

# 数据库基础 与实践技术

( SQL Server 2008 )

何玉洁 编著

DATABASE FUNDAMENTALS  
AND PRACTICAL TECHNOLOGY  
BASED ON SQL SERVER 2008



机械工业出版社  
China Machine Press

高等院校计算机专业人才培养规划教材(应用型)

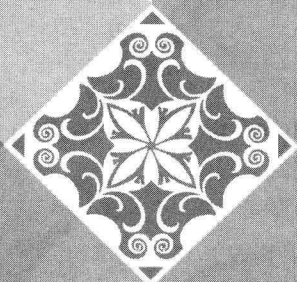
# 数据库基础 与实践技术

( SQL Server 2008 )

何玉洁 编著

D

DATABASE FUNDAMENTALS  
AND PRACTICE TECHNOLOGY  
BASED ON SQL SERVER 2008



机械工业出版社  
China Machine Press

## 图书在版编目 ( CIP ) 数据

数据库基础与实践技术: SQL Server 2008/ 何玉洁编著. —北京: 机械工业出版社, 2013.1  
(高等院校计算机专业人才培养规划教材·应用型)

ISBN 978-7-111-40698-3

I. 数… II. 何… III. 关系数据库系统—高等学校—教材 IV. TP311.138

中国版本图书馆 CIP 数据核字 (2013) 第 025938 号

### 版权所有·侵权必究

封底无防伪标均为盗版

本书法律顾问 北京市展达律师事务所

数据库技术是一门实践性很强的应用性技术, 本书采用目前使用广泛的 Microsoft SQL Server 2008 作为实践平台, 主要介绍数据库技术的应用。

本书包括 14 章和 1 个附录。第 1 章和第 2 章介绍数据库的基础理论, 主要针对数据库初学者。第 3 章~第 14 章为数据库技术应用, 主要包括数据库创建与维护管理、SQL、架构与基本表、数据操作语言、高级查询、索引、视图、存储过程、触发器、函数、游标、安全管理、数据库备份与恢复以及数据传输等内容。附录部分介绍 SQL Server 2008 提供的一些常用系统函数。

本书可作为高等院校计算机及相关专业数据库课程教材, 也可作为数据库开发人员的参考用书。

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)

责任编辑: 马 超

北京市荣盛彩色印刷有限公司印刷

2013 年 3 月第 1 版第 1 次印刷

185mm × 260 mm • 22.25 印张

标准书号: ISBN 978-7-111-40698-3

定 价: 39.80 元

凡购本书, 如有缺页、倒页、脱页, 由本社发行部调换

客服热线: (010) 88378991 88361066

投稿热线: (010) 88379604

购书热线: (010) 68326294 88379649 68995259

读者信箱: hzsj@hzbook.com

# 前言



数据库技术起源于 20 世纪 60 年代末，经过几十年的迅速发展，目前已经建立起一套较完整的理论体系，产生了一大批商用软件产品。随着数据库技术应用领域的不断扩大，对数据库的应用已深入到国民经济和社会生活的各个方面。因此，作为当代的大学生，学习和掌握数据库知识都是非常必要的。

为了提高计算机专业本科生对数据库知识的实践能力，很多高校的计算机专业对数据库的授课做了一些调整，相当一部分学校将数据库的教学安排在两个学期完成，或者分为两个阶段来完成。第一个阶段基本是以讲授数据库基本原理为主，第二个阶段以讲授数据库实践技术为主，作者所在的学校也是如此。但在教材选择上，本人发现适合数据库基础理论教学的教材很多，但适合数据库实践教学教材却比较少，鉴于此，本人结合自己多年数据库基础理论及数据库技术实践的教学经验和科研经验编写了本书，希望能对学生学习数据库实际的管理以及后台编程技术有所帮助。

本书是为学生学习数据库知识而编写的一本教材，其中选用 Microsoft SQL Server 2008 作为实践平台。SQL Server 具有良好的用户操作界面，功能全面而强大，有很高的市场占有率和很好的发展前景，非常适合作为学生的数据库实践平台。

本书主要包括两大部分内容。一部分是数据库管理方面的知识，这部分知识涵盖在第 2、3、5、8、12、13、14 章中，主要包括安装和配置 SQL Server 2008 数据库管理系统、创建与维护数据库、创建与维护关系表、创建分区表、构建索引的技术、安全管理、数据传输、备份和恢复数据库。另一部分是数据库编程方面的知识，这部分知识涵盖在第 4、6、7、9、10、11 章中，主要包括 SQL 基础、基本数据操作语句、高级查询、视图、存储过程、触发器、函数以及游标等。触发器主要是为了增强数据的完整性和一致性；存储过程主要是为了提高数据的操作效率，方便客户端的编程；函数主要是为了能实现一些复杂的数据操作以及模块共享功能；使用游标可以实现对数据的逐行处理。为方便数据库初学者学习和掌握数据库实践技能，本书特意在第 1 章介绍了数据库中最常用的基础知识，初学者在掌握了第 1 章的知识后，便可学习后续章节内容。

本书内容涵盖了常用的数据库管理和编程技术，内容由浅入深，除包含一般数据库实践知识外，还全面介绍了数据库技术发展的新功能、新思想，这些新功能在 SQL Server 2008 中

得到了全面的支持，包括查询结果的并、交、差运算，以及开窗函数、公用表表达式、筛选索引、包含列索引、分区表、索引视图（物化视图）等。

本书实例丰富、图文并茂，并紧密结合实际问题，从问题出发，循序渐进地给出解决思路和方法，使读者能更准确地理解知识并应用。

为方便读者学习和实践所学知识，也为了更好地讲解某些知识，本书为读者提供了示例数据库 MySimpleDB，其数据是从 SQL Server 2008 提供的 AdventureWorks 数据库中选取的，有需要的读者可从华章网站（[www.hzbook.com](http://www.hzbook.com)）上下载该数据库。为方便教师授课，本书为教师提供授课幻灯片、书中的编程代码以及习题答案，这些资料也可从华章网站上下载。

本书的出版得到了机械工业出版社华章公司领导和姚蕾的大力支持与帮助，他们给予我很多的鼓励和很好的建议，帮助我完成了本书的编写。梁琦、田英爱、谷葆春、张鸿斌等对本书部分章节及上机实践题目的编写提供了帮助，同时对本书提出了很多很好的建议和意见，是他们的积极参与和帮助，使本书得以顺利完成。

真诚地希望读者对本书提出自己宝贵的意见和建议，因为我知道在教学探索的道路上没有止境，没有最好，只有更好。

何玉洁

2013 年 1 月



| 教学内容                        | 学习要点及教学要求                                                                                                                                                                  | 参考学时                          |
|-----------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------|
| 第 1 章<br>关系数据库基础            | 1) 了解数据管理的发展历史, 重点体会数据库技术对管理数据带来的本质上的变化<br>2) 重点理解关系数据模型的特点<br>3) 了解概念层数据模型、关系数据模型以及概念层模型向关系模型的转换方法<br>4) 理解数据库三级模式的概念, 理解三级模式的作用<br>5) 了解关系规范化的基本概念, 知道 1NF、2NF 和 3NF 的定义 | 4 ~ 6 学时 (对于有数据库基础的读者, 可跳过此章) |
| 第 2 章<br>SQL Server 2008 基础 | 1) 了解安装 SQL Server 2008 的软硬件要求<br>2) 了解 SQL Server 2008 的安装及配置<br>3) 掌握 SQL Server 配置管理器及 SQL Server Management Studio 工具的使用                                               | 2 学时                          |
| 第 3 章<br>数据库的创建与管理          | 1) 掌握用图形化方法和 T-SQL 语句创建数据库的方法<br>2) 掌握数据库空间的维护方法<br>3) 了解分离和附加数据库以及复制数据库的作用和实现方法                                                                                           | 3 ~ 4 学时                      |
| 第 4 章<br>SQL 基础             | 1) 了解常用的数据类型<br>2) 了解用户定义数据类型的方法<br>3) 掌握变量的定义和赋值方法<br>4) 了解语句批等一些 T-SQL 基础知识                                                                                              | 2 学时                          |
| 第 5 章<br>架构与基本表             | 1) 掌握架构的概念及创建方法<br>2) 掌握基本表的创建方法<br>3) 掌握分区表的概念及创建方法                                                                                                                       | 2 学时                          |
| 第 6 章<br>数据操作语言             | 1) 掌握单表查询、多表连接查询<br>2) 了解自连接、左外连接和右外连接含义<br>3) 掌握分组统计查询<br>4) 了解全外连接含义<br>5) 掌握数据的插入、删除、修改语句的用法                                                                            | 6 学时 (对于有数据库基础的读者, 可跳过本章部分内容) |
| 第 7 章<br>高级查询               | 1) 了解 case 函数的作用并掌握其使用方法<br>2) 掌握嵌套子查询和相关子查询语句<br>3) 掌握查询结果的并、交、差运算<br>4) 了解替代表达式的子查询以及派生表子查询的使用<br>5) 了解开窗函数、公用表表达式语句的使用                                                  | 4 ~ 6 学时                      |
| 第 8 章<br>索引                 | 1) 理解索引的作用<br>2) 了解 B- 树索引的存储结构<br>3) 理解聚集索引、非聚集索引的特点<br>4) 了解包含列索引、筛选索引的含义和作用<br>5) 掌握建立聚集索引、非聚集索引、包含列索引、筛选索引的方法                                                          | 2 学时                          |

(续)

| 教学内容               | 学习要点及教学要求                                                                                                                                                               | 参考学时     |
|--------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------|
| 第 9 章<br>视图        | 1) 理解一般视图的作用并掌握其建立方法<br>2) 了解索引视图的作用<br>3) 理解分区视图的作用并掌握其建立方法                                                                                                            | 2 学时     |
| 第 10 章<br>存储过程和触发器 | 1) 理解存储过程的作用, 并掌握其定义方法及调用方法<br>2) 理解触发器的作用并掌握其定义方法                                                                                                                      | 4 学时     |
| 第 11 章<br>函数和游标    | 1) 理解用户定义的标量函数、内联表值函数、多语句表值函数的定义并掌握其使用方法<br>2) 理解游标的作用及定义过程                                                                                                             | 4 学时     |
| 第 12 章<br>安全管理     | 1) 理解 SQL Server 三个权限认证过程<br>2) 理解登录账户的概念并掌握其建立方法<br>3) 理解数据库用户的含义并掌握其建立方法<br>4) 掌握为用户授权的方法<br>5) 理解角色的概念以及系统提供的主要角色的作用, 掌握构建用户自定义角色及为角色授权的方法, 掌握为角色添加和删除成员的方法          | 2 ~ 3 学时 |
| 第 13 章<br>备份和恢复数据库 | 1) 理解数据库备份的含义和作用<br>2) 了解备份设备的作用和创建方法<br>3) 理解完整备份、差异备份和日志备份的备份内容<br>4) 了解各种备份对数据库恢复模式的要求<br>5) 了解常用的备份策略<br>6) 掌握完整备份、差异备份和日志备份的备份方法<br>7) 理解数据库的恢复顺序<br>8) 掌握数据库的恢复方法 | 2 ~ 3 学时 |
| 第 14 章<br>数据传输     | 1) 了解 ODBC 接口和 OLE DB 接口<br>2) 掌握 Excel、Text、Access 及 SQL Server 之间数据的导入 / 导出方法                                                                                         | 1 ~ 2 学时 |

# 目 录



前言

教学建议

|                      |    |
|----------------------|----|
| 第 1 章 关系数据库基础        | 1  |
| 1.1 数据管理的发展          | 1  |
| 1.1.1 文件管理系统         | 1  |
| 1.1.2 数据库管理系统        | 3  |
| 1.2 数据库系统与数据库管理系统    | 5  |
| 1.2.1 数据库系统的组成       | 5  |
| 1.2.2 数据库管理系统的功能     | 6  |
| 1.3 数据和数据模型          | 7  |
| 1.3.1 数据和数据模型概述      | 7  |
| 1.3.2 数据模型三要素        | 8  |
| 1.3.3 概念层数据模型        | 9  |
| 1.3.4 组织层数据模型        | 11 |
| 1.3.5 E-R 模型向关系模型的转换 | 16 |
| 1.4 数据库系统的结构         | 17 |
| 1.4.1 三级模式结构         | 18 |
| 1.4.2 两级模式映像功能与数据独立性 | 20 |
| 1.5 关系数据库规范化理论       | 20 |
| 1.5.1 函数依赖           | 21 |
| 1.5.2 关系规范化          | 23 |
| 1.6 数据库应用结构          | 25 |
| 1.6.1 集中式结构          | 25 |
| 1.6.2 文件服务器结构        | 26 |
| 1.6.3 客户端/服务器结构      | 26 |
| 1.6.4 互联网应用结构        | 27 |
| 小结                   | 28 |
| 习题                   | 28 |

|                                     |    |
|-------------------------------------|----|
| 第 2 章 SQL Server 2008 基础            | 30 |
| 2.1 SQL Server 2008 的版本和组件          | 30 |
| 2.1.1 SQL Server 2008 的版本           | 30 |
| 2.1.2 SQL Server 2008 的组件           | 31 |
| 2.1.3 SQL Server 2008 各版本支持的功能      | 32 |
| 2.1.4 安装 SQL Server 2008 需要的软件、硬件环境 | 33 |
| 2.1.5 实例                            | 34 |
| 2.2 安装 SQL Server 2008              | 35 |
| 2.3 配置 SQL Server 2008              | 41 |
| 2.4 卸载 SQL Server 2008              | 43 |
| 2.5 SQL Server Management Studio 工具 | 45 |
| 2.5.1 连接到数据库服务器                     | 45 |
| 2.5.2 查询编辑器                         | 46 |
| 2.5.3 SSMS 的常用选项                    | 47 |
| 小结                                  | 56 |
| 习题                                  | 56 |
| 上机练习                                | 56 |
| 第 3 章 数据库的创建与管理                     | 57 |
| 3.1 SQL Server 数据库概述                | 57 |
| 3.1.1 系统数据库                         | 57 |
| 3.1.2 数据库的组成                        | 58 |
| 3.1.3 数据库文件组                        | 59 |
| 3.1.4 数据库文件的属性                      | 60 |
| 3.2 创建数据库                           | 60 |
| 3.2.1 用图形化方法创建数据库                   | 60 |
| 3.2.2 用 T-SQL 语句创建数据库               | 65 |
| 3.3 查看和设置数据库选项                      | 68 |

|                        |     |                            |     |
|------------------------|-----|----------------------------|-----|
| 3.4 修改数据库 .....        | 71  | 第 5 章 架构与基本表 .....         | 107 |
| 3.4.1 扩大数据库空间 .....    | 71  | 5.1 架构 .....               | 107 |
| 3.4.2 收缩数据库空间 .....    | 73  | 5.1.1 创建架构 .....           | 107 |
| 3.4.3 添加和删除数据库文件 ..... | 77  | 5.1.2 在架构间传输对象 .....       | 109 |
| 3.4.4 创建和更改文件组 .....   | 79  | 5.1.3 删除架构 .....           | 109 |
| 3.5 删除数据库 .....        | 81  | 5.2 基本表 .....              | 109 |
| 3.6 分离和附加数据库 .....     | 82  | 5.2.1 用 T-SQL 语句实现 .....   | 109 |
| 3.6.1 分离数据库 .....      | 82  | 5.2.2 用 SSMS 工具实现 .....    | 116 |
| 3.6.2 附加数据库 .....      | 84  | 5.3 分区表 .....              | 121 |
| 3.7 复制数据库 .....        | 87  | 5.3.1 基本概念 .....           | 121 |
| 小结 .....               | 91  | 5.3.2 创建分区表 .....          | 122 |
| 习题 .....               | 91  | 5.3.3 创建分区函数 .....         | 122 |
| 上机练习 .....             | 92  | 5.3.4 创建分区方案 .....         | 123 |
| 第 4 章 SQL 基础 .....     | 93  | 小结 .....                   | 125 |
| 4.1 SQL 概述 .....       | 93  | 习题 .....                   | 125 |
| 4.1.1 SQL 的发展 .....    | 93  | 上机练习 .....                 | 126 |
| 4.1.2 SQL 特点 .....     | 93  | 第 6 章 数据操作语言 .....         | 128 |
| 4.1.3 SQL 功能概述 .....   | 94  | 6.1 数据查询语句 .....           | 128 |
| 4.2 系统提供的数据类型 .....    | 94  | 6.1.1 查询语句的基本结构 .....      | 129 |
| 4.2.1 数值数据类型 .....     | 94  | 6.1.2 单表查询 .....           | 130 |
| 4.2.2 字符串类型 .....      | 95  | 6.1.3 多表连接查询 .....         | 140 |
| 4.2.3 日期和时间数据类型 .....  | 97  | 6.1.4 使用 TOP 限制结果集行数 ..... | 146 |
| 4.2.4 货币类型 .....       | 98  | 6.1.5 将查询结果保存到新表中 .....    | 147 |
| 4.3 用户定义数据类型 .....     | 98  | 6.2 数据更改功能 .....           | 148 |
| 4.3.1 创建用户定义数据类型 ..... | 98  | 6.2.1 插入数据 .....           | 148 |
| 4.3.2 删除用户定义数据类型 ..... | 100 | 6.2.2 更新数据 .....           | 151 |
| 4.4 T-SQL 基础知识 .....   | 100 | 6.2.3 删除数据 .....           | 152 |
| 4.4.1 语句批 .....        | 100 | 小结 .....                   | 153 |
| 4.4.2 脚本 .....         | 101 | 习题 .....                   | 154 |
| 4.4.3 注释 .....         | 101 | 上机练习 .....                 | 154 |
| 4.5 变量 .....           | 102 | 第 7 章 高级查询 .....           | 156 |
| 4.5.1 变量的种类 .....      | 102 | 7.1 CASE 函数 .....          | 156 |
| 4.5.2 变量的声明与赋值 .....   | 102 | 7.1.1 CASE 函数介绍 .....      | 156 |
| 4.6 流程控制语句 .....       | 104 | 7.1.2 CASE 函数应用示例 .....    | 158 |
| 小结 .....               | 106 | 7.2 子查询 .....              | 159 |
| 习题 .....               | 106 | 7.2.1 嵌套子查询 .....          | 160 |
| 上机练习 .....             | 106 |                            |     |

|                         |     |                              |     |
|-------------------------|-----|------------------------------|-----|
| 7.2.2 相关子查询 .....       | 166 | 小结 .....                     | 231 |
| 7.2.3 其他形式的子查询 .....    | 171 | 习题 .....                     | 231 |
| 7.3 查询结果的并、交、差运算 .....  | 172 | 上机练习 .....                   | 231 |
| 7.3.1 并运算 .....         | 172 |                              |     |
| 7.3.2 交运算 .....         | 173 |                              |     |
| 7.3.3 差运算 .....         | 175 |                              |     |
| 7.4 其他查询功能 .....        | 176 |                              |     |
| 7.4.1 开窗函数 .....        | 176 |                              |     |
| 7.4.2 公用表表达式 .....      | 180 |                              |     |
| 7.4.3 MERGE 语句 .....    | 184 |                              |     |
| 小结 .....                | 190 |                              |     |
| 习题 .....                | 191 |                              |     |
| 上机练习 .....              | 191 |                              |     |
| <b>第 8 章 索引</b> .....   | 195 | <b>第 10 章 存储过程和触发器</b> ..... | 234 |
| 8.1 基本概念 .....          | 195 | 10.1 存储过程 .....              | 234 |
| 8.2 索引存储结构及类型 .....     | 196 | 10.1.1 存储过程概念 .....          | 234 |
| 8.2.1 B- 树结构 .....      | 196 | 10.1.2 创建和执行存储过程 .....       | 234 |
| 8.2.2 索引类型 .....        | 197 | 10.1.3 从存储过程返回数据 .....       | 238 |
| 8.3 创建索引 .....          | 204 | 10.1.4 查看和维护存储过程 .....       | 240 |
| 8.3.1 用 SQL 实现 .....    | 204 | 10.1.5 一些系统存储过程 .....        | 241 |
| 8.3.2 用 SSMS 工具实现 ..... | 213 | 10.2 触发器 .....               | 244 |
| 8.3.3 设计索引注意事项 .....    | 215 | 10.2.1 创建触发器 .....           | 245 |
| 小结 .....                | 215 | 10.2.2 DML 触发器示例 .....       | 247 |
| 习题 .....                | 216 | 10.2.3 DDL 触发器示例 .....       | 252 |
| 上机练习 .....              | 216 | 10.2.4 查看和维护触发器 .....        | 252 |
|                         |     | 小结 .....                     | 254 |
| <b>第 9 章 视图</b> .....   | 217 | 习题 .....                     | 254 |
| 9.1 标准视图 .....          | 217 | 上机练习 .....                   | 255 |
| 9.1.1 基本概念 .....        | 217 |                              |     |
| 9.1.2 定义视图 .....        | 217 | <b>第 11 章 函数和游标</b> .....    | 256 |
| 9.1.3 通过视图查询数据 .....    | 222 | 11.1 用户自定义函数 .....           | 256 |
| 9.1.4 修改和删除视图 .....     | 223 | 11.1.1 基本概念 .....            | 256 |
| 9.1.5 视图的作用 .....       | 224 | 11.1.2 创建和调用标量函数 .....       | 256 |
| 9.2 分区视图 .....          | 225 | 11.1.3 创建和调用内联表值函数 .....     | 258 |
| 9.3 索引视图 .....          | 227 | 11.1.4 创建和调用多语句表值函数 .....    | 259 |
| 9.3.1 索引视图基本概念 .....    | 227 | 11.1.5 查看和修改用户自定义函数 .....    | 260 |
| 9.3.2 定义索引视图 .....      | 228 | 11.1.6 删除用户自定义函数 .....       | 261 |
|                         |     | 11.2 游标 .....                | 262 |
|                         |     | 11.2.1 游标概念 .....            | 262 |
|                         |     | 11.2.2 使用游标 .....            | 262 |
|                         |     | 11.2.3 游标示例 .....            | 264 |
|                         |     | 小结 .....                     | 268 |
|                         |     | 习题 .....                     | 268 |
|                         |     | 上机练习 .....                   | 269 |
|                         |     | <b>第 12 章 安全管理</b> .....     | 271 |
|                         |     | 12.1 安全控制概述 .....            | 271 |
|                         |     | 12.2 SQL Server 的安全控制 .....  | 272 |

|                              |     |                      |     |
|------------------------------|-----|----------------------|-----|
| 12.3 管理登录账户.....             | 274 | 13.2.2 恢复模式.....     | 305 |
| 12.3.1 建立登录账户.....           | 274 | 13.2.3 备份类型及策略.....  | 307 |
| 12.3.2 删除登录账户.....           | 278 | 13.2.4 实现备份.....     | 309 |
| 12.4 管理数据库用户.....            | 279 | 13.3 恢复数据库.....      | 314 |
| 12.4.1 建立数据库用户.....          | 279 | 13.3.1 恢复数据库的顺序..... | 314 |
| 12.4.2 删除数据库用户.....          | 281 | 13.3.2 实现还原.....     | 315 |
| 12.5 管理权限.....               | 282 | 小结.....              | 318 |
| 12.5.1 权限的种类.....            | 282 | 习题.....              | 319 |
| 12.5.2 权限的管理.....            | 282 | 上机练习.....            | 319 |
| 12.6 角色.....                 | 289 | 第 14 章 数据传输.....     | 320 |
| 12.6.1 固定的服务器角色.....         | 289 | 14.1 数据访问接口.....     | 320 |
| 12.6.2 固定的数据库角色.....         | 293 | 14.1.1 ODBC.....     | 320 |
| 12.6.3 用户定义的角色.....          | 297 | 14.1.2 OLE DB.....   | 322 |
| 小结.....                      | 300 | 14.2 数据导入和导出.....    | 323 |
| 习题.....                      | 301 | 14.2.1 导出数据.....     | 323 |
| 上机练习.....                    | 301 | 14.2.2 导入数据.....     | 329 |
| 第 13 章 备份和恢复数据库.....         | 303 | 小结.....              | 334 |
| 13.1 备份数据库.....              | 303 | 习题.....              | 334 |
| 13.1.1 为什么要进行数据备份.....       | 303 | 上机练习.....            | 334 |
| 13.1.2 备份内容及备份时间.....        | 303 | 附录 A 系统提供的常用函数.....  | 335 |
| 13.2 SQL Server 支持的备份机制..... | 304 | 参考文献.....            | 346 |
| 13.2.1 备份设备.....             | 304 |                      |     |

# 第1章 关系数据库基础

随着信息管理水平的不断提高，计算机应用范围的日益广泛，数据库技术正逐步渗透到我们日常生活中的方方面面。例如，银行对账户、信贷业务的管理，超市对货物销售、进货情况的管理，飞机、火车订票系统，图书馆对书籍及书籍借阅的管理等，无一不使用了数据库技术。人们在日常生活中有时可能不知不觉地用到了数据库。

有了准确的数据，还需要对数据进行科学的管理，使之能服务于人。数据库就是研究如何对数据进行科学的管理以便为人们提供可共享的、安全的、可靠的数据的技术。

本章主要介绍使用数据库技术管理数据的优势，同时介绍目前主流的关系数据库管理系统的基本概念。

## 1.1 数据管理的发展

从计算机产生之后，人们就希望用计算机存储和管理数据。最初用计算机对数据进行管理是以文件方式进行的，即将数据保存在用户定义好的文件中，然后编写对数据文件进行操作的应用程序。这种数据管理方式对用户的要求比较高，需要有较强的计算机编程技能，同时还要具有数据存储和访问知识，以最大限度地提高数据访问效率。随着数据量的不断增大，计算机处理能力的不断增强，人们对数据的要求越来越多，希望达到的目的也越来越复杂，因此用文件对数据进行管理，通过编程的方法来操作文件中数据的方式已经很难满足人们对数据的需求，由此产生了数据库管理系统，也就是用数据库管理系统来对数据进行维护。综上所述，使用计算机对数据进行管理经历了文件管理和数据库管理两个发展阶段。

本节介绍文件管理和数据库管理在数据管理上的主要差别。

### 1.1.1 文件管理系统

理解当今数据库特征的最好办法是了解一下在使用数据库技术之前数据管理的特点。

早期的数据管理是采用文件方式进行的，即数据保存在文件中，文件是由操作系统和特定的软件或程序共同管理的。在文件管理方式中，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。用户可以通过操作系统和特定的识别这些文件格式的软件对文件进行打开、读、写等操作，也可以通过程序设计语言编写对文件进行操作的程序。

假设现在要用某种程序设计语言编写对学生信息进行管理的系统（注：程序负责对文件的打开、关闭以及对其中数据的读、写操作，具体对磁盘文件操作的实现是由操作系统中的文件管理等部分完成的）。在此系统中，要对学生的基本信息和选课情况进行管理。在学生基本信息管理中要用到学生的基本信息数据，假设此数据保存在 F1 文件中。学生选课情况的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。假设用 F2 和 F3 两个文件分别存储课程的基本信息和学生的选课信息数据，此部分的学生基本信息数据可以使用 F1 文件中的数据。设 A1 为实现“学生基本信息管理”功能的应用程序，A2 为实现“学生选课管理”功能的应用程序。文件管理系统示例如图 1-1 所示。

假设 F1、F2 和 F3 文件分别包含如下信息。

F1：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：课程号、课程名、授课学期、学分、课程性质。

F3: 学号、姓名、所在系、专业、课程号、课程名、修课类型、修课时间、考试成绩。

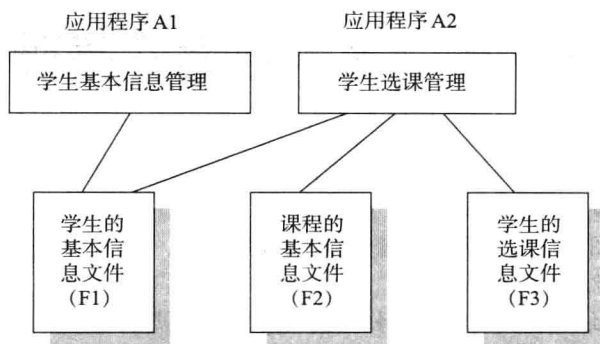


图 1-1 文件管理系统示例

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条记录。

“学生选课管理”的处理过程大致如下：

在学生选课管理中，若有学生选课，需先查 F1 文件，判断有无此学生；若有则访问 F2 文件，判断其所选的课程是否存在；若一切合乎规则，则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，就会发现直接用文件管理数据有如下一些缺点。

### 1. 编写应用程序不方便

应用程序编写者必须对所使用的文件的逻辑及物理结构（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，而对文件的查询、修改等处理都必须在应用程序中编程实现。这样也易造成各应用程序在功能上的重复，例如图 1-1 中的“学生基本信息管理”和“学生选课管理”都要对 F1 文件进行操作。

### 2. 数据冗余不可避免

假设应用程序 A2 需要在 F3 文件中包含学生的所有或大部分信息。例如，除了学号之外，还需要姓名、专业、所在系等信息，而 F1 文件中也包含这些信息，因此 F3 文件和 F1 文件中有重复的信息，由此会造成数据的重复（又称为数据冗余）。

数据冗余所带来的问题不仅仅是存储空间的浪费，更为严重的是会造成数据不一致（inconsistency）。例如，假设学生所在的专业发生了变化，通常情况下，我们可能只记得在 F1 文件中进行修改，而忘记了在 F3 文件中也要进行同样的修改，由此会造成同一名学生在 F1 文件和 F3 文件中的“专业”不一样，也就是数据不一致。人们不能判定哪个文件中的数据是正确的，因此也就失去了数据的可信度。

文件系统中没有维护数据一致性的功能，这完全由用户（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，若要保证数据的一致性，几乎是不可能的。

### 3. 应用程序依赖性

就文件处理而言，程序依赖于文件的格式。文件和记录的逻辑结构通常是应用程序代码的一部分，如 C 语言用 Struct、Visual Basic 用 Type 来定义用户的数据结构。文件结构的每一次修改都将导致应用程序的修改。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。例如，增加一些字段、修改某些字段的长度（如电话号码从 7 位扩展到 8 位）。这些都需要在应用程序中做相应的修改，而（频繁）修改应用程序是很麻烦的，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数据文件的过度依赖造成的，换句话说，文件系统的数据独立性

(data independence) 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, 一般允许多个应用程序并发执行 (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。我们可以想一下, 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对此文件关闭之前, 也想打开此文件, 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的同时对此文件进行修改。上述就是文件系统不支持并发访问的原因。

对于以数据为中心的应用系统来说, 支持多个用户对数据的并发访问是必不可少的功能。

#### 5. 数据间联系弱

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。例如上述的 F1 文件和 F3 文件, 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 (即选课的学生信息必须是已经存在的); 同样, F3 文件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)。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, 必须依靠应用程序来保证这些联系, 也就是必须通过编写程序来手工地保证这些联系。这不但增加了程序编写的工作量和复杂度, 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. 难以按不同用户的愿望表示数据

如果用户需要的信息来自于多个不同文件中部分信息的组合, 就需要对多个文件进行提取、比较、组合和表示。例如, 假设有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 3 个文件: 从 F1 文件中得到“所在班”信息, 从 F2 文件中得到“学分”, 从 F3 文件中得到“考试成绩”; 而“学号”、“姓名”可以从 F1 文件或 F3 文件中得到, “课程名”可以从 F2 文件或 F3 文件中得到。在生成一行数据 (所在班, 学号, 姓名, 课程名, 学分, 考试成绩) 时, 必须对从 3 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。例如, 将从 F1 文件中读取的学号与从 F3 文件中读取的学号进行比较, 学号相同时, 才可以将 F1 文件中的“所在班”、F3 文件中的“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 文件和 F3 文件的组合后, 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, 且涉及的表比较多, 我们可以想象这个过程有多么复杂。因此, 这种大容量且复杂信息的查询, 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权, 如只能读和修改数据而不能删除数据, 或者对文件中的某个或某些字段不能读或修改等。而在实际生活中, 数据的安全性是非常重要的。例如, 在学生选课管理中, 学生对其考试成绩一般只有查看权, 而教师则有录入其所授课程的考试成绩的权限、教务部门对录入有误的成绩有修改权等。

随着人们对数据需求的增加以及计算机科学的不断发展, 如何对数据进行有效、科学、正确、方便的管理已成为人们的迫切愿望。针对文件系统的上述缺陷, 人们逐步发展了以统一管理和共享数据为主要特征的数据库管理系统。

### 1.1.2 数据库管理系统

数据库技术的发展主要源于文件管理系统在管理数据上的诸多缺陷。对于上述学生基本信息管

理和学生选课管理，如果使用数据库技术来实现，其实现方式与文件系统有本质的区别，如图 1-2 所示。

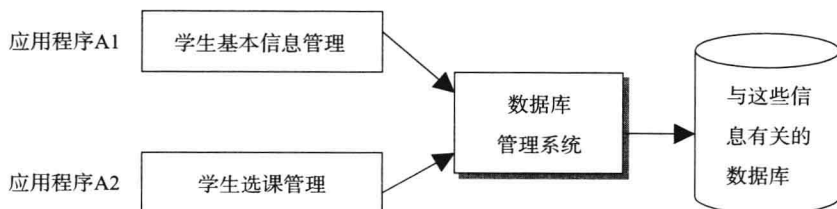


图 1-2 用数据库技术实现示例

比较图 1-1 和图 1-2，可以直观地发现两者有如下差别：

- 使用文件系统时，应用程序直接访问存储数据的文件，而使用数据库系统时则通过数据库管理系统（DataBase Management System, DBMS）访问数据，而且存储数据的文件以数据库的形式展示给客户，这个变化使得应用程序开发人员不再需要关心数据的物理存储方式和存储结构，这些都交给了数据库管理系统来完成，从而极大地简化了应用编程工作。
- 在数据库系统中，数据不再仅仅服务于某个程序或用户，而是看成一定业务范围的共享资源，由一个称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在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）之间增加了一层——数据库管理系统。数据库管理系统实际上是一个系统软件。正是因为有了这个系统软件，才使得以前在应用程序中由开发人员编程实现的很多烦琐的操作和功能交给了数据库管理系统，这样应用程序不再需要关心数据的存储方式。而且数据存储方式的变化也不再影响应用程序，这些变化交给数据库管理系统来处理，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。

与直接用文件管理数据的局限性进行比较，我们就能体会到使用数据库技术管理数据所带来的好处。

### 1. 将相互关联的数据集成在一起

在数据库系统中，所有的数据都存储在数据库中，应用程序可通过 DBMS 访问数据库中的所有数据。

### 2. 数据冗余小

由于数据被统一管理，因此可以从全局着眼，合理地组织数据。例如，将 1.1.1 节中的 F1、F2 和 F3 文件中的重复数据，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（关系数据库的概念在本章后续部分介绍）中，可以将每一种信息存储在一个表中，重复的信息只存储一份，当在学生选课中需要学生的名字时，根据学生选课信息中的学号，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并不影响对信息的提取，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。例如，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”一个地方进行修改即可。

同 1.1.1 节中的问题一样，当检索（所在班，学号，姓名，课程名，学分，考试成绩）信息时，这些信息需要从 3 个地方（关系数据库为 3 张表）得到，也需要对信息进行适当的组合，即学生选课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，学生选课中的课程号也必须与课程基本信息中课程号相同的信息组合在一起。在以前的文件管理系统中，这个工作是由开发者

编程实现的，而现在有了数据库管理系统，这些烦琐的工作完全交给了数据库管理系统来完成。

### 3. 程序与数据相互独立

在用数据库技术管理数据的方式中，所有的数据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，应用程序不再需要处理数据文件的存储结构。

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，或者是顺序和非顺序之间的转换，应用程序不必做任何修改；另一方面是指当数据的逻辑结构发生变化时，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，应用程序并不需要知道数据存储方式或数据项已经发生了变化。

### 4. 保证数据的安全和可靠

数据库技术能够保证数据库中的数据是安全的、可靠的。数据库中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；同时它还有一套完整的备份和恢复机制，以保证当数据遭到破坏（由软件或硬件故障引起的）时，能够很快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现实世界的实际情况，比如人的性别只能是“男”或“女”，人的年龄应该在 0 ~ 150 岁之间（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，在现实世界中显然是不对的。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。保证数据正确性的特征也称为数据完整性。数据完整性是通过在数据库中建立约束来实现的。当用户建立好保证数据正确的约束之后，如果有不符合约束条件的数据进入到数据库中，数据库管理系统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，这个特点针对大型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。

多用户问题是数据库管理系统内部解决的问题，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。可以想象一下火车订票系统，如果多个订票点同时对一列火车进行订票操作，那么必须要保证不同订票点订出票的座位不能相同。

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，可以发现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固有的，而是靠数据库管理系统提供和保证的。

## 1.2 数据库系统与数据库管理系统

本节介绍数据库系统的组成以及数据库管理系统的功能。

### 1.2.1 数据库系统的组成

数据库管理系统是一个系统软件，如 SQL Server、Oracle、DB2 等都是著名的数据库管理系

统软件,但在有了数据库管理系统这个软件之后,并不意味着已经具有了用数据库管理系统管理数据的优点,必须在这个软件基础之上进行一些必要的工作,以把数据库管理系统提供的功能发挥出来。首先应该利用这个系统存放用户自己的数据,让数据库管理系统帮助我们把这些数据管理起来,其次还应有对这些数据进行操作并让这些数据发挥作用的应用程序,最后还需要一个维护整个系统正常运行的管理人员,比如当数据库出现故障或问题时应该如何处理以使数据库恢复正常,这个管理人员称为数据库系统管理员。因此,一个完整的数据库系统是基于数据库的一个计算机应用系统,这样的系统一般包括 4 个主要部分:数据库、数据库管理系统、应用程序和数据库系统管理员,如图 1-3 所示。

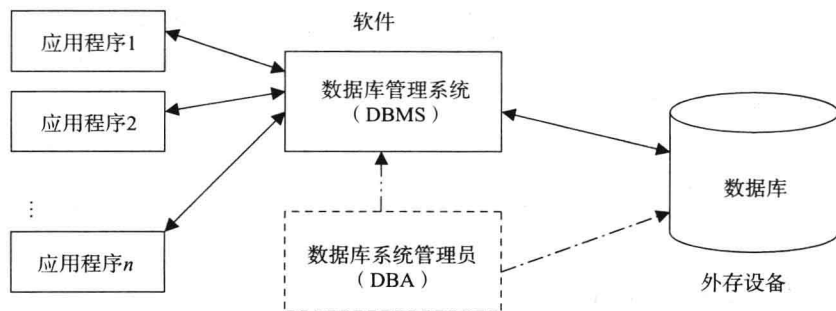


图 1-3 数据库系统的组成

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;数据库管理系统是管理数据库的系统软件,它实现数据库系统的各种功能,是整个数据库系统的核心;数据库系统管理员负责数据库的规划、设计、协调、维护和管理等工作;应用程序指以数据库以及数据库中的数据为基础的应用程序。

数据库系统的运行还要有计算机硬件环境和软件环境的支持,同时还要有使用数据库系统的用户。硬件环境是指保证数据库系统正常运行的最基本的内存、外存等硬件资源;软件环境是指数据库管理系统作为系统软件,需要建立在一定的操作系统环境上,没有合适的操作系统,数据库管理系统是无法正常运转的,比如 SQL Server 2008 的企业版就需要服务器版操作系统的支持。

至此,可以看出,数据库、数据库管理系统和数据库系统是 3 个不同的概念,数据库强调的是数据,数据库管理系统是系统软件,而数据库系统强调的是整个应用系统。

## 1.2.2 数据库管理系统的功能

数据库管理系统 (DBMS) 是数据库系统的核心,上文已经介绍了数据库的许多优点和功能,这些优点和功能并不是数据库中的数据固有的,而是数据库管理系统提供的。数据库管理系统的任务就是对数据资源进行管理,使之能为多个用户共享,同时保证数据的安全性、可靠性、完整性和一致性,还要保证数据的高度独立性。

简单地讲,数据库管理系统应该具备如下功能。

- 1) 数据定义功能: 定义数据的结构、数据与数据之间的关联关系、数据的完整性约束等。
  - 2) 数据更改功能: 实现对数据库中数据的更改,包括插入、删除和修改数据。
  - 3) 数据查询功能: 实现灵活的数据查询功能,使用户可以方便地查看数据库中的数据。
  - 4) 数据控制功能: 实现对数据库数据的安全性控制、完整性控制等各方面的控制功能。
  - 5) 数据管理功能: 实现数据库的备份和恢复。
  - 6) 数据通信功能: 在分布式数据库或提供网络操作功能的数据库中还必须提供数据的通信功能。
- 除上述功能之外,还有性能优化、并发控制等许多其他功能。