

计 算 机 科 学 丛 书

数据库设计教程

Thomas M. Connolly Carolyn E. Begg 著 何玉洁 梁琦 等译

Database Solutions

A step-by-step guide
to building databases

Thomas Connolly
& Carolyn Begg

Learn to use UML for data modeling.

Turn theory into practice with the ready-to-use data models and databases supplied.

Build your own database by following the examples in Oracle and Access.



ADDISON-WESLEY

Database Solutions

A Step by Step Guide to Building Databases



机械工业出版社
China Machine Press

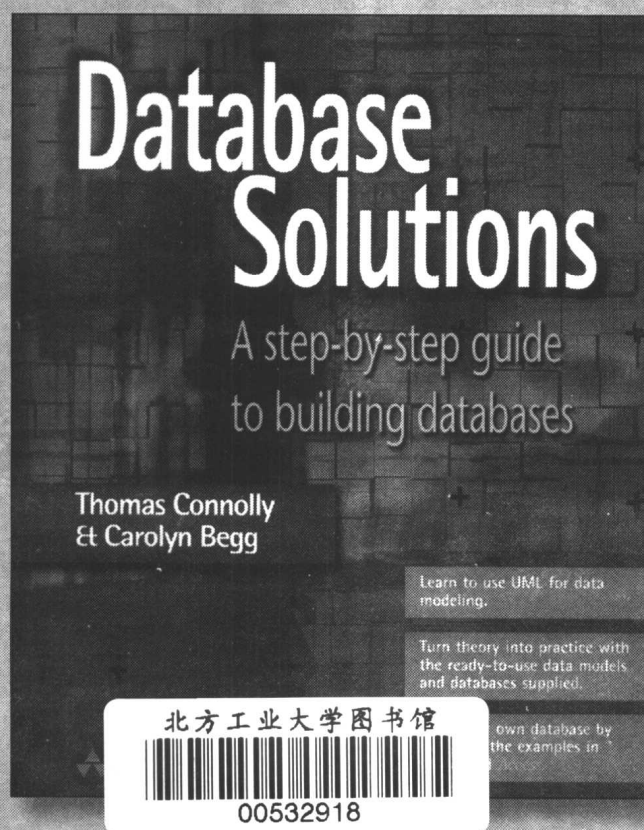


计 算 机 科 学 丛

TP311.13
103D

数据库设计教程

Thomas M. Connolly Carolyn E. Begg 著 何玉洁 梁琦 等译



Database Solutions

A Step by Step Guide to Building Databases



机械工业出版社
China Machine Press

本书详细介绍了数据库的分析、设计、实现方法和过程。本书不仅覆盖了普通数据库书籍中所没有讲到的数据库分析、设计和执行过程,还通过两个贯穿全书的实例描述了数据库的创建、监控和调整的全过程。主要内容包括DBMS和关系数据库的相关知识,数据库分析和设计技术,逻辑数据库设计方法学,物理数据库设计方法学等内容。随书光盘中包含了书中实例所用的脚本。本书深入浅出、内容全面,极具实用性,适合数据库设计、开发和应用的技术人员参考,也适合作为高等院校数据库设计课程的教材。

Thomas M. Connolly, Carolyn E. Begg: Database Solutions: A Step by Step Guide to Building Databases (ISBN:0-201-67476-9).

Copyright © 2000 by Pearson Education Limited.

This translation of Database Solutions: A Step by Step Guide to Building Databases is published by arrangement with Pearson Education Limited.

本书中文简体字版由英国Pearson Education培生教育出版集团授权出版。

版权所有,侵权必究。

本书版权登记号:图字:01-2003-2617

图书在版编目(CIP)数据

数据库设计教程/康诺利(Connolly, T. M.)等著;何玉洁等译.-北京:机械工业出版社,2003.6

(计算机科学丛书)

书名原文:Database Solutions: A Step by Step Guide to Building Databases

ISBN 7-111-12151-1

I. 数… II. ①康… ②何… III. 数据库-程序设计-方案 IV. TP311.13

中国版本图书馆CIP数据核字(2003)第036134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:瞿静华 朱 劼

北京昌平奔腾印刷厂印刷·新华书店北京发行所发行

2003年6月第1版第1次印刷

787mm×1092mm 1/16·17.75 印张

印数:0 001-4 000 册

定价:35.00 元(附光盘)

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、研究及收藏。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：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：hzedu@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

译者序

随着计算机技术的飞速发展和信息管理领域的日益扩大，数据库技术得到了空前的发展，数据库技术的应用领域也日趋扩大和深入。数据库是用科学的方法管理和组织数据的技术，是计算机科学的一个重要分支。

本书讲述了数据库设计方法学。全书由六个部分组成。第一部分由第1章到第3章组成，分别介绍了数据库使用示例，数据库管理系统的功能，关系数据库的一些基本概念，包括关系数据模型和关系完整性，以及数据库应用程序的生命周期等。这部分内容是后续数据库设计部分的预备知识。第二部分由第4章到第6章组成，介绍了进行数据库设计首先要做的事情——事实发现，包括事实发现所使用的技术和方法，整理所发现的事实并建立数据库模型的方法以及如何在数据模型的基础上设计规范化的关系数据库表。第三部分由第7章到第11章组成，第7章全面介绍了本书使用的数据库设计方法学，从第8章到第10章用一个完整的例子介绍了逻辑数据库设计过程中的三个步骤，第11章进一步介绍了一些高级建模技术。第四部分由第12章到第17章组成，其中从第12章到第16章使用前面讲述逻辑数据库设计的例子介绍了物理数据库设计阶段的五个步骤，第17章介绍了使用SQL和QBE实现所举示例的查询。第五部分由第18章和第19章组成，该部分进一步利用了前边章节介绍的数据库设计方法学的步骤，用另一个例子再次说明了本书所介绍的数据库设计方法学的应用。第六部分包含四个附录。附录A介绍了其他一些较常用的表示数据模型的方法。附录B对本书所介绍的数据库设计方法学中的步骤进行了归纳和总结，便于读者快速参考和查阅。附录C介绍了一些常用的数据库文件和索引的组织方式。附录D介绍了15个常用应用领域的数据模型，读者在解决类似问题时可以参考这些模型。

本书的作者从事过多年的数据库方面的工作，具有丰富的数据库设计和使用经验，因此本书将数据库技术与实际应用密切地结合起来，弥补了现在很多数据库图书重视理论介绍，忽视解决实际问题的缺点。本书融合了作者多年的数据库设计、开发和使用的经验，因此所介绍的内容非常有效和实用，非常适合数据库设计人员使用。

本书将抽象枯燥的数据库设计方法用一个很容易理解的例子贯穿起来，一点一点地展开，直至使用完数据库设计方法学的全部步骤，因此很容易激发起读者的阅读兴趣，使读者能更好地理解这些方法学的运用。在读者已经基本熟悉了这些方法学的过程之后，再举出第二个例子，第二个例子快速地将前边介绍的方法学又贯穿一遍，使读者能进一步加深对方法学的理解。此外，本书图文并茂，读者能很直观地理解数据库设计的概念和方法。

本书适合于有一定数据库理论基础，又非常希望能将理论知识应用到实际使用过程中的读者阅读，尤其适合于那些在数据库设计中不知如何下手或者经验不足的数据库设计新手参考。本书可作为学习完数据库理论知识之后的后续学习内容，也可作为数据库设计和开发人员的参考书。本书也可作为大专院校数据库课程的教材或参考读物。

本书全书由何玉洁负责组织翻译和审校，主要由何玉洁和梁琦翻译完成，李艳华、余佐、崔晗、杨永刚、种璟、刘斌、卢立能、陈正林、何玉书、李迎、李钊等参加了本书的部分翻译工作。

限于译者水平，译文中难免有疏漏和不妥之处，欢迎读者批评指正。

译者

2002年11月

前言

背景

如今，数据库是信息系统的基本框架并且从根本上改变了许多公司和个人的工作方式。数据库技术经过多年的发展已经创造出了更强大、更易用的数据库系统，使用户无需具备开发高效系统的知识便可创建数据库并实现其应用。目前的参考文献中，有许多优秀的、阐述了数据库的部分发展生命周期的著作。然而，我们发现很少有书籍介绍数据库的分析、设计以及执行，并且很少有书籍采用专业和非专业读者都易懂的方式描述数据库的开发过程。

因此，本书的创作初衷是尽可能清楚地为读者解释如何分析、设计及实现数据库，包括从仅含几张表的简单数据库到包括数十至数百张表的大型数据库。实际上本书对进行理论研究人士也非常有用，而且简单明了地阐述了数据库设计的方法学。

本书所介绍的用于关系数据库管理系统（当今商业应用中的主要系统）的方法学已经通过业界和学术界多年的检验。这种方法学分为两个阶段：

- 逻辑数据库设计阶段。在这个阶段，我们开发了一个模型来描述所需要的内容，忽略了实现的细节。
- 物理数据库设计阶段。在这个阶段，我们确定如何在目标数据库管理系统（DBMS）实现这些需求，如Access、Paradox、Oracle、Informix或者SQL Server。

我们为每个阶段都提供了易学的步骤，缺乏经验的设计者可以按照所提供的步骤进行学习。此外，一些指导方针也可以用于完成这个学习过程。对于经验丰富的设计者，这种方法学更多地成为了一种框架或者检查列表，而其说明性则极少。

帮助理解数据库的设计

为了帮助读者应用方法学、理解重要的问题，我们提供了一个综合的工作示例，该示例是基于一家称为StayHome的录像出租公司，这个示例将贯穿本书始终。为更完善地补充方法学，我们在第18章和第19章中选取了另外一个名为Perfect Pets的宠物诊所作为示例。

为了帮助读者理解得更加深入，我们在附录D增加了数据库的解决方案（本书配套光盘上包含相应SQL脚本）。每个方案都有一个简短的介绍，读者可以参照这个介绍，在阅读样例之前试着自己进行数据库的设计。

常见的数据模型

在提供数据库设计经验的同时，附录D还提供了许多常见的且有用的数据模型。实际上，据统计约有1/3的数据模型包含了适用于大多数公司的公共结构，而其余的2/3则用于特殊产业或者特殊单位。因此，大多数数据库设计工作包含了重建那些被其他公司创建了多次的结构。这些具有特征性的模型也许并不能完全地满足你的公司的需求，但是，它们可以作为开发的起点，在此基础上，你可以建立更适合本公司特殊需要的模型。我们所提供的模型覆盖了以下的常见业务领域：

- 消费者订购。
- 库存控制。
- 资产管理。

- 项目管理。
- 课程管理。
- 人力资源管理。
- 工资管理。

统一建模语言

越来越多的公司正在采用标准化的方法。使用标准化的方法，他们可以选择一种特定的方法进行数据建模，并将这种方法贯穿于数据库开发过程的始终。本书采用一种常用于逻辑数据库设计的高级数据模型，这种模型是建立在实体 - 关系（ER）概念上的模型。目前，ER模型没有标准的表示法。大多数书籍在提及关系数据库管理系统（DBMS）的数据库设计时会用到以下两种传统的表示法之一：

- Chen氏表示法，它由代表实体的矩形和代表关系的菱形以及连接矩形和菱形的连线组成。
- Crow's Feet（标注）表示法，它由代表实体的矩形以及代表实体间关系的连线组成。连线一端的标注代表一对多的关系。

当今的CASE工具能很好地支持这两种表示法。但是，它们应用起来比较麻烦而且难以解释。在本书中，我们采用最新的面向对象建模语言中的类图表示法，这种建模语言称为UML（Unified Modeling Language，统一建模语言）。UML表示法结合了三种主要的面向对象设计的要素：Rumbaugh的OMT建模，Booch的面向对象分析和设计以及Jacobson的对象论。可以预计，UML将会成为一种标准，而且对象管理组（OMG）已经采用了UML作为对象方法的标准表示法。

我们相信这种表示法更易于理解和使用。为了便于理解，我们在本书光盘中提供了评估版本的可视化建模工具Rational Rose。这种工具支持UML，并且可以创建自己的ER图。

说明怎样执行一个设计

我们认为说明怎样将数据库设计转换为物理实现是非常重要的。在本书中，我们说明了如何用Microsoft Access 97实现第一个示例（名为StayHome的录像租借公司），第二个示例（名为Perfect Pets的宠物诊所）说明了如何用Oracle 8实现数据库设计。

本书的读者对象

哪些人应该阅读本书？我们试图以独立的方式编写这本书，但物理数据库设计是个例外，因为在这一部分，你应该很好地理解目标数据库管理系统是怎样运转的。我们的读者就是那些希望开发数据库的人，包括如下但并非仅仅如下所示的一些人：

- 信息模型设计者和数据库设计者。
- 数据库应用设计者和实现者。
- 数据库制造商。
- 数据和数据库管理者。
- 专门从事数据库设计的学者。
- 数据库专业的学生，包括大学生和研究生。
- 希望设计并开发数据库应用的人。

本书的结构

本书共有六个部分：

- 第一部分——相关知识介绍。在第1章和第2章，我们介绍了数据库管理系统和关系模型。第3章提供了数据库应用生命周期的概述。
- 第二部分——数据库分析与设计技术。我们在第4章讨论了数据库分析技术，说明了怎样用其中的一些技术来分析录像租借公司StayHome的需求。在第5章说明了怎样用UML画实体-关系(ER)图。怎样应用规范化规则的内容将在第6章中阐述。ER模型和规范化都是第三部分描述的方法学中非常重要的技术。
- 第三部分——逻辑数据库设计。我们用循序渐进的方式描述了逻辑数据库的设计。第1步，为每个公司的视图创建局部逻辑数据库。第2步，将每一个模型映射为数据库表。第3步，将局部数据模型组合起来创造本公司的全局数据模型。
- 第四部分——物理数据库设计。我们依然用循序渐进的方式介绍物理数据库的设计。第4步，为目标数据库管理系统设计一些基本表。第5步，选择文件组织和索引。第6步，引入受控冗余以获得改进的性能。第7步，设计安全性控制来保护数据免受未授权者的访问。最后，第8步，监控和调整操作系统。正如刚才所讲，我们将在Microsoft Access 97中说明怎样实现StayHome数据库应用。
- 第五部分——第二个实例。第二个宠物诊所示例贯穿于第18章和第19章。我们将说明怎样在Oracle 8中实现Perfect Pets 数据库的设计。
- 第六部分——附录。附录A说明了两种主要的ER表示法：Chen氏表示法和Crow的Feet表示法。附录B提供了方法学的总结。附录C提供了文件组织和存储结构的背景知识，这些知识可以帮助读者理解第三部分所讲的物理数据库设计方法学的某些内容。附录D提供了15种常见的数据模型。

为增强本书的可读性，本书采用了下列风格和结构：

- 在每章的开头，清晰地给出每章的主题。
- 在每章的结束，概括主要知识点。
- 引入的每一个重要概念都被清晰地定义。
- 采用一系列的注意和提示来显示重点。
- 大量地使用图表，以便更清晰地说明和表示概念。
- 非常实用的方式，每章都包含一些工作示例来说明这些知识点的使用方式。
- 书末附有术语表，可以快速找到相关概念，还提供了每一部分所涉及概念的参考文献。

随书光盘

随书光盘包含下述内容：

- 方便地得到最新的可视化模型工具Rational Rose的评估版本。
- StayHome数据库应用在Microsoft Access 97中的实现。
- 创建Perfect Pets数据库应用实现的SQL脚本。这个脚本可用来在许多关系数据库管理系统(如Oracle、Informix和SQL)中创建数据库。
- 附录D中常见的数据模型的SQL脚本可用于为数据库应用创建相应的基本表。而且，这些脚本还可用来在许多其他关系数据库管理系统中创建数据库。

修正与建议

由于此类型的书籍难免存在错误，因此你的指正对本书以后的再版很重要。请把意见、改正和有建设性的建议发邮件至：thomas.connolly@paisley.ac.uk。

目 录

出版者的话

专家指导委员会

译者序

前言

第一部分 相关知识介绍

第1章 引言	1
1.1 数据库系统使用示例	1
1.2 数据库概述	2
1.2.1 数据库	2
1.2.2 数据库管理系统	3
1.2.3 视图	3
1.2.4 DBMS环境的组成	4
1.2.5 DBMS架构	4
1.3 DBMS的功能	6
1.4 数据库设计	9
1.5 DBMS的优缺点	9
1.6 本章小结	10
第2章 关系模型	12
2.1 数据模型	12
2.2 术语	12
2.2.1 关系数据结构	13
2.2.2 关系表的属性	14
2.2.3 关系键	15
2.2.4 关系数据库的表示	16
2.3 关系完整性	18
2.3.1 空值	18
2.3.2 实体完整性	19
2.3.3 参照完整性	19
2.3.4 业务规则	19
2.4 关系语言	19
2.5 本章小结	21
第3章 数据库应用程序生命周期	22
3.1 软件危机	22
3.2 信息系统生命周期	23
3.3 数据库应用程序生命周期	23

3.4 数据库规划	23
3.5 系统定义	23
3.6 需求的收集与分析	26
3.7 数据库设计	28
3.8 选择DBMS	28
3.9 应用程序设计	28
3.9.1 事务设计	29
3.9.2 用户界面设计	29
3.10 构建原型	30
3.11 实现	30
3.12 数据转换与加载	30
3.13 测试	31
3.14 操作性维护	31
3.15 本章小结	31

第二部分 数据库分析与设计技术

第4章 事实发现	33
4.1 什么时候使用事实发现技术	33
4.2 收集哪些事实	34
4.3 事实发现技术	34
4.3.1 检查文档	35
4.3.2 面谈	35
4.3.3 观察业务的运转	36
4.3.4 研究	36
4.3.5 问卷调查	36
4.4 StayHome案例研究	37
4.4.1 StayHome案例研究——概览	37
4.4.2 StayHome案例研究——数据库规划	40
4.4.3 StayHome案例研究——系统定义	44
4.4.4 StayHome案例研究——需求收集 和分析	46
4.4.5 StayHome案例研究——数据库设计	51
4.5 本章小结	51
第5章 实体-关系建模	52
5.1 实体	52

5.2 关系	53	8.1 为每个视图构建逻辑数据库模型	84
5.2.1 关系的度	54	8.1.1 步骤1.1: 标识实体	85
5.2.2 递归关系	54	8.1.2 步骤1.2: 标识关系	86
5.3 属性	54	8.1.3 步骤1.3: 标识实体或关系的 有关属性	89
5.3.1 简单属性和复合属性	55	8.1.4 步骤1.4: 确定属性域	92
5.3.2 单值属性和多值属性	55	8.1.5 步骤1.5: 确定候选键和主键属性	92
5.3.3 派生属性	55	8.1.6 步骤1.6: 特化/泛化实体	94
5.3.4 键	56	8.1.7 步骤1.7: 删除与关系模型不相容 的特性	94
5.4 强实体和弱实体	57	8.1.8 步骤1.8: 检查模型是否支持 用户事务	100
5.5 关系的多样性约束	57	8.2 本章小结	101
5.5.1 一对一关系	58	第9章 逻辑数据库设计——步骤2	102
5.5.2 一对多关系	59	9.1 步骤2: 为每个局部逻辑数据模型 创建并检查表	102
5.5.3 多对多关系	59	9.1.1 步骤2.1: 为局部逻辑数据模型 创建表	102
5.5.4 非二元关系的多样性约束	60	9.1.2 步骤2.2: 用规范化方法检查表 结构	110
5.5.5 基数约束与参与约束	62	9.1.3 步骤2.3: 检查表是否支持用户 事务	110
5.6 关系上的属性	62	9.1.4 步骤2.4: 定义完整性约束	112
5.7 ER模型中的设计问题	63	9.1.5 步骤2.5: 与用户讨论局部逻辑 数据模型	115
5.7.1 扇形陷阱	63	9.2 本章小结	115
5.7.2 深坑陷阱	64	第10章 逻辑数据库设计——步骤3	116
5.8 本章小结	66	10.1 StayHome的业务视图	116
第6章 规范化	67	10.1.1 用户的需求说明	116
6.1 简介	67	10.1.2 局部逻辑数据模型	117
6.2 数据冗余和更新异常	67	10.2 步骤3: 构建并检查全局逻辑 数据模型	118
6.2.1 插入异常	68	10.2.1 步骤3.1: 合并局部逻辑数据 模型为全局模型	119
6.2.2 删除异常	69	10.2.2 步骤3.2: 检查全局逻辑数据 模型	124
6.2.3 更新异常	69	10.2.3 步骤3.3: 检查未来的可变性	124
6.3 第一范式	69	10.2.4 步骤3.4: 与用户讨论全局逻辑 数据模型	124
6.4 第二范式	71	10.3 本章小结	124
6.5 第三范式	74	第11章 高级建模技术	125
6.6 本章小结	75		
第三部分 逻辑数据库设计			
第7章 方法学概述	77		
7.1 数据库设计方法学简介	77		
7.1.1 什么是数据库设计方法学	77		
7.1.2 数据库设计方法学的目标	78		
7.1.3 为什么构建数据模型	78		
7.1.4 数据库设计中的关键成功因素	80		
7.2 数据库设计方法学概述	80		
7.3 本章小结	82		
第8章 逻辑数据库设计——步骤1	84		

11.1 特化/泛化	125
11.1.1 超类和子类	125
11.1.2 超类/子类关系	125
11.1.3 属性继承	126
11.1.4 特化过程	127
11.1.5 泛化过程	127
11.1.6 超类/子类关系的约束	129
11.2 创建表达特化/泛化的表	130
11.3 本章小结	131

第四部分 物理数据库设计

第12章 物理数据库设计——步骤4	133
12.1 逻辑与物理数据库设计的比较	134
12.2 物理数据库设计方法学概述	134
12.3 步骤4: 为目标DBMS转换全局 逻辑数据模型	135
12.3.1 步骤4.1: 为目标DBMS设计 基本表	135
12.3.2 步骤4.2: 为目标DBMS设计 业务规则	140
12.4 本章小结	143
第13章 物理数据设计——步骤5	144
13.1 理解系统资源	144
13.2 步骤5: 设计物理表示	145
13.2.1 步骤5.1: 分析事务	146
13.2.2 步骤5.2: 选择文件组织方式	150
13.2.3 步骤5.3: 选择索引	151
13.3 使用Microsoft Access 97的StayHome 文件的组织与索引	154
13.3.1 选择索引指南	154
13.3.2 StayHome的索引	155
13.4 本章小结	156
第14章 物理数据库设计——步骤6	157
14.1 步骤6: 受控冗余的考虑	157
14.1.1 步骤6.1: 考虑派生的数据	158
14.1.2 步骤6.2: 同时考虑重复列或 连接表	159
14.2 本章小结	167
第15章 物理数据库设计——步骤7	169
15.1 步骤7: 设计安全机制	169

15.1.1 步骤7.1: 设计用户视图	169
15.1.2 步骤7.2: 设计访问规则	170
15.2 本章小结	174
第16章 物理数据库设计——步骤8	175
16.1 步骤8: 监视并调整操作系统	175
16.2 本章小结	176
第17章 使用SQL和QBE查询 StayHome示例	177
17.1 Microsoft SQL和QBE介绍	177
17.1.1 SQL	177
17.1.2 QBE	178
17.2 StayHome查询示例	178
17.3 本章小结	184

第五部分 第二个实例

第18章 Perfect Pets——逻辑数据库设计	185
18.1 Perfect Pets实例	185
18.1.1 数据需求	185
18.1.2 事务需求	187
18.2 使用逻辑数据库设计方法	187
18.2.1 步骤1.1: 标识实体	188
18.2.2 步骤1.2: 标识关系	188
18.2.3 步骤1.3: 标识实体或关系的 有关属性	189
18.2.4 步骤1.4: 确定属性域	191
18.2.5 步骤1.5: 确定候选键和主键属性	192
18.2.6 步骤1.6: 特化和泛化实体	192
18.2.7 步骤1.7: 删除与关系模型不相容 的特性	192
18.2.8 步骤1.8: 检查模型是否支持 用户事务	194
18.2.9 步骤2.1: 为局部逻辑数据模型 创建表	194
18.2.10 步骤2.2: 用规范化方法检查表	194
18.2.11 步骤2.3: 检查模型是否支持 用户事务	194
18.2.12 步骤2.4: 定义完整性约束	197
18.2.13 步骤2.5: 与用户一起讨论局部 逻辑数据模型	198
第19章 Perfect Pets——物理数据库设计	199

19.1 使用物理数据库设计方法学199

19.1.1 步骤4.1: 为目标DBMS设计基本表...199

19.1.2 步骤4.2: 为目标DBMS设计
业务规则202

19.1.3 步骤5.1: 分析事务204

19.1.4 步骤5.2: 选择文件组织方式204

19.1.5 步骤5.3: 选择索引211

19.1.6 步骤6.1: 考虑派生的数据212

19.1.7 步骤6.2: 同时考虑重复列或
连接表212

19.1.8 步骤7.1: 设计用户视图212

19.1.9 步骤7.2: 设计访问规则213

19.2 后续内容介绍216

第六部分 附 录

附录A 可选的数据建模表示法217

附录B 数据库设计方法学总结222

附录C 文件组织和索引227

附录D 常用数据模型234

术语表256

参考文献262

第一部分 相关知识介绍

第1章 引言

第2章 关系模型

第3章 数据库应用程序生命周期

第1章 引 言

本章主题：

- 数据库系统的一般用途。
- 术语“数据库”的含义。
- 术语“数据库管理系统”(DBMS)的含义。
- DBMS环境的主要组成部分。
- DBMS的典型功能与服务。
- DBMS的优缺点。

现在，数据库已经成为我们每天生活中不可缺少的一部分，虽然有人可能并没意识到这一点。在开始讨论数据库之前，我们简要地介绍一些数据库系统的应用。为了方便讨论，我们认为“数据库”是相关数据的集合，而“数据库管理系统”(Database Management System, DBMS)是管理和控制对数据库进行访问的软件，更概括的术语“数据库系统”是指与数据库相关的应用程序的集合。在1.2节中，我们将提供更精确的定义。在本章末尾，我们将介绍现代DBMS的典型功能并且简要地介绍数据库的优缺点。

1.1 数据库系统使用示例

1. 从超市购物

当你从当地超市购买货物时，就是在访问数据库。收银员使用条形码阅读器来扫描你的每种货物。这就链接了一个使用条形码从产品数据库中查询该项货物价格的应用程序，然后该程序产生这些库存项目的数量，在收银机上显示价格。如果记录产品的数量低于指定的最低极限值，数据库系统可能会自动设置一个订单来获得更多的该产品的库存。

2. 使用信用卡购物

当你使用信用卡购物时，服务人员要检查你是否有足够的剩余金额可以购买该商品。这种检查可以用电话来进行，或者也可以用连接到计算机系统的磁卡阅读器自动完成。无论是哪种情况，都在某一个地方有一个数据库，此数据库中包含了你使用信用卡进行购物的信息。为了检查你的信用卡，存在一个数据库应用程序，此程序使用你的信用卡号码来检查你想购买的商品的价格，以及你这个月已经购买的商品的总额是否在信用限度内。当购买被确认有效后，则这次的购买详细信息又被加到了这个数据库中。在确认此次购买生效之前，这个应用程序也会

访问数据库，检查该信用卡是否在被盗或丢失列表中。还有其他数据库应用程序，它们向每个持卡者每月发送通知，并在收到付款后汇去账单。

3. 在旅行社预定假期

当你咨询某次假期的安排时，旅行社可能访问几个包含假期和飞机的详细信息的数据库。当你预定假期时，数据库系统必须进行所有必要的预定安排。在这种情况下，该系统必须要确保不同的代理没有预定相同的假期或飞机上相同的座位。例如，如果在从伦敦到纽约的飞机上只剩下一个座位，两个代理在同一时间预定这最后一个座位，系统不得不处理这种情况，允许一个预定继续，并通知另一个代理没有位置了。通常，旅行代理都有独立的用来记账的数据库。

4. 使用图书馆

图书馆一般会有一个包含所有图书的详细信息的数据库，其中的信息可能还包括读者的信息、预定信息等等。可能会允许读者基于书名、作者或其他数据查找所需书籍。数据库系统可能会允许读者预定书籍，并在书籍可以借阅时发邮件通知读者。该系统也给没有按期还书的借阅者发提醒通知。一般情况下，系统都有一个条形码阅读器，类似于前边超市中描述的那种，用来记录归还和借出图书馆的书籍。

这只是几种数据库系统应用，毫无疑问，你会知道其他许多应用情况。尽管我们熟知并常用这些应用，但在它们的背后，却隐藏着复杂的高级技术。这种技术的核心就是数据库本身。对于要尽可能有效地支持最终用户需求的系统，需要合适的结构化数据库。构建这种结构就是所说的数据库设计，这也是本书讨论的核心内容。无论要构建的数据库是小型的，还是像上面所说的那种大型数据库，数据库设计都是基础，本书中提出的方法学将相对容易地帮你构建正确的数据库。拥有一个设计优良的数据库可以使你构建符合用户需求的系统，同时，提供可接受的性能。

1.2 数据库概述

在本节中，我们为术语“数据库”和“数据库管理系统”提供更为正式的定义。

1.2.1 数据库

数据库 (Database) 逻辑上相关的可共享的数据（以及数据的描述）集合，用于处理公司所需的信息。

让我们详细分析数据库的定义以便全面理解这个概念。数据库是一个巨大的数据存储地，可以同时被许多部门和用户使用。这些用户所需要的所有数据用最小的冗余集成在一起。更重要的是，数据库通常不属于任何一个部门或用户，而是公司内的共享资源。

系统目录 (System Catalog) 存储数据库中所有对象的数据。

除了存储公司的数据，数据库也存储数据的描述。因此，数据库也定义为完整记录的自描述集合。数据的描述，也就是元数据——关于数据的数据，被称为系统目录或数据字典。正是数据库的自描述特性提供了被称为数据独立性的特点。这就意味着如果有新的数据结构要加入到数据库中，或者要修改数据库中已经存在的数据结构，那么如果应用程序不直接依赖于被修改的部分，则对该数据库的应用程序来说没有影响。例如，如果我们为记录添加了新列或者创建了新表，则对已经存在的应用程序是没有影响的。但是，如果从一个应用程序使用的表中

删除了一列，那么这个应用程序就会受到这个变化的影响，因此也必须相应地进行修改。

我们要解释的数据库定义中的最后一个术语就是“逻辑相关”。当我们分析了公司的信息需求后，我们尽力去标识数据库中要描述的重要对象和这些对象之间的逻辑关系。我们提出的数据库设计方法学将引导你标识这些重要的对象和它们的逻辑关系（方法学参见第7章~第16章）。

1.2.2 数据库管理系统

DBMS（数据库管理系统） 是使用户可以定义、创建和维护数据库以及提供对该数据库有限制的访问的软件系统。

DBMS是与用户、应用程序和数据库进行相互作用的软件。其中，DBMS允许用户从数据库中插入、更新、删除和检索数据。拥有所有数据和数据描述的核心仓库使得DBMS能够提供对数据进行通常查询功能的语言，称为查询语言。这种查询语言的提供（例如SQL），降低了早期系统中所存在的问题，那时，用户不得不使用固定的一组查询，或者用额外的程序来解决主要的软件管理问题。在1.3节中，我们将讨论DBMS的这些典型功能和服务。

结构化查询语言（Structured Query Language, SQL，发音为S-Q-L或者See-Quel） 关系数据库的主要查询语言（SQL的介绍参见2.4节和第17章）。

图1-1说明了数据库的访问形式。它表明销售和库存控制部门使用他们的应用程序，通过DBMS对数据库进行访问。每个部门的应用程序处理数据入口、数据维护和生成报表。数据的物理结构和存储由DBMS管理。

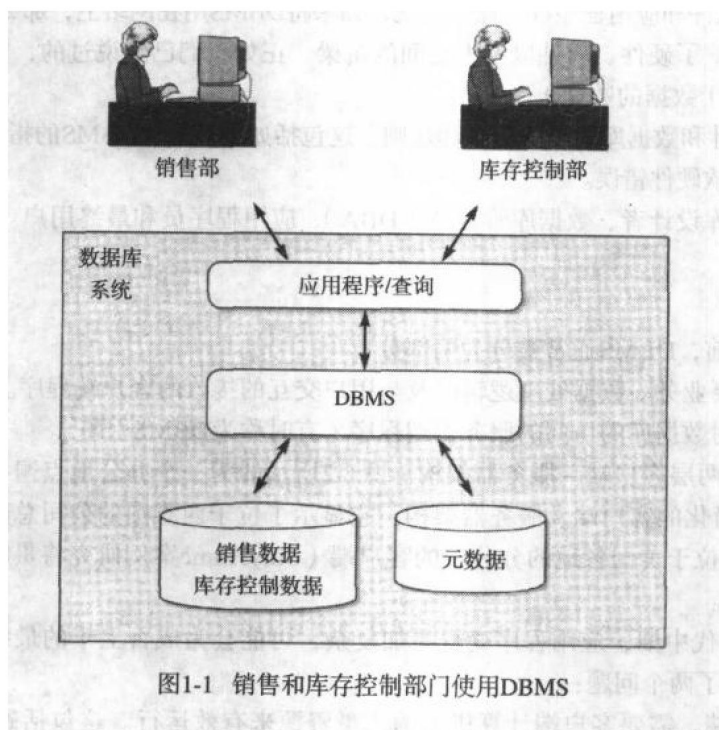


图1-1 销售和库存控制部门使用DBMS

1.2.3 视图

根据前面所描述的功能，DBMS是非常强大的工具。但是，最终用户对系统的任务是复杂还