



教育部大学计算机课程改革项目规划教材

丛书主编 卢湘鸿

数据库应用案例教程

黄都培 主编



清华大学出版社

教育部大学计算机课程改革项目规划教材

丛书主编 卢湘鸿

数据库应用案例教程

黄都培 主编

清华大学出版社

北京

内 容 简 介

本教材主要面向大学非计算机专业学生,故在教材的内容、组织和案例的选取上都做了精心策划。教材选择了一个文科专业学生比较熟悉并可能参与过的大学生计算机设计竞赛为案例模板,实际设计并开发了一套“高校学生竞赛管理系统”的数据库系统应用案例,开篇以任务驱动的方式将整个案例分解、贯穿于各章节讲解之中,每一章开头均针对实际案例给出了“任务描述、知识铺垫、解决思路、能力训练”的要点,不仅讲述了数据库相关的概念、内容和方法,而且通过贯穿各章的综合应用示例使学生能够边学习边实践,并给出具体的操作指导。全书共分9章,由浅入深地介绍了数据库的概念、功能和发展;关系数据库及其应用系统的设计流程和方法;Access 2007—2013的环境和基本操作;通用数据库结构查询语言与使用;Access数据库对象(表、查询、窗体、报表、宏、模块)的功能和应用。此外,本书还详细讲解了设计和开发一个实际数据库应用系统的具体方法。

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

数据库应用案例教程/黄都培主编. —北京:清华大学出版社,2015

教育部大学计算机课程改革项目规划教材

ISBN 978-7-302-38584-4

I. ①数… II. ①黄… 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(2014)第273625号

责任编辑:谢琛

封面设计:常雪影

责任校对:白蕾

责任印制:王静怡

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>, 010-62795954

印装者:三河市李旗庄少明印装厂

经 销:全国新华书店

开 本:185mm×260mm

印 张:18.5

字 数:445千字

版 次:2015年1月第1版

印 次:2015年1月第1次印刷

印 数:1~2500

定 价:32.00元

产品编号:061574-01

序

以计算机为核心的信息技术的应用能力已成为衡量一个人文化素质高低的重要标志之一。

大学非计算机专业开设计算机课程的主要目的是掌握计算机应用的能力以及在应用计算机过程中自然形成的包括计算思维意识在内的科学思维意识,以满足社会就业需要、专业需要与创新创业人才培养的需要。

根据《教育部关于全面提高高等教育质量的若干意见》(教高[2012]4号)精神,着力提升大学生信息素养和应用能力,推动计算机在面向应用的过程中培养文科学生的计算思维能力的文科大学计算机课程改革、落实由教育部高等教育司组织制订、教育部高等学校文科计算机基础教学指导委员会编写的高等学校文科类专业《大学计算机教学要求(第6版——2011年版)》(下面简称《教学要求》),在建立大学计算机知识体系结构的基础上,清华大学出版社依据教高司函[2012]188号文件中的部级项目1-3(基于计算思维培养的文科类大学计算机课程研究)、2-14(基于计算思维的人文类大学计算机系列课程及教材建设)、2-17(计算机艺术设计课程与教材创新研究)、2-18(音乐类院校计算机应用专业课程与专业基础课程系列化教材建设)的要求,组织编写、出版了本系列教材。

信息技术与文科类专业的相互结合、交叉、渗透,是现代科学技术发展趋势的重要方面,是新学科的一个不可忽视的生长点。加强文科类专业(包括文史法教类、经济管理类与艺术类)专业的计算机教育、开设具有专业特色的计算机课程是培养能够满足信息化社会对文科人才要求的重要举措,是培养跨学科、复合型、应用型的文科通才的重要环节。

《教学要求》把大文科的计算机教学,按专业门类分为文史法教类(人文类)、经济管理类与艺术类等三个系列。大文科计算机教学知识体系由计算机软硬件基础、办公信息处理、多媒体技术、计算机网络、数据库技术、程序设计、美术与设计类计算机应用以及音乐类计算机应用等8个知识领域组成。知识领域分为若干知识单元,知识单元再分为若干知识点。

大文科各专业对计算机知识点的需求是相对稳定、相对有限的。由属于一个或多个知识领域的知识点构成的课程则是不稳定、相对活跃、难以穷尽的。课程若按教学层次可分为计算机大公共课程(也就是大学计算机公共基础课程)、计算机小公共课程和计算机背景专业课程等三个层次。

第一层次的教学内容是文科各专业学生应知应会的。这些内容可为文科学生在与专业紧密结合的信息技术应用方面进一步深入学习打下基础。这一层次的教学内容是对文科大学生信息素质培养的基本保证,起着基础性与先导性的作用。

第二层次是在第一层次之上,为满足同一系列某些专业共同需要(包括与专业相结合而不是某个专业所特有的)而开设的计算机课程。其教学内容,或者在深度上超过第一层次的

教学内容中的某一相应模块,或者拓展到第一层次中没有涉及的领域。这是满足大文科不同专业对计算机应用需要的课程。这部分教学内容在更大程度上决定了学生在其专业中应用计算机解决问题的能力与水平。

第三层次,也就是使用计算机工具,以计算机软硬件为背景而开设的为某一专业所特有的课程。其教学内容就是专业课。如果没有计算机作为工具支撑,这门课就开不起来。这部分教学内容显示了学校开设特色专业课的能力与水平。

这些课程,除了大学计算机应用基础,还涉及数字媒体、数据库、程序设计以及与文史哲法教类、经济管理类与艺术类相关的许多课程。通过这些课程的开设,是让学生掌握更多的计算机应用能力,在计算机面向应用过程中培养学生的计算思维及更加宽泛的科学思维能力。

清华大学出版社出版的这套教育部部级项目规划教材,就是根据教高司函[2012]188号文件及《教学要求》的基本精神编写而成的。它可以满足当前大文科各类专业计算机各层次教学的基本需要。

对教材中的不足或错误,敬请同行和读者批评指正。

卢湘鸿

2014年10月于北京中关村科技园

卢湘鸿 北京语言大学信息科学学院计算机科学与技术系教授,原教育部高等学校文科计算机基础教学指导分委员会副主任、秘书长,现任教育部高等学校文科计算机基础教学指导分委员会顾问、全国高等院校计算机基础教育研究会文科专业委员会常务副主任兼秘书长,30多年来一直从事非计算机专业的计算机教育研究。主要成果不

前言

学习和掌握信息技术并利用计算机获取信息为本专业应用服务,已经成为当今社会对人才培养的基本要求和目标。当今信息社会对大学生信息素养与能力培养的宗旨是:信息社会中的每一个大学生都应能够在浩瀚的信息海洋中有效地检索、发现有用的信息,并进行适当的分析和处理,使获取的信息在自己的学习、工作和生活中发挥作用。

本着上述目标和宗旨,教育部高校文科专业计算机教学指导委员会多年来致力于高校非计算机专业,特别是文科专业计算机信息技术教育研究,为全国各高校文科类专业的计算机教学系统化、规范化给出了有效的指导。本教材即是依据《高等学校文科类专业计算机教学基本要求(2013版)》,面向计算机应用(小公共)课程教学所编写。

教材以计算思维为导航,以面向创新、面向应用为准则。教材编写组成员长期坚持在计算机教学第一线,对数据库教学做过较深入的研究和探索,大部分内容依据多年教学讲授和实践总结形成。考虑到文科学生接受知识的思维习惯,教材打破通常教科书的编写思路和格局,尽可能减少纯理论、技术、原理性的内容介绍,选择了一个文科专业学生比较熟悉并可能参与过的大学生计算机设计竞赛为开发模板,实际设计并开发了一套“高校学生竞赛管理系统”的数据库系统应用案例,以任务驱动的方式将整个案例分解、贯穿于各章节讲解之中,每一章开头均针对实际案例给出了“任务描述、知识铺垫、解决思路、能力训练”的要点,并结合案例给出大量可操作的示例和操作指导,使学生处在一个从感性认识—知识理解—实践应用的学习环境之中,由浅入深地掌握数据库的基础知识和开发方法,学会运用相关知识和技术设计和开发实际管理业务数据库系统,提高自己的计算机应用和创新能力。另外,从学生就业的需要出发,教材还结合并参考全国计算机等级考试的内容作了一定的补充。针对文科各专业学生的应用需求,在教材的组织、内容的取舍、讲解的侧重、案例的选取以及编写的方式上做了精心策划和认真安排。在教材编写中尽量体现:贴切对象,突出特点;瞄准前沿,面向应用;概念清晰,指导详尽;示例典范,通俗易懂;讲练结合,操作性强。力求将较先进的计算机数据库技术及其信息处理方法展现给读者,循序渐进地带领读者走进数据库的世界。

参加本教材编写的老师有:中国人民大学的战疆,中国政法大学的李激、王宝珠、宗恒,全书由黄都培审改并统稿。

作者

2014年10月

目 录

第 1 章 数据库——信息处理与管理的好助手	1
1.1 信息处理与数据库	1
1.1.1 数据、信息与信息处理	1
1.1.2 数据库能帮助我们做什么	3
1.1.3 数据管理技术的发展	4
1.1.4 认识和了解数据库	5
1.2 数据模型——数据库设计基础	7
1.2.1 用数据模型来组织数据	7
1.2.2 概念模型与 E-R 图设计	8
1.2.3 基于关系的数据逻辑模型	10
1.2.4 关系模型的优化与规范化设计	12
1.3 数据库案例分析与设计	14
1.3.1 实际管理中数据库设计步骤	14
1.3.2 管理业务的需求分析	15
1.3.3 “高校学生竞赛管理系统”分析与设计	16
本章小结	19
习题与操作练习	19
第 2 章 Access——优秀的关系数据库管理工具	22
2.1 认识和了解 Access	22
2.2 Access(2007-2013 通用版)的操作环境	23
2.2.1 Access 系统环境	24
2.2.2 Access 数据库操作环境	25
2.3 Access 的基本操作	29
2.3.1 创建数据库	29
2.3.2 保存并发布数据库	32
2.3.3 对数据库对象的常用操作	33
2.3.4 数据的导入和导出	36
2.3.5 数据库维护工具	40

2.4 Access 面向对象系统设计与开发	41
2.4.1 面向对象设计的基本概念	42
2.4.2 为对象设置属性	43
2.5 表达式及其使用	44
2.5.1 表达式	44
2.5.2 表达式生成器	48
本章小结	51
习题与操作练习	52
第3章 表——关系数据库的基础与核心	55
3.1 如何把实际管理的数据送到数据库中	55
3.1.1 设计 Access 的数据库表	55
3.1.2 表的各种视图	58
3.1.3 创建表	60
3.1.4 设计字段的属性	63
3.1.5 使用“查阅向导”创建查阅字段	69
3.1.6 向表中输入数据	72
3.2 创建表的主键和关联	73
3.2.1 设置表的主键	73
3.2.2 建立表间关系	74
3.2.3 创建、编辑表间关系操作	75
3.3 对表的编辑和修改	79
3.3.1 修改表结构	79
3.3.2 编辑表中的记录	81
3.4 表中记录的排序与筛选	84
3.4.1 对表中记录重新排序	84
3.4.2 对表中记录筛选	86
3.5 调整表的外观	89
3.5.1 调整表中数据的显示格式	89
3.5.2 调整表的显示样式	90
本章小结	93
习题与操作练习	94
第4章 SQL 语言——通用的数据库访问语言	98
4.1 SQL 语言简介	98
4.1.1 SQL 语言的基本功能和特点	98
4.1.2 SQL 语言的数据类型	99
4.1.3 在 Access 中使用 SQL 语句	100
4.2 使用 SQL 创建和删除数据库表	102

4.2.1	使用 SQL 创建数据表	102
4.2.2	使用 SQL 删除数据表	103
4.3	使用 SQL 语句操纵数据	104
4.3.1	插入数据	104
4.3.2	修改数据	105
4.3.3	删除数据	106
4.4	使用 SQL 语句查询数据	107
4.4.1	查询语句 SELECT 的基本结构	107
4.4.2	单表查询	108
4.4.3	多表连接查询	117
4.4.4	嵌套查询	119
	本章小结	121
	习题与操作练习	121
第 5 章	查询——按我们的需求获取有用信息	125
5.1	查询概述	125
5.1.1	查询的概念	125
5.1.2	查询的功能	126
5.1.3	查询的类型	126
5.1.4	查询的视图	127
5.1.5	查询的创建及运行	128
5.2	使用“查询向导”创建查询	130
5.2.1	使用“简单查询向导”创建查询	130
5.2.2	使用“交叉表查询向导”创建查询	134
5.2.3	使用“查找重复项查询向导”创建查询	137
5.2.4	使用“查找不匹配项向导”创建查询	138
5.3	使用“查询设计”创建查询	141
5.3.1	查询设计视图	141
5.3.2	查询定义窗格中字段的操作	142
5.3.3	不含条件的选择查询	143
5.3.4	含筛选条件的选择查询	145
5.3.5	参数查询	149
5.3.6	分组、汇总查询	152
5.3.7	交叉表查询	155
5.4	创建操作查询	157
5.4.1	生成表查询	157
5.4.2	追加查询	159
5.4.3	更新查询	160
5.4.4	删除查询	162

本章小结·····	164
习题与操作练习·····	164
第6章 窗体——让用户更清晰地浏览数据库 ·····	167
6.1 窗体简介·····	167
6.1.1 窗体的主要功能·····	167
6.1.2 窗体的类型·····	168
6.1.3 窗体视图·····	169
6.1.4 窗体的基本结构·····	170
6.1.5 窗体创建工具·····	171
6.2 自动创建输出数据源的窗体·····	172
6.2.1 使用“窗体”创建窗体·····	172
6.2.2 使用“其他窗体”创建窗体·····	174
6.2.3 使用“窗体向导”创建窗体·····	177
6.2.4 使用“空白窗体”创建窗体·····	177
6.3 面向对象的可视化窗体设计·····	179
6.3.1 “窗体设计工具”选项卡介绍·····	179
6.3.2 窗体中的控件·····	180
6.3.3 设计窗体的常用操作·····	182
6.3.4 窗体的使用·····	185
6.3.5 窗体外观设计·····	186
6.4 利用“设计视图”创建各种自定义窗体·····	186
6.4.1 利用“设计视图”创建捆绑数据源的窗体·····	186
6.4.2 利用“设计视图”创建非捆绑数据源窗体·····	188
本章小结·····	191
习题与操作练习·····	191
第7章 报表——输出标准的统计分析结果 ·····	195
7.1 报表概述·····	195
7.1.1 报表的类型·····	195
7.1.2 报表的视图·····	196
7.2 使用向导创建报表·····	198
7.2.1 自动创建报表·····	198
7.2.2 创建空报表·····	199
7.2.3 使用“报表向导”创建报表·····	200
7.2.4 创建标签报表·····	204
7.3 使用设计视图创建报表·····	206
7.3.1 报表的结构·····	206
7.3.2 创建简单报表·····	208

7.3.3 创建主/子报表	211
7.3.4 创建图表报表	215
7.4 报表的排序、分组和计算	218
7.4.1 报表记录的排序	218
7.4.2 报表数据的分组和计算	220
7.5 编辑报表	223
7.5.1 设置报表格式	223
7.5.2 添加图像和线条	223
7.5.3 添加日期和时间	223
7.5.4 添加页码	224
7.6 报表的打印和预览	224
7.6.1 页面设置	224
7.6.2 预览和打印报表	225
7.7 报表导出成 HTML 文档	226
本章小结	227
习题与操作练习	228
第 8 章 宏——让系统的运行与操作变得更便捷	231
8.1 宏能帮助我们做什么	231
8.1.1 宏的基本概念及功能	231
8.1.2 宏的设计视图	232
8.1.3 常用宏操作命令分类释义	233
8.1.4 宏的类型及设计流程	235
8.2 宏的设计与创建	235
8.2.1 序列宏的创建示例	236
8.2.2 条件宏的创建及示例	238
8.2.3 宏组的创建及示例	241
8.3 宏的运行	242
8.3.1 直接运行宏	242
8.3.2 触发事件运行宏	243
8.3.3 嵌入宏的创建与运行	245
8.3.4 自动运行宏	246
8.3.5 在宏中运行其他宏	247
8.4 宏的编辑与调试	247
8.4.1 宏的编辑	247
8.4.2 宏的调试	248
本章小结	249
习题与操作练习	249

第 9 章 模块——数据库应用程序的摇篮	253
9.1 认识 Access 模块	253
9.1.1 模块的基本功能和类型	253
9.1.2 VBA 模块的开发环境	254
9.1.3 模块的创建与运行	255
9.2 VBA 程序及相关知识	256
9.2.1 VBA 基础知识	256
9.2.2 VBA 程序的基本语句	259
9.2.3 程序的基本结构	261
9.2.4 过程	269
9.3 面向对象的程序设计	270
9.3.1 数据库对象变量	270
9.3.2 数据库对象的方法与事件	271
9.3.3 在模块中执行数据库操作命令	273
9.3.4 面向对象 VBA 程序设计综合示例	275
本章小结	279
习题与操作练习	279

第 1 章

数据库——信息处理与管理的好助手

任务描述：要开发一个“高校学生竞赛管理系统”，其第一步是业务需求分析和数据库设计，即要在实际赛事管理业务中抽取报名、学生、赛场、作品和评审等业务实体中有用的数据以及各实体彼此的联系和需求信息，然后按照关系数据库的基本要求转换成计算机能够处理和管理的数据库。因此，本章的主要任务是：

1. 通过对数据库功能和特点的介绍，了解数据库在赛事管理中的重要作用，明确数据库能帮助我们做什么和怎样做。
2. 对“高校学生竞赛管理系统”进行业务需求分析，抽象出与管理功能相关且有用的数据，并分析数据间的联系。
3. 掌握数据库设计的基本流程和方法，对“高校学生竞赛管理系统”进行概念和逻辑模型设计。

知识铺垫：信息处理与数据库管理技术；数据模型及其功能；业务需求分析方法；概念模型设计(E-R图)；关系数据库及其特点；关系规范理论及概念模型转换关系规则。

解决思路：对“高校学生竞赛管理系统”进行实际管理业务需求分析，抽取相关的数据和功能，借助关系数据库系统设计方法，构建其概念模型和逻辑模型，并将其合理优化，转换为数据库管理系统所能存储和处理的符合规范的关系模型。

能力训练：数据库应用系统的设计；关系数据库优化、E-R转换关系设计。

1.1 信息处理与数据库

今天人类已进入了大数据的信息社会，人们每时每刻都要跟数据和信息打交道，需要使用数据库来进行数据的处理和管理。

1.1.1 数据、信息与信息处理

什么是数据和信息，它们之间又有什么关系？这里，举两个日常接触的小例子来说明这个问题。

【例 1】 天气预报。

我们经常通过天气预报来了解天气，气象台主要依据事先采集、勘测的一组气象数据，它们往往用专门的符号来描述，如气压、云层、温度、湿度、风力……然后把它们进行整理加工、综合分析(与经验值比较、统计、运算、推断)，从而预报天气(阴、晴、刮风、下雨……)。我们把气压、云层、温度、湿度、风力等称为数据，而把经过整理加工和综合分析的结果(天气预报)称为信息。

【例 2】 学生管理。

学校对在校学生进行管理,首先,要把每个学生的学号、姓名、性别、出生日期、年级以及所属院系、专业、成绩、档案、照片等描述学生的基本属性登记入册,然后根据管理的不同需要(如教务处的排课、考试与成绩管理;财务处的助、奖学金管理;资产处的宿舍、设备管理、图书馆的借阅图书管理、校园的一卡通管理……)对它们进行编辑、分类、排序、更新、筛选、检索、统计、汇总等综合处理来实现各种应用管理。可以把学生的学号、姓名、性别,甚至档案、照片等描述学生基本属性的称为数据,而把经分类、检索、统计、计算、汇总等综合处理后得到的管理结果称为信息。

从上述例子中归纳出以下两个概念:

- 数据(Data)——对客观事物特征及属性的一种抽象的、用各种形式描述的符号,是信息的符号表示和载体。
- 信息(Information)——以数据为载体,是具有一定含义的、经过加工处理的、对人类决策有帮助和价值的数据。

对信息系统来说,常用的信息定义是:“信息是已被处理成某种形式的数据,这种形式对接受信息者具有意义,并在当前或未来的行动和决策中,具有实际的和可觉察到的价值。”

显而易见,若想得到信息,首先必须将客观世界中的现象和问题通过数据这种媒体记载下来,这里所说的数据应该是原始的、广义的、可鉴别的抽象符号,它可以描述事物的属性、状态、程度、方式等,这些符号单独表示时没有任何含义,只有把它们放入特定的场合,对它们进行解释和加工,才能使它们具有意义、升华为信息,进而为各项管理服务。所以,数据和信息是两个不同的概念,它们之间的关系如图 1.1 所示。



图 1.1 数据与信息的关系

在图 1.1 中可以清晰地看到数据与信息既有联系又有区别,数据是未加工处理的原始信息,信息是经过加工处理提炼出来的数据,可以借助信息系统把它们联系起来,信息系统在把数据转变为信息的过程中完成了信息处理,即:

信息处理——对数据进行采集、存储、分类、解释、编辑、比较、检索、查询、统计、计算、汇总、传输、复制、分析、预测和输出等,使之生成信息的过程。

实际上,可以将信息处理按照流程大致划分为五大阶段内容:收集、存储、加工、传递、提供。信息系统的基本功能亦可按这一流程分为以下五个方面:

1. 数据的收集

信息处理的第一步工作就是收集数据,因为数据是信息处理的原料,没有实际的数据就不会产生任何实用价值的信息。根据数据的来源不同,可以把数据收集工作分为原始数据收集和二次数据收集两种。原始数据收集是指在数据发生的当时当地,从客观实体中直接把数据取出,并用特定的技术手段在介质上记录下来;二次数据收集则是指收集已记录在某种介质上、与所描述的实体在时间与空间上已分离开的数据。这一步工作的关键是收集的数据的准确性和真实性。

2. 数据的存储

信息系统必须具有存储记忆的功能,特别是在大数据时代,否则就无法突破时间与空间的限制,发挥提供信息、支持决策的作用。数据存储主要应考虑存储量、数据格式、存储方式、使用方式、存储时间和安全保密等问题。简单地说,信息系统的存储功能就是保证已得到的信息能够不丢失、不走样、不外泄,大量保存、长期使用、节省空间、方便调用。

3. 信息的加工

信息加工亦称信息转换或处理,在这一环节中要把数据转换成计算机可处理的形式并生成对用户有用的信息。这一环节的工作大致可分为两大类,一类为数值运算,如加、减、乘、除、乘方等;另一类为非数值运算,如逻辑比较运算、对文字符号的编辑排版、音频、视频信号的模数转换及合成处理等,它们可以是一些较简单的运算,也可以是一些较为复杂的运算。由于这一环节经常要对大量数据进行各种加工处理,因而也是较为烦琐的工作。这一环节中最关键和困难之处在于能把繁杂数据中真正有用的信息提取出来,为此,信息加工不仅要依赖各种运算处理技术,而且要有“去粗取精,去伪存真”的功能。

4. 信息的传递

当信息在地理上有一定的分布时,信息的传递就成为一项很重要的功能,计算机网络目前就是一种信息传递的主要工具。在信息传递中的关键问题是传输通道,它是发送方与接受方之间的联系渠道,这个通道的完善与否将影响着整个信息系统的质量。对通道的要求首先是要保证其传输能力与传输速度;其次还要保证不失真,能抗干扰。

5. 信息的提供

信息的提供(亦称输出)是信息处理的最终环节,也是信息系统与用户的接口和界面。对这一环节的要求主要是按用户的需要、以用户所能接受和清晰易懂的形式及时提供对其有用的信息。

1.1.2 数据库能帮助我们做什么

随着社会和技术的发展,人们逐渐发明创造出各种物理设备来提高信息处理的能力与效率。从古代人类使用的结绳记事、烽火台、驿站,到现代的算盘、计算尺、电报电话、录音录像直至当今的电子计算机、互联网和“云”,人类在不断地改进信息处理的技术和工具,以适应社会发展中日益增多的信息处理的需要。

当今的社会已进入到“大数据”环境,大数据的特点是:

- (1) 数据的增长速度快(以摩尔定律递增);
- (2) 大量的机器数据(存放在计算机存储设备中,存储量从 KB 到 YB);
- (3) 数据应用范围扩大(各行业、学科:舆情分析、经济预测、电子政务、心理测试……);
- (4) 数据的多样化(数据源为各种媒体:文字、数值、声音、图像、影视;静态到动态);
- (5) 数据的复杂性(从结构化到非结构化,从自然科学领域的具体数值到人文社会科学领域的模糊数据);

(6) 数据处理技术(计算机、网络、数据库、流媒体、数据挖掘、云计算……)。

大数据环境对信息处理、数据分析要求极高,为了高效率地完成大数据的处理工作,人们努力研究如何把计算机技术与信息处理更好地结合起来,以实现计算机的普及应用和对数据的组织管理,更好地提高各行各业的业务管理水平。由于计算机数据库能够快速、高效、安全、优质地处理数据,并且其技术在不断地更新发展,为信息处理带来了极大的方便,也使人类社会对信息的利用和开发方式发生了根本性的变革。通俗地说,数据库就是计算机存储各类数据的“仓库”,而数据库系统就是对数据实现高效、优质、智能、安全管理的计算机信息系统。它可以存储和管理大量、复杂、持久、有关联的各种数据,并可以对这些数据进行采集、存储、分析、加工、传递、提供等有效的安全管理,所以被越来越多地应用于各行各业的管理业务和统计、分析、决策中。如国家舆情分析决策系统、国民经济统计预测系统、企业的财务管理系统、学校的教务管理系统、商场的销售管理系统、医院就诊管理系统、法院的案件审判系统、犯罪心理测试与干预系统以及各行业的网站建设等,如图 1.2 所示。本教材就是以“高校学生竞赛管理系统”为例,介绍数据库在竞赛管理业务中的实际应用。

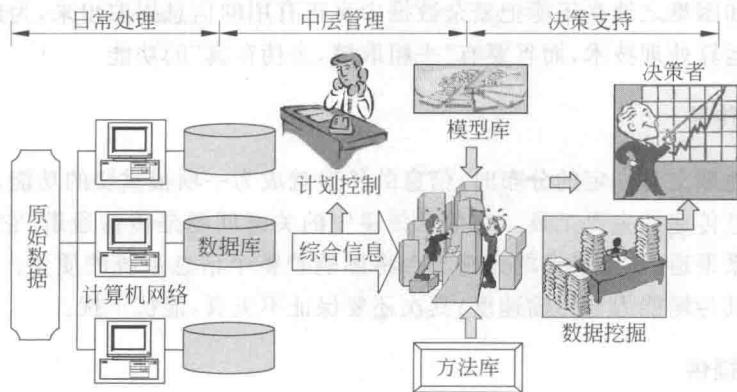


图 1.2 数据库应用与发展

1.1.3 数据管理技术的发展

信息处理的前期和主要任务是对来自客观世界大量待处理的数据进行收集、保存和管理。随着计算机信息处理技术的提高,数据管理技术大致经历了如下几个阶段。

1. 人工管理阶段(早期)

计算机不能对数据进行有效地保存和管理,数据堆放杂乱无序;人工输入、存储、输出数据的工作量大;各个应用程序之间存在大量重复数据,数据存储冗余高。

2. 文件管理阶段(20 世纪 50—60 年代)

计算机用数据文件来存放数据,数据可以保存在磁带或磁盘上;由操作系统提供管理文件的方法;采用文件系统(file system)组织、管理数据。在文件系统中,数据按其内容、结构、用途组成若干个数据文件,一个数据文件与其特定的应用文件存放在一起,而不是独立存放的。从某种程度上看,文件系统中的文件及其所包含的数据反映了一个组织中各业务

活动的基本情况和数据组织的层次结构。比如,商场的销售数据被包含在商场销售业务应用文件中;学生成绩数据被存放在学校教务管理文件中,在对这些应用文件进行处理的同时,即对相应的数据文件使用和管理。这种为每个应用程序分别创建和存储若干个数据文件的数据组织方法称为传统数据管理方法,它的局限性是显而易见的。

(1) 数据文件的设计很难满足多种应用程序的不同要求,使得分散在多个文件中的数据存在着不必要的重复(冗余),造成存储空间的浪费。

(2) 由于数据无法统一管理,在数据的结构、编码、格式、命名以及输出等方面不容易做到规范化、标准化,不能反映数据之间的内在联系;在数据的安全和保密方面难以采取有效的措施,对文件的并发访问也较难控制。

(3) 由于数据与应用程序捆绑在一起,对数据文件结构的修改将导致应用程序的修改,应用程序的维护工作量很大。

(4) 编写应用程序很不方便,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。

3. 数据库管理阶段(20世纪60年代末至今)

针对文件管理的上述缺点,人们逐步研发了以统一管理和共享数据为主要特征的数据库以及数据库系统。显然,数据库是为了更好地组织、保存、管理各种业务中的数据,以帮助人们有效地进行信息处理。如图1.3所示,在数据库系统中,相关的数据集合被独立存放在应用程序之外,不再仅仅服务于某个应用程序或用户,而是被多个应用程序访问和共享,用一个叫做数据库管理系统(Database Management System, DBMS)的软件统一管理。由于有DBMS的统一管理,它对数据库的管理的三重模式以及数据库的物理、逻辑独立性,使应用程序不必直接介入诸如打开、关闭、读、写文件等低级数据操作;用户也不必关心数据存储和其他实现的细节,可以在更高的抽象级别上观察和访问数据、设计应用程序;数据库结构的修改也可以由DBMS独立处理,从而减少应用程序的维护工作量。由于有DBMS的统一管理,人们可以从全局着眼,合理组织数据,减少数据冗余;还可以更好地对数据实施规范化和标准化,保证了并发访问时的数据一致性和数据安全性。

与文件系统相比,数据库系统具有以下特点:

- ① 数据的独立性强,减少了应用程序对数据结构的依赖性。
- ② 数据的冗余度小,避免了存储数据的重复和空间的浪费。
- ③ 数据的共享度高,保证了数据可以为不同的用户所使用。
- ④ 数据库的结构化,有利于对数据统一有效地管理和控制。

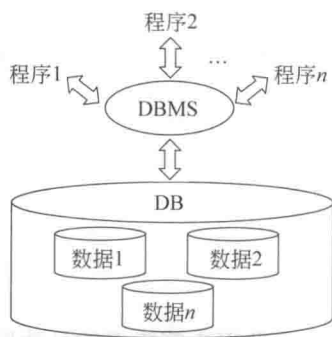


图 1.3 数据库系统中程序与数据的关系

1.1.4 认识和了解数据库

通常我们所称的数据库是一个由多个元素构成的系统,数据库系统是管理大量的、持久的、共享的数据,保证数据库的安全可靠,进而实现信息系统总体目标的一种强有力的信息处理工具,是现阶段数据管理的主要形式。