



普通高等教育“十三五”规划教材

# 数据库基础与应用

## (Access 2010)

◎ 主编 刘卫国

中国工信出版集团



电子工业出版社  
PUBLISHING HOUSE OF ELECTRONICS INDUSTRY  
<http://www.phei.com.cn>

本书包含配套微视频

普通高等教育“十三五”规划教材

# 数据库基础与应用 ( Access 2010 )

刘卫国 主编

電子工業出版社

Publishing House of Electronics Industry

北京 • BEIJING

## 内 容 简 介

本书在适度介绍关系数据库的基本概念和基本原理的基础上,以常用的 Access 2010 作为实践环境,介绍数据库的基本操作和应用开发技术。全书主要内容有数据库基础知识、数据库的操作、表的操作、查询的操作、SQL 查询的操作、窗体的操作、报表的操作、宏的操作、模块与 VBA、应用案例。全书以“教学管理”数据库的操作为主线,设计编排了大量的实例,便于读者学习和提高。本书的附录部分是实验指导,其中设计了 12 个实验,方便读者上机练习。本书中的一些重要概念与操作方法还配套制作了微视频,读者可通过扫描书中二维码观看、学习。

本书既可作为高等院校数据库基础与应用课程的教材,又可供社会各类计算机应用人员与参加各类计算机等级考试的读者阅读参考。

未经许可,不得以任何方式复制或抄袭本书之部分或全部内容。

版权所有,侵权必究。

## 图书在版编目(CIP)数据

数据库基础与应用: Access 2010 / 刘卫国主编. —北京: 电子工业出版社, 2016.12  
ISBN 978-7-121-30180-3

I. ①数… II. ①刘… 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(2016)第 255676 号

策划编辑: 戴晨辰

责任编辑: 戴晨辰

印 刷: 三河市华成印务有限公司

装 订: 三河市华成印务有限公司

出版发行: 电子工业出版社

北京市海淀区万寿路 173 信箱 邮编: 100036

开 本: 787×1 092 1/16 印张: 19 字数: 535.04 千字

版 次: 2016 年 12 月第 1 版

印 次: 2016 年 12 月第 1 次印刷

定 价: 42.00 元

凡所购买电子工业出版社图书有缺损问题, 请向购买书店调换。若书店售缺, 请与本社发行部联系, 联系及邮购电话: (010) 88254888, 88258888。

质量投诉请发邮件至 [zltz@phei.com.cn](mailto:zltz@phei.com.cn), 盗版侵权举报请发邮件至 [dbqq@phei.com.cn](mailto:dbqq@phei.com.cn)。

本书咨询联系方式: [dcc@phei.com.cn](mailto:dcc@phei.com.cn), 192910558 (QQ 群)。

# 前言

## Preface

数据，已经渗透到当今每一个行业和业务职能领域，成为重要的生产要素。人们对于海量数据的挖掘和运用，预示着新一波生产率增长和消费者盈余浪潮的到来。在信息社会，数据已经成为重要的资源。大数据时代改变了人类原有的生活和发展模式，也改变了人类认识世界和价值判断的方式。以数据库技术为基础的数据管理技术，可以对数据进行有效的收集、加工、分析与处理，以释放更多的数据价值，充分发挥数据的作用。

数据库技术是计算机学科的一个重要分支，是各类计算机信息系统的核心技术和重要基础，所有与数据信息有关的业务及应用系统都需要数据库技术的支持。数据库知识是当今大学生信息素养的重要组成部分，数据库基础与应用课程是高等学校一门重要的计算机基础课程。通过对课程的学习，使读者能够准确理解数据库的基本概念、原理和方法，掌握数据库的各种操作技术，具备数据库设计与数据库应用系统开发的基本能力，为今后应用数据库技术管理信息、更好地利用信息打下基础。

Microsoft 公司推出的 Access 数据库管理系统是集成在 Office 套装软件中的一个组件，由于它具有界面友好、易学实用的特点，所以适用于中小型数据管理应用场合，既可以用作本地数据库，也可以应用于网络环境。从教学的角度讲，Access 很适合初学者理解和掌握数据库的概念与操作方法。随着 Microsoft Office 软件的不断更新，Access 先后形成了很多版本，新的版本除了继承和发扬以前版本的优点之外，还增加了许多新的功能，例如全新的界面、方便的模板，以及在功能方面的诸多改善等。本书以常用的 Access 2010 作为实践环境，介绍数据库的基本操作和应用开发技术。

近年来，国内有关 Access 数据库的教材出版了不少。本书力图改变将 Access 教学过分“工具化”的倾向，即不仅仅是将 Access 当成一种数据库操作工具，而是将 Access 教学从数据库工具操作层面上升到数据库应用系统设计与开发层面，体现数据库教学的更高要求，这和当下计算机教育界讨论的“计算思维”的思想是吻合的。在编写过程中，本书适度介绍关系数据库的基本概念和基本原理，体现数据库应用的基本要求，告诉学生如何从现实世界的数据库对象中抽象出可以存放到计算机中的数据，讲清数据库中的“表”是如何来的。全书共 10 章，主要内容有数据库基础知识、数据库的操作、表的操作、查询的操作、SQL 查询的操作、窗体的操作、报表的操作、宏的操作、模块与 VBA、应用案例等。全书以“教学管理”数据库的操作为主线，设计编排了大量的实例，便于读者学习和提高。

数据库基础与应用课程是一门实践性很强的课程，上机实验十分重要。只有通过不断上机实践，才能熟练掌握 Access 的基本操作，充分理解数据库技术的基本思想和方法，并将所学知识应用到系统开发中去。本书的附录部分是实验指导，其中设计了 12 个实验，每个实验包括实验目的、实验内容和实验思考等内容，“实验内容”包括适当的操作提示，以帮助读者完成操作练习；“实验思考”作为实

验内容的扩充,读者可以结合上机操作进行思考,可以根据实际情况从中选择部分内容作为上机练习。

本书既可作为高等学校数据库基础与应用课程的教材,又可供社会各类计算机应用人员与参加各类计算机等级考试的读者阅读参考。

本书的配套教学资源可以通过华信教育资源网 <http://www.hxedu.com.cn> 注册免费下载。书中一些重要的概念与操作方法还配套制作了微视频,读者可利用手机等移动智能终端扫描书中的二维码直接观看。

本书第1~5章、第10章由刘卫国编写,第6~9章及附录由蔡立燕编写,全书由刘卫国主编、定稿。此外,参与部分编写工作的还有熊拥军、王鹰、童键、舒卫真、刘泽星、奎晓燕、蔡旭晖、刘胤宏、周欣然、周克涛、孙士闯等。

由于作者学识水平有限,书中难免存在疏漏或不妥之处,恳请广大读者批评指正。

作者

2016年7月

# 目 录

## Contents

<b>第 1 章 数据库基础知识</b> ..... (1)	2.1.2 Access 2010 工作窗口..... (31)
1.1 数据管理技术..... (1)	2.1.3 Access 2010 数据库的组成 .. (37)
1.1.1 数据与数据管理..... (1)	2.2 数据库的创建与基本操作..... (37)
1.1.2 数据管理技术的变迁..... (2)	2.2.1 创建数据库的方法..... (38)
1.2 数据库系统..... (7)	2.2.2 数据库的基本操作..... (38)
1.2.1 数据库系统的组成..... (8)	2.2.3 更改默认数据库文件夹..... (40)
1.2.2 数据库系统的特点..... (10)	2.3 数据库对象的组织和管理..... (41)
1.3 数据模型..... (10)	2.3.1 导航窗格的操作..... (41)
1.3.1 数据模型的组成要素..... (11)	2.3.2 在导航窗格中对数据库对象 的操作..... (42)
1.3.2 数据抽象的过程..... (11)	2.3.3 数据库视图的切换..... (43)
1.3.3 概念模型..... (12)	2.4 数据库的维护..... (44)
1.3.4 逻辑模型..... (15)	2.4.1 数据库的备份与还原..... (44)
1.4 数据库的体系结构..... (17)	2.4.2 数据库的压缩与修复..... (44)
1.4.1 数据库的三级模式..... (18)	2.4.3 数据库的拆分..... (45)
1.4.2 三级模式间的二级映射..... (19)	2.5 数据库的安全保护..... (45)
1.5 关系数据库..... (19)	2.5.1 设置数据库密码..... (45)
1.5.1 关系的数学定义..... (19)	2.5.2 解密数据库..... (46)
1.5.2 关系运算..... (21)	2.5.3 启用数据库中禁用的内容 .. (46)
1.5.3 关系的完整性..... (23)	习题 2..... (47)
1.6 数据库的设计..... (24)	<b>第 3 章 表的操作</b> ..... (50)
1.6.1 数据库设计的基本步骤..... (24)	3.1 表结构的设计..... (50)
1.6.2 从概念模型到关系模型 的转换..... (25)	3.1.1 字段参数..... (50)
1.6.3 数据库设计实例..... (26)	3.1.2 字段的数据类型..... (51)
习题 1..... (27)	3.2 表的创建..... (53)
<b>第 2 章 数据库的操作</b> ..... (30)	3.2.1 创建表的方法..... (53)
2.1 Access 操作基础..... (30)	3.2.2 设置字段属性..... (57)
2.1.1 Access 2010 的安装、启动 与退出..... (30)	3.3 表中数据的输入..... (62)
	3.3.1 使用数据表视图输入数据 .. (62)



3.3.2	特殊类型字段的输入方法	(63)	4.6	创建操作查询	(104)
3.3.3	创建查阅列表字段	(63)	4.6.1	生成表查询	(104)
3.4	表之间的关联	(66)	4.6.2	删除查询	(105)
3.4.1	创建表之间的关联	(66)	4.6.3	更新查询	(106)
3.4.2	编辑表之间的关联	(67)	4.6.4	追加查询	(107)
3.4.3	设置参照完整性	(68)	习题 4		(109)
3.4.4	在主表中显示子数据表	(68)	<b>第 5 章 SQL 查询的操作</b>		(111)
3.5	表的维护	(69)	5.1	SQL 与 SQL 查询	(111)
3.5.1	表结构的修改	(69)	5.1.1	SQL 的发展与功能	(111)
3.5.2	表中内容的修改	(70)	5.1.2	SQL 视图	(112)
3.5.3	表中数据的查找与替换	(71)	5.1.3	创建 SQL 查询	(113)
3.5.4	表的修饰	(73)	5.2	SQL 数据查询	(113)
3.6	表的操作	(75)	5.2.1	基本查询	(113)
3.6.1	将表中的记录排序	(75)	5.2.2	嵌套查询	(117)
3.6.2	对表中的记录进行筛选	(76)	5.2.3	多表查询	(118)
3.6.3	对表中的行进行汇总统计	(79)	5.2.4	联合查询	(119)
习题 3		(80)	5.3	SQL 数据定义	(120)
<b>第 4 章 查询的操作</b>		(82)	5.3.1	建立表结构	(120)
4.1	查询概述	(82)	5.3.2	修改表结构	(122)
4.1.1	查询的功能	(82)	5.3.3	删除表	(122)
4.1.2	查询的类型	(83)	5.4	SQL 数据操纵	(122)
4.1.3	查询视图	(83)	5.4.1	插入记录	(122)
4.2	查询的条件	(84)	5.4.2	更新记录	(123)
4.2.1	Access 常量	(84)	5.4.3	删除记录	(123)
4.2.2	Access 常用函数	(85)	习题 5		(124)
4.2.3	Access 运算	(86)	<b>第 6 章 窗体的操作</b>		(127)
4.2.4	查询条件举例	(88)	6.1	窗体概述	(127)
4.3	创建选择查询	(88)	6.1.1	窗体的功能	(127)
4.3.1	使用查询向导创建选择查询	(89)	6.1.2	窗体的类型	(127)
4.3.2	在查询设计视图中创建选择 查询	(93)	6.1.3	窗体的视图	(128)
4.3.3	在查询中进行计算	(95)	6.2	创建窗体的方法	(129)
4.4	创建交叉表查询	(99)	6.3	在设计视图中创建窗体	(134)
4.4.1	使用交叉表查询向导创建 交叉表查询	(99)	6.3.1	窗体设计窗口	(134)
4.4.2	在查询设计视图中创建 交叉表查询	(101)	6.3.2	控件的功能与分类	(135)
4.5	创建参数查询	(102)	6.3.3	控件的操作	(137)
4.5.1	单参数查询	(102)	6.4	控件的应用	(141)
4.5.2	多参数查询	(103)	6.4.1	面向对象的基本概念	(141)
			6.4.2	窗体和控件的属性	(141)
			6.4.3	控件应用举例	(144)
			习题 6		(153)

<b>第 7 章 报表的操作</b> .....	(155)	<b>9.3 VBA 程序的流程控制</b> .....	(202)
7.1 报表概述.....	(155)	9.3.1 顺序控制.....	(202)
7.1.1 报表的功能.....	(155)	9.3.2 选择控制.....	(206)
7.1.2 报表的类型.....	(155)	9.3.3 循环控制.....	(209)
7.1.3 报表的视图.....	(156)	9.3.4 辅助控制.....	(212)
7.2 创建报表的方法.....	(156)	<b>9.4 VBA 过程</b> .....	(213)
7.3 使用设计视图创建报表.....	(160)	9.4.1 子过程与函数过程.....	(213)
7.3.1 报表设计窗口.....	(160)	9.4.2 过程参数传递.....	(215)
7.3.2 报表的修饰.....	(162)	9.4.3 变量的作用域和生存期.....	(217)
7.3.3 报表的外观设计.....	(163)	<b>9.5 VBA 数据库访问技术</b> .....	(218)
7.3.4 报表的打印.....	(163)	9.5.1 常用的数据库访问接口 技术.....	(218)
7.4 报表的高级设计.....	(165)	9.5.2 ADO 对象模型.....	(219)
7.4.1 报表统计计算.....	(165)	9.5.3 利用 ADO 访问数据库的 基本步骤.....	(220)
7.4.2 报表排序和分组.....	(166)	<b>9.6 VBA 程序的调试与错误处理</b> .....	(224)
7.4.3 创建子报表.....	(168)	9.6.1 VBA 程序的调试方法.....	(225)
习题 7.....	(170)	9.6.2 VBA 程序的错误处理.....	(226)
<b>第 8 章 宏的操作</b> .....	(172)	习题 9.....	(227)
8.1 宏概述.....	(172)	<b>第 10 章 应用案例</b> .....	(231)
8.1.1 宏的类型.....	(172)	10.1 数据库应用系统的开发过程.....	(231)
8.1.2 宏的操作界面.....	(173)	10.2 需求分析.....	(232)
8.1.3 常用的宏操作命令.....	(174)	10.3 系统设计.....	(232)
8.2 宏的创建.....	(175)	10.3.1 数据库设计.....	(232)
8.2.1 创建独立的宏.....	(175)	10.3.2 系统功能设计.....	(236)
8.2.2 创建嵌入的宏.....	(180)	10.4 系统实现.....	(238)
8.2.3 创建数据宏.....	(181)	10.4.1 创建数据库.....	(238)
8.3 宏的运行与调试.....	(182)	10.4.2 创建窗体.....	(238)
8.3.1 宏的运行.....	(182)	10.4.3 创建查询.....	(242)
8.3.2 宏的调试.....	(184)	10.4.4 创建报表.....	(246)
8.4 宏的应用.....	(184)	10.5 应用系统的集成.....	(248)
习题 8.....	(188)	10.5.1 创建切换面板.....	(248)
<b>第 9 章 模块与 VBA</b> .....	(190)	10.5.2 设置数据库启动选项.....	(252)
9.1 模块与 VBA 概述.....	(190)	习题 10.....	(252)
9.1.1 模块的概念.....	(190)	<b>附录 A 实验指导</b> .....	(254)
9.1.2 VBA 的开发环境.....	(191)	实验 1 Access 2010 操作基础.....	(254)
9.1.3 模块的创建.....	(192)	实验 2 数据库的操作.....	(256)
9.2 VBA 程序的数据描述.....	(195)	实验 3 表的操作.....	(257)
9.2.1 数据类型.....	(195)	实验 4 查询的操作.....	(262)
9.2.2 常量与变量.....	(196)	实验 5 SQL 查询的操作.....	(263)
9.2.3 内部函数.....	(198)		
9.2.4 表达式.....	(201)		





实验 6 窗体的创建 .....	(265)	实验 11 VBA 对象与数据库访问	
实验 7 窗体控件的应用 .....	(267)	技术 .....	(276)
实验 8 报表的操作 .....	(269)	实验 12 数据库应用系统开发 .....	(285)
实验 9 宏的操作 .....	(271)	参考文献 .....	(295)
实验 10 VBA 程序设计基础 .....	(274)		

# 第 1 章 数据库基础知识

在信息社会,数据已经成为重要的资源。大数据时代改变了人类原有的生活和发展模式,也改变了人类认识世界和价值判断的方式。以数据库技术为基础的数据管理技术,可以对数据进行有效的收集、加工、分析与处理,使得释放更多数据价值,充分发挥数据的作用。数据库技术是计算机学科的一个重要分支,是各类计算机信息系统的核心技术和重要基础,所有与数据信息有关的业务及应用系统都需要数据库技术的支持。数据库知识是当今大学生信息素养的重要组成部分,Microsoft 公司推出的 Access 数据库管理系统是集成在 Office 套装软件中的一个组件,它具有自身的特点和应用场合。从教学的角度讲,Access 很适合于初学者理解和掌握数据库的概念与操作方法。本书以常用的 Access 2010 作为实践环境。

本章围绕数据库系统的基础知识而展开。通过本章的学习,要掌握数据管理的概念、数据管理技术的发展、数据库的概念、数据库系统的组成与特点、数据模型的概念、数据库的体系结构、关系数据库的基本知识及数据库的设计方法等内容,这些内容对学习掌握 Access 2010 数据库操作及应用开发是十分必要的。

## 1.1 数据管理技术

人类在长期的社会实践中会产生大量数据,如何对数据进行高效的处理和应用成为迫切的现实需要。只有在计算机成为数据处理的工具之后,才使数据处理现代化成为可能。

### 1.1.1 数据与数据管理

#### 1. 数据和信息

数据(Data)和信息(Information)是数据处理中的两个基本概念,有时可以混用,如平时讲数据处理就是信息处理,但有时必须分清。一般认为,数据是人们用于记录事物情况的物理符号。为了描述客观事物而用到的数字、字符,以及所有能输入到计算机中并能被计算机处理的符号都可以看成数据。例如,章达明老师的基本工资为 2750 元,职称为教授,这里的“章达明”、“2750”、“教授”就是数据。在实际应用中,有两种基本形式的数据,一种是可以参与数值运算的数值型数据,如表示工资、成绩的数据;另一种是由字符组成、不能参与数值运算的字符型数据,如表示姓名、职称的数据。此外,还有图形、图像、声音等多媒体数据,如人的照片、商品的商标等。

信息是数据中所包含的意义。通俗地讲,信息是经过加工处理并对人类社会实践和生产活动产生决策影响的数据。不经过加工处理的数据只是一种原始材料,对人类活动产生不了决策作用,它的价值只是在于记录了客观世界的事实。只有经过提炼和加工,原始数据才发生了质的变化,给人们以新的知识和智慧。例如,收到一条微信,“有聚划算活动,商品 5 折”,就会根据数据“5 折”,再参照商品原价格(原始数据),计算出打折后的商品价格(新的数据),新的数据包含了新的信息,即商品降价了,可作为是否购买商品的依据。



数据与信息既有区别，又有联系。一方面，数据是信息的载体，但并非任何数据都能成为信息，只有经过加工处理之后具有新的内容的数据才能成为信息。另一方面，信息不随表示它的数据形式的改变而改变，它是反映客观现实世界的知识，而数据则具有任意性，用不同的数据形式可以表示同样的信息。例如，一个城市的天气预报情况是一条信息，而描述该信息的数据形式可以是文字、图像或声音等。

## 2. 数据处理和数据管理

数据处理是指将数据转换成信息的过程，它包括对数据的收集、存储、分类、计算、加工、检索和传输等一系列活动。其基本目的是从大量的、杂乱无章的、难以理解的数据中整理出对人们有价值、有意义的数据（信息），从而作为决策的依据。长期以来，人类主要用人工进行数据处理工作，使用计算机才实现了数据处理的自动化，才使数据处理的速度更快、效率更高。没有计算机，就不会有现代数据处理技术的形成和发展。从这个意义上讲，计算机是一个具有程序执行能力的数据处理工具，如图 1-1 所示。计算机数据处理所得到的输出数据（信息），除了取决于输入数据外，还取决于程序。程序不同，完成的数据处理方法不同，得到的结果也不同，即包含的信息也就不同。计算机通过不同的程序完成不同的数据处理任务。

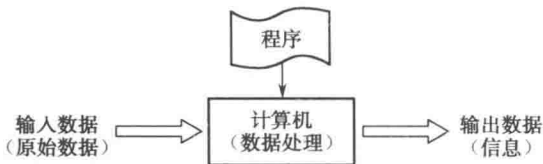


图 1-1 计算机数据处理

以高考成绩处理系统为例，全体考生各门课程的考试成绩记录了考生的考试情况，属于原始数据，对考试成绩进行分析和处理，如按成绩从高到低顺序排列、统计各分数段的人数等，进而可以根据招生人数确定录取分数线，输出的数据即包含了丰富的信息。

数据管理是指数据的收集、组织、存储、检索和维护等操作，这些操作是数据处理的中心环节，是任何数据处理业务中不可缺少的部分。数据管理的基本目的是为了实现在数据共享、降低数据冗余、提高数据的独立性、安全性和完整性，从而能更加有效地管理和使用数据资源。

### 1.1.2 数据管理技术的变迁

计算机技术的发展和数据处理的现实需要，促使数据管理技术得到了很大发展，从而有效地提高了数据处理的应用水平。计算机数据管理技术经历了人工管理、文件管理和数据库管理 3 个发展阶段。

#### 1. 人工管理阶段

20 世纪 50 年代中期以前，计算机主要应用于科学计算，虽然当时也有数据管理的问题，但当时的数据管理是以人工管理方式进行的。在硬件方面，外存储器只有磁带、卡片和纸带等，没有磁盘等直接存取的外存储器；在软件方面，只有汇编语言，没有操作系统，没有对数据进行管理的软件。在人工管理阶段，数据管理有以下特点。

##### (1) 数据不保存

人工管理阶段处理的数据量较少，一般不需要将数据长期保存，只是在计算时将数据随应用程序一起输入，计算完后将结果输出，数据和应用程序一起从内存中被释放。若要再次进行计算，则需重新输入数据和应用程序。

### (2) 由应用程序管理数据

系统没有专用的软件对数据进行管理，数据需要由应用程序自行管理。每个应用程序不仅要规定数据的逻辑结构，而且要设计数据的存储结构及输入/输出方法等，程序设计任务繁重。

### (3) 数据有冗余，无法实现共享

应用程序与数据是一个整体，一个应用程序中的数据无法被其他应用程序使用，因此，应用程序与应用程序之间存在大量的重复数据，数据无法实现共享。

### (4) 数据对应用程序不具有独立性

由于应用程序对数据的依赖性，数据的逻辑结构或存储结构一旦有所改变，则必须修改相应的应用程序，这就进一步加重了程序设计的负担。

以一所学校的信息管理为例，在人工管理阶段，应用程序和数据之间的关系如图 1-2 所示。图中各类数据对应于相应的应用程序，而且不同种类的数据之间会有许多重复的数据，如教务数据可能包含教师和学生的部分数据。

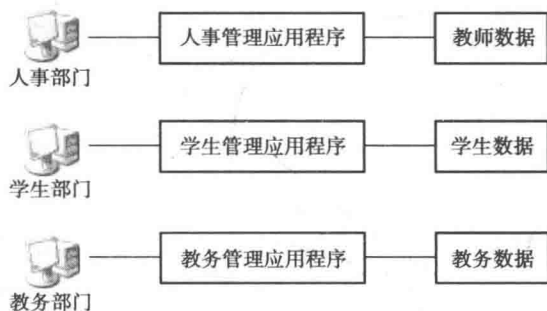


图 1-2 人工管理阶段应用程序和数据之间的关系

## 2. 文件管理阶段

20 世纪 50 年代后期至 60 年代后期，计算机开始大量用于数据管理。在硬件方面，出现了直接存取的大容量外存储器，如磁盘、磁鼓等，这为计算机数据管理提供了物质基础；在软件方面，出现了高级语言和操作系统。操作系统的出现为数据管理提供了技术支持。

应用程序利用操作系统的文件管理功能（由文件系统实现），将相关数据按一定的规则构成文件，通过文件系统对文件中的数据进行存取和管理，实现数据的文件管理方式。其特点有以下两个方面。

### (1) 数据可以长期保存

文件系统为应用程序和数据之间提供了一个公共接口，使应用程序采用统一的存取方法来存取和操作数据。数据可以组织成文件，能够长期保存、反复使用。

### (2) 数据对应用程序有一定的独立性

应用程序和数据不再是一个整体，而是通过文件系统把数据组织成一个独立的数据文件，由文件系统对数据的存取进行管理。程序员只需通过文件名来访问数据文件，不必过多考虑数据的物理存储细节，因此，程序员可集中精力进行算法设计，并大大减少了应用程序维护的工作量。

文件管理使计算机在数据管理方面有了长足的进步，时至今日，文件系统仍是高级语言普遍采用的数据管理方式。然而，当数据量增加、使用数据的用户越来越多时，文件管理便不能适应更有效地使用数据的需要了，其症结表现在以下 3 个方面。

① 数据的共享性差、冗余度大，容易造成数据不一致。由于数据文件是根据应用程序的需要而建立的，当不同的应用程序所使用的数据有相同部分时，也必须建立各自的数据文件，即数据不能共享，



造成大量数据重复。这样不仅浪费存储空间，而且使数据修改变得非常困难，容易产生数据不一致等问题，即同样的数据在不同的文件中所存储的数值不同，造成矛盾。

② 数据独立性差。在文件系统中，尽管数据和应用程序有一定的独立性，但这种独立性主要是针对某一特定应用而言的，就整个应用系统而言，文件系统还未能彻底体现数据逻辑结构独立于数据存储的物理结构的要求。在文件系统中，数据和应用程序是互相依赖的，即应用程序的编写与数据组织方式有关，如果改变数据的组织方式，就必须修改应用程序；而应用程序发生变化，如改用另一种程序设计语言来编写应用程序，也必须修改文件的数据结构。

③ 数据之间缺乏有机的联系，缺乏对数据的统一控制和管理。文件系统中各数据文件之间是相互独立的，没有从整体上反映现实世界事物之间的内在联系，因此，很难对数据进行合理的组织以适应不同应用的需要。在同一个应用项目中的各个数据文件没有统一的管理机构，数据完整性和安全性很难得到保证。

在文件管理阶段，学校信息管理中应用程序和数据文件之间的关系如图 1-3 所示。显然，教务数据中有教师、学生和课程等数据，数据冗余是难免的。各应用程序通过文件系统对相应的数据文件进行存取和处理，但各个数据文件基本上对应于相关的应用程序，而且各数据文件之间是孤立的，数据之间的联系无法体现，例如，因某个学生转学而在学生数据文件中删除了其数据，此时无法在教务数据文件中实现联动操作，因此仍然存在缺陷。

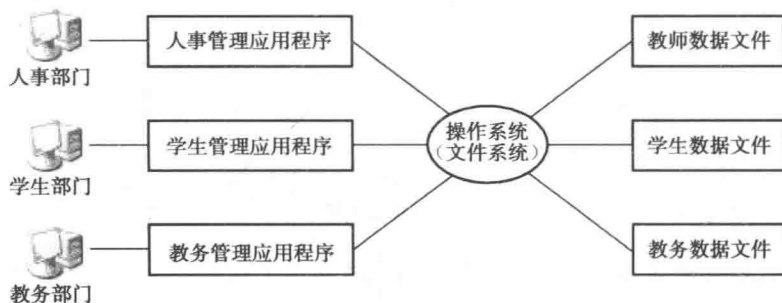


图 1-3 文件管理阶段应用程序和数据文件之间的关系

### 3. 数据库管理阶段

20 世纪 60 年代后期，计算机用于数据管理的规模更加庞大，数据量急剧增加，数据共享性要求更加强烈。同时，计算机硬件价格下降，而软件价格上升，编制和维护软件所需的成本相对增加，其中维护成本更高。这些成为计算机数据管理技术发展的原动力。

数据库 (Database, DB) 是按一定的组织方式存储起来的、相互关联的数据集合。在数据库管理阶段，由一种称为数据库管理系统 (Database Management System, DBMS) 的系统软件来对数据进行统一的控制和管理。数据库管理系统把所有应用程序中使用的相关数据汇集起来，按统一的数据模型存储在数据库中，为各个应用程序所使用。在应用程序和数据库之间保持较高的独立性，数据具有完整性、一致性和安全性高等特点，并且具有充分的共享性，有效地减少了数据冗余。

在数据库管理阶段，学校信息管理中应用程序和数据库之间的关系如图 1-4 所示。有关学校信息管理的数据都存放在一个统一的数据库中，数据库不再面向某个部门的应用，而是面向整个应用系统，实现了数据共享，并且数据库和应用程序之间保持较高的独立性。

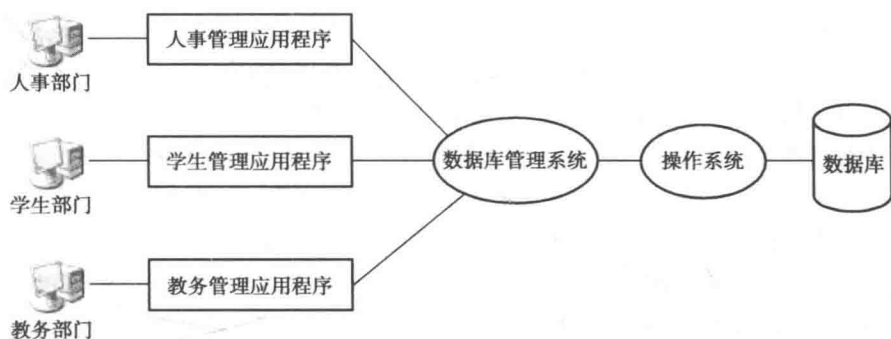


图 1-4 数据库管理阶段应用程序和数据库之间的关系

#### 4. 数据管理技术新发展

数据库技术的发展先后经历了层次数据库、网状数据库和关系数据库。层次数据库和网状数据库可以看作第 1 代数据库系统，关系数据库可以看作第 2 代数据库系统。自 20 世纪 70 年代提出关系数据模型和关系数据库后，数据库技术得到了蓬勃发展，应用也越来越广泛。但随着应用的不断深入，占主导地位的关系数据库系统已不能满足新的应用领域的需求。例如，在实际应用中，除了需要处理数字、字符数据的简单应用之外，还需要存储并检索复杂的复合数据（如集合、数组、结构）、多媒体数据、计算机辅助设计绘制的工程图纸和地理信息系统（Geographic Information System, GIS）提供的空间数据等，对于这些复杂数据，关系数据库无法实现对它们的管理。正是实际应用中涌现出的许多问题，促使数据管理技术不断向前发展，出现了许多不同类型的新型数据管理技术。下面概要性地做一些介绍。

##### （1）分布式数据库系统

分布式数据库系统（Distributed Database System, DDBS）是数据库技术与计算机网络技术、分布式处理技术相结合的产物。分布式数据库系统是系统中的数据地理上分布在计算机网络的不同结点上，但逻辑上属于一个整体的数据库系统。分布式数据库系统不同于将数据存储在服务器上供用户共享存取的网络数据库系统，它不仅能支持局部应用（访问本地数据库），而且能支持全局应用（访问异地数据库）。

分布式数据库系统的主要特点有以下 3 个方面。

① 数据是分布的：数据库中的数据分布在计算机网络的不同结点上，而不是集中在一个结点，区别于数据存放在服务器上由各用户共享的网络数据库系统。

② 数据是逻辑相关的：分布在不同结点的数据逻辑上属于同一数据库系统，数据间存在相互关联，区别于由计算机网络连接的多个独立的数据库系统。

③ 结点的自治性：每个结点都有自己的计算机软硬件资源，包括数据库、数据库管理系统等，因而能够独立地管理局部数据库。局部数据库中的数据可以仅供本结点用户存取使用，也可供其他结点上的用户存取使用，提供全局应用。

##### （2）面向对象数据库系统

面向对象数据库系统（Object Oriented Database System, OODBs）是将面向对象的模型、方法和机制，与先进的数据库技术有机地结合而形成的新型数据库系统。它从关系模型中脱离出来，强调在数据库框架中发展类型、数据抽象、继承和持久性。面向对象数据库系统的基本设计思想是：一方面把面向对象的程序设计语言向数据库方向扩展，使应用程序能够存取并处理对象；另一方面扩展数据库系统，使其具有面向对象的特征，提供一种综合的语义数据建模概念集，以便对现实世界中复杂应



用的实体和联系建模。因此,面向对象数据库系统首先是一个数据库系统,具备数据库系统的基本功能,其次是一个面向对象的系统,针对面向对象的程序设计语言的永久性对象存储管理而设计,充分支持完整的面向对象概念和机制。

### (3) 多媒体数据库系统

多媒体数据库系统 (Multimedia Database System, MDBS) 是数据库技术与多媒体技术相结合的产物。随着信息技术的发展,数据库应用从传统的企业信息管理扩展到计算机辅助设计 (Computer Aided Design, CAD)、计算机辅助制造 (Computer Aided Manufacture, CAM)、办公自动化 (Office Automation, OA)、人工智能 (Artificial Intelligent, AI) 等多种应用领域。这些领域中要求处理的数据不仅包括传统的数字、字符等格式化数据,还包括大量多种媒体形式的非格式化数据,如图形、图像、声音等。这种能存储和管理多种媒体数据的数据库称为多媒体数据库。

多媒体数据库的结构和操作与传统格式化数据库的结构和操作有很大差别。现有数据库管理系统无论是在模型的语义描述能力、系统功能、数据操作上,还是在存储管理、存储方法上,都不能适应非格式化数据的处理要求。综合程序设计语言、人工智能和数据库领域的研究成果,设计支持多媒体数据管理的数据库管理系统已成为数据库领域中一个新的重要研究方向。

在多媒体信息管理环境中,不仅数据本身的结构和存储形式各不相同,而且不同领域对数据处理的要求也比一般事务管理复杂得多,因而对数据库管理系统提出了更高的功能要求。

### (4) 数据仓库技术

随着信息技术的高速发展,数据库应用的规模、范围和深度不断扩大,一般的事务处理已不能满足应用的需要,企业界需要在大量数据基础上的决策支持,数据仓库 (Data Warehouse, DW) 技术的兴起满足了这一需求。数据仓库作为决策支持系统 (Decision Support System, DSS) 的有效解决方案,涉及 3 方面的技术内容:数据仓库技术、联机分析处理 (On-Line Analysis Processing, OLAP) 技术和数据挖掘 (Data Mining, DM) 技术。

数据仓库、联机分析处理和数据挖掘是作为 3 种独立的数据处理技术出现的。数据仓库用于数据的存储和组织,联机分析处理集中于数据的分析,数据挖掘则致力于知识的自动发现。它们都可以分别应用到信息系统的设计和实现中,以提高相应部分的处理能力。但是,由于这 3 种技术内在的联系性和互补性,将它们结合起来就是一种新的决策支持系统架构。这一架构以数据库中的大量数据为基础,系统由数据驱动。

### (5) 大数据技术

大数据 (Big Data) 是规模非常巨大和复杂的数据集,传统数据库管理工具对其处理起来面临很多困难,如对数据库高并发读/写要求、对海量数据的高效率存储和访问需求、对数据库高可扩展性和高可用性的需求。

大数据有 4 个基本特征:数据规模大 (Volume)、数据种类多 (Variety)、要求数据处理速度快 (Velocity)、数据价值密度低 (Value),即所谓的 4V 特性。这些特性使得大数据区别于传统的数据概念。大数据的概念与海量数据不同,后者只强调数据的量,而大数据不仅用来描述大量的数据,还进一步指出数据的复杂形式、数据的快速处理特性,以及对数据分析处理后最终获得有价值信息的能力。

① 数据规模大:大数据聚合在一起的数据量是非常大的,根据国际数据公司 IDC 的定义至少要有超过 100TB 的可供分析的数据,数据量大是大数据的基本属性。

② 数据种类多:数据类型繁多、格式复杂是大数据的重要特性。数据来自多种数据源,非结构化数据大量涌现。由于非结构化数据没有统一的结构属性,难以用表结构来表示,在记录数据数值的同时还需要存储数据的结构,增加了数据存储和处理的难度。

③ 要求数据处理速度快：要求数据的快速处理是大数据区别于传统海量数据处理的重要特性之一。对于大数据而言，许多应用都要求能够实时处理，例如，有大量在线交互的电子商务应用，就具有很强的时效性。在这种情况下，大数据要求快速、持续的实时处理。

④ 数据价值密度低：数据价值密度低是大数据关注的非结构化数据的重要属性。传统的结构化数据，依据特定的应用，对事物进行了相应的抽象，每一条数据都包含该应用需要考虑的信息，而大数据为了获取事物的全部细节，不对事物进行抽象、归纳等处理，直接采用原始的数据，保留了全部数据，从而可以分析更多的信息，但也引入了大量没有意义的信息，这样就使得一方面数据的绝对数量激增，另一方面数据包含有效信息量的比例不断减少，数据价值密度偏低。

大数据处理技术就是从各种类型的数据中快速获得有价值信息的技术。大数据本质也是数据，但它包括非结构化数据，为此给数据管理技术带来了以下新的挑战。

① 大数据的表示面临挑战。用统一的模型对非结构化数据进行分析处理非常困难，传统的数据表示方法不能直观地展现数据本身的含义。为了有效利用数据并挖掘其中的知识，必须寻找最合适且有效的数据表示方法。目前使用的方法是数据标识，标识方法可减轻数据识别和分类的困难，但却给用户增添了预处理工作量。研究既有效又简易的数据表示方法是进行大数据处理首先面临的技术难题之一。

② 数据量的成倍增长给数据存储能力带来了挑战。传统的数据库追求高度的数据一致性和容错性，缺乏较强的扩展性和较好的系统可用性，不能有效存储视频、音频等非结构化和半结构化的数据。大数据及其潜在的商业价值要求使用专门的数据处理技术和专用的数据存储设备。目前，数据存储能力的增长远远赶不上数据的增长，设计最合理的分层存储架构成为信息系统的关键。

③ 数据融合技术面临挑战。数据不整合就发挥不出大数据的巨大价值，大数据面临的一个重要问题是个人、企业和政府机构的各种数据和信息能否方便地融合。如同人类有许多种自然语言一样，作为网络空间中唯一客观存在的数据难免有多种格式，但为了清除大数据处理的障碍，应研究推广不与平台绑定的数据格式。大数据已成为联系人类社会、物理世界和网络空间的纽带，需要通过统一的数据格式构建融合人、机、物三元世界的统一信息系统。

④ 数据跨越组织边界传播，给信息安全带来了巨大挑战。随着技术的发展，大量信息跨越组织边界传播，信息安全问题相伴而生，不仅使没有价值的信息大量出现，保密数据、隐私数据也成倍增长。国家安全、知识产权、个人信息等都面临着前所未有的安全挑战。大数据时代，犯罪分子获取信息更加容易，人们防范、打击犯罪行为更加困难，这对数据存储的物理安全性及数据的多副本与容灾机制提出了更高的要求。想要应对瞬息万变的安全问题，最关键的是算法和特征，如何建立相应的强大安全防御体系来发现和识别安全漏洞是保证信息安全的重要环节。

目前，围绕大数据，一批新兴的数据挖掘、数据存储、数据处理与分析技术不断涌现，使得人们能够将隐藏在海量数据中的信息和知识挖掘出来，从而为人类的社会经济活动提供决策依据。大数据将在商业智能、政府决策、公共服务等领域得到广泛应用。

## 1.2 数据库系统

数据库系统(Database System, DBS)是指基于数据库的计算机应用系统。和一般的应用系统相比，数据库系统有其自身的特点，它将涉及一些相互联系而又有所区别的基本概念。





## 1.2.1 数据库系统的组成

数据库系统是一个用来为用户提供信息服务的计算机应用系统。它通常由计算机硬件、软件、数据库和有关人员 4 部分组成，它们的逻辑关系如图 1-5 所示。

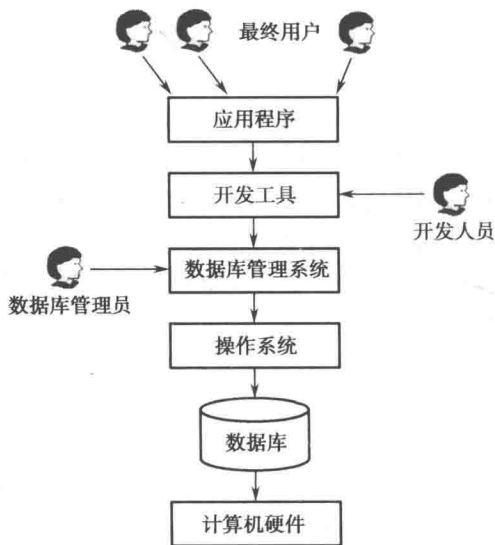


图 1-5 数据库系统

### 1. 计算机硬件系统

计算机硬件系统是数据库系统的物质基础，是存储数据库及运行相关软件的硬件设备，主要包括中央处理器（CPU）、存储设备、输入/输出设备及计算机网络环境。

### 2. 计算机软件系统

计算机软件系统包括操作系统、数据库管理系统、系统开发工具及数据库应用程序等。

#### (1) 操作系统

操作系统是所有软件的核心和基础，是其他软件运行的环境和平台。在计算机硬件层之上，由操作系统统一管理计算机的资源。

#### (2) 数据库管理系统

数据库管理系统在操作系统的支持下工作，是数据库系统的核心软件。常见的数据库管理系统有 Access、Visual FoxPro、SQL Server、Oracle、Sybase 等。

数据库管理系统是用户与数据库的接口，它可以实现数据的组织、存储和管理，提供访问数据库的方法，包括数据库的建立、查询、更新及各种数据控制等。数据库管理系统具有以下几个方面的功能。

① 数据定义功能。数据库管理系统提供数据定义语言（Data Definition Language, DDL），通过它可以方便地对数据库中的数据对象进行定义。

② 数据操纵功能。数据库管理系统提供数据操纵语言（Data Manipulation Language, DML），使用它可以实现对数据库中数据的基本操作，如修改、插入、删除和查询等。

③ 数据库运行管理功能。数据库管理系统通过对数据库的控制以适应共享数据的环境，确保数据库数据正确有效和数据库系统的正常运行。对数据库的控制主要通过以下 4 个方面实现。

● 数据的安全性控制：是指防止数据库的非法使用造成数据的泄露、更改或破坏。例如，系统提