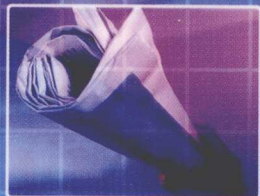




高等学校**应用型特色**规划教材

二级Access 2003

与公共基础知识教程



韩相军 田宜春 主 编
梁艳荣 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 2003 与公共基础知识教程

韩相军 田宜春 主 编

梁艳荣 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订的2008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》编写；本书由浅入深、循序渐进地介绍了二级 Access 2003 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 2003 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 2003 与公共基础知识教程/韩相军，田宜春主编；梁艳荣，乔建斌副主编. —北京：清华大学出版社，2009.11

(高等学校应用型特色规划教材)

ISBN 978-7-302-21185-3

I. 二… II. ①韩… ②田… ③梁… ④乔… III. ①关系数据库—数据库管理系统，Access 2003—高等学校—教材②电子计算机—高等学校—教材 IV. TP3

中国版本图书馆 CIP 数据核字(2009)第 171917 号

责任编辑：彭 欣

装帧设计：杨玉兰

责任校对：王 晖

责任印制：王秀菊

出版发行：清华大学出版社

地 址：北京清华大学学研大厦 A 座

<http://www.tup.com.cn>

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

印 刷 者：北京市世界知识印刷厂

装 订 者：三河市金元印装有限公司

经 销：全国新华书店

开 本：185×260 印 张：29.5 字 数：713 千字

版 次：2009 年 11 月第 1 版 印 次：2009 年 12 月第 2 次印刷

印 数：4001~7000

定 价：43.00 元

本书如存在文字不清、漏印、缺页、倒页、脱页等印装质量问题，请与清华大学出版社出版部联系调换。联系电话：(010)62770177 转 3103 产品编号：033196-01

前 言

随着计算机与网络技术的飞速发展,作为计算机应用的一个重要领域——数据库技术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系统核心,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 2003 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 2003 与公共基础知识。全书共分为 14 章,第 1~10 章介绍二级 Access 2003 数据库的基本概念和使用技术,第 11~14 章介绍公共基础知识,主要内容如下。

第 1 章介绍数据库基础知识,包括数据库基础知识、关系数据库、数据库设计基础和 Access 2003 简介等。

第 2 章介绍 Access 2003 数据库,包括 Access 2003 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,包括创建表、管理与维护表、操作表和数据的导入与导出等。

第 4 章介绍查询,包括查询概述、创建选择查询、创建交叉表查询、创建参数查询、创建操作查询、创建 SQL 查询和查询设计等。

第 5 章介绍窗体,包括认识窗体、创建窗体、自定义窗体和格式化窗体等。

第 6 章介绍报表,包括报表的基本概念、创建报表、编辑报表、复杂报表设计和打印报表等。

第 7 章介绍数据访问页,包括数据访问页的基本概念、创建数据访问页和编辑数据访问页等。

第 8 章介绍宏,包括宏的功能、创建宏和通过事件触发宏等。

第 9 章介绍模块与 VBA 编程基础,包括模块的基本概念、创建模块、VBA 程序设计基础、VBA 流程控制语句、过程调用与参数传递、VBA 程序运行错误处理和程序调试等。

第 10 章介绍 VBA 数据库编程,包括 VBA 的常见操作和数据库编程等。

第 11 章介绍数据结构与算法,包括算法、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、树和二叉树、查找和排序等。

第 12 章介绍程序设计基础,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 13 章介绍软件工程基础,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 14 章介绍数据库设计基础,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 Access 常用事件。

附录 C 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 D 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。

本书是集体智慧的结晶,第 1、2、9 和 10 章由梁艳荣编写,第 3 章由乔建斌编写,第 4 章由苗英凯编写,第 5 章由陈佳编写,第 6 章由张艳玲编写,第 7 章由谢颖丽编写,第 8 章由丁同朝编写,第 11 章由韩勇华编写,第 12~14 章由韩相军编写,全书由韩相军定稿。本书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,敬请读者批评指正。邮箱:dfxyjsj@126.com。

编 者



目 录

第 1 章 数据库系统的基础知识..... 1	2.2.2 数据库管理36
1.1 数据库系统概述..... 1	2.3 Access 2003 数据库安全性37
1.1.1 数据库系统的组成..... 1	2.3.1 设置数据库密码37
1.1.2 数据库系统的特点..... 3	2.3.2 用户级安全机制39
1.1.3 高级数据库阶段..... 5	2.3.3 使用权限40
1.2 数据模型..... 6	2.4 安全机制管理40
1.2.1 数据模型的概念..... 6	2.4.1 增加账户40
1.2.2 数据模型的组成要素..... 7	2.4.2 删除账户41
1.2.3 概念模型..... 7	2.4.3 更改账户权限41
1.2.4 常用的数据模型..... 9	2.4.4 打印账户和组账户列表.....41
1.2.5 数据库的体系结构..... 11	2.5 加密数据库42
1.3 关系数据库..... 12	2.6 习题42
1.3.1 关系的一些术语..... 12	第 3 章 表.....45
1.3.2 关系模型..... 13	3.1 创建表45
1.3.3 关系运算..... 13	3.1.1 数据表结构设计45
1.3.4 关系完整性..... 15	3.1.2 主码52
1.4 数据库设计基础..... 16	3.1.3 索引53
1.4.1 数据库设计步骤..... 17	3.1.4 使用向导创建表54
1.4.2 数据库设计过程..... 18	3.1.5 通过输入数据 创建表54
1.5 Access 2003 数据库简介..... 23	3.2 管理与维护表55
1.5.1 Access 的发展过程..... 23	3.2.1 表间关系的建立55
1.5.2 Access 2003 新增功能..... 24	3.2.2 修改表结构57
1.6 习题..... 26	3.3 操作表58
第 2 章 Access 2003 数据库..... 29	3.3.1 数据输入58
2.1 Access 2003 数据库开发环境..... 29	3.3.2 排序记录60
2.1.1 Access 2003 数据库的安装、 启动与关闭..... 29	3.3.3 筛选记录61
2.1.2 Access 2003 数据库构成..... 30	3.3.4 设置数据表格式62
2.1.3 Access 2003 数据库操作 环境..... 33	3.4 数据的导入与导出63
2.2 Access 2003 数据库操作与管理..... 34	3.4.1 导入、导出数据63
2.2.1 创建数据库..... 35	3.4.2 链接数据64
	3.5 习题65



第4章 查询	72	5.3.3 窗体与对象的属性及 设置方法	121
4.1 查询的种类与应用	72	5.3.4 窗体与对象的事件	127
4.1.1 查询的种类	72	5.3.5 常用控件的创建方法	130
4.1.2 查询的应用	73	5.4 窗体与控件的其他应用设计	138
4.2 查询的建立方法	74	5.4.1 创建计算控件	138
4.2.1 使用查询向导	74	5.4.2 查找记录	140
4.2.2 使用查询设计器	77	5.4.3 显示提示信息	141
4.3 查询条件	79	5.4.4 创建与使用主/子窗体	141
4.3.1 表达式	79	5.4.5 打印与预览窗体	142
4.3.2 标准函数	82	5.5 窗体外观格式设计	143
4.4 查询设计	84	5.5.1 加线条	143
4.4.1 条件选择查询	84	5.5.2 加矩形	143
4.4.2 交叉表查询	85	5.5.3 设置控件格式属性	144
4.4.3 参数查询	88	5.5.4 使用 Tab 键设置控件次序	146
4.4.4 操作查询	89	5.6 习题	146
4.4.5 重复项、不匹配项查询	92	第6章 报表	151
4.5 SQL 查询	92	6.1 报表的作用与组成	151
4.5.1 SQL 的数据定义	93	6.1.1 报表的作用	151
4.5.2 SQL 的数据操纵	94	6.1.2 报表的组成	151
4.5.3 SQL 视图	99	6.1.3 报表的视图	152
4.6 习题	100	6.2 报表的分类	153
第5章 窗体	107	6.2.1 纵栏式报表	153
5.1 窗体基础知识	107	6.2.2 表格式报表	153
5.1.1 窗体的概念与作用	107	6.2.3 图表报表	154
5.1.2 窗体构成	108	6.2.4 标签报表	154
5.1.3 窗体类型	108	6.3 创建报表	154
5.1.4 窗体视图	109	6.3.1 报表向导	155
5.2 创建窗体	110	6.3.2 自动创建报表	155
5.2.1 自动创建窗体	110	6.3.3 创建子报表	156
5.2.2 窗体向导	111	6.3.4 创建多列报表	156
5.2.3 图表向导	113	6.3.5 将窗体转换为报表	157
5.2.4 自动窗体: 数据透视表	114	6.3.6 自定义报表	157
5.2.5 自动窗体: 数据透视图	115	6.4 报表编辑	157
5.3 自定义窗体	116	6.4.1 报表控件及格式设计	158
5.3.1 窗体设计视图的组成与 主要功能	117	6.4.2 排序与分组	161
5.3.2 属性、事件与方法	120	6.4.3 在报表中应用计算	164

6.5 复杂报表设计	166	8.3.1 事件的概念	194
6.5.1 报表属性	166	8.3.2 通过事件触发宏	194
6.5.2 节属性	167	8.4 习题	195
6.5.3 报表加页分割	167	第 9 章 模块与 VBA 编程基础	198
6.5.4 使用报表的预定义格式	168	9.1 模块的基本概念	198
6.5.5 制作图表	168	9.1.1 类模块	198
6.6 打印报表	169	9.1.2 标准模块	199
6.6.1 报表预览	169	9.1.3 将宏转换为模块	199
6.6.2 打印与保存报表	169	9.2 创建模块	199
6.7 习题	170	9.3 VBA 编程基础	200
第 7 章 数据访问页	175	9.3.1 认识 VBA	200
7.1 数据访问页的基本概念	175	9.3.2 面向对象程序设计的基本概念	201
7.1.1 页视图	175	9.3.3 VB 编程环境	202
7.1.2 设计视图	176	9.3.4 数据类型和数据库对象	207
7.2 创建数据访问页	177	9.3.5 变量、常量与数组	210
7.2.1 自动创建数据访问页	177	9.3.6 运算符与表达式	217
7.2.2 使用向导创建数据访问页	178	9.3.7 函数	222
7.2.3 使用设计视图创建数据访问页	180	9.4 VBA 流程控制语句	229
7.3 编辑数据访问页	181	9.4.1 声明语句	230
7.3.1 添加标签	181	9.4.2 赋值语句	230
7.3.2 添加命令按钮	181	9.4.3 标号和 Goto 语句	231
7.3.3 添加滚动文字	182	9.4.4 执行语句	231
7.3.4 设置背景	182	9.5 过程调用与参数传递	238
7.4 习题	183	9.5.1 过程调用	238
第 8 章 宏	186	9.5.2 参数传递	240
8.1 宏的功能	186	9.6 VBA 程序的运行错误处理与调试	241
8.1.1 宏的基本概念	186	9.6.1 程序的运行错误处理	241
8.1.2 设置宏操作	187	9.6.2 程序的调试	243
8.2 宏的创建、调试与运行	190	9.7 习题	245
8.2.1 创建操作序列宏	190	第 10 章 VBA 数据库编程	266
8.2.2 创建宏组	191	10.1 VBA 常见操作	266
8.2.3 创建条件操作宏	191	10.2 VBA 的数据库编程	278
8.2.4 宏调试	192	10.3 习题	297
8.2.5 运行宏	193	第 11 章 数据结构与算法	302
8.3 通过事件触发宏	194	11.1 算法	302



11.1.1	算法的基本概念	302	12.2.3	结构化程序设计原则和 方法的应用	355
11.1.2	算法复杂度	306	12.3	面向对象的程序设计	355
11.2	数据结构的基本概念	309	12.3.1	关于面向对象方法	355
11.2.1	什么是数据	309	12.3.2	面向对象方法的 基本概念	358
11.2.2	数据结构的图形表示	314	12.4	习题	362
11.2.3	线性结构与非线性结构	315	第 13 章	软件工程基础	364
11.3	线性表及其顺序存储结构	316	13.1	软件工程基本概念	364
11.3.1	线性表的基本概念	316	13.1.1	软件定义与软件特点	364
11.3.2	线性表的顺序存储结构	317	13.1.2	软件危机与软件工程	365
11.3.3	顺序表的插入运算	319	13.1.3	软件工程过程与软件 生命周期	367
11.3.4	顺序表的删除运算	320	13.1.4	软件工程的目标与 原则	368
11.4	栈和队列	321	13.1.5	软件开发工具与软件 开发环境	370
11.4.1	栈及其基本运算	321	13.2	结构化分析方法	370
11.4.2	队列及其基本运算	323	13.2.1	需求分析与需求 分析方法	371
11.5	线性链表	326	13.2.2	结构化分析方法	372
11.5.1	线性链表的基本概念	326	13.2.3	软件需求规格 说明书	376
11.5.2	线性链表的基本运算	330	13.3	结构化设计方法	378
11.5.3	循环链表及其基本运算	332	13.3.1	软件设计的基本概念	378
11.6	树与二叉树	333	13.3.2	概要设计	381
11.6.1	树的基本概念	333	13.3.3	详细设计	386
11.6.2	二叉树及其基本性质	335	13.4	软件测试	390
11.6.3	二叉树的存储结构	338	13.4.1	软件测试的目的	391
11.6.4	二叉树的遍历	339	13.4.2	软件测试的准则	391
11.7	查找技术	341	13.4.3	软件测试技术和 方法综述	392
11.7.1	顺序查找	341	13.4.4	软件测试的实施	399
11.7.2	二分法查找	342	13.5	程序的调试	403
11.8	排序技术	342	13.5.1	基本概念	403
11.8.1	交换类排序法	342	13.5.2	软件调试方法	404
11.8.2	插入类排序法	344	13.6	习题	406
11.8.3	选择类排序法	346			
11.9	习题	348			
第 12 章	程序设计基础	351			
12.1	程序设计方法与风格	351			
12.2	结构化程序设计	353			
12.2.1	结构化程序设计的原则	353			
12.2.2	结构化程序的基本结构 与特点	354			

第 14 章 数据库设计基础	407	14.4 数据库设计与管理	432
14.1 数据库系统的基本概念	407	14.4.1 数据库设计概述	432
14.1.1 数据、数据库和数据库 管理系统	407	14.4.2 数据库设计的需求分析	433
14.1.2 数据库系统的发展	411	14.4.3 数据库概念设计	434
14.1.3 数据库系统的基本特点	413	14.4.4 数据库的逻辑设计	438
14.1.4 数据库系统的内部 结构体系	414	14.4.5 数据库的物理设计	439
14.2 数据模型	416	14.4.6 数据库管理	439
14.2.1 数据模型的基本概念	416	14.5 习题	441
14.2.2 E-R 模型	417	附录 A 常用函数	443
14.2.3 层次模型	421	附录 B 常用事件	447
14.2.4 网状模型	422	附录 C 考试大纲	450
14.2.5 关系模型	423	附录 D 习题参考答案	455
14.3 关系代数	426	参考文献	460

第 1 章 数据库系统的基础知识

本章从数据库系统的基础知识入手,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。
- Access 2003 数据库简介。

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,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,例如文字、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,是数据库系统的中心枢纽。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, 例如查询、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。

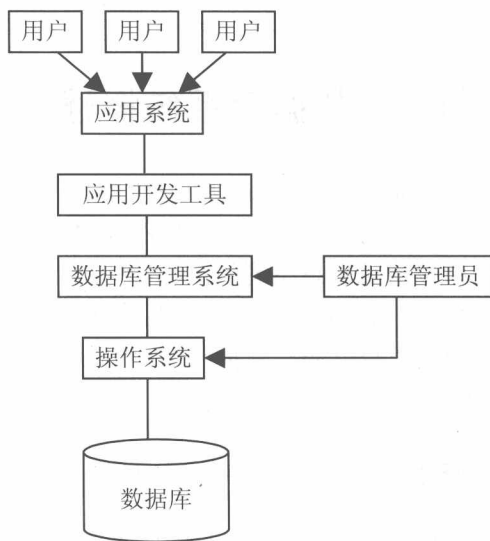


图 1.1 数据库系统的组成

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。

20 世纪 50 年代中期以前, 计算机主要用于科学计算。那时的计算机硬件方面, 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，比如存储结构、存取方法、输入输出方式等。存储结构发生变化，应用程序也要做相应的修改，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性。

(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。

数据库系统阶段的数据管理技术具有以下特点。这个阶段基本实现了数据共享，减少了数据冗余；数据库采用特定的数据模型，具有较高的数据独立性，有统一的数据控制和管理功能。

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的错综复杂的联系。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。另外一级映像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. 数据管理系统对数据进行统一的管理和控制

DBMS 不仅具有基本的数据管理功能,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。

(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。比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。比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,必须采取某种措施,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 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结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结点上。每个结点都有一定的局部数据库管理系统, 使每个结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。网络上每个结点的数据库系统都具有独立处理本地事务的能力, 而且各局部结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。例如, 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, 在个别结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, 只要有一个结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。数据可以存储在临近的常用结点上, 如果本结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发展, 主要出现了下列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。它通过将一个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)。与传统数据库管理系统一样,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,第一类是概念模型,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三个方面的要求。

- (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三部分组成的。

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，例如表中的字段、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；一类是与数据之间联系有关的对象。

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。

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(Entity)

客观存在并可相互区别的事物称为实体。实体可以是实际事物，也可以是抽象事件。比如，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。