

高等学校计算机系列

# 数据库原理与设计

杨海霞 主编  
相 洁 南志红 副主编  
相万让 审

 人民邮电出版社  
POSTS & TELECOM PRESS

TP311.13/307

2007

高等学校计算机系列

# 数据库原理与设计

杨海霞 主编  
相 洁 南志红 副主编  
相万让 审

人民邮电出版社

北 京

## 图书在版编目 (CIP) 数据

数据库原理与设计 / 杨海霞主编. —北京: 人民邮电出版社, 2007.9

高等学校计算机系列

ISBN 978-7-115-16662-3

I. 数… II. 杨… 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 (2007) 第 122541 号

## 内 容 提 要

本书系统地讲述了数据库的基本理论和开发技术。全书分为三篇, 主要内容包括: 数据库系统的组成、数据模型、关系代数、SQL、关系规范化理论、数据库的设计方法、SQL 程序设计、数据库的访问技术、维护技术以及数据库技术新进展等, 每章后均附有习题。

本书语言通俗易懂, 既有理论的概括与探讨, 又有实际应用中的经验方法的总结, 本书可作为高等学校计算机、信息及相关专业的数据库教材, 也适合从事数据库开发人员参考阅读。

高等学校计算机系列

## 数据库原理与设计

◆ 主 编 杨海霞

副 主 编 相 洁 南志红

审 相万让

责任编辑 邹文波

◆ 人民邮电出版社出版发行 北京市崇文区夕照寺街 14 号

邮编 100061 电子函件 315@ptpress.com.cn

网址 <http://www.ptpress.com.cn>

北京市通州大中印刷厂

新华书店总店北京发行所经销

◆ 开本: 787×1092 1/16

印张: 21.5

字数: 516 千字

印数: 1—3 000 册

2007 年 9 月第 1 版

2007 年 9 月北京第 1 次印刷

ISBN 978-7-115-16662-3/TP

定价: 29.80 元

读者服务热线: (010)67170985 印装质量热线: (010)67129223

# 前言

数据库技术一直是计算机科学技术中发展最快的领域和应用最广的技术之一，但长期以来，数据库作为计算机及其相关专业的一门课程，其教学一直处于以理论为主的状态，致使教师感觉抽象难教，学生认为枯燥难学。究其原因，主要是由于单纯地讲解理论，不能把数据库的理论知识依托于一个具体的平台来就事论理。只有把数据库原理与实际的应用开发有机结合，增强学生实践动手能力，才能培养出真正满足社会需求的数据库技术人才。正是在此动机推动下，笔者根据多年的教学及项目开发的实际经验，有感而发，写成此书。

与其他同类教材相比，本书具有以下特点。

(1) 理论依托于具体软件。本书以 SQL Server 2000 作为数据库管理系统，以 Power Designer 为设计工具，把数据库的基本概念、基本理论和基本技术及数据库的设计贯穿于一个学籍管理实例中，通过详尽的分析，使学生感性认识数据库，从而加深对数据库理论知识的理解。

(2) 循序渐进，步步深入。本书从数据库的基本知识开始，首先讲解每个数据库对象的基本功能和基本操作，然后应用前面的知识进行学籍管理数据库的设计，从中领会数据库的设计技巧，最后介绍数据库的访问技术和维护技术。这样，数据库应用系统设计开发的全过程便清晰地展示出来了。

(3) 注重应用，有配套的实验指导。为加强学生的实际操作能力，精心编写了配套的实验指导书，通过实验篇、设计篇、训练篇和实例篇四部分的训练，由点及面，使理论教学与实验教学环环紧扣，既方便教师教学，又有助于学生学习，具有很强的实用性。

全书分为三篇。第一篇知识篇，主要介绍数据库系统的基本概念和基本理论。这部分由第 1 章至第 4 章组成，具体内容包括数据管理的发展过程、数据库系统的组成、关系代数、基本 SQL 语句的使用、数据完整性约束的实现方法、视图、存储过程和触发器的概念及定义方法、关系规范化理论、事务及并发控制、备份恢复机制等。

第二篇设计篇，主要介绍数据库的设计方法和 SQL 程序设计。这部分由第 5 章和第 6 章组成，第 5 章的具体内容包括数据库的概念结构设计、逻辑结构设计、物理结构设计、数据库的维护和实施以及如何通过 PowerDesigner 来实现各个过程；第 6 章介绍了 SQL 流程控制语句、游标的使用、SQL 程序的调试与错误处理以及如何编写存储过程和自定义函数实现复杂的数据查询功能等。

第三篇技术篇，主要介绍数据库的访问技术、维护技术以及数据库技术的新进展。这部分由第 7 章至第 10 章组成，具体内容包括使用 ODBC 访问数据库、使用 Visual Basic 访问数据库、使用 ASP 访问数据库、SQL Server 的数据维护、数据仓库与数据挖掘技术、面向对象数据库系统、分布式数据库系统、并行数据库系统、空间数据库、多媒体数据库系统和模糊数据库系统等。

本书语言通俗易懂，既有理论的概括与探讨，又有实际的经验方法的总结，本书可作为高等学校计算机、信息及相关专业的数据库课程教科书。为便于教师进行教学，本书配有 PPT

格式的课件素材，可到人民邮电出版社的网站下载。

本书由杨海霞、相洁、南志红、殷鹏编写。其中，第1章、第3章由杨海霞编写，第2章由杨海霞、殷鹏编写，第4章、第7章、第8章由南志红编写，第5篇、第6章、第9章、第10章、附录A和附录B由相洁编写。全书由杨海霞统稿，相万让审阅。在编写过程中得到了徐仲安教授、容和平教授、张永奎教授、樊正棠教授、何吉成教授、王建珍教授的支持与帮助，同时，本书的出版还得到了山西财经大学、太原理工大学等高校教师的大力支持，在此一并致以衷心的感谢。

由于作者水平有限，书中难免存在疏漏、欠妥之处，敬请读者批评指正。

编 者

2007年6月

# 目 录

## 第一篇 知 识 篇

|                                  |    |
|----------------------------------|----|
| <b>第 1 章 数据库系统概述</b> .....       | 1  |
| 1.1 引言 .....                     | 1  |
| 1.1.1 数据库管理技术的发展 .....           | 1  |
| 1.1.2 数据库系统的组成 .....             | 4  |
| 1.1.3 数据库技术的发展 .....             | 4  |
| 1.2 数据库系统结构 .....                | 5  |
| 1.2.1 数据库系统的模式结构 .....           | 5  |
| 1.2.2 数据库系统的体系结构 .....           | 8  |
| 1.3 数据模型 .....                   | 10 |
| 1.3.1 数据表示 .....                 | 11 |
| 1.3.2 数据模型的要素 .....              | 11 |
| 1.3.3 概念模型 .....                 | 12 |
| 1.3.4 常用的数据模型 .....              | 15 |
| 1.4 关系数据库 .....                  | 17 |
| 1.4.1 关系模型概述 .....               | 17 |
| 1.4.2 关系数据结构及形式化定义 .....         | 18 |
| 1.4.3 关系的完整性 .....               | 22 |
| 1.4.4 关系代数 .....                 | 24 |
| 1.5 SQL Server 数据库管理系统 .....     | 31 |
| 1.5.1 SQL Server 概述 .....        | 31 |
| 1.5.2 SQL Server 中的数据库对象 .....   | 33 |
| 思考与练习 .....                      | 36 |
| <b>第 2 章 关系数据库语言</b> .....       | 37 |
| 2.1 SQL 概述 .....                 | 38 |
| 2.1.1 SQL 的特点 .....              | 38 |
| 2.1.2 SQL 数据库的体系结构 .....         | 39 |
| 2.1.3 SQL 语句的分类 .....            | 40 |
| 2.1.4 SQL 的语法约定 .....            | 41 |
| 2.2 SQL Server 数据库的基本管理与设置 ..... | 43 |
| 2.2.1 启动和关闭数据库服务 .....           | 43 |
| 2.2.2 创建数据库 .....                | 45 |
| 2.2.3 修改数据库 .....                | 47 |
| 2.2.4 设置数据库选项 .....              | 48 |

|                          |            |
|--------------------------|------------|
| 2.2.5 删除数据库 .....        | 49         |
| 2.3 数据定义 .....           | 49         |
| 2.3.1 表的创建与管理 .....      | 50         |
| 2.3.2 索引的创建与管理 .....     | 55         |
| 2.3.3 视图的创建与管理 .....     | 57         |
| 2.3.4 存储过程的创建与管理 .....   | 60         |
| 2.3.5 触发器的创建与管理 .....    | 62         |
| 2.4 数据查询 .....           | 64         |
| 2.4.1 单表查询 .....         | 65         |
| 2.4.2 连接查询 .....         | 71         |
| 2.4.3 嵌套查询 .....         | 74         |
| 2.4.4 集合查询 .....         | 77         |
| 2.5 数据更新 .....           | 77         |
| 2.5.1 插入数据 .....         | 77         |
| 2.5.2 修改数据 .....         | 78         |
| 2.5.3 删除数据 .....         | 79         |
| 2.5.4 更新操作与数据库的一致性 ..... | 80         |
| 思考与练习 .....              | 80         |
| <b>第3章 关系数据理论</b> .....  | <b>83</b>  |
| 3.1 关系模式设计中的问题 .....     | 83         |
| 3.1.1 规范化理论的主要内容 .....   | 83         |
| 3.1.2 关系模式的操作异常 .....    | 83         |
| 3.2 函数依赖 .....           | 86         |
| 3.2.1 函数依赖的定义及性质 .....   | 86         |
| 3.2.2 函数依赖的分类 .....      | 88         |
| 3.2.3 码 .....            | 89         |
| 3.3 范式 .....             | 89         |
| 3.3.1 第一范式 .....         | 89         |
| 3.3.2 第二范式 .....         | 90         |
| 3.3.3 第三范式 .....         | 91         |
| 3.3.4 BC 范式 .....        | 93         |
| 3.3.5 多值依赖与 4NF .....    | 95         |
| 3.4 关系模式的规范化 .....       | 99         |
| 3.4.1 关系模式规范化的原则 .....   | 99         |
| 3.4.2 关系模式规范化的步骤 .....   | 99         |
| 3.4.3 关系模式规范化的要求 .....   | 100        |
| 思考与练习 .....              | 103        |
| <b>第4章 数据库保护</b> .....   | <b>104</b> |
| 4.1 安全性 .....            | 104        |

|                        |     |
|------------------------|-----|
| 4.1.1 安全性控制概述 .....    | 104 |
| 4.1.2 安全性控制的一般方法 ..... | 105 |
| 4.2 完整性 .....          | 109 |
| 4.2.1 完整性约束的分类 .....   | 110 |
| 4.2.2 完整性控制 .....      | 112 |
| 4.3 并发控制 .....         | 114 |
| 4.3.1 并发控制概述 .....     | 115 |
| 4.3.2 并发操作的调度 .....    | 118 |
| 4.3.3 封锁 .....         | 119 |
| 4.3.4 死锁和活锁 .....      | 120 |
| 4.4 恢复 .....           | 122 |
| 4.4.1 恢复的原理 .....      | 122 |
| 4.4.2 恢复的实现 .....      | 124 |
| 4.5 数据库复制与数据库镜像 .....  | 129 |
| 4.5.1 数据库复制 .....      | 129 |
| 4.5.2 数据库镜像 .....      | 130 |
| 思考与练习 .....            | 130 |

## 第二篇 设计篇

|                             |     |
|-----------------------------|-----|
| 第5章 数据库设计 .....             | 132 |
| 5.1 数据库设计步骤 .....           | 132 |
| 5.2 需求分析 .....              | 133 |
| 5.2.1 需求分析方法 .....          | 134 |
| 5.2.2 数据流图与数据字典 .....       | 135 |
| 5.2.3 实例——学籍管理需求分析 .....    | 138 |
| 5.3 概念结构设计 .....            | 142 |
| 5.3.1 数据抽象与局部视图设计 .....     | 142 |
| 5.3.2 视图的集成 .....           | 143 |
| 5.3.3 实例——学籍管理概念结构设计 .....  | 144 |
| 5.4 逻辑结构设计 .....            | 146 |
| 5.4.1 E-R 图向关系数据模型的转换 ..... | 147 |
| 5.4.2 数据模型的优化 .....         | 148 |
| 5.4.3 设计用户子模式 .....         | 149 |
| 5.4.4 实例——学籍管理逻辑结构设计 .....  | 149 |
| 5.5 物理结构设计 .....            | 151 |
| 5.5.1 分析影响数据库物理设计的因素 .....  | 152 |
| 5.5.2 关系模式存取方法选择 .....      | 152 |
| 5.5.3 确定系统配置 .....          | 154 |
| 5.5.4 评价物理结构 .....          | 154 |



|                                            |            |
|--------------------------------------------|------------|
| 5.6 数据库实施与维护 .....                         | 154        |
| 5.6.1 创建数据库 .....                          | 154        |
| 5.6.2 组织数据入库 .....                         | 154        |
| 5.6.3 数据库试运行 .....                         | 155        |
| 5.6.4 数据库运行与维护 .....                       | 155        |
| 5.7 数据库建模工具 .....                          | 156        |
| 5.7.1 数据库建模工具简介 .....                      | 156        |
| 5.7.2 利用 PowerDesigner 设计学籍管理的概念结构实例 ..... | 158        |
| 5.7.3 利用 PowerDesigner 设计学籍管理的物理结构实例 ..... | 163        |
| 5.7.4 利用 PowerDesigner 进行学籍管理的数据库实施 .....  | 163        |
| 思考与练习 .....                                | 167        |
| <b>第 6 章 SQL 程序设计与开发</b> .....             | <b>168</b> |
| 6.1 批处理与脚本 .....                           | 168        |
| 6.1.1 批处理 .....                            | 168        |
| 6.1.2 脚本 .....                             | 169        |
| 6.2 SQL 程序设计基础 .....                       | 170        |
| 6.2.1 SQL 程序基本成分 .....                     | 170        |
| 6.2.2 SQL 程序编写规范 .....                     | 173        |
| 6.3 流程控制语句 .....                           | 174        |
| 6.3.1 语句块: BEGIN...END .....               | 174        |
| 6.3.2 条件执行: IF...ELSE 语句 .....             | 175        |
| 6.3.3 多分支 CASE 表达式 .....                   | 176        |
| 6.3.4 循环: WHILE 语句 .....                   | 178        |
| 6.3.5 非条件执行: GOTO 语句 .....                 | 179        |
| 6.3.6 调度执行: WAITFOR .....                  | 179        |
| 6.4 游标 .....                               | 180        |
| 6.4.1 游标 .....                             | 180        |
| 6.4.2 声明游标: DECLARE CURSOR 语句 .....        | 181        |
| 6.4.3 打开游标: OPEN 语句 .....                  | 182        |
| 6.4.4 读取数据: FETCH 语句 .....                 | 182        |
| 6.4.5 关闭游标: CLOSE 语句 .....                 | 184        |
| 6.4.6 释放游标: DEALLOCATE 语句 .....            | 184        |
| 6.4.7 游标使用实例 .....                         | 184        |
| 6.5 SQL 程序的调试与错误处理 .....                   | 186        |
| 6.5.1 SQL 程序的错误类型 .....                    | 186        |
| 6.5.2 SQL 程序的错误处理 .....                    | 187        |
| 6.6 SQL 程序实例 .....                         | 187        |
| 6.6.1 自定义函数 .....                          | 188        |
| 6.6.2 存储过程 .....                           | 189        |

|                 |     |
|-----------------|-----|
| 6.6.3 触发器 ..... | 192 |
| 思考与练习 .....     | 193 |

### 第三篇 技 术 篇

|                                        |     |
|----------------------------------------|-----|
| <b>第 7 章 数据库访问技术</b> .....             | 195 |
| 7.1 数据库访问技术概述 .....                    | 195 |
| 7.1.1 数据库访问技术发展概况 .....                | 195 |
| 7.1.2 数据库访问技术 .....                    | 196 |
| 7.2 使用 ODBC 访问数据库 .....                | 199 |
| 7.2.1 ODBC 层次结构 .....                  | 199 |
| 7.2.2 配置 ODBC 数据源 .....                | 200 |
| 7.2.3 通过 ODBC 访问 SQL Server 数据库 .....  | 203 |
| 7.2.4 处理 ODBC 错误 .....                 | 207 |
| 7.3 嵌入式 SQL 与动态 SQL .....              | 208 |
| 7.3.1 嵌入式 SQL .....                    | 208 |
| 7.3.2 动态 SQL 语句 .....                  | 211 |
| 7.4 使用 Visual Basic 访问数据库 .....        | 212 |
| 7.4.1 数据访问对象 .....                     | 212 |
| 7.4.2 使用 RDO 模型访问 SQL Server 数据库 ..... | 215 |
| 7.4.3 使用 ADO 模型访问 SQL Server 数据库 ..... | 218 |
| 7.4.4 使用数据环境设计器访问 SQL Server .....     | 221 |
| 7.5 使用 ASP 访问数据库 .....                 | 224 |
| 7.5.1 ASP 编程基本知识 .....                 | 224 |
| 7.5.2 连接数据库 .....                      | 225 |
| 7.5.3 从数据库到 Web 页 .....                | 226 |
| 7.5.4 Errors 错误集合与 Error 错误对象 .....    | 231 |
| 7.5.5 执行 SQL Server 存储过程 .....         | 233 |
| 7.6 应用实例 .....                         | 234 |
| 7.6.1 功能描述 .....                       | 234 |
| 7.6.2 VB 程序 .....                      | 234 |
| 7.6.3 ASP 程序 .....                     | 240 |
| 思考与练习 .....                            | 245 |
| <b>第 8 章 SQL Server 数据库维护技术</b> .....  | 246 |
| 8.1 安全控制 .....                         | 246 |
| 8.1.1 SQL Server 的安全体系结构和安全认证模式 .....  | 246 |
| 8.1.2 SQL Server 的安全性管理 .....          | 249 |
| 8.2 SQL Server 的数据完整性 .....            | 257 |
| 8.2.1 SQL Server 的数据完整性种类 .....        | 258 |
| 8.2.2 SQL Server 数据完整性实现方式 .....       | 258 |

|                                             |            |
|---------------------------------------------|------------|
| 8.2.3 SQL Server 数据完整性的实现方法 .....           | 258        |
| 8.3 SQL Server 的并发控制 .....                  | 262        |
| 8.3.1 SQL Server 的事务类型 .....                | 262        |
| 8.3.2 SQL Server 锁的粒度和类型 .....              | 262        |
| 8.4 SQL Server 的数据备份与数据恢复 .....             | 263        |
| 8.4.1 SQL Server 备份方式 .....                 | 264        |
| 8.4.2 SQL Server 数据库备份 .....                | 264        |
| 8.4.3 SQL Server 数据库恢复 .....                | 268        |
| 8.5 数据库复制与数据库镜像 .....                       | 270        |
| 8.5.1 SQL Server 数据库复制 .....                | 270        |
| 8.5.2 SQL Server 数据库镜像 .....                | 280        |
| 8.6 数据转换和集成 .....                           | 281        |
| 8.6.1 数据转换服务 DTS .....                      | 281        |
| 8.6.2 DTS 导入/导出向导的使用 .....                  | 282        |
| 思考与练习 .....                                 | 286        |
| <b>第 9 章 数据挖掘与数据仓库技术 .....</b>              | <b>287</b> |
| 9.1 数据挖掘 .....                              | 287        |
| 9.1.1 数据挖掘的主要步骤 .....                       | 287        |
| 9.1.2 不同存储形式的数据挖掘 .....                     | 288        |
| 9.2 数据仓库与 OLAP .....                        | 289        |
| 9.2.1 数据仓库 .....                            | 289        |
| 9.2.2 联机分析处理 (OLAP) .....                   | 291        |
| 9.3 数据仓库的设计与实施 .....                        | 293        |
| 9.3.1 确定主题 .....                            | 293        |
| 9.3.2 确定粒度与维度 .....                         | 293        |
| 9.3.3 数据预处理 .....                           | 294        |
| 9.4 SQL Server 2000 Analysis Services ..... | 295        |
| 9.4.1 安装 .....                              | 295        |
| 9.4.2 创建数据仓库 .....                          | 295        |
| 9.5 使用 Analysis Services 创建数据挖掘模型 .....     | 301        |
| 9.5.1 数据挖掘模型 .....                          | 301        |
| 9.5.2 创建和读取数据挖掘模型 .....                     | 303        |
| 9.6 开发 OLAP 应用 .....                        | 306        |
| 思考与练习 .....                                 | 308        |
| <b>第 10 章 数据库技术新进展 .....</b>                | <b>309</b> |
| 10.1 面向对象数据库系统 .....                        | 309        |
| 10.2 分布式数据库系统 .....                         | 311        |
| 10.3 并行数据库系统 .....                          | 313        |
| 10.4 空间数据库 .....                            | 315        |

---

|                                        |            |
|----------------------------------------|------------|
| 10.4.1 空间数据的特性 .....                   | 315        |
| 10.4.2 空间数据库模型 .....                   | 316        |
| 10.5 其他数据库新技术 .....                    | 316        |
| 10.5.1 多媒体数据库系统 .....                  | 316        |
| 10.5.2 专业数据库 .....                     | 317        |
| 10.5.3 模糊数据库系统 .....                   | 318        |
| 思考与练习 .....                            | 319        |
| <b>附录 A SQL Server 支持的运算符与函数</b> ..... | <b>320</b> |
| <b>附录 B 学籍管理数据表结构</b> .....            | <b>325</b> |
| <b>参考文献</b> .....                      | <b>327</b> |

# 第一篇 知 识 篇

## 第 1 章 数据库系统概述

数据库技术的发展一日千里，其应用已深入到工农业生产、商业、金融、行政管理、科学研究、工程技术、国防和军事等各个领域，成为信息管理、办公自动化、计算机辅助决策、计算机辅助设计、智能信息处理等计算机应用系统的核心部分。作为计算机软件学科中的一个十分活跃而重要的独立分支，数据库技术已形成了一整套的理论与技术体系。

本章将从数据管理技术的发展着手，比较详细地介绍数据库系统的基本知识。

### 1.1 引 言

#### 1.1.1 数据库管理技术的发展

数据管理是指对数据的收集、整理、组织、存储、维护、检索、传送等操作过程。数据处理是指对数据进行收集、存储、加工、传播等一系列活动的总和，其基本目的是从大量的、杂乱无章的、难以理解的数据中抽取并导出对于某些特定的应用来说是有价值的、有意义的数据，借以作为决策的依据。可见，数据管理是任何数据处理业务中必不可少的部分，数据管理和数据处理是密不可分的。

随着计算机科学与技术的发展，利用电子计算机进行数据管理大体上经历了从低级到高级的三个阶段，即人工管理阶段、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，从硬件看，外存储器只有磁带、卡片和纸带，没有磁盘；从软件看，没有操作系统，该阶段的计算机系统还没有支持管理数据的软件，主要应用于科学计算。在程序中不仅规定数据的逻辑结构，而且还要设计物理结构（包括存储结构和存取方法等）。当数据的物理组织和存储设备改变时应用程序必须重新编制，这种特性称为数据与程序不具有独立性，数据的组织是面向应用的，应用程序之间无法共享数据资源，存在

大量的重复数据，维护应用程序之间数据的一致性很困难。应用程序和数据之间的关系如图 1-1 所示。

## 2. 文件系统阶段

20 世纪 50 年代后期至 60 年代中期，计算机不仅用于科学计算，而且开始用于数据处理工作。从硬件来看，外存储器有了磁盘、磁鼓等直接存取的存储设备；软件方面有了操作系统，在操作系统中已经有了专门的管理数据的软件——文件系统。应用程序和数据之间的关系如图 1-2 所示。

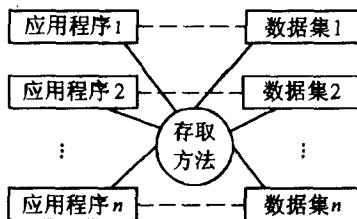
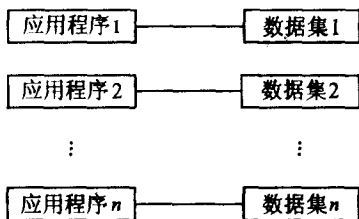


图 1-1 人工管理阶段应用程序与数据之间的对应关系    图 1-2 文件系统阶段应用程序与数据之间的对应关系

用文件系统管理数据具有如下优点。

(1) 数据能够长期保留在外存上，可以反复对其进行查询、修改、插入和删除等操作。

(2) 由软件对数据进行管理，程序和数据之间由文件系统提供存取方法进行转换，程序和数据有了一定的独立性。

然而，文件系统在数据管理方面也存在着如下缺陷。

(1) 数据共享性差，冗余度大

文件系统中的数据文件是为了满足某个部门或某一应用而进行设计的，当不同的应用程序具有部分相同的数据时，也必须建立各自的文件，导致同一数据项可能重复出现在多个文件中，因此数据的冗余度大，容易造成数据的不一致性。

(2) 程序和数据之间的独立性较差，数据联系弱

文件系统中的文件是为某一特定应用服务的，文件的逻辑结构对该应用程序来说是优化的，数据的逻辑结构一旦改变，必须修改应用程序，修改文件结构的定义，使得系统很难扩充。而应用程序的改变（例如应用程序改用不同的高级语言），也将引起文件数据结构的改变。所以程序和数据之间的独立性较差。可见，文件系统仍然是一个无结构的数据集合。文件是面向应用程序的这一特性，决定了文件之间是孤立的，文件不能反映现实世界之间的内在联系。

## 3. 数据库系统阶段

20 世纪 60 年代后期，计算机硬件和软件技术有了较大发展，出现了大容量直接存取设备。计算机在企业管理中的广泛应用，对数据管理提出了更高的要求。为了实现数据的统一管理，达到数据共享的目的，产生了数据库技术。在数据库系统中，应用程序和数据之间的关系如图 1-3 所示。

数据库系统的特点主要表现在如下 4 个方面。

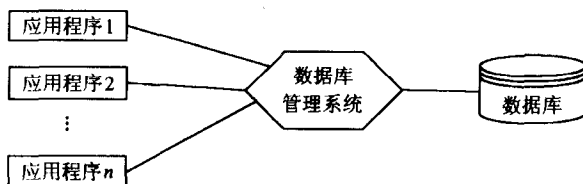


图 1-3 数据库系统阶段应用程序与数据的对应关系

### (1) 数据结构化

数据结构化是数据库系统与文件系统的根本区别。在文件系统中，文件中的记录内部是有结构的，但记录间没有联系，数据通常针对某个局部应用，数据的最小存储单位是记录，不能细到数据项。在数据库系统中，采用一定的数据模型，将整个组织的数据结构化成一个数据整体，数据不再面向应用（程序），而是面向系统，这种整体的结构化使得系统弹性大，有利于实现数据共享。另外，存储数据的方式更灵活，可以存取数据库中的一个数据项、一组数据项、一条记录或一组纪录。

### (2) 数据的共享性高、冗余低、易扩充

由于数据库是从整体角度看待和描述数据，数据不再面向某个应用，而是面向整个系统，数据库中的相同数据不会多次重复出现，数据冗余度降低，从而避免由于数据冗余度大而带来的数据冲突问题；另一方面，某一用户或某个应用程序通常仅使用总体数据的子集，这样，有效地发挥了数据共享的优势；同时，数据是有结构的，很容易增加新的应用，易于扩充。当应用需求改变或增加时，只需重新选择不同的子集，或加上一部分数据便可满足。

### (3) 数据独立性高

数据独立性是指数据的组织和存储方法与应用程序互不依赖、彼此独立的特性，它包括物理独立性和逻辑独立性。物理独立性是指用户的应用程序和存储在磁盘上的数据是相互独立的，用户程序不需要了解数据在磁盘上是如何存储的。逻辑独立性是指用户的应用程序和数据库的逻辑结构是相互独立的，数据的逻辑结构发生改变时，应用程序可以不变。

(4) 数据由数据库管理系统（DataBase Management System, DBMS）统一管理和控制。数据库管理系统必须提供以下 4 个方面的数据控制功能：

- 并发访问控制；
- 数据的安全性控制；
- 数据的完整性检查；
- 数据库恢复。

数据库管理系统在数据库的建立、运行和维护时对数据库进行统一的控制，保证数据的完整性、安全性；在多用户同时使用时进行并发控制；在数据库出现故障后对系统进行恢复。

综上所述，数据库就是长期存储在计算机内、有组织的、可共享的数据集合，通过数据库管理系统进行统一协调和管理。应当指出的是，虽然数据库管理系统是数据库系统的核心技术，但数据库系统的建立、使用和维护等工作只靠一个数据库管理系统是远远不够的，还必须相应的人员和工具。

### 1.1.2 数据库系统的组成

数据库系统是指带有数据库并利用数据库技术进行数据管理的计算机系统。一般由数据库、数据库管理系统（及其开发工具）、应用系统、数据库管理员（DataBase Administrator, DBA）和用户构成。在不引起混淆的情况下人们常常把数据库系统简称为数据库。

数据库系统可以用图 1-4 表示。

需要注意的是：数据库、数据库管理系统和数据库系统是 3 个不同的概念。数据库强调的是相互关联的数据；数据库管理系统强调的是管理数据库的系统软件；而数据库系统强调的是基于数据库技术的计算机系统，它包含数据库、数据库管理系统、操作系统、应用软件、硬件系统、数据库管理员、最终用户和开发人员等。

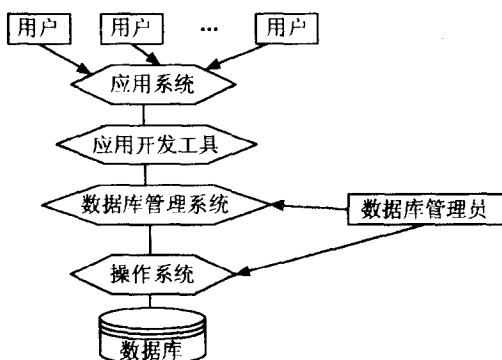


图 1-4 数据库系统

### 1.1.3 数据库技术的发展

数据库技术是计算机科学技术中发展最快的分支之一。20 世纪 70 年代以来，数据库系统从第一代的网状和层次数据库系统发展到第二代的关系数据库系统。目前现代数据库系统正向着面向对象数据库系统发展，并与网络技术、分布式计算、面向对象程序设计技术相结合。

第一代数据库系统为网状和层次数据库系统。1969 年 IBM 公司开发了基于层次模型的信息管理系统（Information Management System, IMS）。20 世纪 60 年代末至 70 年代初，美国数据库系统语言会议（Conference on Data Systems Languages, CODASYL）下属的数据库任务组（DataBase Task Group, DBTG）提出了若干报告（DBTG 报告）。该报告确定并建立了网状数据库系统的许多概念、方法和技术。正是基于上述报告，Cullinet Software 开发了基于网状模型的产品 IDMS（Information Data Management System），IMS 和 IDMS 这两个产品推动了网状和层次数据库系统的发展。

第二代数据库系统为关系数据库系统。1970 年 IBM 公司研究员 E.F.Codd 发表的关于关系模型的论文推动了关系数据库系统的研究和开发。尤其是关系数据库标准语言——结构化查询语言（SQL）的提出，使关系数据库系统得到了广泛的应用。20 世纪 80 年代以来，数据库软件厂商新推出的数据库管理系统几乎都支持关系模型，非关系系统的产品也都加上了关系接口。数据库领域当前的研究工作也都是以关系方法为基础。关系数据库已成为目前应用最广泛的数据库系统，如现在广泛使用的小型数据库系统 Foxpro、Access，大型数据库系统 Oracle、SQL Server、Informix、Sybase 等都是关系数据库系统。

随着数据库系统应用的范围进一步扩大，数据库处理对象的复杂性和灵活性对数据库系统提出了越来越高的要求。例如多媒体数据、CAD 数据、图形图像数据需要更好的数据模型来表达，以便存储、管理和维护。正是在这种形势下，20 世纪 80 年代产生了面向对象数据库系统。面向对象数据模型能完整地描述现实世界复杂的数据结构，并具有封装性和继承性等面向对象技术的优点。



## 1.2 数据库系统结构

虽然数据库系统软件产品种类很多，它们支持不同的数据模型，使用不同的数据库语言，建立在不同的操作系统之上，数据的存储结构也各不相同，但从数据库管理系统的角度看，它们都具有相同的特征，即采用三级模式结构（个别小型数据库系统除外），并提供两级映像功能；从数据库最终用户角度看，又分为单用户结构、主从式结构、分布式结构和客户/服务器结构。

### 1.2.1 数据库系统的模式结构

#### 1. 数据库系统的三级模式结构

数据库系统的三级模式结构是指数据库系统是由外模式、模式和内模式三级构成，如图 1-5 所示。

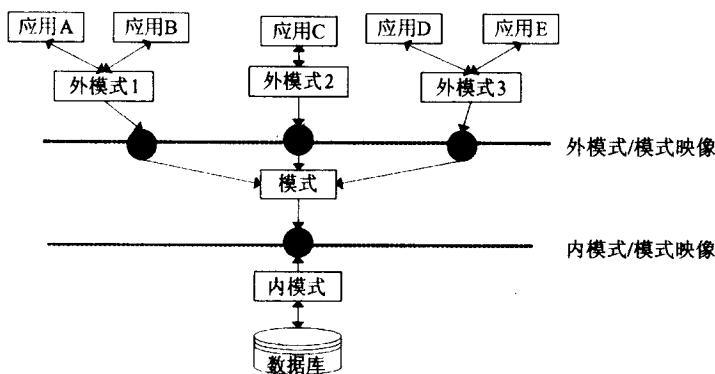


图 1-5 数据库系统的模式结构

#### (1) 模式

模式也称逻辑模式，是数据库中全体数据的逻辑结构和特征的描述，是所有用户的公共数据视图。模式是相对稳定的，模式的一个具体的值称为模式的一个实例（Instance）。同一个模式可以有很多实例。实例反映的是数据库某一时刻的状态。模式是数据库系统模式结构的中间层，不涉及数据的物理存储细节和硬件环境，与具体的应用程序，与所使用的应用开发工具及高级程序设计语言无关。

实际上模式是数据库中的数据在逻辑上的视图。一个数据库只有一个模式。数据库模式以某一种数据模型为基础，统一综合地考虑了所有用户的需求，并将这些需求有机地结合成一个逻辑整体。定义模式时不仅要定义数据的逻辑结构（例如，数据记录由哪些数据项构成，数据项的名字、类型、取值范围等），而且要定义与数据有关的安全性、完整性要求，定义这些数据之间的联系。

#### (2) 外模式

外模式也称子模式或用户模式，它是数据库用户（包括应用程序员和最终用户）看见和