

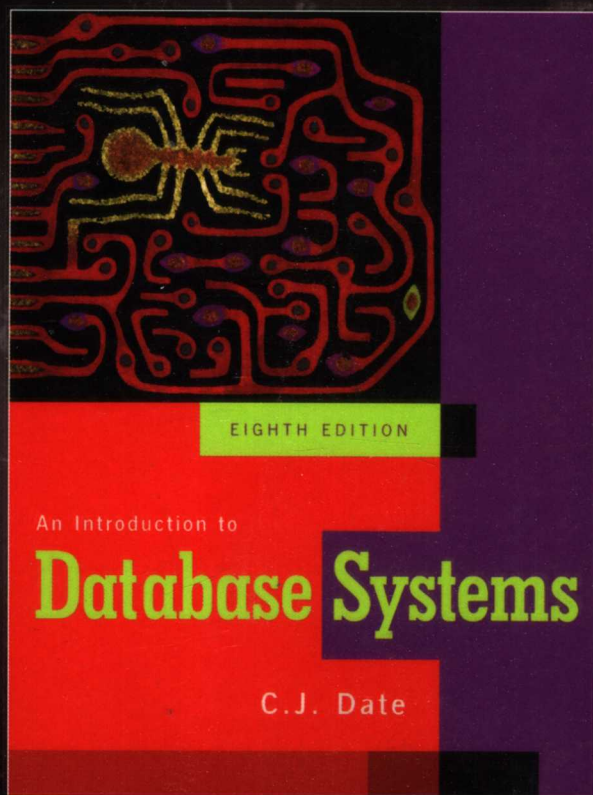


计 算 机 科 学 丛 书

原书第8版

数据库系统导论

(英) C. J. Date 著 孟小峰 王珊 姜芳芃 等译



An Introduction to Database Systems
Eighth Edition



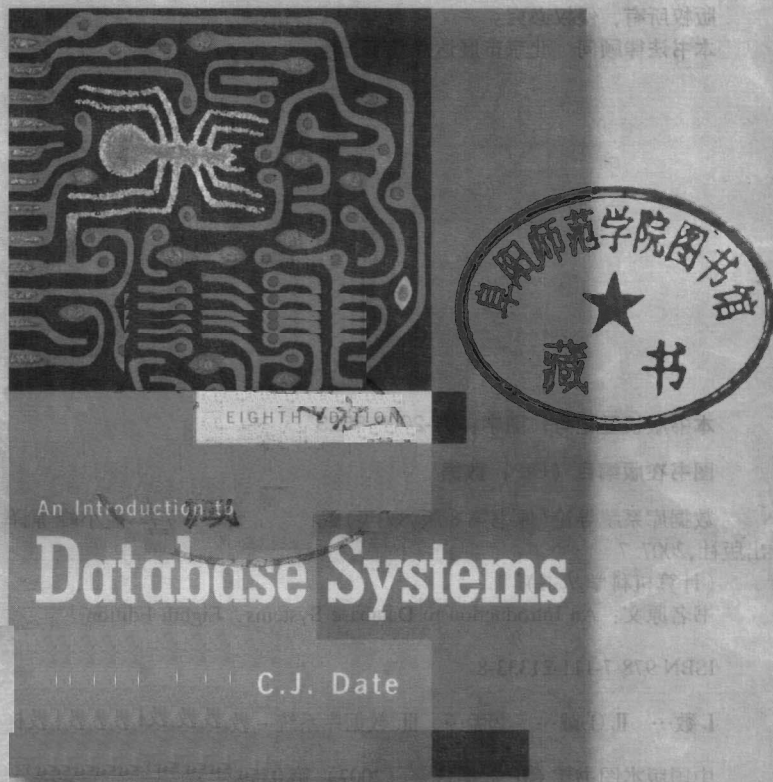
机械工业出版社
China Machine Press

计 算 机 科 学 丛 书

TP392
8-3

数据库系统导论

(英) C. J. Date 著 孟小峰 王珊 姜芳亢 等译



An Introduction to Database Systems
Eighth Edition



机械工业出版社
China Machine Press

本书全面介绍了现在应用广泛的数据库系统,为数据库技术基础知识提供坚实的基础,并对数据库领域的将来发展方向给出看法,本书一直是数据库方面的权威著作。本书整体上可以划分成六个主要部分:基本概念、关系模型、数据库设计、事务管理、高级专题、对象、关系和 XML。第 8 版已经对数据库系统目前的系统的最新发展内容进行了扩充;同时又注重于强调概念的理解,而不仅局限于公式的条陈。

本书可用作计算机展业本科生和研究生学习数据库的教科书,也可供从事数据库研究工作的相关人员作为参考书。

Simplified Chinese edition copyright © 2007 by Pearson Education Asia Limited and China Machine Press.

Original English language title: *An Introduction to Database Systems*, Eighth Edition (ISBN 0-321-19784-4) by C. J. Date, Copyright © 2004.

All rights reserved.

Published by arrangement with the original publisher, Pearson Education, Inc., publishing as Addison Wesley.

本书封面贴有 Pearson Education (培生教育出版集团) 激光防伪标签,无标签者不得销售。

版权所有,侵权必究。

本书法律顾问 北京市展达律师事务所

本书版权登记号:图字:01-2003-8523

图书在版编目(CIP)数据

数据库系统导论(原书第8版)/(英)戴特(Date, C. J.)著;孟小峰等译. - 北京:机械工业出版社, 2007. 7

(计算机科学丛书)

书名原文: *An Introduction to Database Systems*, Eighth Edition

ISBN 978-7-111-21333-8

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

中国版本图书馆 CIP 数据核字(2007)第 055924 号

机械工业出版社(北京市西城区百万庄大街 22 号 邮政编码 100037)

责任编辑:王璐

北京京北制版厂印刷·新华书店北京发行所发行

2007 年 7 月第 1 版第 1 次印刷

184mm × 260mm · 40 印张

定价:75.00 元

凡购本书,如有倒页、脱页、缺页,由本社发行部调换
本社购书热线:(010) 68326294

对本书的赞誉

“C. J. Date 的书是关系理论和数学方法的一面旗帜……也是 SQL 标准的先驱。本书注重内容的翔实以及概念和原理的重要性，从而帮助读者掌握和理解本领域的知识。”

——Carl Eckberg, 圣地亚哥州立大学

“本书第 8 版全面、极佳地展示了当代数据库领域。尤其是，Date 将类型、关系、对象数据库和对象 - 关系数据库等章节综合在一起，给出了关于数据库中对象 - 关系方法的非常清晰并且自包含的解释。”

——Martin K. Solomon, 佛罗里达大西洋大学

“Chris Date 是计算机行业中数据库技术方面最受尊敬的专家和思想家，对于那些想对数据库系统做广泛了解的读者来说，他的这本《数据库系统导论》仍然是一本权威的著作，而且是一本数据库系统当前发展动向的指南。”

——Colin J. White, Intelligent Business Strategies 公司总裁

“这是我所看过的解释并发控制的最好的书，并且这本书囊括了本领域的方方面面。”

——Bruce O. Larsen, 史蒂文斯科技学院

“……是一本不可缺少的读物和参考书。它是信息系统或数据库实践者的必备书籍。”

——Declan Brady, MICS, 富士通公司系统评定和数据库专家

“除了对参考书目翔实的注释之外，作者对该领域的深入洞察、对深奥理论的浅显解释、对一些有争论问题的开放式讨论、以及书中全面的内容等，都使该书成为近 20 多年中数据库领域中最受欢迎的书。”

——Qiang Zhu, 迪尔伯恩, 密歇根大学

“[该书] 最吸引人之处在于其全面性和给出了当前研究发展的最新内容。Date 之所以能够写出该领域的最新动态，得益于他参与了这些发展过程。”

——David Livingstone, 纽斯卡尔, 诺森比亚大学

信息原则（唯一表示的原则）：数据库全部的信息内容有且只有一种表示方式，也就是表中的行列位置有明确的值。

黄金规则：如果一个更新操作将使数据库处于违反自身某个谓词的状态，这样的更新是被禁止的。

（基本和导出的关系变量的）可交换性原则：基本和导出的关系变量没有任意的和不必要的区别。

数据库相对性原则：从用户的角度看，（a）所有的关系变量都是基本关系变量；（b）只要信息等价，数据库是“真实”的这一点是任意的。

规范化原则：

- i. 一个非 5NF 的关系变量应能被分解为一系列的 5NF 的投影关系。
- ii. 通过连接投影能重构最初的关系变量。
- iii. 分解过程应保持独立性。
- iv. 重构过程应使用每一个投影关系。
- v. （……）只要所有的关系变量都符合 5NF，就停止规范化。

正交设计准则：如果 A 和 B 是数据库中任意的两个基关系变量，那么并不存在无损分解将 A 和 B 分为 $A_1, \dots, A_m$ 和 $B_1, \dots, B_n$ （相互独立）使在 $A_1, \dots, A_m$ 中的某些投影 A_i 和在 $B_1, \dots, B_n$ 中的某些投影 B_j 有重叠的含义。

值可替代原则：不管系统在何处需要类型 T 的值，类型 T' （ T' 是 T 的子类型）的值总是可以被替换的。

分布式数据库的基本原则：对用户来说，一个分布式系统看起来应该完全像一个非分布式系统。

译者序

计算机技术的发展过程是一个不断创新的过程。没有任何一项技术的发展能够像计算机技术的发展这样，其创新总是让人应接不暇，其进步总是让人不得歇息的机会。数据库技术是计算机科学技术中发展最快的领域之一，也是应用最广的技术之一，它是计算机信息系统与应用系统的核心技术和重要基础。数据库作为计算机专业及相关专业的核心课程，国内外已有大量与此相关的教科书出版。初步分来，不外以下几类：一是以导论和概论为主的综合类，这些书适合低年级计算机本科的入门教育；二是较为深入的数据库实现技术介绍，这些书适合高年级本科和低年级研究生的提高教育；三是面向某些数据库专门技术的专深内容的讲解，适合数据库研究方向的高年级研究生使用。

近30年来，戴特（C. J. Date）博士的这本《数据库系统导论》（《An Introduction to Database Systems》）一直是数据库方面的权威著作，经久不衰，深受读者的喜爱。作者戴特博士是关系数据库技术领域中非常著名的独立撰稿人、学者和顾问。多年来，戴特先生一直活跃在数据库领域，他是最早认识到卡德博士（E. F. Codd）在关系模型方面所做的开创性贡献的学者之一。此次推出的最新版在继续保持了对数据库技术的基本内容提供翔实讲解的基础上，又对该领域的最新发展进行了介绍。正如Michigan大学的Qiang Zhu博士对该书的评论中提到的“（本书）除了对参考书目的翔实的注释之外，作者对该领域的深入洞察、对深奥理论的浅显的解释、对一些有争论问题的开放式讨论、以及书中有关前沿的内容等，都使该书成为近二十多年中数据库领域中最受欢迎的书。”的确，戴特博士这种深入浅出的写作风格和对大量第一手资料的充分掌握，是本书长久不衰，生命力得以不断延续的根本所在。

大约在20多年前，本书的第2版由科学出版社翻译出版，第7版由机械工业出版社华章分社出版，此次翻译的是本书的第8版。在翻译过程中，本人深感书的内容变化颇大。虽多年从事数据库的教学工作，借本次翻译之机，犹如重新充电。尤其书中对许多传统概念的重新认识和讲解（如关系模型、视图更新等）对本人启发很大。教书不能一成不变，教好书尤其要推陈出新。依据翻译的感受和使用本书一学期的教学实践，归纳起来，我认为本书具有如下的突出的特点：

第一，注重基础知识的讲解。每一本有关数据库的书都有其侧重点，每一个作者也都有其擅长的方面，本书所强调的主要是基础。要想在基础之上做一些事情，必须要牢牢地把握好基础，并且对其有正确的理解。戴特博士一贯认为关系数据模型是现代数据库技术的基础，它使这一领域具有了科学性。如今任何有关数据库技术基础的书籍，若没有包括关系模型，那它是不全面的。同样，如果没有深入理解关系模型，而声称是数据库领域的专家，那也是无法让人接受的。为此本书用长达六章的篇幅来详细介绍关系模型的有关概念（第7版为五章）。

第二，注重用全新的观点和视角认识新旧问题。戴特博士在书的前言中特别引用了一段话“人们可以获得一些确定性的知识，人们还可以从新的发现中更正那些以前所接受的错误的观点，使其更为公正，正是从这个意义上讲，笔者认为自己所一直信奉并且尽力保持的这种不断改变自己观点的哲学方法是科学的。对于笔者说的话，不管是以前说过的还是以后将要说的，笔者并不像神学家信奉他们的信条那样认为它们是真理。笔者所主张的只是在当时为大多数人所接受的较为明智的看法……”我十分钦佩作者的这种精神。比如，关系模型不是一个静态的事物，理解这一点是非常重要的——过去几年里它不断发展和进化，未来仍将如此。作者用他在《对象关系数据库基础：第三版宣言》（1998）中的观点，重新解释了关系模型的含义，并与传统定义相比较。特别地，书中对类型（域）、关系值和关系变量、完整性，谓词和视图等内容都提出了大量新的观点。基于新的解释，作者对视图更新这一传统难题给予了更为清晰的解释，在教与学两个环节上都更便于把握。再比如对传统的规范化理论，增加了新的关系值的属性、逆规范化

(denormalization)、正交设计 (orthogonal design) 和语义模型等有关内容 (包括“商业规则”)。

第三, 注重分析问题的本质, 带给读者以思考的空间。比如第 18 章信息空缺, 先描述空值和 3VL 的基本思想, 然后讨论了这些思想所带来的一些严重后果, 以此来证明空值和 3VL 是非常严重的错误。在关系模型这样清晰规范的系统中应该没有它们的位置, 但作者认为对空值和 3VL 不加介绍是不合适的, 只有通过分析批判, 才能给读者更大的认识空间。同样针对“对象数据库/对象关系数据库为什么不能取代关系数据库”这一问题, 作者用大量的分析比较, 论述的十分精彩。特别是“对象 DBMS 还是一个 DBMS 吗?”, 一语道破对象数据库的本质。

第四, 注重反映数据库的新技术。数据库学科的发展可谓突飞猛进。第 8 版注重补充数据库新的成果, 但又不哗众取宠, 取舍得当是这部分的关键。如第 18 章 (空缺信息), 第 20 章 (类型继承), 第 22 章 (决策支持), 第 23 章 (时态数据库), 第 24 章 (逻辑数据库), 和第 21 章 (分布式数据库), 第 25、26 章 (对象数据库/对象关系数据库) 等, 都是数据库的最新的且相对成熟的技术。

最后, 本书提供了丰富翔实的参考文献和注释。这也是本书一贯的风格和传统。它几乎成为一种书写体例为大家所广泛接受。这些参考书目反映了以前和现在该领域的研究状况, 而且对有兴趣的读者还提供了好的信息资源。我本人就获益匪浅, 当我要查找一个经典文献的出处时, 最先想到的就是戴特博士的这本书。

当然本书也不是没有缺点。可以看出, 戴特博士在尽其所能延续这本书的经典地位, 但毕竟戴特博士脱离数据库学术界和大学教学一线多年, 其对主流技术的把握略显生疏, 比如对数据库实现技术无所涉及, 但好在这方面已有其他很好的教材可以参考。

全书主要包括六个部分和三个附录。第一部分主要介绍了数据库系统中的一般概念和关系数据库系统。第二部分详细介绍了关系模型这一数据库领域的基础理论。第三部分讨论了数据库设计的理论和实践中的一些问题。第四部分主要介绍了事务处理, 包括恢复和并发控制。第五部分讨论了数据库的一些高级主题, 包括: 安全性、优化、类型继承、分布式数据库、时态数据库、逻辑数据库等。最后, 第六部分讲解了对对象关系数据库技术以及 XML 数据管理技术。

本书的翻译和审校由孟小峰和王珊共同组织完成。参加翻译的还有姜芳芳、尹少宜、赖彩凤、李先、凌妍妍、王凌、陈继东、刘伟、周军锋、黄静、贾琳琳等。全书由孟小峰负责统一定稿。

本书涉及面广, 内容丰富, 术语量大, 由于译者水平有限, 译文中不当之处在所难免。诚恳读者批评指正并不吝赐教。如果你有任何建议或意见, 欢迎发 E-mail 至 xfmeng@public.bta.net.cn。

译 者

二〇〇七年五月于北京

译者简介



译者

孟小峰, 1964年生, 博士、教授、博士生导师, 中国人民大学信息学院副院长兼计算机系主任。现为中国计算机学会理事、中国计算机学会普及工委主任、中国计算机学会数据库专委会委员秘书长、电子政务与办公自动化专委会副主任委员、《计算机研究与发展》编委等。主持或参加过二十多项国家科技攻关项目、国家自然科学基金以及国家863项目。获电子部科技进步特等奖1项、北京市科技进步二等奖2项、第七届“中创软件人才奖”、“教育部新世纪优秀人才支持计划”等奖励。研制开发的主要软件产品有国产数据库系统 COBASE、嵌入式移动数据库系统“小金灵”、中文自然语言查询系统 NChql、并行数据库系统 PBASE/1、Native XML 数据库系统 OrientX 等。近年在国内外杂志及国际会议发表论文100余篇, 出版数据库方面的著作七部, 其参与主讲的“数据库系统概论”获评国家精品课程, 并多次应邀担任国际会议 SIGMOD、ICDE、DASFAA、MDM、DEXA 等程序委员及各类主席。近期主要研究领域为 Web 数据集成、XML 数据库、移动数据库等。

<http://www.ccf-dbs.org.cn/idke/xfmeng>

<http://idke.ruc.edu.cn/wamdm/>

本书是数据库系统概论的教材, 全书共分10章。第1章介绍数据库系统概论, 第2章介绍数据库系统组成, 第3章介绍数据库系统性能, 第4章介绍数据库系统安全, 第5章介绍数据库系统维护, 第6章介绍数据库系统应用, 第7章介绍数据库系统开发, 第8章介绍数据库系统管理, 第9章介绍数据库系统测试, 第10章介绍数据库系统部署。

译者

本书是数据库系统概论的教材, 全书共分10章。第1章介绍数据库系统概论, 第2章介绍数据库系统组成, 第3章介绍数据库系统性能, 第4章介绍数据库系统安全, 第5章介绍数据库系统维护, 第6章介绍数据库系统应用, 第7章介绍数据库系统开发, 第8章介绍数据库系统管理, 第9章介绍数据库系统测试, 第10章介绍数据库系统部署。

第8版前言

本书全面介绍了现在广泛应用的数据库系统。本书提供了坚实的数据库技术基础知识，并对数据库领域未来的发展方向给出看法。本书适合作为课本，而不是参考书（当然，在某种程度上本书作为参考书会很有用）。学习本书重点要**理解**，而不是机械的记忆。

本书前提

本书适合任何对计算技术有专业兴趣并在某些方面也想理解数据库系统内涵的读者。作者假设读者应该至少对下面两方面的知识有所了解：

- 1) 现代计算机系统的存储和文件管理（索引等）知识。
- 2) 至少一门高级编程语言的特征（Java、Pascal、PL/I 等）。

关于第一个前提知识，请注意本书的在线附录 D “存储结构和存取方法”中有详细的指南。

结构

必须承认本书比较厚。事实上数据库技术已经发展成了一个非常庞大的领域，以普通的篇幅进行全面介绍是不太可能的（实际上，大多数同类题材的书都在 1000 页左右）。本书整体上可以划分成 6 个主要部分：

- 1) 基本概念。
- 2) 关系模型。
- 3) 数据库设计。
- 4) 事务管理。
- 5) 高级专题。
- 6) 对象、关系和 XML。

每一部分依次被划分成几个章节：

第一部分（共 4 章）全面介绍数据库系统的一般概念，特别是关系数据库系统，并介绍了标准数据库语言 SQL。

第二部分（共 6 章）详尽地介绍关系模型，它不但是关系数据库系统的理论基础，实际上更是整个数据库系统的理论基础。

第三部分（共 4 章）讨论数据库设计的一般问题；其中 3 章介绍设计理论，1 章讲解语义模型和实体/关系模型。

第四部分（共 2 章）涉及事务管理（也就是恢复和并发控制）。

第五部分（共 8 章）内容有点杂，但总体说来是介绍关系的概念和数据库技术未来若干方面的联系——安全性、分布式数据库、时态数据、决策支持等。

最后，第六部分（共 3 章）描述了数据库系统中的对象技术。第 25 章特别介绍了对象系统；第 26 章考虑对象和关系技术共存的可能性，并且讨论对象/关系系统；第 27 章讨论 XML 和数据库的相关性。

本书有 3 个附录。附录 A 概述一种全新的完全不同的实现技术，称作 TransRelational™ 模型；附录 B 给出 SQL 表达式的 BNF 范式；附录 C 是术语表，包括重要的缩略语、首字母缩写词和本书用到的符号。

在线材料^①

- 教师指导手册提供了如何在教数据库课程时使用本书的指导。它包括一系列针对每部分、

① 本教师补充资料只有符合条件的教师才能得到，有需要者请与 Pearson 公司北京办事处联系，联系电话：8610 - 88819178 或 8008100855。——编者注。

每章的笔记、提示和建议，还有附录和其他的辅导资料。

- 习题答案（包含在教师指导手册中）。
- ppt 格式的幻灯片。

每个读者都可以从网址 <http://www.aw.com/cssupport> 得到下面的材料：

- 关于存储结构和存取方法的附录 D（前面提到过）。
- 部分习题的解答。

如何学习本书

读者大体上应该按照本书写作的顺序阅读，但也可以跳过后面的章节，建议第一次阅读本书的顺序是：

- 速读第 1 章和第 2 章。
- 仔细阅读第 3、4 章（4.6 节和 4.7 节可以是例外）。
- 速读第 5 章。
- 仔细阅读第 6、7、9、10 章，但可以跳过第 8 章（8.6 节的 SQL 部分应该看看）。
- 浏览第 11 章。
- 仔细阅读第 12、14 章，^① 可以跳过 13 章。
- 仔细阅读第 15、16 章（可能除了 15.6 节介绍的两阶段提交部分）。
- 可以根据个人的喜好和兴趣，有选择地阅读随后的章节（但要按照前后顺序读）。

每章从引言开始，以小结结束，每章都附有课后习题，本书的配套网站上提供的在线答案还包括习题知识点之外的其他信息。大多数章节还包括一些参考文献，其中很多都有注解。这种结构便于将介绍的主题组织成层次的形式，书的主体包括重要的概念和结果，各种辅助问题和更复杂的方面放在习题、答案或者参考文献中（如果合适的话）。注意：参考文献通过一个方括号中的两部分的数字来区分。例如，参考文献 [3.1] 表示第 3 章最后所列出的参考文献中的第一个，也就是 E. F. Codd 于 1982 年 2 月发表在 CACM 25 第 2 期上的论文。（参考文献中的缩略语，比如“CACM”，其解释参见附录 C。）

和早期版本的比较

本书和以前版本的主要区别是：

- 第一部分：新版的第 1~4 章和第 7 版的第 1~4 章内容大体一致，但在细节上有重大的修订。特别是第 4 章，对 SQL 的介绍升级到当前标准 SQL: 1999，并以标准贯穿整书。（因此，本版与第 7 版相比，有超过半数章节进行了修订。）注意：还有的内容有可能纳入下一个 SQL 标准版本中——2003 年底的时候可能通过——在适当的时候都有所提及。
- 第二部分：第 5~10 章是关于关系模型的内容，和第 7 版的第 5~9 章相比，这部分整体被重写了，还扩充了很多内容，有了很大程度的改进。特别是关于类型（俗称为域）的介绍，被扩充为一章（第 5 章），关于完整性的介绍（第 9 章）我彻底地调整了结构并进行了重写。我在短时间内完成这部分章节的增改，但更新的内容，没有表现在底层概念上，而是表现在我选择如何描述它们的方式上，这些改变是以我在教授这部分内容时描述它们的实践经历为基础的。

注意：这里有必要作进一步的解释。本书的早期版本用 SQL 作为教授关系概念的基础，这是因为我们相信，具体的东西比抽象的东西更容易展示。但不幸的是，SQL 和关系模型之间的鸿沟扩大了，并且这一鸿沟有继续扩大的趋势，最终会达到一种状况，再这样继续下去，有可能会误导我们对 SQL 的使用。可悲的是，SQL 越来越远离作为关系原理的具体体现，它包含了太多不合理的地方——冗长、功能太多，坦率地说我压根都不想讨论它。但是，从商业的角度看 SQL 显然很重要，因此每个数据库专业人员都有必要熟悉它，在本书中忽略这个问题也是不恰当的。因此在书中介绍了我的解决策略：（a）第一部分有一章来介绍 SQL 基础；（b）其他章中有单独的部分描述该章讨论的主题涉及的 SQL 诸方面。通过这种方法本书提供了全面的、广泛的 SQL 材料，这些材料放在我觉得

① 如果愿意可以提前读第 14 章，不过最好先读完第 4 章。

合适的上下文中。

- 第三部分：第 10 ~ 13 章是第 7 版中第 9 ~ 12 章的完美修订。在内容细节上有全面的改善。

注意：一些进一步的解释是有必要的。一些批评家抱怨说早期版本关于数据库设计主题的讨论太靠后了，但我认为是学生们还没有准备好来正确地设计数据库或者来充分理解设计主题，除非他们初步理解了数据库是什么和数据库是如何应用的；换句话说，我认为在向学生提出设计问题前，在关系模型和相关主题上花些时间是很重要的。因此我仍然认为第三部分的位置是合适的。（的确，我知道有些老师愿意提前教授实体/关系模型，因此我尽量使第 14 章自成一体，这样他们就可以在讲完第 4 章后马上讲述第 14 章了。）

- 第四部分：这部分的两章，第 15、16 章，是从第 7 版的第 14、15 章修订而来的，是完全的重写、扩展和改进。特别是第 16 章，现在包括对所谓的事务的 ACID 属性的仔细分析和一些相关的非正统的结论。
- 第五部分：第 20 章关于类型继承和第 23 章关于时态数据库都全部重写，以反映相关领域的最新研究和发展。对其他章节的修订都是些美化工作，不过在解释和举例上有所改善，并且多处引用了新材料。
- 第六部分：第 25 和 26 章是第 7 版的第 24 和 25 章的改善和扩展。第 27 章是全新的。

最后，附录 A 也是新的，附录 B 和 C 分别是对第 7 版的附录 A 和 C 的修订（第 7 版的附录 B 的材料已经合并到本书的主体部分了）。

本书的特色

市场上的每本数据库书籍都有各自的优缺点，并且每个作者都是有针对性地著书。一个集中于事务管理问题，一个强调实体/关系模型；另一个仔细分析 SQL 的各个方面，也有以纯粹“对象”的观点来介绍数据库的，还有专门讨论一些商业数据库产品的，等等。当然我也不例外，我也有把削减题材的斧头：也许是可以被称为基础的斧头。我坚定地相信，当我们在基础之上再有所发展之前，必须打好基础，对它有恰当的理解。这种信念解释了我为什么在本书中如此强调关系模型；特别是第二部分的篇幅——本书中最重要的部分——在这部分中我尽可能谨慎地提出了我对关系模型的理解。我对基础很感兴趣，而不是时尚或流行的东西。产品总是在改变，但原则更持久。

在这方面，我希望能提醒你注意这个事实，对某些很重要的主题（基础），本书使用一个完整、深入的章（或者附录）来介绍。讨论的主题有：

- 类型
- 完整性
- 视图
- 缺失信息
- 继承
- 时态数据库
- TransRelational™ 模型

基于同样的观点（基础的重要性），我必须承认本书的整体叙述方式这几年来已经变了。前几版主要是自然的描述；描述该领域在实践中实际的样子，“尽管有各种缺点与瑕疵”。后来的版本，相对而言说明性更多些；讨论的是该领域应该是什么样子和未来可能的发展方向，如果我们行事正确的话。现在的版本在这个意义上的确是说明性的了（因此这是带有个人观点的教材）。既然“行事正确”的第一部分是教育人什么是正确的事，我希望，本书的新版本能在这方面有所帮助。

另一个相关的观点：也许你也知道，我最近和我的同事 Hugh Darwen 合作出版了另一本“说明性”的书，《Foundation for Future Database Systems: The Third Manifesto》（本书参考文献 [3.3]）^①。我们称该书为《第三个宣言》，或者简称为《宣言》，它基于关系模型为未来的数据库系统提供详细的技术上的建议；它是 Hugh 和我多年教学与思考这些问题的结晶。并且，《宣

① 本书有一个相关站点是 <http://www.thethirdmanifesto.com> 更多的相关材料也见 <http://www.dbdebunk.com>

言》的思想渗透在本书中。这并不是说阅读《宣言》是阅读本书的必要前提，但它的确是和本书涉及的内容有很大的直接相关性，并且在那里经常会发现更进一步的相关信息。

注意：参考文献 [3.3] 使用 **Tutorial D** 语言，用来解释目的，本书也是一样。Tutorial D 的语法和语义多多少少是自解释的（该语言可能是描述特征的，自由松散的，像“Pascal”之类），但是个别的特征在第一次使用的时候会有所解释，如果有必要的话。

结束语

我愿意用下面略微修改过的另一篇序言来结束这个前言——伯兰特·罗素的《伯兰特·罗素关于智力、物质和道德辞典》。

我被指责有不停改变观点的恶习……我自己丝毫不以之（前述的习惯）为耻。活跃于 1900 年的哪个唯物论者能自夸在最后的半个世纪他的观点从没改变过？……我所重视并且努力去追求的哲学是科学，从某种意义上讲就是要获得真实的知识，新的发现会使任何聪慧的头脑都不可避免地承认先前的错误。至于我曾经说过的，无论说得早或晚，我不会像那些空头理论家出于他们的信念那样宣称真理。我只能说我所表达的观点目前来说是明智合理的……如果后续的研究不能表明它需要修改的话，我会很惊讶的。这种观点不是傲慢武断的声明，而是我目前能做到的更接近于清楚和正确思考的最好声明。文字的清晰是我追求的最终目的。

如果你比较过本书第 8 版和以前的版本，你会发现我在许多事情上已经改变了观点（毫无疑问还会继续改变下去）。我希望你能接受上面引用的评论，来恰当地判断这件事的目前状态。我赞同伯兰特·罗素关于科学质询是什么的观点，但是他对此观点的描述比我所能描述的更加雄辩、更有说服力。

致谢

我很乐意再一次感谢在本书写作过程中有过直接或间接参与的许多人：

首先，我必须感谢我的朋友 David McGoveran 和 Nick Tindall，感谢他们在本书出版中花费的心血；David 提供了第 22 章关于决策支持的初稿，Nick 提供了第 27 章关于 XML 的初稿。我也必须感谢我的朋友和同事 Hugh Darwen 对手稿中所有 SQL 部分的主要帮助（多种形式的帮助）。Nagraj Alur 和 Fabian Pascal 提供给我各种各样的技术背景材料。特别感谢 Steve Tarin 发明了附录 A 中描述的技术，并且感谢他帮助我彻底理解了该技术。

其次，本书的内容受益于过去几年中我教过的学生在研讨会上提出的意见，也颇得益于许多朋友和评论家的意见和讨论，包括 IBM 公司的 Hugh Darwen，根特大学的 Guy de Tré，圣地亚哥州立大学的 Carl Eckberg，Rensselaer 工艺美术学院的 Cheng Hsu，密歇根大学迪尔伯特分校的 Abdul-Rahman Itani，波士顿大学的 Vijay Kanabar，史蒂文斯技术研究所的 Bruce O. Larsen，纽卡斯尔 Northumbria 大学的 David Livingstone，Alternative Technologies 的 David McGoveran，IBM 公司的 Steve Miller，独立咨询员 Fabian Pascal，佛罗里达亚特兰大大学的 Martin K. Solomon，Required Technologies 的 Steve Tarin 和 IBM 公司的 Nick Tindall。上述名单中的每一位或者至少校阅了手稿中的某些部分，或者提供了技术材料，或者帮我找到了许多技术问题的答案，我很感谢他们。

我要感谢我的妻子 Lindy 又帮我设计了书的封面，感谢她在我写本书和这么多年来所有其他数据库相关书籍过程中一贯的支持。

最后，我要感谢 Addison-Wesley 公司的 Maite Suarez-Rivas 和 Katherine Harutunian，感谢他们在我写这本书时所给予的鼓励和支持，同时本书设计者 Elisabeth Beller 的辛勤工作。

C. J. Date
Healdsburg, California, 2003

目 录

译者序
译者简介
第8版前言

第一部分 基础知识

第1章 数据库管理概述	2
1.1 引言	2
1.2 数据库系统的构成	4
1.3 数据库的内涵	6
1.4 使用数据库的优点	9
1.5 数据独立性	11
1.6 关系系统及其他数据库系统	14
1.7 小结	16
习题	16
参考文献	17
第2章 数据库系统体系结构	18
2.1 引言	18
2.2 三级体系结构	18
2.3 外部层	19
2.4 概念层	21
2.5 内部层	22
2.6 映像	22
2.7 数据库管理员	23
2.8 数据库管理系统	24
2.9 数据通信	26
2.10 客户/服务器体系结构	26
2.11 实用程序	27
2.12 分布式处理	27
2.13 小结	29
习题	29
参考文献	30
第3章 关系数据库简介	32
3.1 引言	32
3.2 关系模型概述	32
3.3 关系和关系变量	34
3.4 关系的含义	36
3.5 优化	37
3.6 数据字典	38
3.7 基本关系变量和视图	39
3.8 事务	41
3.9 供应商和零件数据库	42

3.10 小结	43
习题	44
参考文献	45
第4章 SQL 简介	47
4.1 引言	47
4.2 SQL 基本操作	47
4.3 目录	49
4.4 视图	50
4.5 事务	50
4.6 嵌入式 SQL	51
4.7 动态 SQL 和 SQL/CLI	55
4.8 SQL 并不完美	57
4.9 小结	57
习题	58
参考文献	59

第二部分 关系模型

第5章 类型	64
5.1 引言	64
5.2 值与变量	64
5.3 类型与表示	66
5.4 类型定义	69
5.5 操作符	70
5.6 类型生成子	74
5.7 SQL 支持	75
5.8 小结	80
习题	81
参考文献	82
第6章 关系	83
6.1 引言	83
6.2 元组	83
6.3 关系类型	86
6.4 关系的值	87
6.5 关系变量	92
6.6 SQL 的支持	96
6.7 小结	100
习题	101
参考文献	102

第 7 章 关系代数	105
7.1 引言	105
7.2 关系封闭性	106
7.3 基本代数: 语法	107
7.4 基本代数: 语义	109
7.5 举例	115
7.6 关系代数的作用	117
7.7 深入讨论	118
7.8 附加的操作符	119
7.9 分组与解组	124
7.10 小结	126
习题	126
参考文献	128
第 8 章 关系演算	130
8.1 引言	130
8.2 元组演算	131
8.3 举例	136
8.4 关系演算与关系代数的比较	138
8.5 计算能力	141
8.6 SQL 语言	142
8.7 域演算	149
8.8 QBE	151
8.9 小结	154
习题	155
参考文献	156
第 9 章 完整性	158
9.1 引言	158
9.2 进一步讨论	159
9.3 谓词和命题	161
9.4 关系变量谓词和数据库谓词	162
9.5 约束检查	162
9.6 内部谓词与外部谓词	163
9.7 正确性与一致性	164
9.8 完整性和视图	165
9.9 约束分类模式	166
9.10 码	167
9.11 触发器	172
9.12 SQL 的支持	174
9.13 小结	177
习题	178
参考文献	180
第 10 章 视图	185
10.1 引言	185
10.2 视图的用途	187
10.3 视图检索	189
10.4 视图更新	189
10.5 快照	199
10.6 SQL 对视图的支持	200

10.7 小结	202
习题	202
参考文献	203

第三部分 数据库设计

第 11 章 函数依赖	209
11.1 引言	209
11.2 基本概念	209
11.3 平凡的函数依赖和非平凡的函数 依赖	211
11.4 依赖集的闭包	211
11.5 属性集的闭包	212
11.6 最小函数依赖集	214
11.7 小结	215
习题	216
参考文献	217
第 12 章 进一步规范化I: 1NF、2NF、 3NF 和 BCNF	219
12.1 引言	219
12.2 无损分解和函数依赖	221
12.3 第一、第二和第三范式	223
12.4 保持函数依赖	228
12.5 BOYCE/CODD 范式	229
12.6 具有关系值属性的关系变量	233
12.7 小结	234
习题	234
参考文献	236
第 13 章 进一步规范化II: 高级范式	238
13.1 引言	238
13.2 多值依赖与第四范式	238
13.3 连接依赖与第五范式	241
13.4 规范化过程小结	244
13.5 逆规范化	246
13.6 正交设计	247
13.7 其他的规范化形式	249
13.8 小结	250
习题	250
参考文献	251
第 14 章 语义建模	256
14.1 引言	256
14.2 总体方法	257
14.3 E/R 模型	258
14.4 E/R 图	261
14.5 基于 E/R 模型的数据库设计	262
14.6 简单分析	265
14.7 小结	267
习题	268
参考文献	269

第四部分 事务管理

第 15 章 恢复	280
15.1 引言	280
15.2 事务	280
15.3 事务恢复	283
15.4 系统恢复	284
15.5 介质恢复	286
15.6 两阶段提交	286
15.7 保存点	287
15.8 SQL 对事务的支持	287
15.9 小结	288
习题	288
参考文献	289

第 16 章 并发	292
16.1 引言	292
16.2 三个并发问题	292
16.3 锁	294
16.4 重提三个并发问题	295
16.5 死锁	297
16.6 可串行性	298
16.7 重提恢复	299
16.8 隔离级别	300
16.9 意向锁	301
16.10 ACID 的不足之处	303
16.11 SQL 的支持	306
16.12 小结	306
习题	307
参考文献	308

第五部分 高级专题

第 17 章 安全性	314
17.1 引言	314
17.2 自主存取控制	315
17.3 强制存取控制	319
17.4 统计数据库	320
17.5 数据加密	324
17.6 SQL 的支持	327
17.7 小结	329
习题	330
参考文献	330

第 18 章 优化	333
18.1 引言	333
18.2 一个启发性的示例	334
18.3 查询处理概述	334
18.4 表达式转换	337
18.5 数据库统计信息	340
18.6 分而治之的策略	341
18.7 关系操作的实现算法	343

18.8 小结	346
习题	347
参考文献	349

第 19 章 信息空缺	362
19.1 引言	362
19.2 3VL 方法概述	363
19.3 上述方案所造成的某些结果	367
19.4 空值和码	369
19.5 外连接	370
19.6 特殊值	372
19.7 SQL 的支持	372
19.8 小结	376
习题	376
参考文献	378

第 20 章 类型继承	381
20.1 引言	381
20.2 类型的层次结构	383
20.3 多态性和可置换性	386
20.4 变量与赋值	388
20.5 约束特化	391
20.6 比较	392
20.7 操作、版本和签名	395
20.8 圆是椭圆吗	397
20.9 再论约束特化	400
20.10 SQL 的支持	401
20.11 小结	404
习题	405
参考文献	406

第 21 章 分布式数据库	409
21.1 引言	409
21.2 预备知识	409
21.3 十二个目标	411
21.4 分布式系统面对的问题	416
21.5 客户/服务器系统	424
21.6 DBMS 独立性	426
21.7 SQL 的支持	429
21.8 小结	429
习题	430
参考文献	430

第 22 章 决策支持	436
22.1 引言	436
22.2 决策支持的特征	437
22.3 决策支持的数据库设计	438
22.4 数据准备	442
22.5 数据仓库和数据集市	443
22.6 联机分析处理	446
22.7 数据挖掘	451
22.8 SQL 的支持	452

22.9 小结	453	25.5 混合性问题	533
习题	453	25.6 小结	538
参考文献	454	习题	539
第 23 章 时态数据库	457	参考文献	540
23.1 引言	457	第 26 章 对象/关系数据库	546
23.2 问题是什么	459	26.1 引言	546
23.3 时间区间	463	26.2 第一个根本性错误	548
23.4 归并和反归并关系	467	26.3 第二个根本性错误	552
23.5 关系操作符推广	474	26.4 实现上的问题	554
23.6 数据库设计	478	26.5 真正融合的好处	556
23.7 完整性约束	482	26.6 SQL 工具	557
23.8 小结	486	26.7 小结	561
习题	487	习题	562
参考文献	488	参考文献	562
第 24 章 基于逻辑的数据库	489	第 27 章 互联网与 XML	569
24.1 引言	489	27.1 引言	569
24.2 综述	489	27.2 万维网和因特网	569
24.3 命题演算	491	27.3 XML 综述	570
24.4 谓词演算	494	27.4 XML 数据定义	577
24.5 数据库的证明理论观点	498	27.5 XML 数据操纵	583
24.6 演绎数据库系统	501	27.6 XML 和数据库	589
24.7 递归查询过程	504	27.7 SQL 的支持	590
24.8 小结	509	27.8 小结	594
习题	510	习题	595
参考文献	511	参考文献	596
第六部分 对象、关系和 XML		附 录	
第 25 章 对象数据库	516	附录 A TransRelational 模型	599
25.1 引言	516	附录 B SQL 表达式	613
25.2 对象、类、方法和消息	518	附录 C 缩略语和符号	620
25.3 进一步的分析	521		
25.4 一个详实的示例	526		

第一部分 基础知识

第一部分包括以下4章：

- 第1章介绍什么是数据库和为什么需要数据库系统。该章还简述了关系数据库系统和其他数据库系统之间的不同之处。
- 第2章介绍数据库系统的一般体系结构，即 ANSI/SPARC 体系结构。该体系结构构成本书其他部分的基本框架。
- 第3章概括介绍了关系系统。其目的是为在第二部分和后续章节中进一步讨论奠定基础。该章还介绍并解释了一个运行实例：供应商和零件数据库。
- 最后，第4章介绍了标准关系语言 SQL（更精确地说是 SQL: 1999）。