



计 算 机 科 学 丛 书

原书第6版

Mc
Graw
Hill Education

数据库系统概念

Abraham Silberschatz Henry F. Korth S. Sudarshan 著

耶鲁大学

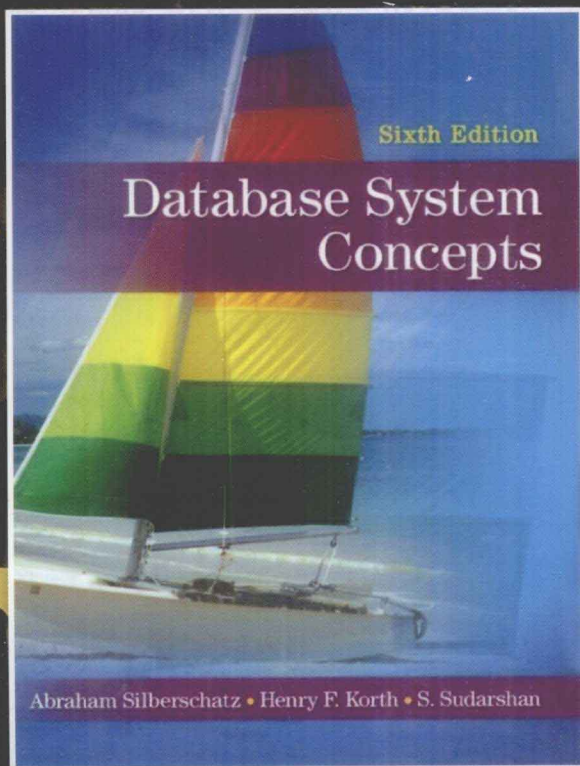
利哈伊大学

印度理工学院

杨冬青 李红燕 唐世渭 等译
北京大学

Database System Concepts

Sixth Edition



机械工业出版社
China Machine Press



计 算 机 科 学 丛 书

原书第6版

Education

数据库系统概念

Abraham Silberschatz Henry F. Korth S. Sudarshan 著

耶鲁大学

利哈伊大学

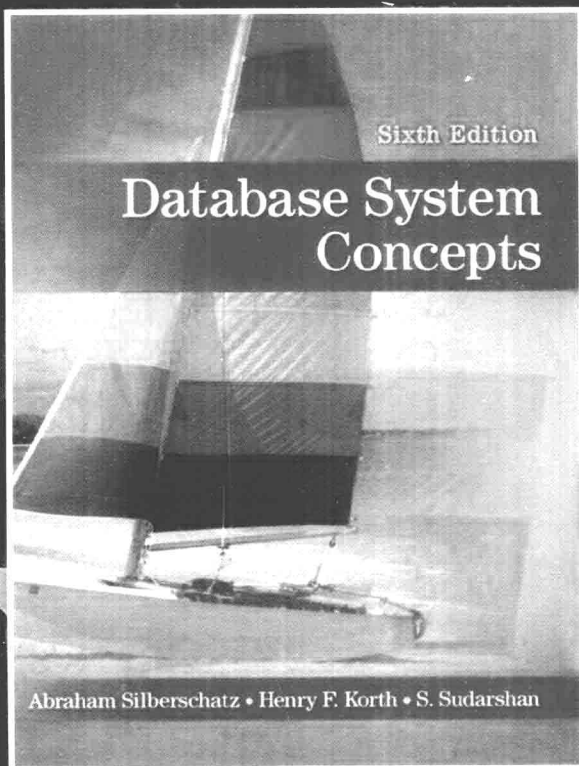
印度理工学院

杨冬青 李红燕 唐世渭 等译

北京大学

Database System Concepts

Sixth Edition



机械工业出版社
China Machine Press

本书是经典的数据库系统教科书《Database System Concepts》的最新修订版，全面介绍数据库系统的各种知识，透彻阐释数据库管理的基本概念。本书内容丰富，不仅讨论了关系数据模型和关系语言、数据库设计过程、关系数据库理论、数据库应用设计和开发、数据存储结构、数据存取技术、查询优化方法、事务处理系统和并发控制、故障恢复技术、数据仓库和数据挖掘，而且对性能调整、性能评测标准、数据库应用测试和标准化、空间和地理数据、时间数据、多媒体数据、移动和个人数据库管理以及事务处理监控器、事务 workflow、电子商务、高性能事务系统、实时事务系统和持续长时间的事务等高级应用主题进行了广泛讨论。

本书既可作为高年级本科生或低年级研究生的数据库课程教材，也可供数据库领域的技术人员参考。

Abraham Silberschatz, Henry F. Korth, S. Sudarshan: Database System Concepts, Sixth Edition (ISBN 978-0-07-352332-3).

Copyright © 2011 by The McGraw-Hill Companies, Inc.

All Rights reserved. No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including without limitation photocopying, recording, taping, or any database, information or retrieval system, without the prior written permission of the publisher.

This authorized Chinese translation edition is jointly published by McGraw-Hill Education (Asia) and China Machine Press. This edition is authorized for sale in the People's Republic of China only, excluding Hong Kong, Macao SAR and Taiwan.

Copyright © 2012 by McGraw-Hill Education (Asia), a division of the Singapore Branch of The McGraw-Hill Companies, Inc. and China Machine Press.

版权所有。未经出版人事先书面许可，对本出版物的任何部分不得以任何方式或途径复制或传播，包括但不限于复印、录制、录音，或通过任何数据库、信息或可检索的系统。

本授权中文简体字翻译版由麦格劳-希尔（亚洲）教育出版公司和机械工业出版社合作出版。此版本经授权仅限在中华人民共和国境内（不包括香港特别行政区、澳门特别行政区和台湾）销售。

版权 © 2012 由麦格劳-希尔（亚洲）教育出版公司与机械工业出版社所有。

本书封面贴有 McGraw-Hill 公司防伪标签，无标签者不得销售。

封底无防伪标均为盗版

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：01-2010-3824

图书在版编目（CIP）数据

数据库系统概念（原书第6版）/（美）西尔伯沙茨（Silberschatz, A.）等著；杨冬青等译. —北京：机械工业出版社，2012.3

（计算机科学丛书）

书名原文：Database System Concepts, Sixth Edition

ISBN 978-7-111-37529-6

I. 数… II. ①西… ②杨… 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（2012）第 028157 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：李荣 谢晓芳

北京诚信伟业印刷有限公司印刷

2012 年 3 月第 1 版第 1 次印刷

185mm × 260mm · 52.25 印张

标准书号：ISBN 978-7-111-37529-6

定价：99.00 元

凡购本书，如有缺页、倒页、脱页，由本社发行部调换

客服热线：(010) 88378991；88361066

购书热线：(010) 68326294；88379649；68995259

投稿热线：(010) 88379604

读者信箱：hzjsj@hzbook.com

文艺复兴以降，源远流长的科学精神和逐步形成的学术规范，使西方国家在自然科学的各个领域中取得了垄断性的优势；也正是这样的传统，使美国在信息技术发展的六十多年间名家辈出、独领风骚。在商业化的进程中，美国的产业界与教育界越来越紧密地结合，计算机学科中的许多泰山北斗同时身处科研和教学的最前线，由此而产生的经典科学著作，不仅擘划了研究的范畴，还揭示了学术的源变，既遵循学术规范，又自有学者个性，其价值并不会因年月的流逝而减退。

近年，在全球信息化大潮的推动下，我国的计算机产业发展迅猛，对专业人才的需求日益迫切。这对计算机教育界和出版界都既是机遇，也是挑战；而专业教材的建设在教育战略上显得举足轻重。在我国信息技术发展时间较短的现状下，美国等发达国家在其计算机科学发展的几十年间积淀和发展的经典教材仍有许多值得借鉴之处。因此，引进一批国外优秀计算机教材将对我国计算机教育事业的发展起到积极的推动作用，也是与世界接轨、建设真正的世界一流大学的必由之路。

机械工业出版社华章公司较早意识到“出版要为教育服务”。自1998年开始，我们就将工作重点放在了遴选、移译国外优秀教材上。经过多年的不懈努力，我们与 Pearson, McGraw-Hill, Elsevier, MIT, John Wiley & Sons, Cengage 等世界著名出版公司建立了良好的合作关系，从他们现有的数百种教材中甄选出 Andrew S. Tanenbaum, Bjarne Stroustrup, Brian W. Kernighan, Dennis Ritchie, Jim Gray, Alfred V. Aho, John E. Hopcroft, Jeffrey D. Ullman, Abraham Silberschatz, William Stallings, Donald E. Knuth, John L. Hennessy, Larry L. Peterson 等大师名家的一批经典作品，以“计算机科学丛书”为总称出版，供读者学习、研究及珍藏。大理石纹理的封面，也正体现了这套丛书的品位和格调。

“计算机科学丛书”的出版工作得到了国内外学者的鼎力襄助，国内的专家不仅提供了中肯的选题指导，还不辞劳苦地担任了翻译和审校的工作；而原书的作者也相当关注其作品在中国的传播，有的还专程为其书的中译本作序。迄今，“计算机科学丛书”已经出版了近两百个品种，这些书籍在读者中树立了良好的口碑，并被许多高校采用为正式教材和参考书籍。其影印版“经典原版书库”作为姊妹篇也被越来越多实施双语教学的学校所采用。

权威的作者、经典的教材、一流的译者、严格的审校、精细的编辑，这些因素使我们的图书有了质量的保证。随着计算机科学与技术专业学科建设的不断完善和教材改革的逐渐深化，教育界对国外计算机教材的需求和应用都将步入一个新的阶段，我们的目标是尽善尽美，而反馈的意见正是我们达到这一终极目标的重要帮助。华章公司欢迎老师和读者对我们的工作提出建议或给予指正，我们的联系方式如下：

华章网站：www.hzbook.com

电子邮件：hzjsj@hzbook.com

联系电话：(010) 88379604

联系地址：北京市西城区百万庄南街1号

邮政编码：100037



华章科技图书出版中心

数据库系统是对数据进行存储、管理、处理和维护的软件系统，是现代计算环境中的一个核心成分。随着计算机硬件、软件技术的飞速发展和计算机系统在各行各业的广泛应用，数据库技术的发展尤其迅速，引人注目。有关数据库系统的理论和技术是计算机科学技术教育中必不可少的部分。《数据库系统概念》是一本经典的、备受赞扬的数据库系统教科书。其内容由浅入深，既包含数据库系统的基本概念，又反映数据库技术的新进展。本书被国际上许多著名大学所采用，并多次再版。

我们先后将本书的第3版、第4版和第5版译成中文，由机械工业出版社分别于2000年、2003年和2006年出版发行。国内许多大学采用《数据库系统概念》作为本科生和研究生数据库课程的教材或主要教学参考书，收到了良好的效果。现在，我们又翻译了该书第6版。第6版保持了前5版的总体风格，同时对内容进行了扩充，对结构进行了调整，以更好地符合数据库教学的需求，反映数据库设计、管理和使用方式的发展和变化。第6版的内容大体上可以分为五个部分。

第1~9章讲述数据库系统的基本概念，包括对数据库系统的性质和目标的综述，对关系数据模型和关系语言的介绍，对数据库设计过程、关系数据库理论以及数据库应用设计和开发（包括基于Web的界面构建数据库应用和应用安全性问题等）的详细讨论。

第10~19章主要讨论数据库系统实现技术，包括数据存储结构、数据存取技术、查询优化方法、事务处理系统的基本概念和并发控制、故障恢复技术，还包括在并行数据库系统和分布式数据库系统中所采用的一些主要策略和技术。

第20~23章主要讨论数据管理与应用的深入话题，包括对数据仓库和数据挖掘概念与技术的较详细的介绍，以及对用于查询文本数据的信息检索技术（包括在Web搜索引擎中使用的基于超链接的技术）的介绍。这一部分还介绍了新型的数据库系统，包括对象-关系数据库模型、数据表示的XML标准，以及XML的查询语言。

第24~26章是一些高级话题，内容包括应用开发中的诸如性能调整、性能评测标准、数据库应用测试和标准化等高级话题，以及空间和地理数据、时间数据、多媒体数据、移动和个人数据库管理中的问题。这一部分还讨论了事务处理监控器、事务工作流、电子商务、高性能事务系统、实时事务系统和持续长时间的事务等高级事务处理问题。

第27~30章对PostgreSQL、Oracle、IBM DB2和Microsoft SQL Server这四个领先的数据库系统进行实例研究，结合这几个具体系统来讨论前面各部分描述的各种实现技术是如何使用到实际系统中的。

上述五大部分中第一部分的主要内容，以及第二、第三、第四部分的部分内容可以作为本科生数据库概论课程的教材或主要参考资料，第二、第三和第四部分的其余内容可以用于研究生的数据库课程教学，第五部分可以作为帮助学生了解实际系统的补充材料。

杨冬青、李红燕、唐世渭组织并参加了本书的翻译和审校工作；参加翻译的还有范红杰、程序、苗高杉、邹淼、陈巍、王婧、王林青、孟必平。

限于译者水平，译文中疏漏和错误难免，欢迎批评指正。

译者

于北京大学

数据库管理已经从一种专门的计算机应用发展为现代计算环境中的一个重要成分，因此，有关数据库系统的知识已成为计算机科学教育中的一个核心的部分。在本书中，我们讲述数据库管理的基本概念。这些概念包括数据库设计、数据库语言、数据库系统实现等多个方面。

本书可作为本科生三年级或四年级数据库入门课程的教科书，也可作为研究生一年级的教科书。除了作为入门课程的基本内容外，本书还包括了可作为课程补充或作为高级课程介绍性材料的高级内容。

我们仅要求读者熟悉基本的数据结构、计算机组织结构和一种高级程序设计语言，例如 Java、C 或 Pascal。书中的概念都以直观的方式加以描述，其中的许多概念都基于我们大学运行的例子加以阐释。本书中包括重要的理论结果，但省略了形式化证明，取而代之的是用图表和例子来说明为什么结论是正确的。对于形式化描述和研究结果的证明，读者可以参考文献注解中列出的研究论文和高级教材。

本书中所包括的基本概念和算法通常是基于当今的商品化或试验性的数据库系统中采用的概念和算法。我们的目标是在一个通常环境下描述这些概念和算法，而没有与某个特定的数据库系统绑定。特定的数据库系统的细节将在第九部分“实例研究”中讨论。

在本书第 6 版中，我们保持了前面版本的总体风格，同时对内容和结构进行了扩展来反映数据库设计、管理和使用的方式所发生的变化。我们还考虑了数据库概念的教学方面的趋势，并在适当的地方做出了推动这种趋势的修改。

本书的组织

本书组织成十个主要部分：

- **综述**（第 1 章）。第 1 章对数据库系统的性质和目标进行了一般性综述。我们解释了数据库系统的概念是如何发展的，各数据库系统的共同特性是什么，数据库系统能为用户做什么，以及数据库系统如何与操作系统交互。我们还引入了一个数据库应用的例子：包括多个系、教员、学生和课程的一个大学机构。这个应用作为贯穿全书的运行实例。这一章本质上是诱导性、历史性和解释性的。
- **第一部分：关系数据库**（第 2 章至第 6 章）。第 2 章介绍了数据的关系模型，包括基本概念，诸如关系数据库的结构、数据库模式、码、模式图、关系查询语言和关系操作等。第 3~5 章主要介绍最具影响力的面向用户的关系语言——SQL。第 6 章介绍形式化的关系查询语言，包括关系代数、元组关系演算和域关系演算。

这部分描述了数据操纵，包括查询、修改、插入和删除（假设已有一个模式设计）。关于模式设计的问题延迟到第二部分讲述。

- **第二部分：数据库设计**（第 7 章至第 9 章）。第 7 章给出了数据库设计过程的概要介绍，主要

侧重于用实体 - 联系数据模型来进行数据库设计。实体 - 联系模型为数据库设计问题, 以及我们在数据模型的约束下捕获现实应用的语义时所遇到的问题提供了一个高层视图。UML 类图表示也在这一章中讲述。

第 8 章介绍关系数据库设计理论。这一章讲述函数依赖和规范化, 重点强调提出各种范式的动机, 以及它们的直观含义。这一章以关系设计的概览开始, 依赖于对函数依赖的逻辑蕴涵的直观理解。这使得规范化的概念可以在函数依赖理论的完整内容之前先作介绍。函数依赖理论将在本章中稍后部分讨论。教师可以只选用 8.1 节至 8.3 节这些较前面的章节, 而不会丢失连贯性。不过, 完整地讲授这一章将有利于学生对规范化概念形成较好的理解, 从而诱导出函数依赖理论中一些较艰深的概念。

第 9 章讲述应用设计和开发。这一章侧重于用基于 Web 的界面构建数据库应用。另外, 这一章还讲述了应用安全性。

- **第三部分: 数据存储和查询** (第 10 章至第 13 章)。第 10 章讨论存储设备、文件和数据存储结构。在第 11 章中介绍多种数据存取技术, 包括 B⁺ 树索引和散列。第 12 章和第 13 章阐述查询执行算法和查询优化。这两章使用户能更好地理解数据库的存储和检索的内部机制。
- **第四部分: 事务管理** (第 14 章至第 16 章)。第 14 章着重介绍事务处理系统的基本概念, 包括原子性、一致性、隔离性和持久性。它还提供了用于保证这些特性的方法的一个概述, 包括封锁和快照隔离性。

第 15 章重点讲述并发控制, 并介绍保证可串行化的几种技术, 包括封锁、时间戳和乐观 (有效性检查) 技术。在这一章中还讨论死锁问题, 并介绍保证可串行化的其他方法, 特别是详细讨论广泛使用的快照隔离方法。

第 16 章讨论在系统崩溃和存储器故障情况下保证事务正确执行的主要技术。这些技术包括日志、检查点和数据库转储。被广泛使用的 ARIES 算法也在这里做了介绍。

- **第五部分: 系统体系结构** (第 17 章至第 19 章)。第 17 章介绍计算机系统体系结构, 并描述了作为基础的计算机系统对于数据库系统的影响。在这一章中讨论了集中式系统、客户 - 服务器系统、并行和分布式体系结构。

在第 18 章关于并行数据库的讨论中, 我们探讨了各种并行技术, 包括 I/O 并行、查询间并行和查询内并行, 以及操作间并行和操作内并行。这一章中还讨论了并行系统设计。

第 19 章讨论分布式数据库系统, 在分布式数据库系统的环境下重新讨论数据库设计、事务管理、查询执行和优化问题。这一章还包括了故障时的系统可用性问题, 并介绍了异构分布式数据库、基于云的数据库和分布式目录系统。

- **第六部分: 数据仓库、数据挖掘与信息检索** (第 20 章和第 21 章)。第 20 章介绍数据仓库和数据挖掘的概念。第 21 章描述用于查询文本数据的信息检索技术, 包括在 Web 搜索引擎中使用的基于超链接的技术。

第六部分使用了第一部分和第二部分的模型和语言概念, 但并不依赖于第三部分、第四部分或第五部分。因此它可以很容易地结合到侧重于 SQL 和数据库设计的课程中。

- **第七部分: 特种数据库** (第 22 章和第 23 章)。第 22 章介绍基于对象的数据库。该章讲述了对对象 - 关系数据模型, 该模型扩展了关系数据模型以支持复杂数据类型、类型继承和对象标识。该章还描述了用面向对象的编程语言来访问数据库。

第 23 章介绍数据表示的 XML 标准, 它正日益广泛地应用于复杂数据交换和存储。这一章

还描述了 XML 的查询语言。

- **第八部分：高级主题**（第 24 章至第 26 章）。第 24 章讨论应用开发中的高级话题，包括性能调整、性能评测标准、数据库应用测试和标准化。

第 25 章介绍空间和地理数据、时间数据、多媒体数据以及移动和个人数据库管理中的问题。

最后，第 26 章讨论高级事务处理。这一章的内容包括事务处理监控器、事务 workflow、电子商务、高性能事务系统、实时事务系统和持续长时间的事务。

- **第九部分：实例研究**（第 27 章至第 30 章）。在这一部分我们对四个领先的数据库系统进行实例研究，包括 PostgreSQL、Oracle、IBM DB2 和 Microsoft SQL Server。这几章中列举了上述每一种系统的独有特性，描述了它们的内部结构，提供了关于各个产品的丰富的有用信息，帮助读者了解前面各部分描述的各种实现技术是如何使用到实际系统中的。这几章中还包括实际系统设计中的几个有趣的方面。

- **第十部分：附录**（附录 A ~ 附录 E）。我们提供 5 个附录，包括一些历史性的和高级的内容；这些附录只在本书的 Web 站点（<http://www.db-book.com>）中联机提供。只有附录 A（详细的大学模式）例外，它给出了我们的大学模式的细节，包括完整的模式、DDL 和所有的表。这个附录出现在纸质版本中。

附录 B（高级关系数据库设计）描述了高级关系数据库设计，包括多值依赖理论、连接依赖、投影连接和域 - 码范式。这个附录是为希望更详细地研究关系数据库设计理论的读者，以及希望在课程中这样做的教师准备的。这个附录同样只是联机提供，就在本书的网站上。

附录 C（其他关系查询语言）描述其他的关系查询语言，包括 QBE Microsoft Access 和 Datalog。

虽然大多数新的数据库应用系统使用关系模型或对象 - 关系模型，但网状的和层次的数据模型在一些遗留应用中也仍然在使用。为了满足希望了解这些数据模型的读者的需要，我们给出了描述网状和层次的数据模型的附录，分别为附录 D（网状模型）和附录 E（层次模型）。

第 6 版

对于本书第 6 版的产生起指导作用的包括我们收到的关于前面几版的许多意见和建议，我们在耶鲁大学、利哈依大学、孟买印度理工学院讲授本课程的体会，以及我们对于数据库技术发展方向的分析。

我们用一个大学的例子替换了原先的银行企业的运行例子。这个例子与学生相关，不仅有助于他们记住这个例子，而且更重要的是，使他们能够更深入地洞察所需进行的各种设计决策。

我们重新组织了这本书，将我们关于 SQL 的介绍都集中在一起，放在本书的开头部分。第 3 章、第 4 章和第 5 章对 SQL 进行完整的介绍。第 3 章给出该语言的基础，而将更高级的特性放在了第 4 章。在第 5 章中，我们介绍从通用的程序设计语言访问 SQL 的 JDBC 和其他方法。我们介绍触发器和递归，最后讨论联机分析处理（OLAP）。入门性的课程可以选择只包括第 5 章的某些节，或者将一些节延迟到讲过了数据库设计之后再讲，这样也不失连续性。

除了这两个主要的改变外，我们修订了每一章中的素材，对旧的素材进行更新，增加对于数据库技术的新的发展的讨论，并且改进学生难于理解的那些内容的描述。我们还增加了新的练习并更新了参考文献。具体的改变如下。

- **较早介绍了 SQL**。许多讲课教师使用 SQL 作为课程实习的核心成分（访问我们的网站 www.db-

book.com 以参考示例课程实习)。为了给学生充足的时间做课程实习, 尤其对于那些四学期制的大学和学院来说, 尽早教授 SQL 非常重要。有了这个认识, 我们在内容组织方面做了几个改动:

- 用新的一章 (第 2 章) 介绍关系模型, 放在 SQL 之前, 打下概念基础, 而不使学生迷失在关系代数的细节中。
- 第 3 章、第 4 章和第 5 章详细介绍 SQL。这几章还讨论不同的数据库系统支持的变体, 从而当学生在实际的数据库系统上执行查询时会少遇到一些问题。这几章涵盖了 SQL 的所有方面, 包括查询、数据定义、约束说明、OLAP 和从各种语言 (包括 Java/JDBC) 内部使用 SQL。
- 形式语言 (第 6 章) 推迟到 SQL 之后, 而且可以略掉不讲, 而不会影响其他章的顺序。只有在第 13 章关于查询优化的讨论依赖于第 6 章介绍的关系代数。
- **新的数据库模式。**我们采用了一个新的模式作为贯穿全书的运行实例, 基于大学的数据。对于学生来说, 这个模式比以前的银行模式更加直观和有意义, 而且在数据库设计的章节中展示了更复杂的设计权衡。
- **对学生的亲手实践提供了更多的支持。**为了推动学生跟着我们的运行实例进行实践, 我们在附录 A 中列出了大学数据库的数据库模式和样例的关系实例, 并在用到大学实例的各个章中也进行了说明。另外, 在我们的 Web 站点 <http://www.db-book.com> 上, 提供了对于整个例子的 SQL 数据定义语句, 以及建立我们的样例关系实例的 SQL 语句。这鼓励学生直接在一个实际的数据库系统里去运行样例查询, 并且试着去修改这些查询。
- **修订了对 E-R 模型的涵盖。**我们修改了第 7 章中 E-R 图的符号表示, 以使得它与 UML 更兼容。这一章还充分利用了新的大学数据库模式来描述更复杂的设计权衡。
- **修订了对关系设计的涵盖。**第 8 章的风格现在更具可读性, 它在阐述函数依赖理论之前提供了对函数依赖和规范化的一个直观的理解, 其结果是理论的动机更清晰了。
- **扩充了关于应用开发和安全性的素材。**第 9 章包括了关于应用开发的新的素材, 以反映该领域的快速变化。特别地, 扩展了关于安全性的介绍, 包括它在当今紧密互联的世界中的至关重要性, 并重点强调了抽象概念外的实际问题。
- **修订和更新了对数据存储、索引和查询优化的涵盖。**对第 10 章进行了新技术方面的更新, 包括对闪存的更多的讨论。
更新了第 11 章对 B⁺ 树的讨论, 以反映实际的实现, 包括对批量载入的介绍, 并且对陈述方式也进行了改进。第 11 章中 B⁺ 树的例子修改成了 $n = 4$, 以避免 (不切实际的) $n = 3$ 时发生的空结点的特殊情况。
第 13 章中有关于高级查询优化技术的新的素材。
- **修订了对事务管理的涵盖。**第 14 章全面涵盖了入门性课程中的基础知识, 而高级的细节在第 15 章和第 16 章中介绍。我们扩展了第 14 章, 以涵盖数据库用户和数据库应用开发人员所面对的事务管理的实际问题。这一章还包括了对第 15 章和第 16 章所讨论话题的扩展的综述, 以确保即使略掉第 15 章和第 16 章, 学生仍能对并发控制和恢复的概念有基本的了解。
第 14 章和第 15 章现在包括了对当今被广泛支持和使用的快照隔离性的详细介绍, 对于使用它的潜在风险也做了介绍。
第 16 章现在对基本的基于日志的恢复进行了简化的描述, 进一步介绍了 ARIES 算法。

- **修订和扩充了对分布式数据库的涵盖。**我们现在涵盖了云数据存储，它在商业应用中正日益受到极大的重视。云存储为企业提供了改进成本管理和增加存储可扩充性的机遇，特别是对于基于 Web 的应用。我们分析了这些优点，也指出了可能的缺点和风险。

我们原来在高级事务处理这一章中讨论多数据库，现在对它更早地进行了介绍，作为分布式数据库这一章中的一部分。

- **推迟了对对象数据库和 XML 的介绍。**虽然面向对象的语言和 XML 在数据库的范围之外被广泛使用，但在数据库中的使用仍然有限，这使得它们对于高级课程或作为入门性课程的补充材料更适合一些。因此这些话题在本书中向后移了，放到了第 22 章和第 23 章。
- **QBE、Microsoft Access 和 Datalog 放在联机附录中。**这些话题原来是“其他关系语言”这一章中的一部分，现在放在联机附录 C 中了。

上面没有列出的所有话题都在第 5 版的基础上做了更新，虽然相对地它们的总体组织没有改变。

回顾材料和习题

每一章都有一个术语回顾表，还有一个总结，这可以帮助读者回顾这一章中涵盖的关键话题。

练习划分成两部分：**实践习题**（practice exercise）和**习题**（exercise）。实践习题的解答在本书的 Web 站点公开发布。我们鼓励学生独立解决这些实践习题，然后用 Web 站点上的解答来检查自己的答案。其他练习题的解答只有讲课教师能得到（参看下面的“讲课教师注意事项”以获取如何得到题解的信息）。

许多章的末尾有一个“工具”节，它提供与该章的话题相关的软件工具的信息；这些工具中的一些可以用于实验室练习。大学数据库和习题中用到的其他关系的 SQL DDL 和样例数据在本书的 Web 站点可以得到，也可以用于实验室练习。

讲课教师注意事项

本书包括基本内容和高级内容，在一个学期内也许不能讲授所有这些内容。我们以符号“*”把某些节标记为高级内容，如果愿意的话可以省略这些节，而仍能保持内容的连续性。较难的（可以忽略的）练习同样用符号“*”标记出来了。

可以用本书各章的不同子集来设计课程。某些章也可以用不同于本书中的顺序来讲授。我们列出一些可能的安排如下：

- 第 5 章可以跳过或推迟到后面讲，而不失连续性。我们期望大多数的课程至少 5.1.1 节较早地讲授，因为 JDBC 很可能会是学生实习的一个有用工具。
- 第 6 章可以在第 2 章之后、讲 SQL 之前讲。也可以把这一章从入门性的课程中省略掉。

我们建议如果课程中讲查询处理的话，那么把 6.1 节也包括在课程中。但是，如果学生在课程中不使用关系演算的话，那么 6.2 节和 6.3 节可以省略。

- 如果教师愿意的话，第 7 章可以在第 3 章、第 4 章和第 5 章之前讲，因为第 7 章完全不依赖于 SQL。
- 第 13 章可以从入门性的课程中省略掉，而对于其他任何章的讲授都没有影响。
- 我们对事务处理的讨论（第 14 章至第 16 章）和对系统体系结构的讨论（第 17 章至第 19 章）都包括一章综述（分别为第 14 章和第 17 章）和后续的两章详细讨论。如果计划把后面几章推

迟到高级课程中去讲授，可以选择使用第 14 章和第 17 章，省略第 15 章、第 16 章、第 18 章和第 19 章。

- 第 20 章和第 21 章介绍数据仓库、数据挖掘和信息检索，它们可以用作自学材料或从入门性课程中删去。
- 对于入门性课程来说，可以不讲第 22 章和第 23 章。
- 第 24 章至第 26 章涵盖高级应用开发，空间、时间和移动数据，以及高级事务处理，更适合于高级课程或学生自学。
- 第 27 章至第 30 章的实例研究适合学生自学。或者，这几章可以用作课堂上讲的前面那些章的概念的实例。

基于本书的样例课程提纲可以在本书的 Web 站点找到。

Web 站点和教学补充材料：

本书的 Web 站点的网址是：

<http://www.db-book.com>

该站点包括以下内容：

- 本书所有各章的幻灯片
- 实践习题的答案
- 五个附录
- 最新勘误表
- 实验素材，包括大学模式和习题中用到的其他关系的 SQL DDL 和样例数据，以及关于建立和使用各种数据库系统和工具的说明书

下列附加材料仅有教师可以获得：

- 包括书中所有习题的答案的教师手册
- 包括额外习题的问题库

关于如何得到教师手册和问题库的进一步信息，请发电子邮件到 customer.service@mcgraw-hill.com。McGraw-Hill 关于本书的网页是

<http://www.mhhe.com/silberschatz>

与我们联系

我们已尽了最大的努力来避免在本书中出现排版错误、内容失误等。然而，与发布新软件类似，错误在所难免；在本书的 Web 站点中有一个最新勘误表。如果你能指出尚未包含在最新勘误表中的本书的疏漏之处，我们将十分感激。

我们很高兴能收到你对改进本书的建议，我们也很欢迎你对本书 Web 站点做出对其他读者有用的贡献，例如程序设计练习、课程实习建议、联机实验室和指南以及讲课要点等。

请将电子邮件发到 db-book-authors@cs.yale.edu。其他来信请寄到 Avi Silberschatz, Department of Computer Science, Yale University, 51 Prospect Street, P. O. Box 208285, New Haven, CT 06520-8285 USA。

致谢

许多人对于第 6 版以及第 6 版所源自的前 5 版提供了帮助。

第 6 版

- Anastassia Ailamaki, Sailesh Krishnamurthy, Spiros Papadimitriou 和 Bianca Schroeder (Carnegie Mellon University), 他们撰写了第 27 章, 描述 PostgreSQL 数据库系统。
- Hakan Jakobsson (Oracle), 他撰写了第 28 章, 描述 Oracle 数据库系统。
- Sriram Padmanabhan (IBM), 他撰写了第 29 章, 描述 IBM DB2 数据库系统。
- Sameet Agarwal, José A. Blakeley, Thierry D'Hers, Gerald Hinson, Dirk Myers, Vaqar Pirzada, Bill Ramos, Balaji Rathakrishnan, Michael Rys, Florian Waas 和 Michael Zwilling (全部来自 Microsoft), 他们撰写了第 30 章, 描述 Microsoft SQL Server 数据库系统。特别地, José Blakeley 对这一章进行了协调和编辑; César Galindo-Legaria, Goetz Graefe, Kalen Delaney 和 Thomas Casey (全部来自 Microsoft) 对于前一版中的 Microsoft SQL Server 这一章做出了贡献。
- Daniel Abadi 审阅了第 5 版的目录表, 并协助进行了新的组织。
- Steve Dolins (University of Florida), Rolando Fernandez (George Washington University), Frantisek Franek (McMaster University), Latifur Khan (University of Texas-Dallas), Sanjay Madria (University of Missouri-Rolla), Aris Ouksel (University of Illinois) 和 Richard Snodgrass (University of Waterloo), 他们作为本书的审阅人, 他们的意见对于我们形成第 6 版帮助很大。
- Judi Paige, 她帮助形成了插图和演示幻灯片。
- Mark Wogahn, 他保证了生成本书文本的软件正常工作, 包括 LaTeX 宏和字体。
- N. L. Sarda, 他的反馈帮助我们改进了好几章, 特别是第 11 章; Vikram Pudi, 他促使我们替换掉原来的银行模式; Shetal Shah, 他对几个章节给出了反馈。
- 耶鲁大学和印度孟买理工学院的学生, 他们对第 5 版, 以及第 6 版的预印本给出了意见。

前面各版

- Chen Li, Sharad Mehrotra, 他们为第 5 版提供了关于 JDBC 和安全性的素材。
- Marilyn Turnamian 和 Nandprasad Joshi 为第 5 版担任秘书工作, Marilyn 还为第 5 版准备了一个封面设计的早期版本。
- Lyn Dupré 对第 3 版进行了审查, Sara Strandman 对第 3 版进行了文字编辑。
- Nilesh Dalvi, Sumit Sanghai, Gaurav Bhalotia, Arvind Hulgeri K. V. Raghavan, Prateek Kapadia, Sara Strandman, Greg Speegle 和 Dawn Bezviner 协助准备了前面各版的教师素材。
- 用船作为封面概念的一部分的想法最初是 Bruce Stephan 向我们建议的。
- 以下人士指出了第 5 版中的错误之处: Alex Coman, Ravindra Guravannavar, Arvind Hulgeri, Rohit Kulshreshtha, Sang-Won Lee, Joe H. C. Lu, Alex N. Napitupulu, H. K. Park, Jian Pei, Fernando Saenz Perez, Donnie Pinkston, Yma Pinto, Rajarshi Rakshit, Sandeep Satpal, Amon Seagull, Barry Soroka, Praveen Ranjan Srivastava, Hans Svensson, Moritz Wiese 和 Eyob Delele Yirdaw。
- 以下人士对本书第 5 版和前面各版提出了建议和意见: R. B. Abhyankar, Hani Abu-Salem, Jamel R. Alsabbagh, Raj Ashar, Don Batory, Phil Bernhard, Christian Breimann, Gavin M. Bierman, Janek Bogucki, Haran Boral, Paul Bourgeois, Phil Bohannon, Robert Brazile, Yuri Breitbart, Ramzi Bualuan, Michael Carey, Soumen Chakrabarti, Tom Chappell, Zhengxin Chen, Y. C. Chin, Jan Chomici, Laurens Damen, Prasanna Dhandapani, Qin Ding, Valentin Dinu, J. Edwards, Christos Faloutsos, Homma Farian, Alan Fekete, Frantisek Franek, Shashi Gadia, Hector Garcia-Molina, Goetz Graefe, Jim Gray, Le Gruenwald, Eitan M. Gurari, William Hankley, Bruce Hillyer, Ron Hitchens,

Chad Hogg, Arvind Hulgeri, Yannis Ioannidis, Zheng Jiaping, Randy M. Kaplan, Graham J. L. Kemp, Rami Khouri, Hyoung-Joo Kim, Won Kim, Henry Korth (Henry F. 的父亲), Carol Kroll, Hae Choon Lee, Sang-Won Lee, Irwin Levinstein, Mark Llewellyn, Gary Lindstrom, Ling Liu, Dave Maier, Keith Marzullo, Marty Maskarinec, Fletcher Mattox, Sharad Mehrotra, Jim Melton, Alberto Mendelzon, Ami Motro, Bhagirath Narahari, Yiu-Kai Dennis Ng, Thanh-Duy Nguyen, Anil Nigam, Cyril Orji, Meral Ozsoyoglu, D. B. Phatak, Juan Altmayer Pizzorno, Bruce Porter, Sunil Prabhakar, Jim Peterson, K. V. Raghavan, Nahid Rahman, Rajarshi Rakshit, Krithi Ramamritham, Mike Reiter, Greg Riccardi, Odinaldo Rodriguez, Mark Roth, Marek Rusinkiewicz, Michael Rys, Sunita Sarawagi, N. L. Sarda, Patrick Schmid, Nikhil Sethi, S. Seshadri, Stewart Shen, Shashi Shekhar, Amit Sheth, Max Smolens, Nandit Soparkar, Greg Speegle, Jeff Storey, Dilys Thomas, Prem Thomas, Tim Wahls, Anita Whitehall, Christopher Wilson, Marianne Winslett, Weining Zhang 和 Liu Zhenming。

本书的制作出版

出版人是 Raghu Srinivasan。策划编辑是 Melinda D. Bilecki。项目经理是 Melissa Leick。市场经理是 Curt Reynolds。作品监督人是 Laura Fuller。设计者是 Brenda Rolwes。封面设计是 Studio Montage, St. Louis, Missouri。文字编辑是 George Watson。校对是 Kevin Campbell。特约索引制作者是 Tobiah Waldron。Aptara team 由 Raman Arora 和 Sudeshna Nandy 组成。

个人注记

Sudarshan 感谢他的妻子 Sita 的爱和支持, 感谢他的孩子 Madhur 和 Advaith 的爱和乐观精神。Hank 感谢他的妻子 Joan, 孩子 Abby 和 Joe 的爱和理解。Avi 感谢 Valerie 在本书修订期间的爱、耐心和支持。

A. S.

H. F. K.

S. S.

出版者的话

译者序

前 言

第 1 章 引言	1
1.1 数据库系统的应用	1
1.2 数据库系统的目标	2
1.3 数据视图	4
1.3.1 数据抽象	4
1.3.2 实例和模式	5
1.3.3 数据模型	5
1.4 数据库语言	6
1.4.1 数据操纵语言	6
1.4.2 数据定义语言	6
1.5 关系数据库	7
1.5.1 表	7
1.5.2 数据操纵语言	8
1.5.3 数据定义语言	8
1.5.4 来自应用程序的数据库访问	8
1.6 数据库设计	9
1.6.1 设计过程	9
1.6.2 大学机构的数据库设计	9
1.6.3 实体-联系模型	10
1.6.4 规范化	11
1.7 数据存储和查询	11
1.7.1 存储管理器	12
1.7.2 查询处理器	12
1.8 事务管理	12
1.9 数据库体系结构	13
1.10 数据挖掘与信息检索	15

1.11 特种数据库	15
1.11.1 基于对象的数据模型	15
1.11.2 半结构化数据模型	16
1.12 数据库用户和管理员	16
1.12.1 数据库用户和用户界面	16
1.12.2 数据库管理员	16
1.13 数据库系统的历史	17
1.14 总结	18
术语回顾	19
实践习题	19
习题	19
工具	20
文献注解	20

第一部分 关系数据库

第 2 章 关系模型介绍	22
2.1 关系数据库的结构	22
2.2 数据库模式	23
2.3 码	24
2.4 模式图	25
2.5 关系查询语言	26
2.6 关系运算	27
2.7 总结	28
术语回顾	29
实践习题	29
习题	30
文献注解	30
第 3 章 SQL	31
3.1 SQL 查询语言概览	31

3.2 SQL 数据定义	32	实践习题	58
3.2.1 基本类型	32	习题	60
3.2.2 基本模式定义	32	工具	61
3.3 SQL 查询的基本结构	34	文献注解	62
3.3.1 单关系查询	34		
3.3.2 多关系查询	35		
3.3.3 自然连接	38		
3.4 附加的基本运算	40	第 4 章 中级 SQL	63
3.4.1 更名运算	40	4.1 连接表达式	63
3.4.2 字符串运算	41	4.1.1 连接条件	63
3.4.3 select 子句中的属性说明	42	4.1.2 外连接	64
3.4.4 排列元组的显示次序	42	4.1.3 连接类型和条件	67
3.4.5 where 子句谓词	42	4.2 视图	67
3.5 集合运算	43	4.2.1 视图定义	68
3.5.1 并运算	43	4.2.2 SQL 查询中使用视图	68
3.5.2 交运算	44	4.2.3 物化视图	69
3.5.3 差运算	44	4.2.4 视图更新	69
3.6 空值	45	4.3 事务	71
3.7 聚集函数	46	4.4 完整性约束	71
3.7.1 基本聚集	46	4.4.1 单个关系上的约束	72
3.7.2 分组聚集	46	4.4.2 not null 约束	72
3.7.3 having 子句	47	4.4.3 unique 约束	72
3.7.4 对空值和布尔值的聚集	48	4.4.4 check 子句	72
3.8 嵌套子查询	48	4.4.5 参照完整性	73
3.8.1 集合成员资格	49	4.4.6 事务中对完整性约束的 违反	75
3.8.2 集合的比较	49	4.4.7 复杂 check 条件与断言	75
3.8.3 空关系测试	50	4.5 SQL 的数据类型与模式	76
3.8.4 重复元组存在性测试	51	4.5.1 SQL 中的日期和时间类型	76
3.8.5 from 子句中的子查询	52	4.5.2 默认值	77
3.8.6 with 子句	53	4.5.3 创建索引	77
3.8.7 标量子查询	53	4.5.4 大对象类型	77
3.9 数据库的修改	54	4.5.5 用户定义的类型	78
3.9.1 删除	54	4.5.6 create table 的扩展	79
3.9.2 插入	55	4.5.7 模式、目录与环境	80
3.9.3 更新	56	4.6 授权	80
3.10 总结	57	4.6.1 权限的授予与收回	81
术语回顾	58	4.6.2 角色	82
		4.6.3 视图的授权	82

4.6.4 模式的授权	83	习题	120
4.6.5 权限的转移	83	工具	121
4.6.6 权限的收回	84	文献注解	122
4.7 总结	85		
术语回顾	85		
实践习题	85		
习题	87		
文献注解	87		
第5章 高级 SQL	88	第6章 形式化关系查询语言	123
5.1 使用程序设计语言访问数据库	88	6.1 关系代数	123
5.1.1 JDBC	89	6.1.1 基本运算	123
5.1.2 ODBC	93	6.1.2 关系代数的形式化定义	128
5.1.3 嵌入式 SQL	95	6.1.3 附加的关系代数运算	129
5.2 函数和过程	97	6.1.4 扩展的关系代数运算	132
5.2.1 声明和调用 SQL 函数和 过程	98	6.2 元组关系演算	135
5.2.2 支持过程和函数的语言 构造	99	6.2.1 查询示例	135
5.2.3 外部语言过程	101	6.2.2 形式化定义	137
5.3 触发器	102	6.2.3 表达式的安全性	137
5.3.1 对触发器的需求	102	6.2.4 语言的表达能力	138
5.3.2 SQL 中的触发器	102	6.3 域关系演算	138
5.3.3 何时不用触发器	105	6.3.1 形式化定义	138
5.4 递归查询**	106	6.3.2 查询的例子	138
5.4.1 用迭代来计算传递闭包	106	6.3.3 表达式的安全性	139
5.4.2 SQL 中的递归	107	6.3.4 语言的表达能力	140
5.5 高级聚集特性**	109	6.4 总结	140
5.5.1 排名	109	术语回顾	140
5.5.2 分窗	110	实践习题	140
5.6 OLAP**	112	习题	142
5.6.1 联机分析处理	112	文献注解	143
5.6.2 交叉表与关系表	114		
5.6.3 SQL 中的 OLAP	116		
5.7 总结	118		
术语回顾	118		
实践习题	119		
		第二部分 数据库设计	
		第7章 数据库设计和 E-R 模型	146
		7.1 设计过程概览	146
		7.1.1 设计阶段	146
		7.1.2 设计选择	147
		7.2 实体-联系模型	148
		7.2.1 实体集	148
		7.2.2 联系集	148
		7.2.3 属性	150

7.3 约束	151	7.10.1 数据约束和关系数据库 设计	174
7.3.1 映射基数	152	7.10.2 使用需求: 查询、性能	175
7.3.2 参与约束	152	7.10.3 授权需求	175
7.3.3 码	153	7.10.4 数据流、工作流	175
7.4 从实体集中删除冗余属性	153	7.10.5 数据库设计的其他问题	176
7.5 实体-联系图	155	7.11 总结	176
7.5.1 基本结构	155	术语回顾	177
7.5.2 映射基数	156	实践习题	177
7.5.3 复杂的属性	156	习题	179
7.5.4 角色	157	工具	180
7.5.5 非二元的联系集	157	文献注解	180
7.5.6 弱实体集	157		
7.5.7 大学的 E-R 图	159		
7.6 转换为关系模式	159	第 8 章 关系数据库设计	181
7.6.1 具有简单属性的强实体集 的表示	160	8.1 好的关系设计的特点	181
7.6.2 具有复杂属性的强实体集 的表示	160	8.1.1 设计选择: 更大的模式	181
7.6.3 弱实体集的表示	161	8.1.2 设计选择: 更小的模式	182
7.6.4 联系集的表示	161	8.2 原子域和第一范式	183
7.7 实体-联系设计问题	163	8.3 使用函数依赖进行分解	184
7.7.1 用实体集还是用属性	163	8.3.1 码和函数依赖	185
7.7.2 用实体集还是用联系集	164	8.3.2 Boyce-Codd 范式	186
7.7.3 二元还是 n 元联系集	165	8.3.3 BCNF 和保持依赖	187
7.7.4 联系属性的布局	165	8.3.4 第三范式	188
7.8 扩展的 E-R 特性	166	8.3.5 更高的范式	189
7.8.1 特化	166	8.4 函数依赖理论	189
7.8.2 概化	167	8.4.1 函数依赖集的闭包	189
7.8.3 属性继承	168	8.4.2 属性集的闭包	191
7.8.4 概化上的约束	168	8.4.3 正则覆盖	191
7.8.5 聚集	169	8.4.4 无损分解	193
7.8.6 转换为关系模式	170	8.4.5 保持依赖	194
7.9 数据建模的其他表示法	171	8.5 分解算法	195
7.9.1 E-R 图的其他表示法	171	8.5.1 BCNF 分解	195
7.9.2 统一建模语言 UML	173	8.5.2 3NF 分解	197
7.10 数据库设计的其他方面	174	8.5.3 3NF 算法的正确性	198
		8.5.4 BCNF 和 3NF 的比较	198
		8.6 使用多值依赖的分解	199