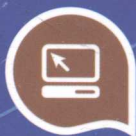


全国高等院校计算机基础教育“十三五”规划教材

数据库技术及应用

SHUJUKU JISHU JI YINGYONG

齐 晖 潘惠勇 主 编
吴 婷 王 琳 郭 飞 副主编



中国铁道出版社
CHINA RAILWAY PUBLISHING HOUSE

全国高等院校计算机基础教育“十三五”规划教材

数据库技术及应用

齐 晖 潘惠勇 主 编
吴 婷 王 琳 郭 飞 副主编

常州大学图书馆
藏书章

中国铁道出版社
CHINA RAILWAY PUBLISHING HOUSE

内 容 简 介

本书是根据教育部高等学校计算机基础课程教学指导委员会制定的《高等学校计算机基本教学发展战略研究暨计算机基础课程教学基本要求》(2009版)、新的《全国计算机等级考试(NCRE)二级 Access 数据库程序设计考试大纲》的要求精心组织编写而成的。

本书基于 Microsoft Access 2010 系统地介绍了关系数据库管理系统的基础知识及 Access 2010 的功能和基本操作, 主要内容包括: 数据库基础概述, Access 2010 系统概述, Access 数据库的建立、使用、维护和管理等, 结构化查询语言 SQL 的使用, 创建数据库的各种对象(包括表、查询、窗体、报表、宏和模块), VBA 编程, 数据库编程, 数据库应用系统开发实例等内容。

本书注重理论与实践的结合, 力求通俗易懂, 并配有适量的图片和示例, 适合作为高等院校 Access 数据库技术及应用课程的教材, 也可作为 Access 数据库应用系统开发人员的参考书或者自学者的自学教材。

图书在版编目(CIP)数据

数据库技术及应用/齐晖, 潘惠勇主编. —北京:
中国铁道出版社, 2017. 12
全国高等院校计算机基础教育“十三五”规划教材
ISBN 978-7-113-24184-1

I. ①数… II. ①齐… ②潘… 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(2017)第 319754 号

书 名: 数据库技术及应用
作 者: 齐 晖 潘惠勇 主编

策 划: 韩丛付 周海燕
责任编辑: 周海燕 包 宁
封面设计: 刘 颖
责任校对: 张玉华
责任印制: 郭向伟

读者热线: (010) 63550836

出版发行: 中国铁道出版社(100054, 北京市西城区右安门西街 8 号)

网 址: <http://www.tdpress.com/51eds/>

印 刷: 三河市宏盛印务有限公司

版 次: 2017 年 12 月第 1 版 2017 年 12 月第 1 次印刷

开 本: 880 mm × 1 230 mm 1/16 印张: 18.5 字数: 592 千

书 号: ISBN 978-7-113-24184-1

定 价: 49.00 元

版权所有 侵权必究

凡购买铁道版图书, 如有印制质量问题, 请与本社教材图书营销部联系调换。电话: (010) 63550836

打击盗版举报电话: (010) 51873659

由于数据库技术在国民经济、科技文化和国防建设等诸多领域的广泛应用,基于数据库技术和数据库管理系统的应用软件的研发和使用已经成为各专业领域和管理人员必备的基础。因此,“数据库技术及应用”课程已经成为高等院校理工、经管、文科三大类学生的计算机基础教学核心课程和必修课程。

Access 是 Microsoft Office 系列应用软件之一,是一个功能强大且易于实现和使用的关系型数据库管理系统,既具有典型的 Windows 应用程序风格,也具备可视化及面向对象程序设计的特点。Access 能有效地组织、管理和共享数据库的数据信息,把数据库和网络结合起来,为用户在网络中共享信息奠定了基础。Access 可以直接开发一个小型的数据库管理系统,也可以作为一个中小型管理信息系统的数据库部分,还可以作为一个商务网站的后台数据库部分,Access 有着相当广泛的用户群。目前已有大量的基于 Access 数据库的应用在 Internet 上发布,并且其数量呈快速上升之势。同时,Access 概念清晰、简单易用、功能完备,尤其适合数据库技术的初学者。

本书是根据教育部高等学校计算机基础课程教学指导委员会制定的《高等学校计算机基本教学发展战略研究暨计算机基础课程教学基本要求》(2009 版)、新的《全国计算机等级考试(NCRE)二级 Access 数据库程序设计考试大纲》的要求精心组织编写而成的。

本书考虑到学生计算机实际的操作技能和学习特点,以应用为目的,以案例为引导,通过对大量实例的分析和讲解,采用图文并茂、通俗易懂的形式,以 Access 2010 为背景,循序渐进地介绍数据库系统的基本知识;Access 数据库的建立、使用、维护和管理等;结构化查询语言 SQL 的使用;VBA 数据库编程基础;一个小型数据库应用系统开发实例等内容。在本书编写和组织上力求避免术语、概念的枯燥讲解和操作的简单堆砌,学生只要参照书中提供的实例进行系统学习,并配合一定的上机实际操作,就能很快掌握 Access 数据库管理系统的基本功能和操作,掌握面向应用开发的系统知识,并能够学以致用地完成简单实用的小型数据库管理系统的开发。

本书注重理论联系实际,条理清晰,概念明确,注重实际操作技能的训练。理论部分的讲解以“学籍管理”的设计为案例展开,全书的最后以“学籍管理”系统的开发为例,详细地介绍了如何构建一个小型数据库应用系统。书中配有大量的习题,包括选择题、判断题、填空题、简答题以及上机操作题等,供学生复习和上机练习使用。在上机练习时,可以举一反三,以达到熟练操作的目的。

本书共分 11 章,从各个方面介绍 Access 2010 的功能。

第 1 章主要介绍数据库技术的发展、数据库的基本概念、关系数据库系统、数据库设计基础等内容,通过这些内容使读者掌握数据库的基本概念、理论和设计方法。

第 2 章介绍 Access 2010 的特点、界面;Access 的 6 种对象,数据库的创建和基本操作。使读者对 Access 2010 从宏观上有一定的了解。

第 3 章、第 4 章、第 5 章、第 6 章、第 7 章重点介绍 Access 数据表的建立,创建和使用查询,窗体、报表的建立,宏等内容,这些是 Access 中最基本的内容。

第 8 章通过典型的实例介绍 Access 的编程语言 VBA,内容主要有模块的概念、模块的建立、VBA 程序设计的基本方法、过程的建立、调用与参数传递、事件驱动的程序设计方法、DoCmd 对象以及 VBA 程序调试和错误处理方法等。这部分是本书的难点,也是开发应用程序的基础。

第 9 章介绍如何使用 Access 2010 提供的安全功能来实现数据库的安全操作和数据库的导入/导出功能,以及

利用 SharePoint 和 Access 2010 对数据库进行发布和协作管理数据库。

第 10 章主要介绍 VBA 数据库编程, 内容包括 VBA 提供的 3 种数据库访问接口; DAO 的引用方法、DAO 的模型层次结构和访问数据库的步骤; ADO 的引用方法、ADO 的模型层次结构和主要 ADO 对象的使用。

第 11 章介绍数据库管理系统开发的一般流程, 并以“学籍管理”系统的建立, 介绍如何将数据库的各个对象有机地联系起来, 构建一个小型数据库应用系统。

本书参考学时不得低于 60 学时, 其中 32 学时理论教学, 28 学时上机实践。本书可作为高等学校 Access 数据库技术及应用课程的教材, 也可作为 Access 数据库应用系统开发人员的参考书或者自学者的自学教材。

本书由中原工学院计算机学院基础教学部的 9 位教师集体写作完成, 他们是杨要科、吴婷、金秋、马宗梅、郭飞、齐晖、王琳、程传鹏、潘惠勇。齐晖、潘惠勇任主编, 吴婷、王琳、郭飞任副主编。全书由齐晖、潘惠勇审阅并统稿。

本书中所有的例题和实例均在 Access 2010 中运行通过。为了便于教师使用本书和学生学习, 本书配有电子教案、MOOC 网站, 以及本书案例中的素材, 有需要者请登录 <http://mooc1.zut.edu.cn/course/87180281.html> 下载 (课程资源下载如有问题, 请联系我们: qihui63@126.com)。

由于编者学识水平所限, 书中难免有不妥之处, 望读者不吝指正。

编 者

2017 年 11 月

第 1 章 数据库基础概述.....1

- 1.1 数据库基本概念.....1
 - 1.1.1 计算机数据管理的发展.....1
 - 1.1.2 数据库系统.....6
 - 1.1.3 数据模型.....7
- 1.2 关系数据库.....9
 - 1.2.1 关系数据模型.....9
 - 1.2.2 关系运算.....11
 - 1.2.3 关系的完整性.....12
- 1.3 数据库设计基础.....13
 - 1.3.1 数据库设计原则.....13
 - 1.3.2 数据库设计步骤.....13
 - 1.3.3 数据库设计过程.....14

习题 1.....16

第 2 章 Access 2010 系统概述.....20

- 2.1 Access 2010 的新功能.....20
- 2.2 Access 2010 数据库对象的组成.....21
 - 2.2.1 表.....21
 - 2.2.2 查询.....22
 - 2.2.3 窗体.....22
 - 2.2.4 报表.....22
 - 2.2.5 宏.....23
 - 2.2.6 模块.....23
- 2.3 Access 2010 的启动和退出.....23
 - 2.3.1 Access 2010 的启动.....23
 - 2.3.2 Access 2010 的退出.....24
- 2.4 Access 2010 的主界面.....24
- 2.5 Access 2010 的命令选项卡.....25
- 2.6 创建数据库.....26
 - 2.6.1 创建空数据库.....26
 - 2.6.2 利用样本模板创建数据库.....27
- 2.7 数据库的打开与关闭.....28
 - 2.7.1 打开数据库文件.....28
 - 2.7.2 关闭数据库文件.....29

习题 2.....29

第 3 章 数据表的设计.....31

- 3.1 数据表概述.....31
 - 3.1.1 字段的命名.....31
 - 3.1.2 字段的数据类型.....32
 - 3.1.3 数据表结构设计实例.....33
- 3.2 创建数据表.....34
 - 3.2.1 创建数据表.....34
 - 3.2.2 字段属性的设置.....37
 - 3.2.3 索引和主键.....40
 - 3.2.4 编辑数据表的结构.....42
 - 3.2.5 编辑数据表的数据.....43
 - 3.2.6 数据表的操作.....44
- 3.3 数据表的关系.....45
 - 3.3.1 表之间关系的建立.....45
 - 3.3.2 维护数据完整性和精确性.....46
 - 3.3.3 关系的查看与编辑.....47
 - 3.3.4 子数据表的属性及使用.....47
- 3.4 数据表的优化与调整.....48
 - 3.4.1 排序表数据.....48
 - 3.4.2 筛选表数据.....49
 - 3.4.3 查找与替换.....51
 - 3.4.4 外观的设置.....52

习题 3.....54

第 4 章 使用查询搜索信息.....57

- 4.1 查询概述.....57
 - 4.1.1 一个查询的例子.....57
 - 4.1.2 查询的主要功能.....58
 - 4.1.3 查询的类型.....58
 - 4.1.4 查询的视图模式.....58
- 4.2 使用向导创建查询.....59
 - 4.2.1 使用向导创建简单查询.....59
 - 4.2.2 使用“查找重复项查询向导”
创建查询.....62

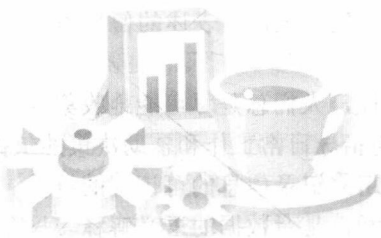


4.2.3 使用“查找不匹配项查询向导” 创建查询.....	64	5.6.1 窗体的设计视图.....	105
4.2.4 使用“交叉表查询向导” 创建查询.....	65	5.6.2 常用控件.....	109
4.3 查询中的表达式.....	68	5.6.3 控件的基本类型.....	110
4.3.1 运算符.....	68	5.6.4 属性和事件.....	110
4.3.2 函数.....	71	5.6.5 控件的用法.....	111
4.4 使用设计视图创建查询.....	74	5.6.6 使用控件自定义窗体实例.....	113
4.4.1 基本查询.....	75	5.7 优化窗体的外观.....	127
4.4.2 含新字段查询.....	76	5.7.1 应用主题.....	127
4.4.3 选择查询.....	79	5.7.2 设置窗体的格式属性.....	127
4.4.4 参数查询.....	81	5.7.3 在窗体中添加图片.....	128
4.4.5 排序查询.....	82	习题 5.....	128
4.4.6 分组统计查询.....	83	第 6 章 报表.....	131
4.4.7 交叉表查询.....	85	6.1 认识报表.....	131
4.5 创建操作查询.....	87	6.1.1 报表的功能.....	131
4.5.1 生成表查询.....	87	6.1.2 报表的类型.....	131
4.5.2 追加查询.....	88	6.1.3 报表的视图.....	133
4.5.3 更新查询.....	89	6.1.4 报表的结构.....	133
4.5.4 删除查询.....	90	6.2 创建报表.....	134
4.6 创建 SQL 查询.....	91	6.2.1 快速创建报表.....	135
4.6.1 SQL 语句简介.....	91	6.2.2 向导创建报表.....	135
4.6.2 SQL 查询语句.....	91	6.2.3 创建标签式报表.....	137
4.6.3 SQL 数据操作语句.....	93	6.2.4 创建图表报表.....	138
4.6.4 SQL 数据定义语句.....	93	6.2.5 创建复杂报表.....	140
习题 4.....	94	6.3 高级报表.....	143
第 5 章 创建和使用窗体.....	98	6.3.1 分组、排序和汇总.....	143
5.1 窗体概述.....	98	6.3.2 报表计算.....	146
5.1.1 窗体的概念.....	98	6.3.3 主/子报表.....	148
5.1.2 窗体的视图.....	99	6.3.4 交叉报表.....	151
5.2 创建窗体.....	99	6.3.5 弹出式模式报表.....	153
5.3 使用“窗体”工具创建窗体.....	100	6.4 报表技巧.....	153
5.4 使用“窗体向导”工具创建窗体.....	101	6.4.1 为控件设置与众不同的名称.....	154
5.5 使用“其他窗体”工具创建窗体.....	102	6.4.2 在打印预览时隐藏其他窗体.....	154
5.5.1 创建多项目窗体.....	102	6.4.3 向报表中添加更多信息.....	154
5.5.2 创建数据表窗体.....	102	6.4.4 隐藏页面页眉.....	155
5.5.3 创建分割窗体.....	103	6.4.5 增加运行时的强调效果.....	155
5.5.4 创建“数据透视表”窗体.....	103	6.4.6 避免空报表.....	156
5.5.5 创建“数据透视图”窗体.....	105	6.5 报表布局.....	156
5.6 使用“窗体设计”工具自定义窗体.....	105	6.5.1 布局的概念.....	156
		6.5.2 布局的创建.....	156
		6.5.3 使用布局快速更改设计.....	157



6.6 美化报表.....	158	8.3.4 输入/输出语句.....	196
6.6.1 报表的外观设置.....	158	8.4 VBA 程序流程控制语句.....	198
6.6.2 在报表中使用条件格式.....	158	8.4.1 顺序结构.....	198
6.6.3 添加图像和线条.....	159	8.4.2 选择结构.....	198
6.6.4 插入日期和时间.....	160	8.4.3 循环结构.....	202
6.6.5 插入页码.....	160	8.5 数组.....	205
6.7 报表的打印.....	160	8.5.1 数组的概念.....	205
6.7.1 报表的页面设置.....	160	8.5.2 数组的声明.....	206
6.7.2 打印报表.....	161	8.5.3 数组的应用.....	206
习题 6.....	161	8.6 过程的创建和调用.....	207
第 7 章 宏.....	163	8.6.1 Sub 子过程.....	208
7.1 宏概述.....	163	8.6.2 Function 函数过程.....	209
7.1.1 宏的基本概念.....	163	8.6.3 过程调用中的参数传递.....	210
7.1.2 常用宏操作.....	164	8.7 事件及事件驱动.....	211
7.1.3 设置宏操作参数.....	165	8.7.1 事件及事件驱动的定义.....	211
7.2 宏的创建.....	165	8.7.2 事件的分类.....	212
7.2.1 创建独立宏.....	165	8.7.3 事件驱动的程序设计方法.....	217
7.2.2 创建嵌入宏.....	168	8.8 DoCmd 对象.....	218
7.2.3 创建数据宏.....	173	8.8.1 程序导航.....	218
7.2.4 宏的编辑.....	175	8.8.2 控制大小和位置.....	220
7.3 宏的运行.....	176	8.9 程序调试和错误处理.....	220
7.3.1 独立宏的运行.....	176	8.9.1 调试工具的使用.....	221
7.3.2 嵌入宏的运行.....	177	8.9.2 调试菜单.....	221
7.3.3 数据宏的运行.....	177	8.9.3 VBA 的错误类型及处理方式.....	224
7.4 宏的调试.....	177	8.9.4 VBA 程序调试的方法.....	225
习题 7.....	178	习题 8.....	232
第 8 章 模块和 VBA 程序设计.....	180	第 9 章 数据库的管理与数据的共享.....	236
8.1 模块的基本概念.....	180	9.1 数据库密码的设置.....	236
8.1.1 模块简介.....	180	9.2 数据库的压缩与修复.....	237
8.1.2 VBA 程序设计概述.....	182	9.2.1 手动压缩和修复 Access 数据库.....	238
8.1.3 模块的编程界面 VBE.....	183	9.2.2 设置关闭时自动压缩数据库.....	238
8.2 VBA 程序设计基础.....	188	9.3 数据库的备份和恢复.....	239
8.2.1 数据类型.....	188	9.3.1 备份数据库.....	239
8.2.2 常量和变量.....	189	9.3.2 用备份副本还原 Access 数据库 文件.....	240
8.2.3 运算符.....	190	9.4 使用信任中心.....	240
8.2.4 VBA 常用函数.....	193	9.4.1 启用禁用内容.....	241
8.3 VBA 常用语句.....	195	9.4.2 使用受信任位置中的 Access 2010 数据库.....	241
8.3.1 语句的书写规则.....	195	9.5 数据库的导入和导出.....	242
8.3.2 注释语句.....	195	9.5.1 数据的导入并链接.....	242
8.3.3 VBA 赋值语句.....	196		

9.5.2 数据的转换.....	246	10.3 ActiveX 数据对象.....	259
9.5.3 数据的导出.....	246	10.3.1 ADO 库的引用方法.....	259
9.6 生成 .accde 文件.....	248	10.3.2 ADO 模型层次结构.....	259
9.7 利用 Access 和 SharePoint 管理数据.....	249	10.3.3 主要 ADO 对象的使用.....	260
9.7.1 SharePoint Server 2010 的安装.....	249	10.4 特殊域聚合函数和 RunSQL 方法.....	264
9.7.2 将数据库发布到 SharePoint 网站.....	250	10.5 综合案例.....	266
9.7.3 链接 SharePoint 列表到 Access 2010.....	252	习题 10.....	268
9.7.4 发布数据库到 SharePoint 网站.....	253	第 11 章 数据库应用系统开发实例.....	271
9.7.5 利用 SharePoint 网站和 Access 2010 协作完成数据的输入.....	254	11.1 系统开发的一般过程.....	271
习题 9.....	255	11.2 系统需求分析.....	272
第 10 章 VBA 数据库编程.....	256	11.3 系统设计.....	272
10.1 VBA 数据库编程相关基础知识.....	256	11.3.1 系统模块设计.....	272
10.1.1 VBA 数据库应用程序一般框架.....	256	11.3.2 数据库的设计.....	273
10.1.2 VBA 数据库访问接口.....	256	11.4 系统模块设计与实现.....	274
10.1.3 VBA 访问的数据类型.....	257	11.4.1 学生信息维护窗体.....	274
10.2 数据访问对象.....	257	11.4.2 学生信息查询窗体.....	274
10.2.1 DAO 库的引用方法.....	257	11.4.3 学生信息打印报表.....	277
10.2.2 DAO 模型层次结构.....	257	11.4.4 创建主控界面.....	277
10.2.3 利用 DAO 访问数据库的一般 步骤.....	258	11.4.5 创建登录窗体.....	281
		11.5 设置启动窗体.....	282
		11.6 生成 ACCDE 文件.....	283
		习题 11.....	284
		参考文献.....	286



第 1 章

数据库基础概述

自 1946 年计算机诞生以来,其主要应用是科学计算,自 20 世纪 60 年代以来,数据库技术也作为计算机数据处理的一门新技术发展了起来,它是计算机科学技术中发展最快的领域之一,经过 50 多年的发展形成了较为完整的理论体系,已被广泛应用于教学管理、科学研究、企业管理和社会服务等各个领域,是最先进的数据管理技术。本章主要介绍数据库中的基本概念、关系数据库、数据库设计等内容。

教学目标

- 了解数据管理的相关概念和含义、数据模型的主要类型与数据库的关系。
- 熟悉数据、数据库及数据库管理系统的含义和数据库设计的原则、步骤及过程。
- 掌握关系数据库及相关的概念、关系运算和关系的完整性。

1.1 数据库基本概念

自从计算机被发明之后,人类社会就进入了高速发展阶段,大量的信息堆积在人们面前。此时,如何组织存放这些信息,如何在需要时快速检索出信息,以及如何让所有用户共享这些信息就成为一个大问题。数据库技术就是在这种背景下诞生的,这也是使用数据库的原因。当今,世界上每一个人的生活几乎都离不开数据库。如果没有数据库,很多事情几乎无法解决。例如,没有学校的图书管理系统,借书会是一个很麻烦的事情,更不用说网上查询图书信息了;没有教务管理系统,学生要查询自己的成绩也不是很方便;没有计费系统,人们也就不能随心所欲地拨打手机;没有数据库的支持,网络搜索引擎就无法继续工作,网上购物就更不用想了。可见,数据库应用已经遍布了人们生活的各个角落。

1.1.1 计算机数据管理的发展

现代意义上的数据库系统出现于 20 世纪 60 年代后期,伴随着计算机硬件系统的飞速发展、价格的逐步下降、操作系统性能的日益提高以及 1970 年前后关系型数据模型的出现,数据库技术正广泛应用于各个领域,可以说我们已经无法离开数据库系统。

1. 数据和信息

在数据处理中,最常用到的基本概念就是数据和信息。

数据是指描述事物的符号记录。数据不仅仅是指传统意义的由 0~9 组成的数字,而是所有可以输入到计算机中并能被计算机处理的符号的总称。

在计算机中可表示数据的种类很多,除了数字以外,文字、图形、图像、声音都是数据。例如学生的基本情况、超市商品的价格、员工的照片、人的指纹、播音员朗诵的佳作、气象卫星云图等都可以是数据。

信息是指以数据为载体的对客观世界实际存在的事物、事件和概念的抽象反映。具体说是一种被加工为特定形式的数据,是通过人的感官(眼、耳、鼻、舌、身)或各种仪器仪表和传感器等感知出来并经过加工而形成的反映现实世界中事物的数据。

例如,在学生档案中,记录了学生的姓名、性别、年龄、出生日期、籍贯、所在系别、入学时间,那么下面的描述:

(李军,男,21,1993,四川,外语系,2012)就是数据。对于这条学生记录,所表述的信息为:

李军是个大学生,1993年出生,男,四川人,2012年考入外语系。

数据是数据库的基本组成内容,是对客观世界所存在的事物的一种表征,人们总是尽可能地收集各种各样的数据,然后对其进行加工处理,从中抽取并推导出有价值的信息,作为指导日常工作和辅助决策的依据。

数据和信息是两个互相联系、互相依赖但又互相区别的概念。数据是用来记录信息的可识别的符号,是信息的具体表现形式。数据是信息的符号表示或载体,信息则是数据的内涵,是对数据的语义解释。只有经过提炼和抽象之后,具有使用价值的信息才能成为信息。

2. 数据处理和数据管理

数据要经过处理才能变为信息,这种将数据转换成信息的过程称为数据处理。数据处理具体是指对信息进行收集、整理、存储、加工及传播等一系列活动的总和。数据处理的目的是从大量的、杂乱无章的甚至是难于理解的原始数据中,提炼、抽取出人们所需要的有价值、有意义的信息,作为科学决策的依据。

可用“信息 = 数据 + 数据处理”简单地表示信息、数据与数据处理的关系。

数据是原料,是输入,而信息是产出,是输出结果。数据处理的真正含义是为了产生信息而处理数据。数据、数据处理、信息的关系如图 1.1 所示。



图 1.1 数据、数据处理、信息的关系

数据的组织、存储、检查和维护等工作是数据处理的基本环节,这些工作一般统称为数据管理。数据管理的核心是数据管理。

数据处理与数据管理是相互联系的,数据管理技术的优劣,将直接影响数据处理的效率。

3. 计算机数据管理的发展阶段

计算机在数据管理方面经历了从低级到高级的发展过程,到目前为止,数据管理大致经历了人工管理、文件系统、数据库系统三个阶段。

1) 人工管理阶段

这一阶段(20世纪50年代中期以前)计算机主要用于科学计算。外部存储器只有磁带、卡片和纸带,软件只有汇编语言,还没有数据管理方面的软件。数据处理的方式基本上是批处理。这个时期数据管理具有以下几个特点。

(1) 数据不保存。因为当时计算机主要用于科学计算,对于数据保存的需求尚不迫切。需要时把数据输入内存,运算后将结果输出。数据并不保存在计算机中。

(2) 没有专用的软件对数据进行管理。在应用程序中,不仅要管理数据的逻辑结构,还要设计其物理结构、存取方法、输入/输出方法等。当存储改变时,应用程序中存取数据的子程序就需随之改变。

(3) 数据不具有独立性。数据的独立性是指逻辑独立性和物理独立性。当数据的类型、格式或输入/输出方式等逻辑结构或物理结构发生变化时,必须对应用程序做出相应的修改。

(4) 数据是面向程序的。一组数据只对应于一个应用程序。即使两个应用程序都涉及某些相同数据,也必须各自定义,无法相互利用。因此,在程序之间有大量的冗余数据。这时期数据与程序的关系如图 1.2 所示。

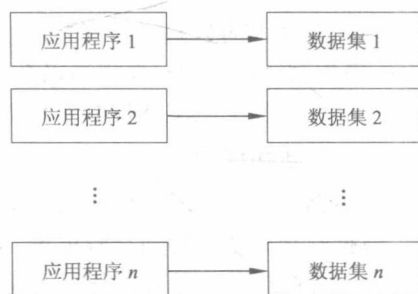


图 1.2 人工管理阶段数据与程序的关系

2) 文件系统阶段

这一阶段(20世纪50年代后期到60年代中期)计算机不仅用于科学计算,还用于信息管理。此时,外部存储器已有磁盘、磁鼓等直接存取的存储设备;软件领域出现了高级语言和操作系统。操作系统中的文件系统是专门的数据管理软件。这时可以把相关的数据组成一个文件存放在计算机中,在需要时只要提供文件



名, 计算机就能从文件系统中找出所要的文件, 把文件中存储的数据提供给用户进行处理。这个时期数据管理具有以下几个特点。

(1) 数据以“文件”形式可长期保存在外部存储器的磁盘上。应用程序可对文件进行大量的检索、修改、插入和删除等操作。

(2) 文件组织已多样化。有索引文件、顺序存取文件和直接存取文件等。因而对文件中的记录可顺序访问, 也可随机访问, 便于存储和查找数据。

(3) 数据与程序间有一定的独立性。数据由专门的软件即文件系统进行管理, 程序和数据间由软件提供的存取方法进行转换, 数据存储发生变化不一定影响程序的运行。

(4) 对数据的操作以记录为单位。这是由于文件中只存储数据, 不存储文件记录的结构描述信息。文件的建立、存取、查询、插入、删除、修改等所有操作, 都要用程序来实现。

在文件系统阶段, 仍有很多缺点。主要表现在以下几个方面。

(1) 数据冗余度大。由于各数据文件之间缺乏有机的联系, 造成每个应用程序都有对应的文件, 有可能同样的数据在多个文件中重复存储, 数据不能共享。

(2) 数据独立性低。数据和程序相互依赖, 一旦改变数据的逻辑结构, 必须修改相应的应用程序。而应用程序发生变化, 如改用另一种程序设计语言来编写程序, 也需修改数据结构。

(3) 数据一致性差。由于相同数据的重复存储、各自管理, 在进行更新操作时, 容易造成数据的不一致。

这样, 文件系统仍然是一个不具有弹性的无结构的数据集合。文件之间是孤立的、不能反映现实世界中事物之间的内在联系。这时期数据与程序的关系如图 1.3 所示。

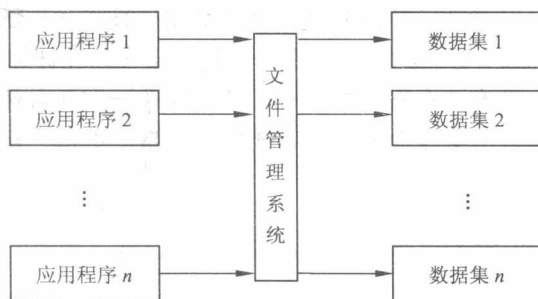


图 1.3 文件系统阶段数据与程序的关系

3) 数据库系统阶段

数据管理技术进入数据库系统阶段是在 20 世纪 60 年代末。由于计算机应用于管理的规模更加庞大, 数据量急剧增加; 硬件方面出现了大容量磁盘, 使计算机联机存取大量数据成为可能; 硬件价格下降, 而软件价格上升, 使开发和维护系统软件的成本增加。文件系统的管理方法已无法适应开发应用系统的需要。为解决多用户、多个应用程序共享数据的需求, 出现了统一管理数据的专门软件系统, 即数据库管理系统。这使利用数据库技术管理数据变成了现实。这时期数据管理的特点有以下几方面。

(1) 数据共享性高、冗余度低。这是数据库系统阶段的最大改进, 数据不再面向某个应用程序而是面向整个系统, 当前所有用户可同时访问数据库中的数据。这样就减少了不必要的冗余, 节约了存储空间, 同时也避免了数据之间的不相容性与不一致性。

(2) 数据结构化。即按照某种数据模型, 将应用的各种数据组织到一个结构化的数据库中。在数据库中数据的结构化, 不仅要考虑某个应用的数据结构, 还要考虑整个系统的数据结构, 并且还要能够表示出数据之间的有机关联。

(3) 数据独立性高。数据的独立性是指逻辑独立性和物理独立性。数据的逻辑独立性是指当数据的总体逻辑结构改变时, 数据的局部逻辑结构不变。由于应用程序是依据数据的局部逻辑结构编写的, 所以应用程序不必修改, 从而保证了数据与程序间的逻辑独立性。数据的物理独立性是指当数据的存储结构改变时, 数据的逻辑结构不变, 从而应用程序也不必改变。

(4) 有统一的数据控制功能。数据库为多个用户和应用程序所共享, 对数据的存取往往是并发的, 即多个用户可以同时存取数据库中的数据, 甚至可以同时存取数据库中的同一个数据。为确保数据库数据的正确有效和数据库系统的有效运行, 数据库管理系统提供下述四方面的数据控制功能。

① 数据的安全性控制: 防止不合法使用数据造成数据的泄露和破坏, 保证数据的安全和机密。例如, 系统提供口令检查或其他手段来验证用户身份, 防止非法用户使用系统; 也可以对数据的存取权限进行限制, 只有通过检查后才能执行相应的操作。

② 数据的完整性控制：系统通过设置一些完整性规则以确保数据的正确性、有效性和相容性。正确性是指数据的合法性，如年龄属于数值型数据，只能包含 0, 1, ..., 9，不能包含字母或特殊符号。有效性是指数据是否在其定义的有效范围内，如月份只能用 1~12 之间的正整数表示。相容性是指表示同一事实的两个数据应相同，否则就不相容，如一个人不能有两个性别。

③ 并发控制：防止多用户同时存取或修改数据库时，因相互干扰而提供给用户不正确的数据，并使数据库受到破坏。

④ 数据恢复：当数据库被破坏或数据不可靠时，系统有能力将数据库从错误状态恢复到最近某一时刻的正确状态。

这时期数据与程序之间的关系如图 1.4 所示。

4. 数据库系统的新技术

随着科学技术和数据库系统的发展，从 20 世纪 80 年代开始数据库研究又出现了许多新的领域，相继研究出了分布式数据库系统、面向对象数据库系统和网络数据库系统。

1) 分布式数据库系统

20 世纪 80 年代，随着数据库技术的广泛应用，并与迅速发展的网络技术相结合，产生了分布式数据库系统。分布式数据库是一个物理上分布在计算机网络的不同结点，但在逻辑上又同属于一个系统的数据集合。在分布式数据库系统中，数据库存储在几台计算机中，这几台计算机之间通过高速网络相互通信，计算机之间没有共享公共的内存或磁盘，系统中每一台计算机称为一个结点。其一般结构如图 1.5 所示。

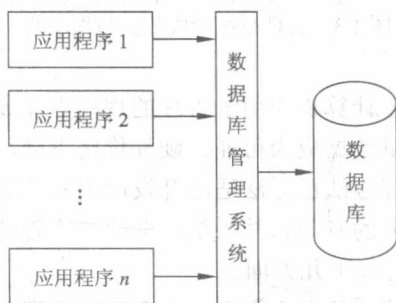


图 1.4 数据库系统阶段数据与程序的关系

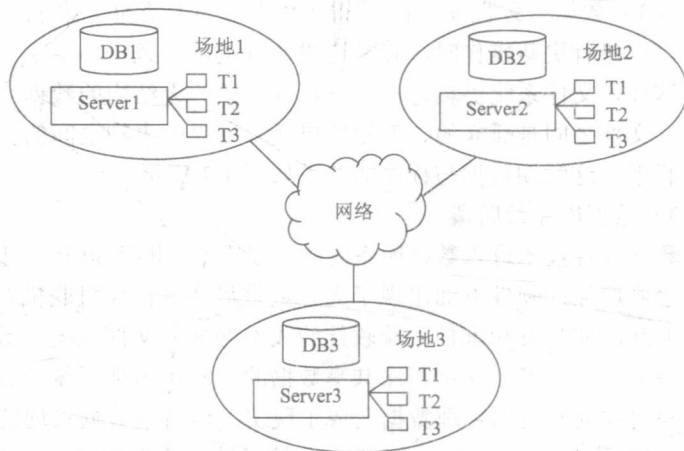


图 1.5 分布式数据库系统

在分布式数据库系统中，应用分为局部应用和全局应用两种。局部应用是指仅操作本地结点上数据库的应用；而全局应用是指需要操作两个或两个以上结点中的数据库的应用。例如，一个银行系统中，有多个分支机构分布在不同的城市，每个分支机构有自己的服务器（结点），用来维护该分支机构的所有账户的数据库。同时有若干客户机，用来完成本地客户的存、取款业务等（局部应用）。分支机构的客户机也可以完成某些全局应用，如不同分支机构中账户之间的转账，就需要同时访问和更新两个结点上的数据库中的数据。不支持全局应用的系统不能称为分布式数据库系统。同时，分布式数据库系统不仅要求数据的物理分布，而且要求这种分布是面向处理、面向应用的。

分布式数据库系统是物理上分散，逻辑上集中的数据库系统，系统中的数据分别存放在计算机网络的不同结点上，网络中的每个结点具有独立处理的能力（称为场地自治），可以执行局部应用，每个结点也可以通过网络通信子系统执行全局应用。

分布式数据库系统主要有如下几个特点。

(1) 数据的物理分布性。数据库中的数据分布在计算机网络的不同结点上，而不是集中在一个结点上。因此它不同于通过计算机网络共享的集中式数据库系统。

(2) 数据的逻辑整体性。分布在计算机网络不同结点上的数据在逻辑上属于同一个系统,因此,它们在逻辑上是相互联系的整体。

(3) 结点的自主性。每个结点有自己的计算机、数据库(即局部数据库,简称 LDB)、数据库管理系统(LDBMS),因而能独立地管理局部数据库。局部数据库中的数据可以供本结点的用户存取(局部应用),也可以供其他结点上的用户存取以供全局应用。

2) 面向对象数据库系统

20 世纪 90 年代,许多从事数据库研究的学者把数据库技术和面向对象技术相结合,研究出一种新的数据库系统——面向对象数据库系统(Object Oriented DataBase System, OODBS),以满足新的应用需要。面向对象数据库系统的研究有两种观点,一种是在面向对象程序设计语言中引入数据库技术,另一种是从关系数据库系统自然地引入面向对象技术而进化到具有新功能的结果。现在一般把前一类数据库系统称为面向对象数据库系统(OODBS),后一类称为对象关系数据库系统(Object Relation DataBase System, ORDBS),这两类统称为对象数据库系统。本书主要介绍对象关系数据库系统。

目前,各个关系数据库厂商都在不同程度上扩展了关系模型,推出了符合面向对象数据模型的数据库系统。面向对象的数据模型吸收了面向对象程序设计方法的核心概念和基本思想,用面向对象的观点来描述现实世界的实体。

对象关系数据库系统可定义为在关系数据模型的基础上,提供元组、数组、集合等丰富的数据类型以及处理新的数据类型的能力,并且具有继承性和对象标识等面向对象特点,这样形成的数据模型称为对象关系数据模型。基于对象关系模型的数据库系统称为对象关系数据库系统。所以对象关系数据库系统除了具有原来关系数据库的各种特点外,还具有以下特点。

(1) 扩充数据类型。以关系数据库和 SQL 为基础,扩展关系数据模型,增加面向对象的数据类型和特性。新的数据类型可定义为原有类型的子类或超类。新的数据类型定义之后,存放在数据库管理系统中,如同基本数据类型一样,可供所有用户共享。

(2) 支持复杂对象。OODBS 中的基本结构是对象而不是记录,一个对象不仅包括描述它的数据,还包括对它操作的方法。它不仅支持简单的对象,还支持由多种基本数据类型或用户自定义的数据类型构成的复杂对象,支持子类、超类和继承的概念,因而能对现实世界的实体进行自然而直接的模拟,可表示诸如某个对象由“哪些对象组成”,有“什么性质”,处在“什么状态”,具有丰富的语义信息,这是传统数据库所不能比拟的。

(3) 提供通用的规则系统。规则在数据库管理系统(DataBase Management System, DBMS)及其应用中是十分重要的,在传统的关系数据库管理系统(Relation DataBase Management System, RDBMS)中用触发器来保证数据库的完整性。触发器可以看成规则的一种形式。OODBS 支持的规则系统将更加通用,更加灵活。例如,规则中的事件和动作可以是任何的 SQL 语句,可以使用用户自定义的函数,规则还能够被继承。这就大大增强了 OODBS 功能,使之具有主动数据库的特性。

面向对象数据库系统的功能要求:①在数据模型方面,引入面向对象的概念,包括对象、类、对象标识、封装、继承、多态性、类层次结构等;②在数据库管理方面,提供与扩展对持久对象、长事务的处理能力以及并发控制、完整性约束等能力;③在数据库界面方面,支持消息传递,提供计算能力完备的数据库语言,解决数据库语言与宿主语言的失配问题,并且数据库语言应具有类似 SQL 的非过程化的查询功能。

除此之外,还要求兼顾对传统的关系数据的管理能力。

面向对象数据库系统主要研究的问题有对象数据模型、高效的查询语言、并发的事务处理技术、对象的存储管理以及版本管理等。

3) 网络数据库系统

随着客户机/服务器结构的出现,使得人们可以最有效地利用计算机资源。在客户机/服务器结构中的服务器又称数据库服务器,主要用于放置数据库管理系统以及存储数据,而客户机则负责应用逻辑与用户界面。它们通过网络互连,当客户机需要访问数据时,向服务器提出某种数据或服务请求,服务器将响应这些请求并把结果或状态信息返回给客户机。通过网络将地理位置分散的、各自具备自主功能的若干台计算机和数据

库系统有机地连接起来的,并且采用通信手段实现资源共享的系统称为网络数据库系统。

但是在网络环境中,为了使一个应用程序能访问不同的数据库系统,需要在应用系统和不同的数据库管理系统之间加一层中间件。所谓中间件是网络环境中保证不同的操作系统、通信协议和数据库管理系统之间进行对话、互操作的软件系统。其中涉及数据访问的中间件,就是 20 世纪 90 年代提出的开放的数据库连接 (Open DataBase Connectivity, ODBC) 技术和 Java 数据库连接 (Java DataBase Connectivity, JDBC) 技术。使用 ODBC 和 JDBC 技术来进行数据库应用程序的设计,可以使应用系统移植性更好,并且能访问不同的数据库系统,共享数据资源。

1.1.2 数据库系统

数据库系统 (DataBase System, DBS) 是指引进数据库技术后的计算机系统,主要包括相应的数据库、数据库管理系统、数据库应用系统、计算机硬件系统、软件系统和用户。

1. 数据库

关于数据库 (DataBase) 的定义,一般认为数据库是长期存储在计算机内、有组织的、可共享的数据集合。数据库中的数据按一定的数据模型组织、描述和存储,具有较小的冗余度、较高的数据独立性和易扩展性,可为各种用户共享,并且还具有完善的自我保护能力和数据恢复能力。

数据库是用来存储数据的。数据库中的数据包括两大类,一类是用户数据,如学生数据库中每个学生的信息。另一类是系统数据,如系统中用户的权限、各种统计信息等。

2. 数据库管理系统

数据库管理系统 (DBMS) 位于用户与操作系统之间,是可借助操作系统完成对硬件的访问,并负责数据库存取、维护和管理的系统软件。它是数据库系统的核心组成部分,用户在数据库中的一切操作,包括定义、查询、更新以及各种控制都是通过 DBMS 进行的。

DBMS 的基本功能如下:

(1) 数据定义功能。在关系数据库管理系统中就是创建数据库、创建表、创建视图和创建索引,定义数据的安全性和数据的完整性约束等。

(2) 数据操纵功能。实现对数据库的基本操作,包括数据的查询处理,数据的更新(增加、删除、修改)等。

(3) 数据库的运行管理。主要完成对数据库的控制,包括数据的安全性控制、数据的完整性控制、多用户环境下的并发控制和数据库的恢复,以确保数据正确有效和数据库系统的正常运行。

(4) 数据库的建立和维护功能。包括数据库的初始数据的装入,数据库的转储、恢复、重组织,系统性能监视、分析等功能。

(5) 数据通信。DBMS 提供与其他软件系统进行通信的功能。它实现用户程序与 DBMS 之间的通信,通常与操作系统协调完成。

3. 数据库应用系统

数据库应用系统 (DBAS) 是指利用数据库系统资源开发的面向实际应用的软件系统。一个数据库应用系统通常由数据库和应用程序组成。它们都是在数据库管理系统支持下设计和开发出来的。

4. 用户

用户是指使用和管理数据库的人,他们可以对数据库进行存储、维护和检索等操作。数据库系统中用户可分为三类。

(1) 终端用户。终端用户主要是指使用数据库的各级管理人员、工程技术人员等,一般来说,他们是非计算机专业人员。

(2) 应用程序员。应用程序员负责为终端用户设计和编制应用程序,以便终端用户对数据库进行操作。

(3) 数据库管理员。数据库管理员 (DBA) 是指对数据库进行设计、维护和管理的专门人员。

数据库系统的组成结构如图 1.6 所示。

1.1.3 数据模型

模型是对现实世界特征的模拟和抽象。如一组建筑设计沙盘，一架精致的航模飞机等都是具体的模型。数据模型是模型的一种，它是现实世界数据特征的抽象。现实世界中的具体事务必须用数据模型这个工具来抽象和表示，计算机才能够处理。

1. 概述

数据模型通常由数据结构、数据操作和数据约束三部分组成。数据结构是所研究的对象类型的集合。数据操作是指对数据库中各种对象（型）的实例（值）允许执行的操作的集合，包括操作及有关的操作规则。数据约束是一组完整规则的集合。通过数据结构、数据操作和数据约束可以完整描述数据模型。

根据模型应用的层次不同，可以将这些模型划分为 3 类。

第 1 类模型是概念数据模型，又称概念模型或信息模型，它是按用户的观点来对数据和信息建模，是用户和数据库设计人员之间进行交流的工具，这一类模型中最著名的就是实体关系模型。实体关系模型直接从现实世界中抽象出实体类型以及实体之间的关系，然后用实体关系图（E-R 图）表示数据模型。E-R 图有下面四个基本成分。

- (1) 矩形框，表示实体类型（问题的对象）；
- (2) 菱形框，表示关系类型（实体之间的联系）；
- (3) 椭圆形框，表示实体类型的属性；
- (4) 连线。实体与属性之间，关系与属性之间用直线连接；关系类型与其涉及的实体类型之间也以直线相连，并在直线端部标注关系的类型（1:1、1:n 或 m:n）。

如图 1.7 所示是一个 E-R 图的示例。

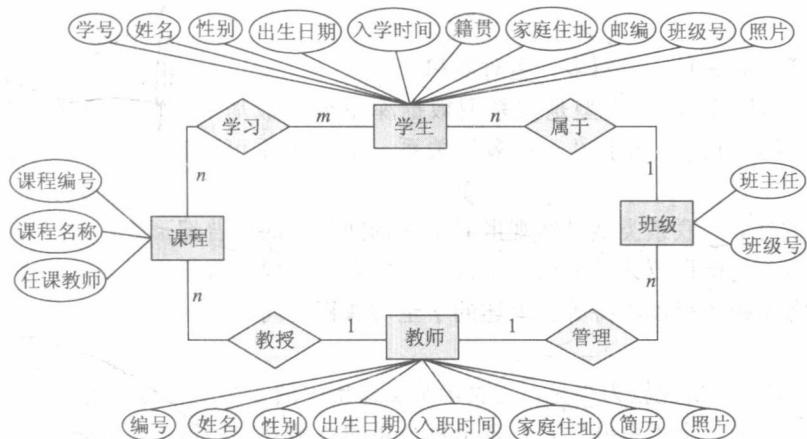


图 1.7 E-R 图

第 2 类模型是逻辑数据模型，又称数据模型，它是一种面向数据库系统的模型，该模型着重于在数据库管理系统一级的实现。概念模型只有转化为数据模型后才能在数据库中得以实现。其主要包括网状模型、层次模型、关系模型等，它是按计算机系统的观点对数据建模，主要用于 DBMS 的实现。

第 3 类模型是物理数据模型，又称物理模型，它是一种面向计算机物理表示的模型，此模型给出了数据模型在计算机上物理结构的表示。

数据模型是数据库系统的核心和基础。各种机器上实现的 DBMS 软件都是基于某种数据模型的。为了把现实世界中的具体事物抽象、组织为某一 DBMS 支持的数据模型，人们常常先将现实世界抽象为信息世界，然后再将信息世界转换为机器世界。也就是说把现实世界中的客观对象抽象为某一种信息结构，这种信息结



图 1.6 数据库系统组成结构图



构并不依赖于具体的计算机系统,不是某一个 DBMS 支持的数据模型,而是概念级的模型;然后再把概念模型转换为计算机上某一 DBMS 支持的数据模型,这一过程如图 1.8 所示。

2. 相关概念

建立数据模型需要掌握以下几个概念。

1) 实体

客观存在,并可相互区别的事物称为实体(Entity)。实体可以是实实在在的客观存在,例如学生、教师、商店、医院;也可以是一些抽象的概念或地理名词,如地震、北京市。

2) 属性

实体所具有的特征称为属性(Attribute)。实体本身并不能被装进数据库,要保存客观世界的信息,必须将描述事物外在特征的属性保存在数据库中。属性的差异能使我们区分同类实体,如一个人可以具备姓名、年龄、性别、身高、肤色、发型、衣着等属性,根据这些属性可以在熙熙攘攘的人群中一眼认出所熟悉的人。

3) 实体集和实体型

具有共性的实体组成的一个集合称为实体集(Entity Set)。一个实体所有属性的集合称为实体型。例如,要管理学生信息,可以存储每一位学生的学号、姓名、性别、出生年月、出生地、家庭住址、各科成绩等,其中学号是人为添加的一个属性,用于区分两个或多个因巧合而属性完全相同的学生。在数据库理论中,这些学生属性的集合就是一个实体集,这些学生所具有的所有的属性就是一个实体型,在数据库应用中,实体集以数据表的形式呈现,实体型以字段名称的形式呈现。

4) 联系

客观事物往往不是孤立存在的,相关事物之间保持着各种形式的联系方式。在数据库理论中,实体(集)之间同样也保持着联系,这些联系同时也制约着实体属性的取值方式与范围。这种实体集之间的对应关系称为联系。

实体的联系方式通常有 3 种,一对多、多对多和一对一。

(1) 一对多。“一对多”联系类型是关系型数据库系统中最基本的联系形式,例如“系”表与“教师”表这两个实体的联系方式就属于“一对多”关系,即一个系可以有 multiple 教师,但一名教师只能属于一个系。

(2) 多对多。“多对多”联系类型是客观世界中事物间联系的最普遍形式,例如在一个学期中,一名学生要学若干门课程,而一门课程要让若干名学生来修;一名顾客要逛若干家商店才能买到称心的商品,而一家商店必须有许多顾客光顾才得以维持等。上述的学生与课程之间、顾客与商店之间的关系均为“多对多”联系。

(3) 一对一。“一对一”情况较为少见,它表示某实体集中的一个实体对应另一个实体集中的一个实体。例如为补充系的信息,添加一个“系办”表,表示每个系的系部办公室地点。从常识得知,一个系只有一个系部办公室,反之一个系部办公室只为一个系所有,这两个实体的联系方式就属于“一对一”关系。

3. 常用模型

数据库领域常用的数据模型经常是按照数据的组织形式划分为,包括层次模型、网状模型、关系模型和面向对象模型 4 种。

1) 层次模型

在层次模型中,实体间的关系形同一棵根在上的倒挂树,上一层实体与下一层实体间的联系形式为一对多。现实世界中的组织机构设置、行政区划关系等都是层次结构应用的实例。基于层次模型的数据库系统存在天生的缺陷,它访问过程复杂,软件设计的工作量较大,现已较少使用。

层次模型具有以下特点。

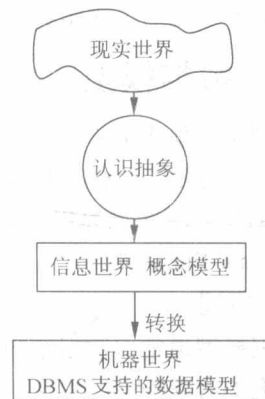


图 1.8 现实世界中客观对象的抽象过程