

计算机系列教材

数据库原理及应用

黄雪华 徐述 曹步文 黄静 编著

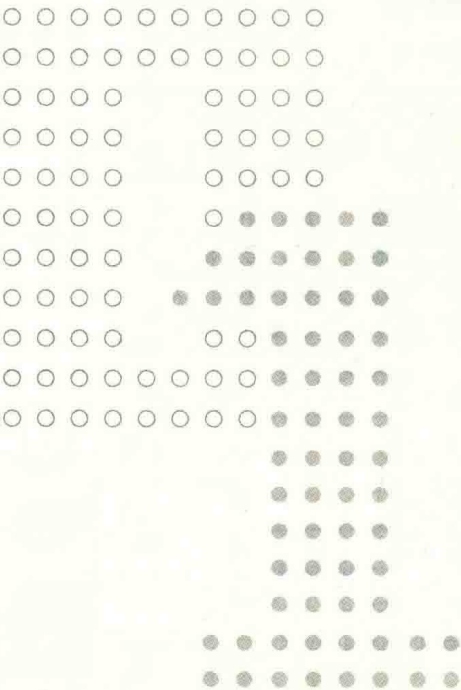


清华大学出版社

计算机系列教材

黄雪华 徐述 曹步文 黄静 编著

数据库原理及应用



清华大学出版社
北京

内 容 简 介

在大数据时代背景下,本书以培养数据管理应用型人才为目标,系统全面地讲述了数据库系统的基础理论知识、基本方法与应用技术。本书总共包含4篇17章,第1篇为基础篇,包括绪论、关系模型数据库、关系代数。第2篇为设计及应用篇,包括使用实体—联系模型进行数据建模、扩展的实体—联系模型、实体—联系模型到关系模型的转换、UML类图建模、关系数据理论、关系数据库标准语言SQL、数据库编程、数据库设计。第3篇为管理篇,包括并发控制、数据库存储技术、关系查询优化、数据库安全、数据库恢复。第4篇为新技术篇,包括数据库的发展及新技术。

本书可以作为高等院校以及高职高专数据库课程的教材,也可以作为从事数据库系统研究、开发、应用和工程技术人员的参考用书。

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

数据库原理及应用/黄雪华等编著. —北京:清华大学出版社,2018

(计算机系列教材)

ISBN 978-7-302-50073-5

I. ①数… II. ①黄… 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(2018)第097047号

责任编辑:白立军 常建丽

封面设计:常雪影

责任校对:李建庄

责任印制:丛怀宇

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 装 者:北京泽宇印刷有限公司

经 销:全国新华书店

开 本:185mm×260mm

印 张:21.5

字 数:498千字

版 次:2018年9月第1版

印 次:2018年9月第1次印刷

定 价:49.00元

产品编号:075932-01

前 言

本书参考了很多优秀的国内外经典教材,如斯坦福大学的 *A First Course in Database*、*Fundamentals of Database Systems*、*Database System Concepts* 以及王珊主编的《数据库系统概论》等。本书根据数据库课程在大学中的教学大纲要求进行编写,以 SQL Server 中文版为实验环境,引入了大量的实例,偏重基础理论及应用实践,是作者多年从事数据库教学经验的总结。

本书分为 4 篇 17 章(书中有 * 号的部分是可选的学习内容),如下表所示。

第 1 篇 基础篇	第 1 章 绪论
	第 2 章 关系模型数据库
	第 3 章 关系代数
第 2 篇 设计及应用篇	第 4 章 使用实体—联系模型进行数据建模
	第 5 章 扩展的实体—联系模型
	第 6 章 实体—联系模型到关系模型的转换
	第 7 章 UML 类图建模
	第 8 章 关系数据理论
	第 9 章 关系数据库标准语言 SQL
	第 10 章 数据库编程
	第 11 章 数据库设计
第 3 篇 管理篇	第 12 章 并发控制
	第 13 章 数据库存储技术
	第 14 章 关系查询优化
	第 15 章 数据库安全
	第 16 章 数据库恢复
第 4 篇 新技术篇	第 17 章 数据库的发展及新技术

本书第 3 篇(第 12~16 章)由徐述编写,其余篇章由黄雪华编写,其中部分插图由刘欢所画,同时,在曹步文、黄静老师的协助下,本书最终顺利完成。相关的配套资源可从清华大学出版社的网站上下载,也可向作者索取。

本书的出版得到清华大学出版社的大力支持和帮助,在此表示诚挚感谢。

由于时间比较紧张，作者水平有限，书中难免有不足之处，望广大读者给予批评指正。
作者的电子邮箱为 107531852@qq.com。

编 者
2018 年 5 月

目 录

第1篇 基础篇

第1章 绪论 /3

1.1 概述 /3

1.1.1 数据库的基本概念 /4

1.1.2 数据库管理系统介绍 /7

1.1.3 示例 /8

1.2 逻辑数据模型 /9

1.2.1 层次模型 /10

1.2.2 网状模型 /11

1.2.3 关系模型 /12

1.3 数据库系统的结构 /12

1.3.1 模式及实例的概念 /13

1.3.2 数据库系统的三级模式结构 /13

1.3.3 数据库系统的体系结构 /16

1.4 数据管理技术的发展历史 /17

1.4.1 人工管理阶段 /18

1.4.2 文件系统管理阶段 /18

1.4.3 数据库管理系统阶段 /20

* 1.5 DBMS 组成 /22

1.6 小结 /22

1.7 习题 /23

第2章 关系模型数据库 /24

2.1 关系模型数据库的数据结构 /24

2.1.1 关系模型的基本概念 /24

2.1.2 关系的性质 /27

2.1.3 关系模型的形式化定义 /29

2.2 关系模型的完整性 /35

2.2.1 实体完整性 /35

2.2.2 参照完整性 /35

2.2.3 用户自定义完整性 /37

- 2.3 本书示例数据库 /38
- 2.4 小结 /39
- 2.5 习题 /39

第3章 关系代数 /40

- 3.1 关系操作 /40
- 3.2 关系操作的语言 /40
- 3.3 关系代数运算 /41
 - 3.3.1 传统的集合运算 /41
 - 3.3.2 专门的关系运算 /45
 - 3.3.3 关系代数表达式应用举例 /52
- 3.4 小结 /53
- 3.5 习题 /53

第2篇 设计及应用篇

第4章 使用实体—联系模型进行数据建模 /57

- 4.1 数据模型 /57
- 4.2 概念模型 /57
- 4.3 实体—联系模型 /58
 - 4.3.1 基本概念 /58
 - 4.3.2 一个完整的示例 /62
 - 4.3.3 E-R图表示法小结 /64
 - 4.3.4 联系的不同表示法 /64
- 4.4 E-R图应用举例 /66
- 4.5 小结 /67
- 4.6 习题 /67

第5章 扩展的实体—联系模型 /68

- 5.1 扩展的实体—联系模型介绍 /68
 - 5.1.1 扩展的E-R模型的基本概念 /68
 - 5.1.2 一个完整的示例 /72
- 5.2 E-R及EER模型的设计步骤 /74
- 5.3 E-R及EER模型的设计原则 /74
- 5.4 EER图应用举例 /77
- 5.5 小结 /77
- 5.6 习题 /77

第 6 章 实体—联系模型到关系模型的转换 /78

- 6.1 E-R 模型到关系模型的转换 /78
 - 6.1.1 实体的映射 /78
 - 6.1.2 二元联系的映射 /79
 - 6.1.3 其他元素的映射 /82
- 6.2 一个完整的 E-R 模型转换示例 /83
- 6.3 EER 模型到关系模型的转换 /85
 - 6.3.1 父类与子类的转换 /85
 - 6.3.2 聚集的转换 /87
- 6.4 一个完整的 EER 模型转换示例 /88
- 6.5 小结 /90
- 6.6 习题 /90

*** 第 7 章 UML 类图建模 /91**

- 7.1 概述 /91
- 7.2 UML 类图表示法 /91
- 7.3 示例 /95
- 7.4 UML 类图到关系模型的转换 /96
- 7.5 数据库设计工具 /97
- 7.6 小结 /99
- 7.7 习题 /99

第 8 章 关系数据理论 /100

- 8.1 规范化理论概述 /100
- 8.2 基本概念 /104
- 8.3 范式 /106
 - 8.3.1 第一范式 /106
 - 8.3.2 第二范式 /107
 - 8.3.3 第三范式 /109
 - 8.3.4 BCNF /110
 - 8.3.5 多值依赖与第四范式 /112
 - * 8.3.6 连接依赖与 5NF /113
 - 8.3.7 规范化小结 /114
- 8.4 Armstrong 公理系统 /114
- 8.5 关系模式分解 /118
- * 8.6 模式分解算法 /121
- 8.7 规范化应用 /122
- 8.8 小结 /124

8.9 习题 /124

第9章 关系数据库标准语言 SQL /126

9.1 SQL 概述 /126

9.2 SQL 定义 /128

9.2.1 数据定义和数据类型 /128

9.2.2 定义约束 /131

9.2.3 模式修改语句 /135

9.2.4 应用举例 /136

9.3 查询 /137

9.3.1 单表查询 /138

9.3.2 多表查询 /143

9.3.3 嵌套查询 /149

9.3.4 集合查询 /153

9.3.5 基于派生表的查询 /155

9.3.6 应用举例 /155

9.4 数据更新 /160

9.4.1 插入数据 /160

9.4.2 修改数据 /161

9.4.3 删除数据 /162

9.4.4 应用举例 /163

9.5 视图 /164

9.5.1 定义视图 /164

9.5.2 查询视图 /166

9.5.3 更新视图 /167

9.5.4 视图的优点 /168

9.5.5 应用举例 /169

9.6 索引 /170

9.7 其他的相关理论 /171

9.8 小结 /172

9.9 习题 /172

第10章 数据库编程 /173

10.1 编程介绍 /173

10.2 嵌入式 SQL /173

10.3 数据库编程语言 /175

10.3.1 基本语法 /175

10.3.2 存储过程与函数 /179

- 10.3.3 触发器 /183
- 10.3.4 游标 /185
- 10.4 数据库接口及访问技术 /186
 - 10.4.1 ADO.NET 编程 /187
 - 10.4.2 JDBC 编程 /189
- 10.5 小结 /191
- 10.6 习题 /191

第 11 章 数据库设计 /192

- 11.1 数据库设计概述 /192
 - 11.1.1 数据库设计方法 /192
 - 11.1.2 数据库设计步骤 /193
- 11.2 需求分析 /195
 - 11.2.1 需求分析的方法 /196
 - 11.2.2 数据流图 /197
 - 11.2.3 数据字典 /200
- 11.3 概念结构设计 /201
 - 11.3.1 概念模型的特点 /201
 - 11.3.2 概念结构设计方法 /202
 - 11.3.3 局部概念模型设计 /203
 - 11.3.4 全局概念模型设计 /205
- 11.4 逻辑结构设计 /209
 - 11.4.1 E-R 模型到关系模型的转换 /209
 - 11.4.2 关系模型的优化 /209
 - 11.4.3 设计用户子模式 /210
- 11.5 物理结构设计 /210
 - 11.5.1 存取方法 /211
 - 11.5.2 存储结构 /212
 - 11.5.3 评价物理结构 /213
- 11.6 数据库的实施 /213
- 11.7 数据库的运行和维护 /215
- 11.8 数据库设计案例——学生成绩管理系统 /216
 - 11.8.1 需求分析 /216
 - 11.8.2 概念结构设计 /217
 - 11.8.3 逻辑结构设计 /217
 - 11.8.4 物理结构设计 /221
 - 11.8.5 相关数据库代码 /221
 - 11.8.6 部分模块界面图 /228

- 11.9 小结 /242
- 11.10 习题 /242

第3篇 管 理 篇

第12章 并发控制 /245

- 12.1 事务 /245
 - 12.1.1 事务的概念 /245
 - 12.1.2 事务的 ACID 性质 /245
- 12.2 并发控制 /246
 - 12.2.1 事务并发执行的必要性 /246
 - 12.2.2 并发操作带来的问题 /247
 - 12.2.3 并发事务调度可串行化 /248
- 12.3 封锁技术 /249
 - 12.3.1 封锁类型 /249
 - 12.3.2 封锁协议 /250
 - 12.3.3 两段锁协议 /251
- 12.4 封锁带来的问题 /252
 - 12.4.1 活锁 /252
 - 12.4.2 死锁 /252
- 12.5 多粒度封锁 /254
 - 12.5.1 多粒度树 /254
 - 12.5.2 意向锁 /255
- 12.6 小结 /256
- 12.7 习题 /257

第13章 数据库存储技术 /258

- 13.1 数据库系统存储结构 /258
 - 13.1.1 数据库磁盘存储器中的数据结构 /258
 - 13.1.2 数据库系统存储介质 /259
- 13.2 数据文件的记录格式 /260
 - 13.2.1 定长记录 /260
 - 13.2.2 变长记录 /260
- 13.3 数据文件格式 /262
 - 13.3.1 文件格式 /262
 - 13.3.2 顺序文件 /262
 - 13.3.3 聚集文件 /263
- 13.4 索引技术 /263

13.4.1	索引的概念	/263
13.4.2	主索引	/263
13.4.3	辅助索引	/265
13.4.4	索引的更新	/265
13.5	B+树索引文件	/266
13.5.1	B+树的结构	/266
13.5.2	B+树的查询	/267
13.5.3	B+树的更新	/268
13.6	散列索引文件	/269
13.6.1	散列技术	/269
13.6.2	静态散列索引	/270
13.6.3	可扩充散列结构	/271
13.7	小结	/273
13.8	习题	/274
第 14 章	关系查询优化	/275
14.1	查询处理	/275
14.1.1	概述	/275
14.1.2	查询代价度量	/276
14.2	查询优化	/276
14.2.1	查询优化概述	/277
14.2.2	代数优化	/277
14.2.3	物理优化	/281
14.3	小结	/284
14.4	习题	/284
第 15 章	数据库安全	/285
15.1	数据库安全概述	/285
15.1.1	TCSEC 标准	/285
15.1.2	CC 标准	/286
15.2	数据库系统安全控制	/287
15.2.1	数据库系统安全模型	/287
15.2.2	用户身份标识与鉴别	/288
15.2.3	存取控制概述	/289
15.3	自主存取控制	/290
15.3.1	授权	/291
15.3.2	角色	/293
15.3.3	视图机制	/294

- 15.4 审计 /295
 - 15.4.1 审计事件 /295
 - 15.4.2 审计的作用 /296
- 15.5 强制存取控制 /296
- 15.6 数据加密 /297
 - 15.6.1 加密技术 /297
 - 15.6.2 数据库中的加密支持 /297
- 15.7 更高安全性保护 /298
- 15.8 小结 /298
- 15.9 习题 /298

第 16 章 数据库恢复 /299

- 16.1 故障类型 /299
 - 16.1.1 事务故障 /299
 - 16.1.2 系统故障 /299
 - 16.1.3 介质故障 /299
- 16.2 恢复的基本原理与实现方法 /300
- 16.3 恢复技术 /300
 - 16.3.1 数据转储 /300
 - 16.3.2 日志文件格式 /301
 - 16.3.3 日志登记原则 /302
 - 16.3.4 使用日志重做和撤销事务 /303
 - 16.3.5 检查点 /305
- 16.4 恢复算法 /306
 - 16.4.1 事务回滚 /306
 - 16.4.2 系统崩溃后的恢复 /306
 - 16.4.3 介质故障后的恢复 /307
- 16.5 小结 /308
- 16.6 习题 /308

第 4 篇 新技术篇

第 17 章 数据库的发展及新技术 /311

- 17.1 数据库系统发展的特点 /311
- 17.2 数据管理技术发展的趋势 /313
- 17.3 面向对象数据库管理系统 /314
 - 17.3.1 面向对象数据库管理系统介绍 /314
 - 17.3.2 对象关系数据库管理系统介绍 /316

- 17.4 分布式数据库 /316
- 17.5 并行数据库 /319
- 17.6 空间数据库 /320
- 17.7 数据仓库与数据挖掘 /322
- 17.8 大数据 /326
- 17.9 小结 /329
- 17.10 习题 /329

参考文献 /330

第1篇

基础篇

本篇介绍数据库系统的基本概念、基础理论知识和数据管理发展历史,并从逻辑数据模型的三个方面详细地阐述关系数据库。本篇是后面章节的基础,也是本课程的入门。

本篇包含以下章节。

第1章 绪论,介绍数据库应用举例,重点阐述数据库系统的基本概念,详细说明了逻辑数据模型的构成及分类,讲述数据库系统的三级模式结构及体系结构,介绍数据管理技术发展的三个阶段以及 DBMS 组成。其中,DBMS 组成为选修内容。

第2章 关系模型数据库,详细介绍了关系模型数据库的数据结构,以及关系模型数据库的三大完整性约束。

第3章 关系代数,介绍关系模型操作语言,重点讲述关系代数运算。

第1章 绪 论

数据库是一门非常重要的课程,其理论知识及技术应用相当广泛。是信息类学科必须掌握的内容,而且近几年它的理论知识及技术发展、更新达到了一个新的高潮。

“数据库”不仅是计算机类专业、信息管理信息系统、计算机网络、大数据等相关专业的专业基础必修课程,也是很多非计算机专业的选修课程。

本章介绍数据库应用举例,重点阐述数据库系统的基本概念,详细说明了逻辑数据模型的构成及分类,讲述数据库系统的三级模式结构及体系结构,介绍数据管理技术发展的三个阶段以及 DBMS 组成。通过本章的学习,可以了解数据库、数据库系统的概貌。

1.1 概述

随着计算机技术、通信技术以及互联网的蓬勃发展,人们对数据的需求日益递增,数据量呈现爆炸式增长。当今是数据的时代,涌现了各种类型的应用型数据,如物联网数据、无线网络数据、GPS 数据、3D 数据、多媒体数据、社交网络数据、自媒体数据以及大规模的历史数据等。这些数据的存储、表示、处理、传输要用到数据管理技术。在人们的生活中,有关数据的应用无处不在,如银行、航空公司、高校、销售、在线销售、医院、人力资源等。

银行需要存储客户的基本信息、账户信息、存款信息、取款信息、转账信息、借款信息等。

航空公司需要存储旅客信息、票务信息、航班信息、订票信息等。

高校需要存储学生信息、教师信息、课程信息、排课信息、选课、成绩,甚至学校的贴吧及论坛,这些信息都保存在学校的数据库中。

销售方面,如沃尔玛等大型国际连锁超市的货物信息、库存信息、销售信息等都保存在数据库中。

在线销售,如大型电子商务网站“淘宝”的商品信息、客户信息、销售记录等数据都保存在数据库中。

医院需要保存病人的基本信息、诊断信息、病例、医生的相关信息等,这些信息都保存在医院的数据库里,以便于随时获取。

人力资源需要保存每个员工的基本信息、休假、奖惩记录、薪资等信息。

可以说,当今时代数据无处不在,人们随时随地都需要查看数据、保存数据、管理数据、搜索数据、分析数据。这些数据一般都保存在专门的数据库里。

商家经常根据顾客的购买历史记录及浏览过的产品为顾客推荐商品,有时候在超市里会看见尿片与啤酒摆放在一起,银行根据客户的信用记录决定是否为客户发放贷款等,这些都是从数据中挖掘出规律然后应用于生活中的例子。