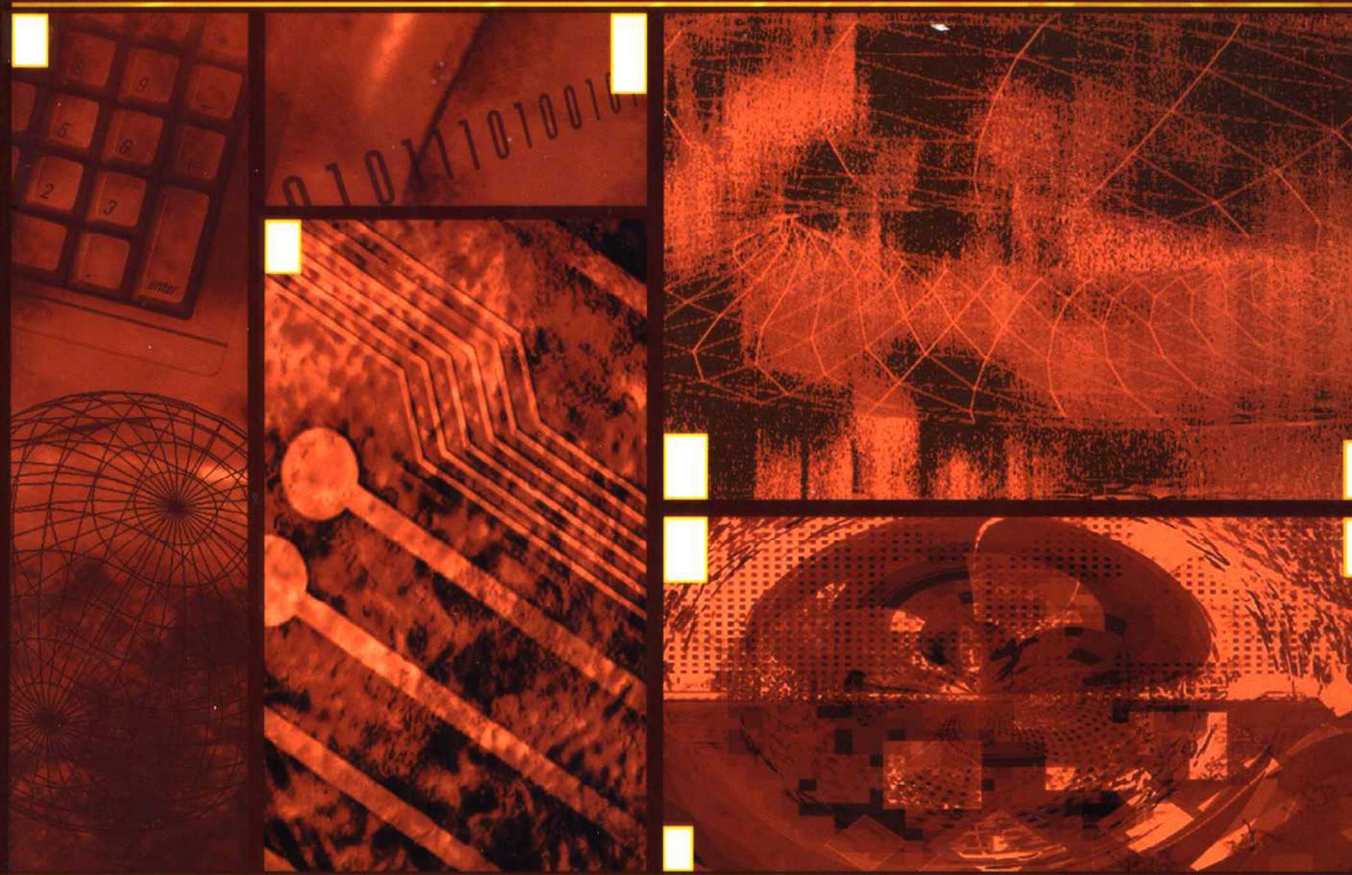


国家示范性软件职业技术学院系列教材

数据库基础 及Access应用

刘远东 何思文 吴斌新 编著

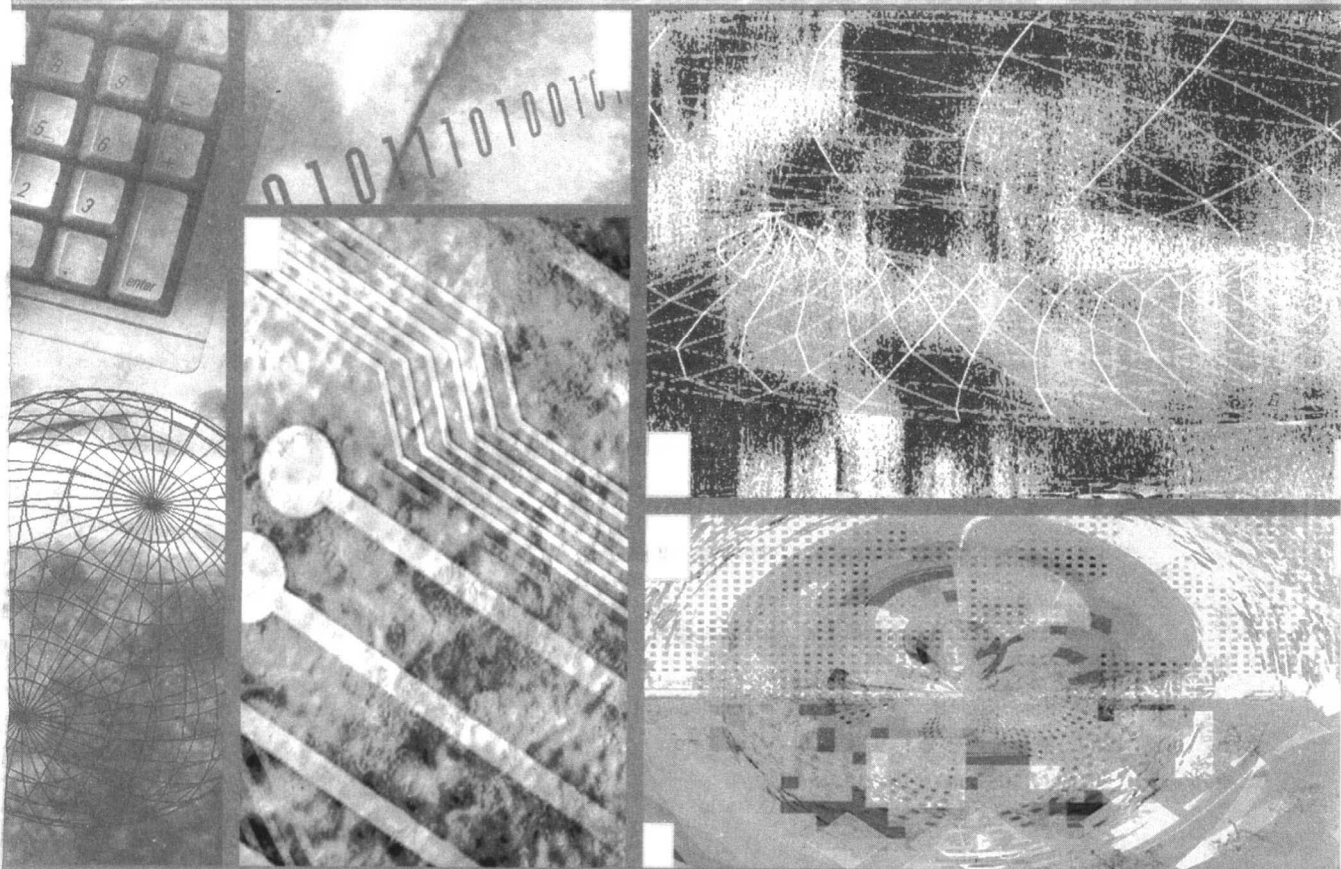


机械工业出版社
China Machine Press

国家示范性软件职业技术学院系列教材

数据库基础 及Access应用

刘远东 何思文 吴斌新 编著



机械工业出版社
China Machine Press

本书根据高职高专教育的特点和要求,以实用、够用为原则,以 Access 2000 为平台,具体结合一个实例“电脑配件信息管理系统”,采用通俗的语言,循序渐进、由浅入深地讲述了关系数据库的基本原理,以及利用关系数据库管理系统进行数据库应用系统开发的基本流程和方法,每章后均附有练习题,便于教学。

本书力求理论与实践紧密结合,兼顾系统学习与实际应用,除可作为大专院校的教科书外,也可供从事计算机开发与应用的工程技术人员自学与参考。

版权所有,侵权必究。

本书法律顾问 北京市展达律师事务所

图书在版编目(CIP)数据

数据库基础及 Access 应用/刘远东等编著. —北京:机械工业出版社, 2005. 6

(国家示范性软件职业技术学院系列教材)

ISBN 7-111-16444-X

I. 数… II. 刘… III. 关系数据库—数据库管理系统, Access—高等学校: 技术学校—教材 IV. TP311.138

中国版本图书馆 CIP 数据核字 (2005) 第 040813 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:王子恢

北京中兴印刷有限公司印刷·新华书店北京发行所发行

2005 年 6 月第 1 版第 1 次印刷

787mm × 1092mm 1/16 · 19.25 印张

印数: 0001—4000 册

定价: 29.00 元

凡购本书,如有倒页、脱页、缺页,由本社发行部调换
本社购书热线: (010) 68326294

出版者的话

近年来,随着我国教育的不断深入,我国高等职业教育发展迅速,无论是办学规模、数量和专业设置,还是办学条件和招生人数,都大大好于历史上任何一个时期。特别是继 2003 年教育部批准 35 所学校成立示范性软件职业技术学院以来,我国的软件人才培养格局得到了进一步改善,使软件人才培养形成了多层次、多样化的结构。两年前的这一重大改革尝试,对于加大软件人才培养力度和提高软件人才培养质量,构建功能配套的软件人才培养体系,尽快满足国家软件产业发展对高素质软件职业技术人才的迫切需求,实现软件产业人才培养方面的跨越式发展,推动高等职业教育办学体制、培养模式的改革,具有重大意义。实践证明,我国软件人才的培养模式和办学体制已发生重大改革,多元化的软件人才培养体系已形成,必将为我国软件产业的持续发展带来源源不断的推动力。

示范性软件职业技术学院的目标,是建设一批能够培养大量具有竞争能力的实用型软件技术人才的基地,面向就业,产学结合,为我国专科层次软件职业技术人才培养以及高等职业教育的持续健康发展起到示范作用。示范性软件职业技术学院有别于传统高职教育的鲜明特点是,使学生在学期间就能够参加实际的软件开发工作,突出以技术应用能力为核心的人才培养特点,以职业技术需要为依据,针对行业按需培养具有良好实践技能的一线应用型人才。

教材作为知识的载体,是人才培养过程中传授知识、训练技能和发展智力的重要工具之一,也是学校教学、科研水平的重要反映。它不仅是教师传授知识的媒体,而且也成为培养学生综合能力的媒介,在教学中起到保证教学质量、主导教学方向的作用。虽然软件职业技术教育才刚刚起步,但其发展速度非常之快,有些课程还来不及编写专用的教材,还不同程度地借用着本科教材或沿用“本科压缩型”的专科教材和使用自编校内讲义,或者理论偏深,或者内容过于肤浅,或者知识过于陈旧,或者缺乏高职特色且实践环节薄弱,缺乏实际工作中的应用实例。显然,这样的教材是不适合软件职业技术学院用来培养大量就业能力强的高质量实用型软件技术人才的。“工欲善其事,必先利其器”。为了实现软件职业技术学院的培养计划,不断推进软件职业技术教育的健康发展,必须使软件职业技术学院的教材建设与之相适应。为此,机械工业出版社华章分社秉承“出版要为教育服务”的理念,组织有关院校富有教学经验的多位教师,认真规划了当前急需的各种教材,从而倾心编写和出版了这套具有软件职业技术学院自身特色的实用教材。首批推出的是最有创新点和最具特色的专业基础课系列教材,其鲜明特色是:

(1) 定位准确。本着厚基础、重能力、求创新的总体思路,优化整合课程内容,从内容选材、教学方法、学习方法、实验和实训配套等方面突出高职教育的特点。按照国际软件产业的发展趋势和专业课教学计划,坚持学科定位,坚持办学方向,大力推动技术应用型人才的培养。突出对学生应用能力培养的特点,摆脱理论分析长而深的模式,增加并充实应用实例的内容。对职业岗位所需知识和能力结构进行恰当的设计安排。在知识的实用性、综合性上多下功夫,理论联系实际,加强操作与实训,把学生应用能力的培养融会贯通于教材之中。以就业为导向,打破传统学科教育的教学模式,以社会需要为目标,根据岗位和岗位群

所需能力与素质,提炼、整合成专业课程体系,构建新型高职教学模式,培养实践技能强、具有良好职业道德的技术应用型人才。

(2) 优势明显。繁荣高等职业教育教材事业,提高教材质量的基础和关键是教师。为此,我们建立了一支老、中、青相结合的教材编写队伍,他们都是知识功底扎实、教学经验丰富的主讲教师,所从事的都是其所在学校有学科优势的专业。他们在处理知识、能力和素质三者辩证统一的关系方面有独到见解。在内容的组织上,以素质教育为核心,正确把握新世纪教学内容和课程体系的改革方向,力争编写出高质量、有特色、有水平的好教材,以实现教材内容和体系的创新。

(3) 特色鲜明。这套教材充分体现了高等职业教育在教学改革建设中不断改进的理論教学方法和具有高职特色的培养模式,既跟踪现代科学技术的最新发展,适当阐述技术原理和依据,更重视案例式教学方法。各门基础理论课内容以必需和够用为度,其广度和深度取决于学习专业课的需要,“够用为主”。对学生向专业高层次发展很重要的那些内容,或扼要概述,或归为自学部分。这为学生根据所在岗位工作实际补充专业知识和进一步学习提供了便利,增强了可持续发展的能力。所选的例题、习题或案例均贴近实际,通过实习实训课练习实际操作,解决实践技能问题,强化动手能力训练,以应对学生的基本训练和能力培养,适应学生的接受能力和知识水平,适应学生的兴趣和需要,从而有利于快速培养满足社会需求的技术型与技能型人才。在这套教材的编写过程中,始终贯彻了能够使学生终身受益的“授之以渔”的教育思想,从而确保这套教材质优实用、特色十足。

(4) 精益求精。为了确保这套教材及时反映科学技术的变化,突出先进性;确保提高教材建设的水平,突出创新性;确保在课程教学大纲的要求范围内,在强化基本理论、基础知识、基本技能上下功夫,突出实用性;确保总结以往教材建设的经验,切实转变和更新教育思想和观念,突出针对性,各参编单位认真组织,精心筹划,从大纲拟定到内容撰写,反复推敲,几易其稿,彻底摒弃了低水平的“剪刀+浆糊”式的“编”书方式;各主审单位高度负责,对每次修改都提出了许多建设性的意见和建议。这对确保教材的品高质优、简洁实用起到了至关重要的决定作用。

(5) 科学严谨。主要体现在教材体系和教学内容两方面。一方面教材体系较好地体现了本门学科的内在科学逻辑,教材自身的章节与章节之间,内容、习题和附录之间等,都体现了科学的逻辑结构,恰当地反映了本学科与其他学科的外在联系。对基本概念、定理、定义的阐述准确无误,符合科学的抽象;对基本原理的论证、公式推导等符合科学的逻辑关系。另一方面,在教材内容安排上,从高职教育的特点出发,由浅入深,由简到繁,由具体到抽象,循序渐进,与教材的整个内容配套和协调,同时还涵盖了本学科领域的新进展、新方法、新技术,展示了本学科当前发展的最新水平,进而有利于培养出富有时代气息、适应未来专业和学科发展需要的现代应用型技术人才。

(6) 配套建设。为了给广大用书的教师和学生提供全方位的服务,在教材的配套建设方面,我们还为广大教师提供了相应的多媒体课件,从而极大地方便了教学。

为了不断改进教材的出版质量,我们将加强对教材的质量跟踪,采取多种形式调查了解教师、学生对使用教材质量的反映。我们热诚欢迎广大读者及时向我们反馈质量信息,以便在后续教材建设(编写出版一批有特色的专业主干课程教材)中时做得更好,为广大师生提供更多、更适用的优秀教材。

国家示范性软件职业技术学院系列教材

◆ 专家指导委员会 ◆

(按姓氏笔画排序)

马在强	四川托普信息技术职业学院
尹作林	内蒙古电子信息职业技术学院
王世卿	郑州大学软件技术学院
王路群	武汉软件职业学院
邓安远	九江学院信息科学与技术学院
包健	杭州电子科技大学软件学院
刘远东	深圳信息职业技术学院
宋改勤	郑州经济管理干部学院
张凤生	河北软件职业技术学院
张舜德	宁波大红鹰职业技术学院软件职业技术学院
李占波	郑州大学软件技术学院
李学向	郑州大学软件技术学院
杨东勇	浙江工业大学软件学院
陈秋劲	武汉软件职业学院
巫家敏	大连东软信息技术职业学院
幸莉珊	河北软件职业技术学院
郑宁	杭州电子科技大学软件学院
胡景德	成都东软信息职业技术学院
赵玉兰	哈尔滨华夏计算机职业技术学院

联络人 王子恢

前 言

随着信息技术的迅速发展与广泛应用，人类社会正处在信息化时代，面对纷繁复杂的大量信息，如何对其进行有效的管理和利用是人类社会面临的重大课题。数据库技术正是为了适应信息社会的需要而发展起来的一门综合性数据管理技术。

数据库技术始于 20 世纪 60 年代，数据库应用也从简单的事务管理扩展到各个应用领域，如用于工程设计的工程数据库、用于因特网的 Web 数据库、用于决策支持的数据仓库技术、用于多媒体技术的多媒体数据库等，但应用最广泛的还是在基于事务管理的各类信息系统领域。随着时代的进步和发展，数据库的应用领域会越来越广泛，数据库技术也将是所有信息技术和信息产业的基础。

本书根据高职高专教育的特点和要求，以实用、够用为原则，以 Access 2000 为平台，采用通俗的语言，循序渐进、由浅入深地讲述了关系数据库的基本原理，以及利用关系数据库管理系统进行数据库应用系统开发的基本流程和方法。

全书共分 11 章，其主要内容有：关系数据库的基础理论，包括数据库系统的组成和结构，概念模型与数据模型，SQL 语言基础、关系运算及规范化；数据库应用系统开发的基本方法；Access 数据库的基本功能介绍和安全管理；Access 数据库各对象的创建、使用和管理，包括表结构的设计、索引与表间关系、数据表视图及记录的操作、选择查询和动作查询、窗体对象设计、报表对象设计、数据访问页实现、宏对象设计、VBA 模块设计，以及集中实训等。

其中，第 1、2、10、11 章由吴斌新编写，第 3、4、5 章由刘远东编写，第 6、7、8、9 章由何思文编写；全书由刘远东统稿，何思文审阅。本书从案例的要求出发，具体结合自主开发的一个实例“电脑配件管理系统”，来讲解相关知识点和操作方法，且每章后均附有练习题，便于教学。

本书内容精练，实例丰富，概念清晰，可操作性强，已经过十几个班级的使用，并在使用过程中不断地进行修改和完善。读者在学习完本书后，会对数据库应用系统的开发有一个全面的认识，同时也掌握了 Access 数据库管理系统的使用。

限于作者水平，书中难免存在遗漏甚至不妥之处，恳请读者批评指正。

编 者
2005 年 1 月
于深圳

目 录

出版者的话
专家指导委员会
前言

第1章 数据库基本原理	1
1.1 数据管理的基本概念	1
1.1.1 信息与数据	1
1.1.2 数据库	1
1.1.3 数据库系统	2
1.1.4 数据库管理系统	2
1.2 数据管理及其发展	3
1.3 E-R图	5
1.4 数据模型	8
1.4.1 层次模型	8
1.4.2 网状模型	8
1.4.3 关系模型	9
1.5 关系运算	11
1.5.1 SQL语言基础	11
1.5.2 常用的SQL语句	11
1.5.3 传统的集合运算	14
1.5.4 选择运算	15
1.5.5 投影运算	16
1.5.6 连接运算	16
1.5.7 外键	18
1.6 关系规范化基础	18
1.6.1 规范化问题	19
1.6.2 属性的联系	20
1.6.3 函数依赖	22
1.6.4 关键字	24
1.7 关系的范式	25
1.7.1 第一范式	25
1.7.2 第二范式	26
1.7.3 第三范式	27
1.7.4 BCNF	28
1.7.5 规范化小结	28
1.8 数据库的保护	30

练习题	30
第2章 数据库设计基础	33
2.1 数据库设计概述	33
2.1.1 数据库应用系统的特点	33
2.1.2 数据库设计阶段	33
2.2 数据库设计的需求分析	34
2.3 数据库概念设计	36
2.3.1 设计局部E-R图	37
2.3.2 视图集成	37
2.4 数据库的逻辑设计	39
2.4.1 导出初始关系模式	39
2.4.2 规范化处理	40
2.4.3 模式评价与修正	40
2.5 数据库的物理设计	41
2.6 数据库实现和运行	42
练习题	42
第3章 关系数据库 Access 概述	43
3.1 关系数据库 Access 2000	43
3.1.1 Access 的特点	43
3.1.2 系统安装	44
3.1.3 启动与退出	45
3.1.4 操作界面	46
3.1.5 Access 的数据库对象	51
3.2 数据库创建	58
3.3 数据库管理	64
3.3.1 压缩和修复数据库	64
3.3.2 加密和解密数据库	65
3.3.3 为数据库设置密码	67
练习题	67
第4章 表对象	69
4.1 表对象和表结构	69
4.1.1 创建新表的五种方法	69
4.1.2 表结构	75
4.2 字段名称和字段说明	76
4.2.1 字段名称	76
4.2.2 字段说明	76

4.3 数据类型	77	5.5.2 生成表查询	146
4.4 字段属性	78	5.5.3 追加查询	149
4.4.1 字段属性的种类	78	5.5.4 删除查询	153
4.4.2 “字段大小”属性的使用	78	练习题	155
4.4.3 “格式”属性的使用	79	第6章 窗体对象	157
4.4.4 其他字段属性的使用	84	6.1 什么是窗体	157
4.5 主键和索引	90	6.1.1 窗体的三种视图	157
4.5.1 定义主键	90	6.1.2 窗体的类别	159
4.5.2 建立索引	91	6.1.3 窗体的结构	160
4.6 建立表间关系	93	6.1.4 窗体的显示特性	160
4.6.1 “关系”工具栏	95	6.2 使用向导建立窗体	161
4.6.2 建立表间关系	95	6.3 使用窗体设计视图创建窗体	164
4.6.3 编辑表间关系	96	6.3.1 进入窗体设计视图的步骤	164
4.7 数据的输入与修改	96	6.3.2 窗体控件工具箱	165
4.7.1 数据表视图	96	6.3.3 实例设计	166
4.7.2 编辑修改	98	6.4 窗体中的常用控件	168
4.7.3 查找和替换	99	6.4.1 窗体属性	168
4.7.4 排序和筛选	101	6.4.2 控件属性	169
4.7.5 修饰数据表视图	104	6.4.3 选项组	170
练习题	108	6.4.4 组合框	172
第5章 查询对象	111	6.4.5 列表框	174
5.1 查询的概念与创建	111	6.4.6 命令按钮	175
5.1.1 查询的概念和种类	111	6.4.7 子窗体	176
5.1.2 创建选择查询	111	6.4.8 设置控件标题	178
5.2 查询的设计	113	6.5 自动套用格式	178
5.2.1 设计视图的使用	113	6.6 外观设计	179
5.2.2 运行查询	116	6.6.1 控件大小和位置的调整	179
5.2.3 条件的使用	116	6.6.2 控件外观的修饰	180
5.2.4 改变字段属性	121	6.6.3 文字的美化	181
5.2.5 在查询中执行计算	122	练习题	181
5.2.6 修改查询	123	第7章 报表对象	183
5.3 几个复杂查询的设计	125	7.1 报表	183
5.3.1 汇总查询	125	7.1.1 什么是报表	183
5.3.2 交叉表查询	128	7.1.2 报表和窗体	183
5.3.3 参数查询	129	7.1.3 报表的形式	183
5.3.4 多表查询	130	7.1.4 报表的三种视图方式	184
5.4 查询向导	132	7.1.5 报表的结构	185
5.4.1 简单查询向导	133	7.2 使用向导创建报表	186
5.4.2 交叉表查询向导	135	7.3 利用设计视图修改报表	189
5.4.3 查找重复项查询向导	138	7.3.1 打开已有报表进行修改	189
5.4.4 查找不匹配项查询向导	140	7.3.2 用设计视图设计新报表	189
5.5 动作查询	143	7.3.3 添加日期和时间	190
5.5.1 更新查询	144	7.3.4 添加页码	191

7.3.5 添加线控件	192	9.4.1 直接执行宏	243
7.4 报表数据的重组	193	9.4.2 单步执行宏	243
7.4.1 排序	193	9.4.3 以事件响应方式执行宏	244
7.4.2 分组汇总	194	9.5 Autoexec 宏对象	245
7.5 表达式	196	练习题	245
7.5.1 什么是表达式	196	第10章 VBA 编程	247
7.5.2 表达式生成器	197	10.1 VBA 基础知识	247
练习题	199	10.1.1 数据类型	247
第8章 数据访问页	201	10.1.2 常量、变量和数组	248
8.1 什么是数据访问页	201	10.1.3 运算符与表达式	253
8.2 创建数据访问页	203	10.2 VBA 开发环境	256
8.2.1 使用“自动创建数据页”	203	10.3 模块和过程	257
8.2.2 使用向导	204	10.3.1 模块与过程的概念	257
8.2.3 使用设计视图	206	10.3.2 模块的建立	258
8.3 在数据访问页中重组记录	207	10.4 运行与调试	263
8.3.1 分组的作用	207	10.4.1 运行代码	264
8.3.2 分组记录	208	10.4.2 错误类型与检测	264
8.3.3 设置或更改记录的排列 顺序	208	10.4.3 调试方法	264
8.4 页对象中的基本控件及其应用	211	10.4.4 排错	265
8.4.1 标签	211	10.5 VBA 程序控制	268
8.4.2 文本框	211	10.5.1 分支结构	268
8.4.3 滚动文字	211	10.5.2 循环结构	271
8.4.4 超链接	212	10.5.3 控件对象的基本属性	274
练习题	213	练习题	279
第9章 宏对象	215	第11章 集中实训	281
9.1 什么是宏对象	215	11.1 实训目的	281
9.2 常用宏操作	218	11.2 实训内容	281
9.2.1 对象操作类	218	11.3 实训步骤	281
9.2.2 数据导入导出类	227	11.3.1 需求分析	281
9.2.3 记录操作类	233	11.3.2 概念结构设计	284
9.2.4 数据传递类	237	11.3.3 逻辑结构设计	285
9.2.5 代码执行类	237	11.3.4 物理设计	286
9.2.6 提示警告类	238	11.3.5 网络应用方案	290
9.2.7 其他类	239	11.3.6 用户权限分配	291
9.3 常用事件	242	11.3.7 系统发布	296
9.4 宏对象的调试与执行	243	11.4 实训小结	297

第1章 数据库基本原理

随着信息技术的迅速发展与广泛应用，人类社会正处在信息化时代。面对纷繁复杂的大量信息，如何对其进行有效的管理和利用是人类社会面临的重大课题。数据库技术正是为了适应信息社会的需要而发展起来的一门综合性数据管理技术。本章概括地介绍数据管理及数据库的发展过程，以及关系数据库的基本概念。

1.1 数据管理的基本概念

1.1.1 信息与数据

数据是描述事物的符号记录，如字符、数字、文本、声音、图形、图表、图像等。这些符号都可以输入到计算机中，由计算机进行存储和管理。

信息是由原始数据经过提炼、加工而得到的用来决定行动、计划和决策的有价值的数
据，或称赋予一定语义的数据。对这些数据进行收集、组织、加工、储存、抽取、传播等工作就是数据处理，它是指将数据转换成信息的过程。我们可以用下面的式子简单地表示信息、数据与数据处理的关系：

$$\text{信息} = \text{数据} + \text{数据处理}$$

数据和信息之间的关系如同原料和成品之间的关系一样，数据是原料，信息则是成品。

数据库技术要解决的主要问题就是如何科学地组织和存储数据，如何高效地获取数据、更新数据和加工处理数据，并保证数据的安全性、可靠性和持久性。数据库技术产生于 20 世纪 60 年代末，随后得到了迅速发展。其应用领域不断扩大，已经由最初的一般性事务处理发展到情报检索、人工智能、计算机辅助设计等各个领域。尤其是 Internet 的出现以及多种信息技术的交汇，给数据库技术的应用提供了更广阔的空间，从而促进了数据库技术的快速发展，数据库的地位也与日俱增。

随着计算机应用的普及，数据库技术已成为人们进行数据处理不可缺少的工具，因此，了解数据库系统的基本原理，掌握数据库设计的基础理论将有助于人们更好地利用数据库技术，提高工作和生活的信息化水平以适应信息社会的要求。

1.1.2 数据库

数据库 (Database, DB) 是数据的集合，具有统一的结构形式并存放于统一的存储介质内，是多种应用数据的集成，并可被各个应用程序所共享。

数据库存放数据是按照数据所提供的数据模式存放的，它能构造复杂的数据结构以建立数据间内在联系与复杂的关系，从而构成数据的全局结构模式。

数据库中的数据具有“集成”、“共享”的特点，也就是数据库集中了各种应用的数据，进行统一的构造与存储，而使它们可被不同应用程序所使用。

1.1.3 数据库系统

数据库系统是计算机系统的重要组成部分，是指引进了数据库技术后的计算机系统。它除了包括支持数据库系统的计算机硬件系统和操作系统之外，还包括以数据为主体的数据库（数据文件的集合）、管理数据库的软件系统（即数据库管理系统）以及基于数据库管理系统而开发的用户应用程序。此外，还包括数据库管理员和用户等。整个数据库系统的层次结构关系如图 1-1 所示。

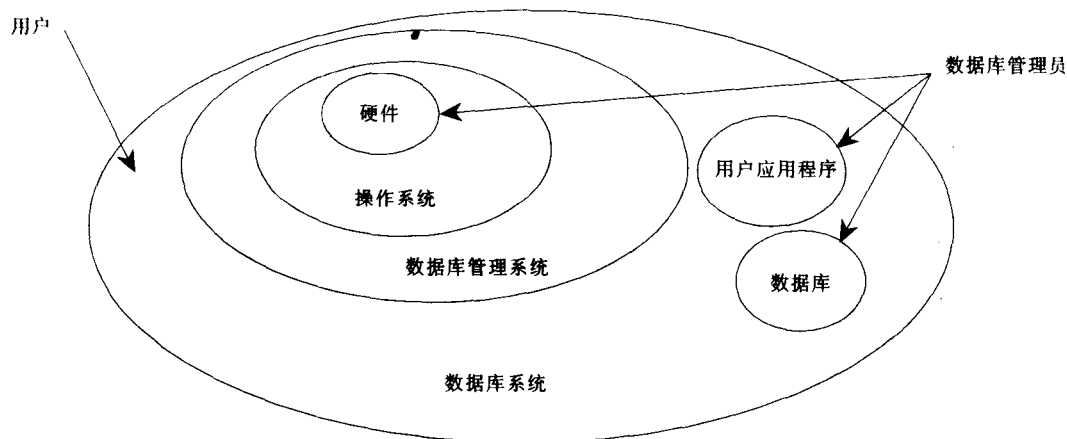


图 1-1 数据库系统的层次结构关系

其中，数据库文件等数据资源存储在外存储器上，数据库管理系统是为数据库的建立、使用和维护而配备的软件，操作系统是数据库管理系统与硬件的接口，用户通过数据库应用系统在逐层系统的支持下使用数据库，数据库管理员（Database Administrator）负责整个系统的管理工作。一般情况下，对于微机数据库系统，高级用户就是 DBA，DBA 应该懂得数据库的基本原理，充分掌握数据库管理系统的性能，有较强的程序设计和维护能力。

1.1.4 数据库管理系统

数据库管理系统（Database Management System, DBMS）是数据库系统的核心软件，它是一种系统软件，负责数据库中的数据组织、数据操纵、数据维护、控制及保护和数据服务等，其主要目标是使数据成为方便用户使用的资源，易于为各类用户所共享，并增进数据的安全性、完整性和可用性。它建立在操作系统的基础之上，对数据库进行统一的管理和控制。具体来说，DBMS 提供以下几大功能：

1) 数据库的定义 提供数据定义语言（Data Definition Language, DDL）或操作命令，以便对各级数据模式进行具体描述。

2) 数据操纵 提供数据操纵语言（Data Manipulation Language, DML），以便用户对数据库中数据进行追加、插入、修改、删除、查询等操作。

3) 数据运行控制 为便于多个用户共享数据库中的数据资源，DBMS 提供数据库的并发

操作控制、数据的安全性和完整性控制功能。

4) 数据字典 数据字典 (Data Dictionary, DD) 中存放着对实际数据库各级模式所作的定义, 即对数据库结构的描述。DD 提供了集中管理数据库数据描述的方法, 并通过查阅数据字典实现对数据库的使用和操作。

1.2 数据管理及其发展

所谓数据管理, 是指如何用计算机管理数据, 是数据处理的中心内容。随着计算机技术的发展, 计算机数据管理经历了由低级到高级的发展过程。主要围绕提高数据独立性、降低数据的冗余度, 在数据共享、数据的安全性和完整性等方面进行改进, 让用户能方便地管理和运用数据资源。计算机数据管理主要经历了人工管理阶段、文件系统阶段、数据库系统阶段和分布式数据库系统阶段。

1. 人工管理阶段

人工管理阶段是指计算机诞生的初期 (20 世纪 50 年代后期之前), 这个时期的计算机主要用于科学计算。在硬件上, 没有磁盘等存储设备, 在软件方面, 没有操作系统和管理数据的软件, 数据处理方式是批处理。这个时期数据管理的特点是:

1) 数据不保存。由于这时计算机主要用于科学计算, 一般不需要将数据长期保存。

2) 没有对数据进行管理的软件。程序员不仅要规定数据的逻辑结构, 而且还要在程序中设计物理结构, 包括存储结构、存取方法、输入输出方式等, 给程序设计人员带来了极大的负担。

3) 数据不独立于程序, 而是面向应用的, 即使两个程序用到相同的数据, 也必须各自定义、各自组织, 数据无法共享 (如图 1-2 所示), 因而产生大量的重复数据, 并且一旦数据发生变化, 所有相关程序都要进行修改。

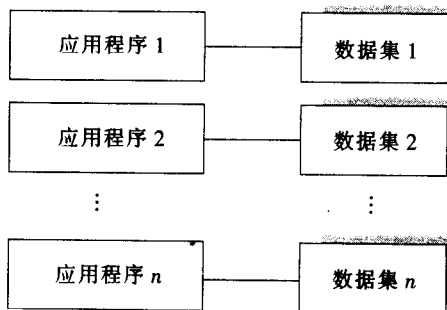


图 1-2 人工管理阶段数据与程序关系图

2. 文件系统阶段

文件系统阶段是指计算机不仅用于科学计算, 而且开始大量用于管理中的数据处理工作阶段 (20 世纪 50 年代后期到 60 年代中期)。在硬件方面, 有了磁盘、磁鼓等直接存取的存储设备; 在软件方面, 操作系统中已经有了专门用于管理数据的软件, 称为文件系统。在处理方式上, 不仅有了文件批处理, 而且有了联机实时处理。这个时期数据管理的特点是:

1) 数据需要长期保留在外存储器中供反复使用。由于计算机大量用于数据处理, 经常对文件进行查询、修改、插入、删除等操作, 所以数据需要长期保留, 便于反复操作。

2) 程序和数据之间有了一定的独立性。操作系统提供了文件管理功能和访问文件的存取方法, 程序和数据之间有了数据存取的接口, 程序可以通过文件名和数据打交道, 而不必寻找数据的物理位置。至此, 数据有了物理结构和逻辑结构的区别, 但此时程序和数据之间的独立性不够充分。

3) 文件的形式已经多样化。由于已经有了直接存取的存储设备, 文件也就不再局限于顺序文件, 还有了索引文件、链表文件等。因此, 对文件的访问可以是顺序访问, 也可以是

直接访问。

4)数据的存取基本上以记录为单位(如图1-3所示)。

由图1-3可以看出,数据和程序之间虽然具有了一定的独立性,但还不够充分,每个文件仍然对应于一个应用程序,数据和程序相互依赖,一旦改变数据的逻辑结构,相应的应用程序也必须修改,并且各个程序之间不能共享相同的数据,因而数据冗余度大,并容易产生数据的不一致性。

3. 数据库系统阶段

数据库系阶段是从20世纪60年代后期开始的。在这个阶段,计算机的应用更加广泛,数据量剧增,共享要求更高。软件价格上升,硬件价格相对下降,文件系统的管理方法已经无法适应开发应用系统的需要。为了解决数据的独立性问题,实现数据的统一管理,达到数据共享的目的,发展了数据库技术,从而出现了数据库系统。数据库中的数据不再是面向某个应用程序而是面向整个应用的,如图1-4所示。

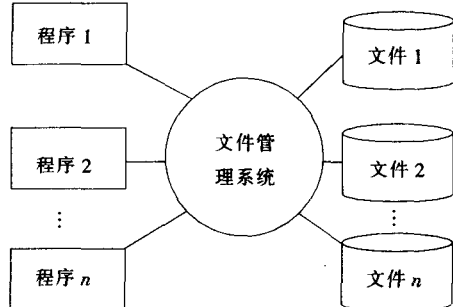


图1-3 文件系统的模型

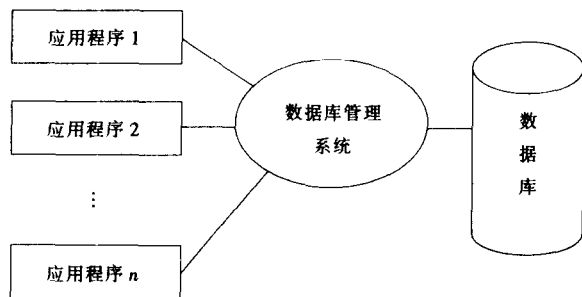


图1-4 数据库管理阶段数据与程序关系

数据库系统的主要特点是:

1)实现了数据共享,减少了数据冗余。在数据库系统中,对数据的定义和描述已经从应用程序中分离出来,通过数据库管理系统统一管理,从而保证了数据和程序的相对独立性。这样,数据可以供用户共享,实现最小冗余度,并能更好地保证数据的安全性和完整性。

2)采用特定的数据模型。将数据按照一定的数据模型组织起来,形成数据库,并由DBMS统一管理,便于用户操作使用。

3)真正统一的数据控制功能。数据库作为多个用户和应用程序的共享资源,对数据的存取往往是并发的,而DBMS恰恰可以提供并发控制功能、数据的安全性控制功能和数据的完整性控制功能。

4. 分布式数据库系统阶段

20世纪70年代之前,数据库系统多数是集中式的。网络技术的发展为数据库提供了越来越好的运行环境,使数据库系统从集中式发展到分布式,从主机-终端结构发展到客户/服务器结构。

分布式数据库系统是数据库技术和网络技术相结合的产物。它是一个逻辑上统一，地域上分布的数据集合，是计算机网络环境中各个结点的局部数据库的逻辑集合，同时受分布式数据库管理系统的控制和管理，其结构如图 1-5 所示。

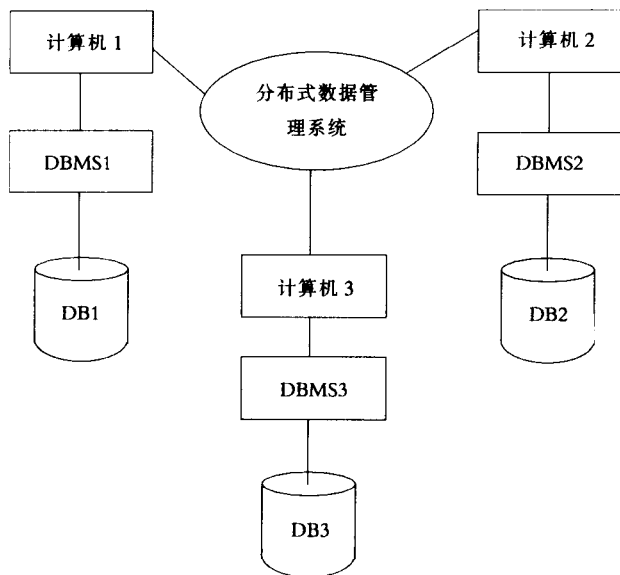


图 1-5 分布式数据库系统

其主要特点是：

1) 局部自主 每个结点有自己的数据库管理系统，可以独立地处理本地的事务，而各个结点之间可以互相访问，通过有效的配合处理复杂事务，如银行的通存通兑及划汇业务、保险业跨地区处理、国际洲际民航订票业务的受理、连锁商店的管理、军事情报系统、决策指挥系统等。

2) 较强的可靠性 个别结点或通信链路发生故障，不会导致整个系统的破坏。

3) 适当的数据冗余 数据可以在不同结点上留有副本，系统根据距离选择用户最近的副本数据进行操作，以减少通信代价，增加系统的可靠性。

4) 灵活方便 分散了工作负荷，缓解了单机容量的压力，数据可以存储在临近的结点上，比集中式数据库更加节省查找时间。

5) 易于扩展 分布式数据库系统扩大规模要比集中式系统更加方便、灵活与经济。

1.3 E-R 图

把客观存在的事物以数据的形式存储到计算机中，通常需经历对现实世界中事物的认识、概念化到计算机数据库里的具体表示等抽象过程，即从现实世界到形成概念，再通过计算机进行处理。

概念世界是现实世界在人们头脑中的反映，是对客观事物及其联系的一种抽象描述。它不是现实世界的简单映像，而要经过选择、命名、分类等抽象过程而产生概念模型。建立概念模型涉及到下面几个名词：

1) 实体 客观存在且相互区别的事物称为实体。实体既可以是实际事物，也可以是抽象事件。例如，一个职员、一个部门是实际事物，而一次商品销售、职员就职于某个部门则是较抽象的事件，它们都是实体。

2) 属性 描述实体的特性称为属性。例如，职工实体可以用职工编号、姓名、性别、出生日期、职称、部门、职务等属性来描述。属性的具体取值称为属性值，用以刻画一个具体实体。例如，属性值的组合（0001、李明、男、08/10/1968、工程师、工艺科、科长）就刻画了李明这个职工。

3) 关键字 如果某个属性或几个属性组合的值能够唯一地标识出实体集中的每一个实体，就可以选作关键字。上例中的“职工编号”可作为关键字。由于可能有重名的职工存在，所以“姓名”不宜作关键字。

4) 联系 实体集之间的对应关系称为联系，它反映现实世界事物之间的相互关联。联系分为两种，一种是实体内部各属性之间的联系，如光的波长与颜色之间存在联系；另一种是实体之间的联系，如教师与学生两个实体之间就存在着教学联系。

存入计算机系统里的数据是将概念世界中的事物数据化的结果。为了准确地反映事物本身及事物之间的各种联系，数据库中的数据必须有一定的结构，这种结构用数据模型来表示。数据模型将概念世界中的实体、实体之间的联系进一步抽象成便于计算机处理的形式。

实体-联系模型（简称 E-R 图）是描述概念世界、建立概念模型的实用工具。数据库设计工作比较复杂，将现实世界的数据库组织成符合具体数据库管理系统所采取的数据模型，一般情况下不可能一步到位。P. P. S. Chen 于 1976 年提出形象的“实体-联系”方法。通过绘制 E-R 图可以描述组织模式，如一个企业的整体数据关联模式。它可以进一步转换为任何一种 DBMS 所支持的数据模型。E-R 图包括三个要素：

- 实体 用矩形框表示，框内标明实体名称。
- 属性 用椭圆形表示，并用连线与实体连接起来。
- 实体之间的联系 用菱形框表示，框内标明联系名称，并用连线将菱形框分别与有关实体相连，且在连线上标明联系类型。

例如，货品实体的属性有货号、货名、规格、单价等，其 E-R 图如图 1-6 所示。

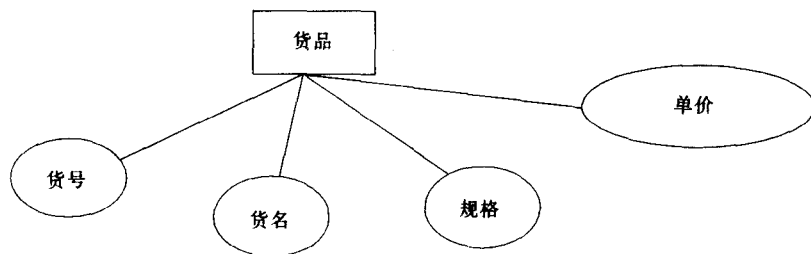


图 1-6 实体与属性

实体与实体间的联系可归结为三种类型：

(1) 一对一联系

设 A 、 B 为两个实体集。若实体集 A 中的每一个实体至多和 B 中的一个实体有联系，而 B 中的每个实体至多和 A 中的一个实体有联系，则称 A 对 B 或 B 对 A 是 1:1 联系。例如，民

航售票系统中, 乘客实体集中的一位乘客只能拥有一张机票, 而机票实体集中, 一张机票也只能对应一位乘客, 这样, 乘客实体与机票实体就构成了一对一联系。

(2) 一对多联系

若实体集 A 中的每一个实体可以和实体集 B 中的几个实体有联系, 而实体集 B 中的每一个实体至多和实体集 A 中的一个实体有联系, 则称实体集 A 对 B 是 $1:m$ 联系。例如, 一个学校的系与学生之间是一对多联系, 一个系可以有多名学生, 而一个学生只能在一个系注册。

(3) 多对多联系

若实体集 A 中的每一个实体都可与 B 中的多个实体有联系, 而实体集 B 中的每一个实体也可与 A 中的多个实体有联系, 则 A 对 B 或 B 对 A 是 $m:n$ 联系。例如, 在单位中, 一个职工可以参加多个项目的工作, 一个项目有多个职工参加, 则职工与项目之间是多对多联系。

下面我们通过一个简单的例子来说明如何用 E-R 图来描述概念模型, 并用 E-R 图来表示一个单位采购的概念模型。在这个系统中涉及的实体有:

- 1) 进货数据记录 属性有进货序列号、进货人、货号、货名、进货单价、进货数量、进货日期。
- 2) 库存数据 属性有货号、买入数量、卖出数量、库存数量。
- 3) 供货商 属性有供货商名称、供货人、地址、电话。
- 4) 销售数据记录 属性有销售序列号、售货人、货号、销售单价、销售数量、销售日期。

实体之间有如下联系:

- 一个仓库可以放多种货品。
- 一种货品可以多次销售。
- 一个供货商可以提供多种货品, 一种货品可由多个供货商提供。

其 E-R 图如图 1-7 所示。

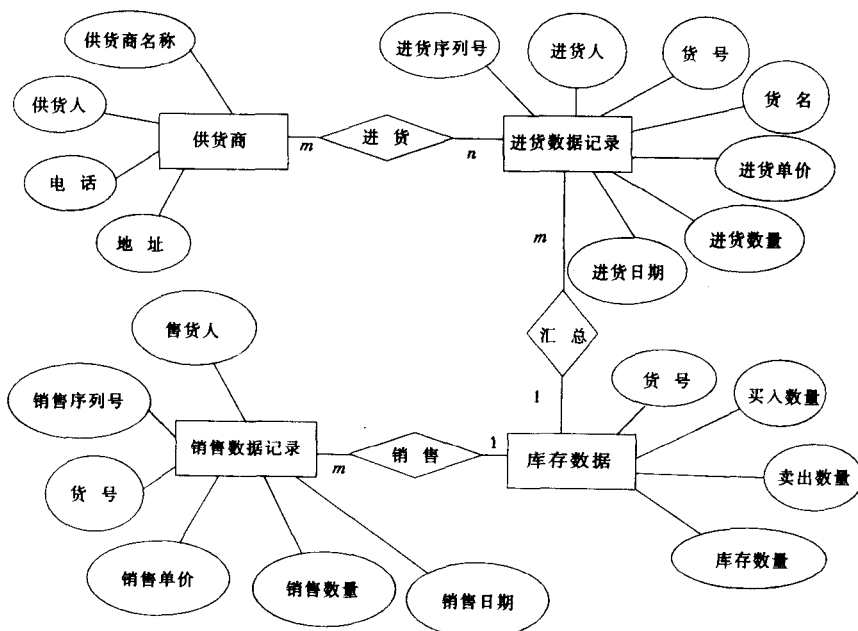


图 1-7 进货、销售、库存概念模型的 E-R 图