

教育部高等学校非计算机专业计算机基础课程教学指导分委员会推荐
计算机基础课程系列教材

Database Technology
and Application

数据库技术及应用 —Access

李雁翎 编著



高等教育出版社
Higher Education Press

教育部高等学校非计算机专业计算机基础课程教学指导分委



推荐

计算机基础课程系列教材

数据库技术及应用

——Access

李雁翎 编著

高等教育出版社

内容提要

本书由数据库理论和基于 Access 应用技术两大部分组成,全书共 16 章,主要内容包括:数据库系统概述、数据库设计、关系代数、关系数据库标准语言 SQL、Access 系统概述、数据库操作、表操作、查询操作、窗体设计、宏、VBA 编程、报表的创建与使用、菜单与工具栏、数据的传递与共享、数据库安全措施、小型应用系统开发。附录部分给出了常用的属性、宏命令和方法,以便于读者查阅。

本书体系完整,结构清晰,实例丰富,讲解详细,易学易懂,适用性强。本书既可作为高等学校计算机和相关专业数据库技术课程的教材,也可作为全国计算机等级考试二级 Access 的培训或自学教材。为了方便教师教学和学生自主学习,本书还配有辅助教材。电子课件可从高等教育出版社网站下载,网址为 <http://4a.hep.com.cn> 或 <http://4a.hep.edu.cn>。

图书在版编目(CIP)数据

数据库技术及应用——Access / 李雁翎编著. —北京:
高等教育出版社,2005.7

ISBN 7-04-017283-6

I. 数... II. 李... III. 关系数据库-数据库管理
系统, Access-高等学校-教材 IV. TP311.138

中国版本图书馆 CIP 数据核字(2005)第 054587 号

策划编辑 陈红英 责任编辑 萧 潇 封面设计 于文燕 责任绘图 朱 静
版式设计 张 岚 责任校对 尤 静 责任印制 陈伟光

出版发行	高等教育出版社	购书热线	010-58581118
社 址	北京市西城区德外大街 4 号	免费咨询	800-810-0598
邮政编码	100011	网 址	http://www.hep.edu.cn
总 机	010-58581000		http://www.hep.com.cn
		网上订购	http://www.landaco.com
经 销	北京蓝色畅想图书发行有限公司		http://www.landaco.com.cn
印 刷	涿州市星河印刷有限公司		
开 本	787×1092 1/16	版 次	2005 年 7 月第 1 版
印 张	22.75	印 次	2005 年 7 月第 1 次印刷
字 数	510 000	定 价	28.40 元

本书如有缺页、倒页、脱页等质量问题,请到所购图书销售部门联系调换。

版权所有 侵权必究

物料号 17283-00

前 言

数据库技术是计算机应用的重要分支,是高等学校非计算机专业的必修课。本书以改革计算机教学、适应 21 世纪人才培养需要为出发点,根据教育部高等学校非计算机专业计算机基础课程教学指导分委员会提出的有关“数据库技术及应用”课程教学要求而编写,编者力图以教学指导分委员会制定的《数据库技术及应用课程大纲》为主线,从培养学生分析问题和解决问题的能力入手,以数据库原理和技术为核心,建构新编教材的体例。

本书由数据库理论和基于 Access 应用技术两大部分组成,其特色主要体现在两个方面,一方面是理论部分与 Access 应用技术部分相辅相成,既要求坚实的理论基础,又强调技术实践的应用;另一方面是用一个实用的应用系统程序贯穿整个教学过程,围绕其编排大量详实的实例,强化讲解。讲解的内容包括:与数据库的基本原理相关的基础知识;数据模型和 E-R 模型的设计;SQL 语言以及 SQL 查询的应用;Access 2003 数据库管理系统技术与应用;ActiveX 控件应用;VBA 程序设计;数据库应用系统开发的方法及步骤;数据库安全技术。全书体系清晰,具有整体性、全面性和实用性。

全书共 16 章,各章内容安排如下:

第 1 章主要介绍与数据库管理系统相关的一些数据库基础理论方面的知识,包括数据、数据模型、数据库、数据库管理系统和数据库系统等概念。

第 2 章主要介绍数据库设计步骤及各设计阶段的任务,包括 E-R 模型、E-R 图、关系模型、关系的规范化及 E-R 模型与关系模型的转换等。

第 3 章主要介绍集合运算(并、差、交、笛卡尔积)和关系运算(投影、选择、连接、除法)等。

第 4 章主要介绍关系数据库标准 SQL 语言,包括 SQL 用于数据定义、数据更新及数据查询操作等。

第 5 章主要介绍 Access 2003 的功能特性、系统安装环境与安装方法、系统环境、系统参数的设置等。

第 6 章主要介绍数据库的建立、数据库的对象、数据库不同版本的转换等。

第 7 章主要介绍表的建立以及对表的操作等。

第 8 章主要介绍查询的建立,包括用查询从已有的数据资源中提取数据、多种查询的应用以及标准 SQL 语言在 Access 2003 中的应用等。

第 9 章主要介绍面向对象的基本概念,包括窗体的创建方法、窗体的属性、事件和方法的定义以及窗体中控件的功能及属性、ActiveX 控件、ADO 数据对象控件的应用等。

第 10 章主要介绍宏的建立以及宏的应用等。

第 11 章主要介绍在 Access 2003 环境下的编程,包括 VBA 程序设计基础、VBA 程序基本结构和 VBA 程序实例等。

第 12 章主要介绍利用报表设计器和报表向导创建各类格式不同报表、报表的修改方法、报表用于统计分析等。

第 13 章主要介绍菜单/工具栏的创建方法、菜单/工具栏选项功能的定义以及菜单的调用方法等。

第 14 章主要介绍数据的传递与共享、数据的导入/导出等。

第 15 章主要介绍数据库安全措施,包括设置数据库密码、创建工作组、设置用户与组的权限等。

第 16 章主要以“阳光超市管理系统”为案例介绍开发设计数据库应用系统的一般过程及主要内容,综合全书各章内容,用一个应用程序实例总结升华全部内容。

本书既可作为高等学校计算机及相关专业数据库技术课程的教材,也可作为全国计算机等级考试二级 Access 的培训或自学教材,还可以作为广大计算机用户和广大计算机技术初学者的自学用书。

为了方便教师教学和学生自主学习,本教程配有辅助教材及相关的教学网站。

本书的编写得到了吉林大学崔恒珠教授的热情指教,王丛林老师以及高等教育出版社编辑同志给予了大力支持,在此一并感谢。

北京对外经贸大学信息学院陈恭和教授审阅了全书,并提出了许多宝贵的意见,在此表示衷心的感谢。

由于作者水平有限,难免有错误和不足之处,欢迎广大读者批评指正。

编 者
2005 年 4 月

目 录

第1章 数据库系统概述	1	2.4.3 E-R模型与关系模型的转换	37
1.1 信息、数据与数据处理	2	2.4.4 关系数据库	38
1.1.1 信息与数据	2	2.5 物理结构设计	40
1.1.2 数据处理	3	2.6 数据库实施	41
1.2 数据描述	6	2.7 数据库使用与维护	42
1.3 数据模型	7	思考题	42
1.3.1 层次模型	8	设计题	43
1.3.2 网状模型	8	第3章 关系代数	44
1.3.3 关系模型	9	3.1 集合运算	44
1.3.4 面向对象模型	10	3.1.1 并运算	44
1.4 数据库系统	10	3.1.2 差运算	46
1.4.1 数据库	11	3.1.3 交运算	46
1.4.2 数据库管理系统	11	3.1.4 笛卡尔积运算	48
1.4.3 数据库系统的体系结构	12	3.1.5 应用实例	49
1.4.4 数据库系统的组成	13	3.2 特殊的关系运算	52
思考题	14	3.2.1 投影运算	53
第2章 数据库设计	15	3.2.2 选择运算	55
2.1 数据库设计概述	15	3.2.3 连接运算	56
2.1.1 数据库设计的任务	16	3.2.4 除法运算	59
2.1.2 数据库设计的步骤	16	思考题	60
2.2 需求分析	17	计算题	60
2.3 概念结构设计	19	第4章 关系数据库标准语言 SQL	63
2.3.1 实体-联系模型	19	4.1 SQL语言概述	63
2.3.2 E-R图	22	4.1.1 SQL语言的特点	64
2.3.3 实体集联系类型	23	4.1.2 SQL语言的功能	65
2.3.4 E-R模型的设计实例	24	4.2 数据定义	66
2.4 逻辑结构设计	25	4.2.1 SQL语言基本数据类型	66
2.4.1 关系模型	26	4.2.2 创建表	67
2.4.2 关系的规范化	32	4.2.3 修改表结构	68

4.2.4 删除表	69	6.1.7 模块	104
4.2.5 建立索引	70	6.2 创建数据库	105
4.2.6 删除索引	70	6.2.1 直接创建空数据库	105
4.3 数据更新	71	6.2.2 利用向导创建数据库	106
4.3.1 插入数据	71	6.3 数据库打开与关闭	110
4.3.2 更新数据	72	6.3.1 打开数据库	110
4.3.3 删除数据	72	6.3.2 关闭数据库	111
4.4 数据查询	73	6.4 数据库压缩/修复	112
4.4.1 Select 语句	73	6.5 数据库转换	113
4.4.2 简单查询	74	思考题	114
4.4.3 连接查询	82	实验题	115
4.4.4 嵌套查询	85	第7章 表操作	116
4.4.5 集合查询	89	7.1 表	116
思考题	90	7.1.1 表的结构	117
设计题	90	7.1.2 数据类型	118
第5章 Access 系统概述	92	7.1.3 表结构的定义	119
5.1 Access 功能及特性	92	7.2 创建表	120
5.2 Access 的安装	93	7.2.1 使用数据表视图创建表	120
5.2.1 安装环境	94	7.2.2 使用表向导创建表	121
5.2.2 安装方法	94	7.2.3 使用设计视图创建表	123
5.3 Access 的集成环境	94	7.3 表中数据的输入	124
5.3.1 标题栏	94	7.4 表的属性设置与维护	125
5.3.2 菜单栏	94	7.4.1 字段的维护	125
5.3.3 工具栏	95	7.4.2 字段的输入/显示格式设置	127
5.3.4 工作区与状态行	99	7.4.3 字段标题的设置	133
5.4 退出 Access	99	7.4.4 字段有效规则的设置	134
思考题	99	7.5 表中数据的编辑	136
实验题	99	7.5.1 数据的修改	136
第6章 数据库操作	100	7.5.2 数据的复制	137
6.1 Access 数据库对象	100	7.5.3 数据的删除	137
6.1.1 表	100	7.5.4 数据的查找/替换	137
6.1.2 查询	101	7.6 使用表	138
6.1.3 窗体	102	7.6.1 记录定位	139
6.1.4 报表	102	7.6.2 记录排序	139
6.1.5 宏	103	7.6.3 记录筛选	141
6.1.6 数据访问页	103	7.6.4 字段隐藏/取消隐藏	141
		7.6.5 字段冻结/解冻	143

7.7 建立表间关联关系	144	9.3.2 对象属性	187
7.7.1 表间关联关系类型	144	9.3.3 对象事件和方法	190
7.7.2 建立索引	146	9.4 使用视视图创建窗体	191
7.7.3 设置主关键字	148	9.4.1 常用的窗体控件	192
7.7.4 创建表间关联关系	149	9.4.2 常用窗体控件的功能	192
7.8 使用子表	152	9.4.3 常用的 ActiveX 控件	195
思考题	153	9.4.4 ADO 数据对象	196
实验题	153	9.4.5 窗体常用控件的操作	198
第8章 查询操作	155	9.5 窗体设计实例	200
8.1 查询概述	155	9.5.1 数据输入窗体	200
8.1.1 查询的作用	155	9.5.2 数据浏览窗体	202
8.1.2 查询的类型	156	9.5.3 “关于”窗体	207
8.2 创建选择查询	156	思考题	209
8.2.1 使用设计视图创建查询	157	实验题	209
8.2.2 使用向导创建查询	160	第10章 宏	210
8.3 创建参数查询	160	10.1 什么是宏	210
8.4 创建动作查询	163	10.2 宏的创建与编辑	211
8.4.1 创建生成表查询	163	10.3 使用宏与宏组	213
8.4.2 创建更新查询	165	10.3.1 直接运行宏或宏组	213
8.4.3 创建追加查询	166	10.3.2 触发事件运行宏或宏组	214
8.4.4 创建新字段查询	168	10.3.3 用宏命令间接运行宏或宏组	219
8.4.5 创建删除查询	169	思考题	220
8.5 修改查询	171	实验题	221
8.6 SQL 查询	172	第11章 VBA 编程	222
8.6.1 Select 语句	172	11.1 标准模块	222
8.6.2 创建 SQL 查询	172	11.2 VBA 程序设计基础	223
8.6.3 SQL 应用	173	11.2.1 数据类型	223
思考题	181	11.2.2 常量	225
实验题	181	11.2.3 变量	225
第9章 窗体设计	182	11.2.4 函数	227
9.1 窗体的组成	182	11.2.5 表达式	232
9.2 创建窗体	183	11.2.6 编码规则	234
9.2.1 使用自动窗体创建窗体	183	11.3 程序基本结构	235
9.2.2 使用窗体向导创建窗体	183	11.3.1 顺序结构	235
9.3 引入面向对象编程的概念	186	11.3.2 分支结构	235
9.3.1 对象	186	11.3.3 循环结构	237

11.3.4 过程	240	实验题	292
11.3.5 自定义函数	242	第14章 数据的传递与共享	293
11.4 VBA 程序实例	243	14.1 数据的导出	293
11.4.1 主页窗体	243	14.1.1 向其他数据库导出数据库对象	294
11.4.2 查询窗体	245	14.1.2 将数据库对象导出到 Excel 中	294
11.4.3 登录窗体	248	14.1.3 将数据库对象导出到 Word 中	295
思考题	252	14.1.4 将数据导出到数据文件中	296
实验题	252	14.2 数据的导入	298
第12章 报表的创建与使用	254	14.2.1 导入 Access 数据库对象	298
12.1 报表的组成	254	14.2.2 向 Access 数据库导入 Excel 数据	299
12.2 创建报表	256	14.2.3 向 Access 数据库导入数据文件	301
12.2.1 使用自动报表创建报表	256	思考题	303
12.2.2 使用报表向导创建报表	257	实验题	304
12.2.3 将窗体转换为报表	261	第15章 数据库安全措施	305
12.3 报表设计	261	15.1 数据库用户密码	305
12.3.1 报表控件的使用	261	15.1.1 设置用户密码	306
12.3.2 报表的页面设置	262	15.1.2 撤销用户密码	306
12.3.3 设计报表布局	263	15.2 工作组	306
12.3.4 使用报表设计视图创建报表	263	15.2.1 创建新的工作组	307
12.3.5 设计汇总报表	265	15.2.2 加入已有的工作组	307
12.3.6 设计分组报表	266	15.3 用户与组	308
12.4 创建其他报表	269	15.4 用户与组的权限	310
12.4.1 创建图表报表	269	15.5 安全向导	311
12.4.2 创建标签报表	272	15.6 编码数据库	315
思考题	275	思考题	316
实验题	275	实验题	316
第13章 菜单与工具栏	277	第16章 小型应用系统开发	317
13.1 菜单系统	277	16.1 应用系统开发概述	317
13.2 创建菜单	278	16.1.1 系统分析阶段	317
13.2.1 使用“命令栏”对象创建菜单	278	16.1.2 系统设计阶段	318
13.2.2 使用宏创建菜单	284	16.1.3 系统实施阶段	319
13.3 工具栏	286	16.1.4 系统维护阶段	319
13.3.1 使用“系统工具”创建工具栏	286	16.2 主要功能模块设计	319
13.3.2 使用 ToolBar 和 ImageList 创建 工具栏	286	16.2.1 主页设计	319
思考题	291		

16.2.2 登录窗口设计	320	思考题	329
16.2.3 菜单设计	320	实验题	330
16.2.4 控制面板设计	321	附录一 字段常用属性	331
16.2.5 数据库设计	322	附录二 对象常用属性	333
16.2.6 工作窗体设计	325	附录三 常用的宏命令	338
16.2.7 报表设计	327	附录四 常用的 DoCmd 方法	341
16.3 设置自动启动窗体	328	附录五 ADO 对象属性与方法	346
16.4 发布 MDE 文件	329		

第 1 章

数据库系统概述

在信息社会,信息系统越来越突显出重要性,数据库技术作为信息系统的核心技术和基础也更加引人注目。处于信息系统管理核心的数据库系统已融入人们的日常工作和生活中,扮演着相当重要的角色。

具体地说,如果将一组相互关联的数据集中存储,就可以称其是数据库,将管理、控制数据的软件称为数据库管理系统,将以数据库管理系统为基础运行的相互作用的应用程序、用户及硬件环境称为数据库系统。比如,一个消费者去超级市场购买商品,就好像身处在一个“数据库系统”之中,正在访问一个商品的“数据库”;收银员使用一个条形码阅读器扫描消费者购买的每一件商品,就启动了一个应用程序根据条形码阅读器获取的有关商品的“数据”,从商品数据库中找出商品价格,从商品库存数据库中减少商品的库存数量,并且要计算消费者的消费额度、增加销售总额、提示系统预定和补充商品等,这些操作就是“数据库系统”的工作。通过以上描述,可见人们对数据库系统并不陌生,常见的数据库系统有信用卡消费系统、图书馆系统、学籍管理系统、学生选课系统、高等院校网络系统、飞机订票系统等。

本章将对有关数据库系统的基本术语给予解释,逐一讲解信息、数据、数据处理、数据库、数据库管理系统功能及数据库系统的构成等基础知识和概念。

1.1 信息、数据与数据处理

进入数据库应用领域,首先遇到的是信息、数据和数据库等基本概念。这些不同的概念和术语,将贯穿在人们进行数据处理的整个过程之中。掌握好这些概念和术语,对更好地学习和使用数据库管理系统,有着重要的意义。这些概念是学习数据库技术应用、学习 Access 数据库管理系统软件必备的基础知识。

1.1.1 信息与数据

1. 信息 (Information)

在人类社会中存在各种各样的事物,每个事物都有其自身的表现特征和存在方式以及与其他事物的相互关联、相互影响和相互作用。人们对各种各样的事物进行管理的行为,就与这些信息的内容有关。

在数据处理领域,信息可定义为人们对于客观事物属性和运动状态的反映。它所反映的是关于某一客观系统中,某一事物的存在方式或某一时刻的运动状态。也可以说,信息是经过加工的,对人类客观行为产生影响的、通过各种方式传播的、可被感知的数据表现形式。

信息是人们进行社会活动、经济活动及生产活动的产物,并可作用于这些活动过程。信息是有价值、可以被感知的。信息可以通过载体传递,可以通过信息处理工具进行存储、加工、传播、再生和增值。

在信息社会中,信息一般可与物质或能量相提并论,它是一种重要的资源。

2. 数据 (Data)

数据是反映客观事物存在方式和运动状态的记录,是信息的载体。对客观事物属性和运动状态的记录是用一定的符号来表达的,因此说数据是信息的具体表现形式。数据所反映的事物是它的内容,而符号是它的形式。

数据表现信息的形式是多种多样的,不仅有数字、文字符号,还有图形、图像和声音等。用数据记录同一信息可以有不同的形式,信息不会随着数据形式的不同而改变其内容和价值。

数据与信息在概念上是有区别的。从信息处理角度看,任何事物的存在方式和运动状态都可以通过数据来表示,数据经过加工处理后,具有知识性并对人类活动产生作用,从而形成信息。

用数据符号表示信息,其形式有许多种类型。常见的有三种类型:数值型数据,即对客观事物进行定量记录的符号,如数量、年龄、价格和度数等;字符型数据,即对客观事物进行定性记录的符号,如姓名、单位、地址等;特殊型数据,即对客观事物进行形象特征和过程记录的符号,如声音、视频、图像等。

从计算机的角度看,数据泛指那些可以被计算机接受并能够被计算机处理的符号,是数据库中存储的基本对象。

总之,信息是有用的数据,数据是信息的表现形式。信息是通过数据符号来传播的,数据如

不具有知识性和有用性则不能称其为信息,也就没有价值。

1.1.2 数据处理

数据处理也称为信息处理。所谓数据处理,实际上就是利用计算机对各种类型的数据进行加工处理。它包括对数据的采集、整理、存储、分类、排序、检索、维护、加工、统计和传输等一系列操作过程。数据处理的目的是从人们收集的大量原始数据中,获得人们所需要的资料并提取有用的数据成分,作为行为和决策的依据。

随着计算机软件、硬件技术的发展,数据处理量的规模日益扩大,数据处理的应用需求越来越广泛,数据管理技术的发展也不断变迁,经历了人工管理、文件系统、数据库系统3个阶段。

1. 人工管理阶段

20世纪50年代中期以前,计算机主要用于数值计算。在这一阶段,硬件方面,外存储器只有卡片机、纸带机、磁带机,没有像硬盘一样可供快速、随机存储的外存储器;软件方面,没有操作系统软件和数据管理软件支持,数据处理方式基本是批处理。在这一管理方式下,应用程序与数据之间不可分割,当数据有所变动时程序则随之改变,数据的独立性差;另外,各程序之间的数据不能相互传递,缺少数据的共享性。

在人工管理阶段,应用程序与数据之间的关系如图1-1所示。

在人工管理阶段数据处理的特点是:

(1) 数据不保存。在人工管理阶段进行数据处理的过程,是将数据与其对应的应用程序一同输入内存,通过应用程序对数据进行加工处理后输出处理结果,计算任务完成,随着应用程序的释放,数据也将从内存中释放。

(2) 应用程序与数据之间缺少独立性。应用程序与数据之间相互依存,不可分割,设计应用程序时不仅要设计数据处理的算法、数据的逻辑结构,还要指明数据在存储器上的存储地址,当数据有所变动时应用程序则随之改变,因而这种管理方式既不灵活也不安全,编程效率低。

(3) 数据不能共享。由于数据与应用程序不具有独立性,一个应用程序只能对应一组数据,各程序之间的数据不能相互传递,若多个应用程序需要使用同一组数据时,仍然需要逐个进行数据定义,不能进行相互调用。数据不能共享,造成应用程序之间的大量数据冗余。

2. 文件系统阶段

20世纪50年代后期至60年代中后期,硬件方面,随着磁鼓、磁盘联机的外存储器的研制并投入使用,软件方面,高级语言和操作系统软件出现,计算机的应用不仅仅用于科学计算,同时也开始以“文件”的方式介入数据处理。

在这一阶段,数据被组织成数据文件,这种数据文件可以脱离应用程序而独立存在,数据文件可长期保存在硬盘中多次存取。由于使用专门的文件管理系统实施数据管理,应用程序与数

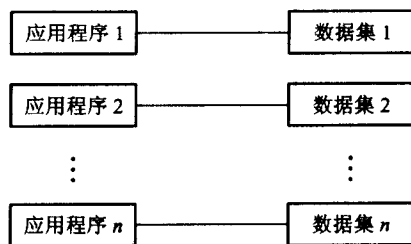


图 1-1 人工管理阶段应用程序与数据的关系

据文件之间具有了一定的独立性,同时数据的逻辑结构与物理结构之间也具有一定的相对独立性。在文件系统阶段,应用程序与数据之间的关系如图 1-2 所示。

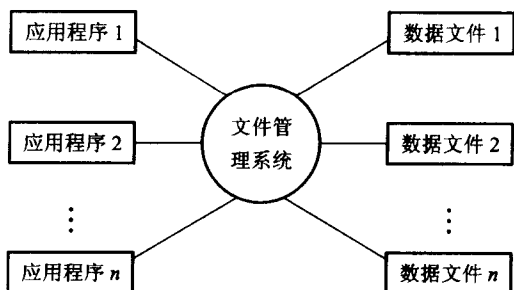


图 1-2 文件系统阶段应用程序与数据的关系

在文件系统阶段数据处理的特点是：

(1) 数据长期保存。在利用计算机进行数据处理过程中,大量数据的存储、数据的检索和数据的维护成为可能,是因为有了外存储器的出现,数据可以以文件形式长期保存在硬盘上,供用户反复调用和进行更新操作。

(2) 应用程序与数据之间有了一定的独立性。在文件系统阶段,操作系统提供了文件管理功能和访问文件的存取方法,应用程序与数据文件之间有了数据存取接口,应用程序可以通过文件名对数据进行访问,不必再寻找数据的物理位置,至此,数据有了物理结构与逻辑结构的区别,因此比人工管理阶段前进了一步。但此时,应用程序是基于特定的物理结构和特定的存取方法进行程序访问的,数据文件与应用程序仍彼此依赖,它们之间的独立性只是相对的“设备独立性”。

(3) 数据文件形式多样化。由于有了直接存取的存储设备,文件的形式不局限于顺序文件,还有了随机文件等,因而,对数据文件的访问可以是顺序访问,也可以是随机访问。但数据文件之间相互独立,数据之间的联系要通过应用程序控制。

(4) 数据文件不再只属于一个应用程序。在文件系统阶段,一个数据文件可被多个应用程序使用,一个应用程序也可使用多个数据文件。由于应用程序对数据文件的访问基于物理结构和特定的存取方法,因此,应用程序对数据结构的依赖不能从根本上改变。

(5) 仍有一定的数据冗余。由于数据文件的设计很难满足多个用户的不同需求,大多数情况下,仍是一个应用程序对应一个数据文件,同样的数据会出现在不同的应用程序中。

(6) 数据的不一致性。由于有一定的数据冗余,在进行数据更新时,就会出现同样数据在多个应用程序中的不一致。

3. 数据库系统阶段

进入 20 世纪 60 年代后期,随着计算机应用领域的日益发展,计算机在数据处理方面的应用越来越广泛,处理的数据量越来越大,仅仅基于文件系统的的海理技术很难满足应用领域的需

求。与此同时,计算机硬件技术也飞速发展,磁盘存储技术取得重要突破,大容量磁盘进入市场,数据处理软件环境的改善成为许多软件公司的重要目标。在实际需求迫切、硬件与软件技术发展趋于成熟的条件下,数据库系统应运而生。

数据库系统克服了文件系统阶段的缺陷,对相关数据实行统一规划管理,形成一个数据中心,构成一个数据“仓库”,实现了整体数据的结构化。

在数据库系统阶段,应用程序与数据之间的关系如图 1-3 所示。

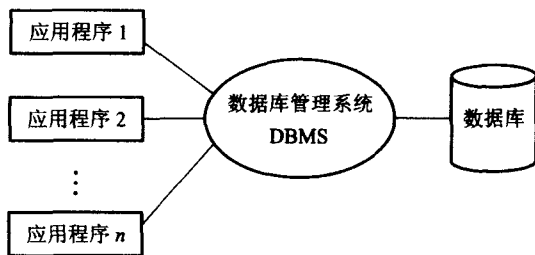


图 1-3 数据库系统阶段应用程序与数据的关系

在数据库系统阶段数据处理的特点是:

(1) 数据整体结构化。在文件系统阶段进行数据结构设计时,考虑的是文件记录内部数据项之间的联系,而不同的文件记录之间没有联系,可以说,文件系统阶段的数据整体是无结构的。在数据库系统中,数据不再针对某一个应用,而是面向全组织,形成整体的结构化。不仅数据是结构化的,而且数据的存取单位即一次可以存取数据的大小也很灵活,可以小到一个数据项,如一个商品的名称;大到一组记录,如成千上万个商品记录。在文件系统中,只能通过对整条记录的读取操作,才能获得其中数据项的信息,而不能直接对记录中的数据项进行数据存取操作。

(2) 数据共享性高。这是数据库系统阶段对文件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。由于数据库包含了所有与应用系统相关的数据成分,每个用户可根据不同的需求访问某一部分数据,而且不同的用户也可以访问同一部分数据。简言之,在数据库系统阶段,由于数据的整体结构化,一个用户可面对的数据资源是多样化的,一部分数据资源可被多种需求的用户访问。

以本书应用案例中的阳光超级市场数据库为例,用户与数据库的关系如图 1-4 所示。

从上面的例子不难看出,由于从整体化的观点定义数据结构,数据是面向系统,而不是面向应用,数据的灵活性大大提高,用户可以针对不同的需求取整体数据集的子集,数据冗余明显减少,实现了数据共享。

(3) 具有很高的数据独立性。在这一阶段,应用程序不再只与一个孤立的数据文件相对应,可以取整体数据集的某个子集作为逻辑文件与其对应,通过数据库管理系统实现逻辑文件与物理数据之间的映射,数据与应用程序之间不存在依赖关系,而是相互独立的。

(4) 完备的数据控制功能。数据库是一个多级的系统结构,为了确保数据库为多个用户共

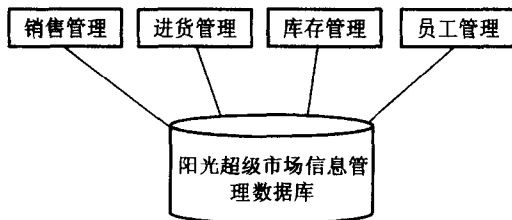


图 1-4 阳光超级市场数据库

享,确保多个用户同时存取一个数据子集,确保数据库中数据的一致性、完整性,需要通过一组专门的软件,对数据资源进行统一管理和统一控制。这一组专门的软件具有数据的安全性控制、数据的完整性控制、并发控制和故障恢复功能。

随着软件环境和硬件环境的不断改善,数据处理应用领域需求的持续扩大,数据库技术与其他软件技术的加速融合,到 20 世纪 80 年代,新的更高级的数据库技术相继出现并得到长足的发展,分布式数据库系统、面向对象数据库系统和并行数据库系统等新型数据库系统应运而生。

在此之前,数据库系统是集中式的,即把数据集中在一个数据库中进行集中管理,减少了数据的冗余和不一致性,实现了数据的共享。但是,集中式数据库也有其缺点,当数据库中数据量过大时,应用系统相当庞大,操作相对复杂,开销也同样增加,数据通信过于频繁。随着计算机硬件成本的降低,网络软件和远程通信的发展,分布式数据库系统便显示出优越性,其特点是多数数据处理任务就地完成;各地计算机由数据通信网络联系构成逻辑整体性。在现实世界中,存在许多复杂的数据结构应用领域,用已有的数据模型(层次模型、网状模型、关系模型)很难描述,由此推出了面向对象数据库系统,其特点是使用面向对象的方法构建数据模型。

新型数据库系统带来了一个又一个数据库技术发展的高潮,但对于中、小数据库用户来说,由于很多高级的数据库系统的专业性要求太强,通用性受到一定的限制,在很大程度上推广使用范围也受到约束。随着基于关系模型的关系数据库系统功能的扩展与改善,对象关系数据库、数据仓库、Web 数据库、嵌入式数据库等数据库技术构成了新一代数据库系统的主流。

1.2 数据描述

在实际应用中,任何数据的处理都是在反映客观事物属性和运动状态以及满足用户需求的基础上进行的。所谓数据描述,就是以数据符号的形式,从满足用户需求出发,对客观事物属性和运动状态进行描述。数据的描述既要符合客观现实,又要适应数据库原理与结构,同时还要适应计算机原理与结构。进一步说,由于计算机不能够直接处理现实世界中的具体事物,所以人们必须将客观存在的具体事物进行有效的描述与刻画,转换成计算机能够处理的数据,这一转换过程可分为三个数据范畴:现实世界、信息世界和计算机世界。

从客观现实到计算机的描述,数据的转换过程如图 1-5 所示。

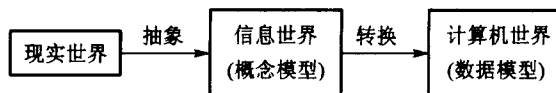


图 1-5 数据的转换过程

1. 现实世界

现实世界是指客观存在的事物及其相互间的联系。在现实世界中,人们可以通过事物不同的属性和运动状态对事物加以区别,描述事物的性质和运动规律。事物可以是个体的特殊事物,也可以是集体的共同事物;事物可以是具体的、可见的实物,也可以是抽象的概念。

例如,在一个商品超市里,琳琅满目的商品存放在货架上和仓库中,是可见的客观存在;这些商品的采购渠道、销售数量、销售利润也是可见的客观存在。若想开发一个“商品超市数据库管理系统”,这些商品信息就成为数据库系统所用到的原始数据,它们存在于人们头脑之外的客观世界,即现实世界。

2. 信息世界

信息世界是人们对客观存在的事物及其相互间的联系的反映。人们将对客观事物的反映通过符号记录下来,事实上是对现实世界的一种抽象描述。

在信息世界中,不是简单地对现实世界进行一种符号记录,而是要通过选择、分类、命名等抽象过程产生出概念模型,用以表示对现实世界的抽象与模拟。

3. 计算机世界

计算机世界是信息世界的数据化。客观存在的事物及其相互间的联系的反映,在这里用数据模型来表示。也就是说,计算机世界的数据模型将信息世界的概念模型进一步抽象,形成便于计算机处理的数据表现形式。

1.3 数据模型

一般而言,模型是对客观存在的事物及其相互间的联系的抽象与模拟。模型可以分为实物模型和抽象模型,如汽车模型、轮船模型是实物模型,是对客观存在事物的外部特征或内在功能的模拟与刻画;而数据模型如“ $s = vt$ ”则是一种抽象模型,它从较高的层面上抽象地描述了距离、速度、时间三者之间的数量关系,是对客观存在事物的某些规律性的、本质的特性进行模拟与刻画。

数据模型是一种抽象模型。现实世界中的客观事物是彼此相互联系的。一方面,某一事物内部的诸多因素和诸多属性根据一定的组织原则相互具有联系,构成一个相对独立的系统;另一方面,某一事物同时也作为一个更大系统的一个因素或一种属性而存在,并与系统的其他因素或属性发生联系。客观事物的这种普遍联系性,决定了作为事物属性记录符号的数据与数据之间也存在着一定的联系性,数据模型是对数据间联系和约束条件的全局性描述。