

国外计算机科学教材系列

# 数据库系统

——设计、实现与管理（第三版）

## Database Systems

A Practical Approach to Design, Implementation, and Management

Third Edition

THOMAS CONNOLLY • CAROLYN BEGG

### DATABASE SYSTEMS

A Practical Approach to  
Design, Implementation,  
and Management



Third Edition



[英] Thomas Connolly 著  
Carolyn Begg

宁洪等译



电子工业出版社

Publishing House of Electronics Industry

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

数据库系统  
作者著作

国外计算机科学教材系列

# 数据库系统

## ——设计、实现与管理

(第三版)

Database Systems

A Practical Approach to Design, Implementation,  
and Management, Third Edition

[美] Thomas Connolly 著  
Carolyn Begg

宁 洪 等译

电子工业出版社  
Publishing House of Electronics Industry  
北京·BEIJING

## 内 容 简 介

本书是数据库领域的经典著作,内容系统、全面、实用,被世界多所大学选为数据库课程的教材。全书共分八个部分,其中的主要内容有:数据库系统和数据库设计的基础知识;关系模型和关系语言;数据库分析和设计的主要技术;数据库设计方法学;数据库安全、事务管理、询问处理与优化和嵌入式SQL;分布式DBMS和基于对象的DBMS,以及一些日益重要的技术,包括DBMS与Web技术的连接、半结构化技术与XML的关系、数据仓库、联机分析处理和数据挖掘等。

本书可作为计算机科学与技术及相关专业高年级本科生及研究生的教材,同时也可作为数据库研究和开发人员的参考书。

© Pearson Education Limited 1995, 2002.

This edition of Database Systems: A Practical Approach to Design, Implementation, and Management, Third Edition, ISBN: 0201708574 by Thomas Connolly, Carolyn Begg is published by arrangement with Pearson Education Limited. All Rights Reserved.

Simplified Chinese edition published by Publishing House of Electronics Industry. Copyright © 2004.

Licensed for sale in mainland territory of the People's Republic of China only, excluding Hong Kong.

本书中文简体字翻译版由 Pearson Education Limited 授予电子工业出版社。未经出版者预先书面许可,不得以任何形式或手段复制或抄袭本书内容。

此版本仅限在中华人民共和国境内(不包括香港、澳门特别行政区以及台湾地区)发行与销售。

版权贸易合同登记号 图字:01-2002-4138

### 图书在版编目(CIP)数据

数据库系统——设计、实现与管理(第三版)/(美)康诺利(Connolly, T.)等著;宁洪等译.

-北京:电子工业出版社,2004.1

(国外计算机科学教材系列)

书名原文: Database Systems: A Practical Approach to Design, Implementation, and Management, Third Edition  
ISBN 7-5053-9418-5

I. 数… II. ①康… ②宁… III. 数据库系统-教材 IV. TP311.13

中国版本图书馆CIP数据核字(2003)第110466号

责任编辑:杜 萌

印 刷:北京兴华印刷厂

出版发行:电子工业出版社

北京市海淀区万寿路173信箱 邮编:100036

经 销:各地新华书店

开 本:787×1092 1/16 印张:58 字数:1485千字

印 次:2004年1月第1次印刷

定 价:79.00元

凡购买电子工业出版社的图书,如有缺损问题,请向购买书店调换;若书店售缺,请与本社发行部联系。联系电话:(010)68279077。质量投诉请发邮件至zlt@phei.com.cn,盗版侵权举报请发邮件至dbqq@phei.com.cn。

## 出版说明

21世纪初的5至10年是我国国民经济和社会发展的关键时期,也是信息产业快速发展的关键时期。在我国加入WTO后的今天,培养一支适应国际化竞争的一流IT人才队伍是我国高等教育的重要任务之一。信息科学和技术方面人才的优劣与多寡,是我国面对国际竞争时成败的关键因素。

当前,正值我国高等教育特别是信息科学领域的教育调整、变革的重大时期,为使我国教育体制与国际化接轨,有条件的高等院校正在为某些信息学科和技术课程使用国外优秀教材和优秀原版教材,以使我国在计算机教学上尽快赶上国际先进水平。

电子工业出版社秉承多年来引进国外优秀图书的经验,翻译出版了“国外计算机科学教材系列”丛书,这套教材覆盖学科范围广、领域宽、层次多,既有本科专业课程教材,也有研究生课程教材,以适应不同院系、不同专业、不同层次的师生对教材的需求,广大师生可自由选择和自由组合使用。这些教材涉及的学科方向包括网络与通信、操作系统、计算机组织与结构、算法与数据结构、数据库与信息处理、编程语言、图形图像与多媒体、软件工程等。同时,我们也适当引进了一些优秀英文原版教材,本着翻译版本和英文原版并重的原则,对重点图书既提供英文原版又提供相应的翻译版本。

在图书选题上,我们大都选择国外著名出版公司出版的高校教材,如Pearson Education培生教育出版集团、麦格劳-希尔教育出版集团、麻省理工学院出版社、剑桥大学出版社等。撰写教材的许多作者都是蜚声世界的教授、学者,如道格拉斯·科默(Douglas E. Comer)、威廉·斯托林斯(William Stallings)、哈维·戴特尔(Harvey M. Deitel)、尤利斯·布莱克(Uyless Black)等。

为确保教材的选题质量和翻译质量,我们约请了清华大学、北京大学、北京航空航天大学、复旦大学、上海交通大学、南京大学、浙江大学、哈尔滨工业大学、华中科技大学、西安交通大学、国防科学技术大学、解放军理工大学等著名高校的教授和骨干教师参与了本系列教材的选题、翻译和审校工作。他们中既有讲授同类教材的骨干教师、博士,也有积累了几十年教学经验的老教授和博士生导师。

在该系列教材的选题、翻译和编辑加工过程中,为提高教材质量,我们做了大量细致的工作,包括对所选教材进行全面论证;选择编辑时力求达到专业对口;对排版、印制质量进行严格把关。对于英文教材中出现的错误,我们通过与作者联络和网上下载勘误表等方式,逐一进行了修订。

此外,我们还将与国外著名出版公司合作,提供一些教材的教学支持资料,希望能为授课老师提供帮助。今后,我们将继续加强与各高校教师的密切联系,为广大师生引进更多的国外优秀教材和参考书,为我国计算机科学教学体系与国际教学体系的接轨做出努力。

电子工业出版社

## 教材出版委员会

- 主任** 杨芙清 北京大学教授  
中国科学院院士  
北京大学信息与工程学部主任  
北京大学软件工程研究所所长
- 委员** 王 珊 中国人民大学信息学院院长、教授
- 胡道元 清华大学计算机科学与技术系教授  
国际信息处理联合会通信系统中国代表
- 钟玉琢 清华大学计算机科学与技术系教授  
中国计算机学会多媒体专业委员会主任
- 谢希仁 中国人民解放军理工大学教授  
全军网络技术研究中心主任、博士生导师
- 尤晋元 上海交通大学计算机科学与工程系教授  
上海分布计算技术中心主任
- 施伯乐 上海国际数据库研究中心主任、复旦大学教授  
中国计算机学会常务理事、上海市计算机学会理事长
- 邹 鹏 国防科学技术大学计算机学院教授、博士生导师  
教育部计算机基础课程教学指导委员会副主任委员
- 张昆藏 青岛大学信息工程学院教授

# 译者序

数据库作为数据管理的基本技术和工具已在科研部门、政府机关、企事业单位广泛使用。然而，正如本书作者在前言所述，“正是由于其简单性，许多用户可能缺乏如何开发正确且高效系统的知识，而创建了数据库及其应用程序”。本书的目标就是尽可能清楚地介绍数据库的基础理论，并给出一个既能为专业人员也能为非专业人员所用的数据库设计方法学。

该书的两位作者均具有丰富的数据库管理系统和数据库应用系统的设计经验，Thomas Connolly 曾参与设计数据库管理系统 RAPPOR 和配置管理工具 LIFESPAN，Carolyn Begg 则是数据库系统在生物领域应用方面的专家。

全书共分为八个部分。第一部分概述数据库系统和数据库设计的各个方面；第二部分讨论关系模型和关系语言；第三部分讨论数据库分析和设计的主要技术及实用方法；第四部分通过一个案例给出数据库设计方法学；第五部分讨论数据库安全、事务管理、询问处理与优化及嵌入式 SQL；第六部分讨论分布式 DBMS 和基于对象的 DBMS；第七部分介绍若干与数据库相关的新技术；第八部分为附录。

该书系统性强、思路清晰，特别注重理论与实践的结合，内容既包括了数据库领域的经典内容，又反映了近期的研究成果，此外采用了清晰且易于理解的表述方法，便于阅读和理解，便于“教”和“学”，是不可多得的一本好书。该书可用做本科生数据库管理或数据库设计的导论性教材，也可作为研究生或高年级本科生这方面课程的教材。本书还可以作为一些 IT 专业人士的参考书，如系统分析和设计人员、应用程序开发人员、系统程序员、数据库从业人员以及独立的自学者。

宁洪全面组织了本书的翻译，参加初译的人员包括闫伟、杨艳萍、李姗姗、孟兆伟、张国敏、高树静、周麓、来森。宁洪、孟兆伟和李姗姗修改了初译稿并复译了部分章节，陈波、闫琪和汪昌健也参加了部分修改工作。最后由宁洪统稿并审校了全书。

限于译者水平，译文中疏漏和错误难免，欢迎批评指正。

# 前 言

## 背景

在过去的30年中,数据库的研究带来了巨大的生产力,使得数据库系统成为软件工程领域最重要的成果。目前,数据库作为信息系统的基本框架,已从根本上改变了许多公司的运作方式。特别是在最近几年里,这项技术本身的发展产生了一些功能更强大、使用更方便的系统。这使得数据库系统的普及度越来越高。遗憾的是,正是由于其简单性,许多用户有可能在缺乏如何开发正确且高效系统的知识的情况下而创建了数据库及其应用程序。这就可能导致所谓的“软件危机”,有时也称为“软件萧条”的继续。

编写本书的最初原因是作者在工业界的工作经历,即为新软件系统中数据库设计提供咨询,间或也解决老系统中存在的种种问题。进入学术界后,作者从学生那里也发现了类似的问题。因此,本书的目标就是给出一本教程,它尽可能清楚地介绍数据库的基础理论,并给出一个既能为专业人员也能为非专业人员所用的数据库设计方法学。

本书针对当前商业应用的主流系统——关系数据库管理系统(DBMS)给出的设计方法学已在学术界和工业界测试和使用了许多年。它包括三个主要的阶段:数据库的概念设计、逻辑设计和物理设计。第一个阶段在不考虑任何物理因素的前提下设计概念数据模型,这个模型在第二阶段被细化为逻辑数据模型,细化过程主要是去除在关系系统中无法表示的结构。在第三阶段,逻辑数据模型被转换成针对目标DBMS的物理设计,在物理设计阶段主要考虑为有效和安全地访问存储在辅助存储器中的数据库所需要的存储结构和访问方法。

方法学按阶段被分为一系列步骤。对于缺少经验的设计者,最好遵从步骤的先后顺序进行设计,可用其提供的指南帮助完成整个过程。对于有经验的设计者,该方法学的指导作用显得较弱一些,经常作为开发框架和检查列表使用。为了帮助读者学习使用上述方法学并理解其要点,整个方法学的描述始终采用一个完整的DreamHome案例。此外,附录B还给出了另外三个案例,供读者自行尝试。

## UML (统一建模语言)

越来越多的公司都在规范各自数据建模的方法,即选择一种特定的建模方法并在整个数据库开发项目中始终如一地使用它。有一种流行的、用于数据库概念/逻辑设计的高级数据模型,也就是在本书所采用的模型,是基于ER(实体联系)模型的概念。由于当前还没有表示ER模型的标准表示方法,相关的大部分书籍在描述关系数据库管理系统的数据库设计时,常常使用下述两种表示方法之一:

- Chen表示方法,即用矩形表示实体,用菱形表示联系,用线段连接矩形和菱形。
- Crow Feet表示方法,仍用矩形表示实体,用实体间的连线表示联系,在一对多联系连线的一端标有一个鸭足标记。

当前,这两种表示方法都有CASE工具。然而,它们都难于使用和解释。本书的前一版本我们曾使用Chen表示方法,而在随后Pearson Education进行的一次问卷调查中,比较一致的意见是应



该使用最新的、称为UML的面向对象建模语言。UML表示方法结合了面向对象设计三大流派的成分：Rumbaugh的OMT建模语言，Booch的面向对象分析和设计，Jacobson的Objectory。

换用表示方法主要有以下三个原因：（1）UML正成为一个工业标准；例如，对象管理组（OMG）已经采纳UML作为对象方法的标准表示方法；（2）UML表达清楚并且易用；（3）UML目前已被学术界用于面向对象分析与设计的教学，在数据库模块的教学中并使用UML更协调。因此，在这个版本中，我们将采用UML的类图表示方法。读者会发现这种表示方法更加容易理解和使用。在实际使用UML之前，我们曾花费了大量时间试验，以检测UML是否适合于数据库设计。我们在Pearson Education出版的“Database Solutions: A Step-by-Step Guide to Building Databases”一书中总结了这项工作。该书将前面提到的方法学用于两个案例的数据库设计和创建，其中一个使用Microsoft Access作为目标DBMS，另一个使用Oracle作为目标DBMS。书中还包含了许多其他示例，各个示例都有简单的解决方案，随书附带的光盘上还赠送了可视化建模工具Rational Rose的有效副本。

### 第三版本的更新之处

我们对本书的第二版做了相应的修改旨在增强可读性、修改和扩充原内容、增加新材料。第三版中主要变动包括：

- 增加了新的一章介绍关系代数和关系演算。
- 增加了新的一章介绍两种最为常用的商业关系DBMS，即Microsoft Access和Oracle。
- 增加了新的一章介绍实况发现技术，以及这些技术如何应用于数据库应用程序开发周期，尤其是早期阶段，如数据库规划、系统定义和需求收集与分析等。
- 使用UML的类图表示法描述实体联系模型以及增强实体联系模型（EER）。
- 改进了数据库设计的方法学，并用四章的篇幅重新进行了描述，通过DreamHome案例充分说明了如何实际使用该方法学。
- 将第二版中两章SQL教程重新组织成三章：其中，一章介绍数据操纵，一章介绍数据定义，另一章包括嵌入式SQL和开放数据库互连（ODBC）。
- 在介绍安全的章节中增加了对Microsoft Access和Oracle安全机制的描述，以及Web环境中的安全问题。
- 介绍事务管理的章节中增加了Oracle的事务管理机制。
- 介绍查询处理的章节中增加了Oracle的查询处理机制。
- 介绍分布式DBMS和复制服务器的章节中增加了Oracle的数据分布和复制机制。
- 介绍面向对象DBMS（OODBMS）的章节被分为两章。其中，一章讨论与OODBMS相关的通用问题和持久编程问题，另一章介绍新的1999对象数据管理组（ODMG）标准并概要介绍了一个商业OODBMS，ObjectStore。
- 介绍对象关系DBMS的章节做了修改，包括了SQL3的正式版本。
- 介绍Web技术和DBMS的章节做了修改。附录中提供了多个将数据库集成到Web环境的编程实例，这些例子涉及JavaScript，PHP和PostgreSQL，CGI和Perl，用到JDBC和SQLJ的Java数据库互连，JSP，ASP，以及Oracle的PL/SQL Server Page。
- 增加了一章专门介绍半结构化数据和XML，以及与之相关的技术和对XML查询语言的一些建议。
- 介绍数据仓库的一章被分为两章。一章介绍数据仓库，另一章介绍使用维度建模技术进行数据仓库的数据库设计。



## 读者对象

本书可作为本科生数据库管理或数据库设计的导论性教材,也可作为研究生或高年级本科生这方面课程的教材,学时可为一到两个学期。一般而言,信息系统、商务IT或计算机科学等专业的教程都需要开设这类课程。

本书还可以作为一些IT专业人士的参考书,如系统分析和设计人员、应用程序开发人员、系统程序员、数据库从业人员及独立的自学者。随着当今数据库系统的广泛使用,这些专业人士可能是来自于需要数据库的任何类型的公司。

读者在学习关于物理数据库设计的第16章和关于查询处理的第20章之前,如果对附录C中介绍的文件组织和数据结构相关概念有清楚的了解,将会有所帮助。理想的情况是这些背景知识已从前导课程中获得。如果不具备这个条件,可以在开始数据库课程后,学完第1章立即学习附录C。

如果对高级编程语言如C有较好的理解,在学习关于嵌入式SQL和动态SQL的第21章和关于ObjectStore的26.3节时会更有效。

## 突出特点

1. 为数据库逻辑设计和概念设计提供了易用、逐步指导的方法学,该方法学基于广泛采用的实体联系模型并可用规范化作为验证技术。此外,通过一个完整的案例来说明如何使用这个方法学。
2. 为数据库物理设计提供了易用、逐步指导的方法学,包括从逻辑设计到物理实现的映射,文件组织方法的选择,对应用有效的各种索引结构以及何时引入可控性冗余。此外,通过一个完整的案例来说明如何使用这个方法学。
3. 安排了单独的章节说明数据库设计阶段如何适配到整个系统开发周期,以及如何使用实况发现技术来获取系统需求。
4. 每章都采用清晰且易于理解的表述方法,如突出显示定义、明确给出各章学习要点,各章最后带有小结。通篇使用了大量示例和图表来说明概念。一个来自实际的DreamHome案例贯穿全书,另给出若干案例供学生选做课程实践题目。
5. 扩充了下列最新的正式标准及事实标准:SQL(结构化查询语言),QBE(仿效实例查询),面向对象数据库的ODMG(对象数据管理组)标准。
6. 用三章的篇幅介绍了SQL标准,包含交互式 and 嵌入式SQL。
7. 专设一章概述两种最常用的商业DBMS:Microsoft Access和Oracle。后续的许多章节都相应地介绍这两种DBMS如何支持讨论中的各种机制。
8. 全面地讨论了与分布式DBMS和复制服务器相关的概念和问题。
9. 全面地介绍了基于对象的DBMS中日益显示其重要性的一些概念和问题。具体包括回顾ODMG标准,给出对象管理机制的指南,该机制已经被加入到SQL的最新版本SQL3中。
10. 扩展了作为数据库应用平台的Web部分的内容,并给出多个Web数据库访问的代码示例。
11. 介绍了半结构化数据、它与XML的关系,以及关于XML查询语言的建议。
12. 全面地介绍了数据仓库、联机分析处理(OLAP)和数据挖掘。
13. 全面地介绍了用于数据仓库数据库设计的维度建模技术,并且通过一个完整的案例来演示如何使用该方法进行数据仓库数据库设计。
14. 介绍了DBMS系统实现的有关概念,包括并发技术和恢复控制、安全、查询处理和查询优化。

## 教学方法

在开始撰写本书之前,作者的目标之一就是写一本让读者容易接受和理解的教材,而不管读者具备怎样的背景知识和经验。根据作者使用教材的经验及从很多同事、客户和学生处吸收的意见,实际存在若干读者喜爱和不喜爱的设计特性。考虑到这些因素,本书决定采用如下的风格和结构:

- 在每章的开头明确说明该章的学习要点。
- 清楚定义每一个重要的概念,并将其用楷体字突出说明。
- 通篇大量使用图表来支持和阐明概念。
- 面向实际应用:为了做到这点,每章都包含了许多实际有效的示例用以说明所描述的概念。
- 每一章后附有小结,涉及该章所有主要的概念。
- 给出课后思考题,问题的答案都可以在书中找到。
- 给出课后习题,使学生能自测对章节的理解,全部习题的答案可以在配套的教辅材料(参见书后的“教学支持说明”)中找到。

## 教辅材料

在 Pearson Education 的要求下,我们还准备了适用于该书的大量教学辅导材料。配套的教师指南包括:

- 课程结构:包括在各类课程中应覆盖内容的建议。
- 教学意见:包括讲授内容建议、教学过程的提示,以及安排学生对所学内容进行课程设计的思路。
- 答案:所有课后思考题和习题的答案。
- 测试题:测试题(与每章后面的课后思考题和习题相似)及答案。
- 幻灯片:一组电子版幻灯片,包括每章的要点、放大的原文中所插图表,旨在帮助教师授课和组织讨论。

有关教辅材料及其有关信息可在 Pearson Education 网站上查找:

<http://www.booksites.net/connolly>

## 本书的组织结构

### 第一部分 背景知识

本书的第一部分介绍数据库系统和数据库设计。

**第1章** 介绍数据库管理的概念。主要阐述了早期各种系统,如基于文件的系统之不足及数据库方法所具备的优势。

**第2章** 总览数据库环境。主要讨论了三层 ANSI-SPARC 体系结构的优点,介绍了目前最通用的数据模型,列出了多用户 DBMS 应提供的各种功能。该章还分析了 DBMS 的基本软件体系结构,这一部分内容在数据库管理初级课程中可以略去。

### 第二部分 关系模型和语言

本书的第二部分介绍关系模型和关系语言,即关系演算和关系代数、QBE 和 SQL。还介绍了两种非常流行的商业系统:Microsoft Access 和 Oracle。

**第3章** 介绍当前最流行的数据模型，也是最经常被选做商业应用标准的模型，关系模型背后的概念。具体安排是首先介绍术语并说明其与数学上的关系的联系，然后讨论关系完整性规则、实体完整性和引用完整性。这一章以概述视图结束，第6章还将对视图展开讨论。

**第4章** 介绍关系演算和关系代数，并用示例加以说明。这部分内容在数据库管理初级课程中可以略去。然而，在第20章学习查询处理和第22章学习分布式DBMS的分段时需要关系代数的知识。此外，虽然不是特别必需，但对过程式的代数与非过程式的演算有个比较性的了解有利于学习第5章和第6章介绍的SQL语言。

**第5章** 介绍SQL的数据操作语句：SELECT，INSERT，UPDATE和DELETE。本章通过一系列有效的示例，以指南的方式说明了这些语句的主要概念。

**第6章** 讨论SQL标准中主要的数据定义机制。这章仍以指南的方式，介绍了SQL的数据类型、数据定义语句、完整性增强特性（IEF）和数据定义语句中更高级的特性，包括访问控制语句GRANT和REVOKE。此外，还介绍了视图及用SQL如何创建视图。

**第7章** 这又是一个实践性章节。主要讨论交互式查询语言QBE，这种语言被誉为非技术型的计算机用户访问数据库信息最容易的方法。书中通过Microsoft Access来展示QBE。

**第8章** 通过介绍两种通用的商业关系DBMS来结束本书的第二部分，即Microsoft Access和Oracle。在本书的后续章节中，我们查看这两种DBMS如何实现各种数据库机制，如安全和查询处理。

### 第三部分 数据库分析与设计技术

本书的第三部分讨论数据库分析和设计的主要技术，以及这些技术如何实际运用。

**第9章** 总览数据库应用生命周期的各个主要阶段。特别强调了数据库设计的重要性并说明这个过程如何被分为三个阶段：概念、逻辑和物理数据库设计。此外，还描述了应用程序的设计（功能方面）对数据库设计（数据方面）的影响。数据库应用生命周期的关键阶段是选择合适的DBMS。这一章讨论了对DBMS的选择过程，提供了一系列方针和建议。最后讨论数据管理和数据库管理的重要性。

**第10章** 讨论数据库开发者何时使用实况发现技术及捕获何种类型的实况。这一章描述了最常使用的实况发现技术及其优缺点。通过DreamHome案例说明在数据库应用生命周期的早期阶段如何应用这些技术。

**第11章和第12章** 介绍了实体联系模型和扩展的实体联系模型（EER），在扩展的实体联系模型中，允许使用更多、更高级的数据建模技术，如子类、超类和分类。EER是一种流行的高级概念数据模型，也是这里讨论的数据库设计方法学的一种基本技术。读者还学习使用UML来表示实体联系图。

**第13章** 阐述了规范化背后的概念，它是逻辑数据库设计方法学中另一项重要技术。通过从一个完整的案例中抽取的几个有效部分，说明如何从一种范式转换到另一种范式，以及使数据库逻辑设计转换为某一更高范式（至第五范式）的优点。

### 第四部分 方法学

本书的第四部分介绍了一种数据库设计方法学。该方法学被分为三个阶段，分别是概念、逻辑和物理数据库的设计。每个部分都使用DreamHome案例加以阐述：

**第14章** 为概念数据库设计给出逐步指导的方法学。说明了如何将设计分解成多个基于各自视图的易于管理的部分,还给出了识别实体、属性、联系和关键字的方法。

**第15章** 为关系模型的逻辑数据库设计给出逐步指导的方法学。阐述了如何将概念数据模型映射到逻辑数据模型,以及如何针对所需的事务使用规范化技术来验证逻辑数据模型。对于有多个用户视图的数据库应用,这一章还介绍了如何将得到的多个数据模型合并为一个能表示所有视图的全局数据模型。

**第16章和第17章** 为关系系统的物理数据库设计给出逐步指导的方法学。阐述了如何将逻辑数据库设计阶段开发的全局数据模型转换成某关系系统的一个物理设计。方法学中还说明如何通过选择文件组织方式和存储结构,以及通过规范弱化引入可控制冗余来改善实现的性能。

## 第五部分 可选的数据库专题

第五部分阐述了作者认为对于现代数据库管理课程必要的四个专题:

**第18章** 讨论数据库的安全问题,不仅仅是DBMS中的安全性,还包括了整个DBMS环境的安全问题。这一章还检查了Microsoft Access和Oracle提供的一些安全保障,阐述了在Web环境下的一些安全问题,并给出了克服这些问题的方法。

**第19章** 集中讨论了数据库管理系统应该提供的三种功能,即事务管理、并发控制及故障恢复。这些技术用于确保当多个用户访问数据库时或出现硬件/软件组件错误时数据库是可靠且一致的。本章还讨论了更适合于长期事务的高级事务模型。最后阐述Oracle中的事务管理。

**第20章** 阐述查询过程和查询优化。本章讨论查询优化的两种主要技术:一种是使用启发式规则排序查询中的操作,另一种是通过比较不同策略的相对代价选择资源耗费最少的策略。本章还介绍了Oracle中的查询处理。

**第21章** 借助C语言示例程序说明嵌入式和动态SQL。本章还介绍开放数据库互连(ODBC),这一标准现在已经成为访问异构SQL数据库的业界标准。

## 第六部分 当前趋势

第六部分阐述分布式DBMS以及基于对象的DBMS。分布式DBMS技术是当前数据库系统领域的一个主要发展方向。本书前面各章主要介绍集中数据库系统,即由单个DBMS控制的位于单个结点的单一逻辑数据库。

**第22章** 讨论分布式DBMS的概念与问题。用户可以通过分布式DBMS访问位于本结点的数据库及存储在远程结点上的数据。

**第23章** 阐述与分布式DBMS相关的各个高级概念。具体地说,重点阐述与分布式事务管理、并发控制、死锁管理以及数据库恢复相关的协议。此外,本章还讨论了X/Open分布事务处理(DTP)协议,利用复制服务器替代分布式DBMS的方案。本章最后分析了Oracle中的数据分布和复制机制。

本书前面各章都基于关系模型和关系系统。其原因在于该类系统在当前传统的商业数据库应用中占主导地位。不过,关系系统并不是没有缺陷,而面向对象DBMS就是数据库系统中试图克服这些缺陷的一个主要发展方向。第24章到第27章就专门叙述这一方面的发展细节。

**第24章** 作为面向对象DBMS的引言,首先归类已出现的各种新兴的数据库应用,说明关系数据模型因其种种弱点对这些新兴的数据库应用无能为力。此外,介绍了面向对象的一些主要概念及在关系数据库中如何存储对象。

**第 25 章** 阐述面向对象 DBMS。首先介绍面向对象数据模型及持久性编程语言。分析了通常 DBMS 所用的两层存储模型与面向对象 DBMS 所用的单层存储模型的区别及对数据访问的影响。本章还讨论了提供编程语言持久性的不同方法、指针混写的不同技术、版本控制、模式进化和面向对象 DBMS 体系结构等问题。本章还简要介绍了如何将本书第四部分介绍的方法学推广到面向对象 DBMS 中。

**第 26 章** 介绍面向对象管理组推荐的新的对象模型，这一模型已成为面向对象 DBMS 的事实标准。本章还介绍了一个商业面向对象数据库——ObjectStore。

**第 27 章** 介绍对象关系 DBMS，并对刚添加到 SQL 新标准——SQL3 中的各种对象管理特性进行了详细描述。本章还讨论了如何扩展查询处理和查询优化机制，以高效处理扩展数据类型。本章最后检查了 Oracle 中对象关系特性。

## 第七部分 新兴研究方向

本书的最后部分讲述五个正蓬勃发展、日益重要的领域，包括将 DBMS 集成到 Web 环境、半结构化数据及其与 XML 的关系、数据仓库、联机分析处理（OLAP）和数据挖掘。

**第 28 章** 阐述将 DBMS 集成到 Web 环境的问题。本章首先简单介绍 Internet 和 Web 技术，然后说明 Web 作为数据库应用平台的适宜性，并讨论这种方法的优缺点。本章讨论了将 DBMS 集成到 Web 环境的几种不同方法，包括脚本语言，CGI，服务器扩展，Java，ASP 和 Oracle 的 Internet Platform。

**第 29 章** 阐述半结构化数据并讨论 XML 及 XML 如何成为 Web 上数据表示和交换的流行标准。本章讨论 XML 相关技术如名字空间，XSL，XPath，XPointer 和 XLink。本章还阐述怎样用 XML 模式定义 XML 文档的内容模型，以及怎样用资源描述框架（RDF）为元数据交换提供框架。本章还讨论了 XML 的查询语言。

**第 30 章** 本章讨论数据仓库，包括它的定义、进化过程及潜在优缺点。本章阐述了数据仓库的体系结构、主要组成部分和相关工具与技术。本章还讨论数据集市及其开发和管理的有关问题。本章最后分析了 DBMS 中的数据仓库机制。

**第 31 章** 本章提供了为决策支持所用数据仓库和数据集市中数据库的设计方法。描述了维度建模技术的基本概念并将其与传统的实体联系（ER）模型技术做了比较。关于数据仓库的设计，给出了逐步指导的方法学并通过扩展的 DreamHome 案例说明如何实际使用该方法学。这一章最后说明如何用 Oracle 的 Warehouse Builder 设计数据仓库。

**第 32 章** 本章考虑联机分析处理（OLAP）和数据挖掘。首先讨论了多维数据库的相关概念，重点分析了三类主要的 OLAP 工具：多维 OLAP（MOLAP），关系 OLAP（ROLAP）和可管理查询环境（MQE）。此外，以 Red Brick Intelligent SQL 为例，分析了如何扩展 SQL 用于提供复杂的数据分析功能。本章也给出了与数据挖掘有关的概念，指出数据挖掘的操作、技术和工具的主要特点，以及数据挖掘与数据仓库的联系。

## 附录

**附录 A** 本附录给出 DreamHome 案例，它在全书通篇使用。

**附录 B** 本附录给出另外三个案例，供学生课程设计选用。

**附录 C** 本附录介绍文件组织和存储结构相关概念，它们对理解第 16 章讨论的物理数据库设计和第 20 章讨论的查询处理是必要的。

附录 D 本附录给出 Codd 的关于关系 DBMS 的 12 准则，它是判断 RDBMS 的标志。

附录 E 本附录给出有别于 UML 的另两种建模表示法，即 Chen 表示方法和 Crow Feet 表示方法。

附录 G 本附录讨论对 Oracle 数据库如何估计磁盘空间需求。

附录 H 本附录提供若干 Web 脚本示例，补充第 28 章关于 Web 和 DBMS 的讨论。

本书的逻辑组织及建议的阅读路线请见图 P.1。

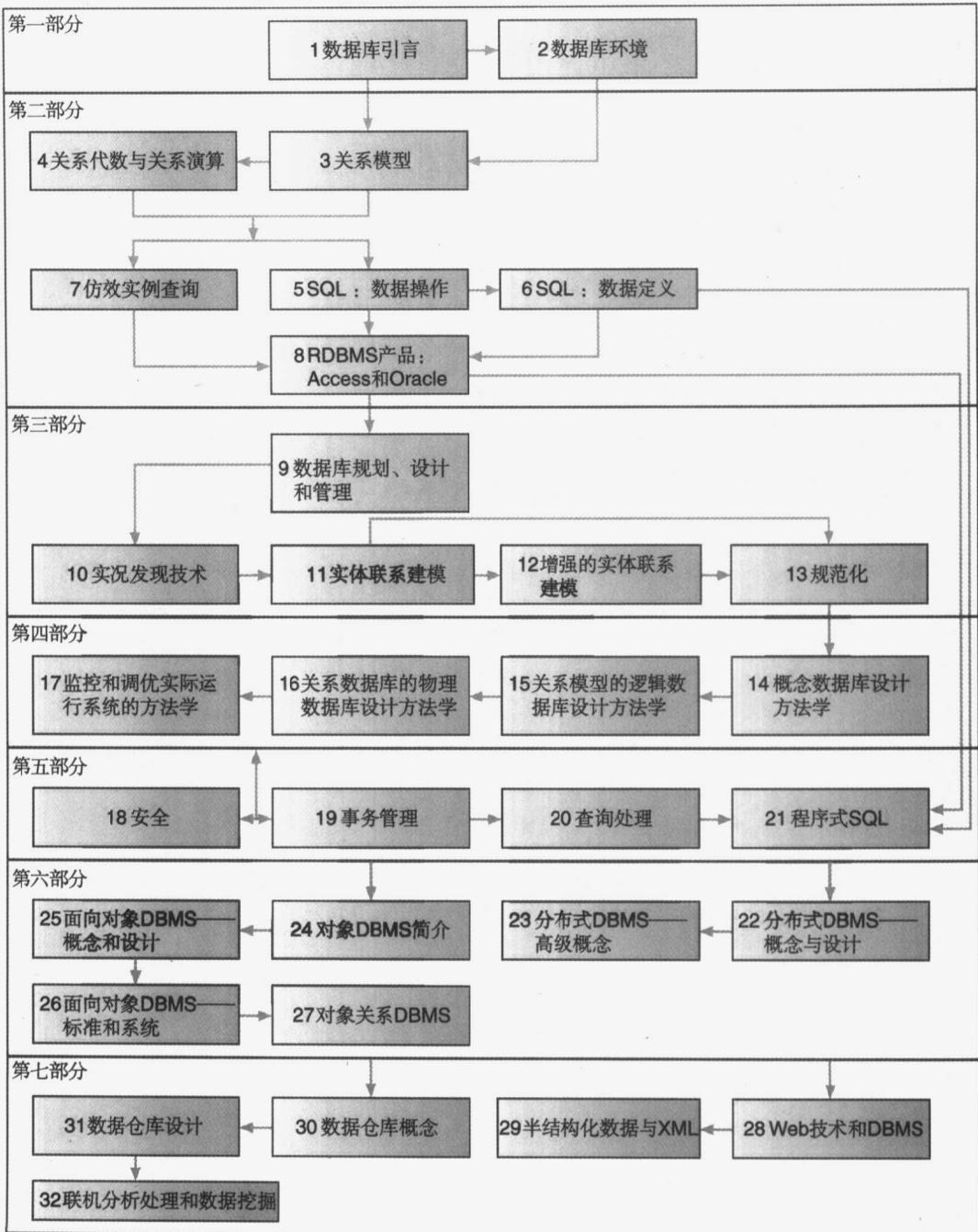


图 P.1 本书的逻辑组织及建议的阅读路线

# 目 录

<b>第一部分 背景知识</b> .....	1
<b>第 1 章 数据库引言</b> .....	2
1.1 简介 .....	3
1.2 传统的基于文件的系统 .....	4
1.3 数据库方法 .....	10
1.4 数据库环境中的角色 .....	15
1.5 数据库管理系统的历史 .....	17
1.6 DBMS 的优点和缺点 .....	18
<b>第 2 章 数据库环境</b> .....	24
2.1 ANSI-SPARC 三层体系结构 .....	25
2.2 数据库语言 .....	29
2.3 数据模型和概念建模 .....	31
2.4 DBMS 的功能 .....	34
2.5 DBMS 的组件 .....	38
2.6 多用户 DBMS 体系结构 .....	40
2.7 系统目录 .....	44
<b>第二部分 关系模型和语言</b> .....	49
<b>第 3 章 关系模型</b> .....	50
3.1 关系模型简史 .....	50
3.2 基本术语 .....	52
3.3 关系完整性 .....	59
3.4 视图 .....	61
<b>第 4 章 关系代数与关系演算</b> .....	64
4.1 关系代数 .....	64
4.2 关系演算 .....	73
4.3 其他语言 .....	78
<b>第 5 章 SQL: 数据操作</b> .....	81
5.1 SQL 介绍 .....	82
5.2 书写 SQL 命令 .....	84
5.3 数据操作 .....	85
<b>第 6 章 SQL: 数据定义</b> .....	113
6.1 ISO SQL 数据类型 .....	114
6.2 完整性增强特性 (IEF) .....	117
6.3 数据定义 .....	120
6.4 视图 .....	126



6.5	事务 .....	135
6.6	访问控制 .....	136
<b>第7章</b>	<b>仿效实例查询 .....</b>	<b>143</b>
7.1	Microsoft Access 查询简介 .....	143
7.2	用 QBE 创建选择查询 .....	145
7.3	使用高级查询 .....	150
7.4	使用动作查询改变表的内容 .....	155
<b>第8章</b>	<b>RDBMS 产品: Access 和 Oracle .....</b>	<b>161</b>
8.1	Microsoft Access 2000 .....	161
8.2	Oracle 8/8i .....	173
<b>第三部分</b>	<b>数据库分析与设计技术 .....</b>	<b>195</b>
<b>第9章</b>	<b>数据库规划、设计和管理 .....</b>	<b>196</b>
9.1	信息系统生存周期 .....	197
9.2	数据库应用的生存周期 .....	197
9.3	数据库规划 .....	199
9.4	系统定义 .....	199
9.5	需求收集和分析 .....	200
9.6	数据库设计 .....	203
9.7	DBMS 选择 .....	206
9.8	应用程序设计 .....	209
9.9	建立原型 .....	211
9.10	实现 .....	212
9.11	数据转换和加载 .....	212
9.12	测试 .....	212
9.13	运行维护 .....	213
9.14	计算机辅助软件工程工具 .....	213
9.15	数据管理和数据库管理 .....	215
<b>第10章</b>	<b>实况发现技术 .....</b>	<b>219</b>
10.1	何时使用实况发现技术 .....	219
10.2	收集实况的类型 .....	220
10.3	实况发现技术 .....	220
10.4	使用实况发现技术的实例 .....	223
<b>第11章</b>	<b>实体联系建模 .....</b>	<b>238</b>
11.1	实体类型 .....	240
11.2	联系类型 .....	240
11.3	属性 .....	244
11.4	强实体类型和弱实体类型 .....	247
11.5	联系的属性 .....	247
11.6	结构化约束 .....	248
11.7	ER 模型的问题 .....	253

第 12 章	增强的实体联系建模 .....	259
12.1	特殊化/概化 .....	259
12.2	聚合 .....	268
12.3	组合 .....	269
第 13 章	规范化 .....	272
13.1	规范化的目标 .....	273
13.2	数据冗余和更新异常 .....	273
13.3	函数依赖 .....	275
13.4	规范化过程 .....	280
13.5	第一范式 .....	280
13.6	第二范式 .....	283
13.7	第三范式 .....	286
13.8	第二范式和第三范式的一般化定义 .....	288
13.9	Boyce-Codd 范式 .....	288
13.10	规范化过程的总结 .....	291
13.11	第四范式 .....	296
13.12	第五范式 .....	297
第四部分	方法学 .....	303
第 14 章	概念数据库设计方法学 .....	304
14.1	数据库设计方法学简介 .....	305
14.2	数据库设计方法学概述 .....	306
14.3	概念数据库设计方法学 .....	307
第 15 章	关系模型的逻辑数据库设计方法学 .....	320
15.1	关系模型的逻辑数据库设计方法学 .....	320
第 16 章	关系数据库的物理数据库设计方法学 .....	347
16.1	逻辑与物理数据库设计的比较 .....	348
16.2	物理数据库设计方法学概述 .....	348
16.3	关系数据库的物理数据库设计方法学 .....	349
第 17 章	监控和调优实际运行系统的方法学 .....	369
第五部分	可选的数据库专题 .....	379
第 18 章	安全 .....	380
18.1	数据库安全 .....	380
18.2	对策——基于计算机的控制 .....	383
18.3	Microsoft Access DBMS 的安全机制 .....	388
18.4	Oracle DBMS 的安全机制 .....	390
18.5	DBMS 和 Web 安全 .....	392
第 19 章	事务管理 .....	400
19.1	事务支持 .....	401
19.2	并发控制 .....	403