

河南省高职高专计算机规划教材

# 数据库原理

主 编 苏 玉



西北大学出版社  
NORTHWEST UNIVERSITY PRESS

河南省高职高专计算机规划教材

# 数据库原理

主 编 苏 玉

副主编 卢中宁 李 丹

编 者 (以姓氏笔画为序)

丁爱萍 马金星 王志强

卢中宁 许绘香 严 峻

苏 玉 李 弋 李 丹

赵 靛

西北大学出版社

## 图书在版编目(CIP)数据

数据库原理 / 苏玉主编. — 西安: 西北大学出版社, 2006.2

高职高专计算机规划教材

ISBN 7-5604-2090-7

I. 数... II. 苏... III. 数据库系统 - 高等学校: 技术学校 - 教材 IV. TP311.13

中国版本图书馆 CIP 数据核字(2005)第 139620 号

## 数据库原理

主 编 苏玉

出版发行 西北大学出版社

通讯地址 西安市太白北路 229 号 邮编 710069

经 销 新华书店经销

印 刷 河南新华印务有限公司

开 本 787mm × 1092mm 1/16

印 张 21.25

字 数 484 千字

版 次 2006 年 2 月第 1 版 2006 年 2 月第 1 次印刷

书 号 ISBN 7-5604-2090-7/TP · 40

定 价 32.00 元

## 出版说明

当前,我国正在走新型工业化道路,坚持以信息化带动工业化,以工业化促进信息化,加快发展现代服务业,全面建设小康社会。走新型工业化道路,既需要一大批拔尖创新人才,也需要数以千万计的专门人才和数以亿计的高素质劳动者。根据劳动力市场技能型人才短缺的状况,国家决定实施技能型紧缺人才培养培训工程,其中计算机人才的培养培训是其重要的组成部分。

为适应高职高专计算机教育发展的需要,促进教育教学改革和教材建设,满足经济和社会发展对计算机人才的需求,根据教育部等六部门印发的《关于实施职业院校制造业和现代服务业技能型紧缺人才培养培训工程的通知》精神,按照教育部《关于制定〈2004~2007年职业教育教材开发编写计划〉的通知》要求,在教育部有关部门的支持和指导下,经河南省教育厅批准,我们组织有关专家,对计算机高职高专教育的培养目标和模式、课程体系、教学内容、教学方法和手段、教学实践等方面,进行了广泛而深入的调研。

在充分调研的基础上,在教育部有关部门和河南省教育厅的大力支持下,我们组织有关专家召开了计算机高职高专教育教研研讨会、教学大纲审定会和主编人会议,确定了教材编写的指导思想、原则和要求,组织全省近40所院校的一线教师,吸取了最新的计算机高职高专教育教学经验和成果,编写了这套教材。本套教材充分体现了以就业为导向,以职业技能培养为本的编写指导思想,突出了思想性、科学性、先进性、可读性和适用性的编写原则,较好地处理了“三基”(基本理论、基本知识、基本技能)关系,学历教育与职业认证、职业准入的关系。

这套教材虽经广泛调研与精心编撰,但一定还会存在这样或那样的不足,我们诚挚欢迎广大读者,尤其是选用该教材的教师和学生对其中的不足之处给予批评指正,以便我们在重印和修订时采纳有益的建议和意见,使之不断完善。

河南省高等学校计算机教育研究会

2006年1月

# 河南省高职高专计算机规划教材

## 参加编写学校

(以汉语拼音排序)

安阳师范学院

华北水利水电学院水利职业学院

黄河水利职业技术学院

河南财政税务高等专科学校

河南纺织高等专科学校

河南公安高等专科学校

河南工业大学

河南工业大学化学职业学院

河南经贸职业学院

河南农业职业学院

河南司法警官职业学院

河南商业高等专科学校

河南质量工程职业学院

济源职业技术学院

焦作大学

开封教育学院

开封人民警察学校

漯河职业技术学院

洛阳师范学院

平顶山工业职业技术学院

商丘师范学院

铁道警官高等专科学校

许昌职业技术学院

信阳农业高等专科学校

中原工学院广播影视职业学院

中州大学

郑州航空工业管理学院信息统计职业学院

郑州经济管理干部学院

郑州交通职业学院

郑州牧业工程高等专科学校

郑州轻工业学院

郑州轻工业学院轻工职业学院

郑州师范高等专科学校

郑州铁路职业技术学院

# 前 言

数据库原理及应用是数据管理的最新技术,是计算机科学的重要分支。近年来,数据库技术以及数据库的应用正得到突飞猛进的发展,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都有必要学习和掌握数据库知识。通过本课程的学习,使学生掌握数据库系统的基本知识,培养学生开发简单的数据库应用系统的能力,为以后学习大型数据库系统,如 SQL Server、ORACLE 等打下基础。

全书共分 12 章:第 1 章概要介绍了数据库技术,包括数据库管理技术的作用及相关概念、数据管理技术和数据库技术的发展、数据库系统体系结构及数据模型;第 2 章介绍了关系数据库模式、关系运算和优化查询;第 3 章介绍了数据库语言的分类,重点介绍了关系数据库标准语言 SQL 的组成、功能和特点;第 4 章介绍了关系数据库规范化理论;第 5 章介绍了数据库设计,包括数据库系统生存期和软件生存期,并重点介绍了概念模型的基本概念、实体-联系方法、实体-联系方法设计 E-R 模型的过程以及把 E-R 模型转换为关系模型的方法等;第 6 章介绍了数据库的安全性和完整性;第 7 章介绍了事务和并发控制;第 8 章介绍了数据库恢复技术;第 9 章介绍了网络环境下数据库系统的体系结构和安全问题,并详细介绍了分布式数据库和 Web 数据库的体系结构和技术要点;第 10 章介绍了数据库技术的发展趋势以及面向对象数据库、并行数据库、数据仓库技术等新兴的数据库系统;第 11 章介绍了 ORACLE 数据库管理系统、DB2 数据库管理系统和 SQL Server 2000 数据库管理等三个主流关系型数据库管理系统;第 12 章分别介绍了 Delphi、PowerBuilder、Visual BASIC 三种较为成熟的数据库开发工具的特点、与 SQL 数据库的连接方式及访问机制和数据库应用程序开发的全过程。

本书的最大特点是内容全面,既包括了数据库的基础理论知识、数据库技术发展趋势及数据库新技术,又包括了数据库的前端和后端的应用技术,如 SQL Server 2000、ORACLE、DB2 等,是目前广泛应用的后台数据库管理系统,Delphi、PowerBuilder、Visual BASIC 等也是非常普及、方便的可视化编程环境将这些内容结合在一本书中是作者的一种尝试,目的是使读者能够比较系统、全面地学习数据库的能力。本书的教学包括上机实践,建议总学时为 72 学时,其中理论教学 60 学时,实践教学 12 学时。

本书由苏玉任主编,负责统稿、修改及定稿。其中第 1、4 章由卢中宁编写,第 2、11 章由赵靓编写,第 3 章和实验部分由本戈编写,第 5 章由李丹编写,第 6 章由丁爱萍编写,第 7 章由苏玉编写,第 8 章由严峻编写,第 9 章由马金星编写,第 12 章由许绘香编写,第 10 章由苏玉、李丹和许绘香共同编写,王志强参与了部分章节的编写和修改工作。在本书的编写过程中,参考了大量国内外相关的资料,其主要部分列在书后参考文献中,在此对他们表示衷心地感谢!

由于时间仓促,加之作者水平有限,书中难免有不妥之处,望广大同仁予以批评指正。

编 者  
2005 年 6 月

# 目 录

|                     |    |
|---------------------|----|
| 第1章 数据库系统概论         | 1  |
| 1.1 数据库技术的发展        | 1  |
| 1.1.1 基本概念          | 1  |
| 1.1.2 数据管理技术的发展     | 2  |
| 1.2 数据库系统的基本概念      | 7  |
| 1.2.1 基本概念          | 7  |
| 1.2.2 数据库与数据库管理系统   | 9  |
| 1.2.3 数据库系统         | 9  |
| 1.2.4 数据库系统的作用      | 12 |
| 1.3 数据模型            | 13 |
| 1.3.1 数据模型概述        | 13 |
| 1.3.2 概念数据模型        | 14 |
| 1.3.3 结构数据模型        | 16 |
| 1.4 数据库系统体系结构       | 20 |
| 1.4.1 数据库三级模式结构     | 20 |
| 1.4.2 数据库的二级映像功能    | 22 |
| 1.4.3 数据库管理系统的功能与特征 | 23 |
| 小 结                 | 25 |
| 习 题                 | 25 |
| 第2章 关系数据库           | 27 |
| 2.1 关系数据库概述         | 27 |
| 2.1.1 关系数据库的基本概念    | 27 |
| 2.1.2 关系数据库模式       | 28 |
| 2.1.3 关系的完整性        | 32 |
| 2.2 关系运算(一)——关系代数   | 33 |
| 2.2.1 关系代数的基本概念     | 33 |
| 2.2.2 关系代数的五种基本操作   | 34 |
| 2.2.3 关系代数的其他操作     | 36 |
| 2.3 关系运算(二)——关系演算   | 38 |
| 2.3.1 元组关系演算        | 38 |
| 2.3.2 域关系的演算        | 39 |
| 2.3.3 关系运算的安全约束和等价性 | 40 |
| 2.4 优化查询            | 40 |

|       |                       |    |
|-------|-----------------------|----|
| 2.4.1 | 基本概念 .....            | 40 |
| 2.4.2 | 关系代数表达式的优化问题 .....    | 41 |
| 2.4.3 | 优化的策略 .....           | 41 |
| 2.4.4 | 优化算法 .....            | 42 |
| 小 结   | .....                 | 42 |
| 习 题   | .....                 | 43 |
| 第3章   | 关系数据库标准语言——SQL .....  | 44 |
| 3.1   | 数据库语言 .....           | 44 |
| 3.1.1 | 数据库语言概述 .....         | 44 |
| 3.1.2 | SQL 语言概述 .....        | 44 |
| 3.2   | 数据定义语言 .....          | 46 |
| 3.2.1 | 创建数据库 .....           | 46 |
| 3.2.2 | 创建表 .....             | 48 |
| 3.2.3 | 修改基本表和删除表 .....       | 49 |
| 3.2.4 | 定义和删除索引 .....         | 50 |
| 3.3   | 数据查询 .....            | 51 |
| 3.3.1 | 简单查询 .....            | 52 |
| 3.3.2 | 复杂查询 .....            | 56 |
| 3.4   | 数据操纵语言 .....          | 61 |
| 3.4.1 | 插入数据 .....            | 61 |
| 3.4.2 | 修改数据 .....            | 61 |
| 3.4.3 | 删除数据 .....            | 62 |
| 3.5   | 视 图 .....             | 62 |
| 3.5.1 | 定义视图 .....            | 63 |
| 3.5.2 | 删除视图 .....            | 64 |
| 3.5.3 | 查询视图 .....            | 64 |
| 3.5.4 | 操纵视图 .....            | 65 |
| 3.6   | 数据控制语言 .....          | 65 |
| 3.6.1 | 权限的类型 .....           | 65 |
| 3.6.2 | 授予权限 .....            | 66 |
| 3.6.3 | 回收权限 .....            | 67 |
| 3.7   | SQL 语言高级功能 .....      | 67 |
| 3.7.1 | 约束 .....              | 67 |
| 3.7.2 | 触发器 .....             | 68 |
| 3.7.3 | 嵌入式 SQL 和动态 SQL ..... | 68 |
| 3.7.4 | 游标 .....              | 69 |
| 小 结   | .....                 | 69 |
| 习 题   | .....                 | 70 |



|                    |     |
|--------------------|-----|
| <b>第4章 关系数据理论</b>  | 72  |
| 4.1 概 述            | 72  |
| 4.1.1 异常问题         | 72  |
| 4.1.2 关系模式规范化的必要性  | 74  |
| 4.2 函数依赖           | 75  |
| 4.2.1 函数依赖的定义      | 75  |
| 4.2.2 函数依赖和键码      | 76  |
| 4.3 范式与规范化         | 77  |
| 4.3.1 基本概念         | 77  |
| 4.3.2 第一范式(1NF)    | 78  |
| 4.3.3 第二范式(2NF)    | 80  |
| 4.3.4 第三范式(3NF)    | 81  |
| 4.3.5 BC 范式(BCNF)  | 82  |
| 4.3.6 多值依赖和4NF     | 83  |
| 4.4 关系模式的分解        | 86  |
| 4.4.1 无损连接         | 86  |
| 4.4.2 保持依赖         | 89  |
| 小 结                | 90  |
| 习 题                | 90  |
| <b>第5章 数据库设计</b>   | 92  |
| 5.1 概 述            | 92  |
| 5.1.1 数据库设计        | 92  |
| 5.1.2 软件生存期        | 95  |
| 5.1.3 数据库系统生存期     | 95  |
| 5.2 数据库设计的需求分析     | 96  |
| 5.2.1 需求分析的任务和目标   | 97  |
| 5.2.2 需求分析的步骤和方法   | 98  |
| 5.2.3 需求分析阶段的输入和输出 | 99  |
| 5.2.4 需求分析说明书      | 101 |
| 5.3 数据库的概念设计       | 102 |
| 5.3.1 概念设计的必要性     | 102 |
| 5.3.2 概念设计的策略和方法   | 102 |
| 5.3.3 E-R图的基本概念    | 104 |
| 5.3.4 E-R图的主要技术    | 105 |
| 5.3.5 采用E-R方法的概念设计 | 106 |
| 5.4 数据库的逻辑设计       | 110 |
| 5.4.1 逻辑设计的基本概念    | 110 |
| 5.4.2 E-R图向关系模式转换  | 111 |

|              |                               |            |
|--------------|-------------------------------|------------|
| 5.4.3        | 关系模式的规范化 .....                | 113        |
| 5.4.4        | 确定完整性约束 .....                 | 113        |
| 5.4.5        | 关系视图设计 .....                  | 114        |
| 5.5          | 数据库的物理设计 .....                | 115        |
| 5.5.1        | 物理设计的步骤 .....                 | 115        |
| 5.5.2        | 确定数据的存储结构 .....               | 115        |
| 5.5.3        | 确定数据的访问方式 .....               | 116        |
| 5.6          | 应用程序设计 .....                  | 117        |
| 5.7          | 数据库系统的实现与维护 .....             | 118        |
| 5.7.1        | 数据库系统的实现 .....                | 118        |
| 5.7.2        | 数据库系统的运行与维护 .....             | 119        |
| 小 结          | .....                         | 120        |
| 习 题          | .....                         | 120        |
| <b>第 6 章</b> | <b>数据库的安全性和完整性</b> .....      | <b>121</b> |
| 6.1          | 数据库的安全性 .....                 | 121        |
| 6.1.1        | 数据库安全性问题的提出 .....             | 121        |
| 6.1.2        | 数据库系统中安全措施 .....              | 121        |
| 6.1.3        | SQL 中的安全机制 .....              | 124        |
| 6.2          | 数据库完整性 .....                  | 136        |
| 6.2.1        | 数据库完整性的定义 .....               | 137        |
| 6.2.2        | 完整性约束条件 .....                 | 137        |
| 6.2.3        | 完整性控制 .....                   | 139        |
| 6.2.4        | 完整性规则和完整性策略 .....             | 139        |
| 6.2.5        | SQL 的完整性约束机制 .....            | 139        |
| 小 结          | .....                         | 141        |
| 习 题          | .....                         | 142        |
| <b>第 7 章</b> | <b>事务和并发控制</b> .....          | <b>143</b> |
| 7.1          | 事务 .....                      | 143        |
| 7.1.1        | 事务的基本概念和性质 .....              | 143        |
| 7.1.2        | 事务的调度 .....                   | 145        |
| 7.1.3        | 事务并发操作引发的问题 .....             | 146        |
| 7.1.4        | 事务并发调度的可串行性 .....             | 147        |
| 7.2          | 并发控制 .....                    | 149        |
| 7.2.1        | 事务的并发控制 .....                 | 149        |
| 7.2.2        | 封锁 .....                      | 149        |
| 7.2.3        | 封锁协议 .....                    | 151        |
| 7.2.4        | 活锁和死锁 .....                   | 156        |
| 7.3          | SQL Server 2000 的并发控制技术 ..... | 158        |

|                                 |     |
|---------------------------------|-----|
| 小 结.....                        | 159 |
| 习 题.....                        | 160 |
| <b>第8章 数据库恢复技术</b> .....        | 161 |
| 8.1 数据库系统的故障 .....              | 161 |
| 8.2 数据库的备份 .....                | 163 |
| 8.2.1 数据库备份的必要性 .....           | 167 |
| 8.2.2 数据库备份的种类 .....            | 167 |
| 8.2.3 数据库备份计划 .....             | 168 |
| 8.3 数据库的恢复 .....                | 171 |
| 8.3.1 数据库恢复的实现技术 .....          | 171 |
| 8.3.2 恢复策略 .....                | 172 |
| 8.3.3 数据库的恢复与镜像 .....           | 173 |
| 小 结.....                        | 175 |
| 习 题.....                        | 175 |
| <b>第9章 网络与数据库</b> .....         | 176 |
| 9.1 概述 .....                    | 176 |
| 9.1.1 网络环境下数据库应用系统的安全问题 .....   | 176 |
| 9.1.2 网络环境下数据库应用系统的体系结构 .....   | 178 |
| 9.1.3 网络环境下的数据库应用规划 .....       | 180 |
| 9.2 分布式数据库 .....                | 181 |
| 9.2.1 分布式数据库的概念 .....           | 181 |
| 9.2.2 分布式数据库的体系结构 .....         | 183 |
| 9.2.3 分布式数据库的查询处理 .....         | 185 |
| 9.2.4 分布式数据库的事务管理 .....         | 187 |
| 9.3 Web 数据库 .....               | 188 |
| 9.3.1 Web 概述 .....              | 189 |
| 9.3.2 Web 数据库技术要点 .....         | 191 |
| 9.3.3 Web 数据库应用的构成及一般结构 .....   | 193 |
| 小 结.....                        | 195 |
| 习 题.....                        | 195 |
| <b>第10章 数据库技术发展趋势与新技术</b> ..... | 196 |
| 10.1 数据库技术的发展趋势.....            | 196 |
| 10.1.1 数据库技术面临的挑战 .....         | 196 |
| 10.1.2 数据库技术的未来 .....           | 197 |
| 10.2 数据库新技术概述.....              | 198 |
| 10.2.1 数据模型的新特征 .....           | 198 |
| 10.2.2 数据库系统的新特征 .....          | 199 |
| 10.2.3 现代数据库系统的范畴 .....         | 200 |

|        |                                |     |
|--------|--------------------------------|-----|
| 10.2.4 | 现代数据库主要领域 .....                | 201 |
| 10.2.5 | 现代数据库实现途径 .....                | 204 |
| 10.3   | 面向对象数据库 .....                  | 205 |
| 10.3.1 | 概述 .....                       | 205 |
| 10.3.2 | 面向对象数据库系统 OODBS 的基本特征 .....    | 206 |
| 10.3.3 | 面向对象数据模型 .....                 | 207 |
| 10.3.4 | 面向对象数据库管理系统 OODBMS .....       | 209 |
| 10.3.5 | 面向对象数据库系统的基本功能和实现途径 .....      | 211 |
| 10.3.6 | 对象 - 关系数据库系统 .....             | 211 |
| 10.4   | 并行数据库系统 .....                  | 213 |
| 10.4.1 | 并行数据库系统的体系结构 .....             | 213 |
| 10.4.2 | 并行数据库系统中数据的划分技术 .....          | 216 |
| 10.4.3 | 并行查询处理技术 .....                 | 217 |
| 10.5   | 数据仓库技术 .....                   | 218 |
| 10.5.1 | 概述 .....                       | 219 |
| 10.5.2 | 数据挖掘和数据中心库 .....               | 220 |
| 10.5.3 | 数据仓库的体系结构 .....                | 222 |
| 10.5.4 | 数据仓库的关键技术 .....                | 222 |
| 10.5.5 | 数据仓库的实现步骤 .....                | 222 |
| 10.6   | 其他新型的数据库系统 .....               | 223 |
| 10.6.1 | 多媒体数据库系统 .....                 | 223 |
| 10.6.2 | 主动数据库系统 .....                  | 224 |
| 10.6.3 | 模糊数据库系统 .....                  | 224 |
| 小 结    | .....                          | 225 |
| 习 题    | .....                          | 226 |
| 第 11 章 | 主要数据库产品 .....                  | 227 |
| 11.1   | ORACLE 数据库 .....               | 227 |
| 11.1.1 | ORACLE 系统概述 .....              | 227 |
| 11.1.2 | 数据库和实例的启动和关闭 .....             | 229 |
| 11.1.3 | 数据字典 .....                     | 230 |
| 11.2   | DB2 数据库 .....                  | 231 |
| 11.2.1 | DB2 数据库的特点 .....               | 231 |
| 11.2.2 | DB2 的安装类型和软硬件要求 .....          | 233 |
| 11.2.3 | DB2 控制中心 .....                 | 234 |
| 11.3   | SQL Server 数据库 .....           | 235 |
| 11.3.1 | SQL Server 2000 主要特点 .....     | 235 |
| 11.3.2 | SQL Server 2000 版本及环境需求 .....  | 236 |
| 11.3.3 | SQL Server 2000 完整的管理工具集 ..... | 238 |

|                                            |            |
|--------------------------------------------|------------|
| 小 结.....                                   | 240        |
| 习 题.....                                   | 240        |
| <b>第 12 章 数据库开发工具 .....</b>                | <b>241</b> |
| 12.1 Delphi/SQL Server 数据库开发 .....         | 241        |
| 12.1.1 Delphi 简介 .....                     | 241        |
| 12.1.2 通过 Delphi 访问 SQL Server 数据库 .....   | 242        |
| 12.1.3 开发实例 .....                          | 249        |
| 12.2 PowerBuilder/SQL Server 数据库开发 .....   | 257        |
| 12.2.1 PowerBuilder 简介 .....               | 257        |
| 12.2.2 通过 PowerBuilder 访问 SQL 数据库 .....    | 258        |
| 12.2.3 开发实例 .....                          | 268        |
| 12.3 Visual BASIC/SQL Server 数据库开发 .....   | 277        |
| 12.3.1 Visual BASIC 简介 .....               | 277        |
| 12.3.2 通过 Visual BASIC 访问 SQL Server ..... | 278        |
| 12.3.3 开发实例 .....                          | 284        |
| 小 结.....                                   | 293        |
| 习 题.....                                   | 293        |
| <b>实 验.....</b>                            | <b>294</b> |
| 实验 1 SQL Server 的安装及其管理工具的使用 .....         | 294        |
| 实验 2 创建数据库及基本表 .....                       | 296        |
| 实验 3 修改基本表、索引的创建与删除 .....                  | 305        |
| 实验 4 数据查询 .....                            | 313        |
| 实验 5 视图的定义与操作 .....                        | 315        |
| 实验 6 数据控制操作 .....                          | 318        |
| <b>参考文献.....</b>                           | <b>321</b> |

# 第 1 章 数据库系统概论

数据库处理在信息系统的研究中一直都是一个非常重要的课题,数据库技术是数据管理的最新技术,是计算机科学技术中发展最快的领域之一,它已成为计算机信息系统的核心技术和重要基础。数据库已经成为人们储存数据、管理信息、共享资源的最先进最常用的技术。

本章介绍数据库系统的基本概念,包括数据管理技术的发展过程、数据库系统的基本概念、数据模型及数据库系统的体系结构等。读者从中可以学习到为什么要使用数据库技术以及数据库技术的重要性。本章是学习后面各章节的预备和基础。

## 1.1 数据库技术的发展

数据库技术研究的是如何科学正确地组织、存储数据,如何高效地获取和处理数据,是由文件管理系统发展起来的一种理想数据管理技术。发展数据库技术的目的是充分地开发与利用大量信息资源,对信息进行识别,存储、处理与传递,推导人类的行为,或为决策提供依据。

### 1.1.1 基本概念

#### 1. 信息处理

信息处理是对信息进行收集、保存、传输、加工和利用的过程,最终体现为对数据的处理。除暗示外,几乎所有的信息都可表示为计算机能识别的字符串或位串,这为信息的传播提供了可能。使人们对信息、数据的利用和处理实现自动化、网络化和社会化,整个人类社会成为信息化社会。

#### 2. 数据处理

数据处理是对各种形式的数据进行收集、储存、加工和传播的一系列活动的总和。其目的是从大量原始数据中抽取、推导出有价值的信息作为行动和决策的依据。数据处理的三个阶段:

(1) 手工处理阶段(原始社会—19 世纪)

工具:算盘,计算尺,微分机,基于齿轮的六位加法器

特点:计算工具极低级,精度差,处理能力低,离不开手工。

## (2) 机械处理阶段(1890—1946)

工具:机械计算工具(卡片制表机)

特点:机械代替了手工操作,数据处理能力、效率有所提高,但受机械性能限制。

## (3) 电子处理阶段(1946— )

工具:计算机

特点:速度快,存储量大,传输速度快,逻辑推理能力强、不疲劳、可靠性好、精度高,不但能进行科学计算而且能进行数据处理。

### 3. 数据管理(技术)

数据管理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。在应用需求的推动下,在计算机硬件、软件发展的基础上,它经历了人工管理、文件系统、数据库系统三个阶段(用计算机管理)。

#### 1.1.2 数据管理技术的发展

本节将以一个学校的教务处对学生、课程和成绩的管理来阐述数据管理各阶段的特点。在没有使用计算机的时候,学校教务处的有关人员将有关学生的信息抄写在一张张的卡片上,为了方便查找,将同一个系、同一个年级、同一个班级的学生的卡片存放在相邻的地方,并对不同的系、年级和班级做上标签。每门课程的信息也是抄写在卡片上,将同一个专业的卡片放在一起并做上标签。每个学期末将同一个班的各门课的成绩单收集起来存放在档案中。

当查找一个学生的信息时,如果知道他所在的系和班级,按照建好的标签可以很快找到该学生的卡片。如果只知道他的姓名,那只有在所有的学生卡片中一个一个查找,需要花费很多的时间。

当计算一个学生某个学期的平均成绩时,首先在档案中找到该学生所在班级这个学期的所有成绩单,从中找出该学生各门课程的成绩,再算平均成绩。统计某一门课的成绩分布时也要进行类似的处理和计算。

##### 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。当时的硬件状况是,外存只有纸带、卡片、磁带,没有磁盘等直接存取的存储设备;软件状况是,没有操作系统,没有管理数据的软件,因此称这样的数据管理方式为人工管理数据。人工管理数据具有如下特点:

###### (1) 数据不保存

由于当时计算机主要用于科学计算,一般不需要将数据长期保存,只是在计算某一课题时将数据输入,用完就撤走。

如果要用计算机统计分析全校每一门课的成绩,就要编写统计分析程序,在运行该程序时读入相应的学生选修课程成绩单等数据,计算完成后数据和程序都撤走,不再在计算机中保存。

###### (2) 应用程序管理数据

数据需要由应用程序自己管理,没有相应的软件系统负责数据的管理工作。应用程序中不仅要规定数据的逻辑结构,而且要设计物理结构,包括存储结构、存取方法、输入方式等,因此程序员负担很重。

(3)数据不共享

数据是面向应用的,一组数据只能对应一个程序。当多个应用程序涉及某些相同的数据时,由于必须各自定义,无法互相利用、互相参照,因此程序与程序之间有大量的冗余数据。例如教务处要统计分析某一门课的成绩和统计分析某一个学生的成绩,要编写两个程序,尽管都要使用学生选修课程成绩单,但是每个程序要分别定义两个成绩单数据,分别输入,分别使用,如图 1.1 所示。

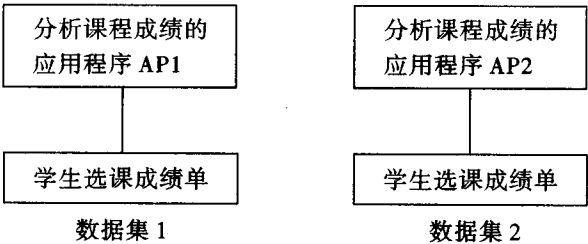


图 1.1 两个应用程序使用同一数据

(4)数据不具有独立性

数据的逻辑结构或物理结构改变后,必须对应用程序做相应的修改,这就进一步加重了程序员的负担。例如,学生成绩由 5 级记分制改为百分制时,上面两个统计分析程序都要修改。

在人工管理阶段,程序与数据之间的一一对应关系如图 1.2 所示。

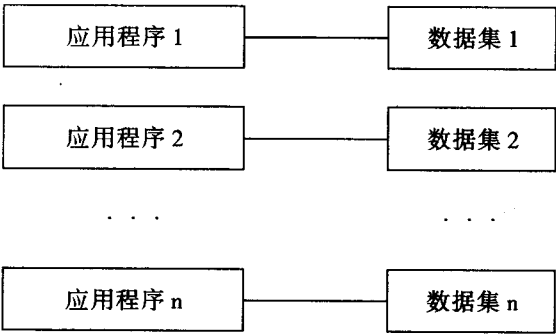


图 1.2 人工管理阶段应用程序与数据之间的对应关系

2. 文件系统阶段

20 世纪 50 年代后期到 60 年代中期,硬件方面已有了磁盘、磁鼓等直接存取存储设备;软件方面,操作系统中已经有了专门的数据管理软件——文件系统。这时可以把相关



的数据组织成一个文件存放在计算机中,在需要的时候只要提供文件名,计算机就能从文件系统中找出所要的文件,把文件中存储的数据提供给用户进行处理。

例如,在学校教务处对学生学籍的管理中,为了改变查找、计算工作量大及花费时间长的被动局面,将学生卡片、课程卡片和学生成绩单中的内容存放到文件 student、course、study 中。并对每个文件编写一组程序用于数据维护,包括增加、删除、修改一条记录和查询一条记录。在此基础上根据需要编写一些查询和报表打印程序,例如根据学生的姓名、学生编号查找学生的信息,统计某学期某个学生的平均成绩,统计某门课的平均成绩等。

管理系统投入应用后,教务处的工作效率大大提高。例如,每学期末将各门课的考试成绩输入计算机以后,可以很快计算出学生的平均成绩,并打印出需要补考的学生的名单。

但经过一段时间的使用后,教务人员发现有时必须修改程序和文件结构才能适应工作的需要。例如,学校领导要求统计不同生源地的学生成绩,因为原来的学籍管理软件中没有实现这个功能,必须编写一段程序来实现这个功能。又如当需要在 student 文件中增加“个人网址”属性时,这涉及到改变文件的结构,需要若干步骤才能完成:

第一步,建立一个新文件 student-new,其结构是在 student 的结构中加入“个人网址”这一项。

第二步,编写一个程序将文件 student 中的数据转存到 student-new 中。

第三步,删除文件 student。

第四步,将文件 student-new 重命名为 student。

这项工作到此并没有结束,因为文件中保存的是数据,不保存数据的结构,数据结构是在程序中定义的,文件 student 的结构写到了所有使用它的程序中,必须一一修改这些程序以适应新的文件结构,否则程序运行就会出错。

从这个例子可以看到用文件系统管理数据的优点和不足。一般地讲,用文件系统管理数据具有如下特点:

(1) 数据可以长期保存

数据可以组织成文件长期保存在计算机中反复使用。

(2) 由文件系统管理数据

文件系统把数据组织成内部有结构的记录,实现“按文件名访问,按记录进行存取”的管理技术。

文件系统使应用程序与数据之间有了初步的独立性,程序员不必过多的考虑数据存储的物理细节。例如,文件系统中可以有顺序结构文件、索引结构文件、Hash 等。数据在存储上的不同不会影响程序的处理逻辑。如果数据的存储结构发生改变,应用程序的改变很小,节省了程序的维护工作量。但是,文件系统仍存在以下缺点:

(1) 数据共享性差,冗余度大

在文件系统中,一个(或一组)文件基本上对应于一个应用(程序),即文件是面向应用的。当不同的应用(程序)使用部分相同的数据时,也必须建立各自的文件,而不能共享相同的数据。因此数据的冗余度大,浪费存储空间。同时由于相同数据的重复存储、各