



面向 21 世纪 课程 教材
Textbook Series for 21st Century

数据库系统概论

(第三版)

萨师煊 王 珊



高等教育出版社
HIGHER EDUCATION PRESS

面向 21 世纪课程教材
Textbook Series for 21st Century

数据库系统概论

(第三版)

萨师煊 王 珊



01153509



高等教育出版社
HIGHER EDUCATION PRESS

内容提要

本书是教育部“高等教育面向 21 世纪教学内容和课程体系改革计划”的研究成果，是面向 21 世纪课程教材和教育部高等学校计算机科学与技术学科“九五”规划教材。

本书是在第二版基础上修订而成的，与第二版相比较，在整体结构上进行了适当的调整，增加了数据库新技术方面的内容。全书内容包括：数据库模型、数据库系统结构、关系数据库系统、SQL 语言、复杂数据理论、数据库维护、数据库设计、关系数据库管理系统、数据库技术新进展、面向对象数据库系统、分布式数据库系统、并行数据库系统等。本书曾获国家优秀教材奖，并且是教育部“九五”重点教材。

本书可作为高等学校计算机有关专业的数据库课程教材，也可供从事计算机软件工作的科技人员和工程技术人员以及其他有关部门人员参阅。

图书在版编目(CIP)数据

数据库系统概论/萨师焯,王珊编著. —3 版. —北
京:高等教育出版社,2000.2(2003 重印)

ISBN 7-04-007494-X

I. 数… II. ①萨… ②王… 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 (2000) 第 01158 号

数据库系统概论(第三版)

萨师焯 王珊

出版发行	高等教育出版社	购书热线	010-64054588
社 址	北京市东城区沙滩后街 55 号	免费咨询	800-810-0598
邮政编码	100009	网 址	http://www.hep.edu.cn
传 真	010-64014048		http://www.hep.com.cn
经 销	新华书店北京发行所	版 次	1983 年 8 月第 1 版
印 刷	北京中科印刷有限公司		1991 年 4 月第 2 版
开 本	787×960 1/16		2000 年 2 月第 3 版
印 张	30	印 次	2003 年 5 月第 10 次印刷
字 数	540 000	定 价	25.10 元

凡购买高等教育出版社图书，如有缺页、倒页、脱页等
质量问题，请在所购图书销售部门联系调换。

版权所有 侵权必究

23



面向 21 世纪课程教材

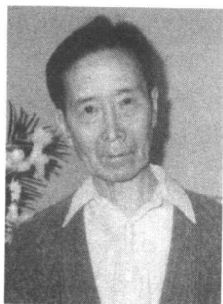


普通高等教育“九五”
国家教委重点教材



SJS312/01

作者简介



萨师焯教授，我国最早从事数据库教学和科研的专家之一。是中国数据库学科的奠基人。在国内外发表许多论文和著作，历任中国人民大学信息系主任、中国计算机学会常务理事，数据库学组组长等职务。

萨师焯教授是我国数据库学术活动的倡导者和领导者，对我国数据库技术的发展、应用和学术交流起了很大的推动作用，对我国数据库技术跟踪国际前沿，缩短我国数据库技术与国际差距，做出了很大贡献。



王珊教授，中国人民大学信息学院院长，中科院计算所兼职研究员，博士生导师。中国计算机学会副理事长，中国计算机学会数据库专委会主任。

自 1981 年以来一直从事数据库方向的教学、科研和开发工作。1984~1986 年在美国马里兰大学工作，设计和开发可扩展的关系数据库管理系统 XDB。主持和承担了七五、八五、九五国家科技攻关项目、863 高科技计划项目、国家自然科学基金项目等 20 多项；在国内外杂志发表论文百余篇，有数据库方面的著作十余部。获多项国家级和省部级科技进步奖以及全国优秀教师和全国五一劳动奖章等荣誉称号。

第三版前言

数据库技术从 20 世纪 60 年代中期产生到今天仅仅 30 多年的历史，已经历了三代演变，造就了 C.W.Bachman、E.F.Codd 和 James Gray 三位图灵奖获得者；发展了以数据建模和 DBMS 核心技术为主，内容丰富的一门学科；带动了一个巨大的软件产业 DBMS 产品及其相关工具和解决方案。30 多年成就辉煌。

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，它已成为计算机信息系统与应用系统的核心技术和重要基础。本书系统地阐述数据库系统的理论、技术和方法，是《数据库系统概论》的第三版。第一版 1983 年出版(1987 年获国家级优秀教材奖)，第二版 1991 年出版。

针对数据库技术的发展和我国应用水平的提高，我们对第二版做了较大的调整、修改和增删，但原书的基本宗旨和风格不变，保持讲述数据库的基本概念、基本理论和基本技术为主的特点。第三版主要的修改是：

1. 根据我国实际情况，网状、层次数据库系统已很少使用，因此把它们删去了，有关的主要概念放在第一章数据模型中介绍。

2. 进一步加强了关系数据库系统的讲解，特别是 SQL 语言的介绍，以适应当前广泛使用关系数据库系统的需要。

3. 把第二版第八章“数据库保护”中的安全性，完整性，并发控制和恢复四节扩展为本书“系统篇”中的四章，内容作了适度的加宽和加深。随着大型数据库系统的普遍使用，这些知识和技术是运行和维护数据库系统必不可少的。本书从使用和管理的角度讲解这些知识而不是讨论实现这些功能的内部技术。

4. 为了反映数据库技术的发展，增加了新技术篇。第十二章全面介绍了数据库发展的总体轮廓，从数据模型、新技术内容、应用领域三个方面，阐述了新一代数据库系统及其相互关系。并选择当前较重要的新技术在后面三章中介绍，它们是面向对象数据库系统，分布式数据库系统和并行数据库系统。

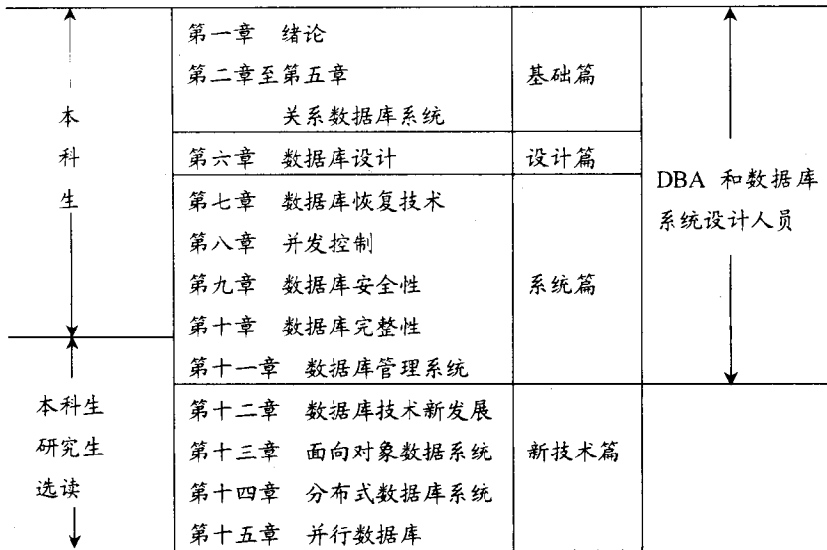
5. 每章后配有小结，习题及阅读参考文献，并对许多文献作了简要的注释，以便读者进一步参考。

此外，为了辅助教学和加强上机实习，人大金仓信息技术有限公司开发了与本书配套的 CAI 软件和 Easybase 数据库系统（个人版、网络版和企业版），

有兴趣的读者和单位可以向金仓公司购买(电话 010-62515575)。

全书分为四篇共十五章。第一章至第十章,是计算机软件专业本科生的基本教程(书中有*号的部分除外),第十一章至第十五章是高级教程。

下图给出了本书各章之间的联系和读者对象示意。



本书内容丰富,讲授时可根据学生及专业情况酌情取舍。例如,对于计算机专业本科学生,第二章 2.5 节“关系演算”,第五章 5.4 节“模式分解”可适当压缩,新技术篇中的章节,教师可以选择部分内容进行讲解。

本书由王珊教授执笔,萨师焯教授审定,陈红副教授和研究生曹会萍、王静等参与了内容讨论和书稿校阅工作。

中国科技大学研究生院邵佩英教授和北京大学杨冬青教授详细审阅了全稿并提出了许多有益的意见,在此向他们表示衷心的感谢。

我们在编写本书的过程中,努力跟踪数据库学科的新发展、新技术,把它们纳入到教材中来,力求反映当代新技术,以保持本书的先进性和实用性。但由于学识浅陋,见闻不广,必有许多不足之处,希学术同仁不吝赐教。

萨师焯 王珊

1999 年仲夏于中国人民大学

第二版前言

《数据库系统概论》第一版出版于 1983 年，距今已 6 年。在这 6 年中，不仅数据库技术有了很大进展，而且国内计算机专业的学生和技术人员的水平也有显著提高。因此在第二版中，我们针对这些情况对原书从结构到内容做了较大的调整、修改和增删；但原书的基本宗旨和风格不变，仍以国家教育委员会颁布的《数据库系统概论教学大纲》作为本书编写的基本依据，保持讲述数据库的基本概念、基本理论和基本技术为主的特点。

全书分为两大部分，共十一章。第一部分，第一至九章，是计算机软件专业本科生的基本教程。第二部分，第十、十一章是高级教程。

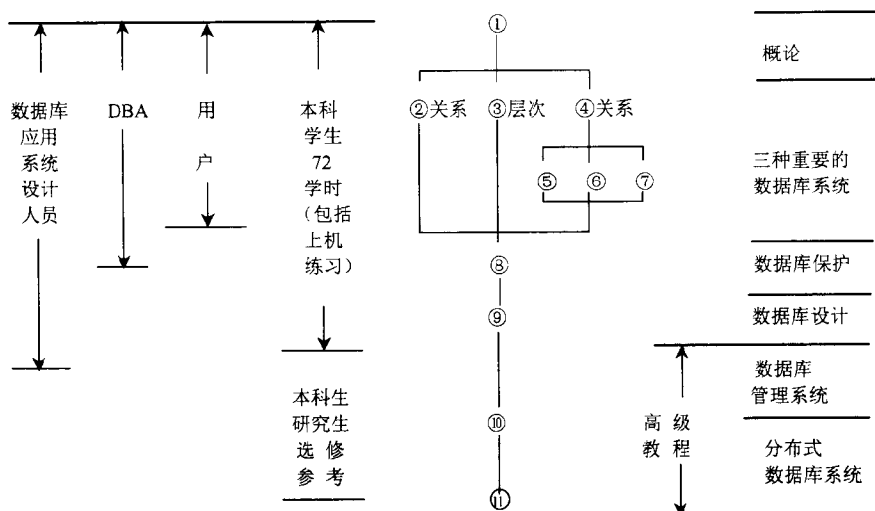
第一章绪论，概述了数据管理的进展、数据模型和数据库系统构成的一般概念。第二至七章介绍三种重要的数据库系统。它们是，网状数据库 DBTG 系统、层次数据库 IMS 和关系数据库系统。鉴于关系数据库具有许多优点并已在应用中日趋成熟，我们把重点放在这一部分，进一步充实了关系数据库的内容，共计有四章（第四至七章）。第四章概述关系模型的基本概念、关系代数 and 关系演算；第五章详细介绍关系数据库标准语言 SQL；第六章讨论关系系统及其查询优化；第七章讲述关系数据理论。

第八章“数据库保护”（包括数据库的安全性、完整性、并发控制和恢复）和第九章“数据库设计”都作了较大的变动，充实了内容，增强了实用性。

第二部分是新增的。包括第十章“数据库管理系统”和第十一章“分布式数据库系统”。这是为了加强读者对 DBMS 的了解，适应对“分布处理”日益普遍的需要，引导读者从学习本书开始向某些数据库的重要新领域过渡。

第一版中有附录 I “文件组织”与附录 II “IMAGE 数据库系统”。前者是基础性教材，由于读者水平的提高已无存在的必要；后者原作为实习课的参考教材，鉴于近几年引进多种商品化的关系数据库系统且应用甚为广泛，因此我们已另行编写了一本《实用数据库系统汇编》作为实习教材（高等教育出版社出版，1990 年出版），其中介绍了五种关系系统和 IMAGE 系统。因此本书删去了第一版中的附录。

下图给出了本书各章之间的联系和读者对象示意。



本书由王珊执笔，萨师煊审定，杜小勇、李曦老师和研究生唐元昌、武志文、任永杰等参与了内容讨论和校阅工作。

中国科技大学研究生院罗晓沛教授、清华大学俞盘祥副教授和北京大学杨冬青副教授详细审阅了全稿并提出了许多有益的意见，在此向他们致以衷心的感谢。

我们在改编过程中，尽可能引入新的观点和方法，力求能反映当代技术水平。但由于学识浅陋，见闻不广，必有许多不足之处，希同行指正。

萨师煊 王珊

1989年12月

于中国人民大学数据工程与知识工程研究所

第一版前言

数据库是数据管理的最新技术，是计算机科学的重要分支。十余年来，数据库管理系统已从专用的应用程序包发展成为通用的系统软件。由于数据库具有数据结构化、最低冗余度、较高的程序与数据独立性、易于扩充、易于编制应用程序等优点，较大的信息系统都是建立在数据库设计之上的。因此，不仅大型计算机及中小型计算机，甚至微型机都配有数据库管理系统。目前，对数据库各种模型的研究以及理论上的探讨都还在蓬勃开展，其应用也从一般管理扩大到计算机辅助设计、人工智能以及科技计算等领域。国外高等学校计算机科学系、计算机应用与信息系统系等都开设有数据库系统方面的课程。近年来，我国在计算机科学教育中也对数据库予以应有的重视。1983年教育部部属高等学校计算机软件专业教学方案将数据库概论列为四年制本科的必修课程，并已于1983年6月通过了教学大纲，我们这本书便是按照该大纲编写的一本教材。

全书共分十章和两个附录。附录I的“文件组织”系学习数据库的预备知识，但已散见于数据结构等先行课中，列在这里，仅供参考用，一般不必作为正式讲授内容。第一、二、三章为有关数据库的一般概念；第四至八章介绍三种重要数据库模型。其中第八章的公理系统与模式分解的大部分内容已超出大纲范围，讲授时可酌情取舍。第九章为数据库控制；第十章为数据库设计。附录II的“IMAGE3000数据库系统”介绍了如何在HP-3000计算机的IMAGE3000数据库管理系统上建立数据库和使用数据库。对于能够用到HP-3000计算机的单位，这一部分可作为实习教材；对于不能用上HP-3000计算机的单位则可作为参考，借以从中了解如何在一个实际的数据库管理系统上建立一个数据库应用系统。

本书是编者根据多年来进行数据库讲演及在高等学校为研究生、本科生开设数据库系统概论课程的经验，并在多次编写的内部交流讲义的基础上修改而成的，由于我们水平不高，而数据库作为计算机科学中的一个新兴分支发展又非常迅速，因此，在本书中存在的问题一定不少，希望读者提出批评意见。

中国科技大学研究生院罗晓沛同志、清华大学计算机科学与工程系杨德元同志、俞盘祥同志与沈金发同志详细阅读了全稿，并提出许多有益的意见，在此谨向他们致以衷心的感谢。

作为内部交流讲义阶段，有50余所兄弟院校采用作为教本，其中不少同

志在使用以后曾以口头或书面形式提出意见，我们也对他们表示感谢。

在讲稿整理过程中，我校信息管理系的刘怡同志、前 78 级、79 级研究生吴鸥琦等同志，82 级研究生张秉训、刘伶、杨明伟等同志以及资料室的王小平等同志，都从不同方面分别做了一些工作，对此作者一并表示诚挚的谢意。

萨师焯 王珊

1983 年 4 月于中国人民大学

目 录

第一篇 基础篇

第一章 绪论	3	2.1 关系模型概述	45
1.1 数据库系统概述	3	2.2 关系数据结构及形式化定义 ...	47
1.1.1 数据、数据库、数据库 管理系统、数据库系统	3	2.2.1 关系	47
1.1.2 数据管理技术的产生 和发展	6	2.2.2 关系模式	51
1.1.3 数据库系统的特点	9	2.2.3 关系数据库	52
1.2 数据模型	13	2.3 关系的完整性	52
1.2.1 数据模型的组成要素	14	2.4 关系代数	55
1.2.2 概念模型	15	2.4.1 传统的集合运算	56
1.2.3 最常用的数据模型	21	2.4.2 专门的关系运算	58
1.2.4 层次模型	22	2.5 关系演算	64
1.2.5 网状模型	27	2.5.1 元组关系演算语言 ALPHA	64
1.2.6 关系模型	30	*2.5.2 元组关系演算	69
1.3 数据库系统结构	33	2.5.3 域关系演算语言 QBE	72
1.3.1 数据库系统模式的概念 ...	33	*2.5.4 域关系演算	78
1.3.2 数据库系统的三级 模式结构	33	2.6 小结	79
1.3.3 数据库的二级映象功能 与数据独立性	35	习题	80
1.4 数据库系统的组成	36	本章参考文献	82
1.5 数据库技术的研究领域	39	第三章 关系数据库标准语言 SQL ...	84
1.6 小结	40	3.1 SQL 概述	84
习题	41	3.1.1 SQL 的特点	85
本章参考文献	42	3.1.2 SQL 语言的基本概念	86
第二章 关系数据库	45	3.2 数据定义	87
		3.2.1 定义、删除与修改 基本表	88
		3.2.2 建立与删除索引	90

3.3 查询	91	基本准则	153
3.3.1 单表查询	92	4.2 关系数据库系统的	
3.3.2 连接查询	101	查询优化	157
3.3.3 嵌套查询	106	4.2.1 关系系统及其查询优化	157
3.3.4 集合查询	114	4.2.2 一个实例	159
3.3.5 SELECT 语句的		4.2.3 查询优化的一般准则 ...	161
一般格式	115	4.2.4 关系代数等价变换规则	162
3.4 数据更新	117	4.2.5 关系代数表达式的	
3.4.1 插入数据	117	优化算法	164
3.4.2 修改数据	118	4.2.6 优化的一般步骤	165
3.5 视图	121	4.3 小结	166
3.5.1 定义视图	121	习题	166
3.5.2 查询视图	125	本章参考文献	167
3.5.3 更新视图	126	第五章 关系数据理论	169
3.5.4 视图的作用	128	5.1 问题的提出	169
3.6 数据控制	129	5.2 规范化	172
3.6.1 授权	130	5.2.1 函数依赖	172
3.6.2 收回权限	132	5.2.2 码	173
3.7 嵌入式 SQL	133	5.2.3 范式	174
3.7.1 嵌入式 SQL 的		5.2.4 2NF	174
一般形式	133	5.2.5 3NF	176
3.7.2 嵌入式 SQL 语句与		5.2.6 BCNF	176
主语言之间的通信	134	5.2.7 多值依赖	178
3.7.3 不用游标的 SQL 语句	137	5.2.8 4NF	181
3.7.4 使用游标的 SQL 语句	141	5.2.9 规范化小结	182
3.7.5 动态 SQL 简介	146	5.3 数据依赖的公理系统	183
3.8 小结	148	* 5.4 模式的分解	188
习题	148	5.4.1 模式分解的三个定义 ...	188
本章参考文献	149	5.4.2 分解的无损连接性和	
第四章 关系系统及其查询优化	151	保持函数依赖性	190
4.1 关系系统	151	5.4.3 模式分解的算法	192
4.1.1 关系系统的定义	151	5.5 小结	195
4.1.2 关系系统的分类	152	习题	196
*4.1.3 全关系系统的十二条		本章参考文献	199

第二篇 设计篇

第六章 数据库设计	203	6.4.1 E-R图向关系模型 的转换	230
6.1 数据库设计概述	203	6.4.2 数据模型的优化	231
6.1.1 数据库和信息系统	203	6.4.3 设计用户子模式	232
6.1.2 数据库设计的特点	204	6.5 数据库的物理设计	233
6.1.3 数据库设计方法简述	205	6.5.1 数据库的物理设计 的内容和方法	234
6.1.4 数据库设计的基本步骤	205	6.5.2 关系模式存取方法选择	234
6.2 需求分析	209	6.5.3 确定数据库的存储结构	237
6.2.1 需求分析的任务	209	6.5.4 评价物理结构	237
6.2.2 需求分析的方法	210	6.6 数据库的实施和维护	238
6.2.3 数据字典	211	6.6.1 数据的载入和应用程序 的调试	238
6.3 概念结构设计	213	6.6.2 数据库的试运行	239
6.3.1 概念结构	213	6.6.3 数据库的运行和维护	239
6.3.2 概念结构设计的方法 与步骤	214	6.7 小结	241
6.3.3 数据抽象与局部视图 设计	216	习题	241
6.3.4 视图的集成	224	本章参考文献	242
6.4 逻辑结构设计	229		

第三篇 系统篇

第七章 数据库恢复技术	247	7.8 Oracle 的恢复技术	260
7.1 事务的基本概念	247	7.9 小结	261
7.2 数据库恢复概述	249	习题	261
7.3 故障的种类	249	本章参考文献	262
7.4 恢复的实现技术	251	第八章 并发控制	264
7.4.1 数据转储	252	8.1 并发控制概述	264
7.4.2 登记日志文件 (Logging)	253	8.2 封锁(Locking)	266
7.5 恢复策略	255	8.3 封锁协议	267
7.5.1 事务故障的恢复	255	8.4 活锁和死锁	270
7.5.2 系统故障的恢复	255	8.5 并发调度的可串行性	272
7.5.3 介质故障的恢复	256	8.6 两段锁协议	274
7.6 具有检查点的恢复技术	256	8.7 封锁的粒度	276
7.7 数据库镜像	259	8.7.1 多粒度封锁	276
		8.7.2 意向锁	277

8.8 Oracle 的并发控制	278	*第十一章 数据库管理系统	315
8.9 小结	279	11.1 DBMS 概述	315
习题	280	11.1.1 DBMS 的目标	315
本章参考文献	281	11.1.2 DBMS 的基本功能	317
第九章 数据库安全性	283	11.1.3 DBMS 和操作系统	318
9.1 计算机安全性概论	283	11.1.4 DBMS 和第四代应用 开发环境	318
9.1.1 计算机系统的三类 安全性问题	283	11.2 DBMS 的进程结构和 多线索机制	319
9.1.2 可信计算机系统 评测标准	284	11.2.1 N 方案: DBMS 与应用 程序相融合的方案	319
9.2 数据库安全性控制	288	11.2.2 2N 方案: 一个 DBMS 进程 对应一个用户进程	320
9.2.1 用户标识与鉴别 (Identification & Authentication)	288	11.2.3 N+1 方案: 一个 DBMS 进程对应所有用户 进程	321
9.2.2 存取控制	289	11.2.4 N+M 方案: M 个 DBMS 进程对应 N 个用户 进程	323
9.2.3 自主存取控制 (DAC) 方法	290	11.2.5 多线索 (Multi_Threated) DBMS 的概念	324
9.2.4 强制存取控制 (MAC) 方法	292	11.3 DBMS 系统结构	326
9.2.5 视图机制	293	11.3.1 DBMS 程序模块的组成	327
9.2.6 审计 (Audit)	294	11.3.2 DBMS 的层次结构	329
9.2.7 数据加密	294	11.3.3 DBMS 的运行过程示例	330
9.3 统计数据库安全性	295	11.4 语言处理	332
9.4 Oracle 数据库的安全性 措施	296	11.4.1 语言翻译处理层的任务 和工作步骤	332
9.5 小结	299	11.4.2 解释方法	334
习题	300	11.4.3 预编译方法	335
本章参考文献	301	11.5 数据存取层	336
第十章 数据库完整性	303	11.5.1 数据存取层的系统结构	337
10.1 完整性约束条件	303	11.5.2 数据存取层的功能 子系统	338
10.2 完整性控制	306	11.6 缓冲区管理	341
10.3 Oracle 的完整性	309		
10.4 小结	313		
习题	313		
本章参考文献	313		

11.7 数据库物理组织	343	习题	347
11.8 小结	347	本章参考文献	348
第四篇 新技术篇			
第十二章 数据库技术新发展	353	13.5 对象-关系数据库	406
12.1 数据库技术三个发展阶段	353	13.6 小结	408
12.1.1 第一代数据库系统	354	习题	408
12.1.2 第二代数据库系统 ——关系数据库系统	355	本章参考文献	409
12.1.3 新一代数据库技术的 研究和发展的	357	*第十四章 分布式数据库系统	412
12.2 数据库新技术	363	14.1 概述	412
12.2.1 数据模型的发展	363	14.1.1 什么是分布式数据库 系统	412
12.2.2 数据库技术与其他 相关技术相结合	366	14.1.2 分布式数据库系统 的特点	415
12.2.3 面向应用领域的 数据库新技术	375	14.1.3 分布式数据库系统 的目标	417
12.3 小结	389	14.2 分布式数据库系统的 体系结构	418
习题	389	14.2.1 分布式数据库系统 的模式结构	418
本章参考文献	390	14.2.2 数据分片	419
第十三章 面向对象数据系统	392	14.2.3 分布透明性	421
13.1 面向对象程序设计方法	393	14.2.4 分布式数据库管理 系统	423
13.2 面向对象数据模型	394	14.3 查询处理和优化	425
13.2.1 OO 模型的核心概念	394	14.3.1 一个实例	426
13.2.2 对象与对象标识	395	14.3.2 查询处理和优化要 解决的问题	428
13.2.3 类和类层次	397	14.3.3 查询优化的目标	429
13.2.4 继承	399	14.3.4 连接查询的优化	430
13.2.5 滞后联编	400	14.4 分布事务管理	432
13.2.6 对象的嵌套	401	14.4.1 分布事务的恢复	432
13.3 面向对象数据库语言	402	14.4.2 并发控制	433
13.4 面向对象数据库的模式 演进	403	14.5 分布式数据库的发展 前景和应用趋势	435
13.4.1 模式的一致性	403		
13.4.2 模式演进操作	404		
13.4.3 模式演进的实现	405		

14.6 小结	436	15.4.3 并行操作算法	448
习题	436	15.4.4 并行查询优化	451
本章参考文献	437	15.5 并行数据库的物理组织 ...	452
*第十五章 并行数据库系统	440	15.5.1 并行数据库的数据	
15.1 并行数据库系统概述	440	划分	452
15.2 并行数据库系统的目标	441	15.5.2 并行数据库的数据	
15.3 支持并行数据库的		复制	454
并行结构	441	15.5.3 并行数据库的数据	
15.3.1 SM 并行结构	442	重组	455
15.3.2 SD 并行结构	443	15.6 并行数据库系统与分布式	
15.3.3 SN 并行结构	444	数据库系统的区别	455
15.3.4 三种并行结构比较	445	15.7 小结	456
15.4 并行数据库的并行		习题	458
查询处理技术	445	本章参考文献	458
15.4.1 并行粒度	446	全书参考文献	460
15.4.2 并行化形式	447		