



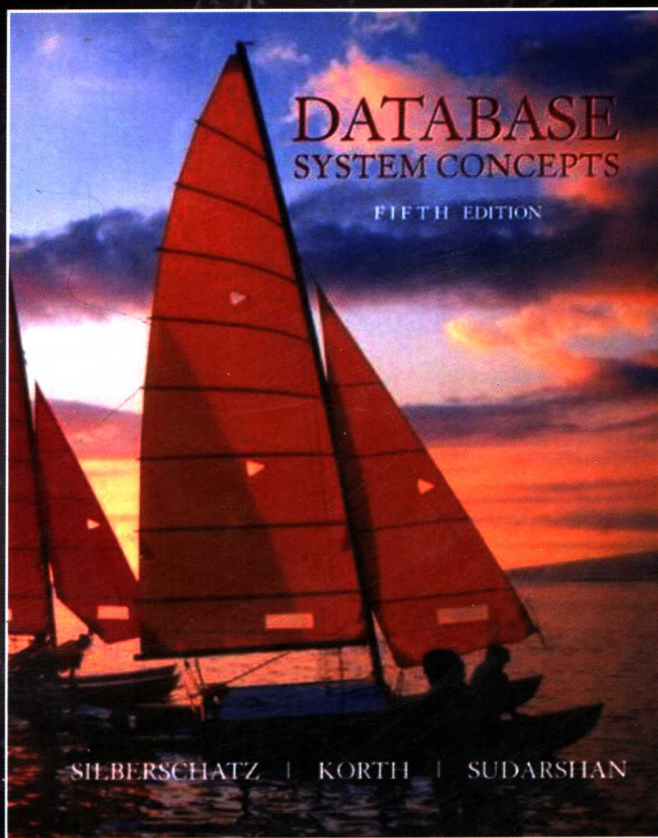
Mc
Graw
Hill

计 算 机 科 学 丛 书

原书第5版

数据库系统概念

Abraham Silberschatz Henry F. Korth S. Sudarshan 著 杨冬青 马秀莉 唐世渭 等译
耶 鲁 大 学 里 海 大 学 印 度 理 工 学 院 北 京 大 学



Database System Concepts
Fifth Edition



机械工业出版社
China Machine Press

计 算 机

丛 书

原书第5版

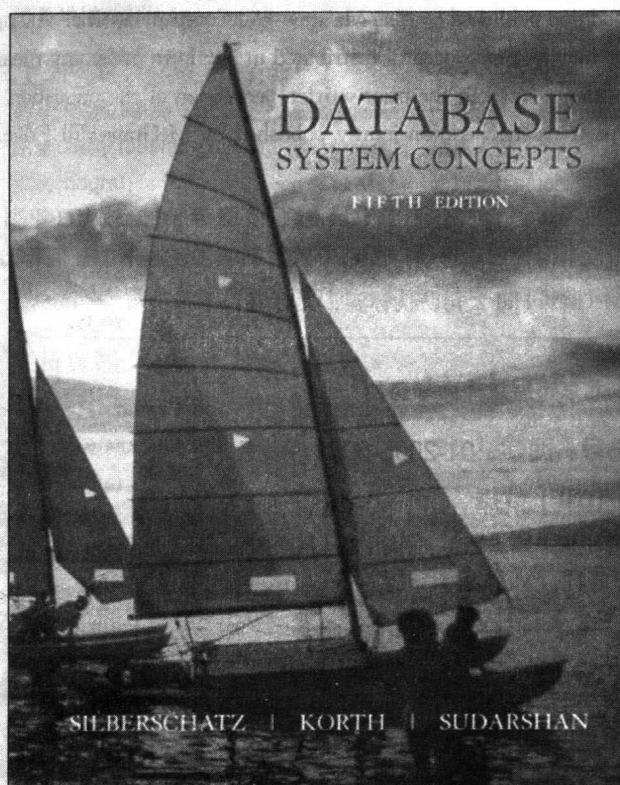
TP311.13

33=4

2006

数据库系统概念

Abraham Silberschatz Henry F. Korth S. Sudarshan 著 杨冬青 马秀莉 唐世渭 等译
耶 鲁 大 学 里 海 大 学 印 度 理 工 学 院 北 京 大 学



Database System Concepts

Fifth 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,不仅讨论了数据库查询语言、模式设计、数据仓库、数据库应用开发、基于对象的数据库和 XML、数据存储和查询、事务管理、数据挖掘与信息检索以及数据库系统体系结构等方面的内容,而且对性能评测标准、性能调整、标准化以及空间与地理数据、事务处理监控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, Fifth Edition (ISBN 0-07-295886-3).

Copyright © 2006, 2002, 1999, 1991, 1986 by The McGraw-Hill Companies, Inc.

Original English edition published by The McGraw-Hill Companies, Inc. All rights reserved. No part of this publication may be reproduced or distributed in any form or by any means, or stored in a database or retrieval system, without the prior written permission of the publisher.

Simplified Chinese translation edition jointly published by McGraw-Hill Education (Asia) Co. and China Machine Press.

本书中文简体字翻译版由机械工业出版社和美国麦格劳-希尔教育(亚洲)出版公司合作出版。未经出版者预先书面许可,不得以任何方式复制或抄袭本书的任何部分。

本书封面贴有 McGraw-Hill 公司防伪标签,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:01-2005-6031

图书在版编目(CIP)数据

数据库系统概念(原书第5版)/西尔伯沙茨(Silberschatz, A.)等著;杨冬青等译.
—北京:机械工业出版社,2006.9

(计算机科学丛书)

书名原文:Database System Concepts, Fifth Edition

ISBN 7-111-19687-2

I. 数… II. ①西… ②杨… III. 数据库系统 IV. TP311.13

中国版本图书馆 CIP 数据核字(2006)第 087537 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 玉 刘立卿

北京瑞德印刷有限公司印刷·新华书店北京发行所发行

2006 年 10 月第 1 版第 1 次印刷

184mm×260mm·50.25 印张

定价:69.50 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

出版者的话

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，还揭开了学术的源变，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、从业人员较少的现状下，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图文信息有限公司较早意识到“出版要为教育服务”。自1998年开始，华章公司就将工作重点放在了遴选、移译国外优秀教材上。经过几年的不懈努力，我们与Prentice Hall, Addison-Wesley, McGraw-Hill, Morgan Kaufmann等世界著名出版公司建立了良好的合作关系，从它们现有的数百种教材中甄选出Tanenbaum, Stroustrup, Kernighan, Jim Gray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、研究及度藏。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，为进一步推广与发展打下了坚实的基础。

随着学科建设的初步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步入一个新的阶段。为此，华章公司将加大引进教材的力度，在“华章教育”的总规划之下出版三个系列的计算机教材：除“计算机科学丛书”之外，对影印版的教材，则单独开辟出“经典原版书库”；同时，引进全美通行的教学辅导书“Schaum's Outlines”系列组成“全美经典学习指导系列”。为了保证这三套丛书的权威性，同时也为了更好地为学校和老师服务，华章公司聘请了中国科学院、北京大学、清华大学、国防科技大学、复旦大学、上海交通大学、南京大学、浙江大学、中国科技大学、哈尔滨工业大学、西安交通大学、中国人民大学、北京航空航天大学、北京邮电大学、中山大学、解放军理工大学、郑州大学、湖北工学院、中国国家信息安全测评认证中心等国内重点大学和科研机构在计算机的各个领域的著名学者组成“专家指导委员会”，为我们提供选题意见和出版监督。

这三套丛书是响应教育部提出的使用外版教材的号召，为国内高校的计算机及相关专业

的教学度身订造的。其中许多教材均已为M. I. T., Stanford, U.C. Berkeley, C. M. U. 等世界名牌大学所采用。不仅涵盖了程序设计、数据结构、操作系统、计算机体系结构、数据库、编译原理、软件工程、图形学、通信与网络、离散数学等国内大学计算机专业普遍开设的核心课程,而且各具特色——有的出自语言设计者之手、有的历经三十年而不衰、有的已被全世界的几百所高校采用。在这些圆熟通博的名师大作的指引之下,读者必将在计算机科学的宫殿中由登堂而入室。

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,但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。教材的出版只是我们的后续服务的起点。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:

电子邮件: hzjsj@hzbook.com

联系电话: (010) 68995264

联系地址: 北京市西城区百万庄南街1号

邮政编码: 100037

专家指导委员会

(按姓氏笔画顺序)

尤晋元
石教英
张立昂
邵维忠
周克定
郑国梁
高传善
裘宗燕

王 珊
吕 建
李伟琴
陆丽娜
周傲英
施伯乐
梅 宏
戴 葵

冯博琴
孙玉芳
李师贤
陆鑫达
孟小峰
钟玉琢
程 旭

史忠植
吴世忠
李建中
陈向群
岳丽华
唐世渭
程时端

史美林
吴时霖
杨冬青
周伯生
范 明
袁崇义
谢希仁

译者简介

杨冬青 1969年毕业于北京大学数学力学系数学专业,现任北京大学信息科学技术学院教授,博士生导师,网络与信息系统研究所副所长,数据库与信息系统研究室主任,中国计算机学会数据库专委会委员。多年来承担并完成973、863国家科技攻关、国家自然科学基金等多项国家重点科研项目;曾获国家科技进步二等奖、三等奖和多项省部级奖励;在国内外科技杂志及会议上发表论文百余篇,著译作十余部。目前的主要研究方向为数据库系统实现技术、Web环境下的信息集成与共享、数据仓库和数据挖掘、典型应用领域的数据库技术等。

马秀莉 博士,1972年出生,2003年毕业于北京大学信息科学技术学院,获理学博士学位。先后参加过多项国家自然科学基金项目、国家重点基础研究发展规划(973)课题等科研项目及典型应用领域的应用研究项目;在国内外科技杂志及国际会议发表论文近20篇。主要研究领域为数据库系统实现技术、数据仓库、联机分析处理、数据挖掘等。

唐世渭 1964年毕业于北京大学数学力学系计算数学专业,毕业后留校任教至今,现为北京大学信息科学技术学院教授,博士生导师,中国计算机学会数据库专委会副主任。多年来承担并完成973、863国家科技攻关、国家自然科学基金等多项国家重点科研项目;曾获国家科技进步二等奖、三等奖各1项;省部级科技进步奖多项;在国内外科技杂志及会议上发表论文百余篇,著译作多部。至今已培养硕士、博士、博士后60余名。目前,主要的研究方向为数据库系统、数据仓库和数据挖掘、Web环境下的信息集成与共享、典型应用领域的信息系统等。

译者序

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版和第4版译成中文，由机械工业出版社分别于2000年初和2003年初出版发行。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5版。第5版保持了前几版的总体风格，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5版的内容大体上可以分为四个部分。

第1~10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Web应用和安全问题等）的详细讨论。这一部分还介绍了新型的数据库系统——对象-关系数据库，以及XML数据库设计和查询的相关内容，包括对XML Schema和XQuery更深入的描述，以及SQL/XML标准的介绍。

第11~22章主要讨论数据库系统实现技术和数据管理与应用的深入话题，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。这一部分还包括了对数据挖掘、联机分析处理和信息检索的较详细的介绍。

第23~25章是一些高级话题，主要包括数据库性能调整 and 性能测试标准、遗产系统应用的转移、时间数据、空间和地理数据、多媒体数据、移动式和个人数据库管理等高级数据类型和新应用，以及事务处理监控器、事务工作流、电子商务、高性能事务系统、实时事务系统、长周期事务及多数据库系统中的事务管理等高级事务处理问题。

第26~29章对PostgreSQL、Oracle、IBM DB2和Microsoft SQL Server四个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。

授课老师可根据需要适当裁剪前三部分内容作为本科数据库概论课程的教材或主要参考资料，或用于研究生的数据库课程教学，第四部分可作为帮助学生了解实际系统的补充材料。

杨冬青、马秀莉、唐世渭组织并参加了本书的翻译和审校工作；参加翻译的还有陈冠华、帅猛、李双峰、李希婷、田枫、刘晓芳、周林宏、袁征、周沫等。

限于译者水平，译文中疏漏和错误难免，欢迎批评指正。

译者
2006年于北京大学

前 言

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组成部分，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，这些概念涉及数据库设计、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。本书不仅可以作为入门课程的基本教材，还可作为课程补充性或高级课程介绍性材料。

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描述，其中许多概念都基于银行运营的例子加以阐释。本书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都基于当今的商品化或试验性数据库系统中所采用的概念和算法。我们的目标是在一个通用环境下描述这些概念和算法，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 5 版中，我们保持了前面版本的总体风格，同时扩展了前几版的内容和结构来反映数据库设计、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。在详细描述本书的内容之前，我们先重点谈谈第 5 版中增加的一些特性。

- **较早介绍了 SQL。**许多讲课教师使用 SQL 作为课程项目的核心成分。为了给学生充足的时间做课程项目，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：

1. 把对实体-联系模型的介绍推迟到第二部分“数据库设计”中介绍。
2. 通过把关系演算推迟到第 5 章介绍，同时，在第 2 章中保留对关系代数的介绍，使对关系模型的介绍更加自然流畅。
3. 较早用两章专门介绍 SQL。第 3 章涵盖了 SQL 的基本特性，包括数据定义和操纵。第 4 章涵盖了一些高级特性，包括完整性约束、动态 SQL 和过程化结构。这章新的内容包括对 JDBC 介绍的扩展、SQL 中的过程化结构、SQL 中的递归以及 SQL: 2003 中的新特性。这章还包括对授权机制的概要介绍；对授权机制的详细介绍推迟到第 8 章进行。

这些改动使得学生在课程早期就可以开始写 SQL 查询，并且开始熟悉数据库系统的使用。这也可以使学生形成对数据库设计的直观印象，从而有利于本书第二部分关于设计方法学的教学。我们发现按这种组织结构讲解，学生更容易接受数据库设计的概念。

- **新的部分（第二部分）专门讨论数据库设计。**本书的第二部分包括三章内容，都致力于讨论数据库设计和数据库应用。其中第 6 章讨论实体-联系模型，该章包括了所有来自第 4 版相应章节（第 2 章）的材料，并做了一些重要的更新。我们还在第 6 章概要介绍了数据库设计的过程。愿意以 E-R 模型作为课程开始的教师可以从这章开始，这样也不会丢失连贯性，因为我们已经力求避免对任何较前章节的依赖（除第 1 章外）。

第 7 章是关于关系设计的，包括了第 4 版第 7 章中涵盖的材料，不过是以一种全新

的、更可读的方式进行介绍。在介绍用函数依赖、多值依赖和规范化进行设计的正式方法之前，我们先来自 E-R 模型的设计概念来建立对整个关系设计问题的直观概览。本章还包括了新的一节，主要讲述当前数据库设计的一些问题。

第二部分第 8 章是全新的一章，该章描述了数据库应用的设计和开发，包括 Web 应用、servlet、JSP、触发器以及一些安全问题。为了与日益增加的软件保护需要同步，关于安全问题的介绍比第 4 版扩展了很多。

- **彻底改写和更新了基于对象的数据库和 XML 的相关内容。**第三部分第 9 章关于对象数据库的讨论做了大量修改，该章重点讲解 SQL 的对象-关系特性，取代了第 4 版中用两章分别讨论面向对象数据库和对象-关系数据库。该章没有介绍学生比较熟悉的关于面向对象的介绍性材料，也没有介绍现已废止的 ODMG 标准的语法细节。然而，关于面向对象数据库的一些重要概念都保留了，包括关于在 Java 中增加持久性的 JDO 标准的材料。

第三部分还包括了关于 XML 数据设计和查询的一章（第 10 章），这是从第 4 版的对应章彻底改写得来的。这一章包括对 XML Schema 和 XQuery 的更深入的描述，以及对 SQL/XML 标准的介绍，并给出了包括 Web 服务在内的更多的 XML 应用实例。

- **重新组织了数据挖掘和信息检索的内容。**数据挖掘和联机分析处理是目前核心的数据库应用——已经不再仅仅是“高级主题”了。因此，我们把这些主题的内容组织到新的部分——第六部分，其中包括关于数据挖掘与分析的一章，以及关于信息检索的一章。
- **实例研究部分涵盖了 PostgreSQL。**PostgreSQL 是近年来非常受欢迎的开源数据库系统。除了作为可以在其上建立数据库应用的平台以外，在强调数据库内核的课程中，PostgreSQL 的源代码还可以用于研究和扩展。因此第九部分增加了关于 PostgreSQL 的实例研究，第 4 版中的三个实例研究（Oracle、IBM DB2 以及 Microsoft SQL Server）还保留，但进行了更新，以反映各个软件的最新版本。

上面没有列出的内容，包括事务处理（并发与恢复）、存储结构、查询处理以及分布式和并行数据库也都分别对第 4 版中的相应内容进行了更新，尽管它们的整体结构大体上没有变化。对第 5 章关于 QBE 的内容进行了修改，删除了与任何实际应用都不对应的聚集和更新的语法细节，但保留了 QBE 的核心概念。

本书的组织

本书组织成九个主要部分，以及三个附录（本书附录是在线内容，没有包含在纸质印刷版中。——编者注）：

- **引言（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。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。这一章还引入了一个数据库应用的例子：由多个分行组成的一个银行企业。这个例子用作贯穿全书的运行实例。
- **第一部分：关系数据库（第 2 章至第 5 章）。**第 2 章介绍了数据的关系模型，包括基本概念以及关系代数。该章还给出了对完整性约束的简要介绍。第 3 章和第 4 章重点介绍最具影响力的面向用户的关系语言——SQL。第 3 章给出了对 SQL 的基本介绍，而第 4 章则描述了 SQL 的一些更高级的特性，包括编程语言和支持 SQL 的数据库之间如何接口。第 5 章介绍其他的一些关系语言，包括关系演算、QBE 和 Datalog。

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 6 章至第 8 章）。第 6 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 类图表示也在这章讲述。

第 7 章介绍了关系数据库设计理论。这章讲述函数依赖和规范化，重点强调提出各种范式的原因，以及它们的直观含义。这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。在完整描述函数依赖理论之前先介绍了规范化概念，函数依赖理论在第 7 章稍后部分讨论。教师可以只选用 7.1 节至 7.3 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 8 章讲述了应用设计和开发。这章侧重于用基于 Web 的界面构建数据库应用。另外，这章还讲述了应用安全性。

- **第三部分：基于对象的数据库和 XML**（第 9 章和第 10 章）。第 9 章介绍基于对象的数据库。该章讲述了对象 - 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 10 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。

- **第四部分：数据存储和查询**（第 11 章至第 14 章）。第 11 章讨论磁盘、文件和文件系统结构。第 12 章介绍多种数据存取技术，包括散列和 B⁺ 树索引。第 13 章和第 14 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。
- **第五部分：事务管理**（第 15 章至第 17 章）。第 15 章着重介绍事务处理系统的基本概念，包括事务的原子性、一致性、孤立性和持久性，以及可串行化概念。

第 16 章重点讲述并发控制，并介绍了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。在这一章中还讨论了死锁问题。

第 17 章讨论了在系统崩溃和磁盘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。

- **第六部分：数据挖掘与信息检索**（第 18 章和第 19 章）。第 18 章介绍了数据仓库的概念以及解释了什么是数据挖掘和联机分析处理（OLAP），包括 SQL: 1999 中对数据挖掘和数据仓库的支持。第 19 章描述了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被结合到侧重于 SQL 和数据库设计的课程上。

- **第七部分：系统体系结构**（第 20 章至第 22 章）。第 20 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，以及网络类型。

第 21 章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。这一章还讨论了并行系统设计。

第 22 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 LDAP 目录系统。

- **第八部分：其他主题**（第 23 章至第 25 章）。第 23 章讨论了性能评测标准、性能调整、标准化和遗产系统应用的转移。

第 24 章讨论高级数据类型和新应用，包括时间数据、空间和地理数据、多媒体数据，以及移动式和个人数据库管理中的问题。

最后，第 25 章讨论高级事务处理，包括事务处理监控器、事务工作流、电子商务、高性能事务系统、实时事务系统、长周期事务以及多数据库系统中的事务管理。

- **第九部分：实例研究**（第 26 章至第 29 章）。这一部分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。这几章还包括了实际系统设计中的几个相关的方面。
- **联机附录**。虽然大多数新的数据库应用系统使用关系模型或面向对象模型，但网状的和层次的数据模型在一些遗产应用中也仍在使用。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的和层次的数据模型的附录，分别为附录 A 和附录 B。这两个附录只是联机提供 (<http://codex.cs.yale.edu/avi/db-book/>)。

附录 C 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。

第 5 版说明

对于本书第 5 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。

我们基本上是重写每一章，将较旧的内容修改为最新的内容，添加关于数据库技术当前进展的讨论，并改进学生难于理解的内容描述。现在每一章都有一个术语回顾，它可以帮助读者温习该章讨论的关键话题。我们还在大多数章后面增加了“工具”小节，提供与该章内容有关的软件工具的信息。我们还增加了新的练习以及更新了参考文献。

在第 5 版中我们把习题划分成两部分：实践习题和习题。实践习题的解决方案在本书的网页可以公开得到。我们鼓励学生独立解决这些实践习题，然后用网页上的解决方案来检查自己的答案。其他习题的解决方案只有讲课教师能得到（参见下面的“教师注意事项”）。

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。以“*”符号标记的节为高级内容，如果愿意的话可以省略这些节，略过这些节也能保持内容的连续性。较难的习题（可以忽略）同样用符号“*”标记了出来。

可以用本书各章的不同子集来设计课程。以下列出一些推荐的安排：

- 对于介绍性的课程，第 4 章的 4.6 节以后的章节可以忽略。
- 如果学生在本课程中不使用关系演算、QBE 或 Datalog，可以不讲第 5 章。
- 对于入门性课程来说，可以不讲第 9 章、第 10 章 (XML) 和第 14 章。
- 我们对事务处理的讨论（第 15 章至第 17 章）和对数据库系统体系结构的讨论（第 20 章

至第 22 章) 都包括一章综述(分别为第 15 章和第 20 章)和后续的两章详细讨论。如果你计划把后续的详细讨论推迟到高级课程中去讲授, 可以选择使用第 15 章和第 20 章, 省略第 16、17、21 和 22 章。

- 第 18 章和第 19 章涵盖了数据挖掘和信息检索, 它们可以作为自学材料或从介绍性课程中删去。
- 第 23 章至第 25 章更适合于高级课程, 或学生自学。
- 第 26 章至第 29 章的实例研究适合学生自学。

Web 主页和教学补充材料

本书的 WWW 主页的网址是:

<http://codex.cs.yale.edu/avi/db-book/>

该主页包括以下内容:

- 本书所有章节的幻灯片
- 实践习题的答案
- 实验材料
- 三个附录
- 最新勘误表
- 本书用户贡献的补充材料

下列附加材料只对教师提供:

- 包括书中所有习题的答案的教师手册
- 包括额外习题的习题库

关于如何得到教师手册的进一步信息, 请发电子邮件到 customer.service@mcgraw-hill.com (国内教师请联系麦格劳-希尔教育出版公司教师服务中心(北京), 联系方式见书后插页。——编者注)。在美国可以打电话 800-338-3987。McGraw-Hill 关于本书的网页是

<http://www.mhhe.com/engcs/compsci/silberschatz-bridge/index.mhtml>。

与我们和其他读者联系

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 WWW 主页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, 我们也很欢迎在本书主页上提供以下内容: 例如程序设计练习、课程实习建议、联机实验室、个别指导以及讲课要点等, 这些内容都能对其他读者很有帮助。

请将电子邮件发到 db-book@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

使用本书前 4 版的大量学生提供了很多有用的意见, 对于我们形成第 5 版帮助很大。此外, 许多人口头或书面地与我们关于此书交换过意见, 对本书提出过建议和评论。在此我们不可能一一列出, 我们谨对下列人员致以特别的感谢:

- DePaul University 的 Hani Abu-Salem, Grand Valley State University 的 Jamel R. Alsabbagh, Notre Dame University 的 Ramzi Bualuan, Omaha 的 University of Nebraska 的 Zhengxin Chen, SUNY Buffalo University 的 Jan Chomick, Harrisburg 的 Penn State University 的 Qin Ding, McMaster University 的 Frantisek Franek, Drexel University 的 Randy M. Kaplan, University of Central Florida 的 Mark Llewellyn, Western Illinois University 的 Marty Maskarinec, Brigham Young University 的 Yiu Kai Dennis Ng, Purdue University 的 Sunil Prabhakar, Old Dominion University 的 Stewart Shen, Foothill College 的 Anita Whitehall, University of Oregon 的 Christopher Wilson, San Antonio 的 University of Texas 的 Weining Zhang, 他们是本书的审阅者。他们的意见对于我们形成第 5 版帮助很大。
- Anastassia Ailamaki, Sailesh Krishnamurthy, Spiros Papadimitriou 和 Bianca Schroeder (Carnegie Mellon University), 感谢他们编写了描述 PostgreSQL 数据库系统的内容。
- Oracle 的 Hakan Jakobsson, 感谢他编写的关于 Oracle 数据库系统的内容。
- IBM 的 Sriram Padmanabhan, 感谢他编写的关于 IBM DB2 数据库系统的内容。
- Microsoft 的 Sameet Agarwal, José A. Blakeley, Thierry D' Hers, Gerald Hinson, Dirk Myers, Vaqar Pirzada, Bill Ramos, Balaji Rathakrishnan, Michael Rys, Florian Waas 和 Michael Zwillling, 感谢他们编写的关于 Microsoft SQL Server 数据库系统的内容。José Blakeley, 感谢他还参与了第 29 章的调整和编辑。感谢 Microsoft 的 César Galindo-Legaria, Goetz Graefe, Kalen Delaney 和 Thomas Casey 在前面版本编写关于 Microsoft SQL Server 的章节的贡献。
- Chen Li 和 Sharad Mehrotra, 感谢他们给出关于 JDBC 和安全性的材料, 这些材料对第 8 章的更新和扩展帮助很大。
- Valentin Dinu, Goetz Graefe, Bruce Hillyer, Chad Hogg, Nahid Rahman, Patrick Schmid, Jeff Storey, Prem Thomas, Liu Zhenning 和 N. L. Sarda, 他们的反馈有助于我们准备第 5 版。
- Rami Khouri, Nahid Rahman 和 Michael Rys, 感谢他们对本书第 5 版草稿的反馈。
- Raj Ashar, Janek Bogucki, Gavin M. Bierman, Christian Breimann, Tom Chappell, Y.C. Chin, Laurens Damen, Prasanna Dhandapani, Arvind Hulgeri, Zheng Jiaping, Graham J.L. Kemp, Hae Choon Lee, Sang-Won Lee, Thanh-Duy Nguyen, D.B. Phatak, Juan Altmayer Pizzorno, Rajarshi Rakshit, Greg Riccardi, N.L. Sarda, Max Smolens, Nikhil Sethi 和 Tim Wahls 感谢他们指出第 5 版中的错误。
- Marilyn Turnamian, 她做了很好的秘书工作, 对于第 5 版的按时完成起了至关重要的作用。

本书的出版人是 Betsy Jones。主编是 Kelly Lowery。项目编辑是 Melinda D. Bilecki。项目经理是 Peggy Selle。执行营销经理是 Michael Weitz。营销经理是 Dawn Bercier。封面图案绘制和封面设计是 JoAnne Schopler。文字编辑是 George Watson。校对是 Judy Gantenbein。设计者是 Laurie Janssen。索引制作者是 Tobiah Waldron。

本书的这一版是基于前 4 版写成的, 我们再一次感谢帮助我们完成前 4 版的人们, 包括 R. B. Abhyankar, Don Batory, Phil Bernhard, Haran Boral, Paul Bourgeois, Phil Bohannon, Robert Brazile, Yuri Breitbart, Michael Carey, Soumen Chakrabarti, J. Edwards, Christos Faloutsos, Homma Farian, Alan Fekete, Shashi Gadia, Jim Gray, Le Gruenwald, Eitan M. Gurari, Ron Hitchens, Yannis Ioannidis, Hyoun-Joo Kim, Won Kim, Henry Korth (Henry F 的父亲), Carol Kroll, Gary Lindstrom, Irwin Levinstein, Ling Liu, Dave Maier, Keith Marzullo, Fletcher Mattox, Sharad Mehrotra, Jim Melton, Alberto Mendelzon, Hector Garcia-Molina, Ami Motro, Bhagirath Narahari, Anil Nigam,

Cyril Orji, Meral Ozsoyoglu, Bruce Porter, Jim Peterson, K.V. Raghavan, Krithi Ramamritham, Mike Reiter, Odinaldo Rodriguez, Mark Roth, Marek Rusinkiewicz, Sunita Sarawagi, N.L. Sarda, S.Seshadri, Shashi Shekhar, Amit Sheth, Nandit Soparkar, Greg Speegle, Dilys Thomas 和 Marianne Winslett。

Marilyn Turnamian 和 Nandprasad Joshi 为第 4 版担任秘书工作, Marilyn 还为第 4 版的封面设计准备了一份早期草稿。Lyn Dupré 编辑了第 3 版原稿, Sara Strandtman 编辑了第 3 版的文字。Nilesh Dalvi, Sumit Sanghai, Gaurav Bhalotia, Arvind Hulgeri K.V. Raghavan, Prateek Kapadia, Sara Strandtman, Greg Speegle 和 Dawn Bezviner 帮助我们准备前面几版的讲课教师参考手册。新的封面是前面四版封面的演进, 最先向我们建议用船作为封面概念的一部分的是 Bruce Stephan。

最后, Sudarshan 感谢他的妻子 Sita 的爱和支持, 感谢他的儿子 Madhur 的爱。Hank 感谢他的妻子 Joan, 孩子 Abby 和 Joe 的爱和理解。Avi 感谢 Valerie 在本书修订期间的耐心和支持。

A.S.
H.F.K.
S.S.

目 录

出版者的话	
专家指导委员会	
译者简介	
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	3
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	9
1.6 数据库设计	9
1.6.1 设计过程	9
1.6.2 银行企业的数据库设计	10
1.6.3 实体-联系模型	10
1.6.4 规范化	11
1.7 基于对象数据库和半结构化数据库	12
1.7.1 基于对象数据模型	12
1.7.2 半结构化数据模型	12
1.8 数据存储和查询	12
1.8.1 存储管理器	13
1.8.2 查询处理器	13
1.9 事务管理	13
1.10 数据挖掘与分析	14
1.11 数据库体系结构	15
1.12 数据库用户和管理员	17
1.12.1 数据库用户和用户界面	17
1.12.2 数据库管理员	17

1.13 数据库系统的历史	18
1.14 小结	19
术语回顾	20
实践习题	20
习题	20
文献注解	20

第一部分 关系数据库

第2章 关系模型	24
2.1 关系数据库的结构	24
2.1.1 基本结构	24
2.1.2 数据库模式	25
2.1.3 码	27
2.1.4 查询语言	28
2.2 关系代数基本运算	29
2.2.1 选择运算	29
2.2.2 投影运算	29
2.2.3 关系运算的组合	30
2.2.4 并运算	30
2.2.5 集合差运算	30
2.2.6 笛卡儿积运算	31
2.2.7 更名运算	33
2.2.8 关系代数的形式化定义	34
2.3 附加的关系代数运算	34
2.3.1 集合交运算	35
2.3.2 自然连接运算	35
2.3.3 除运算	36
2.3.4 赋值运算	37
2.4 扩展的关系代数运算	37
2.4.1 广义投影	38
2.4.2 聚集函数	38
2.4.3 外连接	39
2.5 空值	41
2.6 数据库的修改	42
2.6.1 删除	42
2.6.2 插入	42
2.6.3 更新	43
2.7 小结	43

术语回顾	44	3.12 小结	74
实践习题	44	术语回顾	74
习题	45	实践习题	75
文献注解	46	习题	76
第 3 章 SQL	47	文献注解	77
3.1 背景	47	第 4 章 高级 SQL	79
3.2 数据定义	48	4.1 SQL 的数据类型与模式	79
3.2.1 基本域类型	48	4.1.1 SQL 中内建的数据类型	79
3.2.2 SQL 中的基本模式定义	48	4.1.2 用户定义类型	80
3.3 SQL 查询的基本结构	50	4.1.3 大对象类型	81
3.3.1 select 子句	50	4.1.4 模式、目录与环境	81
3.3.2 where 子句	51	4.2 完整性约束	82
3.3.3 from 子句	51	4.2.1 单个关系上的约束	82
3.3.4 更名运算	52	4.2.2 not null 约束	82
3.3.5 元组变量	52	4.2.3 unique 约束	83
3.3.6 字符串运算	53	4.2.4 check 子句	83
3.3.7 排列元组的显示次序	54	4.2.5 参照完整性	84
3.3.8 重复	54	4.2.6 断言	86
3.4 集合运算	55	4.3 授权	86
3.4.1 union 运算	55	4.4 嵌入式 SQL	87
3.4.2 intersect 运算	55	4.5 动态 SQL	90
3.4.3 except 运算	56	4.5.1 ODBC	90
3.5 聚集函数	56	4.5.2 JDBC	92
3.6 空值	58	4.6 函数和过程化结构**	95
3.7 嵌套子查询	59	4.6.1 SQL 函数和过程	95
3.7.1 集合成员资格	59	4.6.2 过程化结构	96
3.7.2 集合的比较	60	4.6.3 外部语言例程	98
3.7.3 测试是否为空关系	60	4.7 递归查询**	99
3.7.4 测试是否存在重复元组	61	4.7.1 使用迭代的传递闭包	99
3.8 复杂查询	62	4.7.2 SQL 中的递归	100
3.8.1 派生关系	62	4.8 高级 SQL 特性**	101
3.8.2 with 子句	62	4.8.1 create table 的扩展	102
3.9 视图	63	4.8.2 关于子查询的更多内容	102
3.9.1 视图定义	64	4.8.3 数据库更新的高级结构	103
3.9.2 用其他视图定义视图	65	4.9 小结	103
3.10 数据库的修改	66	术语回顾	104
3.10.1 删除	66	实践习题	104
3.10.2 插入	67	习题	105
3.10.3 更新	68	文献注解	106
3.10.4 视图的更新	69	第 5 章 其他关系语言	107
3.10.5 事务	70	5.1 元组关系演算	107
3.11 连接关系**	71	5.1.1 查询的例子	107
3.11.1 举例	71	5.1.2 形式化定义	109
3.11.2 连接类型和条件	72	5.1.3 表达式的安全性	109