



国外经典教材·计算机科学与技术

PEARSON
Prentice
Hall

Database Concepts

Second Edition

数据库原理 (第2版)

(美) David M. Kroenke 著
郭 平 译



清华大学出版社

国外经典教材·计算机科学与技术

数据库原理

(第2版)

(美)David M. Kroenke 著

郭 平 译

清华大学出版社

北 京

Simplified Chinese edition copyright © 2005 by PEARSON EDUCATION ASIA LIMITED and TSINGHUA UNIVERSITY PRESS.

Original English language title from Proprietor's edition of the Work.

Original English language title: Database Concepts, Second Edition by David M. Kroenke, Copyright © 2005

EISBN: 0-13-145141-3

All Rights Reserved.

Published by arrangement with the original publisher, Pearson Education, Inc., publishing as Prentice Hall.

This edition is authorized for sale only in the People's Republic of China (excluding the Special Administrative Region of Hong Kong and Macao).

本书中文简体翻译版由培生教育出版集团授权给清华大学出版社在中国境内(不包括中国香港、澳门特别行政区)出版发行。

北京市版权局著作权合同登记号 图字: 01-2004-6473

版权所有, 翻印必究。举报电话: 010-62782989 13501256678 13801310933

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。

图书在版编目(CIP)数据

数据库原理(第2版)/(美)克罗恩科(Kroenke, D.M.)著; 郭平译. —北京: 清华大学出版社, 2005.6

书名原文: Database Concepts, Second Edition

(国外经典教材·计算机科学与技术)

ISBN 7-302-10581-2

I. 数… II. ①克…②郭… III. 数据库系统—教材 IV. TP311.13

中国版本图书馆 CIP 数据核字(2005)第 015154 号

出版者: 清华大学出版社 地址: 北京清华大学学研大厦
http://www.tup.com.cn 邮编: 100084
社总机: 010-62770175 客户服务: 010-62776969

组稿编辑: 曹 康

文稿编辑: 李 阳

封面设计: 久久度文化

版式设计: 康 博

印刷者: 北京牛山世兴印刷厂

装订者: 三河市新茂装订有限公司

发行者: 新华书店总店北京发行所

开本: 185×260 印张: 16.75 字数: 347 千字

版次: 2005 年 6 月第 1 版 2005 年 6 月第 1 次印刷

书号: ISBN 7-302-10581-2/TP·7176

印数: 1~5000

定 价: 29.00 元

出版说明

近年来,我国的高等教育特别是计算机学科教育,进行了一系列大的调整 and 改革,急需一批门类齐全、具有国际先进水平的计算机经典教材,以适应当前我国计算机科学的教學需要。通过使用国外先进的经典教材,可以了解并吸收国际先进的教学思想和教学方法,使我国的计算机科学教育能够跟上国际计算机教育发展的步伐,从而培育出更多具有国际水准的计算机专业人才,增强我国计算机产业的核心竞争力。为此,我们从国外知名的出版集团 Pearson 引进这套“国外经典教材·计算机科学与技术”教材。

作为全球最大的图书出版机构,Pearson 在高等教育领域有着不凡的表现,其下属的 Prentice Hall 和 Addison Wesley 出版社是全球计算机高等教育的龙头出版机构。清华大学出版社与 Pearson 出版集团长期保持着紧密友好的合作关系,这次引进的“国外经典教材·计算机科学与技术”教材大部分出自 Prentice Hall 和 Addison Wesley 两家出版社。为了组织该套教材的出版,我们在国内聘请了一批知名的专家和教授,成立了一个专门的教材编审委员会。

教材编审委员会的运作从教材的选题阶段即开始启动,各位委员根据国内外高等院校计算机科学及相关专业的现有课程体系,并结合各个专业的培养方向,从 Pearson 出版的计算机系列教材中精心挑选针对性强的题材,以保证该套教材的优秀性和领先性,避免出现“低质重复引进”或“高质消化不良”的现象。

为了保证出版质量,我们为这套教材配备了一批经验丰富的编辑、排版、校对人员,制定了更加严格的出版流程。本套教材的译者,全部来自于对应专业的高校教师或拥有相关经验的 IT 专家。每本教材的责编在翻译伊始,就定期不间断地与该书的译者进行交流与反馈。为了尽可能地保留与发扬教材原著的精华,在经过翻译、排版和传统的三审三校之后,我们还请编审委员或相关的专家教授对文稿进行审读,以最大程度地弥补和修正在前面一系列加工过程中对教材造成的误差和瑕疵。

由于时间紧迫和受全体制作人员自身能力所限,该套教材在出版过程中很可能还存在一些遗憾,欢迎广大师生来电来信批评指正。同时,也欢迎读者朋友积极向我们推荐各类优秀的国外计算机教材,共同为我国高等院校计算机教育事业贡献力量。

清华大学出版社

国外经典教材·计算机科学与技术

编审委员会

主任委员:

孙家广

清华大学教授

副主任委员:

周立柱

清华大学教授

委员(按姓氏笔画排序):

王成山

天津大学教授

王 珊

中国人民大学教授

冯少荣

厦门大学教授

冯全源

西南交通大学教授

刘乐善

华中科技大学教授

刘腾红

中南财经政法大学教授

吉根林

南京师范大学教授

孙吉贵

吉林大学教授

阮秋琦

北京交通大学教授

何 晨

上海交通大学教授

吴百锋

复旦大学教授

李 彤

云南大学教授

杨宗源

华东师范大学教授

沈钧毅

西安交通大学教授

邵志清

华东理工大学教授

陈 纯

浙江大学教授

陈 钟

北京大学教授

陈道蓄

南京大学教授

周伯生

北京航空航天大学教授

孟祥旭

山东大学教授

姚淑珍

北京航空航天大学教授

徐佩霞

中国科学技术大学教授

徐晓飞

哈尔滨工业大学教授

秦小麟

南京航空航天大学教授

钱培德

苏州大学教授

曹元大

北京理工大学教授

龚声蓉

苏州大学教授

谢希仁

中国人民解放军理工大学教授

前 言

Colin Johnson 是西雅图一家小型制造厂的产品管理员。几年前，Colin 计划构建一个数据库来跟踪产品包中的产品。一开始，他运用电子制表软件来完成这一工作，却无法通过该表获取所需的报表。Colin 听说了 Microsoft Access，希望该软件能解决问题。经过几天的尝试，他发现无法掌握 Access 的使用方法，于是购买了一些流行的 Access 书籍并努力学习。不过最终他还是放弃了，于是他聘请了一个顾问，该顾问创建了一个大致可满足 Colin 要求的应用程序。一段时间以后，Colin 想对此应用程序作一些改动，但他却不敢作这样的尝试。

Colin 是个成功的商人，他能够主动地去实现他的目标。作为一个老练的 Windows 用户，他可以通过自学掌握 Excel、PowerPoint 以及大量面向产品的应用程序。但他在使用 Access 来解决问题时却停滞不前。Colin 认为“我确信我可以做到，但我没有更多的时间”。这样的事件非常引人注目，因为在过去的 10 年内上演了无数次。

Microsoft、Oracle 和其他数据库管理系统(DBMS)厂商都意识到了这一情况，他们投入了数百万美元来创建更好的图形界面、数以百计的多面板向导，以及众多示例应用程序。遗憾的是，这样的努力无法从根本上解决问题。事实上，许多用户并不清楚向导可以实现哪些功能。一旦这些用户需要修改数据结构或是组件(例如窗体和查询)，就会陷入麻烦中，这令他们措手不及。如果了解底层结构，这些用户就只能绞尽脑汁，却徒劳无功。最终也只能得到一些设计糟糕的数据库和应用程序，而无法满足用户的要求。

那么为什么像 Colin 这样的用户可以学会使用文字处理软件或电子数据表这样的产品，却无法学会使用 DBMS 产品呢？其中一个主要原因就是许多人都不熟悉数据库的基本概念。每个人可能都知道段落和边距，却无法理解关系的概念。其次，他们可能觉得使用 DBMS 产品一定比了解数据库概念本身更容易。“我们想做的只是跟踪一些内容，为什么实现起来却这么困难？”如果了解关系模型，在存储数据之前将一张销售发票分成 5 个单独的表就可能使业务用户产生迷惑。

基本概念

当今的技术特点是如果不学习基本概念，那么就不可能成功地利用 DBMS。凭

借多年来为业务用户开发数据库的经验，我认为数据库的基本概念主要包括以下内容：

- 关系模型的基本概念
- 结构化查询语言(SQL)
- 数据建模
- 数据库设计
- 数据库管理

像 Colin 这样的用户(或是将接受类似工作的学生)不需要像信息系统专家那样深入地了解这些主题。因此，本书只介绍了一些基本的概念——这些内容对于像 Colin 这样的用户(创建和使用小型数据库的用户)来说至关重要。我已经在 *Database Processing* 一书中重写、简化和删除了一些内容¹。不过在本书中，也力求讨论准确，不会产生误导。即使学生们已学习过更高级的数据库课程，也仍然可从本书中获益。

独立于 DBMS 产品的概念

本书假设学生没有使用过任何特定的 DBMS 产品。虽然附录中介绍了如何使用 Access，但所有的概念都适合于 DBMS 产品。通过这种方式，学生们可以理解所有数据库的基础知识——从小型的 Access 数据库到大型的 Oracle 或 DB2 数据库。

而且这一方法也避免了一个常见的问题。在同时介绍概念和产品时，学生容易将概念与产品特性、功能相混淆。例如参照完整性约束。在讲授基础理论时，学生都知道在某些情况下，一个表中的列值必须总是由另一个表中的列值提供。学生也将知道这一约束如何出现在关系定义的上下文中，以及 DBMS 或应用程序如何强制执行这一约束。如果结合具体的 DBMS 教学，如 Access，那么学生将学到的知识就可能是：在某些情况下选取复选框，而在其他情况下不选取的具体操作。这样很容易导致在介绍产品特性的同时，使学生们淡忘了数据库的基础理论。

然而这并不是说在本书中将不使用 DBMS。相反，学生们可以通过使用企业版 DBMS 产品来熟悉这些基本概念。本书假定您将参阅其他书籍或资料来学习如何使用 DBMS 的特性和功能，Prentice-Hall 提供了有关 Microsoft Access 2003 的大量辅助读物，可以结合本书一起学习。

1 David Kroenke 所著的 *Database Processing*(第 9 版)一书(Upper Saddle River, NJ: Prentice-Hall,2004)。

复习题、练习题和实践项目

学生能否学会最终运用所学的知识非常重要，因此每章都提供了一些复习题、练习题和三个贯穿本书始终的项目。如果学生阅读并理解了每一章的内容，就可以知道复习题的答案。练习题要求学生将每章所讲的概念应用到具体的小问题或任务中。

第一个项目 Garden Glory 介绍了一个向个人或企业提供园艺服务的合伙公司的数据库，并讨论了该数据库的开发和使用。第二个项目 James River 珠宝行分析了一家零售店为支持针对的购买者而设计的程序的数据库需求。第三个项目介绍了 Mid-Western 大学化学系实验设备的调度。本书的所有章节和附录中都包括了这 3 个项目。在每个实例中，都要求学生将各章中学到的知识运用到项目中。教师们将从教师手册中获得有关这些项目的使用的更多信息，可以从 Web 站点(www.prenhall.com/kroenke)的教师专区获取相关数据库和示例数据。

对第 1 版的改进

与第 1 版相比，本书主要有两点改进。首先，关于规范化化的讨论已经转向于如何为关系的规范化提供说明性的过程。在本书的第 1 版中就曾经介绍过规范化的原则，但将原则的应用留给了学生。本版的第 2 章介绍了学生使用规范化的 4 个步骤。这一改变不仅可以使规范化更加简单，还可以使规范化的原则更易于理解。

其次，新添的第 7 章介绍了如下高级主题的概述。

- Web 数据库处理
- 使用 XML 处理数据库
- 分布式数据库处理
- OLAP 和数据挖掘
- 对象关系数据库管理

所有这些讨论的目的都只是将这些主题的本质介绍给学生，从而增强学生的数据库基本知识，并且鼓励他们在重要的主题领域继续努力学习。

主要内容

本书共包括 7 章和 1 个附录。第 1 章解释了使用数据库的原因、数据库的组成以及开发数据库的方法。学生们将学习数据库及其应用程序的用途、数据库相对于电子表格列表的差别和优势。第 2 章介绍了关系模型，定义了基本的关系术语，同

时还介绍了规范化原则的基本概念，并描述了规范化过程。

第 3 章讲述了基本的 SQL 语句。描述了定义数据的基本 SQL 语句，如 SQL SELECT 和数据修改语句。本书并不介绍高级的 SQL 语句，只讲述一些核心的语句。

接下来的 3 章讨论了数据库设计和管理。第 4 章使用实体-关系(E-R)模型解决数据建模问题，其中包括对数据建模的需求、基本的 E-R 术语和概念，还提供了一个简短的 E-R 建模示例应用程序。第 5 章讲述了数据库设计，解释了规范化的基本概念。第 4 章示例中的数据模型在第 5 章中则被转换为关系设计。

第 6 章讨论了数据库管理。该章讲述了数据库管理的需求，概述了并发控制、安全性和备份恢复技术。这些主题对所有的数据库都很重要，对个人的单用户数据库也是如此。事实上，这些主题在某些方面对于私人数据库更为重要，因为它们没有专业的数据库管理员来确保关键任务的执行。

最后，第 7 章介绍了前面所描述的高级主题。在附录中介绍了 Microsoft Access。这一简短的介绍可以帮助学生结合 Access 技术文档文件创建简单的数据库和应用程序组件。

回顾与展望

在过去的 30 多年中，我发现数据库和数据库应用程序的开发是一项令人愉快、有益的工作。我相信，未来数据库的数量、大小以及重要性都将有大幅度的提高，该领域将取得更大的成就。在此希望本书中所提及的概念、知识和技术有助于学生更好地进行数据库项目的开发。

David M. Kroenke
西雅图，华盛顿区

目 录

第 I 部分 数据库基础

第 1 章 数据库简介	3
1.1 使用数据库的原因	3
1.1.1 关于列表的问题	4
1.1.2 使用关系数据库	6
1.1.3 关系表的处理	11
1.2 数据库系统的概念	12
1.2.1 数据库	13
1.2.2 DBMS	14
1.2.3 应用程序	15
1.2.4 桌面数据库系统和企业数据库系统的比较	18
1.3 小结	20
1.4 复习题	21
1.5 练习题	22
1.6 Garden Glory 项目	23
1.7 James River 珠宝行项目	23
1.8 Mid-Western 大学化学系项目	24
第 2 章 关系模型	25
2.1 关系	25
2.1.1 关系与非关系的示例	26
2.1.2 术语说明	27
2.2 键的类型	28
2.2.1 复合键	29
2.2.2 主键与候选键	29
2.2.3 外键与参照完整性约束	30
2.2.4 代理键	33
2.3 函数依赖与规范化	36
2.3.1 函数依赖	36
2.3.2 再论主键与候选键	38
2.3.3 规范化	38

2.3.4	关系设计原则	39
2.3.5	规范化过程	40
2.3.6	规范化的示例	42
2.4	空值问题	47
2.5	小结	48
2.6	复习题	49
2.7	练习题	51
2.8	Garden Glory 项目	51
2.9	James River 珠宝行项目	53
2.10	Mid-Western 大学化学系项目	54
第 3 章	结构化查询语言	57
3.1	示例数据库	57
3.2	用于数据定义的 SQL 语句	59
3.2.1	使用表约束定义主键	61
3.2.2	使用表约束定义外键	62
3.2.3	向 DBMS 提交 SQL 语句	63
3.2.4	DROP 语句	64
3.3	SQL 关系查询语句	64
3.3.1	从单个表中读取指定列	65
3.3.2	从单个表中读取指定行	66
3.3.3	从单个表中读取指定行和指定列	67
3.3.4	在 WHERE 子句中指定范围、使用通配符和空值	68
3.3.5	对结果进行排序	70
3.3.6	SQL 内置函数	71
3.3.7	内置函数和分组	73
3.3.8	使用子查询处理多个表	74
3.3.9	使用连接处理多个表	75
3.3.10	外部连接	78
3.4	修改关系数据的 SQL 语句	80
3.4.1	插入数据	80
3.4.2	修改数据	81
3.4.3	删除数据	81
3.5	小结	82
3.6	复习题	83
3.7	练习题	85

3.8 Garden Glory 项目	86
3.9 James River 珠宝行项目	87
3.10 Mid-Western 大学化学系项目	88

第 II 部分 数据库设计和管理

第 4 章 数据建模与实体-关系模型	93
4.1 需求分析阶段	94
4.2 实体-关系数据模型	94
4.2.1 实体	95
4.2.2 属性	96
4.2.3 标识符	96
4.2.4 关联	96
4.2.5 二元关联的 3 种类型	97
4.2.6 递归关联	98
4.2.7 实体-关系图	98
4.2.8 弱实体和 ID 依赖实体	99
4.3 UML 实体-关系模型	101
4.3.1 UML 形式的实体和关系	101
4.3.2 UML 形式的弱实体	102
4.4 开发 E-R 图示例	102
4.4.1 Heather Sweeney Designs 公司的数据库	103
4.4.2 培训课的客户列表	103
4.4.3 给客户的信函模板	105
4.4.4 销售发货单	107
4.4.5 属性说明	111
4.4.6 业务规则	111
4.4.7 验证数据模型	111
4.5 小结	112
4.6 复习题	113
4.7 练习题	114
4.8 Garden Glory 项目	114
4.9 James River 珠宝行项目	115
4.10 Mid-Western 大学化学系项目	115

第 5 章 数据库设计	117
5.1 用关系模型表示实体	117
5.1.1 ITEM 实体的表示	117
5.1.2 CUSTOMER 实体的表示	118
5.1.3 非规范化	119
5.1.4 SALES-COMMISSION 实体的关系设计	120
5.1.5 弱实体的表示	122
5.2 关联的表示	123
5.2.1 一对一关联的表示	123
5.2.2 存在问题的一对一关联	125
5.2.3 一对多关联的表示	126
5.2.4 多对多关联的表示	128
5.2.5 递归关联的表示	130
5.3 Heather Sweeney Designs 数据库设计	134
5.3.1 弱实体	135
5.3.2 关联	135
5.3.3 强制参照完整性	136
5.4 小结	138
5.5 复习题	138
5.6 练习题	139
5.7 Garden Glory 项目	140
5.8 James River 珠宝行项目	140
5.9 Mid-Western 大学化学系项目	141
第 6 章 数据库管理	143
6.1 数据库处理环境	144
6.1.1 查询、窗体和报表	144
6.1.2 Internet 应用程序处理	145
6.1.3 客户机/服务器以及传统的应用程序处理	145
6.1.4 存储过程和触发器	146
6.1.5 控制、安全和可靠性的必要性	147
6.2 并发控制	148
6.2.1 使用原子事务的必要性	148
6.2.2 并发事务处理	149
6.2.3 丢失更新问题	150
6.2.4 并发问题：脏读取、非一致读取和幻象读取	151

6.2.5	资源锁定	152
6.2.6	串行化事务	153
6.2.7	死锁	153
6.2.8	乐观锁定和悲观锁定	154
6.2.9	声明锁定特征	155
6.2.10	一致事务	156
6.2.11	事务隔离级别	157
6.3	数据库安全	158
6.3.1	处理权限和责任	158
6.3.2	DBMS 安全	160
6.3.3	应用程序安全	161
6.4	数据库备份与恢复	162
6.4.1	通过重新处理恢复	163
6.4.2	通过回滚和前滚恢复	163
6.5	DBA 的其他职责	167
6.6	小结	167
6.7	复习题	169
6.8	练习题	171
6.9	Garden Glory 项目	171
6.10	James River 珠宝行项目	172
6.11	Mid-Western 大学化学系项目	173
第 7 章	高级主题	175
7.1	Web 数据库处理	175
7.1.1	使用 Active Server Pages 进行 Web 处理	176
7.1.2	使用 JSP 进行 Web 处理	179
7.1.3	ODBC 和 JDBC	181
7.1.4	Web 数据库处理面临的挑战	183
7.2	数据库处理和 XML	183
7.2.1	XML 模式文件	184
7.2.2	XML 和数据库处理	185
7.2.3	XML Web Services	186
7.3	分布式数据库处理	187
7.3.1	分布式数据库的类型	187
7.3.2	分布式数据库面临的挑战	188

7.4	OLAP 和数据挖掘	189
7.4.1	OLAP	190
7.4.2	数据挖掘	192
7.5	对象关系型数据库	193
7.6	小结	194
7.7	复习题	196
7.8	练习题	197
附录 A	使用 Microsoft Access 处理数据库	199
A.1	Carbon River Construction 的数据库模式	199
A.1.1	Carbon River Construction 的数据结构图	200
A.1.2	列设计	201
A.2	用 Access 创建表和关联	202
A.2.1	创建表	203
A.2.2	创建关联	205
A.3	创建查询和 SQL 语句	208
A.3.1	运用图形工具创建简单的 SQL 查询	208
A.3.2	操作查询	213
A.4	创建 Access 的数据输入窗体	215
A.4.1	创建简单的窗体	215
A.4.2	在窗体中使用默认值和组合框	217
A.4.3	查找非键值数据	221
A.5	创建 Access 报表	223
A.5.1	带式报表生成器	223
A.5.2	报表中的计算	226
A.5.3	参数化报表	228
A.6	小结	230
A.7	复习题	231
A.8	练习题	233
A.9	Garden Glory 项目	233
A.10	James River 珠宝行项目	234
	术语表	237

P A R T

I

数据库基础

第 I 部分首先介绍数据库的基本概念和关系数据库管理技术,共包括 3 章,其中第 1 章解释了数据库技术、使用数据库的原因,并描述了数据库系统的组成部分。第 2 章介绍关系模型并定义了一些主要的关系术语,同时讲述了一些简单的关系数据库设计原理。最后,第 3 章介绍了结构化查询语言(Structured Query Language, 简写 SQL),这是用于创建和处理关系数据库的国际标准语言。

通过第 I 部分学习并掌握了关于数据库的基本技术后,我们将在第 II 部分学习数据库的设计和管理技术。

