

计算机基础教育课程体系规划教材



数据库技术·应用 及实验指导

何玉洁 张俊超 编著



机械工业出版社
China Machine Press

计算机基础教育课程体系规划教材



数据库技术·应用 及实验指导

何玉洁 张俊超 编著



机械工业出版社
China Machine Press

本套教材根据教育部高等学校非计算机专业计算机基础课程教学指导分委员会最新提出的《关于进一步加强高校计算机基础教学的几点意见》中的课程体系和教学基本要求组织编写,本书根据其中有关“数据库技术及应用”课程的教学要求编写而成。

本书全面阐述数据库系统的基本理论、技术和方法,涵盖了目前流行的Microsoft Access数据库管理系统和Visual Basic可视化编程环境。另外,书中各章后附有适量的习题,以便读者实践该章所介绍的技巧和概念。

本书概念清楚、重点突出、章节安排合理,可作为高等学校非计算机专业数据库课程的教材,也可供数据库初学者自学参考。

版权所有,侵权必究。

法律顾问 北京市展达律师事务所

图书在版编目(CIP)数据

数据库技术·应用及实验指导 / 何玉洁, 张俊超编著. - 北京: 机械工业出版社, 2005.8

(计算机基础教育课程体系规划教材)

ISBN 7-111-16982-4

I. 数… II. ①何… ②张… III. 数据库系统-高等学校-教材 IV. TP311.13

中国版本图书馆CIP数据核字(2005)第081348号

机械工业出版社(北京市西城区百万庄大街22号 邮政编码 100037)

策划编辑: 温莉芳 姚 蕾

责任编辑: 迟振春

北京瑞德印刷有限公司印刷·新华书店北京发行所发行

2005年8月第1版第1次印刷

787mm×1092mm 1/16·18.5印张

印数: 0 001-5 000册

定价: 29.00元

凡购本书,如有倒页、脱页、缺页,由本社发行部调换
本社购书热线:(010) 68326294

丛书序

根据1997年教育部高教司颁发的“加强非计算机专业计算机基础教学工作的几点意见”（简称155号文件）的要求，各校的计算机基础条件已经明显改善，计算机基础教学进入了一个新阶段。

本届非计算机专业计算机基础课程教学指导分委员会分析了当前高校遇到的计算机基础教学的新形势，根据人才培养的基本要求，针对计算机基础教学中普遍存在的问题，提出了“关于进一步加强高等学校计算机基础教学的意见”（俗称白皮书）。并在其附件“计算机基础教学内容的知识结构与课程设置”中提出了“1+X”的课程方案，即1门“大学计算机基础”（必修）加上几门重点课程（必修或选修）。白皮书及附件自2003年底在高校征求意见以来，受到了普遍的关注，引起巨大反响。教指委根据征求到的意见做了进一步的修改，不久将正式发布。这无疑将直接影响今后高校计算机基础教学的整体架构，也将推动新一轮的计算机基础教材的面世。

机械工业出版社以其敏锐的眼光和雄伟的魄力，怀着为计算机基础教学做贡献的责任感，遵循白皮书提出的理念，于2004年在全国范围内邀请计算机基础教学一线教师，组织编写“1+X”中规定的6门核心课程及其若干门整合课程。本丛书参考白皮书对于教材建设所提出的建议，努力在以下几个方面做出特色：

1. 对于重点核心课程的教材，体现课程内容的基础性和系统性，基本概念、基本技术与方法的讲解要准确明晰。

2. 体现非计算机专业计算机基础教材的特点，内容要激发学生学习兴趣，通俗易懂，理论联系实验，每一门课都要使学生真正学到一些有用的知识和技术。

3. 保证教材内容的先进性，特别对于技术性、应用性的内容更应如此。

4. 重视实验教材的建设，重点教材都要配备实验教材。

我们希望本丛书的出版对推动计算机基础教育有所帮助，并在使用中不断改进。书中不足之处恳望读者不吝指正。

丛书主编 冯博琴

2005.4

前 言

进入21世纪之后,我国明显加快了建设世界教育大国的步伐,对高等学校学生的教育质量特别是计算机知识的教育越来越重视,要求也越来越高。数据库技术作为计算机科学中的一个非常重要的部分,正以日新月异的速度发展,因此作为现代的大学生,无论是计算机专业的学生还是非计算机专业的学生,学习和掌握数据库知识都是非常必要的。

本书是为非计算机专业学生学习数据库知识而编写的一本教材。

本书由五部分组成。第一部分介绍数据库的基本理论和基本概念。该部分由第1章~第6章组成,具体内容包括:数据库基本概念、数据模型与数据库系统结构、关系数据模型与关系代数、数据操作语句与数据库对象定义语句、数据安全性与完整性控制、数据库的可靠性机制以及数据库的设计过程。

第二部分介绍数据库的应用。该部分由第7章~第9章组成,具体内容包括:微软的Access数据库管理系统的使用、通用的数据库访问接口以及Visual Basic中提供的数据库访问控件的使用,最后在第9章给出了在Visual Basic环境中,使用数据库访问控件通过通用的数据库访问接口访问Access数据库的三个示例。

第三部分介绍数据库技术的发展。该部分由第10章组成,主要介绍数据库系统结构的发展、面向对象数据库中的数据模型以及面向对象数据库的优点,最后介绍了数据库的一些新技术、研究方向和新的应用领域。

第四部分为Access实验部分。该部分由第11章~第15章组成,以Access 2003为例介绍Access中的数据库操作、数据表操作、查询对象、窗体以及报表的功能。

第五部分为Visual Basic实验部分。该部分由第16章组成,主要介绍利用ADO数据控件浏览和修改单表数据、查询多表数据的方法,并在最后给出了一个实验练习。

本书的最大特点是内容全面,既包括数据库的基础理论知识,又包括数据库的应用技术。Access是目前应用广泛的个人数据库管理系统,其功能强大,且具有很好的图形化操作界面、对机器的要求不高、占用的资源不多等,便于非计算机专业的学生学习和使用。Visual Basic是使用非常普及、方便的可视化编程环境,目前一般的学校都开设Visual Basic课程,这样可以使学生在一个比较熟悉的环境中进一步提高其对数据库的访问和操作能力,从而降低学习难度。

由于时间仓促加之本人水平所限,书中难免有不妥之处,望广大同仁给予批评指正。

何玉洁

2005年6月

目 录

丛书序 前言

第一部分 数据库基本理论

第1章 数据库基础	2
1.1 数据	2
1.1.1 数据与信息	2
1.1.2 数据密集型应用的特点	3
1.2 数据管理的发展	3
1.2.1 文件管理系统	3
1.2.2 数据库管理系统	6
1.3 数据库系统的组成	8
1.4 小结	9
习题	9
第2章 数据模型与数据库系统结构	10
2.1 数据模型	10
2.1.1 概念层数据模型	12
2.1.2 组织层数据模型	14
2.2 数据库系统的结构	17
2.2.1 数据库系统模式的概念	17
2.2.2 三级模式结构	18
2.2.3 二级映像功能和数据独立性	20
2.3 数据库管理系统的功能	21
2.4 小结	22
习题	23

第3章 关系数据库	24
3.1 关系数据模型的组成	24
3.1.1 关系数据结构	24
3.1.2 关系操作	25
3.1.3 关系完整性约束	25
3.2 关系数据模型的基本术语与形式化定义	25
3.2.1 关系模型的基本术语	26
3.2.2 关系数据结构及其形式化定义	27
3.3 关系数据库模型的完整性约束	29
3.3.1 实体完整性	29
3.3.2 参照完整性	30
3.3.3 用户定义的完整性	31
3.4 关系代数	31
3.4.1 传统的集合运算	32
3.4.2 专门的关系运算	33
3.5 小结	38
习题	38
第4章 SQL语言基础	39
4.1 基本概念	39
4.1.1 SQL语言特点	40
4.1.2 SQL语言功能概述	40
4.2 SQL的数据类型	40
4.2.1 SQL-99标准的数据类型	41
4.2.2 Microsoft SQL Server的数据类型	41
4.3 数据定义功能	42

4.3.1 基本表的定义与删除	42
4.3.2 修改表结构	45
4.4 数据查询功能	45
4.4.1 查询语句的基本结构	47
4.4.2 简单查询	47
4.4.3 多表连接查询	58
4.4.4 子查询	62
4.5 数据更改功能	64
4.5.1 插入数据	65
4.5.2 更新数据	65
4.5.3 删除数据	66
4.6 视图	67
4.6.1 视图的概念	67
4.6.2 定义视图	67
4.6.3 删除视图	69
4.6.4 视图的作用	69
4.7 索引	70
4.7.1 建立索引	70
4.7.2 删除索引	71
4.8 小结	71
习题	72

第5章 数据库保护73

5.1 数据库安全性	73
5.1.1 安全控制模型	73
5.1.2 存取控制	74
5.1.3 数据库权限的种类及用户的分类	75
5.1.4 操作权限定义	75
5.2 完整性约束	77
5.2.1 完整性约束条件的作用对象	77
5.2.2 实现数据完整性	78
5.3 数据库备份与恢复	80
5.3.1 数据库故障的种类	80
5.3.2 数据库备份	81
5.3.3 数据库恢复	82
5.4 小结	83
习题	83

第6章 数据库设计84

6.1 数据库设计概述	84
6.1.1 数据库设计的特点	85
6.1.2 数据库设计方法概述	86
6.1.3 数据库设计的基本步骤	87
6.2 数据库需求分析	87
6.2.1 需求分析的任务	87
6.2.2 需求调查	88
6.3 数据库结构设计	89
6.3.1 概念结构设计	89
6.3.2 逻辑结构设计	92
6.3.3 物理结构设计	99
6.4 数据库行为设计	101
6.4.1 功能需求分析	101
6.4.2 功能设计	102
6.4.3 事务设计	102
6.5 数据库的实施和维护	103
6.5.1 数据库数据的加载和试运行	103
6.5.2 数据库的运行和维护	104
6.6 小结	104
习题	105

第二部分 数据库应用

第7章 Access数据库管理系统108

7.1 Access简介	108
7.1.1 Access的特点	109
7.1.2 Access的对象及功能	109
7.2 创建及管理Access数据库	110
7.2.1 创建数据库	110
7.2.2 打开数据库	115
7.2.3 管理数据库	116
7.2.4 设置数据库密码	119
7.3 创建表	120
7.3.1 使用表向导创建表	120
7.3.2 使用表设计器创建表	124
7.3.3 保存表的定义	134
7.3.4 建立表之间的关联	134
7.4 查询对象	136
7.4.1 查询对象概述	136

7.4.2 查询对象的种类	137
7.4.3 建立查询对象	137
7.5 小结	153
习题	153

第8章 数据库应用开发技术155

8.1 数据访问接口	155
8.1.1 ODBC	155
8.1.2 OLE DB和ADO	159
8.1.3 ADO为OLE DB带来了 什么	161
8.2 ADO数据控件	162
8.2.1 ADO数据控件概述	162
8.2.2 ADO数据控件的主要 属性	163
8.2.3 ADO数据控件的主要 方法	167
8.2.4 ADO数据控件的主要 事件	167
8.3 Recordset对象的主要属性和 主要方法	168
8.3.1 Recordset对象的主要 属性	168
8.3.2 Recordset对象的主要 方法	169
8.4 数据绑定控件	170
8.4.1 DataGrid控件	171
8.4.2 DataList控件和DataCombo 控件	177
8.5 小结	179
习题	180

第9章 VB数据库应用编程示例181

9.1 示例1	181
9.2 示例2	182
9.3 示例3	185
9.4 数据窗体向导	187
9.4.1 添加数据窗体向导	187
9.4.2 使用数据窗体向导	187
9.5 小结	192
习题	192

第三部分 数据库技术的发展

第10章 数据库技术的发展194

10.1 数据库系统结构的发展	194
10.1.1 集中式结构	194
10.1.2 文件服务器结构	195
10.1.3 客户/服务器结构	196
10.1.4 Web数据库应用架构	196
10.2 面向对象技术与数据库技术的 结合	197
10.2.1 新的数据库应用和新的 数据类型	197
10.2.2 面向对象数据模型	198
10.2.3 面向对象数据库的优点	202
10.2.4 对象关系数据库与对象 数据库	202
10.3 数据库技术的新发展	202
10.3.1 数据库技术的三个发展 阶段	203
10.3.2 数据库技术的研究 方向	204
10.3.3 面向应用领域的数据库 新技术	206
10.4 数据库技术面临的挑战	207
10.5 小结	209
习题	209

第四部分 Access实验

第11章 创建及打开数据库212

11.1 实验目的	212
11.2 创建数据库	212
11.2.1 实验任务	212
11.2.2 操作指导	212
11.3 打开数据库	214
11.3.1 实验任务	214
11.3.2 操作指导	214
11.4 关于数据库窗口的一些操作	215
11.4.1 实验任务	215
11.4.2 操作指导	215

第12章 数据表操作	217	第13章 查询	240
12.1 建立和修改数据表	217	13.1 实验目的	240
12.1.1 实验目的	217	13.2 利用简单查询向导建立	
12.1.2 建立数据表	217	查询对象	240
12.1.3 修改表结构	222	13.2.1 实验任务	240
12.1.4 定义实体完整性	223	13.2.2 操作指导	240
12.1.5 定义参照完整性	224	13.3 利用设计视图建立无条件查询	
12.2 输入和修改列数据	226	对象	243
12.2.1 实验目的	226	13.3.1 实验任务	243
12.2.2 输入数据	226	13.3.2 操作指导	243
12.2.3 级联更新数据和级联删除			
数据	227		
12.3 操作数据	229		
12.3.1 实验目的	229		
12.3.2 修改和添加记录	229		
12.3.3 排序记录	230		
12.3.4 查找记录	230		
12.3.5 筛选记录	231		
12.3.6 排除筛选记录	232		
12.4 更改数据表的显示方式	232		
12.4.1 更改列的显示顺序	233		
12.4.2 隐藏/显示列	235		
12.4.3 冻结/解冻列	236		
12.4.4 改变字体样式	237		
12.4.5 设置单元格效果和样式	238		
		第14章 窗体	246
		14.1 实验目的	246
		14.2 利用窗体向导建立窗体	246
		14.2.1 实验任务	246
		14.2.2 操作指导	246
		14.3 利用设计视图建立窗体	251
		14.3.1 实验任务	251
		14.3.2 操作指导	252
		第15章 报表	254
		15.1 实验目的	254
		15.2 利用报表向导建立报表	254
		15.2.1 实验任务	254
		15.2.2 操作指导	254
		15.3 利用设计视图建立报表	262
		15.3.1 实验任务	262
		15.3.2 操作指导	262
		15.4 实验结果	264
		第五部分 Visual Basic实验	
		第16章 VB数据库访问	266
		16.1 实验目的	266
		16.2 基于单表的数据浏览和操作	266
		16.2.1 实验任务	266
		16.2.2 操作指导	266
		16.3 基于查询的数据操作	278
		16.3.1 实验任务	278
		16.3.2 操作指导	279
		16.4 实验练习	283
		参考文献	285

第一部分

数据库基本理论

第 1 章

数据库基础

随着信息管理水平的不断提高,信息资源已成为企业的重要财富和资源,用于管理信息的数据库技术也得到了很大的发展,其应用领域也越来越广泛。数据库以多种方式出现,从小型事务处理到大型信息系统,从联机事务处理到联机分析处理,从一般企业管理到计算机辅助设计与制造(CAD/CAM)、地理信息系统等,都应用了数据库技术。数据库技术已经渗透到我们日常生活中的方方面面,比如用信用卡购物,飞机、火车订票系统,图书馆对书籍及借阅的管理等,无一不使用了数据库技术。目前,数据库的建设规模、数据库中信息量的大小以及使用的程度已经成为衡量企业信息化程度的重要标志。

简单地说,数据库技术就是研究如何对数据进行科学管理以便为人们提供可共享的、安全的、可靠的数据的技术。数据库技术一般包含数据管理和数据处理两部分内容。

数据库系统实质上是一个用计算机存储数据的系统,数据库本身可以看做是一个电子文件柜,也就是说,数据库是收集数据文件的仓库或容器。

1.1 数据

1.1.1 数据与信息

信息与数据有很多相似之处,但其表达的内容实际上是有差别的。数据由原始事实(如学生的姓名、年龄、性别、所在系等)组成,这些事实数据具有不同的数据类型。而信息是按特定方式组织在一起的事实数据的集合,它具有超出事实数据本身之外的额外价值。例如,销售管理人员可能会发现了解按月汇总的总销售额比了解每笔销售数据更符合需要,也就是对管理人员来说,按月汇总的数据更有用、更有价值,那么按月汇总的销售数据就称为信息。

数据代表真实世界的客观事物,是简单的客观事实,除它本身之外没有其他价值。可以把数据看做是一块一块的木头,除了作为一个单独的物体而言,木头本身并没有其他价值,但如果在各个木头之间定义了相互的关系,它们就有了更多价值。例如,如果将木头以某种

方式组合在一起,那么,它们就可以做成盒子、梯子、凳子等有用的东西。信息就是这些定义了关系的木头。规则和关系能够将数据组织起来使其成为有价值的信息。

信息的类型视数据之间定义的关系而定,例如,制作梯子的木头之间的关系与制作凳子的木头之间的关系显然是不同的。通过增加新的数据,更新数据之间的组织关系,从而可以生成新的信息。例如,经理可以将产品类型的数据加入到销售数据中去,从而生成按产品类型划分的销售信息。

将数据转换为信息的过程称为处理,即实施一系列逻辑上相关的任务,以完成某项预定的输出。

在某些情况下,对数据的组织和处理是用手工完成的,在另一些情况下,则是利用计算机进行处理。例如,对前面的例子,管理人员可以用手工计算每个销售人员的月销售额,也可以用计算机来计算。重要的不是数据的来源,也不是处理数据的方式,而是处理后的结果是否有价值。信息的价值与其如何帮助决策者实施组织目标直接相关。

1.1.2 数据密集型应用的特点

数据密集型的应用也就是我们所说的以数据为中心的应用,它具有如下三个特点:

1. 涉及的数据量大

以银行的信息管理为例,如果要将全部信息保存起来,则数据量是非常大的,这些大量的数据需要存放在辅助存储设备上,并需要有高效的处理方法。

2. 数据一般需要长期保存

需要长期保存的数据称为持久性数据。比如图书馆和银行的信息,必须要持久地保存,这些数据就是持久性数据。持久保存的数据是有价值的,人们可以通过对积累的数据进行分析,制定出合适的方针和决策。比如通过分析一段时间内哪些图书借出的次数比较多,可以帮助图书管理人员决定下次要进书的种类。这就是我们经常说的辅助决策支持功能。

3. 数据共享

在数据密集型应用中,数据一般不是某个用户专有的,而是可被许多用户使用,而且还必须允许多个用户同时使用这些数据。比如飞机订票系统,有成百上千个订票点,我们不可能在一个订票点工作时,不允许其他的订票点工作,否则设立这么多个订票点就没有意义了。

如何很好地管理这种大量的、持久的、共享的数据是计算机科学技术领域中一个重要的技术和研究课题。

1.2 数据管理的发展

在数据库技术出现之前,人们对数据的管理一般是采用文件系统来完成的,下面我们分别介绍文件管理系统和数据库管理系统管理数据的方式,并比较文件管理系统和数据库管理系统管理数据的本质区别。

1.2.1 文件管理系统

要理解现代数据库的特征,最好先看一下在数据库技术产生之前,人们是如何对数据进行管理的。

早期的数据是采用文件系统进行管理的,即将数据保存在文件中,用户的应用程序直接操作文件中的数据。在文件系统中,数据按其内容、结构和用途分成若干个命名的文件。文

件一般为某一用户或用户组所有，但也可以指定与其他用户共享。用户可以通过操作系统对文件进行打开、读、写、关闭等操作。

假设现在用文件系统来实现两个功能，即对学生的基本信息和选课情况进行管理，每个功能是一个系统，用一个程序实现。在学生基本信息管理中要用到学生的基本信息数据，假设这些数据保存在F1文件中；学生选课情况管理包括：学生的基本信息、课程的基本信息和学生的选课信息，假设后两类数据分别保存在F2和F3两个文件中，此部分的学生基本信息数据可以使用“学生基本信息管理”系统中的F1文件。假设实现“学生基本信息管理”功能的程序叫A1，实现“学生选课情况管理”功能的程序叫A2，则学生的基本信息和选课情况可用图1-1表示。

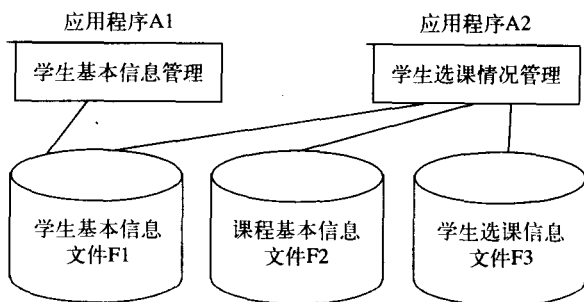


图1-1 文件管理系统示例

假设F1、F2和F3文件分别包含如下信息：

- F1包含学号、姓名、性别、出生日期、所在系、专业、所在班、特长、家庭住址。
- F2包含课程号、课程名、授课学期、学分、课程性质。
- F3包含学号、姓名、专业、课程号、课程名、选课类型、选课时间、考试成绩。

我们将文件中所包含的每一个子项称为文件结构中的字段或列，将每一行数据称为一个记录。

“学生选课情况管理”的处理过程大致为：在学生选课情况管理系统中，若有学生选课，则先查F1文件，判断有无此学生；若有此学生，则再访问F2文件，判断其所选的课程是否存在；若课程也存在，就将学生选课信息写到F3文件中。

这看起来似乎很好，但仔细分析一下，就会发现文件管理系统有如下一些缺点。

1. 编写应用程序不方便

应用程序编写者必须对所用的文件的逻辑结构及物理结构（文件中包含多少个字段，每个字段的数据类型，采用何种存储结构，比如链表或数组等）有清楚的了解。操作系统只提供了打开、关闭、读、写等几个低级的文件操作命令，而文件的查询、修改等处理都必须在应用程序中通过编程实现。这样也容易造成各应用程序在功能上的重复，比如图1-1中的“学生基本信息管理”和“学生选课情况管理”都要对F1文件进行操作，但这两个功能相同的操作却很难共享。

2. 数据冗余不可避免

假设A2需要在F3文件中包含学生的所有或大部分信息，比如，除了学号之外，还需要姓名、性别、专业、所在系等信息，而F1中也包含了这些信息，因此F3和F1文件中有重复的信息。但这些重复的信息只是不同文件的部分内容，因此很难在两个文件中共用这些公共信息，从而造成数据的重复（或叫数据的冗余）。

数据冗余不仅造成了存储空间的浪费（其实，随着计算机硬件技术的飞速发展，存储容量不断扩大，空间问题已经不是我们解决问题时需要关心的主要问题），更为严重的是造成了

数据的不一致。例如,假设某个学生所学的专业发生了变化,我们一般只会想到在F1文件中进行修改,而往往忘记了在F3文件中要进行同样的修改。这样就会造成同一个学生在F1文件和F3文件中的“专业”不一样,也就是数据不一致。人们不能判定哪个数据是正确的,尤其当数据冗余很多的时候,情况更是如此。这样数据就失去了其可信性。

文件系统本身不具备维护数据一致性的功能,这些功能完全由用户(应用程序开发者)负责维护。这在简单的系统中还可以勉强应付,但在复杂的系统中,若让开发者来保证数据的一致性,几乎是不可能的。

3. 应用程序有依赖性

就文件系统而言,应用程序要依赖于文件的结构。文件和记录的结构通常是应用程序代码的一部分,如C语言的struct。文件结构每进行一次修改,比如添加字段、删除字段甚至是修改字段的长度(如电话号码从7位扩到8位),都要对应用程序进行相应的修改,因为我们在打开文件读取数据时,必须要将文件记录中的不同字段的值对应到程序变量中。而随着应用环境和需求的变化,修改文件的结构是不可避免的事情,这样就需要在应用程序中进行相应的修改,而(频繁)修改应用程序是很麻烦的。这是因为首先要熟悉原有程序,修改后还需要对程序进行测试、安装等。

所有这些都是由于应用程序对文件结构过分依赖造成的,换句话说,文件系统的数据独立性不好。

4. 不支持对文件的并发访问

在现代计算机系统中,为了有效地利用计算机资源,系统一般允许多个应用程序并发运行(尤其是在现在的多任务操作系统环境中)。文件最初是作为程序的附属数据出现的,它一般不支持多个应用程序同时对同一个文件进行访问。我们可以回忆一下,假设某个用户打开了一个Excel文件,如果第二个用户在第一个用户没有关闭此文件之前就想打开此文件,他会得到什么信息?他只能以只读的方式打开此文件,而不能在第一个用户打开文件的同时对此文件进行修改。我们再回忆一下,如果我们用C语言编写一个修改某文件内容的程序,其过程是先以写的方式打开文件,然后写入新内容,最后再关闭文件。在文件关闭之前,无论是在其他的程序中还是在同一个程序中都是不能再打开此文件的,这就是文件系统不支持并发访问的含义。

对于以数据为中心的应用系统来说,必须要支持多个用户对数据的并发访问。

5. 数据间的联系弱

在文件系统中,文件与文件之间是彼此独立、毫不相干的,文件之间的联系必须通过程序来实现。比如在上述F1和F3文件中,F3文件中的学号、姓名等学生的基本信息必须是F1文件中已经存在的(即选课的学生必须是已经存在的学生);同样,F3中的课程号等与课程有关的基本信息也必须是F2文件中已经存在的(即学生选的课程也必须是已经存在的课程)。这些数据之间的联系是实际需求当中所要求的很自然的联系,但文件系统本身不具备自动实现这些联系的功能,所以必须通过应用程序来保证这些联系,也就是说必须编写代码来手工地保证这些联系。这样不但增加了编写代码的工作量和复杂度,而且当联系很复杂时,也难以保证其正确性。因此,文件系统不能反映现实世界事物间的联系。

6. 难以按不同用户的需要表示数据

如果用户需要的信息来自多个不同文件的部分数据内容的组合,我们就需要对多个文件进行提取、比较、组合和表示。假设用户希望得到如下信息:

所在班,学号,姓名,课程名,学分,考试成绩

这些信息涉及了三个文件：从F1文件中可以得到“所在系”数据，从F2文件中可以得到“学分”，从F3文件中可以得到“考试成绩”；而“学号”、“姓名”可以从F1或F3文件中得到，“课程名”可以从F2或F3文件中得到。我们在生成一行数据时，必须对从三个文件中读取的数据进行比较，然后组合成一行有意义的数据。比如，将从F1文件中读取的“学号”与从F3中读取的“学号”进行比较，学号相同时，才可以将F1中的“所在系”、F3中的“考试成绩”以及当前所对应的学号和姓名组合成一行数据的内容。同样，在处理完F1和F3文件的组合后，我们可以在组合的结果中再与F2文件中的内容进行比较，找出课程号相同的课程的学分，再与已有的结果组合起来。如果数据量很大、涉及的表比较多，可以想象这个过程有多复杂。因此，这种大容量复杂信息的查询，在文件管理系统中是很难处理的。

7. 无安全控制功能

在文件管理系统中，很难控制某个人对文件的操作，比如控制某个人只能读和修改文件，不能删除文件，或者不能读或修改文件中的某个或者某些字段等等。而在实际应用中，数据的安全性是非常重要且不可缺少的。比如，在学生选课情况管理系统中，我们不允许学生修改他的考试成绩。在银行系统中，更是不允许一般用户修改其存款数额。

随着人们对数据需求的增加以及计算机科学技术的不断发展，对数据进行有效、科学、正确、方便的管理就成为人们的迫切需求。针对文件系统的这些缺陷，人们逐步开发出了以统一管理和共享数据为主要特征的数据库管理系统。

1.2.2 数据库管理系统

数据库技术的出现主要是为了克服文件管理系统在管理数据上的诸多缺陷，满足人们对数据管理的需求。对于上述的学生基本信息管理和学生选课情况管理系统来说，如果使用数据库来管理，其实现方式与文件系统有很大区别。用数据库进行管理的实现过程如图1-2所示。

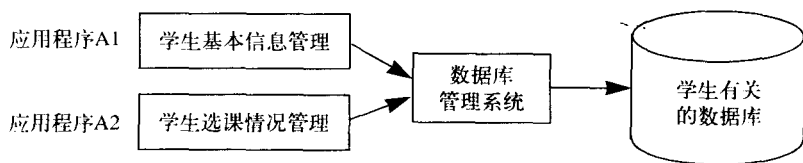


图1-2 数据库管理系统示例

比较一下图1-1和图1-2，可以发现两者有如下差别：

- 在文件系统中，应用程序直接访问存储数据的文件，而在数据库系统中，应用程序则是通过数据库管理系统（DataBase Management System，简称DBMS）来访问数据。
- 在数据库系统中，数据不再仅仅服务于某个程序或用户，存储数据的文件也不再需要直接被应用程序管理，而是由一个叫做数据库管理系统的软件统一管理。

与文件系统相比，数据库管理系统实际上是应用程序和存储数据的数据库（在某种意义上也可以把数据库看成是一些文件的集合）之间的接口。数据库管理系统实际上是一个系统软件。不要小看这个变化，有了这个系统软件后，以前在应用程序中由开发人员实现的很多繁琐的操作和功能，都可以交给这个系统软件。这样应用程序不再需要关心数据的存储方式，而且数据的存储方式的变化也不再影响应用程序。将需要进行的修改交给数据库管理系统，经过数据库管理系统处理后，应用程序感觉不到这些变化，因此，应用程序也不需要任何修改。

与文件系统管理数据的局限性进行比较,可以发现数据库系统具有以下优点。

1. 将相互关联的数据集成在一起

在数据库系统中,所有相关的应用数据都存储在一个称为数据库的环境中,应用程序可通过DBMS访问数据库中的所有数据。

2. 较少的数据冗余

由于数据是统一管理的,因此可以从全局着眼,合理地组织数据。例如,将1.2.1节的F1、F2和F3文件中的重复数据挑选出来,单独进行管理,这样就可以形成如下所示的几部分信息:

学生基本信息:学号、姓名、性别、出生日期、所在系、专业、所在班、特长、家庭住址。

课程基本信息:课程号、课程名、授课学期、学分、课程性质。

学生选课信息:学号、课程号、修课类型、修课时间、考试成绩。

在关系数据库中,可以将每一种信息存储在一个表中(关系数据库的概念我们在后面介绍),重复的信息只存储一份。若在学生选课中需要学生的名字,则根据学生选课中的学号,可以很容易地在学生基本信息中找到此学号对应的名字。因此,消除数据的重复存储不影响我们对信息的提取,同时还可以避免由于数据重复存储而造成的数据不一致问题。比如,当某个学生所学的专业发生变化时,我们只需在“学生基本信息”中进行修改即可。

同1.2.1节中的问题一样,当我们要检索(所在班,学号,姓名,课程名,学分,考试成绩)信息时,这些信息也需要从三个地方(关系数据库中为三张表)得到,也需要对信息进行适当的组合,即学生选课信息中的学号只能与学生基本信息中学号相同的信息组合在一起,同样,学生选课信息中的课程号也必须与课程基本信息中的课程号相同的信息组合在一起。在文件管理系统中,这项工作是由开发者编程实现的,而有了数据库管理系统后,这些繁琐的工作就完全交给了数据库管理系统来完成。

因此,数据库管理系统减少了数据冗余和开发者的负担。

3. 程序与数据相互独立

在数据库中,数据所包含的所有数据项以及数据的存储格式都与数据一起存储在数据库中,它们通过DBMS而不是应用程序来访问和管理,应用程序不再需要包含要处理的文件和记录格式。

程序与数据相互独立有两个方面的含义。一方面是指当数据的存储方式发生变化(这里包括逻辑存储方式和物理存储方式)时,比如从链表结构改为哈希结构,或者从顺序存储转换为非顺序存储,应用程序不必进行任何修改。另一方面是指当数据的结构发生变化时,比如增加或减少了一些数据项,如果应用程序与这些修改的数据项无关,则应用程序也不用进行修改。这些变化都由DBMS负责维护。大多数情况下,应用程序并不知道数据存储方式或数据项已经发生了变化。

4. 保证数据的安全可靠

数据库技术能够保证数据库中的数据是安全可靠的。它有一套安全控制机制,可以有效地防止数据库中的数据被非法使用或非法修改。数据库中还有一套完整的备份和恢复机制,当数据遭到破坏时(由软件或硬件故障引起的),能够很快地将数据库恢复到正确的状态,并使数据不丢失或只有很少的丢失,从而保证系统能够连续、可靠地运行。

5. 最大限度地保证数据的正确性

保证数据的正确性是指存放到数据库中的数据必须符合实际情况,比如人的性别只能是“男”和“女”,人的年龄应该在0~150之间(假设没有年龄超过150岁的人),如果我们在“性

别”中输入了其他的值，或者将一个负数输入到“年龄”中，显然是不符合实际情况的。数据库系统能够保证进入到数据库中的数据都是正确的数据，这就是数据完整性。数据完整性是通过在数据库中建立约束来实现的。当我们建立好保证数据正确性的约束之后，如果有不符合约束条件的数据进入到数据库中，数据库就能主动拒绝这些数据。

6. 数据可以共享并能保证数据的一致性

数据库中的数据可以被多个用户共享，共享是指允许多个用户同时操作相同的数据。当然这个特性是针对大型的多用户数据库系统而言的，对于单用户系统，在任何时候最多只有一个用户访问数据库，因此不存在共享的问题。

多用户系统问题是数据库管理系统在内部解决的问题，它对用户是不可见的。这就要求数据库能够对多个用户进行协调，保证多个用户之间对数据进行的操作不发生矛盾和冲突，即在多个用户同时使用数据库时，能够保证数据的一致性和正确性。可以设想一下，在火车订票系统中，如果多个订票点同时对同一车次进行订票，那么必须保证不同订票点订出的座位不能相同。

数据集成与数据共享是大型环境中数据库系统的主要优点。

数据库技术发展到今天已经是一门比较成熟的技术，经过上面的讨论，我们可以概括出数据库具备如下特征：

数据库是相互关联的数据的集合，它用综合的方法组织数据，具有较小的数据冗余，可供多个用户共享，具有较高的数据独立性，具有安全控制机制，能够保证数据的安全可靠，允许并发地使用数据库，能有效、及时地处理数据，并能保证数据的一致性和完整性。

需要再次强调的是，所有这些特征并不是数据库中的数据所固有的，而是由数据库管理系统提供和保证的。

1.3 数据库系统的组成

前面我们介绍了数据库系统所具有的各种特征，那么什么是数据库系统呢？简单地说，数据库系统就是基于数据库的计算机应用系统。数据库系统一般包括三个部分：数据库、数据库管理系统、应用程序，如图1-3所示。其中，数据库是数据的汇集，它以一定的组织形式保存于存储介质上；DBMS是管理数据库的系统软件，它实现数据库系统的各种功能；应用程序是指以数据库以及数据库数据为基础的应用程序。

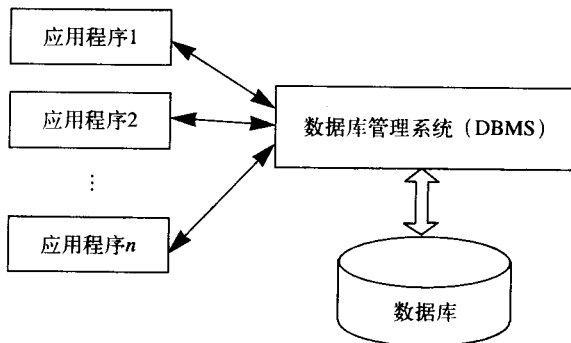


图1-3 数据库系统简图

除此之外，数据库系统还包括支持系统运行的计算机的硬件和操作系统环境。硬件环境是