整性	353
11.2.2 表间参照完整性	355
11.3 SQL Server 安全设置实验	356
11.3.1 创建登录账号	356
11.3.2 数据库用户设置	357
11.3.3 SQL Server 角色管理	358
11.3.4 SQL Server 语句及对象授权实验	359
11.4 数据库系统管理实验	359
11.4.1 故障恢复实验	359
11.4.2 数据库服务器及性能设置实验	362
本章小结	363
习题	364
附录	365
附录 A 英汉术语对照表	365
附录 B MS SQL Server 2000 安装	374
附录 C SQL Server 样例库表结构	380
参考文献	385

第 1 章

CHAPTER

数据库系统概述

为使读者深入理解“数据库”(Database, DB)的内容,本章将以全新的角度,从最原始的观点出发,来介绍“数据库系统”(Database System, DBS)的核心——“数据库管理系统”(Database Management System, DBMS)出现的背景。以此为线索,逐渐自然地引出与此紧密相关的重要概念和内容,并最终得出由数据库的“原理”(Principle)、“应用”(Application)和“设计”(Design)3大部分组成的、数据库系统的总体结构图。该总体结构图的提出,将数据库系统中常常相互交叉、不容易理顺关系的原理、应用与设计这3部分内容,有机地组织起来,使读者既清楚它们之间的联系,又能明了它们之间的区别,从而为学好数据库这门课程开一个好头。

数据库应用系统的开发技术日新月异,本章将对其中的C/S和B/S的两层及多层系统结构模式、组件开发技术、分布式对象技术、CORBA/DCOM/J2EE、CGI/ISAPI/NSAPI、ASP/JSP/PHP等进行较全面的介绍,使读者能对目前的系统应用开发有一个完整的了解。

数据库领域有其自己显著的特点,涉及相当多且新的理论及概念。本章将通过与背景联系的方式,逐步引出这些概念,使读者知晓这些概念的由来,从而加深对概念的理解。数据库技术的发展,归根结底是由实际应用需求推动的,因此,本章将从需求推动的视角来阐述数据库技术的发展。

本章学习目的和要求:

- 数据库管理系统出现的背景;
- 数据库管理系统的基本功能;
- 数据库系统的总体结构;
- 理解数据库原理、应用与设计3部分间的关系;
- 数据库应用系统开发技术;
- 数据库系统中的术语与基本概念;
- 数据库技术的发展。