



普通高等教育“十一五”国家级规划教材

研究型教学模式系列教材

数据库技术及应用 (第2版)

马涛 主编 唐好魁 闫明霞 朱连江 编 蒋宗礼 主审

<http://www.phei.com.cn>

 电子工业出版社
PUBLISHING HOUSE OF ELECTRONICS INDUSTRY

普通高等教育“十一五”国家级规划教材
研究型教学模式系列教材

数据库技术及应用

(第2版)

	马 涛		主编
唐好魁	闫明霞	朱连江	编
	蒋宗礼		主审

电子工业出版社

Publishing House of Electronics Industry

北京·BEIJING

内 容 简 介

本书是普通高等教育“十一五”国家级规划教材，根据教育部对高等学校非计算机专业计算机基础系列课程的教学基本要求，从实用性和先进性出发，全面介绍有关数据库的基础知识和应用技术。

本书分为理论和实验两部分。理论部分共 7 章，主要内容包括：数据库技术的基础理论和基本概念、SQL Server 2000 数据库管理系统的功能及用法、SQL 语言、数据库设计的方法、数据库保护的基础理论及应用、数据库新技术和国产数据库介绍。实验部分设计了 8 个实验，便于读者根据课程教学的进度设计和上机操作。本书附录为读者进行管理信息系统的开发提供了实用工具。本书提供教学用多媒体电子课件和实例数据库 EDU_D，并配套网络教学平台。

本书可作为高等学校非计算机专业的计算机基础课教材，也可作为高职高专院校计算机相关专业的教材，还可供从事数据库开发的读者和计算机技术爱好者学习参考。

未经许可，不得以任何方式复制或抄袭本书之部分或全部内容。
版权所有，侵权必究。

图书在版编目 (CIP) 数据

数据库技术及应用 / 马涛主编. —2 版. —北京: 电子工业出版社, 2010.5

(研究型教学模式系列教材)

ISBN 978-7-121-10828-0

I. ①数… II. ①马… III. ①数据库系统—高等学校—教材 IV. ①TP311.13

中国版本图书馆 CIP 数据核字 (2010) 第 082812 号

责任编辑: 王羽佳

印 刷: 北京京师印务有限公司
装 订:

出版发行: 电子工业出版社

北京市海淀区万寿路 173 信箱 邮编 100036

开 本: 787×1092 1/16 印张: 15.5 字数: 420 千字

印 次: 2010 年 5 月第 1 次印刷

印 数: 5 000 册 定价: 29.00 元

凡所购买电子工业出版社图书有缺损问题，请向购买书店调换。若书店售缺，请与本社发行部联系，联系及邮购电话：(010) 88254888。

质量投诉请发邮件至 zltz@phei.com.cn，盗版侵权举报请发邮件至 dbqq@phei.com.cn。

服务热线：(010) 88258888。

出版说明

随着科教兴国战略的实施和社会信息化进程的加快,我国高等教育事业的发展迎来了新的机遇,高等学校的计算机基础教育也得到蓬勃发展。经过在多年教学实践中的不断探索,我们总结出适合高等学校非计算机专业学生计算机教育的研究型教学模式。

研究型教学模式的基本形式为:精讲多练,以学生在课题研究中探索式的学习为主、以网络教学平台答疑讨论为辅、以试题库在线测验为补充的教学模式。

研究型教学模式的操作,重点突出以下三个方面:

① 加强自学和实践。课堂教学主要精讲重点内容,而不是面面俱到。在教师的指导下,学生通过自学教材,并借助网络教学平台上的多媒体课件或其他多种学习资料进行学习。同时增加上机实验教学的学时比例,充分利用上机练习掌握所学的内容。

② 以实际训练提高教学效果。在上课前给每个学生(或几个学生为一组)布置一项实际操作或软件开发课题。课题力求既结合实际,又能涵盖课程教学的内容,明确具体要求和进度。学生结合课程进度在规定时间内完成该课题后,教师进行考核。

③ 充分重视辅助教学手段在课程教学中的作用。建设在线考试环境,学生可以随时登录进行在线测试。根据教学进度的安排,每个重要学习单元都组织学生在线测试。另外,在教学平台的辅导答疑论坛,安排专人主持,负责解答学生提出的各种问题,根据学生在答疑论坛发表见解的次数和深度,评定答疑讨论分,并计入平时成绩。

总之,研究型教学模式在重视教学过程的每个环节的同时,把调动学生学习的积极性放到了重要位置,把培养学生数字化学习的能力、自主学习的意识和培养学生创新思维的意识有机地融合到平时的教学过程之中。

为了更好地探索研究型教学模式,2006年我们组织编写了这套系列教材,使用3年以来,结合教学过程中的实际需求和各位同仁的反馈意见,我们对这套教材进行了修订。修订后本系列教材主要包括《信息技术基础(第2版)》、《C语言程序设计(第2版)》、《数据库技术及应用(第2版)》和《计算机网络技术与应用(第2版)》等。同时开发了与本套教材相配合的网络化教学平台软件,已在济南大学的非计算机专业学生中试用,收到了较好的教学效果。本套教材还配有习题解答、实验指导及教学用多媒体电子课件,以利于教师备课和学生自学,请登录华信教育资源网(<http://www.hxedu.com.cn>)注册下载。

非计算机专业学生的计算机教育,在教学目的、教学内容和教学方法等方面都不同于计算机专业教育。对非计算机专业的学生,计算机教育的重点应该是计算机应用能力的培养。为此,本套教材从应用出发,以应用为目的,更强调实用性,在确保概念严谨的同时,做到通俗易懂、例题丰富、便于自学。我们希望这套教材能使广大非计算机专业的学生受益,并通过研究型教学模式的应用使他们能更好地灵活掌握信息技术的相关知识和技能。

这套教材得到了济南大学教材建设委员会及各方人士的指导、支持和帮助,在此我们表示衷心的感谢。

教材中还可能存在不足之处,竭诚欢迎广大读者和同行批评指正。

研究型教学模式系列教材编写组

《研究型教学模式系列教材》编委会

主 任 杨 波

副 主 任 陈月辉

委 员 曲守宁 董吉文 刘明军 马 涛

奚 越 唐好魁 徐龙玺 张苏青

韩玫瑰 郭庆北

前 言

在信息化社会,数据库技术已经成为信息处理技术的核心支撑技术之一,被普遍应用于社会、政治、经济活动中,如办公自动化系统、决策支持系统、电子商务系统、证券交易系统、物流管理系统、教学管理系统等管理信息系统。

作为高等学校非计算机专业的学生,对数据库技术的应用已经成为其基本技能之一。这一技能到底要掌握到什么程度,是值得深思的问题。显然,按照计算机专业要求的深度是不切实际的。

数据库系统主要由数据库管理员、系统分析员、数据库设计人员、应用程序员和用户负责维护和使用,其中部分工作通常是由从事某种行业的非计算机专业人员承担。考察一下在推广或应用中失败的管理信息系统,不难发现,其失败的原因不全是由于数据库系统本身存在问题或者其他计算机专业方面产生的问题。需求分析的偏差导致数据库设计方面的缺陷,进一步造成了系统功能的缺陷,常常成为诸多管理信息系统失败的主要原因。

非计算机专业的学生学习数据库技术,是为以此来改进自己所属专业的工作,他们在需求分析方面占有明显的优势。根据工作实际的需要,通过使用包括数据库技术在内的相关信息技术,进行本行业的信息化建设,推进管理水平和工作效率的提高,这就是非计算机专业的学生学习数据库技术的目的。

本书的作者通过多年对教学模式的研究,总结了一套适合非计算机专业的学生学习数据库技术的方法——基于网络化教学平台,精讲多练,以学生在课题研究中探索式的学习为主,以网站答疑讨论为辅,以试题库在线测验为补充的研究型教学模式。作者希望通过本书、网络教学平台和研究型教学模式的结合,使学生更好地掌握数据库技术。

本书根据难易程度和研究型教学模式的需要,对每节的内容进行了划分:✍表示内容比较简单,以自主学习为主;📖表示精讲多练,是重点内容;📖表示读者可以根据自己的兴趣和需要进一步探讨、研究和学习。

本书是普通高等教育“十一五”国家级规划教材。本书根据教育部对高等学校非计算机专业计算机基础系列课程的教学基本要求编写,从实用性和先进性出发,全面介绍了有关数据库的基础知识和应用技术。本书分为理论和实验两部分。理论部分共7章,第1章和第2章介绍了数据库技术的基础理论和基本概念,第3章通过数据库实例 EDU_D 介绍了 Microsoft SQL Server 2000 数据库管理系统的功能及用法,第4章对 SQL 语言进行了重点讲解,第5章通过实例讲述了数据库设计的方法,第6章介绍了数据库保护的基础理论与技术,第7章通过数据库新技术和国产数据库的介绍,开阔了读者的视野。实验部分共设计了8个实验,便于读者根据课程教学的进度开展设计操作和上机实践操作。附录为读者进行管理信息系统的开发提供了实用工具。

本书在章节安排上也改变了传统的章节次序,将 SQL Server 2000 及 SQL 语言的讲解与练习安排在数据库设计之前,目的是先让读者对抽象的数据库有一个感性的认识,然后再学习数据库设计理论。

对规范化理论部分的讲解,怎样能够让非计算机专业学生听懂,一直是数据库理论教学的难点之一。作者根据多年的教学实践,总结出采用函数依赖图和二维表直观展示精心设

计的实例的方法来讲解规范化理论，收到了良好的效果，非计算机专业的学生学习起来比较轻松。

同样，第2章对关系和关系运算的介绍，也把重点偏向学生对二维表的感性认识上，使学生能够感觉到关系模型既亲切又熟悉。

读者通过对附录A、附录B的自学，可以对数据库及管理信息系统的开发有个整体的认识，并能够开发一个管理信息系统。

通过本教材的学习，你可以：

- 学习数据库技术的基础知识
- 掌握 SQL Server 2000 的基本用法
- 熟练掌握 SQL 语言
- 设计一个简单的关系数据库
- 通过自主学习开发管理信息系统

本书可作为高等学校非计算机专业数据库技术及应用课程的教材，也可作为高职高专计算机相关专业的教材，还可供从事数据库开发的读者和计算机技术爱好者学习参考。

本书为使用本书作为教材的教师提供教学用多媒体电子课件、实例数据库 EDU_D 和习题参考答案，请登录华信教育资源网 (<http://www.hxedu.com.cn>) 免费下载。另外，本书配套网络教学平台 (<http://cc.ujn.edu.cn>)，请与作者联系索取。

本书由马涛主持编写并统稿。第1章、附录A和附录B由唐好魁编写，第3章、第4章、附录C和附录D由马涛编写，第2章、第5章和实验篇由闫明霞编写，第6章和第7章由朱连江编写，多媒体电子课件由闫明霞制作。

以下老师也参与了本书的编写工作：王君、曲守宁、奚越、董吉文、徐龙玺、刘明军、韩玫瑰、蒋彦、李英俊、孙志胜、邢静波、杨雪梅、张晓丽、张苏青、王卫峰、黄艺美、潘玉奇、王信堂、郭庆北、王亚琦、段春笋、董梅、马莉、范玉玲、张芊茜、张璘。

北京工业大学的蒋宗礼教授倾注了大量心血对本书进行了审阅。山东建筑大学的李盛恩教授对本书进行了全面、认真的修改，并提出许多宝贵意见。济南大学的杨波教授、曲守宁教授和董吉文教授参与了本书编写的组织与管理工 作，并在技术上给予大力支持，在内容上给予诸多指导。济南大学的郑艳伟老师提出了诸多有价值的建议。在此一并表示衷心的感谢！













本书在编写过程中，参考了大量近年来出版的相关技术资料，吸取了许多同仁和专家的宝贵经验，在此深表谢意！

站在非计算机专业学生的角度，编写一本能够使他们感兴趣且容易学习的教材一直是我们的愿望。但由于编写时间仓促，水平有限，书中难免出现错误或不妥之处，我们诚恳地希望读者和同行批评指正。

作 者

目 录

第 1 章 绪论	(1)
1.1 数据库系统概述	(2)
1.1.1 信息与社会	(3)
1.1.2 数据库的基本概念	(5)
1.1.3 数据库系统的特点	(6)
1.1.4 数据库管理系统的功能	(9)
1.2 数据库模型	(10)
1.2.1 概念模型	(10)
1.2.2 数据模型	(13)
1.2.3 常用数据模型	(13)
1.3 数据库系统结构	(16)
1.3.1 数据库系统的三级模式结构	(16)
1.3.2 二级映像与数据独立性	(17)
1.4 数据库系统的组成	(18)
1.4.1 硬件支撑环境	(18)
1.4.2 软件系统	(18)
1.4.3 数据库	(19)
1.4.4 人员	(19)
本章小结	(21)
习题 1	(21)
第 2 章 关系数据库	(25)
2.1 关系数据模型	(26)
2.1.1 关系数据模型概述	(26)
2.1.2 关系数据模型的基本概念	(26)
2.1.3 关系数据模型的组成	(29)
2.2 关系运算简介	(33)
2.2.1 集合运算	(33)
2.2.2 选择运算	(33)
2.2.3 投影运算	(34)
2.2.4 连接运算	(34)
本章小结	(36)
习题 2	(36)
第 3 章 SQL Server 2000	(39)
3.1 SQL Server 2000 的特点	(40)

3.2	SQL Server 2000 的安装 	(41)
3.2.1	SQL Server 2000 的运行环境要求	(41)
3.2.2	SQL Server 2000 的安装	(41)
3.3	SQL Server 的启动 	(47)
3.4	系统数据库与数据库对象简介 	(47)
3.4.1	系统数据库简介	(47)
3.4.2	系统数据表简介	(48)
3.4.3	系统存储过程简介	(49)
3.4.4	数据库对象简介	(49)
3.5	SQL Server 2000 常用工具 	(50)
3.5.1	SQL Server 企业管理器	(50)
3.5.2	SQL 查询分析器	(65)
3.5.3	导入与导出数据	(66)
3.5.4	数据库的附加与分离	(68)
	本章小结	(69)
	习题 3	(69)
第 4 章	关系数据库语言 SQL	(70)
4.1	SQL 简介 	(71)
4.1.1	SQL 语言的特点	(71)
4.1.2	SQL 语言的主要功能	(71)
4.1.3	SQL 语句的书写准则	(72)
4.2	查询语句 	(72)
4.2.1	基本查询	(73)
4.2.2	使用列表表达式	(75)
4.2.3	WHERE 从句的进一步使用	(77)
4.2.4	数据汇总	(79)
4.2.5	连接查询	(83)
4.2.6	嵌套查询	(84)
4.2.7	联合查询	(87)
4.2.8	使用系统内置函数的查询	(88)
4.3	数据更新 	(93)
4.3.1	插入数据	(93)
4.3.2	修改数据	(93)
4.3.3	删除数据	(94)
4.4	数据定义	(95)
4.4.1	定义基本表 	(96)
4.4.2	修改基本表 	(98)
4.4.3	删除基本表 	(99)
4.4.4	视图 	(99)
4.4.5	索引 	(103)

本章小结	(105)
习题 4	(106)
第 5 章 数据库设计	(110)
5.1 数据库设计概述	(111)
5.1.1 数据库和信息系统	(111)
5.1.2 数据库设计的内容	(111)
5.1.3 数据库设计的基本阶段	(112)
5.2 规范化	(113)
5.2.1 问题的提出	(113)
5.2.2 规范化	(115)
5.2.3 范式	(117)
5.2.4 范式在工程化设计中的实际应用	(122)
5.2.5 规范化小结	(123)
5.3 需求分析	(124)
5.3.1 需求分析的基本内容	(124)
5.3.2 需求分析的方法	(125)
5.3.3 数据流程图简介	(126)
5.3.4 数据字典简介	(126)
5.4 概念结构设计	(130)
5.4.1 概念结构设计任务	(130)
5.4.2 概念结构设计方法与步骤	(130)
5.4.3 局部 E-R 模型设计过程	(131)
5.4.4 全局概念结构设计	(132)
5.4.5 实例	(133)
5.5 逻辑结构设计	(136)
5.5.1 逻辑结构设计任务	(136)
5.5.2 E-R 图向关系模型的转换	(136)
5.5.3 数据模型的优化	(138)
5.5.4 实例	(139)
5.6 数据库的物理设计	(139)
本章小结	(140)
习题 5	(140)
第 6 章 数据库保护	(143)
6.1 事务	(144)
6.1.1 事务的概念	(144)
6.1.2 事务的特性	(145)
6.1.3 SQL Server 2000 事务应用	(146)
6.1.4 事务的状态转换	(147)
6.2 数据库恢复技术	(147)

- 6.2.1 数据库可能出现的故障..... (148)
- 6.2.2 数据库的恢复原理..... (148)
- 6.2.3 SQL Server 2000 中数据备份与恢复的实现..... (150)
- 6.3 并发控制..... (155)
 - 6.3.1 并发操作带来的不一致问题..... (155)
 - 6.3.2 并发控制——封锁及封锁协议..... (156)
- 6.4 数据库的完整性..... (157)
 - 6.4.1 数据库的完整性介绍..... (157)
 - 6.4.2 SQL 中的完整性约束..... (157)
 - 6.4.3 SQL Server 2000 中完整性约束的实现..... (159)
- 6.5 数据库的安全性..... (161)
 - 6.5.1 计算机系统的安全性问题..... (161)
 - 6.5.2 权限..... (162)
 - 6.5.3 数据库的安全性控制..... (162)
 - 6.5.4 SQL Server 2000 中系统安全性的实现..... (164)
- 本章小结..... (170)
- 习题 6..... (171)

第 7 章 数据库新技术及国产数据库介绍..... (173)

- 7.1 数据库技术的发展..... (174)
 - 7.1.1 数据库技术的发展..... (174)
 - 7.1.2 面向对象数据库系统介绍..... (174)
 - 7.1.3 分布式数据库技术介绍..... (177)
 - 7.1.4 主动数据库技术介绍..... (178)
 - 7.1.5 并行数据库技术介绍..... (179)
 - 7.1.6 数据仓库及数据挖掘技术..... (180)
- 7.2 国产数据库介绍..... (182)
 - 7.2.1 金仓数据库管理系统..... (182)
 - 7.2.2 达梦数据库管理系统..... (185)
- 本章小结..... (186)
- 习题 7..... (186)

第 8 章 实验..... (188)

- 实验 1 SQL Server 2000 环境的熟悉和数据库的创建..... (189)
- 实验 2 数据库与数据表的创建、删除与修改..... (189)
- 实验 3 单表 SQL 查询语句练习..... (190)
- 实验 4 多表 SQL 查询语句练习..... (191)
- 实验 5 嵌套查询和集合查询..... (192)
- 实验 6 SQL Server 2000 中视图的创建和使用..... (192)
- 实验 7 SQL Server 2000 中数据的控制与维护..... (193)
- 实验 8 数据定义和数据更新..... (193)

附录 A Delphi/SQL Server 开发与编程	(194)
A.1 Delphi 数据库应用程序结构	(194)
A.2 ADO 组件	(195)
A.3 数据源组件和数据显示/编辑组件	(198)
A.4 设计数据模块连接数据库	(200)
附录 B ASP/SQL Server 开发与编程	(204)
B.1 ASP 技术概述	(204)
B.2 ASP 的内建对象	(205)
B.3 使用 ADO 操作 SQL Server 数据库	(207)
B.4 通过 ODBC 访问数据库	(222)
附录 C 数据类型	(224)
C.1 整数数据类型	(224)
C.2 浮点数据类型	(225)
C.3 二进制数据类型	(226)
C.4 逻辑数据类型	(226)
C.5 字符数据类型	(226)
C.6 文本和图形数据类型	(227)
C.7 日期和时间数据类型	(227)
C.8 货币数据类型	(228)
C.9 特定数据类型	(229)
C.10 用户自定义数据类型	(229)
C.11 新数据类型	(229)
附录 D SQL Server 2000 常用内置函数	(230)
D.1 数据转换函数	(230)
D.2 字符串函数	(230)
D.3 算术函数	(231)
D.4 文本 (text) 与图像 (image) 函数	(231)
D.5 日期与时间函数	(231)
D.6 系统函数	(232)
参考文献	(233)

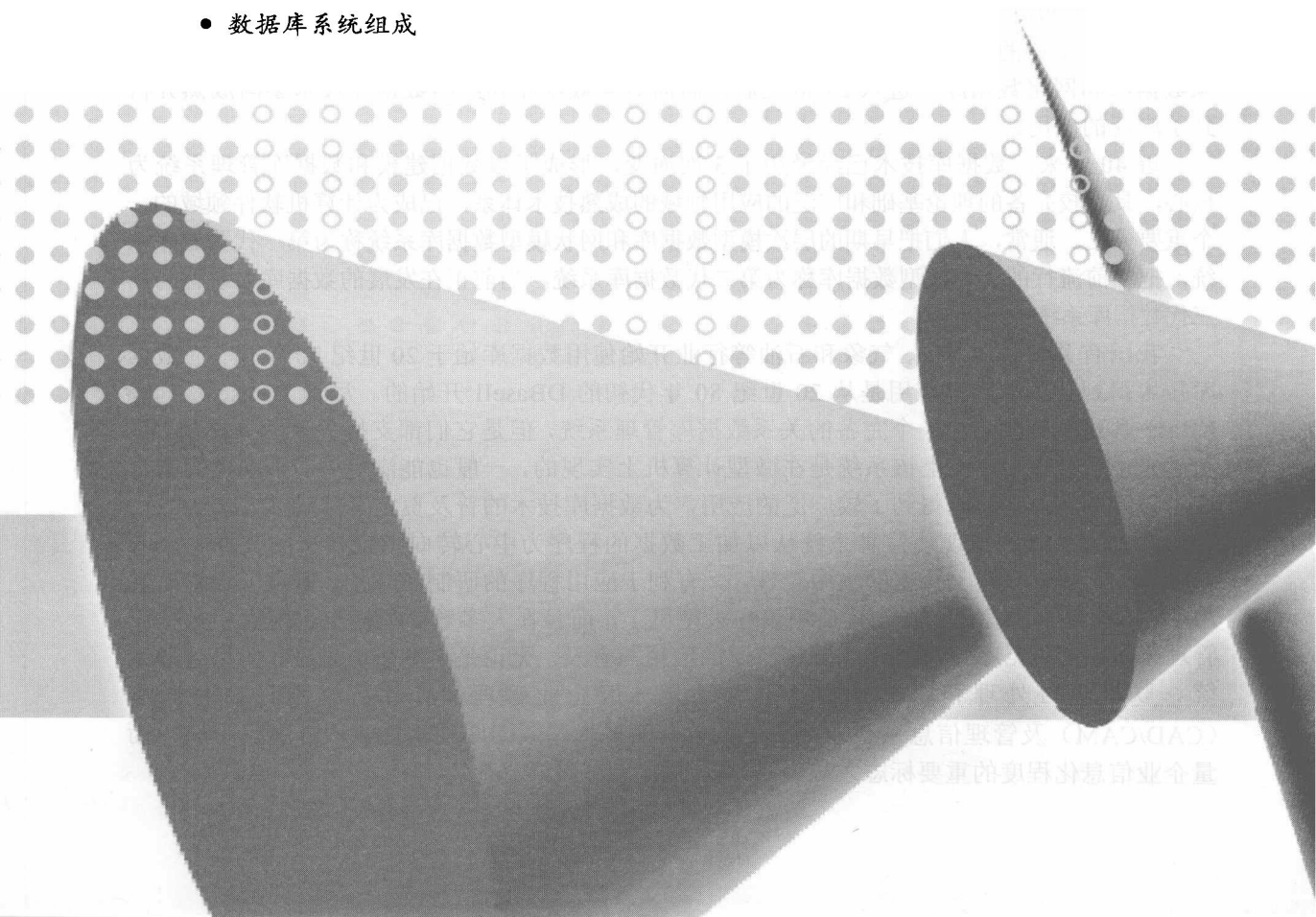
第 1 章 绪 论

本章主要介绍与数据库技术有关的基本概念与术语。通过学习，读者可以初步掌握数据库的基本概念、数据模型及其三要素等知识。通过对数据库系统三级模式和两级映像功能的理解，对数据库的系统结构将会有个总体上的认识，进而对数据库宏观的理解和把握。通过对数据库的概念及其系统组成的学习，有助于读者准确定位自己在将来工作中应承担的角色，有利于开展有目的的自主学习。

本章列出的一些管理信息系统的课题，是读者进行自主学习的驱动目标，便于在对其中某课题的探索中主动地理解和掌握数据库的知识与技能。

本章导读：

- 数据库的基本概念和术语
- 数据模型及其三要素
- 数据库系统的三级模式、两级映像、数据独立性
- 数据库系统组成



1.1 数据库系统概述

自从第一台计算机面世以来，计算机在生产、生活中的应用发生了很大变化。从20世纪50年代开始，计算机的应用领域由科学计算逐渐扩展到广义的数据处理的各个领域。到20世纪60年代末，数据库技术作为数据处理的一种新手段迅速发展起来，成为应用最广泛的计算机应用的技术之一，也是计算机信息系统和应用系统的核心技术和重要基础。

数据库的概念最初产生于20世纪50年代，当时美国为了战争的需要，把各种情报集中起来存储在计算机中，被称为 Information Base 或 Database。在20世纪60年代的软件危机中，数据库技术作为软件技术的分支得到了进一步的发展。

1968年IBM公司推出了层次模型的IMS（Information Management System）数据库系统，1969年美国数据系统语言协会的数据库任务小组（DBTG）发表的系列报告提出了网状模型，1970年IBM研究中心的研究人员发表了关于关系模型的著名论文。这些事件奠定了现代数据库技术的基础。

20世纪70年代和80年代是数据库蓬勃发展的时期，不仅推出了一些网状模型数据库系统和层次模型数据库系统，还围绕关系数据模型进行了大量的研究和开发工作，关系数据库理论和关系模型数据库系统日趋完善。因为关系模型数据库本身具有的优点，它逐渐取代了网状模型数据库和层次模型数据库。到目前为止，关系模型数据库系统仍然是最重要的数据库系统。

20世纪90年代，关系模型数据库技术又有了进一步的改进。由于受到计算机应用领域及其他分支学科的影响，数据库技术与面向对象技术、网络技术相互渗透，产生了面向对象数据库和网络数据库。进入21世纪后，面向对象数据库和网络数据库技术逐渐成熟并得到了广泛的应用。

近40年来，数据库技术已经经历了3次演变，形成了以数据建模和数据库管理系统为核心，具有较完备的理论基础和广泛的应用领域的成熟技术体系，已成为计算机软件领域的一个重要分支。通常，人们把早期的层次模型数据库和网状模型数据库系统称为第一代数据库系统，把当前流行的关系模型数据库称为第二代数据库系统，当前正在发展的数据库系统称为第三代数据库系统。

我国有关部委、国防、气象和石油等行业开始使用数据库始于20世纪70年代。而数据库技术得到真正的广泛应用是从20世纪80年代初的DBaseII开始的。尽管DBase系列和XBase系列都不能称为一个完备的关系数据库管理系统，但是它们都支持关系数据模型，使用起来也非常方便，加上该系统是在微型计算机上实现的，一般也能满足中、小规模的管理信息系统的需要，所以得到了较广泛的应用，为数据库技术的普及奠定了基础。

数据库系统的出现使信息系统从以加工数据的程序为中心转向围绕共享的数据为中心的新阶段。这样既便于数据的集中管理，又有利于应用程序的研制和维护，提高了数据的利用率和相容性。20世纪80年代后不仅在大型机上，而且在大多数微型机上也配置了数据库管理系统，使数据库技术得到了更加广泛的应用与普及。无论是小型事务处理、信息处理系统、联机事务处理和联机分析处理，还是一般企业管理和计算机辅助设计和制造（CAD/CAM）及管理信息系统，都应用了数据库技术。数据库技术的应用程度已经成为衡量企业信息化程度的重要标志之一。

1.1.1 信息与社会

计算机所处理的数据在计算机中的存储方式与在现实生活中人们所面对的事物是有区别的。人们在现实生活中所面对的所有事物都是能够看得见真实存在的，如何把现实中能够“看得见”、“摸得着”的事物变成计算机能够处理的数据，这中间需要一个复杂的转换过程。比如，如何认识、理解、整理、描述和加工现实生活中的这些事物。从数据转化的顺序来说，数据从现实世界进入到数据库需要经历3个阶段，即现实世界阶段、信息世界阶段和机器世界阶段。

现实世界就是人们所生活的客观世界，客观世界存在着的形形色色的事物。

虽然现实世界的事物不能被改变，但仍然可以利用这些事物为人类的生活或生产服务，也就是说，可以在掌握和理解这些事物的基础上抽象出一些特殊的、有意义的信息。这些信息与现实世界的客观事物的根本区别就在于它们是经人类抽象和概念化了的、反映在人们心目中的信息，其中只包含着人们关心的那部分信息。这些信息就构成了信息世界。

尽管信息世界的信息是经过抽象和概念化了的，但它们仍然是计算机无法识别的，所以要对这些信息重新进行加工和转换，使它们能够被计算机所识别，使它们成为计算机能够处理和操作的符号。这些符号，又叫作数据。这些数据构成了机器世界或称为数据世界。

对信息社会而言，事物、信息、数据分别对应着现实世界、信息世界和机器世界，它们之间的对应关系如图1-1所示。

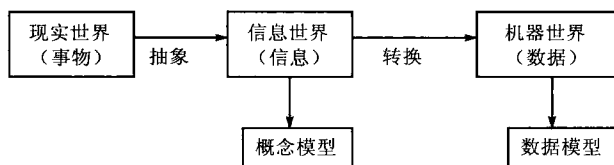


图 1-1 3 个世界之间的对应关系

1. 现实世界

在现实世界中客观存在着各种运动着的事物，各种事物及事物之间也存在着复杂的联系。不同事物之间，存在着不同的特征，这些特征包括静态的和动态的特征。所有的这些特征就是区别于不同事物的标志。在这些特征中，可以抽取一些有意义的特征来描述不同的事物个体。比如，常选择姓名、学号、班级、籍贯等特征来描述一个学生，而描述一名教师的信息则常选择姓名、年龄、性别、籍贯、所属院系等特征。利用这些特征，就可以在表征各类不同事物的同时，将不同的事物区别开。

世界上的各种事物虽然千差万别，看起来相互独立，但实际上它们之间是互相联系的。因为事物的多样性，事物之间的联系也是多方面的。在应用中，人们只选择那些有意义或感兴趣的联系，而没有必要选择所有的联系。例如，在教学管理系统中，教师和学生之间可以仅选择了“教学”这种有意义的联系。有时又称这种联系为关联。

2. 信息世界

现实世界中的事物及其联系由人们的感观所感知，经过大脑的分析、归纳、抽象形成信息。对这些信息进行记录、整理、归纳和格式化后就构成了信息世界。为了正确直观地反映客观事物及其联系，有必要对所研究的信息世界建立一个抽象的模型，称为信息模型（概念模型）。

在信息世界中，数据库技术涉及以下概念。

（1）实体（Entity）

在现实世界客观存在并可以相互区别的事物被抽象为实体。一个实体对应了现实世界中的一个事物。实体可以是具体的人、事、物，如一本书、一件衣服、一次借书、一次服装展示等，可包含很多我们感兴趣的信息，也可以是抽象的概念或联系，如教师与学院的工作关系（即某位教师在某学院工作）也可以被抽象为一个实体。

（2）实体集（Entity Set）

性质相同的同类实体组成的集合，称为实体集。在现实世界中的事物有很多，有一些事物具有被关注的一些共同的特征和性质，它们可以有类似的描述，可以被放在一起进行研究和处理。例如，一个学校的所有学生，当利用学籍管理系统进行管理时，这些学生的姓名、学号、班级和成绩等就是要关注的特征或性质，把这些学生的上述性质一起研究和处理，则这些学生就构成一个实体集。

（3）属性（Attribute）

客观存在的不同的事物，具有不同的特性。从客观世界抽象出来的不同实体，也具有其各自不同的特性。

实体所具有的某些特性称为属性。

可以用若干个属性来刻画一个实体。例如，对于大学生，有很多特性，如学校、学院、专业、班级、学号、姓名、身高、年龄、籍贯、成绩、入学时间等，这些属性组合起来共同表征了一个具体的学生。

也就是说，在信息世界里，人们对某个实体的认识和理解是通过属性来实现的。所以，要正确、全面地描述或者刻画某一个实体，就必须根据不同事物的特征，合理、全面地抽象出不同事物的属性，使人们通过这些属性，就能够对某个事物有一个全面的理解和把握。而且最重要的是，能够通过其中某一个或一些属性把握不同个体之间的本质的区别。

3. 机器世界

用计算机管理信息，必须对信息进行数字化，即将信息用字符和数字来表示。数字化后的信息称为数据，数据是能够被计算机识别并处理的。

当前多媒体技术的发展使计算机能够识别和处理图形、图像、声音等数据。数字化是信息世界到机器世界转换的关键，为数据管理打下了基础。信息世界的信息在机器世界中以数据形式存储。

机器世界对数据的描述常用到如下4个概念。

① 字段（Field）：又叫数据项，它是可以命名的最小信息单位。字段的定义包括字段名（字段的名称）、字段类型（描述该字段的数据类型）、字段长度（限定该字段值的长度）等。

② 记录（Record）：字段的有序集合称为记录，一般对应信息世界中的一个具体的实体。它是对一个具体对象的描述，如（2002178002，男，178），描述了一个学号为2002178002，性别为男，身高为178cm的学生。

③ 文件（File）：同类的记录汇集成文件。文件是描述实体集的。例如，所有图书记录组成了一个图书文件。

④ 关键字（Key）：能唯一标识文件中每个记录的字段或字段集。例如，学生的学号可以作为学生记录的关键字，如果一个字段不能唯一确定一条记录，则可以用多个字段作为关键字来唯一标识一条记录。例如，学号与课程号可以作为学生选课记录的关键字。

机器世界和信息世界术语是相互对应的，它们的对应关系见表1-1。

在数据库中，每个概念都有类型（Type，简称型）和值（Value）的区别。例如，“学生”是一个实体的型，而具体的（张三，男，信息学院，28）是实体的值。又如，“姓名”是属性的型，而“张三”是属性的值。记录也有记录的型和值。有时在不引起误解的情况下，可以不去仔细区别型和值。

表 1-1 信息世界和机器世界的概念的对应关系

信息世界	机器世界
实体	记录
属性	字段
实体集	文件
码	关键字

为了理解上的方便，图 1-2 以学生为例表示了信息在 3 个世界中的有关概念及其联系。需要特别注意的是，实体与属性、型和值的区别，以及 3 个世界中各概念的相应关系。

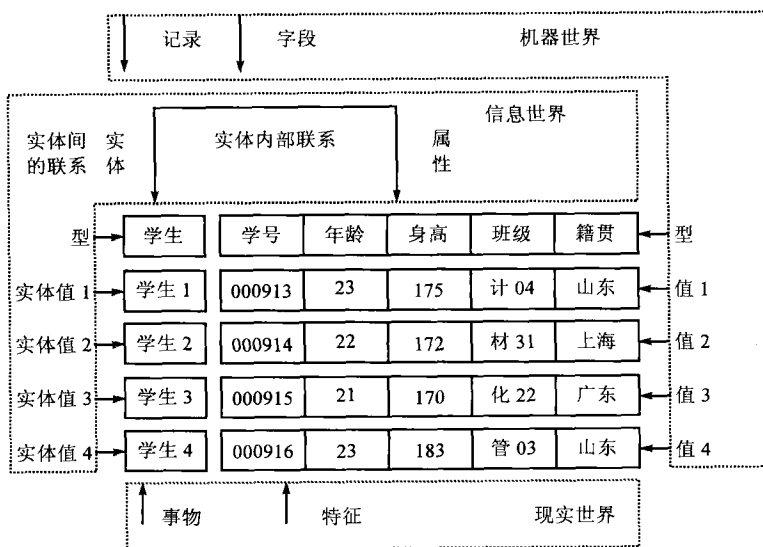


图 1-2 信息 3 个世界的术语联系

1.1.2 数据库的基本概念

1. 数据 (Data)

数据是在数据库中存储的基本对象，是用来记录现实世界的信息，并可以被机器识别的符号。

在计算机领域里，数据这个概念已经不局限于普通意义上的数字了，凡是在计算机中用于描述事物特征的记录都可以称为数据，如文字、图形、图像、声音等。例如，当用书号、书名、单价、作者、出版社这几个特征来描述某本书时，(000913, 数据库, 47.00, 张三, 电子)就是一本书的数据。于是，就可以从这一数据的含义中得到数据库这本书的有关信息。

数据有一定的格式，如书号一般为长度不超过 13 位的数字字符，单价为小数位数为 2 位的实数。这些格式的规定就是数据的语法，而数据的含义就是数据的语义。通过解释、推理、归纳、分析和综合等，从数据所获得的有意义的内容称为信息。因此，数据是信息存在的一种形式。只有通过解释或处理的数据才能成为有用的信息。

2. 数据库 (DB, DataBase)

数据库是以一定的组织形式存储在一起的，能为多个用户所共享，相互关联的数据集