

DATABASE AND DATA PROCESSING

数据库与数据处理

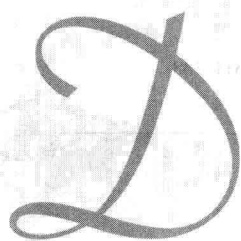
Access 2010实现

张玉洁 孟祥武 编著



机械工业出版社
China Machine Press

计算机基础课程系列教材

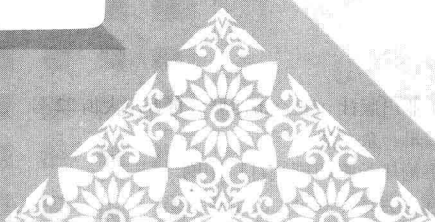
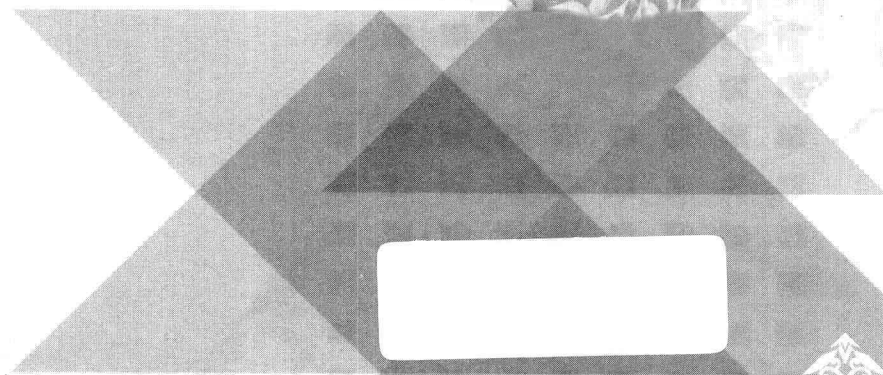
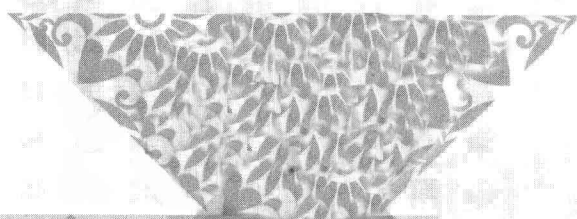


DATABASE AND DATA PROCESSING

数据库与数据处理

Access 2010实现

张玉洁 孟祥武 编著



机械工业出版社
China Machine Press

图书在版编目 (CIP) 数据

数据库与数据处理: Access 2010 实现 / 张玉洁, 孟祥武编著. —北京: 机械工业出版社, 2013. 1
(计算机基础课程系列教材)

ISBN 978-7-111-40611-2

I. 数… II. ①张… ②孟… III. 关系数据库系统—数据库管理系统—教材 IV. TP311.138

中国版本图书馆 CIP 数据核字 (2012) 第 283594 号

版权所有·侵权必究

封底无防伪标均为盗版

本书法律顾问 北京市展达律师事务所

全书共分为 14 章, 包括数据库理论和数据库应用两大部分。其中, 第 1~6 章为理论部分, 主要介绍与数据处理和数据库应用密切相关的核心概念、基本原理和方法, 内容包括: 数据处理与数据库概述、实体-联系数据模型、关系数据库基础、SQL、关系数据库设计、数据库保护; 第 7~14 章为应用部分, 以 Access 2010 为具体的 DBMS, 详细介绍其主要功能和使用方法, 内容包括: 认识 Access 2010、数据库的创建与管理、表与表操作、查询、窗体、报表、宏、模块与 VBA 编程。各章安排有例题讲解、重要提示和本章内容小结以及适量的习题。

本书既可作为高等院校非计算机专业本科、专科学生的数据库课程的教科书, 也可用作 Access 数据库使用人员及管理人士的参考书。

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: 余 洁

三河市杨庄长鸣印刷装订厂印刷

2013 年 1 月第 1 版第 1 次印刷

185mm×260mm·18.75 印张

标准书号: ISBN 978-7-111-40611-2

定 价: 35.00 元

凡购本书, 如有缺页、倒页、脱页, 由本社发行部调换

客服热线: (010) 88378991 88361066

投稿热线: (010) 88379604

购书热线: (010) 68326294 88379649 68995259

读者信箱: hzsj@hzbook.com

前 言

数据库技术是计算机领域中最重要技术之一,近年来,越来越多的高校将数据库技术课程列入了非计算机专业的计算机基础课程的教学范畴。与此同时,随着社会各个领域对信息管理应用人才需求的迅猛增长,数据库技术课程的普及呈现出进一步扩展和延伸的趋势。

数据库技术包括数据库系统原理与数据库系统应用两大部分,只有在数据库理论的指导下,才能设计出一个好的数据库应用系统。本书结合理论与实践,在参考《中国高等院校计算机基础教育课程体系 2008》(中国高等院校计算机基础教育改革课题调研组编制)的基础上,根据作者多年从事非计算机专业数据库相关课程的教学经验编写而成。

本书立足于“以理论为基础,以实例为引导,以应用为目的”,按照数据库设计主要阶段的先后顺序安排内容的讲述,所选内容经典而丰富,力求深入浅出、循序渐进,以应用为目的,用任务来驱动,使学生掌握基本的数据库系统设计和实现方法,并培养分析和解决问题的能力。

全书包括数据库理论和数据库应用两大部分,共分为 14 章。其中,前 6 章为理论部分,将数据库系统原理中最重要、最核心的内容提炼出来,进行循序渐进、深入浅出的介绍,用于指导数据库应用系统的设计;后 8 章为应用部分,以 Access 2010 为操作环境,详细介绍了 Access 的主要功能和使用方法。Access 是目前广泛流行的、采用面向对象技术进行数据库应用系统开发的一种小型数据库管理系统,具有界面友好、易学实用的特点,非常适合作为非计算机专业数据库相关课程的教学实践环境,便于初学者快速掌握数据库的基本概念和操作方法。

本书各章均安排有例题讲解、重要提示和本章内容小结以及适量的习题。此外,在应用部分各章节中还安排了精心设计的上机练习题,旨在引导学生主动思考,树立创新意识,培养实际操作和实践动手能力。

在本书的出版过程中,机械工业出版社华章分社的领导和编辑同志为本书出版付出了辛勤的劳动,在此对他们表示衷心的感谢。另外,在本书的编写过程中,编者参考了书末所列出的有关文献,在此也向这些文献的作者一并表示感谢。

由于编者水平有限,书中难免有不妥之处,竭诚欢迎读者多提宝贵意见。

编 者

2012 年 10 月于北京

教学建议

教学章节	教学要求	课时
第 1 章 数据处理与数据库概述	理解数据处理、数据管理的基本概念； 了解数据管理技术的发展； 掌握数据库、数据库管理系统、数据库系统、数据库管理员、数据字典的概念； 理解数据抽象与数据视图的概念和关系； 熟练掌握数据模型的相关概念、三要素、分类； 熟练掌握数据库系统的特点和体系结构。	2~4
第 2 章 实体-联系数据模型	掌握 E-R 模型的相关概念； 掌握利用 E-R 图进行概念数据库设计的原则和方法。	2~4
第 3 章 关系数据库基础	熟练掌握关系模型的相关概念； 掌握 E-R 模型转换为关系模型的方法； 掌握关系代数及关系运算。	4
第 4 章（选讲） 关系数据库标准语言 SQL	了解 SQL 的特点和组成； 掌握 SQL 查询的基本结构； 掌握聚集函数的使用； 熟练掌握数据定义、数据查询、数据更新、数据控制的功能和使用。	0~4
第 5 章 关系数据库设计	了解关系数据库设计中出现的问题； 理解并掌握函数依赖、规范化、范式的相关概念； 熟练掌握关系规范化的过程； 了解数据库设计的过程。	2~4
第 6 章（选讲） 数据库的保护	掌握事务的基本概念、事务的四个特性； 掌握数据库的并发控制、数据库的恢复机制； 了解数据库的完整性、数据库的安全性机制。	0~4
第 7 章 认识 Access 2010	了解 Access 2010 的运行环境和操作环境； 了解 Access 2010 的数据库对象。	1
第 8 章 数据库的创建与管理	了解 Access 2010 提供的各种数据库模板； 理解 Access 2010 中数据库的概念； 熟练掌握创建数据库、数据库的操作与维护的各种方法。	2
第 9 章 表与表操作	理解并掌握表与表结构、表的视图、字段的数据类型和属性、主键等相关概念； 熟练掌握创建数据表的各种方法； 熟练掌握创建表间关系的方法； 掌握表的复制、重命名、删除，修改表结构，表外观的调整与修饰等操作； 掌握表中记录的添加与删除、复制与修改、排序与筛选、查找与替换等操作； 掌握导出表的方法。	3~5
第 10 章 查 询	掌握查询类型、查询视图的相关概念，以及查询条件的设置方法； 熟练掌握选择查询、参数查询、交叉表查询、操作查询的创建方法； 了解 SQL 查询和嵌套查询的创建方法。	4~8

(续)

教学章节	教学要求	课时
第 11 章 窗 体	掌握窗体的类型、窗体的视图、窗体的组成和窗体的属性等相关概念； 熟练掌握常用控件的概念和使用方法； 掌握创建窗体的方法，熟练掌握运用窗体设计器进行窗体设计和修改的方法； 掌握创建主/子窗体的意义和方法。	4~6
第 12 章 报 表	掌握报表的组成、类型以及报表的各种视图； 掌握创建报表的方法，尤其要熟练掌握运用报表设计器进行报表设计和修改的方法； 掌握编辑报表的各种方法，尤其要熟练掌握排序与分组、计算和汇总以及设置报表属性和格式的方法； 掌握创建子报表的方法。	2
第 13 章 宏	掌握宏的定义、结构、类型等相关概念，以及宏生成器的用法； 了解常用的宏操作； 熟练掌握创建简单宏、宏组和条件宏的方法和步骤； 掌握宏的运行方法以及调试方法； 熟练掌握将宏与控件的事件过程进行绑定的方法； 了解将宏转换为 VB 代码的过程。	2
第 14 章 模块与 VBA 编程	掌握模块的组成和分类、模块和过程的概念以及创建方法； 熟悉掌握 VBA 编程语言的各个组成部分和相关语法； 熟练掌握使用 VBA 编写事件过程的方法； 熟练掌握使用 VBA 创建窗体模块的方法。	4~6
总课时	第 1~14 章建议课时	32~56
	上机练习建议课时（不包括自由上机）	16~32

说明：1) 建议 Access 部分的课堂教学尽量在多媒体机房内完成，实现“讲—练”结合。

2) 实际教学中，建议理论部分教学学时与应用部分教学学时的比例为 1:2。

3) 不同学校可以根据各自的教学要求和计划学时数对教学内容进行取舍。

目 录

前言

教学建议

第1章 数据处理与数据库概述	1
1.1 数据处理	1
1.1.1 数据与信息	1
1.1.2 数据处理与数据管理	1
1.1.3 数据管理技术的发展	2
1.2 数据库基本概念	3
1.2.1 数据库	3
1.2.2 数据库管理系统	3
1.2.3 数据库系统	4
1.2.4 数据库管理员	5
1.2.5 数据字典	5
1.3 数据视图	5
1.3.1 数据抽象	5
1.3.2 模式与实例	6
1.4 数据模型	6
1.4.1 数据模型的三要素	6
1.4.2 数据模型的分类	7
1.4.3 逻辑数据模型	8
1.5 数据库系统的体系结构	10
1.5.1 数据库系统的三级模式结构	10
1.5.2 数据独立性与二级映像	12
1.6 小结	13
习题	13
第2章 实体-联系数据模型	14
2.1 E-R模型概述	14
2.1.1 E-R模型的基本概念	14
2.1.2 完整性约束	15
2.1.3 E-R图	17
2.2 基于E-R模型的概念数据库设计	20
2.2.1 设计问题	20
2.2.2 设计原则	21
2.2.3 设计方法	24

2.3 小结	27
习题	27
第3章 关系数据库基础	29
3.1 关系数据模型基础	29
3.1.1 关系数据模型概述	29
3.1.2 关系及相关概念	30
3.1.3 关系的完整性约束	33
3.2 E-R模型转换为关系模型	35
3.2.1 将实体转换为关系	35
3.2.2 将联系转换为关系	36
3.3 关系代数及关系运算	40
3.3.1 关系代数基础	40
3.3.2 传统的集合运算	40
3.3.3 专门的关系运算	42
3.3.4 关系运算举例	44
3.4 小结	46
习题	46
第4章 关系数据库标准语言 SQL	48
4.1 SQL概述	48
4.1.1 SQL的特点	48
4.1.2 SQL的组成	48
4.1.3 SQL的几个基本概念	49
4.2 SQL查询的基本结构	49
4.2.1 SELECT-FROM-WHERE结构	49
4.2.2 SELECT子句	50
4.2.3 FROM子句	51
4.2.4 WHERE子句	51
4.2.5 其他子句	52
4.3 聚集函数	54
4.3.1 常用的聚集函数	54
4.3.2 聚集函数举例	55
4.4 数据定义	55
4.4.1 SQL模式的创建与撤销	55
4.4.2 基本表操作	56
4.5 数据查询	58

4.5.1 简单查询	58	6.2 数据库的并发控制	86
4.5.2 集合查询	59	6.2.1 并发控制概述	86
4.5.3 连接查询	60	6.2.2 封锁技术	88
4.5.4 嵌套查询	61	6.3 数据库的恢复	90
4.6 数据更新	63	6.3.1 数据库故障	91
4.6.1 插入数据	64	6.3.2 恢复的实现技术	91
4.6.2 修改数据	64	6.4 数据库的完整性	93
4.6.3 删除数据	65	6.4.1 完整性约束条件	93
4.7 视图和索引的应用	65	6.4.2 完整性控制	95
4.7.1 视图与索引概述	65	6.5 数据库的安全性	95
4.7.2 视图操作	66	6.6 小结	98
4.7.3 建立与删除索引	67	习题	99
4.8 数据控制	67	第7章 认识 Access 2010	100
4.8.1 授权	68	7.1 Access 2010 概述	100
4.8.2 收回授权	68	7.1.1 Access 2010 的发展	100
4.9 小结	68	7.1.2 Access 2010 的功能	100
习题	69	7.2 Access 2010 的运行环境	101
第5章 关系数据库设计	71	7.2.1 硬件环境	101
5.1 关系数据库设计中出现的问题	71	7.2.2 软件环境	101
5.2 函数依赖	72	7.2.3 Access 2010 的启动与退出	101
5.2.1 函数依赖的定义	72	7.3 Access 2010 的操作环境	101
5.2.2 函数依赖的分解和结合	73	7.3.1 Backstage 视图	102
5.2.3 平凡函数依赖与非平凡函数依赖	73	7.3.2 功能区	103
5.2.4 完全函数依赖与部分函数依赖	74	7.3.3 快速访问工具栏	106
5.2.5 传递函数依赖	75	7.3.4 导航窗格与工作区	107
5.3 关系模式的规范化	75	7.4 Access 2010 的数据库对象	108
5.3.1 范式与规范化	75	7.4.1 数据表	108
5.3.2 第一范式	76	7.4.2 查询	109
5.3.3 第二范式	76	7.4.3 窗体	110
5.3.4 第三范式	77	7.4.4 报表	110
5.3.5 BCNF	78	7.4.5 宏	111
5.4 数据库设计过程	79	7.4.6 模块	112
5.4.1 数据库设计概述	79	7.5 SharePoint 网站	113
5.4.2 数据库设计的基本步骤	79	7.6 小结	114
5.5 小结	82	习题	114
习题	83	上机练习题	114
第6章 数据库的保护	84	第8章 数据库的创建与管理	115
6.1 事务	84	8.1 数据库模板	115
6.1.1 事务的基本概念	84	8.1.1 Web 数据库模板	115
6.1.2 事务的四个特性	85	8.1.2 客户端数据库模板	116

8.2 创建数据库	117	9.6.3 记录的复制与修改	150
8.2.1 利用模板创建数据库	117	9.6.4 记录的排序与筛选	150
8.2.2 利用 Office.com 上的模板创建 数据库	119	9.6.5 数据的查找与替换	152
8.2.3 创建空数据库	119	9.7 设置表的显示格式	152
8.3 数据库的操作与维护	120	9.7.1 调整行高和列宽	152
8.3.1 打开数据库	120	9.7.2 隐藏列与显示隐藏列	152
8.3.2 保存与关闭数据库	122	9.7.3 冻结列与取消冻结	153
8.3.3 压缩和修复数据库	123	9.7.4 文本格式设置	153
8.3.4 备份数据库	124	9.8 导出表	153
8.3.5 查看和编辑数据库属性	124	9.9 小结	154
8.4 小结	125	习题	154
习题	125	上机练习题	154
上机练习题	125	第 10 章 查询	155
第 9 章 表与表操作	126	10.1 查询概述	155
9.1 表概述	126	10.1.1 查询的概念	155
9.1.1 表与表结构	126	10.1.2 查询的类型	155
9.1.2 表的视图	127	10.1.3 查询的视图	156
9.2 设计数据表结构	129	10.1.4 查询的创建方法	157
9.2.1 字段的命名	129	10.1.5 查询的保存、运行与结果显示	158
9.2.2 字段的数据类型	129	10.2 设置查询条件	158
9.2.3 字段的常规属性	132	10.2.1 表达式	158
9.2.4 主键	136	10.2.2 运算符	159
9.2.5 表结构设计实例	136	10.2.3 函数	159
9.3 创建数据表	137	10.2.4 查询中计算的设置	160
9.3.1 数据表创建的方法	137	10.3 创建选择查询	161
9.3.2 利用字段模板和输入数据创建表	138	10.3.1 使用向导创建选择查询	161
9.3.3 利用数据导入的方式创建表	141	10.3.2 使用设计视图创建选择查询	164
9.3.4 利用表设计器创建表	142	10.4 创建参数查询	166
9.4 创建表间关系	144	10.4.1 参数查询概述	166
9.4.1 表间关系的类型	144	10.4.2 参数查询条件的设置	166
9.4.2 表间关系的创建	144	10.4.3 使用设计视图创建参数查询	167
9.4.3 表间关系的维护	147	10.5 创建交叉表查询	170
9.5 表的维护	148	10.5.1 交叉表查询概述	170
9.5.1 表的复制与重命名	148	10.5.2 创建交叉表查询的方法	170
9.5.2 修改表结构	149	10.5.3 使用向导创建交叉表查询	171
9.5.3 表的删除	149	10.5.4 使用设计视图创建交叉表查询	172
9.6 表内容的操作	150	10.6 创建操作查询	173
9.6.1 记录的定位与选择	150	10.6.1 操作查询概述	173
9.6.2 记录的添加与删除	150	10.6.2 生成表查询的创建	174
		10.6.3 追加查询的创建	175

10.6.4 更新查询的创建	176	12.2.1 使用“报表”工具创建报表	220
10.6.5 删除查询的创建	178	12.2.2 使用“报表向导”创建报表	221
10.7 SQL 查询	179	12.2.3 使用“空报表”工具创建报表	223
10.7.1 QBE 查询与 SQL 查询	179	12.2.4 使用“报表设计”创建报表	224
10.7.2 SQL 聚合函数	180	12.2.5 使用“标签”向导创建报表	228
10.7.3 常用的 SQL 语句	180	12.3 报表的编辑	230
10.7.4 数据定义查询	183	12.3.1 设置报表的属性和格式	230
10.7.5 创建联合查询	184	12.3.2 设置报表的格式	231
10.8 创建嵌套查询	185	12.3.3 排序与分组	231
10.9 小结	186	12.3.4 使用计算和汇总	231
习题	187	12.3.5 加入分页符	232
上机练习题	187	12.3.6 添加页码和日期时间	233
第 11 章 窗体	188	12.3.7 报表中节的操作	233
11.1 窗体概述	188	12.4 创建子报表	234
11.1.1 窗体的类型	188	12.5 报表的打印与导出	237
11.1.2 窗体的视图	189	12.5.1 打印报表	237
11.1.3 窗体的组成	192	12.5.2 导出报表	237
11.1.4 窗体的属性	193	12.6 小结	238
11.2 窗体控件	193	习题	238
11.2.1 控件概述	193	上机练习题	238
11.2.2 窗体中的常用控件	194	第 13 章 宏	239
11.2.3 控件的使用	195	13.1 宏概述	239
11.3 创建窗体	199	13.1.1 宏的结构	239
11.3.1 使用“窗体”工具创建新窗体	200	13.1.2 宏生成器	239
11.3.2 使用向导创建窗体	201	13.1.3 宏的类型	242
11.3.3 使用“空白窗体”工具创建窗体	201	13.2 常用的宏操作	242
11.3.4 使用“其他窗体”工具创建窗体	202	13.3 创建宏	243
11.3.5 使用设计视图创建窗体	206	13.3.1 创建简单宏	243
11.4 创建主/子窗体	210	13.3.2 创建宏组	244
11.5 小结	215	13.3.3 创建条件宏	246
习题	215	13.4 宏的运行与调试	247
上机练习题	215	13.4.1 宏的运行	247
第 12 章 报表	217	13.4.2 宏的调试	248
12.1 报表概述	217	13.5 宏的应用	249
12.1.1 报表的组成	217	13.6 将宏转换为 VB 代码	252
12.1.2 报表的类型	219	13.7 小结	254
12.1.3 报表的视图	219	习题	254
12.2 创建报表	219	上机练习题	254
		第 14 章 模块与 VBA 编程	256
		14.1 模块概述	256

14.1.1 模块的组成	256	14.3.1 数据类型	269
14.1.2 模块的分类	259	14.3.2 常量与变量	269
14.1.3 模块和过程的创建方法	260	14.3.3 表达式	271
14.2 VBA 编程语言概述	261	14.3.4 数组	274
14.2.1 VB 与 VBA	261	14.3.5 控制结构	275
14.2.2 对象和对象名	262	14.4 模块编程实例	281
14.2.3 对象的属性	262	14.5 VBA 程序的调试	283
14.2.4 对象的方法	264	14.6 小结	286
14.2.5 对象的事件	267	习题	286
14.2.6 VBA 的编写环境	267	上机练习题	286
14.3 VBA 语法	268	参考文献	287

第1章 数据处理与数据库概述

计算机早期主要用于科学计算，数据类型单一。随着计算机技术的发展以及计算机的日益普及，计算机应用已经远远超出了这个范畴。如今面对着各种类型的海量数据，利用计算机做得更多的是进行数据处理。数据处理不仅广泛应用于电信、银行、证券、航空、教育、出版、气象等领域，而且在地质勘探测绘、仓库管理、技术情报管理、销售、制造、智能交通、电子商务等领域也呈现出勃勃生机。

数据处理离不开软件的支持，常用的数据处理软件包括：用于管理数据的文件系统和数据库管理系统，用于编写各种处理程序的高级程序设计语言及其编译、解释程序，以及各种数据处理方法的应用软件包等。

1.1 数据处理

数据和信息的关系非常密切，在多数情况下没有严格的区分。信息处理从根本上离不开数据，因此信息处理实质上就是数据处理。而数据处理的最终结果是以信息或知识的方式展示给用户，所以数据处理也称为信息处理。但在某些特定的环境下，数据和信息还是两个不同的概念，不能混用，如不能将数据文件说成信息文件等。

1.1.1 数据与信息

1. 数据

数据是对客观世界中各种事物的一种抽象、符号化的表示。它采用一种人为规定的符号来表示从现实世界中观察和收集到的现象和事实。数据的表现形式很多，可以是数字、文字、时间，也可以是图形、图像、动画、声音等多媒体形式。

从计算机的角度看，数据泛指可以被计算机接受并能被计算机处理的符号；从数据库的角度看，数据就是数据库中存储的基本对象。数据有型与值之分，数据的型给出了该数据所属数据类型的说明，如整型、字符型、布尔型等；数据的值给出了符合给定型的数值，值是型的一个实例。数据的型相当于程序设计语言中变量的类型说明，数据的值相当于变量的取值。数据的型基本上相对稳定，数据的值则是不断变化的。

2. 信息

信息源于拉丁文“Information”，是指一种陈述或一种解释、理解等。数据经过解释并赋予一定的含义之后，就成为了信息，即信息是根据需要对数据进行加工处理后得到的结果。

3. 数据与信息的关系

数据是信息的符号表示，是信息的具体表现形式，信息只有通过数据的形式表示出来才能被理解和接受。信息是数据的内涵，即数据的语义，信息在计算机中的存储即为数据。信息是观念上的，受制于人对客观事物变化规律的认知。例如，一个数字40，它的语义可能是年龄为40岁、体重为40公斤、价格为40元、考试成绩为40分、40个馒头、40本书等，也可能是高烧40度、水深40米、雨量达到40毫米、血压低压为40等。又比如，文字“黎明”，它的语义可能是一个词语，表示天快要亮或刚亮的时候，也可能是名称，如某人的姓名、壁画的名字、绘本的名字等。

数据要符合其语义，数据与其语义是不可分的。数据库系统要保证数据库中的数据符合其语义。

1.1.2 数据处理与数据管理

数据处理的发展及应用的广度和深度，极大地影响着人类社会发展的进程。数据处理，也称信

息处理,是将数据加工成信息的过程。具体指利用计算机对各种数据(包括数值和非数值的)进行收集、整理、存储、分类、排序、检索、维护、加工、统计、传输等一系列活动的总和。主要目的之一是从大量无序、难以理解的数据中,抽取并推导出有用的数据成分,作为行为和决策的依据。

数据处理贯穿于社会生产和生活的各个领域。通常,数据处理的计算方法和过程比较简单,但处理的数据量很大,数据结构复杂,因此,数据处理的重点不是计算,而是数据管理。

数据管理是数据处理的核心,主要功能包括数据的收集和分类、数据的表示和存储、数据的定位与查找、数据的维护和保护、提供数据访问接口和数据服务(如性能检测分析、可视化界面服务等)等。

数据处理与数据管理密切相关,数据管理技术的优劣直接影响数据处理的效果,数据库技术正是针对这一目标进行研究、发展并逐渐完善起来的专门技术。数据是数据库技术的研究目标,数据处理是数据库技术的应用方向,数据管理则是数据库技术研究的主要内容。

1.1.3 数据管理技术的发展

数据管理作为计算机应用领域中最大的一类应用,随着应用需求和计算机软硬件的发展,主要经历了人工管理、文件管理和数据库管理三个发展阶段。

1. 人工管理阶段

20世纪50年代中期之前,计算机主要用于科学计算。数据存储设备主要是卡片、纸带和磁带。没有操作系统和数据管理软件,数据需要人工管理。数据不保存,随用随丢。应用程序和数据不可分割,数据完全依赖于应用程序,不具有独立性,因而数据无法共享。该阶段应用程序与数据之间的对应关系如图1-1所示。

2. 文件管理阶段

20世纪50年代后期至60年代中期,计算机技术有了很大的发展,开始广泛应用于信息处理。数据存储设备主要是磁盘、磁鼓。磁盘是一种随机存取设备,允许用户直接访问数据,摆脱了磁带顺序访问的限制。该阶段出现了操作系统,并使用专门的管理软件,即文件系统(操作系统中的文件管理功能)来实施数据管理。数据可以长期保存在磁盘上,应用程序和数据有了一定的独立性,数据文件有了一定的共享性,但存在较大的数据冗余。该阶段应用程序与数据之间的对应关系如图1-2所示。

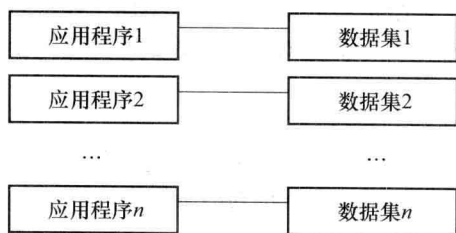


图 1-1 人工管理阶段应用程序与数据之间的对应关系

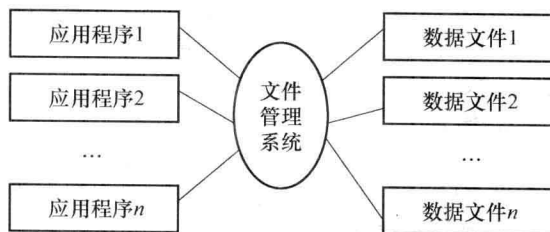


图 1-2 文件管理阶段应用程序与数据之间的对应关系

文件管理阶段的主要缺点体现在以下几个方面:

- 1) 数据冗余较大,相同的数据在多个文件中重复存储。
- 2) 程序和数据之间虽有一定的独立性,但是独立性较差,应用程序依赖于文件的结构。文件结构的每一次修改,都要修改相应的应用程序。
- 3) 不具备自动实现数据之间联系的功能,文件与文件之间相互独立,文件之间的联系必须通过应用程序来构造。
- 4) 数据冗余度高,容易造成数据的不一致。文件系统本身不具备维护数据一致性的功能,完全由用户负责维护。

5) 文件系统对于文件只提供了几个低级的文件操作命令, 文件的查询、修改需要通过编写应用程序来实现, 功能相同的操作很难共享。

3. 数据库管理阶段

20 世纪 60 年代后期, 随着应用需求、软硬件技术发展的日趋成熟, 计算机用于信息处理的规模越来越大, 对数据管理技术的要求也越来越高, 原有的文件系统已经不能胜任数据管理的任务。与此同时, 计算机网络系统和分布式系统的相继出现, 急需一种新的、能在多用户环境下进行数据共享和处理的数据管理软件。在这个背景下, 数据库管理系统 (Database Management System, DBMS) 应运而生。

在数据库管理阶段, 数据由 DBMS 统一管理和控制, 包括数据的安全性控制、数据的完整性控制、并发控制以及数据库恢复等; 实现整体数据的结构化, 数据结构使用数据模型来描述, 无需程序定义和解释。数据面向全系统, 可以被多个用户或应用程序共享, 提高了数据的共享性, 减少了数据冗余, 保证了数据的一致性和完备性。数据与应用程序相对独立, 减少了应用程序开发和维护的成本。在数据库管理阶段, 应用程序与数据之间的对应关系如图 1-3 所示。

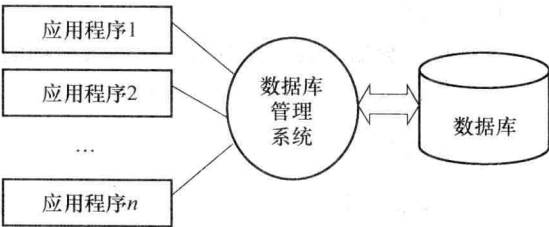


图 1-3 数据库管理阶段应用程序与数据之间的对应关系

1.2 数据库基本概念

本节主要介绍数据库中的几个基本概念, 包括: 数据库、数据库管理系统、数据库系统、数据库管理员以及数据字典。

1.2.1 数据库

数据库 (Database, DB) 是相互关联的数据集合。具体地说, 数据库是指长期存储在计算机内、有组织的、可共享的大量数据集合。数据库具有如下特征:

- 1) 在数据库中, 不仅能表示数据本身, 还要能表示数据与数据之间的联系。如教学管理数据库中有学生和课程两类数据, 学生数据通过选课与课程数据产生联系。
- 2) 数据通过一定的数据模型进行组织、描述和存储, 如层次模型、网状模型和关系模型等。
- 3) 具有较小的冗余度、较高的数据独立性和易扩展性。其中, 数据独立性是指数据的组织和存储方法不依赖于具体的应用程序。
- 4) 用户和应用程序可以共享数据。
- 5) 数据的各种操作, 如查询、修改、删除等都由 DBMS 统一管理。

1.2.2 数据库管理系统

上述数据库的特征蕴藏在一种专业软件中, 即数据库管理系统 (DBMS) 中。DBMS 是一种介于用户与操作系统之间, 专门用于数据管理的系统软件 (如图 1-4 所示), 它提供了有效建立、管理并安全持久地保存大量数据的能力。DBMS 是数据库系统的核心组成部分, 它完成“科学组织和存储数据, 高效获取和处理数据”的任务。用户通过 DBMS 访问数据库中的数据, 数据库管理员也通过 DBMS 进行数据库的维护工作。

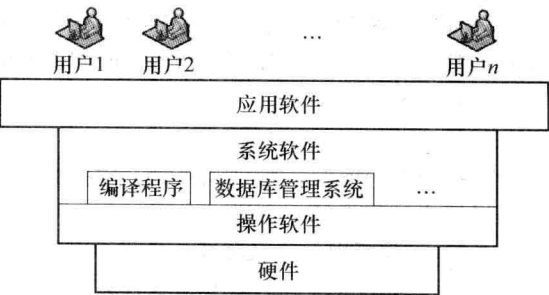


图 1-4 数据库管理系统在计算机系统的位置

DBMS 的主要功能包括:

- 1) 提供了数据定义语言 (Data Definition Language, DDL) 及其翻译程序, 供用户定义数据库的结构、定义数据库中的各种数据对象, 以及数据完整性和其他约束条件。
- 2) 提供了数据操纵语言 (Data Manipulation Language, DML), 供用户完成对数据库中数据的插入、修改、删除和查询操作。
- 3) 提供了数据控制语言 (Data Control Language, DCL), 可以对数据库进行统一管理和控制, 实现存取控制、安全控制、完整性控制, 以及多用户环境下的并发控制、事务管理和发生故障后的系统恢复等。
- 4) 提供数据库初始数据的装入、转换和卸出, 数据存储、备份和恢复, 以及数据库的性能分析和监测等数据维护功能。
- 5) 提供数据字典, 存放数据库各级结构的描述。
- 6) 提供数据库接口即访问数据库的方法, 供用户或应用程序访问数据库。

1.2.3 数据库系统

在具体应用中, 用户实际面对的是数据库系统, 而不是数据库或数据库管理系统。

1. 数据库系统的组成

数据库系统 (Database System, DBS) 就是基于数据库的计算机应用系统。DBS 由计算机软硬件、数据库 (包括物理数据库和描述数据库)、DBMS、数据库应用系统、数据库管理员以及用户组成。

其中, 数据库应用系统是实现业务逻辑的应用软件系统, 它是以 OS (Operating System, 操作系统)、DBMS、高级程序设计语言和实用程序为软件环境, 以某一领域的数据库管理需求为应用背景, 采用数据库概念和技术编写的一个可实际运行的软件。它提供一个友好的图形用户界面, 使用户可以方便地访问数据库中的数据。常见的数据库应用系统有信息管理系统、办公自动化系统、情报检索系统、高校教学管理系统、财务管理系统、商业交易系统。

2. 数据库系统的特点

(1) 数据的结构化

数据的结构化是指数据之间相互联系, 面向整个系统。具体地说就是, 数据库系统不仅描述数据本身, 而且还描述数据之间的联系, 实现了整体数据的结构化。数据的结构化也是数据库系统与文件系统的主要区别之一。

(2) 数据共享程度高、易扩充、冗余度低

共享是指数据可以被多个用户、多个应用程序同时使用。冗余度是指同一数据被重复存储的程度。共享程度高可以大大减少数据冗余, 节约存储空间, 避免数据的不一致性。此外, 由于数据库设计是面向系统而不是面向某个应用, 主要考虑的是数据的结构化, 因而容易扩充。

(3) 数据的独立性高

数据库系统的一个重要目标就是要使程序和数据真正分离, 使它们能独立发展。数据的独立性就是指数据独立于应用程序, 两者之间互不影响。数据的独立性包括逻辑独立性和物理独立性, 具体内容见 1.5 节。

(4) 数据控制能力较强

数据由 DBMS 统一管理和控制, 包括数据的并发控制、安全性控制、完整性控制以及故障恢复等, 详细内容将在第 6 章进行介绍。



重要提示 数据库、数据库系统、数据库管理系统是不同的三个概念, 在不引起混淆的情况下, 数据库通常作为数据库系统或数据库管理系统的简称。比如, 人们常说的数据库应用领域、数据库的三级模式结构, 实际上是指数据库系统的应用领域、数据库系统

的三级模式结构。又比如，甲骨文公司的 Oracle 是一款关系数据库管理系统产品，人们通常简称为 Oracle 数据库。学校里常用的教学管理数据库系统，也常常称为教学管理数据库或教学管理系统。在实际学习中，到底是指哪一个概念，要根据上下文环境来理解。

1.2.4 数据库管理员

数据库是共享资源，可以供多个用户和应用程序共享，需要统一规划、管理、协调、监控，同时数据库也是重要的信息资源，需要对数据库中数据的并发性、安全性、完整性以及故障恢复进行维护。数据库管理员（Database Administrator，DBA）就是从事上述工作的人员或机构。

DBA 是负责对数据库进行全面管理和控制的人员，负责完成整个系统的建立、控制、管理、维护和协调工作。DBA 的工作很繁重也很关键，所有与数据库有关的事宜一般均由 DBA 来决定。

DBA 可能是一名人员，也可能是一组人员。对于一个规模较大的数据库系统，DBA 通常指数据库管理机构或部门。

1.2.5 数据字典

数据字典中存放的是数据库系统中的一些说明信息，即数据的数据，称为元数据，如对各级模式的描述、索引、完整性约束、安全性要求、数据库的使用人员等说明信息。数据字典类似于日常生活中使用的字典或书的目录，能够帮助 DBMS 按照用户指定的数据对象名称，从中快速找到所需要的信息。数据字典在数据库系统的设计、实现、运行和维护各阶段起着非常重要的作用，是数据管理和控制的有力工具。

数据字典提供了对描述数据进行集中管理的手段，可以将它看成是数据库系统自身的小的、专门的数据库，区别于真正的物理数据库，常称为描述数据库或数据库的数据库。这种数据库只能由数据库系统本身来访问和修改。

1.3 数据视图

数据库系统的主要目的之一是通过抽象来屏蔽数据存储和维护细节，为用户提供数据的抽象视图，以简化用户与数据库系统的交互。

1.3.1 数据抽象

1. 三个世界

在数据处理的过程会涉及三个不同的世界：现实世界、信息世界和计算机世界。

现实世界是存在于人们头脑以外的客观世界，狭义上讲，现实世界就是客观存在的每个事物和现象。

信息世界是现实世界在人们头脑中的反映和解释，它将现实世界的事物用文字和符号记录下来，是现实世界的概念化。信息世界是对现实世界的抽象，并从现实世界中抽取能反映现实本质的概念和基本关系。信息世界作为现实世界通向计算机世界的桥梁，起着承上启下的作用。

计算机世界是对信息世界的进一步抽象，反映数据、数据特征和数据之间的联系，是信息世界的形式化和数据化。信息世界的信息在计算机世界中以数据的形式进行存储，有时将计算机世界也称为数据世界。

上述三个世界中的常用术语对照如表 1-1 所示。

表 1-1 三个世界中对应数据和数据特征的术语

现实世界	信息世界	计算机世界
个体	实体	记录或元组
特征	属性	字段或数据项

2. 抽象层次

抽象层次由高向低依次划分为：视图层、逻辑层和物理层。

1) 视图层：因为大多数用户并不需要访问数据库中的全部数据，因而，视图层仅描述整个数据库的部分数据，为用户提供屏蔽了数据类型等细节的一组应用程序。并且从安全性考虑，在视图层，系统定义了多个不同的视图使不同的用户访问不同的数据。

2) 逻辑层：程序设计人员在这个抽象层次上使用某种高级程序设计语言进行工作。逻辑层描述整个数据库所存储的数据以及数据之间的关系。DBA 通常在这个抽象层次上工作。

3) 物理层：描述数据的实际存储情况。DBA 可能需要了解某些数据物理结构的细节，而对于程序设计人员可能没有必要了解这些细节。数据库系统为数据库程序设计人员屏蔽了许多物理层的存储细节。

3. 视图

视图的本意是指一个人看到某个物体所得到的图像。物体有全局的概念，而视图具有局部的含义。将视图的概念引入数据库，数据库相当于一个全局的事物，每个用户从数据库中看到的数据就形成了视图。

图 1-5 给出了与三个抽象层次对应的视图：用户视图、概念视图和存储视图。用户视图也称外部视图，呈现的是数据库的局部结构。概念视图呈现的是数据库的全局结构。存储视图也称内部视图，呈现的是存储记录的物理顺序和彼此关联的方式。

视图	数据抽象
用户视图	视图层
概念视图	逻辑层
存储视图	物理层

图 1-5 数据抽象与视图

1.3.2 模式与实例

对数据库的描述或者说对包含在一个数据库中所有实体的描述定义，称为数据库模式，在多个抽象级别进行这些定义就形成了多个级别的数据库模式（或称数据库系统模式）。特定时刻存储在数据库中的数据集合称为数据库的一个实例。同一个模式可以有很多实例。

模式在数据库设计阶段就要确定下来，而且一般不会频繁修改，因而是相对稳定的，而实例是变化的，因为数据库中的数据总在不断更新。模式反映的是数据的结构及其联系，而实例反映的是数据库某一时刻的状态。关于模式将在 1.5 节进行详细介绍。

1.4 数据模型

数据模型是数据抽象的工具，根据数据抽象的 3 个层次，数据模型分为概念数据模型、逻辑数据模型和物理数据模型。物理数据模型是指具体的 DBMS 在实现其所支持的逻辑数据模型时，所用到的具体的物理存储结构。物理数据模型在数据库应用系统的开发中很少涉及，因此本节不做讨论。

1.4.1 数据模型的三要素

数据模型是对数据、数据的特征及数据之间联系的模拟、组织和抽象。数据模型是数据库系统的基础和核心，是数据库的框架，这个框架表示了信息及其联系的组织和表达方式，同时反映了存取路径，是对数据库如何组织的一种模型化表示。

数据模型有型和值的概念。型是对数据库中全体数据的逻辑结构和属性的描述，称为数据模式。值是一个具体值，称为实例。在同一数据模式下，可以有很多实例。



重要提示 通常情况下，对数据库模式、数据模式、数据库系统模式不做区分，本质是相同的，都属于型，是一种框架、结构，相对稳定，变化不大。它们只是针对