

Database:
Principles, Programming, and Performance
(Second Edition)

数据库原理、 编程与性能

(原书第2版)

HZ BOOKS
华章教育

国外经典教材



Classical Texts From Top Universities

(美) Patrick O'Neil 著
Elizabeth O'Neil

周傲英 俞荣华 季文赟 钱卫宁 等译

 机械工业出版社
China Machine Press

 MORGAN
KAUFMANN



北京华章图文信息技术有限公司

国外经典教材



Classical Texts From Top Universities

(原书第2版)

数据库原理、 编程与性能

*Database:
Principles, Programming,
and Performance (Second Edition)*

(美) Patrick O'Neil 著
Elizabeth O'Neil

周傲英 俞荣华 季文赞 钱卫宁 等译



机械工业出版社
China Machine Press

本书是在波士顿马萨诸塞大学数据库入门和提高等一系列教材的基础上写成的，从理论和实际两方面详细介绍了数据库的设计和实现。本书把重点放在对象-关系模型上，介绍了ORACLE、DB2和INFORMIX系统中普遍采用的新概念，并在结合数据库的基本原理和主要的商业数据库产品的基础上介绍了SQL-99。本书涵盖了关系数据库理论、SQL语言、数据库设计以及数据库完整性、视图、安全性、索引、事务管理等各个方面的内容。

本书是计算机专业学生的一本理想教材，对于数据库设计者和实现者也是一本优秀的参考书。

Patrick O'Neil and Elizabeth O'Neil: Database-Principles, Programming, and Performance Second Edition

Copyright ©1994, 2001 by Morgan Kaufmann Publishers, Inc.

Chinese edition published by arrangement with Morgan Kaufmann.

All rights reserved.

本书中文简体字版由美国Morgan Kaufmann公司授权机械工业出版社独家出版。未经出版者书面许可，不得以任何方式复制或抄袭本书内容。

版权所有，侵权必究。

本书版权登记号：图字：01-2000-2808

图书在版编目（CIP）数据

数据库原理、编程与性能 /（美）奥尼尔（O'Neil, P.），（美）奥尼尔（O'Neil, E.）著；周傲英等译. -北京：机械工业出版社，2002.1

（国外经典教材）

书名原文：Database-Principles, Programming, and Performance Second Edition

ISBN 7-111-09310-0

I. 数… II. ①奥…②奥… ③周… III. 数据库系统 IV. TP311.13

中国版本图书馆CIP数据核字（2001）第067409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：杨海玲

北京市密云县印刷厂印刷·新华书店北京发行所发行

2002年1月第1版第1次印刷

787mm×1092mm1/16·38.25印张

印数：0 001-5 000册

定价：55.00元

凡购本书，如有倒页、脱页、缺页，由本社发行部调换

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，还揭橥了学术的源变，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，有的还专诚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：针对本科生的核心课程，剔抉外版菁华而成“国外经典教材”系列；对影印版的教材，则单独开辟出“经典原版书库”；定位在高级教程和专业参考的“计算机科学丛书”还将保持原来的风格，继续出版新的品种。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。

“国外经典教材”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本科教学

度身订造的。在广泛地征求并听取丛书的“专家指导委员会”的意见后，我们最终选定了这20多种篇幅内容适度、讲解鞭辟入里的教材，其中的大部分已经被M.I.T.、Stanford、U.C. Berkley、C.M.U.等世界名牌大学采用。丛书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，我们的联系方法如下：

电子邮件：hzedu@hzbook.com

联系电话：(010) 68995265

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

(按姓名笔画顺序)

尤晋元	王 珊	冯博琴	史忠植	史美林
张立昂	李伟琴	李师贤	李建中	杨冬青
周克定	周傲英	孟小峰	岳丽华	范 明
高传善	梅 宏	程 旭	程时端	谢希仁
石教英	吕 建	孙玉芳	吴世忠	吴时霖
邵维忠	陆丽娜	陆鑫达	陈向群	周伯生
郑国梁	施伯乐	钟玉琢	唐世渭	袁崇义
裘宗燕	戴 葵			

译者序

本书是一本出色的关于数据库原理与应用的教材。它是作者在波士顿马萨诸塞大学的数据库入门和提高这两门课程教材的基础上编纂而成的。和其他数据库系统教材不同，本书理论与实践并重，每一部分都从基本概念入手，并详细地阐述各个概念在真实的数据库系统中的实现与应用。同时，书中详细比较了各种数据库管理系统的技术差异，并在附录A中给出了两种数据库管理系统的使用方法。于是，学生可以在学习了数据库系统的基本原理之后快速地运用所学的知识。这种对实践的重视同样体现在习题的设计中。作为一本教材，本书在详细介绍了数据库经典技术的同时准确地把握了数据库系统的发展趋势，着重介绍对象-关系模型。于是，在阅读本书以后，读者可以很容易地充分利用现有主流数据库管理系统提供的主要功能。

本书的作者Patrick O'Neil和Elizabeth O'Neil教授都是数据库界知名的学者。他们在数据库领域有很多重要的研究成果，其中部分成果也被编入本书。作者波士顿马萨诸塞大学的计算机科学教授，他们在这本书中融入了多年的教学经验；同时，由于和工业界有着紧密的联系，他们在撰写本书时非常注意理论和实践相结合，并准确地把握了数据库系统的发展趋势。组织严谨、注重实践、内容新颖是本书的三大特色。

本书共分11章和4个附录，依次为：第1章数据库概论，第2章关系模型，第3章基本SQL查询语言，第4章对象-关系SQL，第5章访问数据库的程序，第6章数据库设计，第7章完整性、视图、安全性和目录，第8章索引，第9章查询处理，第10章更新事务，第11章并行和分布式数据库，附录A教学指导，附录B编程细节，附录C SQL语法，附录D集合查询计数。最后，作者还给出了部分习题的解答。

参加本书翻译的人员有：周傲英、俞荣华、季文赞、钱卫宁、贺奇、梁宇奇、王焱、郑仕辉、钱海蕾、张龙、魏黎、范晔、胡江滔。全书由周傲英审校。

限于译者水平，若有疏漏，敬请读者批评指正。

译者
复旦大学计算机系
2001年6月

译者简介

周傲英 计算机软件教授、博士生导师，复旦大学计算机系主任。1993年获博士学位。现任中国计算机学会理事、数据库专业委员会委员、上海计算机学会数据库专业委员会副主任。近年来，曾先后担任PAKDD， DAFAA等国际学术会议程序委员会委员和WAIM'2000的程序委员会主席。目前WAIM已成为在中国召开的一个系列的国际数据库学术会议。周傲英作为发起人之一担任会议顾问委员会成员。作为项目负责人或主要研究人员主持或参加了国家973计划、863计划、国家自然科学基金、国家博士点基金、国防科工委预研计划、上海市科技发展基金及对外合作项目的多项研究和开发工作，取得一系列成果，先后四次获上海市和国家教委科技进步奖。近五年来发表了一批高水平论文以及出版教材/著作五本。1995年入选上海市青年科技启明星计划，1997年获上海高校优秀青年教师荣誉称号，2000年获得国务院特殊津贴，2000年入选教育部“跨世纪优秀人才培养计划”。

俞荣华 硕士研究生。专业为计算机软件与理论，研究方向为数据质量和数据清洗。

季文赟 博士研究生。主要研究方向为数据库和Web数据管理。

钱卫宁 博士生。专业为计算机软件与理论，研究兴趣为数据挖掘和Web挖掘。

序

从1994年问世开始，O'Neil的数据库书籍就已经成为学习、设计或管理关系数据库的标准教材和参考书。本书从理论和实践两方面详细介绍了数据库的设计和实现，包括关系理论，数据库设计，数据库的实现以及性能优化等问题。本书在介绍时，先给出一般的概念，然后结合实际数据库系统中的具体例子加以说明。

第2版反映了自第1版发行后的六年里数据库领域取得的实质性进展。本书把重点放在对象-关系模型；介绍了在Oracle、DB2和Informix系统中普遍采用的新概念；更新了对隔离技术的表述方式。用最新的方法提出性能的问题。对象-关系的介绍是本书的重点：这被认为是SQL数据库语言自最初标准问世以来发生的最重要的变化。本书介绍了SQL-99的设计以及SQL-99与基本原理及主要的商用数据库的关系。

Pat教授和Elizabeth O'Neil教授对数据库设计问题有着广博而精深的见解，在过去的三十年中，他们在数据库领域做出了卓越的贡献。他们桃李满天下，发表了很多基础研究的论文，并与厂商合作开发了一些数据库产品。他们不断创新——在本书中，他们就试图用统一的眼光来看待数据库领域中许多完全不同的概念和趋势。第2版提出一种全新的见解。

《数据库原理、编程与性能》对初学者来说是一本极好的教材，对我们这些不注意数据库领域最新动态的人来说则是易于阅读的最新动态资料，而对需要接受适时教育的数据库设计者和实现者来说，这又是一本不可多得的参考书。

Jim Gray
Microsoft Research

前言

本书第1版介绍了数据库理论的基本原理,探讨了理论和商业应用之间的联系和裂缝。笔者认为有必要让读者了解最新的数据库应用情况,因为在过去的五年中,各种商业数据库产品都发生了巨大的变化。正是数据库领域的这些重要的变化促使我们出版了第2版,但力求保持对原书基本目标的追求。

在第1版中,我们介绍了最新的SQL和商业数据库系统中的应用程序。本书广泛适用于交互式SQL用户、应用程序员、数据库管理员和对数据库感兴趣的学生。在第2版中,我们进一步加强了实际应用与理论基础之间的联系。

一般的书往往只简单罗列了嵌入式SQL的特性,不能满足程序员编写复杂数据库应用软件的需要。厂商的手册能对SQL和数据库编程进行很好的介绍,但缺乏学生和专业人员适应数据库系统和语言变化所需的基础理论。在许多介绍数据库的书籍或手册中并未很好地介绍如下原理:死锁终止(事务需要重做)隐含的问题;实体-联系模型的建立和规范化(以及数据库设计到实际表的转化);在执行访问公共数据的事务时,用户的交互问题(事务活动时进行交互是危险的);索引和查询优化如何影响查询性能方面的考虑——一个数据库管理员应该重视的问题;等等。

本书可以看做是一本完整的技术导论,数据库管理员、应用程序员和资深的SQL用户都会需要它。虽然数据库管理员要求比应用程序员了解更多的东西,但是,如果后者对前者必需掌握的知识有一个总体的把握,那么他的工作会变得更有效。类似的情况也适用于那些认真的交互式SQL用户。任何与数据库相关的工作人员除了牢固掌握SQL外,还应该了解数据库的逻辑设计、数据的物理存储、索引、安全和性能价格比等方面的知识。

本书的作者都是计算机科学的教授,在与数据库公司合作和数据库应用方面有着丰富的经验。有兴趣的读者可以访问主页 www.cs.umb.edu/~poneil和www.cs.umb.edu/~eoneil。

本书的使用

本书是在波士顿马萨诸塞大学的数据库入门和提高这两门课程的基础上写成的。第一门课程介绍数据库的基本原理,本书的前六章取自该课程。第二门课程介绍更高深的数据库概念和性能价格比的问题,本书第7~第10章的内容取自该课程。

在学习本书的时候,读者可根据学习的经验和目的自己安排学习的顺序和重点。举例来说,有基础的读者可以根据自己的兴趣从后面的章节开始学习,只在需要的时候查阅前六章的内容(图1显示了各章之间的关系)。本书在读者掌握前面的概念的基础上才展开新的知识,所以有经验的读者可以根据需要从适当的地方学起。

本书既可用作专业人员的参考书和指南书,也可用作本科学生数据库课程一学期或两学期的入门教材。它涵盖了数据库领域从基础理论和基本概念到发展前沿的所有内容。本书将基本的SQL语言和关系数据库基础放在一起介绍。本书的例子取自ORACLE, INFORMIX和DB2等数据库,我们用它们解释概念,并通过比较这些成功系统中采用的不同方法阐明性能-价格

比问题。每章中的关键性问题通过编程示例和习题得以深入。书后的四个附录提供了补充性的背景介绍。为了方便自学,本书最后还提供了部分习题的答案。

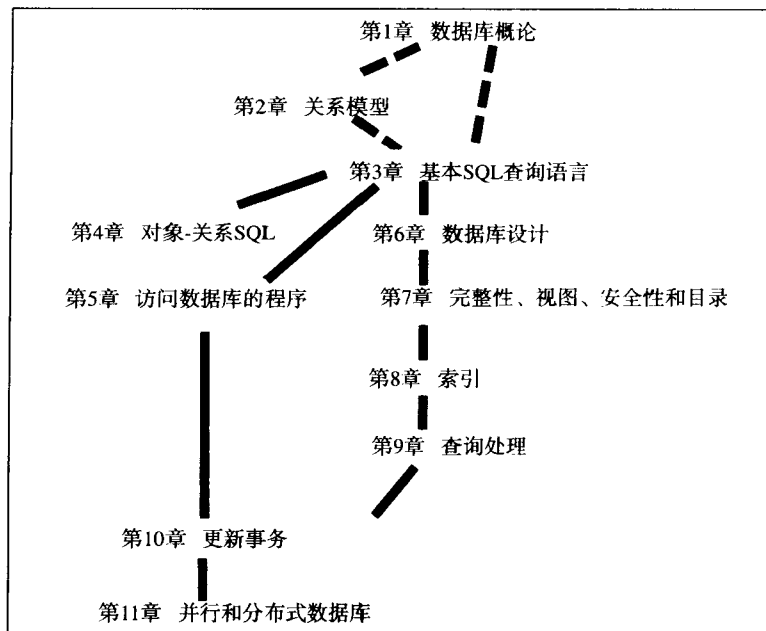


图1 各章之间的联系

新增内容讨论: 对象-关系模型

自从本书第1版出版以来,SQL语言增加了许多新的特性。第3章中新增的几节介绍了这些特性。从根本上说,对象-关系模型已经变成新的数据库标准。所有新的对象-关系数据库管理系统产品都向后兼容,以便支持关系模型,我们认为这一变化正在彻底改变数据库产业。数据库表的设计以及利用SQL语句访问数据的方法将发生很大变化。

对象-关系模型成为主流的过程是值得研究的。INFORMIX公司从1996年收购Illustra产品起,就着力研究把对象-关系特性融入到INFORMIX产品中。(Illustra,后来被称做Montage,参见第1版3.11节;在第2版中,我们大大扩充了这方面的内容,独立编写成第4章。)1997年,ORACLE公司推出的ORACLE第8版,支持一整套对象-关系特性,该数据库被普遍认为是具有革命性意义的产品。从那时起,在数据库产业界,从关系模型到对象-关系模型的转变就已是大势所趋了。最近,一向以良好的对象-关系编程接口著称的IBM的DB2 UDB开始将对象-关系数据模型融入设计和交互层面^[2]。不幸的是,由于还没有一个统一的对象-关系SQL标准,各个产品在语法上差别很大,应用程序不可移植。所幸的是,经过多年的努力,ANSI SQL-3终于在1999年发布了对象-关系数据库的最终标准——SQL-99^[3]。

由于有两种完全不同的表述数据的模型,整个知识领域就被分割开来,这在教学上是个很大的挑战。现在,绝大部分的商业数据库软件使用纯关系模型。这意味着如果不仔细对对象-关系概念与关系概念加以区分,用户对前者会产生混淆。和第1版一样,在第2版中我们仍然在第2章中介绍关系模型,而仅在第3章中介绍最新的关系SQL的表现形式。同时,我们扩充了第1版3.11节中对象-关系领域的内容,单独设为第4章,在这一章中,我们分别介绍了

ORACLE和INFORMIX中对象-关系的标准。因为这两个产品间存在很大差异，我们把第4章中每一节分成两个部分，并行地介绍ORACLE和INFORMIX。我们没有介绍DB2 UDB关于对象-关系的语法规则，因为在编写第4章的时候，DB2对象-关系模型正处于开发阶段。

在详细介绍了对象-关系的概念和产品的使用方法之后，接下去的几章在内容上与第4章基本无关。这样做的原因是：迄今为止有些问题在对象-关系领域还没完全搞明白，如第6章数据库的逻辑设计；而有些问题看起来与对象-关系的概念（PL/SQL和SPL除外）无关，如第10章的更新事务。一个专业人员在仔细阅读第4章之后，就能用对象-关系SQL的知识扩充大部分后续章节的内容；而那些目前不与对象-关系数据库打交道的读者，可能不想立即接触这些内容。从图1中可以看到，第4章与后续各章没有前后联系。各章的内容相对独立。（第7章涉及对象-关系模式下对象目录表的内容，但读者在学习时可以跳过这一段。）在下一版中，对象-关系的概念可能会遍布每个章节，因为那时对象-关系数据库将会广泛使用，不过现在还为时过早。

在编写第2版的时候，我们不得不做一些困难的选择，最难的莫过于选择合适的数据库产品。我们舍弃了面向对象数据库管理系统（OODBMS），因为虽然面向对象数据库的潜在商业价值被一致看好，但却从来没有真正实现过。现有的OODBMS产品只是存在而已。虽然Microsoft SQL Server正越来越引起大家的注意，我们还是舍弃了对它的介绍，因为我们的主要目的是介绍新的对象-关系数据库产品的特性，而SQL Server不支持这些特性。INGRES系统是第1版的重点内容，因为它是校园中使用最广泛的系统。在第2版中，我们删去了这部分内容，因为许多流行的数据库产品都提供了针对校园市场的低价版本。

每一章的变化

在下面对各章的描述中，我们简要介绍了在第2版中第一次引用的材料。

第1章 数据库概论。新增了对对象-关系模型的探讨。

第2章 关系模型。除了澄清第1版的若干细节外，本章基本上未作改动。我们在第2、3章中介绍关系模型，从第4章开始介绍对象-关系模型。

第3章 基本查询语言SQL。在本版中，我们介绍了很多SQL的新特性。3.6节介绍了一些现有数据库产品尚不支持的“高级SQL语法”。它包括INTERSECT[ALL]和EXCEPT[ALL]操作和新的tableref定义（图3-11），提供了诸如[INNER|LEFT|RIGHT|FULL][OUTER]]连接和对ON搜索条件的连接。因为没有统一的SQL标准，本书采用了一种一般的形式，我们称之为“基本SQL”，它提供被所有的数据库产品普遍采用的一系列特性。

第4章 对象-关系SQL。这是全新的一章。我们列表说明各节和各小节的主要内容。每节的最后是对ORACLE和INFORMIX的每一个特性的分别说明和相互比较。

4.1 引言：对象-关系数据库的定义和历史。

4.2 ORACLE中的对象类型，INFORMIX中的行类型，用对象(行)类型定义表，对象嵌套，访问列的点标号，对象数据缺乏封装，ORACLE中REF的介绍，INFORMIX中的类型继承。

4.3 汇集(collection)类型。ORACLE中有两种汇集类型：嵌套表(表类型的列值)和VARRAY(数组类型的列值)。INFORMIX有三种汇集类型：集合、多重集合(无

序,但允许有重值)和列表(有序)。这两种数据库都允许即席查询,从汇集中检索数据,并插入和更新汇集。

4.4 用户自定义函数(UDF)和方法。ORACLE支持一种过程性SQL语言PL/SQL,而INFORMIX支持SPL。在这两个数据库中,UDF可以用过程性语言编写(或用一种嵌入式SQL语言,比如Java)。UDF也被称为SQL中的内置函数。方法是ORACLE支持的一种UDF,它被定义成对象类型的一部分。

4.5 外部函数和用户定义类型包(UDT)。我们概述了数据库系统中打包一组外部函数的用户定义类型的功能。外部函数用类C语言编写并在数据库服务器上执行的UDF。这样一种包在ORACLE中称为Catridge,在INFORMIX中称为DataBlade,而在DB2 UDB中称为Extender。

以下各章的章号比第1版中的相应章号增加了1,因为在第2版中增加了新的第4章。

第5章 访问数据库的程序。在第2版中,我们提供了更多的编程实例,特别是关于事务方面的。我们改进了错误处理的内容。我们介绍了最新的ORACLE语法,特别是ORACLE动态SQL语法(包括SQLSTATE)。同时还将介绍DB2 UDB(代替第1版中的INGRES)。

第6章 数据库设计。这是指数据库的逻辑设计,包括E-R模型和规范化。本章阐述了很多定义和证明,增加了很多实例说明。

第7章 完整性、视图、安全性和目录。本章新介绍了大量Create Table和Alter Table语句的标准子句,重写了触发器部分,其中的例子来自于ORACLE和DB2 UDB。对可更新视图的限制的介绍变化较大,对查询视图不再有任何限制。系统目录部分稍有扩充,增添了一小节对象-关系目录的内容。

第8章 索引。磁盘速度、容量、价格和内存价格都已更新。还更新了ORACLE中Create Tablespace语句。对INGRES索引功能(ISAM、散列等)不再做专门介绍。新增的内容有ORACLE索引组织表和表聚簇(主要是散列聚簇)。为了处理新的ORACLE散列聚簇结构,对溢出链接做了一些修改。新增了DB2 UDB的索引结构。

第9章 查询处理。虽然在这一章的前几节中已经介绍了新的DB2产品的特性,在一些审阅者的坚持下,本书中我们保留了第1版介绍IBM大型机上DB2的查询特性的内容。正是这些特性导致了集合查询基准测试的测试结果。因为大部分DB2查询功能仍代表了当今的最高技术,而第1版中对查询过程的解释即使在现在看来仍是非常详尽的,所以这部分内容仍然用来向学生解释如何完成查询的重要概念。事实上,查询处理器的内容仍与当前大型机DB2产品(DB2 for OS/390)有关。

第10章 本章的一些定义和证明做了较大改动。在10.5节隔离层次定义中,新增了推荐读物[1]中新的研究成果。

第11章 只做了一些细微的改动。

万维网上的支持

本书的主页 http://www.mkp.com/books_catalog/1-55860-438-3.asp 提供以下功能:链接作者的主页;提供为几种数据库产品编写的创建和载入数据库的脚本;提供本书最新的勘误表,示例程序和讲课幻灯片;为教师提供不带黑点标记(·)的习题的答案。读者的建议、意见和勘误

可通过发电子邮件到 poneil@cs.umb.edu 或者 eoneil@cs.umb.edu 与我们联系。

推荐读物

- [1] Hal Berenson, Phil Bernstein, Jim Gray, Jim Melton, Elizabeth O'Neil, and Patrick O'Neil. "A Critique of ANSI SQL Isolation Levels," *ACM SIGMOD Proceedings*, May 1995, pp.1-10.
- [2] Michael Carey, Don Chamberlin, et al. "O-O, What Have They Done to DB2?" *Proceedings of the 1999 VLDB Conference*.
- [3] Andrew Eisenberg and Jim Melton. "SQL:1999, formerly known as SQL3." *SIGMOD Record*, vol.28, no.1, March 1999, pp.131-138.

目 录

出版者的话	
专家指导委员会	
译者序	
序	
前言	
第1章 数据库概论	1
1.1 基本的数据库概念	1
1.2 数据库用户	4
1.3 关系数据库管理系统和对象-关系数据库管理系统概述	5
1.4 小结	15
第2章 关系模型	16
2.1 CAP 数据库	16
2.2 数据库各部分的命名	18
2.3 关系规则	21
2.4 键、超键和空值	23
2.5 关系代数	27
2.6 集合运算	28
2.7 自然关系运算	31
2.8 运算依赖	41
2.9 综合例子	42
2.10 其他关系运算	46
推荐读物	49
习题	49
第3章 基本SQL查询语言	53
3.1 引言	53
3.2 创建数据库	56
3.3 简单的Select语句	58
3.4 子查询	65
3.5 UNION运算符和FOR ALL 条件	74
3.6 高级SQL语法	80
3.7 SQL中的集合函数	86
3.8 SQL中的行分组	90
3.9 SQL Select 语句的完整描述	95
3.10 Insert、Update和Delete语句	105
3.11 Select 语句的能力	107
推荐读物	116
习题	116
第4章 对象-关系SQL	123
4.1 引言	123
4.2 对象和表	125
4.2.1 ORACLE 中的对象类型	125
4.2.2 INFORMIX 中的对象行类型	136
4.2.3 对象和表小结	140
4.3 汇集类型	142
4.3.1 ORACLE中的汇集类型	142
4.3.2 INFORMIX 中的汇集类型	153
4.3.3 汇集类型小结	158
4.4 过程SQL、用户定义函数和方法	159
4.4.1 ORACLE PL/SQL过程、用户定义函数和方法	160
4.4.2 INFORMIX中的用户定义函数	171
4.4.3 用户定义函数小结	179
4.5 外部函数和打包的用户定义类型	180
推荐读物	183
习题	184
第5章 访问数据库的程序	186
5.1 C 语言中嵌入式SQL 的介绍	188
5.2 条件处理	195
5.3 一些通用的嵌入式SQL语句	202
5.4 事务的编程	206
5.5 过程性SQL程序的能力	218
5.6 动态SQL	221
5.7 一些高级的编程概念	229
推荐读物	232
习题	233
第6章 数据库设计	238
6.1 E-R概念介绍	239

6.2 E-R模型的细节	245	9.7 连接表的方法	416
6.3 其他E-R概念	250	9.8 磁盘排序	426
6.4 案例学习	253	9.9 查询性能基准程序：样例研究	430
6.5 规范化：基础知识	255	9.10 查询性能测量	434
6.6 函数依赖	259	9.11 性能价格比评估	445
6.7 无损分解	272	推荐读物	448
6.8 范式	276	习题	448
6.9 其他设计考虑	288	第10章 更新事务	456
推荐读物	289	10.1 事务经历	459
习题	290	10.2 交错的读写操作	463
第7章 完整性、视图、安全性和目录	296	10.3 可串行化和前趋图	467
7.1 完整性约束	296	10.4 用来保证可串行性的锁机制	472
7.2 建立视图	313	10.5 隔离级别	476
7.3 安全性：SQL中的Grant语句	320	10.6 事务恢复	482
7.4 系统目录和模式	322	10.7 恢复细节：日志格式	484
推荐读物	329	10.8 检查点	489
习题	329	10.9 介质恢复	493
第8章 索引	335	10.10 性能：TPC-A 基准测试	494
8.1 索引的概念	335	推荐读物	501
8.2 磁盘存储	338	习题	502
8.3 B 树索引结构	348	第11章 并行和分布式数据库	505
8.4 聚簇索引和非聚簇索引	361	11.1 多CPU 体系结构	505
8.5 散列主索引	367	11.2 CPU价格与性能曲线	508
8.6 向靶上的空位随机投掷飞镖问题	373	11.3 无共享式数据库体系结构	509
推荐读物	377	11.4 查询并行性	515
习题	378	推荐读物	517
第9章 查询处理	382	习题	517
9.1 基本概念	383	附录A 教学指导	518
9.2 表空间扫描和I/O代价	387	附录B 编程细节	529
9.3 DB2 中的简单索引存取	391	附录C SQL语法	534
9.4 过滤因子和统计	399	附录D 集合查询计数	557
9.5 匹配索引扫描、复合索引	402	习题解答	559
9.6 多重索引存取	409		

第1章 数据库概论

本章介绍本书的主要思想和定义。我们描述了数据库的基本概念及数据库系统的典型用户，然后概述与数据库管理系统相关的概念和特性。

1.1 基本的数据库概念

数据库管理系统——简称为数据库系统或DBMS——是一种软件产品，它把一个企业的数以记录的形式在计算机中保存起来。举例来说，批发商往往会用一个数据库管理系统保存销售记录（交易的操作数据）；大学可以用数据库管理系统保存学生的记录（学费、成绩等）；大部分大型图书馆利用数据库系统保存藏书清单和出借记录，提供主题、作者和题目等多种类型的索引；所有的航空公司都利用数据库系统管理航班和提供订票服务；州机动车管理部门利用数据库系统管理驾驶员执照、登记车辆。Tower唱片公司用数据库系统来管理库存，打印所有的磁带和CD，并为顾客提供查询唱片的功能。像这样为一个共同的目的而保存起来的所有数据的集合称为数据库。数据库中的记录通常保存在磁盘上（一种在断电时不丢失保存信息的低速存取介质），一般只在访问时才把记录从磁盘载入内存。

一个数据库管理系统能同时管理多个数据库。举例来说，一所大学可能拥有一个登记学生的数据库和一个图书馆数据库。两个数据库之间没有共享的数据（虽然可能有一部分重复信息，因为一部分读者是学生）。不同的用户可以通过同一个数据库管理系统访问这两个数据库。

1. 数据库的历史

为了访问数据库中的信息，已经开发出许多方法。回顾历史，有两个产品为组织信息提供了两种截然不同的数据模型：1968年IBM发布的IMS和20世纪70年代Cullinet Software 的IDMS。IMS提出了不同类型的记录通过层次结构相互联系的层次数据模型。例如，一个银行数据库系统可以把公司实体记录和诸如总部地址、电话号码这样的信息放在层次结构的顶部；接下来是银行的各个业务部门；在每个部门分支下，是该部门的出纳员和其他职员记录。当要查询某个出纳员时，程序就会沿着各个分层导航。另一方面，在CODASYL委员会数据库任务组1971年发表的报告的基础上诞生的IDMS，被称为网状模型。网状模型是层次模型的一个推广，某一级的一个记录集合在上一级中可能对应两个不同的包含层次（containing hierarchy）。

当然，IMS和IDMS还有许多我们没有提到的特性。简单地说，层次模型把数据组织成一棵根在上、叶在下的有向树。网状模型把数据组织成无环有向图，使得网状模型更容易表达现实世界中的数据结构。这些产品的主要缺点是对数据的查询很难执行，一般需要熟悉复杂的数据导航结构的专业程序员编写相应程序。今天，仍然有相当多的公司在使用这两种数据库。IMS仍然是IBM重要的利润来源。但是，这些使用中的IMS和IDMS已经是“遗产系统”了，而且，很难把这些系统转化成现代的数据模型。虽然某些公司用原有的系统已经足够，但任何想安装新系统的公司都会选择一个支持更新的数据模型的数据库管理系统。