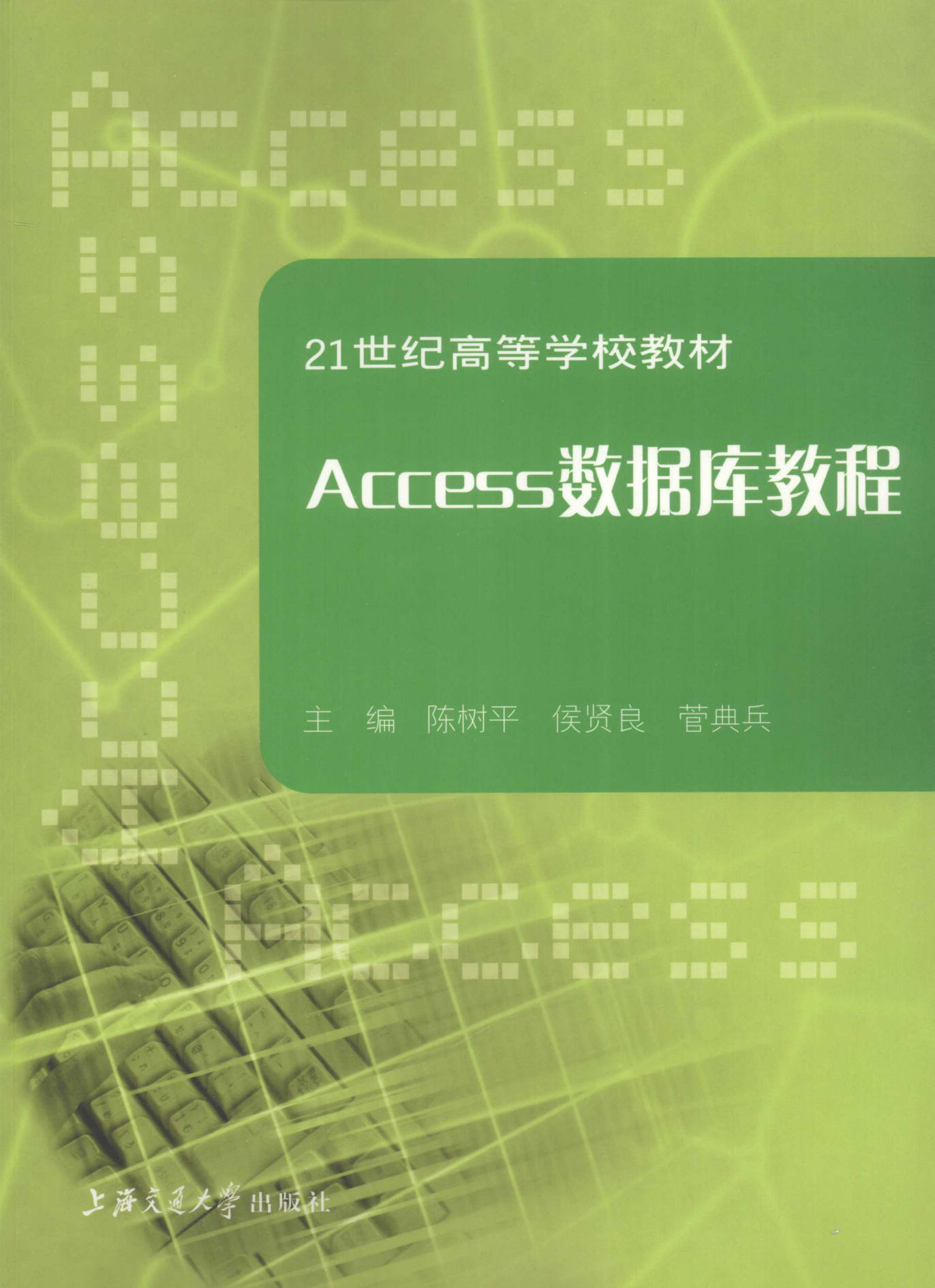




21世纪高等学校教材

Access数据库教程

主 编 陈树平 侯贤良 菅典兵

上海交通大学出版社

21 世纪高等学校教材

主要内容

Access 数据库教程

主编 陈树平 侯贤良 菅典兵

图书在版编目(CIP)数据

Access 数据库教程 / 陈树平, 侯贤良, 菅典兵主编. — 上海: 上海交通大学出版社, 2009

21 世纪高等学校教材

ISBN 978-7-313-02918-2

I. A. 1. ①陈... ②侯... ③菅... II. ①陈... ②侯... ③菅... III. 数据库—教材 IV. TP311.138

中国版本图书馆 CIP 数据核字 (2009) 第 123923 号

Access 数据库教程

主编 陈树平 侯贤良 菅典兵

上海交通大学出版社出版

上海市番禺路 981 号 邮政编码 200030

电话: 64071208 电邮: jk@sjtu.edu.cn

常熟文化印刷有限公司印刷 全国新华书店经销

开本: 787mm × 1085mm 1/16 印张: 19.75 字数: 483 千字

2009 年 8 月第 1 版 2009 年 8 月第 1 次印刷

ISBN 978-7-313-02918-2 定价: 31.00 元

上海交通大学出版社

内 容 提 要

本书根据全国计算机等级考试(二级 Access)大纲,本着“厚基础、重能力、求创新”的基本思路,由多年从事计算机教学的教师编写。

全书共分 10 章。主要内容有数据库基础知识、Access 2003 数据库系统概述、表的建立和使用、数据查询、窗体设计、报表、数据访问页、宏、模块与 VBA、学生成绩管理系统的设计。本书内容丰富、图文并茂、通俗易懂,侧重于对基础知识、基本理论和基本方法的叙述,注重基本操作的训练。全书以实例教学驱动展开,以加强学生对操作技能的掌握。

本书可作为高等院校非计算机专业数据库应用技术课程教学用书,也可供各培训机构作为 Access 数据库应用教材和参加全国计算机等级考试(二级 Access)参考书。

图书在版编目(CIP)数据

Access 数据库教程/陈树平,侯贤良,菅典兵主编.
—上海:上海交通大学出版社,2009
21 世纪高等学校教材
ISBN 978-7-313-05918-5

I. A ... II. ①陈 ... ②侯 ... ③菅 ... 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(2009)第 127957 号

Access 数据库教程

陈树平,侯贤良 菅典兵 主编

上海交通大学出版社出版发行

(上海市番禺路 951 号 邮政编码 200030)

电话:64071208 出版人:韩建民

常熟文化印刷有限公司 印刷 全国新华书店经销

开本:787mm×1092mm 1/16 印张:19.75 字数:483 千字

2009 年 8 月第 1 版 2009 年 8 月第 1 次印刷

ISBN978-7-313-05918-5/TP 定价:31.00 元

版权所有 侵权必究

前 言

随着我国改革开放的不断深入,高等教育得到了较快发展并走向大众化教育。时代的进步与社会的发展对高等学校计算机教育提出了更高、更新的要求。抓好计算机专业课程以及计算机公共基础课程的教学,是提高计算机教育质量的关键。在掌握计算机基础操作的同时掌握现代数据管理技术是一项重要技能。Access 是一个功能强大并且易于使用的关系数据库管理系统,利用 Access 数据库技术可以提高数据管理能力,帮助人们高效、快速地组织、查询和获取信息。

本书根据全国计算机等级考试(二级 Access)大纲,同时又兼顾到基本理论的完整性,由多年从事计算机教学的教师编写。

本书具有下列特点:

- (1) 实用性和先进性:介绍目前较流行的 Access 2003 的使用和操作。
- (2) 操作性:注重基本操作能力的培养,以学生管理系统为主线,以实例操作为主,只要读者按照书中内容操作,均可掌握相应软件的使用。

- (3) 通俗性:力求文字流畅、图文并茂、版面活跃、通俗易懂。

- (4) 理论性:注重基本理论的完善性。

本书共分 10 章。主要内容包括数据库基础知识、Access 2003 数据库系统概述、表的建立和使用、数据查询、窗体设计、报表、数据访问页、宏、模块与 VBA、学生成绩管理系统的设计。每章都配有一定的思考题与习题,供读者练习使用。

本书由陈树平教授策划、统稿,陈树平、侯贤良、菅典兵任主编,梁咏梅、张志宏、陈红梅任副主编。本书第 1 章由陈树平编写;第 2 章由匡友华编写;第 3 章、第 4 章由侯贤良编写;第 5 章由陈红梅编写;第 6 章、第 7 章由梁咏梅编写;第 8 章、第 10 章由张志宏编写,第 9 章由菅典兵编写。附录部分由陈树平整理。

本书配备电子课件,有意者与作者联系。

作者陈树平电子信箱:shuping1576@126.com 或 shuping1576@yahoo.com.cn

由于作者水平有限,书中有不妥之处,恳请读者批评指正,以便改进。

编 者

2009 年 6 月

目 录

第1章 数据库基础知识	1
1.1 数据与数据处理	1
1.1.1 数据与信息	1
1.1.2 数据处理技术的发展	2
1.1.3 数据库管理系统	3
1.2 数据模型	5
1.2.1 数据模型的组成要素	5
1.2.2 概念模型	6
1.2.3 层次模型	7
1.2.4 网状模型	8
1.2.5 关系模型	9
1.3 关系数据库	11
1.3.1 关系的基本概念	11
1.3.2 关系的特点	13
1.3.3 关系运算	13
1.4 数据库系统	16
1.4.1 数据库系统的结构	16
1.4.2 数据库系统的组成	18
1.5 关系的规范化	20
1.5.1 相关概念	20
1.5.2 关系规范化	21
1.6 数据库系统设计步骤	23
1.6.1 需求分析	23
1.6.2 概念结构设计	24
1.6.3 逻辑结构设计	24
1.6.4 物理结构设计	25
1.6.5 系统实施	25
1.6.6 运行和维护	25
本章小结	25
思考题与习题	26

第2章 Access 2003 数据库系统概述	28
-------------------------------	----

2.1 Access 的发展和特点	28
-------------------------	----

2.1.1 Access 的发展	28
2.1.2 Access 2003 的特点	28
2.2 Access 2003 的启动和退出	29
2.2.1 启动 Access 2003	29
2.2.2 退出 Access	30
2.3 Access 2003 的用户界面	30
2.3.1 标题栏和菜单栏	30
2.3.2 工具栏	31
2.3.3 任务窗格	32
2.3.4 工作区和状态栏	32
2.4 Access 2003 的数据对象	32
2.4.1 数据库	32
2.4.2 表	32
2.4.3 查询	33
2.4.4 窗体	33
2.4.5 报表	33
2.4.6 数据访问页	34
2.4.7 宏	34
2.4.8 VBA 模块	34
2.5 数据库的建立	35
2.5.1 创建空数据库	35
2.5.2 利用向导创建数据库	36
2.6 使用数据库	37
2.6.1 打开数据库	37
2.6.2 关闭数据库	38
2.6.3 数据库的复制与删除	38
2.6.4 转换 Access 2003 数据库到低版本	38
本章小结	39
思考题与习题	40

第3章 表的建立和使用	42
-------------------	----

3.1 Access 2003 的数据类型	42
3.2 创建表	43
3.2.1 表的属性	43

3.2.2 使用设计器创建表	44	4.4.3 自定义计算	88
3.2.3 通过输入数据创建表	46	4.4.4 查找重复项查询向导	89
3.2.4 使用向导创建表	47	4.4.5 查找不匹配项查询向导	91
3.2.5 导入表	48	4.5 创建参数查询	92
3.2.6 链接表	50	4.5.1 单条件查询	92
3.3 编辑数据表	52	4.5.2 复合条件查询	93
3.3.1 复制、删除和重命名表对象	52	4.6 创建操作查询	94
3.3.2 修改字段名、字段类型及字段属性	53	4.6.1 创建生成表查询	94
3.3.3 插入、删除和移动字段	54	4.6.2 创建删除查询	95
3.3.4 编辑表中记录	55	4.6.3 创建追加查询	96
3.3.5 调整表的外观	56	4.6.4 创建更新查询	97
3.4 建立数据表之间的关系	57	4.7 创建 SQL 查询	98
3.4.1 表间关系的有关概念	58	4.7.1 SQL 简介	98
3.4.2 建立表间关系	58	4.7.2 SQL 的优点	98
3.5 数据表的使用	60	4.7.3 SQL 查询	99
3.5.1 查找和替换数据	60	4.7.4 单表查询	100
3.5.2 排序记录	61	4.7.5 联合查询	104
3.5.3 筛选记录	62	4.7.6 子查询	106
3.5.4 数据的导出	64	4.7.7 数据定义查询	109
本章小结	65	4.7.8 传递查询	110
思考题与习题	65	4.8 编辑查询	110
第 4 章 数据查询	69	4.8.1 编辑查询中的字段	110
4.1 查询基本知识	69	4.8.2 调整查询中的列宽	111
4.1.1 查询的概念	69	4.8.3 编辑查询中的数据源	111
4.1.2 查询的功能	70	4.8.4 查询结果的使用	112
4.1.3 查询的类型	71	本章小结	113
4.1.4 查询的基本操作与创建方式	72	思考题与习题	114
4.1.5 查询表达式	74	第 5 章 窗体设计	117
4.2 创建选择查询	76	5.1 窗体概述	117
4.2.1 利用查询向导创建查询	76	5.1.1 窗体的概念和功能	117
4.2.2 利用查询设计视图创建查询	77	5.1.2 窗体的组成和结构	117
4.2.3 设计带查询条件的查询	78	5.1.3 窗体的分类	118
4.3 创建交叉表查询	80	5.1.4 窗体的视图	121
4.3.1 使用向导创建交叉表查询	80	5.1.5 窗体的属性	121
4.3.2 创建交叉表查询	82	5.2 创建窗体	122
4.4 在查询中进行计算	83	5.2.1 使用“自动创建窗体”创建窗体	122
4.4.1 Access 查询计算类型	83	5.2.2 使用“图表向导”创建窗体	123
4.4.2 预定义计算	84	5.2.3 使用“窗体向导”创建窗体	125

5.2.4 窗体设计视图	127	6.2.3 使用“标签向导”创建报表	162
5.3 窗体中控件及其应用	128	6.2.4 使用“图表向导”创建报表	164
5.3.1 工具箱的使用	128	6.3 报表设计视图的使用	166
5.3.2 标签控件及应用	130	6.3.1 报表设计视图简介	166
5.3.3 文本框控件及应用	131	6.3.2 使用设计视图创建报表	168
5.3.4 选项组控件及应用	133	6.3.3 丰富报表的内容	170
5.3.5 组合框控件及应用	136	6.4 报表中数据的排序与分组	172
5.3.6 命令按钮控件及应用	138	6.4.1 排序数据	172
5.3.7 选项卡控件及应用	140	6.4.2 数据分组	173
5.3.8 图像控件及应用	143	6.5 创建子报表	175
5.3.9 ActiveX 控件及应用	144	6.5.1 在已有报表中创建子报表	175
5.4 窗体和控件的事件	144	6.5.2 将某个已有报表添加为子报表	177
5.4.1 键盘事件	144	6.6 打印报表	178
5.4.2 鼠标事件	145	6.6.1 报表页面设置	178
5.4.3 窗口事件	145	6.6.2 分页打印报表	179
5.4.4 对象事件	145	6.6.3 分列打印报表	179
5.4.5 操作事件	146	6.6.4 指定打印机及执行打印报表	179
5.5 美化窗体	146	本章小结	180
5.5.1 调整控件布局	146	思考题与习题	181
5.5.2 使用自动套用格式	146	第7章 数据访问页	183
5.5.3 添加日期和时间	147	7.1 数据访问页视图	183
5.6 窗体的操作	147	7.1.1 设计视图	183
5.6.1 在窗体中浏览和定位记录	148	7.1.2 页面视图	184
5.6.2 在窗体中编辑记录	148	7.1.3 网页预览	184
5.6.3 在窗体中排序记录	149	7.2 创建数据访问页	185
5.6.4 在窗体中筛选记录	149	7.2.1 自动创建数据访问页	185
5.6.5 窗体的打印和打印预览	151	7.2.2 使用向导创建数据访问页	186
本章小结	151	7.2.3 在设计视图中创建数据访问页	187
思考题与习题	152	7.3 编辑数据访问页	189
第6章 报表	155	7.3.1 设置数据访问页格式	189
6.1 报表概述	155	7.3.2 修改数据访问页属性	191
6.1.1 报表的类型	155	7.3.3 在数据访问页内添加对象	193
6.1.2 报表的组成	156	7.4 在 IE 中查看数据访问页	198
6.1.3 报表与窗体的区别	157	本章小结	198
6.1.4 报表的视图	157	思考题与习题	199
6.2 快速创建报表	157	第8章 宏	201
6.2.1 自动创建报表	158	8.1 宏的概述	201
6.2.2 使用“报表向导”创建报表	159		

8.1.1 宏与宏组	201	9.4.1 DoCmd 对象及常用操作	237
8.1.2 宏的基本功能与类型	201	9.4.2 事件及常用操作	239
8.2 创建宏	203	9.5 VBA 的程序结构	239
8.2.1 宏设计视图	203	9.5.1 算法描述	239
8.2.2 在宏设计视图中创建宏	204	9.5.2 顺序结构语句	240
8.2.3 创建事件	207	9.5.3 选择结构语句	242
8.2.4 创建条件宏	209	9.5.4 循环结构语句	246
8.2.5 创建宏组	211	9.6 VBA 编程环境	248
8.3 宏的运行与调	213	9.6.1 进入 VBE	248
8.3.1 宏的运行方式	213	9.6.2 VBE 窗口的组成	249
8.3.2 宏的调试	216	9.7 VBA 的数据库编程	252
8.3.3 创建启动窗体	217	9.7.1 数据库引擎与接口	252
8.4 宏的其他操作	219	9.7.2 数据访问对象(DAO)	252
8.4.1 将宏转换为 VBA 程序代码	219	9.7.3 ActiveX 数据库对象	254
8.4.2 更改宏的安全性级别	220	9.7.4 程序设计实例	255
本章小结	221	9.8 VBA 的运行与调试	256
思考题与习题	221	9.8.1 VBA 模块的错误类别	256
第 9 章 模块与 VBA	224	9.8.2 VBA 的错误处理	257
9.1 模块的概念	224	9.8.3 调试工具的使用	258
9.1.1 模块分类	224	9.9 Access 数据库的管理	259
9.1.2 创建模块	224	9.9.1 制作 MDE 文件	259
9.1.3 子过程与函数	225	9.9.2 设定和撤消密码	261
9.1.4 参数传递	227	9.9.3 数据库的备份与还原	262
9.1.5 宏转换为模块	228	9.9.4 数据库的压缩与修复	263
9.2 VBA 程序设计基础	229	9.9.5 切换面板	264
9.2.1 程序设计概述	229	本章小结	267
9.2.2 数据类型	230	思考题与习题	267
9.2.3 VBA 中的变量	230	第 10 章 学生成绩管理系统的设计	271
9.2.4 VBA 中的常量和数组	232	10.1 需求分析	271
9.2.5 VBA 中的运算符	233	10.1.1 系统任务的提出	271
9.2.6 运算符的优先级	234	10.1.2 系统需求	271
9.3 常用函数	235	10.2 概念结构设计	272
9.3.1 数学函数	235	10.3 逻辑结构设计	272
9.3.2 字符串函数	235	10.4 物理设计和运行维护	273
9.3.3 日期和时间函数	236	10.4.1 物理设计	273
9.3.4 格式输出函数	236	10.4.2 数据库实施运行和维护	274
9.3.5 类型转换函数	236	10.5 系统设计	274
9.4 DoCmd 对象及事件的常用操作	237	10.5.1 主窗体的设计	274

10.5.2 设计表间的关系	275	10.7 设计文档	290
10.5.3 功能模块的详细设计	276	本章小结	291
10.5.4 编辑各表的窗体	276	思考题与习题	291
10.5.5 设置查询和查询窗体	278	附录 I	292
10.5.6 报表设计	282	附录 II	297
10.5.7 设计主窗体	284	参考文献	304
10.6 用户安全机制的设置与管理	286		
10.6.1 创建用户组和用户	286		
10.6.2 设置用户与组的访问权限	287		
10.6.3 用户级安全机制的设置	287		

第1章 数据库基础知识

随着电子技术的飞速发展,计算机技术取得了较快的发展与广泛的普及,已经应用到了社会生活的各个领域,尤其在数据处理、信息管理等领域更是获得了广泛的应用,取得了较好的社会 and 经济效益。为了对数据有效的管理和利用,数据库技术应运而生,它是人类进入数字时代的重要标志。据统计,目前计算机应用 80% 是数据处理,而数据处理最有效的方式是数据库技术。数据库技术的应用水平已经成为衡量一个国家发达程度的重要标志,数据库技术是数据处理的核心技术,要掌握数据处理技术必须了解和掌握数据库的相关知识。

本章主要内容包括数据与数据处理、数据模型、关系数据库、数据库系统、关系规范化、数据库系统设计步骤等。

1.1 数据与数据处理

数据库技术是信息时代重要的基础技术之一,是计算机科学技术领域发展最为迅速的重要分支。数据库技术是一门综合技术,它涉及程序设计、操作系统、数据结构、算法设计等相关知识,它是计算机及其相关专业的一门重要课程。

1.1.1 数据与信息

1. 数据与信息

数据是数据库中存储的基本对象。现实世界中人们通常是用各种各样的物理符号来表示客观世界事物的特征和联系,这些符号及其组合就是数据。数据是对现实世界符号化的描述,从数据的描述来看,数据有两方面的含义:其一是描述事物特征的数据内容;其二是存储在某种媒体上的具体形式。数据的形式是多种多样的,可以是数字、字符、文字、图形、符号、声音、影像等。

信息是数据经过加工处理后所获得的有用知识,数据只有经过加工才能转换为信息,这种加工过程就是数据处理。数据处理包括对数据的采集、整理、存储、分类、排序、检索、加工、统计、传输等一系列的操作过程。数据和信息既有联系又有区别,一方面数据是信息的具体表现形式,信息是数据有意义的表现;另一方面信息和数据又很难区分,信息本身就是数据化的。如图 1-1 所示为数据和信息之间的关系。

现代社会是信息社会,信息的价值已经引起人们的重视。信息与物质、能源一起被称为三大资源,人们利用信息提升了思维,信息加工后的结果形成了人们的知识。一个部门的信息化水平反映了该部门应用信息处理问题的能力,一个国家的信息化程度在很大程度上反映了一个国家的发达程度。

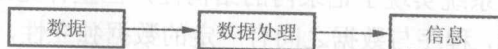


图 1-1 数据和信息之间的关系

2. 数据库

数据库(DataBase, DB)顾名思义是存储数据的仓库,但这个仓库是存储在计算机外部存储设备上并按一定的格式组织的。数据库是长期存储在计算机内、有组织、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,并可被各种用户共享的数据集合。目前,一个国家的数据库建设规模(指数据库的个数、种类)、数据库信息量的大小和使用频度已经成为衡量一个国家信息化程度的重要标志之一。

统计表明,目前全世界 80%以上的计算机主要从事事务处理工作。人们用计算机进行事务处理的主要目的是为了从大量数据中提取有用信息。因此,在进行数据处理时必须在计算机中存储大量的数据,并采用一定的方法对其进行必要的加工、处理和传播。20 世纪 70 年代后,数据库技术得到了飞速发展,越来越受到部门的重视和人们的关注,从事数据处理技术人员的队伍不断扩大。目前,从事数据处理技术的人员已经成为计算机及相关技术队伍的主力军。

1.1.2 数据处理技术的发展

数据库技术是随着数据管理任务的需求而产生的。数据库技术是伴随着计算机软件、硬件技术的发展而发展起来的,数据处理技术经历了人工管理、文件系统、数据库系统 3 个发展阶段。

1. 人工管理阶段

20 世纪 50 年代中期以前,计算机外存主要是卡片、纸带和磁带,这些设备只能顺序存取,没有像磁盘这样的直接存取设备。在软件上,没有专门的数据管理软件,对数据的处理没有一定的格式,数据处理都是用手工进行的。

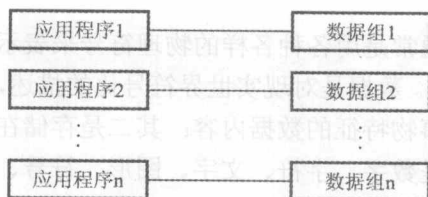


图 1-2 应用程序与数据之间的对应关系

当时的数据处理过程是一组程序对应一组数据,一个程序中的数据不能被其他应用程序使用,使得程序之间存在大量的重复数据,称为数据冗余。如图 1-2 所示为应用程序与数据之间的对应关系,由于数据与程序的这种对应,使得数据的独立性很差,当数据改变时,应用程序必须改变,并且数据不能长期保存。

2. 文件系统阶段

20 世纪 50 年代后期至 60 年代中期,计算机开始用于大量数据管理。硬件上出现可直接存取的外存储器,磁盘、磁鼓等,为计算机管理数据提供了物质基础。在软件上出现了操作系统和高级语言,操作系统中文件系统是专门管理外存的管理软件。

在文件系统中,由于有了专门的文件管理软件进行数据管理,数据可以长期存储在外存储器上,进行反复地查询、修改、插入和删除。文件可以“按文件名访问,按记录存取”,文件系统实现了记录内的结构性,但整体无结构。程序和数据可由文件系统进行存取方法的转换,程序与数据之间有一定的数据独立性,可分开存储,并有了程序文件和数据文件的区别。

尽管文件系统带来了方便,但文件系统也存在一些不足。由于文件系统中的文件是为某一特定应用服务的,文件的逻辑结构对该应用程序是优化的,要想对现有的数据再增加一些新的应用很困难。文件系统不容易扩充,一旦数据的结构改变,必须修改应用程序,修改文件结构的定义。应用程序的改变将引起文件的数据结构改变。因此,数据和程序之间缺乏独立性。如图 1-3 所示为文件系统中程序与数据之间的关系。



图 1-3 文件系统中程序与数据之间的关系

3. 数据库系统阶段

20 世纪 60 年代后期以来,由于大容量磁盘的诞生、计算机硬件价格下降,计算机软件相对成熟,计算机的应用范围也不断扩大。计算机处理的数据量增加,管理的规模也越来越大,文件系统已经无法满足开发应用系统的需要。为了实现对数据的统一管理,达到数据共享的目的,发展了数据库技术。

数据库技术的主要目的是有效地管理和存取大量的数据资源,包括提高数据的共享性,使多个用户能同时访问数据库中的数据;减小数据的冗余度,以提高数据的完整性、一致性;提高数据与应用程序的独立性,从而减少应用程序的开发和维护代价。

用数据库系统管理数据比文件系统有明显的优势,从文件系统到数据库系统,标志着数据管理技术的飞跃。如图 1-4 所示为在数据库管理系统的支持下,数据与程序之间的关系。



图 1-4 数据库系统中数据与程序之间的关系

1.1.3 数据库管理系统

数据库管理系统(DataBase Management System, DBMS)是帮助用户建立、使用和管理数据库的软件系统,它是位于用户与操作系统之间的一层数据库管理软件。如图 1-5 所示为 DBMS 的功能和组成。

从图中可以看出, DBMS 具有下列主要功能:数据库的定义,数据库的操纵,数据库的建立、组织和存储,数据库的控制。

1. 数据库的定义

DBMS 提供了数据定义语言(Data Definition Language, DDL),用户可以通过 DDL 方便的对数据库中的数据对象进行定义。

2. 数据库的操纵

DBMS 提供了数据操纵语言(Data Manipulation Language, DML),用户可以通过 DML 方便的对数据库中的数据对象进行检索、修改、删除、更新等基本操作。

3. 数据库的建立、组织和存储

DBMS 提供了数据库数据的初始化功能、转换数据和确定文件的存储路径和存取方法,提供了数据的存储功能。有效、合理地组织数据,提高存储空间的利用率。

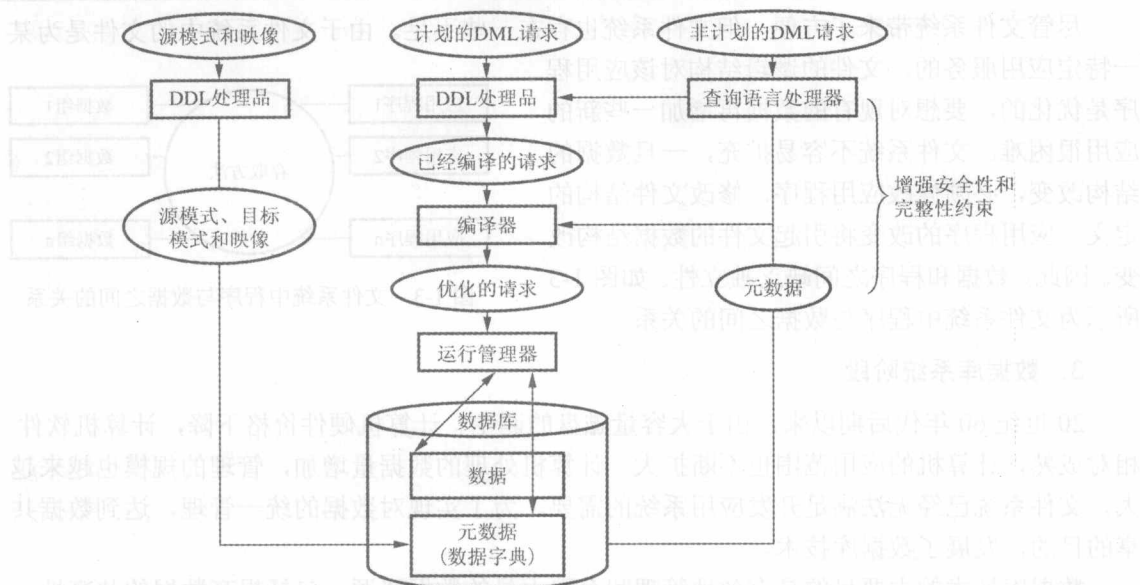


图 1-5 DBMS 的功能和组成

4. 数据库的控制

1) 数据库的安全性

DBMS 监控用户对数据库的请求、控制访问权限，防止用户对数据库的非法访问，保证数据的安全。

2) 数据库的完整性

DBMS 提供了数据的完整性检查机制，可以对数据的语法、语义进行检查，保证用户数据的正确性。

3) 并发控制

当多个用户同时访问数据库时，可能会出现冲突，如一个用户要对数据进行读操作，同时另一个用户要对同一个数据执行写操作，此时可能会出现执行错误，DBMS 提供并发控制机制，能够保证数据并发的正确性。

4) 数据恢复

数据库操作的过程中，可能会出现数据遭到破坏，此时，DBMS 提供恢复功能，将数据恢复到故障前的状态，有效地恢复数据。

5. 其他管理功能

DBMS 对数据进行统一控制和有效管理，主要包括：

1) 数据字典(Data Dictionary, DD)

保存数据库中数据的描述信息，以及访问、控制、管理数据的信息库。

2) 查询处理和优化

用户要对数据库进行高效地访问，必须经过优化。实践表明，优化的数据查询可能会成百上千倍地提高查询效率。DBMS 的基本功能之一是查询优化，优化效率直接影响到 DBMS 功能的实现。

3) 数据通信

DBMS 支持操作系统的联机分布处理、远程作业录入以及网络软件的通信功能。

1.2 数据模型

数据库是某个企业、组织或部门所涉及数据的集合，它不仅要反映数据自身，而且要反映数据之间的联系。数据模型是对现实世界特征的模拟和抽象。由于计算机不可能直接处理现实世界中的具体事物，所以人们需要事先把具体事物转换成计算机能够处理的数据，在数据库中采用数据模型来抽象、表示和处理现实世界中的数据和信息。

数据模型是数据库系统的核心和基础，DBMS 软件都是基于某种数据模型或是支持某种数据模型。为了将现实世界中的具体事物抽象，组织成 DBMS 支持的数据模型，人们要将现实世界抽象成为信息世界，用概念模型或者信息模型来描述，这种模型直观、清晰、容易被人们理解；然后再将信息世界进一步转换为机器世界，主要用于 DBMS 支持的层次数据模型、网状数据模型和关系数据模型来描述。如图 1-6 所示为从现实世界到机器世界的抽象过程。

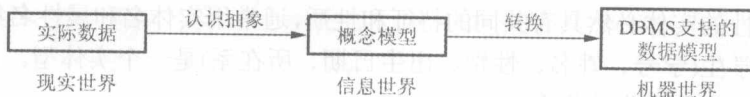


图 1-6 现实世界到机器世界的抽象过程

数据模型通常要满足 3 个条件：

- (1) 能够真实地反映现实世界。
- (2) 容易被人们理解。
- (3) 容易在计算机上实现。

1.2.1 数据模型的组成要素

数据模型通常由 3 部分组成：数据结构、数据操纵和数据的完整性约束。

1. 数据结构

数据结构描述数据库的组成对象及对象之间的联系，是对系统静态特征的描述。例如学生信息系统中，学生的基本信息包括学号、姓名、性别、出生日期、所学专业等。学号、姓名、性别、所学专业用字符型，出生日期用日期型。课程信息包括课程号、课程名、学时、学分、任课教师等。课程号、课程名、任课教师用字符型，学时、学分用数值型。学生选课信息包括学号、课程号、成绩。在上述信息中学号与学生关系中的学号相对应，课号与课程关系中的课号相对应，这样两个对象之间存在着关联才有意义，这种数据关联要存储到数据库中。数据模型就是按照数据结构分类的，按照数据结构类型的不同将数据模型分为层次数据模型、网状数据模型和关系数据模型。

2. 数据操纵

数据操纵是指对数据库中各种实例对象操作的集合，是描述系统的动态特征。例如对数据库中数据进行插入、修改、删除、查询等各种操作。

3. 数据的完整性约束

数据的完整性约束是在特定的数据模型中,数据及其联系所遵守的一组规则。这种规则限制了数据库状态及其状态变化,数据之间的约束保证了数据库的正确性、相容性和一致性。例如学生信息系统中,学号的前两位为入校年份,前两位值只能取 05、06、07、08,表示本科生的入校年份只能是 2005 年、2006 年、2007 年和 2008 年(对于四年制本科),否则就违背了完整性约束。

1.2.2 概念模型

1. 实体及其描述

实体(Entity)是客观存在并可以相互区别的事物,可以是具体的人、事、物,也可以是抽象的事件,如学生、课程、订货、借书等。

属性(Attribute)是对实体特征的描述。如学生实体用学号、姓名、性别、出生日期、所在系等属性描述;课程实体用课程号、课程名、学时等属性描述。

具有相同属性的实体必然具有共同的特征和性质,通常用实体名和属性名集合起来描述,称为实体型。如学生(学号、姓名、性别、出生日期、所在系)是一个实体型。同型实体的集合称为实体集,如全体学生的集合。

2. 实体间的联系

客观世界中事物间有一定的联系,这些联系反映在信息世界中就是实体内部的联系和实体之间的联系。实体内部的联系是指实体内属性之间的联系,而实体间的联系则是指不同实

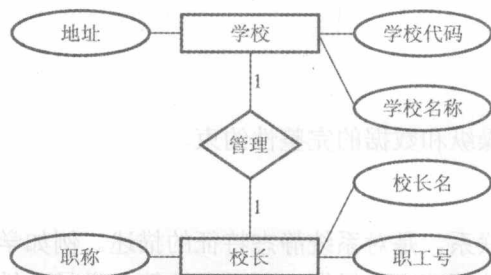


图 1-7 校长与学校一对一联系

体集之间的联系。实体间的联系有 3 种：一对一联系、一对多(或多对一)联系和多对多联系。

一对一联系：对于实体集 A 中的每一个实体,实体集 B 中至多有一个(也可以没有)实体与之联系;反之,对于实体集 B 中的每一个实体,实体集 A 中至多有一个(也可以没有)实体与之联系,则称实体集 A 与实体集 B 具有一对一联系,记作 1:1。如图 1-7 所示,在学校中,一个学校有一个校长,而一个校长也只能在一个学校任职,校长与学校就是一对一联系。

一对多联系：对于实体集 A 中的每一个实体,实体集 B 中有多个实体与之联系,反之,对于实体集 B 中的每一个实体,实体集 A 中至多只有一个实体与之联系,则称实体集 A 与实体集 B 具有一对多联系,记作 1:n。如图 1-8 所示,在学校中,一个系有多个教师,而一个教师只能在一个系工作,系与教师之间就是一对多联系。

多对多联系：对于实体集 A 中的每一个实体,实体集 B 中有多个实体与之联系,反之,对于实体集 B 中的每一个实体,实体集 A 中也有多个实体与之联系,则称实体集 A 与实体集 B 具有多对多联系,记作 m:n。如图 1-9 所示,在学校中,一个学生可以选修多门课程,一门课程也可以有多个学生选修,学生与课程就是多对多联系。

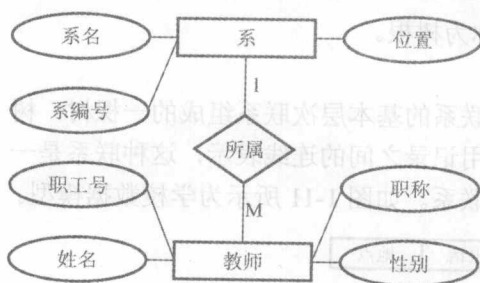


图 1-8 系与教师一对多联系

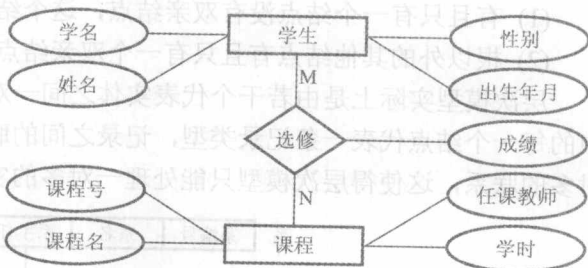


图 1-9 学生和课程多对多联系

实际上，多对多联系是一种普遍现象，一对一联系是一对多联系的特例，而一对多联系又是多对多联系的特例。

3. 实体联系图(Entity Relation)

概念模型的表示方法很多，其中最著名、最常用的是 1976 年 P. P. Chen 提出的实体联系方法，该方法用 E-R 图描述概念模型。

E-R 图方法如下：

- (1) 实体：用矩形表示，矩形框内标记实体名。
- (2) 属性：用椭圆表示，椭圆框内标记属性名，并用直线与相应的实体连接起来。
- (3) 联系：用菱形框表示，菱形框内标记联系名，并用直线与相应的实体连接起来，同时在直线旁标记联系类型，联系的属性也用椭圆表示。

例如，上述校长与学校一对一联系、系与教师一对多联系以及学生和课程多对多联系分别用图 1-7、图 1-8 和图 1-9 表示。

实际问题中，实体之间除两个实体间的联系外，也存在多个实体间的联系。例如考虑供应商、项目、零件之间的关系，一个供应商向多个项目供应多种零件，一个项目可以使用多个供应商供应的多种零件，一种零件由多个供应商供应并用于不同的工程。如图 1-10 所示为供应商、项目、零件之间的 E-R 图。

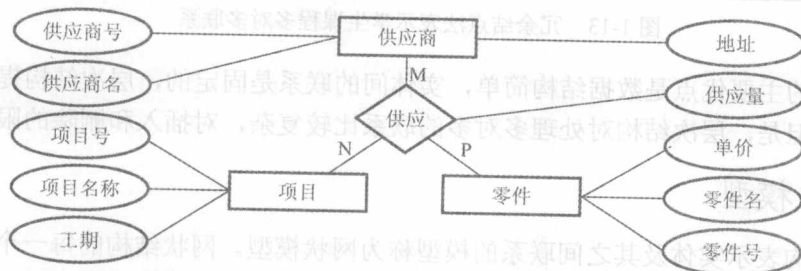


图 1-10 供应商、项目、零件之间的 E-R 图

1.2.3 层次模型

用树型结构表示实体与实体之间联系的模型称为层次模型。层次模型也是一种常见的数据模型，如单位组织机构。层次模型简单、直观。

层次模型有两点限制：

- (1) 有且只有一个结点没有双亲结点, 这个结点称为树根。
- (2) 根以外的其他结点有且只有一个双亲结点。

层次模型实际上是由若干个代表实体之间一对多联系的基本层次联系组成的一棵树, 树中的每一个结点代表一条记录类型, 记录之间的联系用记录之间的连线表示, 这种联系是一对多的联系, 这使得层次模型只能处理一对多的实体联系。如图 1-11 所示为学校数据库模型。

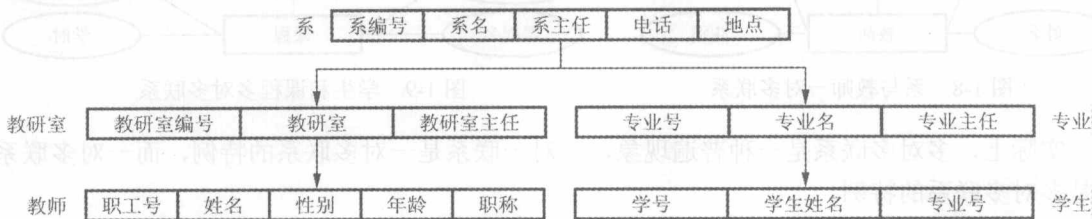


图 1-11 学校数据库模型

那么, 对于多对多的联系层次数据模型是否能处理呢? 答案是肯定的, 但处理起来比较复杂。例如, 如图 1-9 所示的学生和课程之间多对多的联系(这里只用部分属性来描述), 层次模型可以采用两种方式进行处理, 一是采用设置虚拟结点的方法, 如图 1-12 所示为虚拟结点法表示的学生课程之间多对多的联系; 另一种是用冗余结点的方法, 如图 1-13 所示为冗余结点法表示的学生课程之间多对多的联系。

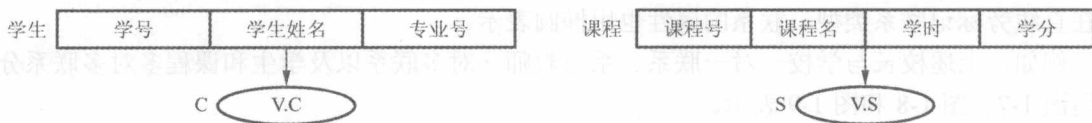


图 1-12 虚拟结点法表示学生课程多对多联系

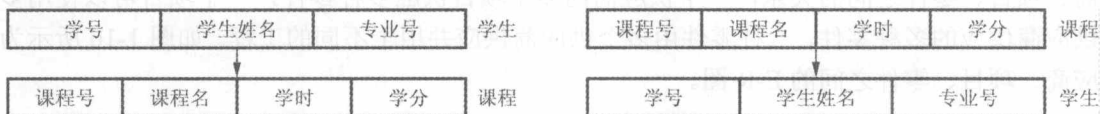


图 1-13 冗余结点法表示学生课程多对多联系

层次模型的主要优点是数据结构简单, 实体间的联系是固定的, 层次结构提供了良好的完整性支持。但是, 层次结构对处理多对多的联系比较复杂, 对插入和删除的限制较多。

1.2.4 网状模型

用网状结构表示实体及其之间联系的模型称为网状模型。网状结构的每一个结点代表一个实体, 网状结构突破了层次结构的两点限制: 允许一个结点有多个双亲结点, 可以有两个以上的结点没有双亲结点。因此, 网状模型可以方便地表示各种联系。如图 1-14 所示为学校网状模型。

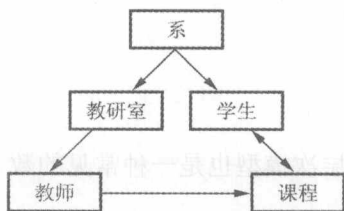


图 1-14 学校网状模型

网状模型的优点是能够直接描述现实世界, 具有良好的性能, 存取效率高。但是网状结构比较复杂, 而且随着应用环境的扩大, 数据库的结构变得越来越复杂, 用户使用麻烦。