

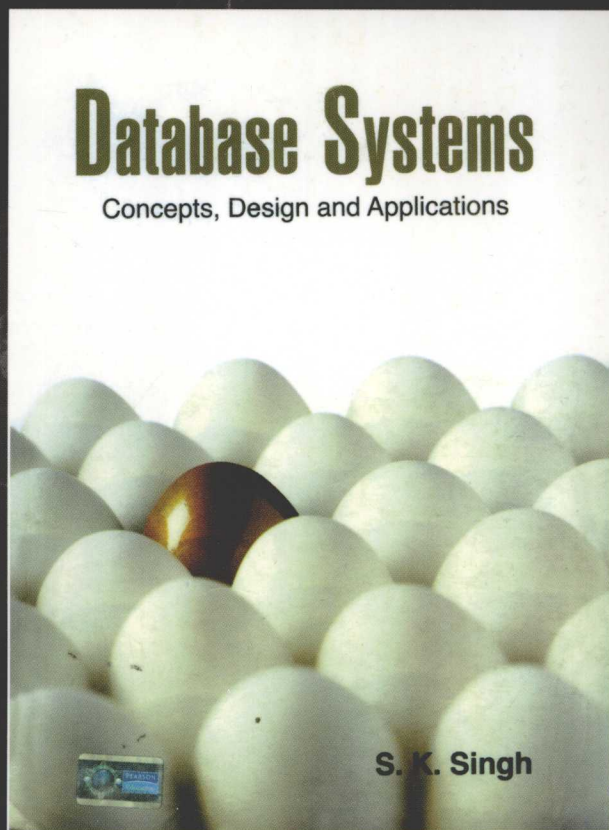
HZ BOOKS
华章教育
PEARSON
Education

计 算 机 科 学 丛 书

数据库系统

概念、设计及应用

(印) S. K. Singh 著 何玉洁 王晓波 车 蕾 等译



Database Systems
Concepts, Design and Applications



机械工业出版社
China Machine Press

计 算 机 科 学 丛 书

数据库系统 概念、设计及应用

(印) S. K. Singh 著 何玉洁 王晓波 车 蕾 等译

Database Systems

Concepts, Design and Applications

Database Systems

Concepts, Design and Applications



机械工业出版社
China Machine Press

本书全面地介绍了数据库系统的概念、设计和应用,共27章,分8个部分:第一部分数据库概念(第1~3章)、第二部分关系模型(第4~7章)、第三部分数据库设计(第8~10章)、第四部分查询、事务和安全管理(第11~14章)、第五部分基于对象的数据库(第15~16章)、第六部分高级和新兴的数据库概念(第17~21章)、第七部分案例学习(第22章)、第八部分商业数据库(第23~27章)。

本书适合作为计算机及相关专业本科生或研究生的教材,同时还是一本关于数据库技术非常全面的参考书籍,对数据库应用人员和IT领域的相关人员也有很好的参考价值。

Authorized translation from the English language edition, entitled DATABASE SYSTEMS: CONCEPTS, DESIGN AND APPLICATIONS, 9788177585674 by SINGH, S. K., published by Dorling Kindersley (India) Pvt. Ltd., publishing as Pearson Education, Copyright © 2006 Dorling Kindersley (India) Pvt. Ltd.

All rights reserved. No part of this book may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording or by any information storage retrieval system, without permission from Pearson Education, Inc.

CHINESE edition published by PEARSON EDUCATION ASIA LTD., and CHINA MACHINE PRESS Copyright © 2010.

This edition is manufactured in the People's Republic of China, and is authorized for sale and distribution in the People's Republic of China exclusively (except Taiwan, Hong Kong SAR and Macau SAR).

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:01-2009-1734

图书在版编目(CIP)数据

数据库系统:概念、设计及应用/(印)辛赫(Singh, S. K.)著;何玉洁等译. —北京:机械工业出版社,2009.11

(计算机科学丛书)

书名原文:Database Systems: Concepts, Design and Applications

ISBN 978-7-111-27958-7

I. 数… II. ①辛… ②何… III. 数据库系统 IV. TP311.13

中国版本图书馆CIP数据核字(2009)第138506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:迟振春

北京京师印务有限公司印刷

2010年1月第1版第1次印刷

184mm×260mm·39.25印张

标准书号:ISBN 978-7-111-27958-7

定价:89.00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:hzsj@hzbook.com

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分社较早意识到“出版要为教育服务”。自1998年开始，华章分社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。华章分社欢迎老师和读者对我们的工作提出建议或给予指正，我们的联系方法如下：

华章网站：www.hzbook.com
电子邮件：hzsj@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

译者序

数据库技术是计算机科学中一个非常重要的部分，数据库技术以及数据库的应用正以日新月异的速度发展，现代大学生（包括计算机专业和非计算机专业）学习和掌握数据库知识是非常必要的。本书作为一本内容全面的数据库专业技术图书，具有以下几个独到之处：

第一，内容丰富。本书涵盖了数据库基础理论的方方面面，从数据库的物理和逻辑结构，到数据库的分析与设计，再到数据库的安全与可靠等。此外，还包括了新的数据库技术和发展方向，比如分布式数据库、并行数据库、面向对象数据库等。同时，也详细地介绍了目前几个主流的商业数据库管理系统，如 Oracle、IBM DB2、SQL Server、MySQL 等。这体现了作者广博的知识以及深厚的理论和实践功底。

第二，通俗易懂。采用大量图表和示例讲解枯燥抽象的理论概念，使读者易于阅读和理解。

第三，易于教学。本书每章的最后都包含大量习题，且习题形式多样，包括复习题、判断题、选择题、填空题等，这既减轻了教师的出题工作量，又便于学生了解自己的学习情况。本书还提供了习题答案^①，以方便读者自我验证对知识的掌握程度。

本书作为一本教科书，非常适合计算机及相关专业本科生使用，同时它还是一本关于数据库技术的非常全面的参考书籍，对数据库应用人员和 IT 领域的相关人员也有很好的参考价值。

本书由何玉洁组织翻译并负责审校，王晓波、车蕾、张泽虹、赵延、张宏旭、于小倩、范洁、张俊超、赵凯、吴春飞、卢立能、李迎、何青、迟育强、靳秀罗等参加了本书部分章节的翻译工作。具体情况如下：王晓波翻译了第 1、2、5 章，张泽虹翻译了第 3、4 章，车蕾翻译了第 6、7、8、9、10、13、14 章，张宏旭翻译了第 11 章，于小倩翻译了第 12 章，卢立能翻译了第 15 章，范洁翻译了第 16 章，张俊超翻译了第 17 章，赵凯翻译了第 18 章，吴春飞翻译了第 19 章，赵延翻译了第 20 章，李迎翻译了第 21 章，何青翻译了第 22 章，迟育强翻译了第 23 章，靳秀罗翻译了第 24 章，其他部分由何玉洁翻译。是大家的认真工作使本书得以顺利完成，在此表示衷心的感谢。

限于译者水平，难免有疏漏和不妥之处，恳请读者批评指正。

何玉洁

^① 习题答案请登录华章网站（www.hzbook.com）下载。——编辑注

序

数据库已经成为无处不在、几乎涉及任何活动的技术，今天每个 IT 应用都以某种形式使用数据库。数据库已经对所有的应用造成了巨大的影响，并且使健康、教育、娱乐、工业以及银行等领域产生了根本的变化。

数据库系统的发展经历了从 20 世纪 60 年代的层次和网状模型到今天的关系模型，从基于文件的系统（基本上是数据的仓库，提供非常简单的检索功能）到现在用于复杂的环境，在用户友好的环境中提供丰富的功能。经过学术界坚持不懈的努力，数据库的服务在不断改善。在这些复杂软件包的背后是数学以及其他研究成果，这些研究成果为数据库系统提供了框架和基本构建块。在动态和灵活的环境中以用户友好的方式提供优质的数据库服务是一个挑战，理解数据库系统的基础知识是设计良好应用的关键。

数据库系统是美国大多数 B. Tech. /MCA/IT 教学计划中的核心课程，本书考虑了学生和教师的需要，广泛地覆盖这些课程的大纲，同时书中的大量实例也可满足专业开发人员的需要。

我相信学生、教师以及数据库系统开发人员都能从本书中获益。

S. K. Gupta

IIT Delhi 计算机科学与工程系教授

前言

本书涵盖了数据库系统的重要领域，完整地呈现了数据库管理系统的原理，以及数据库的发展方向。

本书可作为计算机科学与工程、信息技术以及 BCA、MCA 和 MBA 课程的本科生和研究生的教科书，也可作为工程技术人员的参考手册和实用指南。

本书全面地介绍了数据库系统的概念、设计和应用，共 27 章，分 8 个部分：

第一部分（数据库概念：第 1~3 章）概述数据库系统、数据库架构和物理数据组织的概念。

第二部分（关系模型：第 4~7 章）介绍了关系模型，并讨论了关系系统、查询语言以及实体-联系（E-R）模型。

第三部分（数据库设计：第 8~10 章）覆盖了数据库系统设计的各个方面，并讨论了通过函数分解和规范化处理获得最小数据冗余的方法。

第四部分（查询、事务和安全管理：第 11~14 章）讨论了查询、事务和数据系统中的恢复管理技术，包括查询优化技术（用于选择一个有效的执行计划来最小化运行时间）、数据库事务的主要性质、数据库恢复以及在发生故障时可用于确保数据库一致性的技术，同时包括如何防止数据安全的潜在威胁以及未授权的访问。

第五部分（基于对象的数据库：第 15~16 章）讨论了面向对象数据库的关键概念、面向对象语言，以及商业面向对象数据库管理系统中的新类。

第六部分（高级和新兴的数据库概念：第 17~21 章）介绍了高级和新兴的数据库概念，比如并行数据库、分布式数据库管理、决策支持系统、数据仓库以及数据挖掘。本部分还介绍了新兴的数据库技术，比如可用于 Web 的数据库、移动数据库、多媒体数据库、空间数据库以及数字图书馆。

第七部分（案例学习：第 22 章）提供了 6 个与数据库系统设计有关的案例，这些案例帮助学生回顾概念并加深对概念的理解。

第八部分（商业数据库：第 23~27 章）讨论目前在市场上可用的商业数据库，包括 DB2 Universal Database、Oracle、Microsoft SQL Server、Access 和 MySQL。这将帮助学生从理论向实际应用过渡。

全书包含大量图示，在每章的最后包含大量示例，以增强本书的可读性。欢迎对本书提出改进建议。

感谢

感谢在我完成这本书的过程中，以各种方式帮助我、赋予我灵感并给予我精神上的支持和鼓励的同事和朋友。感谢很多学生和工程师给我提供的有用的注释和建议。特别感谢本书手稿的评审者。

感谢 Tata Steel 的高级管理人员，感谢他们的鼓励和支持，没有他们我将不能完成此书。特别感谢 Tata Steel 前任管理者 J. A. C. Saldanha 先生，他给予我持续不断的灵感。

特别感谢我的妻子 Meena 和我的孩子，感谢他们的忍耐、理解和鼓励。永远感谢我的父母，感谢他们对我的爱、支持和鼓励。

深深地感谢 Tata Steel 的前任管理主任 J. J. Irani 博士，感谢他对本书的关心和鼓励。

最后，感谢 Pearson Education 的编辑团队给予我的协助。

目 录

出版者的话
译者序
序
前言

第一部分 数据库概念

第 1 章 数据库系统概述	1
1.1 引言	1
1.2 基本概念和定义	1
1.2.1 数据	2
1.2.2 信息	3
1.2.3 数据与信息	3
1.2.4 数据仓库	5
1.2.5 元数据	5
1.2.6 系统目录	6
1.2.7 数据项或字段	7
1.2.8 记录	7
1.2.9 文件	7
1.3 数据字典	7
1.3.1 数据字典的组件	9
1.3.2 主动和被动数据字典	11
1.4 数据库	11
1.5 数据库系统	13
1.6 数据库管理员	15
1.7 数据库管理员	16
1.8 面向文件系统与数据库系统	16
1.8.1 学习面向文件系统的优点	18
1.8.2 面向文件系统的缺点	18
1.8.3 数据库方法	20
1.8.4 数据库系统环境	20
1.8.5 DBMS 的优点	22
1.8.6 DBMS 的缺点	24
1.9 数据库系统的历史观点	25
1.10 数据库语言	27
1.10.1 数据定义语言	27
1.10.2 数据存储定义语言	28
1.10.3 视图定义语言	29
1.10.4 数据操纵语言	29
1.10.5 第四代语言	31
1.11 事务管理	31
习题	32
第 2 章 数据库系统架构	39
2.1 引言	39
2.2 模式、子模式和实例	39
2.2.1 模式	40
2.2.2 子模式	42
2.2.3 实例	43
2.3 ANSI-SPARC 的三层 数据库架构	43
2.3.1 内部层	45
2.3.2 概念层	46
2.3.3 外部层	46
2.3.4 三层架构的优点	46
2.3.5 三层架构的特性	47
2.4 数据独立性	47
2.4.1 物理数据独立性	48
2.4.2 逻辑数据独立性	48
2.5 映射	48
2.5.1 概念模式/内模式映射	48
2.5.2 外模式/概念模式映射	48
2.6 DBMS 的结构、组成和功能	49
2.6.1 DBMS 的结构	49
2.6.2 DBMS 的执行步骤	49
2.6.3 DBMS 的组成	50
2.6.4 DBMS 的功能和服务	51
2.7 数据模型	53
2.7.1 基于记录的数据模型	53
2.7.2 基于对象的数据模型	53
2.7.3 物理数据模型	54
2.7.4 层次数据模型	54
2.7.5 网状数据模型	56
2.7.6 关系数据模型	57
2.7.7 实体-联系数据模型	58
2.7.8 面向对象数据模型	60
2.7.9 不同数据模型之间的比较	61

2.8 数据库系统的类型	61	4.4 关系代数	111
2.8.1 集中式数据库系统	62	4.4.1 选择操作	112
2.8.2 并行数据库系统	62	4.4.2 投影操作	113
2.8.3 客户/服务器数据库系统	64	4.4.3 连接操作	114
2.8.4 分布式数据库系统	65	4.4.4 外连接操作	115
习题	66	4.4.5 并运算	115
第3章 物理数据组织	70	4.4.6 差运算	116
3.1 引言	70	4.4.7 交运算	117
3.2 物理存储媒体	70	4.4.8 笛卡儿积运算	117
3.2.1 主存储设备	71	4.4.9 除运算	118
3.2.2 二级存储设备	71	4.4.10 关系代数中的查询示例	119
3.2.3 三级存储设备	71	4.5 关系演算	121
3.2.4 高速缓存	72	4.5.1 元组关系演算	121
3.2.5 主存储器	72	4.5.2 域关系演算	123
3.2.6 闪存	72	习题	124
3.2.7 磁盘存储器	73	第5章 关系查询语言	130
3.2.8 光存储器	75	5.1 引言	130
3.2.9 磁带存储器	76	5.2 Codd的规则	130
3.3 RAID技术	77	5.3 基于语言的信息系统 (ISBL)	131
3.3.1 使用并行机制改进性能	77	5.3.1 ISBL 查询举例	132
3.3.2 RAID技术的优点	78	5.3.2 ISBL的局限性	133
3.3.3 RAID技术的缺点	78	5.4 查询语言	133
3.3.4 使用冗余来提高可靠性	78	5.4.1 QUEL 查询举例	134
3.3.5 RAID级别	79	5.4.2 QUEL的优点	135
3.3.6 RAID级别的选择	81	5.5 结构化查询语言	135
3.4 文件的基本概念	83	5.5.1 SQL的优点	136
3.4.1 文件类型	83	5.5.2 SQL的缺点	136
3.4.2 缓冲区管理	84	5.5.3 基本的SQL数据结构	137
3.5 文件组织	84	5.5.4 SQL数据类型	137
3.5.1 记录和记录型	85	5.5.5 SQL运算符	138
3.5.2 文件组织技术	89	5.5.6 SQL数据定义语言	139
3.6 索引	96	5.5.7 SQL数据查询语言	143
3.6.1 主索引	97	5.5.8 SQL数据操纵语言	145
3.6.2 二级索引	97	5.5.9 SQL数据控制语言	148
3.6.3 基于树的索引	98	5.5.10 SQL数据管理语句	149
习题	99	5.5.11 SQL事务控制语句	149
		5.6 嵌入式结构化查询语言	149
		5.7 基于例子的查询	150
		5.7.1 基于一个关系的QBE查询 (单表检索)	151
		5.7.2 基于几个关系的QBE查询 (多表检索)	155
		5.7.3 利用QBE对数据库进行修改 (更新、删除和插入)	156
第二部分 关系模型			
第4章 关系代数和关系演算	107		
4.1 引言	107		
4.2 关系模型的历史观点	107		
4.3 关系数据库的结构	108		
4.3.1 域	108		
4.3.2 关系的键	109		

5.7.4	Microsoft Access 中的 QBE 查询	156
5.7.5	QBE 的优点	161
5.7.6	QBE 的缺点	161
	习题	161
第 6 章	实体 - 联系模型	170
6.1	引言	170
6.2	E-R 模型的基本概念	170
6.2.1	实体	170
6.2.2	联系	171
6.2.3	属性	174
6.2.4	约束	177
6.3	E-R 模型向关系的转换	178
6.4	E-R 模型存在的问题	185
6.4.1	扇形陷阱	186
6.4.2	深坑陷阱	186
6.5	E-R 图符号	187
	习题	189
第 7 章	增强的实体 - 联系模型	195
7.1	引言	195
7.2	超类和子类实体型	195
7.2.1	超类和子类的符号	196
7.2.2	属性继承	197
7.2.3	超类/子类联系的使用条件	198
7.2.4	使用超类和子类的优点	198
7.3	特殊化和一般化	199
7.3.1	特殊化	199
7.3.2	一般化	200
7.3.3	特殊化和一般化的约束	201
7.4	分类	203
7.5	EER 图的例子	203
	习题	204

第三部分 数据库设计

第 8 章	数据库设计介绍	207
8.1	引言	207
8.2	软件开发生命周期	207
8.2.1	软件开发成本	208
8.2.2	结构化系统分析与设计	209
8.3	数据库开发生命周期	210
8.4	自动化设计工具	216
8.4.1	手工数据库设计的局限性	216
8.4.2	计算机辅助软件工程工具	217
	习题	219

第 9 章	函数依赖和模式分解	221
9.1	引言	221
9.2	函数依赖	221
9.2.1	函数依赖图及其示例	222
9.2.2	完全函数依赖	225
9.2.3	函数依赖的 Armstrong 公理	225
9.2.4	多余的函数依赖	226
9.2.5	函数依赖集的闭包	226
9.3	模式分解	228
9.3.1	有损分解	229
9.3.2	无损连接分解	230
9.3.3	保持函数依赖的分解	231
	习题	232
第 10 章	规范化	235
10.1	引言	235
10.2	规范化	235
10.3	范式	236
10.3.1	第一范式	237
10.3.2	第二范式	239
10.3.3	第三范式	242
10.4	Boyce-Codd 范式	244
10.5	多值依赖和第四范式	246
10.5.1	多值依赖的性质	249
10.5.2	第四范式	250
10.5.3	多值依赖和 4NF 的问题	251
10.6	连接依赖和第五范式	251
10.6.1	连接依赖	251
10.6.2	第五范式	252
	习题	253

第四部分 查询、事务和安全管理

第 11 章	查询处理及优化	259
11.1	引言	259
11.2	查询处理	259
11.3	语法分析器	260
11.4	查询分解	261
11.4.1	查询分析	261
11.4.2	查询规范化	263
11.4.3	语义分析器	264
11.4.4	查询简化器	265
11.4.5	查询重构	267
11.5	查询优化	267
11.5.1	启发式查询优化	268
11.5.2	转换规则	270

15.3.9 OO 概念的优点	349	17.4.4 加锁	374
15.4 面向对象 DBMS	350	17.5 查询并行机制	375
15.4.1 OODBMS 特性	350	17.5.1 I/O 并行机制	
15.4.2 OODBMS 的优点	350	(数据分区)	375
15.4.3 OODBMS 的缺点	351	17.5.2 查询内并行机制	377
15.5 对象数据管理组和面向		17.5.3 交互查询并行机制	378
对象语言	351	17.5.4 操作内并行机制	378
15.5.1 对象模型	351	17.5.5 交互操作并行机制	378
15.5.2 对象定义语言	352	习题	379
15.5.3 对象查询语言	354	第 18 章 分布式数据库系统	384
习题	356	18.1 引言	384
第 16 章 对象-关系数据库	359	18.2 分布式数据库	384
16.1 引言	359	18.2.1 并行数据库与分布式数据库	
16.2 对象-关系 DBMS 的历史	359	之间的区别	384
16.2.1 RDBMS 的缺点	359	18.2.2 分布式数据库的期望性质	385
16.2.2 复杂对象	359	18.2.3 分布式数据库的类型	385
16.2.3 ORDBMS 的出现	360	18.2.4 分布式数据库的期望功能	387
16.3 ORDBMS 查询语言 (SQL3)	361	18.2.5 分布式数据库的优点	387
16.4 ORDBMS 设计	363	18.2.6 分布式数据库的缺点	388
16.4.1 ORDBMS 的挑战	364	18.3 分布式数据库的架构	388
16.4.2 ORDBMS 的特点	364	18.3.1 客户/服务器架构	388
16.4.3 ORDBMS 和 OODBMS 的		18.3.2 合作服务器系统	389
比较	364	18.3.3 中间件系统	390
16.4.4 ORDBMS 的优点	365	18.4 分布式数据库系统设计	391
16.4.5 ORDBMS 的缺点	366	18.4.1 数据分解	391
习题	366	18.4.2 数据分配	394
		18.4.3 数据复制	396
第六部分 高级和新兴的数据库概念		18.5 分布式查询处理	397
第 17 章 并行数据库系统	369	18.6 分布式数据库中的并发控制	400
17.1 引言	369	18.6.1 分布式加锁	401
17.2 并行数据库	369	18.6.2 分布式死锁	401
17.2.1 并行数据库的优点	370	18.6.3 时间戳	402
17.2.2 并行数据库的缺点	370	18.7 分布式数据库中的恢复控制	402
17.3 并行数据库的架构	370	18.7.1 两阶段提交	402
17.3.1 共享内存的多 CPU 并行		18.7.2 三阶段提交	404
数据库架构	370	习题	404
17.3.2 共享磁盘的多 CPU 并行		第 19 章 决策支持系统	410
数据库架构	371	19.1 引言	410
17.3.3 无共享的多 CPU 并行		19.2 决策支持系统的历史	410
数据库架构	371	19.3 决策支持系统的定义	411
17.4 并行数据库处理的关键要素	373	19.3.1 DSS 的性质	412
17.4.1 加速比	373	19.3.2 DSS 的优点	412
17.4.2 可扩展性	373	19.3.3 DSS 的组件	413
17.4.3 同步	374	19.4 操作型数据与 DSS 数据	414

22.6.4	需求定义中的更改 (附加)	478	23.5	安装 DB2 客户端的必备条件	515
22.6.5	修改的表定义	479	23.5.1	安装必备条件: DB2 客户端 (Windows)	515
22.6.6	模式求精	479	23.5.2	安装必备条件: DB2 查询巡视 器客户端 (Windows)	516
22.6.7	修改实体 - 联系图	480	23.6	安装和配置 DB2 UDB 服务器	517
22.6.8	逻辑数据库设计: 示例表数据	480	习题	524
22.7	客户订购仓库的数据库设计	481	第 24 章	Oracle	528
22.7.1	需求定义及分析	481	24.1	引言	528
22.7.2	概念设计: 实体 - 联系图	481	24.2	Oracle 的历史	528
22.7.3	逻辑数据库设计: 表定义	481	24.2.1	Oracle 家族	529
22.7.4	逻辑数据库设计: 示例表数据	482	24.2.2	Oracle 软件	531
22.7.5	函数依赖图	482	24.3	Oracle 特性	531
22.7.6	逻辑记录结构及访问路径	484	24.3.1	应用程序开发特性	531
习题	484	24.3.2	通信特性	533
第八部分 商业数据库			24.3.3	分布式数据库特性	535
第 23 章	IBM DB2 通用数据库	487	24.3.4	数据移动特性	536
23.1	引言	487	24.3.5	性能特性	537
23.2	DB2 产品	487	24.3.6	数据库管理特性	538
23.3	DB2 通用数据库	489	24.3.7	备份和恢复特性	539
23.3.1	DB2 UDB 的配置	489	24.3.8	Oracle Internet 开发 人员套件	540
23.3.2	其他 DB2 UDB 相关的产品	496	24.3.9	Oracle Lite	542
23.3.3	DB2 通用数据库的 主要组件	500	24.4	SQL * Plus	542
23.3.4	DB2 UDB 的特性	503	24.4.1	SQL * Plus 的特性	542
23.4	安装 DB2 UDB 服务器的 必备条件	505	24.4.2	使用 SQL * Plus	542
23.4.1	安装必备条件: DB2 UDB 个人版 (Windows)	505	24.4.3	编辑器命令	543
23.4.2	安装必备条件: DB2 工作组 服务器版和无分区的 DB2 企 业服务器版 (Windows)	506	24.4.4	SQL * Plus 帮助系统以及其他 有用的命令	543
23.4.3	安装必备条件: 分区的 DB2 企业服务器版 (Windows)	508	24.4.5	格式化输出	544
23.4.4	安装必备条件: DB2 连接 个人版 (Windows)	510	24.5	Oracle 的数据字典	544
23.4.5	安装必备条件: DB2 连接 企业版 (Windows)	512	24.5.1	数据字典表	544
23.4.6	安装必备条件: DB2 查询 巡视器服务器 (Windows)	513	24.5.2	数据字典视图	545
23.4.7	安装必备条件: DB2 多维 视图 (Windows)	514	24.6	Oracle 系统架构	546
			24.6.1	存储管理和进程	546
			24.6.2	逻辑数据库结构	548
			24.6.3	物理数据库结构	548
			24.7	Oracle 9i 安装指南	549
			习题	554
			第 25 章	Microsoft SQL Server	557
			25.1	引言	557
			25.2	Microsoft SQL Server 设置	557
			25.2.1	SQL Server 2000 版本	557
			25.2.2	SQL Server 2005 版本	558

25.2.3	Microsoft SQL Server 的特性	558	27.2.3	MySQL 表大小	592
25.3	SQL Server 中的存储过程	559	27.2.4	MySQL 开发途径	593
25.3.1	存储过程的优点	559	27.2.5	MySQL 4.0 中可使用的 特性	594
25.3.2	存储过程的结构	559	27.2.6	嵌入的 MySQL 服务器	595
25.4	安装 Microsoft SQL Server 2000	560	27.2.7	MySQL 4.1 的特性	595
25.4.1	安装步骤	560	27.2.8	MySQL 5.0: 下一个 开发的版本	596
25.4.2	启动和停止 SQL Server	560	27.2.9	MySQL 邮件发送清单	596
25.4.3	自动启动 SQL Server 服务	563	27.2.10	MySQL 支持的操作系统	597
25.4.4	到 Microsoft SQL Server 数据库 系统的连接	564	27.3	PHP 简介	598
25.4.5	数据的源	564	27.3.1	PHP 语言语法	598
25.4.6	安全性	569	27.3.2	PHP 变量	599
25.5	Microsoft SQL Server 的 数据库操作	571	27.3.3	PHP 操作	599
25.5.1	数据库连接	571	27.3.4	安装 PHP	601
25.5.2	创建数据库	572	27.4	MySQL 数据库	601
习题		575	27.4.1	创建自己的第一个数据库	601
第 26 章	Microsoft Access	577	27.4.2	MySQL 连接	601
26.1	引言	577	27.4.3	选择使用的数据库	602
26.2	Access 数据库	577	27.4.4	MySQL 表	602
26.2.1	表	577	27.4.5	创建 MySQL 表	603
26.2.2	查询	578	27.4.6	在 MySQL 表中插入数据	604
26.2.3	报表	579	27.4.7	MySQL 查询	604
26.2.4	窗体	579	27.4.8	从 MySQL 中检索信息	605
26.2.5	宏	581	27.5	在 Windows 中安装 MySQL	606
26.3	Access 中的数据库操作	582	27.5.1	Windows 系统需求	607
26.3.1	创建窗体	582	27.5.2	选择一个安装包	607
26.3.2	创建一个简单的查询	582	27.5.3	使用自动的安装程序 安装 MySQL	607
26.3.3	修改查询	585	27.5.4	使用 MySQL 安装向导	607
26.4	Microsoft Access 的特性	588	27.5.5	下载与启动 MySQL 安装向导	608
习题		589	27.5.6	MySQL 安装步骤	608
第 27 章	MySQL	590	27.5.7	建立许可和密码	609
27.1	引言	590	习题		611
27.2	MySQL 概览	590	参考文献		612
27.2.1	MySQL 特性	590			
27.2.2	MySQL 稳定性	592			

第一部分 数据库概念

第 1 章 数据库系统概述

1.1 引言

在当今竞争的环境下，数据（或信息）以及对它们的有效管理是一个组织非常重要的业务目标。我们现在正处在一个信息爆炸的时代，人们正在被各种数据轰炸着，从各种数据中得出正确的信息从而在适当的时机做出正确的决定是一件很困难的事情。因此，现在一个组织要获得成功，会比以往更依赖于一种能力，即能获取商业或业务方面的准确、可靠、及时的数据，从而为其做出高效的决策提供帮助。

数据库系统是一个工具，它能简化上面提到的管理数据的任务，并及时地从中抽取出有用的信息。它分析并指导一个组织的行动或业务目标。在一个组织的信息系统中，它是数据的中心仓库，对组织的正常运转起着基本的支撑作用，它维护这些数据，帮助用户在做决策的时候解释这些数据。管理者们为了在竞争中占据优势，正在探索使用从数据库中产生的知识。例如，确定顾客的购货模式、跟踪销售、支持顾客关系管理（customer relationship management, CRM）、在线购物、雇员关系管理、实现决策支持系统（decision support system）、库存管理等。为了满足不断变化的组织需求，数据库结构必须要灵活，以便接受新的数据，适应新的关系，从而支持新的决策。

随着计算技术的快速发展和它在现代社会各个领域的广泛应用，数据库已经成为人们日常生活当中一个完整的组成要素。在日常生活中我们会遇到一些与数据库交互的行为，比如，存取款的银行数据库、航空或铁路的票务预订数据库、图书馆查找特定图书的数据库、超市保存存货清单的货物数据库、使用信用卡购物时检查是否有足够信用余额的数据库，等等。

事实上，数据库和数据库管理系统（DataBase Management System, DBMS）在管理商业、政府、银行、大学等方面已经变得非常重要。因此，它们是当今软件工业中非常重要的成员，软件工业要支持这些需求并管理日益增长的令人畏缩的海量数据。

本章介绍数据库和数据库管理系统的基本概念以及 DBMS 的目标、数据模型的类型和存储管理系统。

1.2 基本概念和定义

随着计算机使用的不断普及，那些以数据作为基本资源的各种组织正快速地将其人工系统向计算机化的信息系统转换。因此，对数据的正确组织和管理是一个组织有效运作所必需的。数据的有效使用对于一个组织的计划制定、产品控制、市场营销、货品计价、工资表、会计以及其他事务有着重要的影响。本节给出数据库中所使用术语的形式化定义。

1.2.1 数据

可以将数据定义为一个已知的事实，它能够被记录并具有固定的含义。数据是原始、孤立的事实，从中我们可以得到需要的信息。

数据是独特的信息片段，通常需要按特定的方法将其转换为某种格式。它们用二进制表示，是存储在计算机中的逻辑实体。一个数据表示我们感兴趣的一个事实。对于一个工业组织，它可以是 Thomas Mathew 的雇员（或社会保险）号 106519，可以是该组织的最大铸件材料提供商的地址 Indore，还可以是某个重要的客户 M/s Elbee 公司的电话号码 001-732-3931650。类似地，对于一个研究和开发机构来说，数据或许是新产品的最大数量 100；对于一个培训机构来说，数据或许是数据库管理课程的最大报名人数。因此，一个数据就是我们所关心的周围环境中有关某情况的一个事实。

数据可以在用户的环境中以各种形式存在，它具有某种含义，比如某篇论文中的数据或文本，存储在计算机内存中的位或字节，存在某人头脑中的事实。数据也可以是物体，比如文档、图片甚至是视频片段。数据的例子如表 1-1 所示。

表 1-1 数据的例子

从销售人员观点	在电子供应商的领域	在职工头脑中
客户名	客户名	职工名
客户账号	客户编号	标识号
地址	地址	部门
电话号码	电话号码	出生日期
	单位消耗量	资格
	应付数量	技能类型

通常人们用很多事实来描述所感兴趣的事物。例如，我们来考虑这样的事实，作为 M/s Elbee 公司的经理，我们感兴趣的可能是雇员 Thomas Mathew，我们希望记住他的雇员号是 106519，他的基本工资是每月 2 000 000Rs.（4000 美元），他所在的城市是 Jamshedpur，他所在的国家是印度，他的出生日期是 1957 年 9 月 6 日，他的结婚纪念日是 5 月 29 日，他的电话号码是 0091-657-2431322，等等。我们知道这些事实是为了每月处理 Mathew 的工资单，在生日和结婚纪念日发贺卡给他，打印他的工资条，万一遇到紧急情况通知他的家人，等等。显然，收集所有有关 Mathew 的事实（或数据），并将它们保存在一起确实是很有道理的。表 1-2 显示了有关 Thomas Mathew 的工资单和其他相关情况的所有信息。

表 1-2 Thomas Mathew 的工资单事实

Employee Number	Employee Last Name	Employee First Name	Basic Salary (US \$)	Home Town	Home Country	Date of Birth	Marriage Anniversary	Telephone Number
106519	Mathew	Thomas	4000	Jamshedpur	India	06-09-1957	29.05	0091-657-2431322

数据（data）表示一个信息片段。但实际上，数据一词既可作为单数也可作为复数来使用。术语“数据”一般区分机器可读的（二进制）信息与人类可读的（文本的）信息。例如，一些应用程序区分数据文件（它包含二进制数据）和文本文件（它包含 ASCII 数据）。但数字、字符或它们二者都能表示数据。

三层数据架构

为了将分散在整个组织的数据集中起来，并能对有效决策提供支持，一般将数据组织成如