



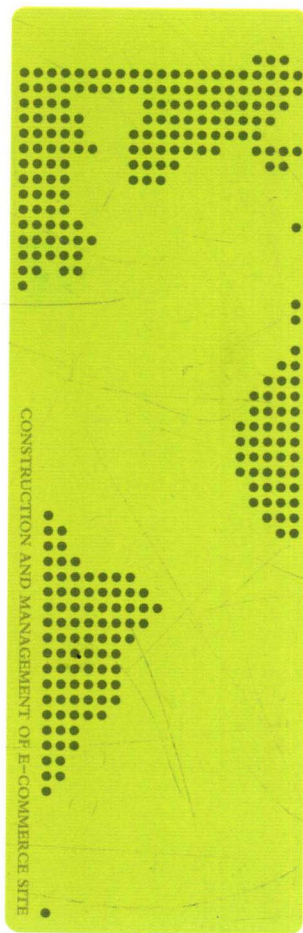
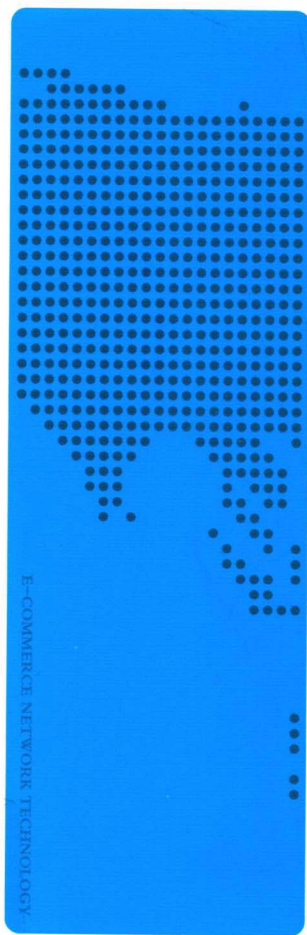
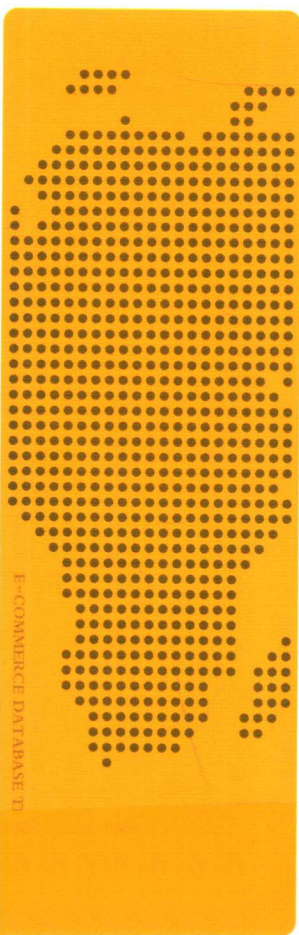
全国高职高专院校“十二五”规划教材
电子商务专业项目式教学课题成果

电子商务

数据库技术

E-COMMERCE DATABASE TECHNOLOGY

主 编 樊颖军 副主编 张 光
主 审 殷锋社



中国水利水电出版社
www.waterpub.com.cn

全国高职高专院校“十二五”规划教材

电子商务专业项目式教学课题成果

电子商务数据库技术

主 编 樊颖军

副主编 张 光

主 审 殷锋社



中国水利水电出版社
www.waterpub.com.cn

内 容 提 要

随着计算机的广泛应用和迅速普及,使得人们在日常生活中管理的信息量急剧增加,因而数据库技术已经逐渐成为信息技术的重要组成部分。目前,数据库技术已经拓展到各行各业的多种应用当中。对于高等院校各专业的学生而言,学习一种数据库管理系统的应用技术,掌握相应的数据库应用系统开发技能是信息技术发展对学生的要求。本教材为读者全面了解数据库技术在电子商务中的应用,运用计算机网络从事经济商务活动、应用、维护和开发电子商务网站打下了坚实的基础。

本书配有电子教案,读者可以从中国水利水电出版社网站和万水书苑上免费下载,网址为: <http://www.waterpub.com.cn/softdown/>或 <http://www.wsbookshow.com>。

图书在版编目(CIP)数据

电子商务数据库技术 / 樊颖军主编. -- 北京: 中国水利水电出版社, 2012.9

全国高职高专院校“十二五”规划教材 电子商务专业项目式教学课题成果

ISBN 978-7-5170-0117-1

I. ①电… II. ①樊… III. ①数据库系统—应用—电子商务—高等职业教育—教材 IV. ①F713.36②TP311.13

中国版本图书馆CIP数据核字(2012)第204473号

策划编辑: 杨庆川 责任编辑: 张玉玲 加工编辑: 庞 焱 封面设计: 李 佳

书 名	全国高职高专院校“十二五”规划教材 电子商务专业项目式教学课题成果 电子商务数据库技术
作 者	主 编 樊颖军 副主编 张 光 主 审 殷锋社
出版发行	中国水利水电出版社 (北京市海淀区玉渊潭南路1号D座 100038) 网址: www.waterpub.com.cn E-mail: mchannel@263.net (万水) sales@waterpub.com.cn
经 售	电话: (010) 68367658 (发行部)、82562819 (万水) 北京科水图书销售中心(零售) 电话: (010) 88383994、63202643、68545874 全国各地新华书店和相关出版物销售网点
排 版	北京万水电子信息有限公司
印 刷	三河市铭浩彩色印装有限公司
规 格	184mm×260mm 16开本 17.25印张 430千字
版 次	2012年10月第1版 2012年10月第1次印刷
印 数	0001—3000册
定 价	32.00元

凡购买我社图书,如有缺页、倒页、脱页的,本社发行部负责调换

版权所有·侵权必究

前 言

本教材以 Access 2003 为主要应用平台, 利用情境教学法重点介绍数据库系统的基本知识、关系数据库、关系数据库标准语言 SQL、数据控制技术等重要内容。具体内容及结构如下: 学习情境 1 数据库认知、学习情境 2 关系数据库系统结构与关系数据库设计、学习情境 3 Access 2003 数据库开发、学习情境 4 关系数据库标准语言——SQL、学习情境 5 数据库控制技术、学习情境 6 管理信息系统。每个学习情境后都附有拓展知识和训练题, 希望通过理论结合实际操作的方式让读者掌握数据库的核心内容。

本教材以学习数据库重要理论与实习操作并重的方式展开各个环节的学习, 以 Access 2003 为主要软件让读者操作。根据诸多专家的观点及作者一线教学的经验, 学生在学习数据库课程时, 对于实际操作的内容往往掌握较好, 因此, 对于数据库设计、关系数据库语言 SQL 等课程重要内容, 本教材都会安排大量练习让学生动手操作, 这样既能加深学生对数据库基本知识的掌握, 也能帮助学生进一步认识数据库软件的使用方法。

本教材由陕西工业职业技术学院的樊颖军任主编, 张光任副主编, 并由樊颖军、张光负责教材的总体策划、统稿和定稿工作。具体编写分工如下: 学习情境 1 由殷锋社编写, 学习情境 2 由朱忠军编写, 学习情境 3 由樊颖军编写, 学习情境 4、学习情境 5 由张光编写, 学习情境 6 由王化冰编写。

本教材在编写过程中参阅和借鉴了国内外相关著作, 在此谨向其作者表示诚挚的谢意。由于时间仓促及编者水平有限, 教材中难免存在错漏和不当之处, 恳请广大读者批评指正。

编 者

2012 年 8 月

目 录

前言	
学习情境 1 数据库认知	1
情境描述	1
学习目标	1
子学习情境一	2
任务一 数据管理技术的产生和发展	2
任务二 信息、数据与数据处理	6
任务三 数据库与数据库管理系统	8
子学习情境二 数据库系统	11
任务一 数据库系统的特点和功能	12
任务二 数据库系统的组成	17
子学习情境三 数据模型	19
任务一 数据模型的组成要素	20
任务二 概念模型	21
任务三 层次模型	23
任务四 网状模型	25
任务五 关系模型	26
任务六 数据模型设计	27
子学习情境四 数据库系统结构	29
任务一 数据库系统的三级模式结构	29
任务二 三级模式结构的特征	30
任务三 数据库系统的二级映像	31
拓展知识 关系数据库的由来	32
基础训练	32
技能训练	33
学习情境 2 关系数据库系统结构与关系数据库设计	34
情境描述	34
学习目标	34
子学习情境一 关系模型的数据模式与关系代数	35
任务一 关系数据结构	35
任务二 关系操作和关系数据语言	38
任务三 传统的集合运算	39
任务四 专门的关系运算	40
子学习情境二 关系的完整性规则及关系数据库的规范化	45
任务一 关系的完整性规则	45
任务二 关系数据库的规范化	48
子学习情境三 数据库存储技术	52
任务一 物理存储介质	52
任务二 记录的存储结构	53
任务三 文件组织	55
子学习情境四 数据库设计的基本步骤	56
任务一 需求分析	59
任务二 概念设计与 E-R 模型	62
任务三 逻辑结构设计	66
任务四 物理结构设计	68
子学习情境五 数据库的完整性	71
任务一 完整性约束的类型	71
任务二 完整性约束的表达与控制	72
任务三 完整性约束的实施	73
子学习情境六 数据库的安全性	73
任务一 数据库安全性措施的层次	74
任务二 数据库的安全机制	74
任务三 数据加密与跟踪审查	78
拓展知识 数据库应用程序编制及调试	78
基础训练	80
技能训练	80
学习情境 3 Access 2003 数据库开发	82
情境描述	82
学习目标	82
子学习情境一 初识 Access 2003	83
任务一 Access 2003 的启动与退出	83
任务二 Access 2003 的用户界面	84
任务三 Access 2003 的主要数据对象	86
任务四 Access 2003 的主要数据类型	88

任务五 Access 2003 的功能选项设置.....89	子学习情境九 Access 2003 中的数据库管理..... 136
子学习情境二 利用 Access 2003 创建数据库.....92	任务一 数据的导入、导出及链接..... 136
任务一 创建 Access 数据库的不同方式.....93	任务二 数据库的备份、压缩与修复..... 139
任务二 打开已有数据库.....94	任务三 设置数据库密码..... 140
任务三 转换 Access 数据库到较低版本.....95	任务四 用户级安全机制..... 140
任务四 数据库的复制与删除.....95	任务五 管理安全机制..... 142
子学习情景三 Access 2003 表操作.....95	拓展知识 VBA 在 Access 2003 中的应用..... 143
任务一 创建表.....95	基础训练..... 154
任务二 表的导入和链接..... 101	技能训练..... 156
任务三 编辑表..... 101	学习情境 4 关系数据库标准语言——SQL..... 158
任务四 数据表关联..... 104	情境描述..... 158
子学习情境四 Access 2003 查询操作..... 105	学习目标..... 158
任务一 查询的种类和创建方法..... 105	子学习情境一 认识 SQL 语言..... 159
任务二 利用查询向导创建查询..... 106	任务一 SQL 语言的产生与发展..... 159
任务三 利用设计视图创建查询..... 109	任务二 SQL 语言的特点..... 159
子学习情境五 Access 2003 窗体操作..... 112	任务三 SQL 语言的功能..... 160
任务一 窗体概述..... 112	子学习情境二 SQL 的数据定义功能..... 161
任务二 利用窗体向导创建窗体..... 113	任务一 定义、删除和修改基本表..... 161
任务三 自动创建窗体..... 114	任务二 视图..... 163
任务四 利用窗体设计视图创建窗体..... 114	任务三 索引..... 165
任务五 窗体的修改..... 115	子学习情境三 SQL 数据查询..... 166
任务六 主/子窗体..... 117	任务一 SQL 语句的单表查询操作..... 168
子学习情境六 Access 2003 的报表操作..... 118	任务二 SQL 语句的复杂查询操作..... 170
任务一 报表概述..... 118	子学习情境四 SQL 数据更新..... 172
任务二 报表的创建..... 119	任务一 插入数据..... 172
任务三 利用设计视图创建或修改报表..... 121	任务二 修改数据..... 173
任务四 子报表..... 123	任务三 删除数据..... 173
任务五 报表的预览和打印..... 125	子学习情境五 SQL 数据控制..... 174
子学习情境七 Access 2003 的数据访问页..... 125	任务一 授权..... 174
任务一 数据访问页概述..... 125	任务二 收回权限..... 175
任务二 创建数据访问页..... 126	拓展知识 嵌入式 SQL..... 175
任务三 编辑数据访问页..... 127	基础训练..... 178
任务四 访问数据访问页..... 129	技能训练..... 178
子学习情境八 Access 2003 的宏操作..... 129	学习情境 5 数据库控制技术..... 179
任务一 宏概述..... 129	情境描述..... 179
任务二 运行宏的条件..... 131	学习目标..... 179
任务三 创建宏..... 132	子学习情境一 事务..... 180
任务四 宏的运行与调试..... 135	任务一 事务的概念..... 180

任务二 事物的特性·····	180	子学习情境二 管理信息系统的开发综述·····	212
子学习情境二 并发控制技术·····	181	任务一 管理信息系统的开发方法概述·····	212
任务一 并发操作带来的不一致性·····	181	任务二 结构化方法的开发步骤·····	214
任务二 基于封锁的并发控制技术·····	183	子学习情境三 管理信息系统的总体规划·····	218
任务三 活锁和死锁·····	186	任务一 管理信息系统总体规划概述·····	218
子学习情境三 数据库恢复技术·····	188	任务二 管理信息系统总体规划的方法·····	221
任务一 数据库可能出现的故障种类·····	188	任务三 可行性分析·····	229
任务二 数据库恢复技术·····	189	子学习情境四 管理信息系统的系统分析·····	231
任务三 数据库恢复策略·····	191	任务一 管理信息系统分析概述·····	231
拓展知识 数据库新技术·····	192	任务二 现行系统分析·····	239
基础训练·····	198	任务三 数据仓库的应用·····	253
技能训练·····	198	任务四 新系统的逻辑模型·····	258
学习情境6 管理信息系统·····	199	任务五 系统分析报告·····	259
情境描述·····	199	子学习情境五 管理信息系统的系统设计·····	260
学习目标·····	199	任务一 系统设计的内容·····	261
子学习情境一 管理信息系统概述·····	200	任务二 功能模块设计·····	261
任务一 管理信息系统的概念·····	200	拓展知识 开发实例·····	264
任务二 管理信息系统的结构·····	202	基础训练·····	268
任务三 系统结构模式·····	207	技能训练·····	268
任务四 管理信息系统的类型·····	209	参考文献·····	269
任务五 系统应用现状和发展趋势·····	212		

学习情境

1

数据库认知



情境描述

数据库技术是计算机技术的重要分支，是数据管理的实用技术。数据库技术是研究数据库的结构、存储、设计和使用的软件学科。它产生于 20 世纪 60 年代中期。人类社会随着计算机技术、通信技术和网络技术的发展，已经进入了信息化时代。信息资源已成为最重要和最宝贵的资源之一，建立一个行之有效的信息系统是企业组织生存和发展的重要条件。

根据计算机应用的调查，在计算机的三大主要应用领域（科学计算、数据处理与过程控制）中，数据处理约占 70% 左右的比重。正因为如此，为适应数据处理需求的不断增长，满足各类信息系统对数据管理的要求，随着数据库技术的发展，数据管理在文件系统的基础上不断发展，产生了目前普遍采用的数据库管理系统。现在，数据库技术已成为计算机领域中最重要技术之一，是软件学科的一个独立分支。

数据库方法原本是针对事务处理中大量数据管理需求的，但是它的应用范围不断被扩大，不仅应用于事务处理，而且进一步应用到情报检索、人工智能、专家系统和计算机辅助设计等，涉及非数值计算各个方面的应用。作为信息系统核心技术和重要基础的数据库技术有了飞速发展，并得到了越来越广泛的应用。应用范围的扩大又进一步促进了数据库技术的深入发展，可以说数据库系统已成为当代计算机系统的重要组成部分。

数据库理论研究主要集中在关系的规范化理论和关系数据理论。近年来，随着人工智能与数据库的结构、数据库与逻辑、逻辑演绎与知识推理等的理论研究，演绎数据库、知识库系统的研制都已成为新的研究方向。

近半个世纪以来，数据库在理论上和实现技术上均得到了很大的发展，出现了许多商用数据库管理系统，使得计算机应用渗透到各行各业的各类管理工作中。例如，管理信息系统、办公自动化系统和决策支持系统等都是使用了数据库管理系统或数据库技术的典型的计算机应用系统。从某种意义上讲，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。

电子商务本质就是信息化商务，作为电子商务专业的学生，有关数据库的知识和技能必须掌握，要设计和开发数据库应用系统，首先应该了解数据库技术的产生和发展情况、熟悉数据系统的组成与工作过程并且掌握数据模型的组成要素，在本学习情境中，我们将接触这些必须掌握的基础知识。

学习目标

- 信息、数据、数据处理和数据管理等数据库技术相关概念的定义。

- 数据库、数据库管理系统、数据库系统的定义及其功能描述。
- 数据模型的定义、组成及特点描述。
- 概念模型及其表示方法。
- 常用的三种结构数据模型的表示及特征描述。
- 数据库系统的三级模式和两级映象结构。

子学习情境一

在系统地介绍数据库理论之前，首先介绍一些与数据库有关的基础知识。

任务一 数据管理技术的产生和发展

对数据进行收集、保存、加工和传播等一系列活动的总和称为数据处理，数据处理的核
心问题就是数据管理。数据管理技术是指对数据分类、编码、存储、检索和维护的技术。数据
管理技术的发展是和计算机技术及其应用的发展联系在一起的，经历了如下 3 个阶段：人工管
理阶段、文件系统阶段和数据库系统阶段。

1. 人工管理阶段

20 世纪 50 年代中期以前，计算机主要用于科学计算。这是计算机数据处理的初级阶段，
对数据的处理是由程序员个人考虑和安排的。程序员在编制程序时，还要考虑数据的逻辑定义
和物理组织，以及数据存放的存储设备、存储方式和地址分配。在这一阶段，程序和数据混为
一体，两者相互依赖，数据成为程序不可分割的一部分，根本谈不上数据的独立性。如图 1-1
(a) 所示，当程序之间出现重复数据时，这些数据也不能共享，数据的处理是分散的。计算
机在数据处理中没有发挥应有的作用，因此严重地影响了计算机使用效率的发挥和提高。当时，
在硬件方面，外存储器只有卡片、纸带、磁带，没有像磁盘这样的可以随机访问、直接存取
的外部存储设备；在软件方面，没有专门管理数据的软件，数据由计算或处理它的程序自行携带，
数据处理方式基本是批处理。这一时期数据管理的特点是：

(1) 数据不保存。因为计算机主要用于科学计算，对于数据保存的需求尚不迫切。一般
的做法是，在进行某一课题计算时将原始数据随程序一起输入主存，运算处理后将结果数据输
出。随着计算任务的完成，用户作业退出计算机系统，数据空间随同程序空间一起被释放。

(2) 没有专用的软件对数据进行管理。每个应用程序都要包括存储结构、存取方法和输
入输出方式等，这些完全由应用程序员编程实现，从而要求应用程序员必须对各种数据存储结
构的性能、实现和维护有较深的了解。

(3) 程序中的存储子程序随着存储结构的改变而改变，因而数据与程序不具有独立性，
即一组数据对应一组程序。这就使得程序依赖于数据，如果数据的类型、格式或者数据量、存
取方法、输入输出方式等改变了，程序必须作相应的修改。此外，由于程序是直接面向存储结
构的，因此不存在逻辑结构与物理结构的区别。

(4) 只有程序的概念，没有文件的概念。即使有文件，也大多是顺序文件，其他组织方
式必须由程序员自行设计与安排。

(5) 数据面向于应用。即一组数据对应于一个程序。由于各应用程序处理的数据不会毫
无关系，程序之间可能有重复数据。

(6) 对数据的存取以记录为单位, 灵活性差。

2. 文件系统阶段

20 世纪 50 年代后期至 60 年代中后期, 随着计算机硬件性能的改进和软件技术的发展, 一个应用程序独享一台计算机资源的情况显然是一种极大的浪费, 于是出现了多道程序和分时系统。这时, 如果仍然让用户自己来安排其数据的存储设备、物理存储方式和物理地址分配, 则会造成灾难性的后果。因此, 出现了文件管理系统, 如图 1-1 (b) 所示, 它作为应用程序和数据文件的接口。在硬件方面, 可直接存取的磁鼓、磁盘成为联机的主要外存。在软件方面, 出现了高级语言和操作系统。操作系统中的文件系统 (有的也称为信息处理模块) 是专门管理外存的数据管理软件。数据处理方式有批处理, 也有联机实时处理。这一阶段数据管理的特点是:

(1) 数据可长期保存在外存的磁盘上。用户可随时通过程序对文件进行查询、修改和增删等处理。由于计算机应用从计算转向管理, 数据处理的工作量增大。

(2) 数据的物理结构与逻辑结构有了区别, 但较简单。程序与数据之间有设备独立性, 由操作系统的文件系统提供存取方法。

(3) 文件的形式已多样化。有索引文件、链接文件和直接存取文件等, 因而对文件的记录可顺序访问, 也可随机访问。但文件之间是独立的, 其共享性差, 联系要通过程序去构造。

(4) 有了存储文件以后, 数据不再仅仅属于某个特定的程序, 而可以重复使用。但文件结构的设计仍然是基于特定的用途, 程序仍然是基于特定的物理结构和存取方法编制的, 因此, 数据结构与程序之间的依赖关系并未根本改变。

(5) 对数据的存取基本上还是以记录为单位。

文件系统阶段是数据管理技术发展中的一个重要阶段。在这一阶段中, 得到充分发展的各种数据结构和算法大大丰富了计算机科学, 今天的数据库技术也正是在文件系统的基础上发展起来的。在文件系统中, 数据按其内容、结构和用途组成若干命名的文件。

应用程序通过文件系统进行数据文件的建立、存取、修改和撤消等操作。虽然文件系统允许各个应用程序所建立的文件共享, 但由于各应用程序根据其自身的需要建立的数据文件有其自身的特点, 尽管这些文件包含了别的应用程序所需的数据, 这些数据仍很难为别的用户或应用程序所使用。因此, 文件系统所管理的数据文件基本上是分散的、相互独立的。以此为基础的数据处理存在以下缺点:

(1) 数据冗余度大。数据冗余是指不必要的重复存储, 同一个数据项重复出现在多个文件中。当文件的设计很难满足多种应用程序的不同要求时, 数据冗余往往是不可避免的。在文件系统下, 数据文件基本上与各自的应用程序相对应, 数据不能以记录和数据项为单位共享。即使有部分数据相同, 只要逻辑结构不同, 用户都必须各自建立自己的文件, 这不仅浪费存储空间, 而且增加更新开销。例如, 在一所学校里, 教师的姓名、所在系等数据可能同时出现在人事档案、科研档案和工资等文件中。这种数据冗余大大降低了存储空间的有效利用率。

(2) 数据的不一致性。由于同一数据重复存储在不同的应用程序中, 如果不能保证其更新的同时性和正确性, 势必造成同一数据在不同的文件中出现不同的值。随着应用环境和需求的变化, 修改文件的结构是常有的事, 如在某些文件的记录中增加一些字段、扩充某些字段的长度、改变某些字段的计量单位或表示格式等。例如, 某个教师因为某项科研工作有重大贡献而被晋升一级工资, 但人事部门未及时通知财务部门, 从而造成该教师的工资在人事档案和工

资档案中有不同的值。

(3) 程序与数据的相互依赖。文件系统中的数据文件是为了满足特定业务领域某部门的专门需要而设计的,它服务于某一特定应用程序。数据和程序相互依赖,如果改变数据的逻辑结构或文件的组织方法,必须修改相应的应用程序。同样道理,如果修改应用程序,改用另一种程序设计语言来编写程序,也将影响数据文件的结构。

(4) 数据无集中管理。除了对记录的存取由文件系统承担以外,文件没有统一的管理机制,其安全性与完整性无法保障。数据的维护任务仍然由应用程序来承担。由于数据缺少统一管理,在数据的结构、编码、表示格式、命名以及输出格式等方面不容易做到规范化、标准化;在数据的安全和保密方面,也难以采取有效的措施。

(5) 编写应用程序很不方便。应用程序的设计者必须对所用的文件的逻辑及物理结构有清楚的了解。操作系统只提供打开、关闭、读、写等几个低级的文件操作命令,对文件的查询、修改等处理都需要在应用程序内解决。

文件是无弹性、无结构的数据集合。无弹性是指由于记录的内部结构由应用程序自己定义,而不是由系统来统一管理,因此对现有数据文件的应用不易扩充、不易移植,也难以增删数据项以适用新的应用要求。无结构是指各个数据文件之间是独立的,缺乏联系,不能反映现实世界事物之间的联系。这些问题阻碍了数据处理技术的发展,不能满足日益增长的信息要求,这是数据库技术产生的原动力,也是数据库系统产生的背景。应用需求和计算机技术的发展促使人们研究一种新的数据管理技术——数据库技术。

3. 数据库系统阶段

计算机的软件工作者针对文件管理方式存在的缺点,经过长期不懈的努力,提出了数据库的概念。数据库技术为数据管理提供了一种较为完善的高级管理方式。它克服了文件管理方式下分散管理数据的弱点,对所有的数据实行统一、集中的管理,使数据存储独立于使用它的程序,从而实现数据共享。数据库管理方式如图 1-1 (c) 所示。

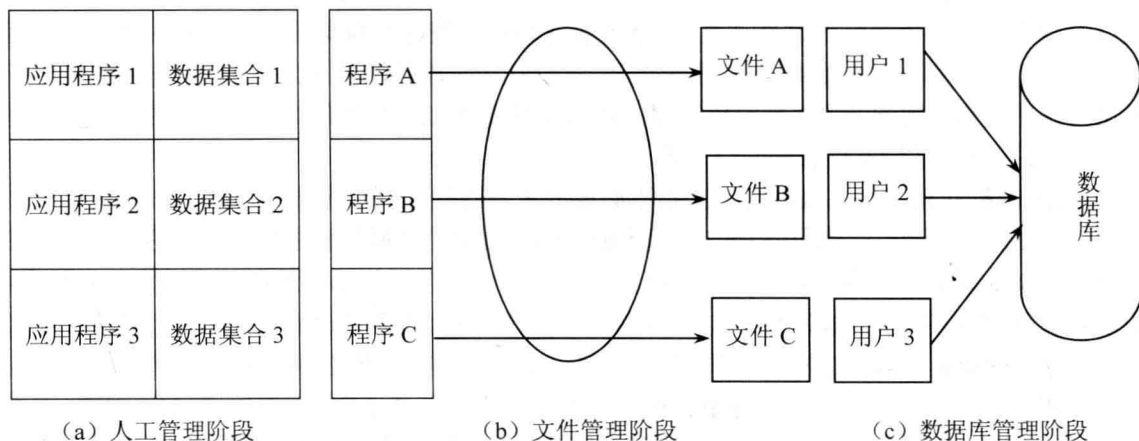


图 1-1 数据库技术的发展阶段

数据库是通用化的相关数据集合,它不仅包括数据本身,而且包括数据之间的联系。数据库中的数据不是只面向某一项特定应用,而是面向多种应用,可以被多个用户、多个应用程序共享。例如,某个企业、组织或行业所涉及的全部数据的汇集。其数据结构独立于使用数据



的程序,数据的增加、删除、修改和检索由系统进行统一的控制,而且数据模型也有利于将来应用的扩展。

为了让多种应用程序并发地使用数据库中具有最小冗余度的共享数据,必须使数据与程序具有较高的独立性。这就需要有一个软件系统对数据实行专门管理,提供安全性和完整性等统一控制机制,方便用户以交互命令或程序方式对数据库进行操作。

为数据库的建立、使用和维护而配置的软件称为数据库管理系统 DBMS (Database Management System),它是在操作系统支持下运行的。目前流行的数据库管理系统包括 Oracle、Informix、Sybase 等,主要应用在微机上的数据库管理系统有 dBASE、FoxBASE、FoxPro、FoxPro for Windows、Access、Clipper、WinBASE 等。

由于 DBMS 不是为某个应用程序服务,而是为整个单位或部门服务的,DBMS 做得复杂一些是可以接受的。许多在文件系统中难以实现的功能,在 DBMS 中都一一实现了。例如,适合不同类型用户的多种用户界面、保证并发访问时的数据一致性的并发控制、增进数据安全性的访问控制、在故障情况下保证数据一致性的恢复功能、保证数据在语义上的一致性的完整性约束检查功能等。随着计算机应用的发展,DBMS 的功能越来越强,规模越来越大,复杂性和开销也随之增加。

世界上第一个 DBMS 是由美国通用电气公司 Bachman 等人于 1964 年开发成功的 IDS (Integrated Data Store)。IDS 奠定了网状数据库 (Network Database) 的基础,并得到了广泛的发行和应用。IBM 公司于 20 世纪 60 年代末推出了第一个商品化的层次数据库管理系统 IMS (Information Management System)。层次和网状数据库描述数据的模型是从过去应用程序处理数据时所用的数据结构概括而来的,尽管有一定的通用性,但其中保留了不少实现的细节,例如指针等。层次和网状数据库是 20 世纪 60 年代技术条件下的合理产物。它们为数据库技术奠定了基础,但数据库技术最有意义的成就是关系数据库的发展。1970 年 Codd 提出了关系数据模型 (Relational Data Model),以关系或表作为描述数据的基础。在其后的几年中,Codd 又发表了一系列文章,奠定了关系数据库的理论基础。

4. 数据库系统与文件系统的关系

数据库系统与文件系统既有联系,又有显著的区别。首先,两者都有一定的数据管理功能和数据检索与更新等数据操作功能;其次,数据库最终还是通过文件系统以文件形式存放在物理存储设备上的。因此,文件系统是数据库系统的基础,数据库系统是文件系统的提高和发展。两者的主要区别有以下几点:

(1) 控制方式。文件系统中的数据文件由各个应用程序员根据需要组织起来,并由各应用程序负责建立、使用和维护,因此其控制方式是分散的;而数据库中的数据文件是由数据库系统统一规划,按照一定的数据模型组织和建立的,由系统统一管理和集中控制。

(2) 数据结构。文件系统中的文件结构简单,不能反映它所代表的现实世界中各事物之间的内在联系;数据库中的数据是一个组织的所有应用所需数据的集成,反映了各数据所代表的现实世界各事物之间存在的内在联系。

(3) 数据独立性。文件系统中的数据文件是面向应用的,它依赖于应用程序的存在而存在。换句话说,文件系统是数据相关的。相反,数据库系统是面向数据的系统,它具有高度的数据独立性。

(4) 数据的维护。数据库系统很容易发现故障,可以排除故障并恢复到故障前的状态;

而文件系统本身无完整性约束的定义和检查功能,所以它没有故障恢复功能。

任务二 信息、数据与数据处理

在信息化的社会中,人们在各个领域的各种活动中都将产生大量的信息,经常对信息进行加工、处理、交流和使用。随着计算机技术的迅速发展,计算机的高速性和巨大存储能力,使得人们可对大量存在的信息进行保存和加工处理。为了记载信息,人们使用各种符号和它们的组合来表示信息,这些符号及其组合就是数据。数据是信息的具体表示形式,信息是数据有意义的体现。

1. 信息

(1) 信息的定义。信息是现实世界事物的存在方式或运动状态的反映。信息具有可感知、可存储、可加工、可传递和可再生等自然属性,信息又是社会上各行各业不可缺少的资源,这也是信息的社会属性。

(2) 信息的特征。

- 信息源于物质和能量。一切事物,包括自然界和人类都产生信息。信息是物质和能量形态的反应,它不可能脱离物质而存在。信息传递需要物质载体,信息的获取和传递要消耗能量。
- 信息是可以感知的。人类对客观事物的感知,可以通过感觉器官,也可以通过各种仪器、仪表和传感器等,不同的信息源有不同的感知方式。
- 信息是可以存储的。动物用大脑存储信息,叫做记忆。文字的发明,提供了长期存储信息的可靠手段。近年来,计算机存储器、录音、录像等技术的发展,进一步扩大了信息存储的范围。
- 信息是可以加工、传递和再生的。电子计算机是信息处理领域中最先进的工具之一,人类对收集到的信息可以进行取舍整理。从客观世界收集到或加工后的信息,可以通过各种形式传递,例如语言、表情、动作、文字、电报、电话、电视、计算机通信和卫星通信等。

(3) 信息系统。信息系统是指为了某些明确的目的而建立的,由人员、设备、程序和数据集合构成的统一整体。信息系统的主要功能是提供信息,以支持一个组织机构的运行、管理和决策。更确切地说,信息系统将不适用的数据形式加工成可利用的形式。

信息系统可分为3类:数据处理系统、管理信息系统和决策支持系统。

- 数据处理系统。用计算机代替繁杂的手工处理事务工作,其目的是提高数据处理的准确性、及时性,节约人力,提高工作效率。例如,计算机运行会计核算软件,对会计的“簿记”事务进行常规处理,提供数据查询、会计报表等功能,使会计部门的日常工作自动化。
- 管理信息系统。由若干子系统构成的一个集成的人机系统,从组织的全局出发,实现数据共享,提供分析、计划、预测、控制等方面的综合信息。其目的是发挥系统的综合效益,提高管理水平。
- 决策支持系统。为决策过程提供有效的信息和辅助决策手段的人机系统。其主要目的是帮助决策者提高决策的科学性及有效性。决策支持系统只能对决策提供支持,并不能由计算机代替人自动地作出决定,人是决策行动的主体。

2. 数据

几乎和信息同样广泛使用的另一个概念是“数据”。所谓数据，通常指用符号记录下来的可加以鉴别的信息。数据的概念包括两个方面：其一，数据内容是事物特性的反映或描述；其二，数据是符号的集合。

数据是数据库中存储的基本对象。广义的理解，数据的种类很多，文字、图形、图像、声音、学生的档案记录、货物的运输情况等，这些都是数据。可以对数据做如下定义：数据是描述现实世界事物的符号记录，是指用物理符号记录下来的可以鉴别的信息。物理符号包括：数字、文字、图形、图像、声音及其他特殊符号。数据的多种表现形式，都可以经过数字化后存入计算机。

信息与数据的关系为：数据是信息的符号表示，或称载体；信息是数据的内涵，是数据的语义解释。信息与数据是密切相关联的，因此，在某些不需要严格区分的场合，也可以把两者不加区别地使用，即信息处理也可以说成数据处理。例如，每届毕业生人数是信息，而 6000~7000 是数据。数据能表示信息，但并非任何数据都能表示信息，信息只是人们消化理解了的数据。数据是信息存在的一种形式，只有通过解释或处理才能成为有用的信息。信息是抽象的，不随数据设备所决定的数据形式而改变，数据的表示方式具有可选择性。

3. 数据处理和数据管理

所谓数据处理就是指对各种类型的数据进行收集、存储、分类、排序、计算或加工、检索、传输、递交等工作。数据处理通常也称为信息处理。由于在数据处理过程中所遇到的数据是有组织的，相互之间存在一定的联系，因此数据处理的效率往往与数据的存储结构和处理方式有密切的联系。

数据管理和数据处理一样，都是计算机系统的最基本的支撑技术。尽管计算机科学技术经历了飞速的发展，但数据管理的这一地位没有变化。可以预期，数据管理将作为计算机科学技术的一个重要分支一直发展下去，社会越信息化，对数据管理的要求也越高。推动这门学科发展的动力是计算机应用，发展这门学科的基础是新的硬件、软件技术。数据库是现阶段数据管理的主要形式。

4. 信息处理

由于信息是现实世界中事物的存在方式和运动状态的反映，而现实世界的事物常常是相互关联的，这就使得人们在了解、掌握事物之间的固有联系和运动规律的基础上，可以从一些已知的信息出发，经过演绎推理，导出新的信息，为人类社会生活的各种需要服务，这就是我们常说的“信息处理”。例如，学号“B050201024”，“B”表示是本科生，“05”表示 05 级，“02”表示所在学院的编号，“01”表示专业编号，最后三位“024”表示顺序号，这样，根据学生的学号，我们可以推算出学生的类别、所在学院、入学年份、所学专业。

信息处理的基本环节包括：人们将原始信息表示成数据，称为源数据，然后对这些源数据进行汇集、存储、综合、推导，从这些原始的、杂乱的、难以理解的数据中抽取或推导出新的数据，这些新的数据称为结果数据，结果数据对某些特定的人们来说是有价值的、有意义的，它表示了新的信息，可以作为某种决策的依据或用于新的推导。这一过程通常称为数据处理或信息处理。信息处理活动的基本环节如图 1-2 所示。

信息是有价值的，信息的价值与它的准确性、及时性、完整性和可靠性有关。因为信息的价值必须通过使用信息的决策者的行为结果来体现，所以，为了提高信息的价值，就要用科

学的方法来管理信息，这种科学的方法就是数据库技术。

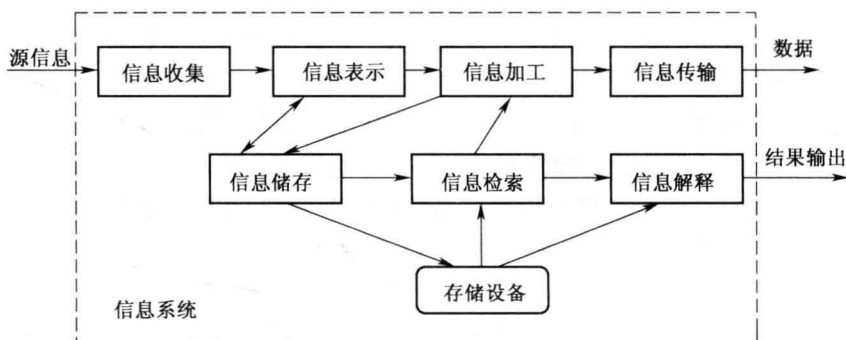


图 1-2 信息处理的基本环节

5. 信息的 3 个领域

作为客观世界的一切真实的信息从客观事物出发，经过筛选流经数据库，通过控制决策机构又回到客观事物。信息的这一循环经历了 3 个领域：现实世界、观念世界和数据世界。

(1) 现实世界。现实世界又称事实世界，指存在于人们头脑之外的客观世界，就是我们赖以生存的生活环境。事实及其相互联系就处于这个世界之中。它是具体事物和抽象概念的总和，是数据处理的源泉。事物可分成“对象”与“性质”两大类，又可分为“特殊事物”与“共同事物”两个重要级别。

(2) 观念世界。观念世界又称信息世界，它是现实世界在人们头脑中的反映。现实世界中的事实反映到人的头脑中来，人的脑子对这些事实有个认识过程，经过选择、命名、分类之后进入观念世界。观念世界的主要对象是实体以及实体间的相互联系。

(3) 数据世界。数据世界又称计算机世界或机器世界，它是数据库系统的处理对象。现实世界中的事实经过观念世界转换成信息，并经过加工、编码进入计算机世界。计算机世界的处理对象是数据，它是载荷信息的各种物理符号，是观念世界中信息的数据化。现实世界中事实及其联系在这里用数据模型来描述。由于数据库技术的崛起，有了处理数据世界的工具。今后只有出现智能计算机之后，才能找到解决用于直接处理现实世界的工具。

任务三 数据库与数据库管理系统

数据库是指存放数据的仓库，只不过这个仓库是在计算机存储设备上，而且数据是按一定的格式存放的。人们收集并抽取出一个应用所需要的大量数据之后，应将其保存起来以供进一步加工处理，进一步抽取有用信息。在科学技术飞速发展的今天，人们的视野越来越广，数据量急剧增加，过去人们把数据存放在文件柜里，现在人们借助计算机和数据库技术科学地保存和管理复杂的大量的数据，以便能方便而充分地利用这些宝贵的信息资源。

1. 数据库

所谓数据库是指长期存储在计算机内的、有组织的、可共享的数据集合。数据库中的数据按一定的数据模型组织、描述和储存，具有较小的冗余度、较高的数据独立性和易扩展性，并可为各种用户共享。较为全面的定义是：数据库是为满足某部门各种用户的多种应用需要，在计算机系统中按照一定数据模型组织、存储和使用的互相关联的数据集合。



目前,具有数百万甚至数十亿字节信息的数据库已经普遍存在于科学技术、工业、农业、商业、服务业和政府部门的信息化系统中。由于数据库应用领域的不断扩大,数据库不仅广泛应用于管理,而且已开始应用到工程设计、图形图像和声音等多介质处理、自动控制和计算机辅助设计等新的应用领域。这些新的应用领域所处理的数据和管理领域中数据的格式有极大的区别,如声音、图像,我们称之为非格式化的数据。处理的要求也不同,因而研究这些新的应用领域中的数据库方法和技术是一个新的课题。它不仅涉及应用系统的设计方法,而且涉及数据库系统的模型实现技术等各种新的问题,面向对象的数据库系统、扩展的数据库系统、多介质的数据库等研究方向的兴起就是基于这些新的需求和应用背景。

2. 数据库管理系统的特点

数据库管理系统(DBMS)是指数据库系统中对数据进行管理的软件系统,它是数据库系统的核心组成部分,数据库系统的一切操作,包括查询、更新及各种控制,都是通过DBMS进行的。DBMS总是基于某种数据模型,因此可以把它看成是某种数据模型在计算机系统上的具体实现。根据所采用数据模型的不同,DBMS可以分成网状型、层次型、关系型、面向对象型等。

DBMS是数据库系统的基础,研制DBMS的基本目标是扩大功能、提高性能和可用性,从而提高用户的生产率。研制以DBMS为核心的一组相互联系的软件系统已成为当前数据库软件产品的方向,这些在DBMS基础上运行的软件系统有:数据通信软件、表格软件、数据字典、报表书写、图形系统等。

但是在不同的计算机系统中,由于缺乏统一的标准,即使同样数据模型的DBMS,它们在不同用户接口、系统功能等方面也常常是不相同的。用户要对数据库进行操作,是由DBMS把操作从应用程序带到外部级、概念级,再导向内部级,进而操纵存储器中的数据。一个DBMS的主要目标是使数据作为一种可管理的资源来处理。DBMS应使数据易于为各种不同的用户所共享,应该增进数据的安全性、完整性和可用性,并提供高度的数据独立性。

3. 数据库管理系统的主要功能

数据库管理系统是一个非常复杂的系统软件,是为数据库的建立、使用和维护而配置的。它把用户程序的数据操作语句转换成对系统存储文件的操作;它又像一个向导,把用户对数据库的一次访问从用户级带到概念级,再导向物理级。DBMS的职能就是有效地实现数据库三级之间的转换,其主要功能包括以下几个方面:

(1) 数据定义功能。DBMS提供数据定义语言(Data Definition Language, DDL),用户通过它可以方便地对数据库中的数据对象进行定义。用DDL可以定义数据库的结构(包括外模式、概念模式、内模式及其相互之间的映象)、数据的完整性约束和保密限制等约束条件。定义工作是由DBA完成的。因此在DBMS中包括DDL的编译程序,它把用DDL编写的各种源模式编译成相应的目标模式。这些目标模式是对数据库的描述,而不是数据本身,它们是数据库的结构,并被保存在数据字典中,供以后进行数据操纵或数据控制时查阅使用。

(2) 数据操纵功能。DBMS还提供数据操纵语言(Data Manipulation Language, DML)实现对数据库的操作,用户可以使用DML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。DML有两类:一类是嵌入在宿主语言中使用,如嵌入在C语言中;另一类是可以独立地交互使用的DML,称为自主型或自含型DML。因而DBMS中必须包括DML的编译程序或解释程序。

(3) 数据库的建立和维护功能。它包括数据库初始数据的载入、转换功能、数据库的转

储功能、数据库的重组功能和性质监视、分析功能等。这些功能大都由各个实用程序来完成。例如装配程序、重组程序、日志程序和统计分析程序等。

(4) 数据库的运行管理。这是 DBMS 运行时的核心部分, 包括并发控制、存取控制(安全性检查)、完整性约束条件的检查和执行、数据库内部的维护(如索引、数据字典的自动维护)等。

(5) 数据字典。数据字典中存放着数据库三级结构的描述。对于数据库的操作都要通过查阅数据字典进行。现有大型系统中, 把数据字典单独抽出来自成系统, 成为一个软件工具, 使得数据字典成为一个比 DBMS 更高级的用户和数据库之间的接口。

DBMS 的其他控制功能还有系统缓冲区的管理和数据存储的某些自适应调节机制等。DBMS 的功能随不同的计算机系统而有所差异, 通常大型计算机系统功能较强、较全, 而小型计算机系统功能较弱。目前微机上的许多 DBMS 就没有并发控制功能、恢复功能、重组功能、性能检测分析功能等。

4. 数据库管理系统的工作过程

应用程序通过 DBMS 读取数据库中的一个记录的过程如下:

(1) 用户在其应用程序中安排一条读记录的 DML 语句。该语句给出被涉及的外模式记录类型名及欲读记录的关键码值。当计算机执行该 DML 语句时, 立即启动 DBMS, 并向 DBMS 发出读记录的命令。

(2) DBMS 接到命令后, 首先访问该用户对应的外模式, 检查该操作是否在合法授权范围内, 若不合法则拒绝执行并向应用程序状态返回区发出不成功的状态信息; 若合法则执行下一步。

(3) DBMS 读取相应的模式描述, 并从外模式映象到概念模式。也就是把外模式的外部记录格式映象到概念模式的概念记录格式, 决定概念模式应读入哪些记录。

(4) DBMS 调用相应的内模式描述, 并从概念模式映象到内模式, 即把概念模式的概念记录格式映象到内模式的内部记录格式, 确定应读入哪些物理记录以及具体的地址信息。DBMS 向操作系统发出从指定地址读取物理记录的命令。

(5) 操作系统执行读取命令, 按指定地址从数据库中把记录读入到数据库的系统缓冲区, 并在操作结束后向 DBMS 作出回答。

(6) DBMS 收到操作系统读操作结束的回答后, 参照概念模式将读入系统缓冲区中的内容转换成概念记录, 再参照外模式转换成用户要求读取的外部记录。

(7) DBMS 把导出的外部记录从系统缓冲区送到应用程序的“程序工作区”中。

(8) DBMS 向运行日志数据库发出读一条记录的信息, 以备以后查询使用数据库的情况。

(9) DBMS 将操作执行成功与否的状态信息返回给用户。

(10) 应用程序根据返回的状态信息决定是否使用工作区中的数据。

如果用户需要修改一条记录的内容, 其过程与此类似。这时首先读出目标记录, 并在用户工作区中用主语言的语句进行修改, 然后向 DBMS 发出写回修改记录的命令。DBMS 在系统缓冲区进行必要的转换后向操作系统发出写命令, 即可达到修改数据的目的。

在数据库管理系统的支持下, 按照应用要求为某一部门或组织设计一个结构良好、使用方便、效率较高的数据库及其应用系统, 这是数据库设计的主要含义。在这一领域内, 主要的研究课题是数据库设计方法学和设计工具的探索。其中包括数据库设计方法、设计工具和理论