



高等学校**应用型特色**规划教材

二级Access 2010

与公共基础知识教程 (第2版)



韩相军 梁艳荣 主 编
韩勇华 乔建斌 副主编

赠 送
电子教案

- 以基础理论—实用技术—实训为主线
- 用任务来驱动，按教与学的实际需要取材谋篇
- 每一章都精心设置了上机实训和等级考试真题模拟，应用与应试并重
- 配备丰富的免费教学资源——电子教案、素材、等级考试系统与试题库



清华大学出版社

高等学校应用型特色规划教材

二级 Access 2010 与公共基础知识教程

(第 2 版)

韩相军 梁艳荣 主 编
韩勇华 乔建斌 副主编

清华大学出版社
北 京

内 容 简 介

本书根据教育部高等学校计算机科学与技术教学指导委员会发布的《关于进一步加强高等学校计算机基础教学的意见》，并紧密结合教育部考试中心制定的 2013 年版《全国计算机等级考试二级 Access 数据库程序设计考试大纲》而编写。书中由浅入深、循序渐进地介绍了二级 Access 2010 数据库与公共基础知识，内容丰富、结构清晰、语言简练、图文并茂，具有很强的实用性和可操作性。

本书既可作为高等院校非计算机专业学习 Access 2010 数据库课程的教材，也可作为全国计算机等级考试二级 Access 数据库程序设计的培训教材和参考用书。

本书封面贴有清华大学出版社防伪标签，无标签者不得销售。

版权所有，侵权必究。侵权举报电话：010-62782989 13701121933

图书在版编目(CIP)数据

二级 Access 2010 与公共基础知识教程/韩相军等主编. --2 版. --北京：清华大学出版社，2013
高等学校应用型特色规划教材

ISBN 978-7-302-34476-6

I. ①二… II. ①韩… III. ①关系数据库系统—高等学校—教材 IV. ①TP311.138

中国版本图书馆 CIP 数据核字(2013)第 270013 号

责任编辑：吴艳华 郑期彤

封面设计：杨玉兰

责任校对：周剑云

责任印制：何 芊

出版发行：清华大学出版社

网 址：<http://www.tup.com.cn>, <http://www.wqbook.com>

地 址：北京清华大学学研大厦 A 座 邮 编：100084

社 总 机：010-62770175 邮 购：010-62786544

投稿与读者服务：010-62776969, c-service@tup.tsinghua.edu.cn

质 量 反 馈：010-62772015, zhiliang@tup.tsinghua.edu.cn

课 件 下 载：<http://www.tup.com.cn>, 010-62791865

印 装 者：清华大学印刷厂

经 销：全国新华书店

开 本：185mm×260mm 印 张：30.75 字 数：748 千字

版 次：2009 年 9 月第 1 版 2013 年 11 月第 2 版 印 次：2013 年 11 月第 1 次印刷

印 数：1~3600

定 价：48.00 元

产品编号：053175-01

前 言

随着计算机与网络技术的飞速发展，数据库技术作为计算机应用的一个重要领域得到了广泛的应用与发展。数据库技术是现代信息科学与技术的重要组成部分，是计算机数据处理与信息管理的核心，掌握数据库知识已经成为各类科技人员和管理人员的基本要求。

Access 2010 关系型数据库管理系统是 Microsoft Office 系列应用软件的一个重要组成部分。它界面友好、功能全面且操作简单，不仅可以有效地组织与管理、共享与开发应用数据库信息，而且可以把数据库和程序设计相结合，使读者通过学习数据库与 VBA 编程能设计简单的数据库应用系统，实现理论与实践相结合。

本书从教学实际需求出发，由浅入深、循序渐进地讲解二级 Access 2010 数据库与公共基础知识。全书共分为 13 章，第 1~9 章介绍二级 Access 2010 数据库的基本概念和使用技术，第 10~13 章介绍公共基础知识，各章主要内容如下。

第 1 章介绍数据库基础知识，内容包括数据库系统、数据模型、关系数据库、数据库设计基础等。

第 2 章介绍 Access 2010 数据库，内容包括 Access 2010 数据库开发环境、数据库操作与管理以及安全性等。

第 3 章介绍表，内容包括表结构设计、表的创建、表的关系、表的操作和数据的导入与导出等。

第 4 章介绍查询，内容包括查询的种类与应用、查询的创建方法、查询条件、设置查询的计算、查询设计和 SQL 查询等。

第 5 章介绍窗体，内容包括窗体基础知识、创建窗体、设计与修饰窗体和定制系统控制窗体等。

第 6 章介绍报表，内容包括报表的基本知识、创建报表、报表排序和分组以及使用计算型控件等。

第 7 章介绍宏，内容包括宏的功能、宏的创建、运行与调试，通过事件触发宏等。

第 8 章介绍模块与 VBA 编程基础，内容包括模块的基本概念、创建模块、VBA 编程基础、VBA 流程控制语句、过程调用与参数传递、VBA 程序的运行错误处理与调试等。

第 9 章介绍 VBA 数据库编程，内容包括 VBA 的常见操作和数据库编程等。

第 10 章介绍数据结构与算法，内容包括算法与数据结构的基本概念、线性表及其顺序存储结构、栈和队列、线性链表、树与二叉树、查找和排序等。

第 11 章介绍程序设计基础，内容包括程序设计方法与风格、结构化程序设计和面向对象的程序设计等。

第 12 章介绍软件工程基础，内容包括软件工程基本概念、结构化分析方法、结构化设计方法、软件测试和程序的调试等。



第 13 章介绍数据库设计基础,内容包括数据库系统的基本概念、数据模型、关系代数和数据库设计与管理等。

附录 A 为 Access 常用函数。

附录 B 为窗体属性及其含义。

附录 C 为控件属性及其含义。

附录 D 为 Access 常用事件。

附录 E 为常用宏操作命令。

附录 F 为全国计算机等级考试二级 Access 数据库程序设计考试大纲。

附录 G 为习题参考答案。

本书内容丰富、条理清晰、图文并茂、易教易学。在讲解每个知识点时都配有相应的综合实例和练习,方便读者上机实践。本书每章都配有学习重点、习题和实验案例,以便于学生复习和练习。建议本书讲授 96 学时,实验课 42 学时,学生课后自主上机练习至少 50 学时。

本书由韩相军、梁艳荣任主编,韩勇华、乔建斌任副主编。具体编写分工为:第 1、2 章由梁艳荣编写,第 3、4 章由乔建斌编写,第 5 章由苗英恺编写,第 6 章由陈佳编写,第 7 章由张艳玲编写,第 8 章由韩勇华编写,第 9 章由丁同朝编写,第 10 章由王常策编写,第 11 章由谢颖丽编写,第 12、13 章由韩相军编写,全书由韩相军定稿。本书在编写过程中得到北京中医药大学东方学院田宜春院长的大力支持,东方学院计算机基础教研室的老师们也提出了许多宝贵的意见和建议,在此表示衷心感谢!

由于编者水平有限,书中错误与疏漏在所难免,敬请广大读者批评指正。邮箱:dfxyjsj@126.com。

编 者

目 录

第 1 章 数据库基础概述	1	2.4.1 设置数据库密码	41
1.1 数据库系统概述.....	1	2.4.2 压缩和修复数据库	42
1.1.1 数据库系统的组成.....	1	2.4.3 生成 ACCDE 文件.....	43
1.1.2 数据库系统的特点.....	3	2.4.4 备份数据库	44
1.1.3 高级数据库阶段.....	4	2.5 习题.....	45
1.2 数据模型.....	6	第 3 章 表	47
1.2.1 数据模型的概念.....	6	3.1 表结构设计.....	47
1.2.2 数据模型的组成要素.....	6	3.1.1 表结构设计概述	47
1.2.3 概念模型.....	7	3.1.2 数据表结构设计	54
1.2.4 常用的数据模型.....	8	3.2 创建表.....	55
1.2.5 数据库系统的体系结构.....	10	3.2.1 创建表的方法	55
1.3 关系数据库.....	12	3.2.2 使用数据表视图创建表	56
1.3.1 关系的一些术语.....	12	3.2.3 使用设计视图创建表	57
1.3.2 关系的特点.....	13	3.2.4 修改表的结构	59
1.3.3 关系模型.....	13	3.2.5 设置和取消表的主键	60
1.3.4 关系完整性.....	13	3.3 表之间的关系.....	61
1.4 数据库设计基础.....	15	3.3.1 表之间关系类型的确定	61
1.4.1 数据库设计步骤.....	15	3.3.2 建立表之间的关系	62
1.4.2 数据库设计过程.....	17	3.3.3 【编辑关系】对话框中的 关系选项	64
1.5 习题.....	22	3.3.4 修改表之间的关系	64
第 2 章 Access 2010 数据库	25	3.3.5 删除表之间的关系	65
2.1 Access 2010 数据库简介	25	3.3.6 子表	65
2.1.1 Access 发展简介	25	3.4 表的基本操作.....	66
2.1.2 Access 数据库的系统结构	25	3.4.1 打开和关闭表	66
2.2 Access 2010 数据库开发环境	28	3.4.2 修改表结构	66
2.2.1 Access 2010 数据库的安装、 启动与关闭.....	29	3.4.3 在表中添加记录	67
2.2.2 Access 2010 用户界面	30	3.4.4 在表中修改记录	70
2.3 数据库操作.....	35	3.4.5 在表中删除记录	70
2.3.1 创建数据库.....	35	3.4.6 表中记录排序	70
2.3.2 数据库的打开与关闭.....	40	3.4.7 表中记录筛选	72
2.4 数据库安全性.....	41	3.4.8 设置表的外观	75
		3.4.9 查找表中的数据	80



3.4.10 替换表中的数据.....	81	5.1.1 窗体的概念与作用	134
3.4.11 表的重命名.....	82	5.1.2 窗体构成	135
3.4.12 删除表.....	82	5.1.3 窗体类型	136
3.4.13 复制表.....	82	5.1.4 窗体视图	137
3.5 导入表、导出表与链接表.....	83	5.2 创建窗体.....	138
3.5.1 导入表.....	83	5.2.1 自动创建窗体	139
3.5.2 导出表.....	85	5.2.2 创建图表窗体	141
3.5.3 链接表.....	86	5.2.3 使用【空白窗体】按钮创建 窗体	143
3.6 习题.....	87	5.2.4 使用向导创建窗体	144
第4章 查询.....	94	5.3 设计窗体.....	147
4.1 查询的种类与应用.....	94	5.3.1 窗体设计视图的组成与主要 功能	147
4.1.1 查询的种类.....	94	5.3.2 常用控件的功能	149
4.1.2 查询的应用.....	95	5.3.3 常用控件的使用	152
4.2 查询的创建方法.....	96	5.3.4 窗体和控件的属性	159
4.2.1 使用查询向导.....	96	5.4 修饰窗体.....	166
4.2.2 使用查询设计视图.....	98	5.4.1 主题的应用	166
4.3 查询条件.....	100	5.4.2 条件格式的使用	167
4.3.1 表达式.....	100	5.4.3 提示信息的添加	168
4.3.2 表达式应用.....	105	5.4.4 窗体的布局	169
4.3.3 设置查询的组合条件.....	106	5.5 定制系统控制窗体.....	170
4.4 设置查询的计算.....	109	5.5.1 创建切换窗体	170
4.4.1 设置查询的总计计算.....	109	5.5.2 创建导航窗体	175
4.4.2 设置查询的分组总计计算.....	110	5.5.3 设置启动窗体	176
4.4.3 设置查询的自定义计算.....	112	5.6 习题.....	177
4.5 查询设计.....	113	第6章 报表.....	182
4.5.1 选择查询.....	113	6.1 报表的基本概念与组成.....	182
4.5.2 交叉表查询.....	114	6.1.1 报表的基本概念	182
4.5.3 参数查询.....	116	6.1.2 报表设计区	183
4.5.4 操作查询.....	117	6.2 创建报表.....	184
4.6 SQL 查询.....	120	6.2.1 使用【报表】按钮创建 报表	185
4.6.1 SQL 的数据定义.....	120	6.2.2 使用【报表设计】按钮创建 报表	186
4.6.2 SQL 的数据操纵	122	6.2.3 使用【空报表】按钮创建 报表	190
4.6.3 SQL 视图	127		
4.7 习题.....	128		
第5章 窗体.....	134		
5.1 窗体基础知识.....	134		

6.2.4 编辑报表.....	192	8.3.3 VB 编程环境.....	222
6.3 报表排序和分组.....	194	8.3.4 数据类型和数据库对象	227
6.3.1 记录排序.....	194	8.3.5 变量、常量与数组	229
6.3.2 记录分组.....	196	8.3.6 运算符与表达式	236
6.4 使用计算型控件.....	197	8.3.7 函数	241
6.4.1 报表添加计算型控件.....	197	8.4 VBA 流程控制语句	249
6.4.2 报表统计计算.....	198	8.4.1 声明语句	249
6.4.3 报表常用函数.....	199	8.4.2 赋值语句	249
6.5 习题.....	200	8.4.3 标号和 Goto 语句	250
第 7 章 宏	204	8.4.4 执行语句	250
7.1 宏的功能.....	204	8.5 过程调用与参数传递.....	258
7.1.1 宏的基本概念.....	204	8.5.1 过程调用	258
7.1.2 设置宏操作.....	205	8.5.2 参数传递	260
7.2 宏的创建、运行与调试.....	206	8.6 VBA 程序的运行错误处理与调试	262
7.2.1 创建独立宏.....	206	8.6.1 程序的运行错误处理	262
7.2.2 创建含子宏的独立宏.....	206	8.6.2 程序的调试	263
7.2.3 创建含有 If 块的独立宏	207	8.7 习题.....	265
7.2.4 创建嵌入宏.....	208	第 9 章 VBA 数据库编程	286
7.2.5 创建数据宏.....	210	9.1 VBA 常见操作.....	286
7.2.6 设置宏的操作参数.....	212	9.2 VBA 的数据库编程	297
7.2.7 运行宏.....	212	9.3 习题.....	315
7.2.8 调试宏.....	213	第 10 章 数据结构与算法.....	321
7.3 通过事件触发宏.....	213	10.1 算法.....	321
7.3.1 事件的概念.....	213	10.1.1 算法的基本概念	321
7.3.2 通过事件触发宏的示例.....	214	10.1.2 算法复杂度	325
7.4 习题.....	215	10.2 数据结构的基本概念.....	327
第 8 章 模块与 VBA 编程基础.....	218	10.2.1 什么是数据	328
8.1 模块的基本概念.....	218	10.2.2 数据结构的图形表示	333
8.1.1 类模块.....	218	10.2.3 线性结构与非线性结构	334
8.1.2 标准模块.....	218	10.3 线性表及其顺序存储结构.....	335
8.1.3 将宏转换为模块.....	219	10.3.1 线性表的基本概念	335
8.2 创建模块.....	219	10.3.2 线性表的顺序存储结构	336
8.3 VBA 编程基础	220	10.3.3 顺序表的插入运算	337
8.3.1 认识 VBA	220	10.3.4 顺序表的删除运算	338
8.3.2 面向对象程序设计的基本 概念.....	221	10.4 栈和队列.....	339
		10.4.1 栈及其基本运算	339



10.4.2 队列及其基本运算.....	341	12.1.4 软件工程的目标与原则	384
10.5 线性链表.....	344	12.1.5 软件开发工具与软件开发 环境	386
10.5.1 线性链表的基本概念.....	344	12.2 结构化分析方法.....	386
10.5.2 线性链表的基本运算.....	348	12.2.1 需求分析与需求分析方法 ...	386
10.5.3 循环链表及其基本运算.....	350	12.2.2 结构化分析方法	388
10.6 树与二叉树.....	351	12.2.3 软件需求规格说明书	392
10.6.1 树的基本概念.....	351	12.3 结构化设计方法.....	393
10.6.2 二叉树及其基本性质.....	353	12.3.1 软件设计的基本概念	393
10.6.3 二叉树的存储结构.....	356	12.3.2 概要设计	396
10.6.4 二叉树的遍历.....	357	12.3.3 详细设计	401
10.7 查找技术.....	359	12.4 软件测试.....	405
10.7.1 顺序查找.....	359	12.4.1 软件测试的目的	405
10.7.2 二分法查找.....	359	12.4.2 软件测试的准则	406
10.8 排序技术.....	360	12.4.3 软件测试技术和方法综述 ...	406
10.8.1 交换类排序法.....	360	12.4.4 软件测试的实施	413
10.8.2 插入类排序法.....	362	12.5 程序的调试.....	417
10.8.3 选择类排序法.....	364	12.5.1 基本概念	417
10.9 习题.....	366	12.5.2 软件调试方法	419
第 11 章 程序设计基础	368	12.6 习题.....	420
11.1 程序设计方法与风格.....	368	第 13 章 数据库设计基础	422
11.2 结构化程序设计.....	370	13.1 数据库系统的基本概念.....	422
11.2.1 结构化程序设计的原则.....	370	13.1.1 数据、数据库和数据库管理 系统	422
11.2.2 结构化程序的基本结构 与特点.....	370	13.1.2 数据库系统的发展	426
11.2.3 结构化程序设计原则 和方法的应用.....	372	13.1.3 数据库系统的基本特点	428
11.3 面向对象的程序设计.....	372	13.1.4 数据库系统的内部结构 体系	429
11.3.1 关于面向对象方法.....	372	13.2 数据模型.....	431
11.3.2 面向对象方法的基本概念....	375	13.2.1 数据模型的基本概念	431
11.4 习题.....	379	13.2.2 E-R 模型.....	432
第 12 章 软件工程基础	380	13.2.3 层次模型	436
12.1 软件工程的基本概念.....	380	13.2.4 网状模型	436
12.1.1 软件的定义与特点.....	380	13.2.5 关系模型	437
12.1.2 软件危机与软件工程.....	381	13.3 关系代数.....	440
12.1.3 软件工程过程与软件生命 周期.....	383	13.4 数据库设计与管理.....	446
		13.4.1 数据库设计概述	446

13.4.2 数据库设计的需求分析.....	447	附录 C 控件属性及其含义	463
13.4.3 数据库概念设计	448	附录 D 常用事件	465
13.4.4 数据库逻辑设计	451	附录 E 常用宏操作命令	468
13.4.5 数据库物理设计	453	附录 F 考试大纲	472
13.4.6 数据库管理	453	附录 G 习题参考答案	477
13.5 习题	454	参考文献	482
附录 A 常用函数	456		
附录 B 窗体属性及其含义	460		

第 1 章 数据库基础概述

本章从数据库的基础知识入手,为进一步学习与使用数据库打下必要的基础。通过本章的学习,读者应该掌握以下内容。

- 数据库系统的概念与组成。
- 数据模型。
- 关系数据库。
- 数据库设计方法与步骤。

1.1 数据库系统概述

数据库技术产生于 20 世纪 60 年代末,是数据管理的最新技术,也是计算机科学的重要分支。在当今信息社会中,信息已成为各个行业、部门的重要财富和资源,信息系统也越来越显示出它的重要性。数据库技术是信息系统的核心和基础,它的出现极大地促进了计算机应用向各行各业的渗透。从一般的小型事务处理到大型的信息系统,越来越多的领域开始采用数据库技术存储与处理其信息资源。数据库的建设规模、数据库信息量的大小和使用频度已成为衡量一个国家信息化程度的重要标志。

1.1.1 数据库系统的组成

数据(Data)是数据库中存储的基本对象。数据的种类很多,如文字、图形、图像和声音等都是数据。

数据可定义为描述事物的符号记录。数据有多种形式,它们均可以经过数字化后存储在计算机中。在描述事物的过程中,数据与其解释是密不可分的。

数据库(Database)是指长期存储在计算机内的、有组织的、可共享的数据集合。数据库中的数据是按一定的数据模型组织、描述和存储的,具有较小的冗余度、较高的数据独立性和易扩展性,并且可以被多个用户、多个应用程序共享。

数据库管理系统(Database Management System, DBMS)是位于用户与操作系统(Operating System, OS)之间的一层数据管理软件,是数据库系统的枢纽。数据库管理系统能科学地组织和存储数据,高效地获取和维护数据。用户对数据库进行的各种操作,如数据库的建立、使用和维护,都是在 DBMS 的统一管理和控制下进行的。

数据库管理系统的主要功能有以下几个方面。

(1) 数据定义功能。提供数据定义语言(Data Definition Language, DDL),用于定义数据库中的数据对象。

(2) 数据操纵功能。提供数据操纵语言(Data Manipulation Language, DML),用于操纵数据,实现对数据库的基本操作,如查询、插入、删除和修改等。

(3) 数据库的运行管理。保证数据的安全性、完整性、多用户对数据的并发使用及发



生故障后的系统恢复。

(4) 数据库的建立和维护功能。提供数据库数据输入、批量装载、数据库转储、介质故障恢复、数据库的重组织及性能监视等功能。

数据库系统(Database System, DBS)是指在计算机系统中引入数据库之后组成的系统,是用来组织和存取大量数据的管理系统。数据库系统是由计算机系统(硬件和软件系统)、数据库、数据库管理系统、数据库管理员和用户组成的具有高度组织性的整体。

通常情况下,把数据库系统简称为数据库。数据库系统组件之间的关系如图 1.1 所示。

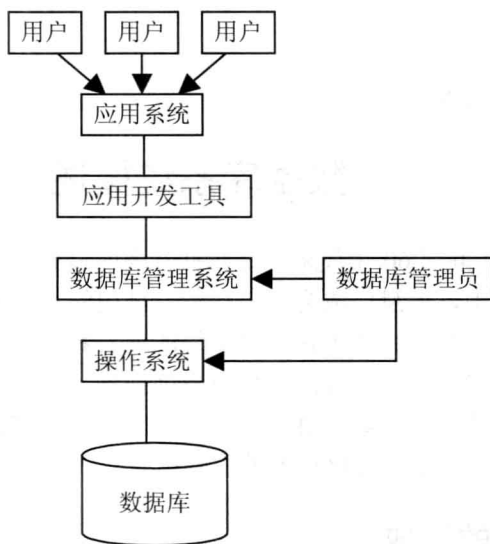


图 1.1 数据库系统的组成

数据库技术的核心任务是数据处理。数据处理是指对各种数据进行收集、存储、加工和传播等一系列活动的总和。数据管理则是指对数据进行分类、组织、编码、存储、检索和维护,它是数据处理的中心问题。

数据管理技术的发展,与计算机硬件(主要是外部存储器)、系统软件及计算机应用的范围有着密切的联系。数据管理技术的发展经历了以下几个阶段:人工管理、文件系统、数据库系统、分布式数据库系统、面向对象数据库系统、数据仓库和数据挖掘阶段。

20 世纪 50 年代中期以前,计算机主要用于科学计算。那时在计算机硬件方面,外存只有卡片、纸带及磁带,没有磁盘等可直接存取的存储设备;软件方面,只有汇编语言,没有操作系统和高级语言,更没有管理数据的软件;数据处理的方式是批处理。这些决定了当时的数据管理只能依赖于人工来进行。

人工管理阶段的数据管理技术具有以下特点。

(1) 数据不进行保存。当时的计算机主要用于科学计算,一个程序对应一个数据。在计算某一问题时,把程序和对应的数据装入,计算完后就退出,没有将数据长期保存的必要。

(2) 没有专门的数据管理软件。数据需要由应用程序管理,因此应用程序的设计者不仅要考虑数据的逻辑结构,还要考虑数据的物理结构,如存储结构、存取方法、输入/输出方式等。存储结构发生变化,应用程序也要作相应的修改,程序员的负担非常重,数据的

独立性也很差。

(3) 数据面向应用。一组数据对应一组程序,倘若多个程序使用相同的数据,必须各自定义,不能共享,所以程序之间存在大量的数据冗余。

(4) 只有程序的概念,基本上没有文件的概念。

20 世纪 60 年代,计算机技术水平有了很大的提高,计算机的应用范围不断扩大,不仅用于科学计算,还大量用于管理。这时的计算机硬件已经有了磁盘、磁鼓等直接存取的外存设备;软件则有了操作系统、高级语言,操作系统中的文件系统专门用于数据管理;处理方式不仅有批处理,还增加了联机实时处理。

文件系统阶段的数据管理技术具有以下特点。

(1) 数据可以长期保存在磁盘上。用户可以反复地对文件进行查询、修改、插入和删除等操作。

(2) 文件系统提供了数据与程序之间的存取方法。应用程序和数据有了一定的独立性,数据存储结构的改变也不一定反映在程序上,大大减轻了程序员的负担。

(3) 数据冗余量大。文件系统中,文件仍然是面向应用的,一个文件基本上对应于一个应用程序,不能对数据进行共享,因此数据冗余大,存储空间浪费。由于数据可能有多个副本,对其中之一进行修改时容易造成数据的不一致。

(4) 文件之间缺乏联系,相对孤立,仍然不能反映客观世界各个事物之间错综复杂的联系。

1.1.2 数据库系统的特点

20 世纪 60 年代末以来,计算机的应用更为广泛,用于数据管理的应用系统规模也更为庞大,由此带来数据量的急剧膨胀;计算机磁盘技术有了很大的发展,出现了大容量的磁盘;在处理方式上,联机实时处理的要求更多。这些变化促使了数据管理手段的进步,数据库技术应运而生。与人工管理和文件系统相比,数据库系统的特点主要有以下几个方面。

1. 数据的结构化

在文件系统中,只考虑了同一文件记录内部数据项之间的联系,而不同文件的记录之间是没有联系的,从整体上看数据是无结构的,不能反映客观世界各种事物之间错综复杂的联系。在数据库系统中,实现了整体数据的结构化,把文件系统中简单的记录结构变成了记录和记录之间的联系所构成的结构化数据。在描述数据时,不仅要描述数据本身,还要描述数据之间的联系。

2. 数据的共享性

数据库系统从整体角度看待与描述数据,使数据不再面向某个应用,而是面向整个系统。这些数据可以供多个部门使用,实现了数据的共享。各个部门的数据基本上没有重复的存储,数据的冗余量较小。



3. 数据的独立性

数据库系统有三层结构：用户(局部)数据的逻辑结构、整体数据的逻辑结构和数据的物理结构。在这三层结构之间，数据库系统提供了两层映像功能。首先是用户数据逻辑结构和整体数据逻辑结构之间的映像，这一映像保证了数据的逻辑独立性，当数据库的整体逻辑结构发生变化时，通过修改这层映像可使局部的逻辑结构不受影响，因此不必修改应用程序。其次是整体数据逻辑结构和数据物理结构之间的映像，它保证了数据的物理独立性，当数据的存储结构发生变化时，通过修改这层映像可使数据的逻辑结构不受影响，因此应用程序同样不必修改。

4. 数据的存储粒度

在文件系统中，数据存储的最小单位是记录；而在数据库系统中，数据存储的粒度可以小到记录中的一个数据项。因此数据库中数据存取的方式非常灵活，便于对数据的管理。

5. DBMS 对数据进行统一的管理和控制

DBMS 不仅具有基本的数据库管理功能，还具有如下控制功能。

(1) 保证数据的完整性。数据的完整性是指数据的正确性、有效性和相容性，要求数据在一定的取值范围内或相互之间满足一定的关系。例如，规定考试的成绩在 0~100 分之间，血型只能是 A、B、AB、O 中的一种，等等。

(2) 保证数据的安全性。让每个用户只能按指定的权限访问数据，防止不合法地使用数据，造成数据的破坏和丢失。例如，学生对于课程的成绩只能进行查询，不能修改。

(3) 并发控制。对多用户的并发操作加以协调和控制，防止多个进程同时存取、修改数据库中的数据时发生冲突，造成错误。例如，在学生选课系统中，某门课只剩下最后一个名额，但有两个学生在两台选课终端上同时发出了选这门课的请求，此时必须采取某种措施，以确保两名学生不能同时拥有这最后的一个名额。

(4) 数据库的恢复。当数据库系统出现硬件或软件故障时，DBMS 具有把数据库恢复到最近某个时刻的正确状态的能力。

6. 为用户提供了友好的接口

用户可以使用交互式的命令语言，如 SQL(Structured Query Language, 结构化查询语言)，对数据库进行操作；也可以把普通的高级语言(如 C++语言等)和 SQL 结合起来，从而把对数据库的访问和对数据的处理有机地结合在一起。

1.1.3 高级数据库阶段

1. 分布式数据库系统

20 世纪 70 年代后期之前，数据库系统多数是集中式的。分布式数据库系统是数据库技术和计算机网络技术相结合的产物，在 20 世纪 80 年代中期已有商品化产品问世。分布式数据库是一个逻辑上统一、地域上分布的数据集合，是计算机网络环境中各个节点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。

分布式数据库在逻辑上像一个集中式数据库系统,实际上数据存储处于不同地点的计算机网络的各个节点上。每个节点都有一定的局部数据库管理系统,使每个节点具有很高的独立性。用户可以由分布式数据库管理系统(网络数据库管理系统),通过网络通信相互传输数据。分布式数据库系统具有高度透明性,每台计算机上的用户不需要了解他所访问的数据究竟在什么地方,就像在使用集中式数据库一样。其主要优点如下。

(1) 局部自主。网络上每个节点的数据库系统都具有独立处理本地事务的能力,而且各局部节点之间也能够互相访问、有效地配合处理更复杂的事务。因此,分布式数据库系统特别适合各个部门的地理位置分散的组织机构,如银行业务、飞机订票和企业管理等。

(2) 可靠性和可用性。分布式系统比集中式系统具有更高的可靠性,在个别节点或个别通信链路发生故障的情况下可以继续工作。一个局部系统发生故障不至于导致整个系统停顿或破坏,只要有一个节点上的数据备份可用,数据就是可用的。

(3) 效率和灵活性。分布式系统分散了工作负荷,缓解了单机容量的压力。数据可以存储在邻近的常用节点上,如果本节点的数据子集包含要查询的全部内容,显然会比集中式数据库在全集上查找节省时间。

(4) 系统易于实现扩展。例如,一个单位要增加新的机构,分布式数据库系统能够在对现有系统影响较小的情况下实现扩充。因此,扩大系统规模比集中式系统更加方便、经济和灵活。

2. 数据库的发展方向

目前,数据库技术的发展方向主要有两个:一是改造和扩充关系数据库,以适应新的应用要求;二是改用新的数据库模型。目前,这两个方面都取得了很大的发展,主要出现了以下几项新的数据库技术。

1) 并行数据库技术

并行数据库技术包括对数据库的分区管理和并行查询。它通过将数据库任务分割成多个子任务的方法由多个处理机协同完成这个任务,从而可极大地提高事务处理能力,并且通过数据分区可以实现数据的并行 I/O 操作。

2) 数据仓库和数据挖掘技术

所谓数据仓库(Data Warehouse, DW),就是按决策目标将传统的事务型数据库中的数据重新组织划分,由此组成一种面向主题的、集成的、稳定的及随时间发展的数据集合。数据仓库与传统数据库的区别在于存储的数据容量大,存储的数据时间跨度大,存储的数据来源复杂,可用于企业与组织的决策分析处理等。所谓数据挖掘(Data Mining, DM),就是从大型数据库或数据仓库的数据中提取人们感兴趣的、隐含的、事先未知的、潜在的知识。数据挖掘方法的提出使人们有能力从过去若干年时间里积累的、海量的、以不同的形式存储的、十分繁杂的数据资料中认识数据的真正价值。目前,数据挖掘的研究已与数据库的研究结合起来。

3) 多媒体数据库

人们通常把能够管理数值、文本、图形、图像和声音等媒体类型的数据库称为多媒体数据库(Multimedia Database)。与传统 DBMS 一样,多媒体 DBMS 也要进行数据的处理、查询和事务的管理等,但是多媒体数据库有不同的用户接口和存储构造,在多媒体 DBMS



中特别强调“媒体独立性”(所谓媒体独立性是指不论管理多媒体数据的媒体如何变化,都不需改变 DBMS)。

4) 模糊数据库

传统的数据库仅允许对精确的数据进行存储和处理,而客观世界中许多事物是不精确的。模糊数据库技术的研究和实践就是为了解决模糊数据的表达和处理问题,使得数据库描述的模型更自然、更贴切地反映客观世界。

5) 网络数据库

网络数据库是数据库技术与 Web 技术相互融合的技术。

此外,还有其他一些新的数据库技术,如模糊演绎数据库、主动数据库、集中式工程数据库和面向对象数据库等。

1.2 数据模型

模型是现实世界特征的模拟和抽象。数据模型也是一种模型,只不过它模拟的对象是数据。根据模型应用的不同层次和目的,可以将模型分为两类,一类是概念模型,按用户的观点来对数据和信息建模,主要用于数据库设计;另一类是数据模型,主要包括网状模型、层次模型和关系模型等,它是按计算机系统的观点对数据建模。

1.2.1 数据模型的概念

数据模型是现实世界数据特征的抽象。数据模型是工具,是用来抽象、表示和处理现实世界中的数据和信息的工具。在数据库中用数据模型这个工具来抽象、表示和处理现实世界的数据和信息,现有数据库系统均是基于某种数据模型的。

数据模型应满足以下 3 个方面的要求。

- (1) 能够比较真实地模拟现实世界。
- (2) 容易被人理解。
- (3) 便于在计算机系统中实现。

1.2.2 数据模型的组成要素

数据模型是由数据结构、数据操作和数据的约束条件 3 个部分组成的。

(1) 数据结构是所研究对象的集合,这些对象是数据库的组成成分,如表中的字段、名称等。数据结构分为两类,一类是与数据类型、内容、性质有关的对象;另一类是与数据之间联系有关的对象。

(2) 数据操作是指对数据库中各种对象的实例(值)允许执行的操作的集合,包括操作及有关的操作规则。数据库的操作主要有检索和更新两大类。数据模型必须定义数据操作的确切含义、操作符号、操作规则以及实现操作的语言。

(3) 数据的约束条件是一组完整性规则的集合。完整性规则是给定的数据模型中数据及其联系所具有的制约和依存规则,用以限定符合数据模型的数据库状态以及状态的变化,以保证数据的正确、有效和相容。

数据模型给出了在计算机系统上描述和动态模拟现实数据及其变化的一种抽象方法,数据模型不同,描述和实现方法亦不相同,相应的支持软件,即数据库管理系统也就不同。

严格地讲,一个数据模型应由上述三部分组成,但数据结构是数据模型的本质标志。

1.2.3 概念模型

概念模型是现实世界到信息世界的第一层抽象,是现实世界到计算机的一个中间层次。概念模型是数据库设计的有力工具和数据库设计人员与用户之间进行交流的语言。它必须具有较强的语义表达能力,能够方便、直接地表达应用中的各种语义知识,且简单、清晰,易于被用户理解。

在现实世界中,事物之间的联系是客观存在的。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。概念世界不是现实世界的简单录像,而是把现实世界中的客观对象抽象为某一种信任结构,这种信任结构不是某一个 DBMS 支持的数据模型,而是概念模型。

建立概念模型涉及以下几个术语。

1) 实体

客观存在并可相互区别的事物称为实体(Entity)。实体可以是实际事物,也可以是抽象事件。例如,一个职工、一个部门属于实际事物;一次订货、借阅若干本图书、一场演出是比较抽象的事件。

同一类实体的集合称为实体集。例如,全体学生的集合、全馆图书等。用命名的实体型表示抽象的实体集。例如,实体型“学生”表示全体学生的概念,并不具体指学生甲或学生乙。

2) 属性

描述实体的特性称为属性(Attribute)。例如,学生实体用若干个属性(学生编号、姓名、性别、出生日期、籍贯等)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。

3) 关键字

如果某个属性或属性组合能够唯一地标识出实体集中的各个实体,可以将其选作关键字(Key),也称为码。

4) 联系

实体集之间的对应关系称为联系(Relationship),它反映了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;另一种是实体之间的联系。

5) E-R 图

概念模型的表示方法有很多,常用实体短横线联系方法(E-R 方法或 E-R 图)来描述现实世界的概念模型。E-R 方法也称为 E-R 模型。

E-R 图有以下 3 个要素。

(1) 实体:用矩形并在框内标注实体名称来表示。

(2) 属性:用椭圆形表示,并用连线将其与相应的实体连接起来。

(3) 联系:用菱形表示,菱形框内写明联系名,并用连线分别与有关实体连接起来,同时在连线上标上联系的类型(1:1、1:n 或 m:n)。图 1.2 所示为 E-R 图的示例。