

# 数据库 技术及应用

—— Microsoft SQL  
Server 2008+Java

马忠贵 曾广平 编



国防工业出版社

National Defense Industry Press

# 数据库技术及应用

## ——Microsoft SQL Server 2008 + Java

马忠贵 曾广平 编

国防工业出版社

· 北京 ·

## 内 容 简 介

本书用通俗的语言将抽象的数据库理论具体化,以目前最流行的大型关系数据库 Microsoft SQL Server 2008 为背景,介绍了数据库的基本原理和主要技术,具体内容包数据库的基本概念、关系数据模型、关系数据理论、关系数据库标准语言 SQL、Microsoft SQL Server 2008 的使用、数据库的安全和维护、数据库设计方法、基于 Java 的数据库系统开发方法和实验指导等。

本书内容循序渐进、深入浅出,以一个读者耳熟能详的学生管理信息系统为例贯穿全书,并配有大量的实例、习题和实验项目。内容取材新颖,采用数据库基本理论与实际应用相结合的原则,在注重理论性、系统性、科学性的同时,兼顾培养读者的自主创新学习能力。

本书可作为高等学校非计算机专业高年级本科生和专科生的教材,也可供从事相关专业的工程技术人员和科研人员参考。

### 图书在版编目(CIP)数据

数据库技术及应用: Microsoft SQL Server 2008 +  
Java / 马忠贵, 曾广平编. —北京: 国防工业出版社,  
2012. 7  
ISBN 978-7-118-08180-0

I. ①数… II. ①马… ②曾… III. ①关系数据库 -  
数据库管理系统②JAVA 语言 - 程序设计 IV. ①TP311.138  
②TP312

中国版本图书馆 CIP 数据核字(2012)第 108401 号

※

国防工业出版社 出版发行  
(北京市海淀区紫竹院南路 23 号 邮政编码 100048)  
北京市李史山胶印厂印刷  
新华书店经售

\*

开本 787 × 1092 1/16 印张 14¼ 字数 324 千字  
2012 年 7 月第 1 版第 1 次印刷 印数 1—4000 册 定价 32.00 元

(本书如有印装错误,我社负责调换)

国防书店:(010)88540777  
发行传真:(010)88540755

发行邮购:(010)88540776  
发行业务:(010)88540717

# 前 言

数据库技术是研究数据库的结构、存储、设计、管理和使用的一门学科,已在当代的社会生活中获得了广泛的应用,渗透到工农业生产、商业、行政管理、科学研究、教育、工程技术和国防军事等各行各业,而且已围绕数据库技术形成了一个巨大的软件产业,即数据库管理系统和各类工具软件的开发与经营。随着信息技术的飞速发展,各行各业数据的采集、存储、处理和传播的数量也与日俱增,出现了各种应用广泛的数据库应用系统,如国家经济信息系统、企业管理信息系统、行政机关的办公自动化系统(Office Automation, OA)以及企业资源规划系统(Enterprise Resource Planning, ERP)等,都需要数据库技术的支持,而且数据已成为各行各业的宝贵资产。从应用的层次和深度看,可将数据库应用系统由低到高分电子数据处理(Electrical Data Processing, EDP)、管理信息系统(Management Information System, MIS)和决策支持系统(Decision Support System, DSS)3类。在 market 需求的驱动下,数据库技术及应用已成为当前高等院校计算机专业的必修课程及非计算机专业选修的核心课程之一。

本书用通俗的语言将抽象的数据库理论具体化,结合目前流行的数据库管理系统——Microsoft SQL Server 2008 讲述了数据库的基本理论与应用。目前,Microsoft SQL Server 2000 的使用范围已经很小,绝大多数用户使用的是 Microsoft SQL Server 2005。Microsoft SQL Server 2008 目前逐渐取代上一版本,成为主流的 SQL Server 产品,这也是本书选择 Microsoft SQL Server 2008 作为背景的原因。尽管本书重点介绍 Microsoft SQL Server Management Studio 中的数据库引擎服务,但有了这方面的知识后,可以很容易地学习其他的服务。

本书以大型关系数据库 Microsoft SQL Server 2008 为背景,介绍数据库的基本原理和主要技术,全书共分 8 章来进行论述。第 1 章为绪论,从数据管理技术的产生和发展引出数据库概念,围绕着数据库系统的组成介绍有关名词术语;然后介绍了数据模型的基本概念、表示方法以及数据模型的三要素;最后介绍了数据库系统的三级模式结构和二级映像功能。第 2 章主要介绍目前最重要的一种数据模型,即关系模型,内容包括关系模型的基本概念与术语、完整性约束、关系的运算、关系表达式的等价变化、关系的查询优化等。第 3 章对 Microsoft SQL Server 2008 系统进行概述,以使读者对该系统有整体的认识 and 了解,对 Microsoft SQL Server 2008 系统在易用性、可用性、可管理性、可编程性、动态开发、安全性等方面有一个初步的理解。内容包括 Microsoft SQL Server 2008 的身份验证模式,Microsoft SQL Server 2008 的使用,使用 Microsoft SQL Server Management Studio 和 Transact-

SQL 语句 2 种方式创建、修改和删除数据库, Transact-SQL 程序设计基础等。第 4 章介绍关系数据库标准语言 SQL, 内容包括 SQL 语言的发展过程和基本特点, DDL、DQL、DML、DCL, 视图, 存储过程及触发器。第 5 章介绍关系数据库规范化理论。首先是关系规范化的提出; 接着引入函数依赖和范式等基本概念; 然后介绍关系模式等价性判定和模式分解的方法; 最后简要介绍 2 种数据依赖的概念。第 6 章介绍数据库的安全和维护。内容包括数据库的安全性、数据完整性、并发控制和恢复技术 4 个方面的内容, 并以 Microsoft SQL Server 2008 为例进行具体说明。第 7 章主要介绍数据库设计的任务和特点、设计方法及设计步骤。以概念结构设计和逻辑结构设计为重点, 介绍每一个阶段的方法、技术以及注意事项。第 8 章以开发一个学生管理信息系统为例, 介绍使用 Java 语言进行 Microsoft SQL Server 2008 数据库系统的开发方法。内容有所取舍, 配有丰富的数据库实验项目和大量的实例。通过课堂教学与上机实践相结合的学习方式, 使读者系统地掌握数据库的基本原理和技术, 掌握关系型数据库管理系统 Microsoft SQL Server 2008 的使用和操作方法, 掌握数据库设计方法和步骤, 具有设计数据库模式以及开发数据库应用系统的基本能力。

本书采用数据库基本理论与实际应用相结合的原则, 在注重理论性、系统性、科学性的同时, 兼顾培养读者的自主创新学习能力。为此, 本书通过目前流行的数据库管理系统 Microsoft SQL Server 2008 的学习, 掌握数据库技术的基本原理, 并且使用高级程序设计语言 Java 和 Microsoft SQL Server 2008 开发 C/S 体系结构的具体管理系统, 旨在培养读者的综合实践与创新能力。

本书在编写过程中参考了大量与数据库相关的技术资料, 在此向其作者表示感谢。书中的全部 Transact-SQL 语句和 Java 程序都上机调试通过。由于编者水平和时间有限, 书中不妥之处在所难免, 恳请同行专家和广大读者批评指正。

马忠贵

2012 年 3 月于北京

# 目 录

|                          |    |
|--------------------------|----|
| 第1章 绪论 .....             | 1  |
| 1.1 数据、信息与数据处理 .....     | 1  |
| 1.2 数据管理技术的产生与发展 .....   | 2  |
| 1.2.1 人工管理阶段 .....       | 2  |
| 1.2.2 文件系统阶段 .....       | 3  |
| 1.2.3 数据库系统阶段 .....      | 4  |
| 1.3 数据库系统的组成 .....       | 6  |
| 1.4 数据模型 .....           | 9  |
| 1.4.1 概念模型 .....         | 10 |
| 1.4.2 数据模型 .....         | 12 |
| 1.5 数据库系统的结构 .....       | 16 |
| 1.5.1 数据库系统的3级模式结构 ..... | 16 |
| 1.5.2 数据库系统的2级映像 .....   | 18 |
| 1.5.3 数据库系统的体系结构 .....   | 19 |
| 1.6 小结 .....             | 20 |
| 习题 .....                 | 21 |
| 第2章 关系数据库 .....          | 22 |
| 2.1 关系数据模型的基本概念 .....    | 22 |
| 2.2 关系数据模型的集合论定义 .....   | 24 |
| 2.2.1 笛卡儿积 .....         | 24 |
| 2.2.2 关系 .....           | 24 |
| 2.2.3 基本关系的性质 .....      | 25 |
| 2.2.4 关系模式 .....         | 25 |
| 2.3 关系模型的完整性约束 .....     | 25 |
| 2.3.1 实体完整性 .....        | 26 |
| 2.3.2 参照完整性 .....        | 26 |
| 2.3.3 域完整性 .....         | 26 |
| 2.3.4 用户定义完整性 .....      | 26 |
| 2.4 关系代数 .....           | 27 |
| 2.4.1 传统的集合运算 .....      | 27 |
| 2.4.2 专门的关系运算 .....      | 29 |
| 2.5 小结 .....             | 31 |



|                                                     |           |
|-----------------------------------------------------|-----------|
| 习题 .....                                            | 32        |
| <b>第3章 Microsoft SQL Server 2008 数据库基础 .....</b>    | <b>33</b> |
| 3.1 Microsoft SQL Server 2008 简介 .....              | 33        |
| 3.2 Microsoft SQL Server 2008 的登录 .....             | 36        |
| 3.3 Microsoft SQL Server Management Studio 简介 ..... | 38        |
| 3.3.1 对象资源管理器 .....                                 | 39        |
| 3.3.2 查询编辑器 .....                                   | 39        |
| 3.3.3 已注册服务器 .....                                  | 41        |
| 3.4 Microsoft SQL Server 2008 数据库的创建与维护 .....       | 42        |
| 3.4.1 Microsoft SQL Server 2008 数据库的构成 .....        | 42        |
| 3.4.2 系统数据库 .....                                   | 43        |
| 3.4.3 创建数据库 .....                                   | 45        |
| 3.4.4 修改数据库 .....                                   | 49        |
| 3.4.5 删除数据库 .....                                   | 51        |
| 3.5 Transact-SQL 程序设计基础 .....                       | 51        |
| 3.5.1 SQL 与 Transact-SQL .....                      | 51        |
| 3.5.2 Microsoft SQL Server 2008 标识符 .....           | 53        |
| 3.5.3 Microsoft SQL Server 2008 数据类型 .....          | 54        |
| 3.5.4 常量、变量和运算符 .....                               | 57        |
| 3.5.5 批处理 .....                                     | 61        |
| 3.5.6 程序注释语句 .....                                  | 61        |
| 3.5.7 函数 .....                                      | 62        |
| 3.5.8 控制流语句 .....                                   | 66        |
| 3.6 小结 .....                                        | 70        |
| 习题 .....                                            | 70        |
| <b>第4章 关系数据库标准语言 SQL .....</b>                      | <b>72</b> |
| 4.1 SQL 的3级模式结构 .....                               | 72        |
| 4.2 SQL 的数据定义 .....                                 | 73        |
| 4.2.1 基本表的定义 .....                                  | 73        |
| 4.2.2 基本表的修改与删除 .....                               | 75        |
| 4.2.3 索引的建立和删除 .....                                | 76        |
| 4.3 SQL 数据查询 .....                                  | 78        |
| 4.3.1 单表无条件查询 .....                                 | 80        |
| 4.3.2 单表有条件查询 .....                                 | 82        |
| 4.3.3 分组查询和排序查询 .....                               | 83        |
| 4.3.4 多表查询 .....                                    | 84        |
| 4.3.5 嵌套查询 .....                                    | 88        |
| 4.4 SQL 的数据操纵 .....                                 | 91        |
| 4.4.1 插入数据 .....                                    | 91        |

|            |                       |            |
|------------|-----------------------|------------|
| 4.4.2      | 修改数据                  | 92         |
| 4.4.3      | 删除数据                  | 92         |
| 4.5        | 视图                    | 93         |
| 4.5.1      | 定义视图                  | 93         |
| 4.5.2      | 删除视图                  | 94         |
| 4.5.3      | 查询视图                  | 95         |
| 4.5.4      | 更新视图                  | 95         |
| 4.5.5      | 视图的作用                 | 96         |
| 4.6        | 存储过程                  | 96         |
| 4.6.1      | 创建存储过程                | 97         |
| 4.6.2      | 执行存储过程                | 99         |
| 4.6.3      | 存储过程的修改               | 100        |
| 4.6.4      | 存储过程的删除               | 101        |
| 4.6.5      | 存储过程的查看               | 101        |
| 4.7        | 触发器                   | 101        |
| 4.7.1      | 创建触发器                 | 102        |
| 4.7.2      | inserted 表和 deleted 表 | 104        |
| 4.7.3      | 修改触发器                 | 105        |
| 4.7.4      | 删除触发器                 | 105        |
| 4.8        | 数据控制                  | 105        |
| 4.8.1      | 数据控制简介                | 105        |
| 4.8.2      | 授权                    | 106        |
| 4.8.3      | 收回权限                  | 107        |
| 4.9        | 小结                    | 107        |
|            | 习题                    | 108        |
| <b>第5章</b> | <b>关系数据库规范化理论</b>     | <b>110</b> |
| 5.1        | 关系规范化的提出              | 110        |
| 5.1.1      | 问题的提出                 | 110        |
| 5.1.2      | 解决的方法                 | 111        |
| 5.1.3      | 关系模式规范化               | 112        |
| 5.2        | 函数依赖的基本概念             | 112        |
| 5.2.1      | 函数依赖                  | 112        |
| 5.2.2      | 函数依赖的3种基本情形           | 113        |
| 5.2.3      | 码的函数依赖                | 113        |
| 5.3        | 关系模式的规范化              | 114        |
| 5.3.1      | 第1范式                  | 114        |
| 5.3.2      | 第2范式                  | 115        |
| 5.3.3      | 第3范式                  | 116        |
| 5.3.4      | BCNF 范式               | 117        |



|                                              |            |
|----------------------------------------------|------------|
| 5.3.5 多值依赖与第4范式 .....                        | 118        |
| 5.4 关系模式规范化步骤 .....                          | 120        |
| 5.5 小结 .....                                 | 121        |
| 习题 .....                                     | 122        |
| <b>第6章 数据库的安全和维护 .....</b>                   | <b>123</b> |
| 6.1 数据库的安全性 .....                            | 123        |
| 6.1.1 数据库安全性概述 .....                         | 123        |
| 6.1.2 数据库安全性控制的一般方法 .....                    | 123        |
| 6.1.3 Microsoft SQL Server 2008 安全管理 .....   | 127        |
| 6.2 数据库的完整性 .....                            | 140        |
| 6.2.1 数据库完整性概述 .....                         | 140        |
| 6.2.2 完整性约束条件 .....                          | 141        |
| 6.2.3 完整性控制 .....                            | 142        |
| 6.2.4 Microsoft SQL Server 2008 的完整性 .....   | 142        |
| 6.3 数据库的并发控制 .....                           | 144        |
| 6.3.1 事务 .....                               | 144        |
| 6.3.2 并发控制概述 .....                           | 145        |
| 6.3.3 封锁 .....                               | 146        |
| 6.3.4 Microsoft SQL Server 2008 的并发控制 .....  | 148        |
| 6.4 数据库的备份和恢复技术 .....                        | 150        |
| 6.4.1 数据恢复的基本原则 .....                        | 151        |
| 6.4.2 故障类型和恢复策略 .....                        | 151        |
| 6.4.3 需要备份的数据 .....                          | 153        |
| 6.4.4 Microsoft SQL Server 2008 的备份和恢复 ..... | 154        |
| 6.5 小结 .....                                 | 161        |
| 习题 .....                                     | 162        |
| <b>第7章 数据库设计 .....</b>                       | <b>163</b> |
| 7.1 数据库设计概述 .....                            | 163        |
| 7.1.1 数据库设计的任务 .....                         | 163        |
| 7.1.2 数据库设计的特点 .....                         | 164        |
| 7.1.3 数据库设计的方法 .....                         | 164        |
| 7.1.4 数据库设计的步骤 .....                         | 165        |
| 7.2 需求分析 .....                               | 166        |
| 7.2.1 需求分析的任务 .....                          | 166        |
| 7.2.2 需求调查 .....                             | 167        |
| 7.2.3 需求分析 .....                             | 168        |
| 7.2.4 需求分析的结果 .....                          | 169        |
| 7.3 概念结构设计 .....                             | 170        |
| 7.3.1 概念结构设计概述 .....                         | 170        |

|            |                                                      |            |
|------------|------------------------------------------------------|------------|
| 7.3.2      | 概念结构设计的方法 .....                                      | 171        |
| 7.4        | 逻辑结构设计 .....                                         | 174        |
| 7.4.1      | 概念模型转换为关系数据模型 .....                                  | 175        |
| 7.4.2      | 关系模型的优化 .....                                        | 175        |
| 7.4.3      | 设计外模式 .....                                          | 176        |
| 7.5        | 物理结构设计 .....                                         | 178        |
| 7.5.1      | 物理结构设计概述 .....                                       | 178        |
| 7.5.2      | 关系模式的存取方法选择 .....                                    | 178        |
| 7.5.3      | 数据库存储结构的确定 .....                                     | 179        |
| 7.5.4      | 物理结构设计的评价 .....                                      | 180        |
| 7.6        | 数据库的实施 .....                                         | 180        |
| 7.6.1      | 建立实际的数据库结构 .....                                     | 180        |
| 7.6.2      | 数据载入 .....                                           | 180        |
| 7.6.3      | 编制与调试应用程序 .....                                      | 180        |
| 7.6.4      | 数据库试运行 .....                                         | 181        |
| 7.7        | 数据库的运行与维护 .....                                      | 181        |
| 7.8        | 小结 .....                                             | 182        |
|            | 习题 .....                                             | 182        |
| <b>第8章</b> | <b>Java 与 Microsoft SQL Server 2008 编程实例 .....</b>   | <b>184</b> |
| 8.1        | Java 概述 .....                                        | 184        |
| 8.1.1      | JDK 的安装 .....                                        | 184        |
| 8.1.2      | 环境变量的设置 .....                                        | 185        |
| 8.1.3      | Java 语言的特点 .....                                     | 185        |
| 8.2        | Java 开发环境 .....                                      | 186        |
| 8.3        | JDBC .....                                           | 187        |
| 8.3.1      | JDBC 驱动程序与安装 .....                                   | 187        |
| 8.3.2      | JDBC 的数据库访问模型 .....                                  | 188        |
| 8.3.3      | JDBC API 介绍 .....                                    | 189        |
| 8.4        | 数据库连接步骤 .....                                        | 190        |
| 8.4.1      | 加载驱动程序 .....                                         | 190        |
| 8.4.2      | 建立连接对象 .....                                         | 190        |
| 8.4.3      | 创建语句对象 .....                                         | 192        |
| 8.4.4      | 结果集 ResultSet 对象 .....                               | 194        |
| 8.4.5      | 关闭有关对象 .....                                         | 195        |
| 8.5        | 数据库连接示例 .....                                        | 195        |
| 8.5.1      | 使用 JDBC-ODBC 桥访问 Microsoft SQL Server 2008 数据库 ..... | 196        |
| 8.5.2      | 使用 JDBC 直接访问 Microsoft SQL Server 2008 数据库 .....     | 199        |
| 8.5.3      | 数据库添加、修改和删除 .....                                    | 202        |
| 8.6        | 可滚动和可更新的 ResultSet .....                             | 203        |

|                                              |     |
|----------------------------------------------|-----|
| 8.6.1 可滚动的 ResultSet .....                   | 203 |
| 8.6.2 可更新的 ResultSet .....                   | 204 |
| 8.7 小结 .....                                 | 206 |
| <b>附录 A 实验指导</b> .....                       | 207 |
| 实验 1 Microsoft SQL Server 2008 管理工具的使用 ..... | 207 |
| 实验 2 数据库和表的创建及维护 .....                       | 209 |
| 实验 3 数据库查询 .....                             | 213 |
| 实验 4 Microsoft SQL Server 2008 的安全性管理 .....  | 214 |
| 实验 5 数据库系统开发(综合设计型实验) .....                  | 215 |
| <b>附录 B 实验报告模板</b> .....                     | 217 |
| <b>参考文献</b> .....                            | 218 |

# 第1章 绪 论

随着计算机技术、通信技术和网络技术的发展,人类社会已经进入了信息化时代。当今,信息资源已成为各行各业最重要和最宝贵的财富和资源,建立一个行之有效的信息系统是各行各业生存和发展的重要条件。作为信息系统核心技术和重要基础的数据库技术正是瞄准这一目标应运而生、飞速发展起来的专门技术,并得到了越来越广泛的应用。

数据库技术是一门应用广泛、实用性强的技术,产生于20世纪60年代末,是数据管理的最新技术,是计算机科学的重要分支。目前,数据库技术作为信息系统的核心和基础,已成为计算机应用的主流,它的出现极大地促进了计算机应用向各行各业的渗透。伴随着云计算和物联网的蓬勃发展,使信息的收集变得更加全面、更加智能和更加深入,数据库仍可以为物联网的广泛应用提供基石。数据库的建设规模、数据库信息量的大小和使用频度已成为衡量该国家信息化程度的重要标志。

伴随着计算机技术,特别是计算机网络和互联网的飞速发展,数据库技术已应用到社会生活的各个领域,引发了“信息爆炸”。IDC(International Data Cooperation)报告指出,全球2007年产生的数据量为 $2.81 \times 10^{11}$  GB。到2008年底,全球数据内容的总量激增为 $4.87 \times 10^{11}$  GB。在如此浩瀚的信息空间中,如何实现信息的存储、访问、共享以及安全是一个至关重要的问题。数据库系统就是研究如何科学地组织数据和存储数据的理论和方法,如何高效地检索数据和处理数据,以及如何既减少数据冗余又能保证数据安全,实现数据共享的计算机应用技术。数据库可以提供高效存储、高效访问、数据共享和数据安全。

本章从数据管理技术的产生和发展引出数据库概念,围绕着数据库系统的组成介绍有关名词术语;然后介绍数据模型的基本概念、表示方法以及数据模型的三要素;最后介绍数据库系统的三级模式结构和二级映像功能。

## 1.1 数据、信息与数据处理

### 1) 数据

数据是数据库中存储的基本对象。数据在大多数人脑中的第一反应就是数字,其实数据不只是简单的数字,还可以是文字、图形、图像、声音、视频、学生档案(如41051001,王强,男,1990-1-1,内蒙,通信工程1001班)、工作日志、货物的运输情况等。

数据是描述事物的符号记录,是信息的符号表示或载体。数据=量化特征描述+非量化特征描述。例如,天气预报中,温度的高低可以量化表示,而“刮风”或“下雨”等特征则需要用文字或图形符号进行描述,它们都是数据,只是数据类型不同而已。

### 2) 信息

信息是数据的内涵,是数据的语义解释,是对现实世界中各种事物的存在方式、运动

状态或事物间联系形式的综合反映。信息是可以被感知、存储、加工、传递和再生。数据与信息是密不可分的,如 1980,若描述一个人的出生日期,表示 1980 年;若描述一根钢管的长度表示 1980mm。

3) 数据处理

数据处理是将数据转换成信息的过程,包括对数据收集、存储、分类、加工、检索、维护等一系列活动,其目的是从大量的原始数据中抽取和推导出有价值的信息。数据、信息及数据处理之间的关系如图 1-1 所示。

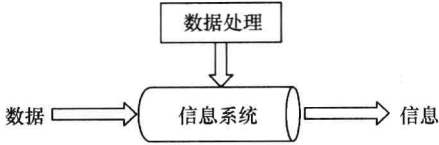


图 1-1 数据、信息及数据处理之间的关系

1.2 数据管理技术的产生与发展

数据库技术是因数据管理任务的需要应运而生。数据管理就是对数据进行分类、组织、编码、存储、检索、传播和利用的一系列活动的总和,是数据处理的核心。数据管理技术是随着计算机技术的发展而发展的,发展可以大体归为人工管理、文件系统和数据库系统 3 个阶段。在计算机软、硬件的发展和应用需求的推动下,每一阶段的发展都以数据存储冗余不断减小、数据独立性不断增强、数据操作更加方便和简单为标志。下面简单描述数据管理技术发展的 3 个阶段。

1.2.1 人工管理阶段

在计算机出现之前,人们运用常规的手段从事记录、存储和对数据加工,也就是利用纸张记录和利用计算工具(算盘、计算尺)进行计算,并主要通过人的大脑来管理和利用这些数据。20 世纪 50 年代中期以前属于人工管理阶段。在这一阶段,计算机主要用于科学计算。在硬件方面,外部存储器只有磁带、卡片和纸带等,还没有磁盘等直接存取数据的存储设备;在软件方面,只有汇编语言,既没有操作系统,也无数据管理方面的软件,而且数据处理方式基本是批处理,因此从计算机内记录的数据上看,数据量小且无结构,用户直接管理,且数据间缺乏逻辑组织,数据仅依赖特定的应用,缺乏独立性。

在人工管理阶段,数据管理有如下几个特点:

(1) 数据不单独保存。因为该阶段计算机主要应用于科学计算,对于数据保存的需求尚不迫切,且数据与程序是一个整体,数据只为本程序所使用,因此所有程序的数据均不单独保存。

(2) 系统没有专用的软件对数据进行管理。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。因此,每个应用程序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,程序员负担很重。

(3) 数据不共享。数据是面向程序的,1 组数据只能对应 1 个程序。多个应用程序

涉及某些相同的数据时,也必须各自定义,因此程序之间有大量的冗余数据。

(4) 数据不具有独立性。程序依赖于数据,如果数据的类型、格式或输入/输出方式等逻辑结构或物理结构发生变化,必须对应用程序做出相应的修改。数据脱离了程序就无任何存在的价值,数据无独立性。

人工管理阶段程序与数据之间的关系如图 1-2 所示。

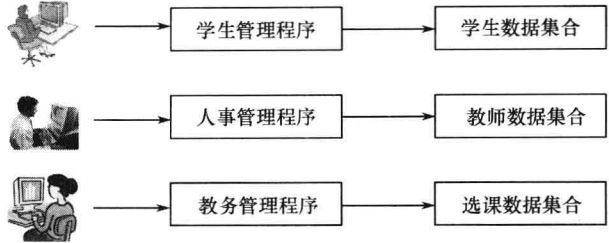


图 1-2 人工管理阶段程序与数据之间的关系

### 1.2.2 文件系统阶段

从 20 世纪 50 年代后期到 60 年代中期,计算机不仅用于科学计算,而且还大量应用于信息管理,大量的数据存储、检索和维护成为紧迫的需求。在硬件方面,有了磁盘、磁鼓等直接存储设备;在软件方面,出现了高级语言和操作系统,且操作系统中有了专门管理数据的软件,一般称为文件系统;在处理方式方面,不仅有批处理,也有联机实时处理。

用文件系统管理数据的特点如下:

(1) 数据以文件形式可以长期保存在外存储设备上。由于计算机大量用于数据处理,数据需要长期保存在外存储设备上,以使用户可随时对文件进行查询、修改、增加和删除等处理。

(2) 文件系统可对数据的存取进行管理。有专门的软件即文件系统进行数据管理,文件系统把数据组织成相互独立的数据文件,利用“按名访问,按记录存取”的管理技术,对文件进行修改、增加和删除的操作。

(3) 数据共享性差,冗余度大。由于数据的基本存取单位是记录,因此,程序员之间很难明白他人数据文件中数据的逻辑结构。理论上,1 个用户可通过文件管理系统访问很多数据文件,然而实际上,1 个数据文件只能对应于同一程序员的 1 个或几个程序,不能共享,即文件仍然是面向应用的。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,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。由于相同数据的重复存储、各自管理,在进行更新操作时,容易造成数据的不一致性。

(4) 数据独立性差。文件系统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,若要对现有的数据增加一些新的应用会很困难,系统不容易扩充。数据和程序相互依赖,一旦改变数据的逻辑结构,必须修改相应的应用程序。而应用程序发生变化,如改用另一种程序设计语言来编写程序,也需修改数据结构。因此,数据和程序之间缺乏独立性。

文件系统阶段程序与数据之间的关系如图 1-3 所示。



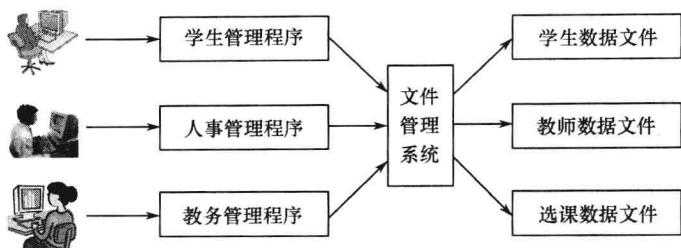


图 1-3 文件系统阶段程序与数据之间的关系

### 1.2.3 数据库系统阶段

20 世纪 60 年代后期,计算机硬件、软件有了进一步的发展。计算机应用于管理的规模更加庞大,数据量急剧增加;硬件方面出现了大容量磁盘,使计算机联机存取大量数据成为可能;硬件价格下降,而软件价格上升,使开发和维护系统软件的成本增加。文件系统的管理方法已无法适应开发应用系统的需要。为解决多个用户、多个应用程序共享数据的需求,出现了统一管理数据的专门软件系统,即数据库管理系统。用数据库系统管理数据比文件系统具有明显的优点,从文件系统到数据库系统,标志着数据管理技术的飞跃。

数据库系统的特点如下:

(1) 数据结构化。有了数据库系统后,数据库中的任何数据都不属于任何应用。数据是公共的,结构是全面的。在描述数据时不仅要描述数据本身,还要描述数据之间的联系。它是按照某种数据模型,将某一领域的各种数据有机地组织到一个结构化的数据库中。数据结构化是数据库的主要特征之一,也是数据库系统与文件系统的本质区别。

例如,要建立学生学籍管理系统,该系统包含学生(学号、姓名、性别、系)、课程(课程号、课程名、学分、教师)、成绩(学号、课程号、成绩)等数据,分别对应 3 个文件。若采用文件处理方式,因为文件系统只表示记录内部的联系,而不涉及不同文件记录之间的联系,若查找某个学生的学号、姓名、所选课程的名称和成绩,必须编写一段程序来实现。而采用数据库方式,数据库系统不仅描述数据本身,还描述数据之间的联系,上述项目可以非常容易地联机查到。

(2) 数据共享性高、冗余度低,易扩展。数据库系统从全局角度看待和描述数据,数据不再面向某个应用程序而是面向整个系统,因此数据可以被多个用户、多个应用共享使用。这样便减少了不必要的数据冗余,节约存储空间,同时也避免了数据之间的不相容性与不一致性。由于数据面向整个系统,是有结构的数据,不仅可被多个应用共享使用,而且容易增加新的应用,这就使得数据库系统弹性大,易于扩展,可以适应各种用户的要求。

(3) 数据独立性高。数据的独立性是指数据的逻辑独立性和数据的物理独立性。数据的逻辑独立性是指用户的应用程序与数据库的逻辑结构是相互独立的,即当数据的总体逻辑结构改变时,数据的局部逻辑结构不变,由于应用程序是依据数据的局部逻辑结构编写的,所以应用程序不必修改,从而保证了数据与程序间的逻辑独立性。例如,在原有的记录类型之间增加新的联系,或在某些记录类型中增加新的数据项,均可确保数据的

逻辑独立性。

数据的物理独立性是指用户的应用程序与存储在磁盘上的数据库中数据是相互独立的,即当数据的存储结构改变时,数据的逻辑结构不变,从而应用程序也不必改变。例如,改变存储设备和增加新的存储设备,或改变数据的存储组织方式,均可确保数据的物理独立性。数据独立性是由数据库系统的二级映像功能来实现的,将在 1.5 节中讲解。

(4) 数据由数据库管理系统统一管理和控制。数据库为多个用户和应用程序所共享,对数据的存取往往是并发的,即多个用户可以同时存取数据库中的数据,甚至可以同时存取数据库中的同一个数据。为确保数据库数据的正确有效和数据库系统的有效运行,数据库管理系统提供下述 4 个方面的数据控制功能。

① 数据的安全性控制。数据的安全性是指保护数据以防止不合法使用数据造成数据的泄密和破坏,保证数据的安全和机密,使每个用户只能按规定对某些数据以某些方式进行使用和处理。例如,系统提供口令检查或其他手段来验证用户身份,防止非法用户使用系统;也可以对数据的存取权限进行限制,只有通过检查后才能执行相应的操作。

② 数据的完整性控制。数据的完整性是指系统通过设置一些完整性规则以确保数据的正确性、有效性和相容性。完整性控制将数据控制在有效的范围内,或保证数据之间满足一定的关系。正确性是指数据的合法性,如年龄属于数值型数据,只能含 0,1,⋯,9,不能含字母或特殊符号;有效性是指数据是否在其定义的有效范围,如月份只能用 1~12 之间的正整数表示;相容性是指表示同一事实的 2 个数据应相同,否则就不相容,如 1 个人不能有 2 个性别。

③ 数据的并发控制。多用户同时存取或修改数据库时,可能会发生相互干扰而提供给用户不正确的数据,并使数据库的完整性受到破坏。因此,必须对多用户的并发操作加以控制和协调,防止相互干扰而得到错误的结果。

④ 数据恢复。计算机系统出现各种故障是很正常的,数据库中的数据被破坏和丢失也是可能的。当数据库被破坏或数据不可靠时,系统有能力将数据库从错误状态恢复到最近某一时刻的正确状态。

数据库系统阶段程序与数据之间的关系如图 1-4 所示。

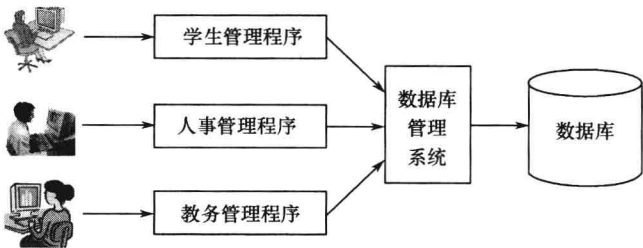


图 1-4 数据库系统阶段程序与数据之间的关系

从文件系统管理发展到数据库系统管理是信息处理领域的一个重大变化。在文件系统阶段人们关注的是系统功能的设计,因此程序设计处于主导地位,数据服从于程序设计;而在数据库系统阶段,数据的结构设计成为信息系统首先关心的问题。

数据库技术经历了以上 3 个阶段的发展,已有了比较成熟的数据库技术,但随着计算

机软、硬件的发展,数据库技术仍在不断向前发展。

20 世纪 70 年代,层次、网状、关系 3 大数据库系统奠定了数据库技术的概念、原理和方法。80 年代以来,数据库技术在商业领域的巨大成功刺激了其他领域对数据库技术需求的迅速增长,这些新的领域为数据库应用开辟了新的天地;另一方面,在应用中提出的一些新的数据管理的需求也直接推动了数据库技术的研究和发展,尤其是面向对象数据库系统。另外,数据库技术不断与其他计算机分支结合,向高一级的数据库技术发展。例如,数据库技术与分布处理技术相结合出现了分布式数据库系统,数据库技术与并行处理技术相结合出现了并行数据库系统以及图形图像数据库等。

### 1.3 数据库系统的组成

数据库系统(Database System,DBS)是指在计算机系统中引入数据库后的系统,是可运行、可维护的软件系统,一般由数据库、数据库管理系统(及其开发工具)、应用系统、数据库管理员和用户构成,如图 1-5 所示。

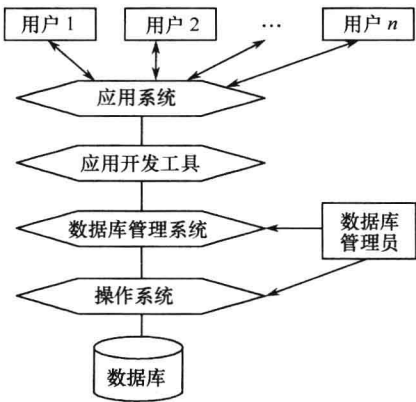


图 1-5 数据库系统的组成

#### 1) 数据库

顾名思义,数据库(Database,DB)是存放数据的仓库。只不过这个仓库在计算机存储设备上,按一定格式存放数据。数据是自然界事物特征的符号描述,而且能够被计算机处理。数据存储的目的是为了从大量的数据中发现有价值的信息,这些有价值的信息就是信息。

数据库是长期存储在计算机内大量的、有结构的、可共享的数据集合。数据库的基本特征:数据库中的数据按一定的数据模型组织、描述和储存,具有较小的冗余度、较高的数据独立性和易扩展性,并为各种用户共享(多个用户同时使用同一个数据库中的数据),而且数据库本身不是独立存在的,它是组成数据库系统的一部分。在实际应用中,人们面对的是数据库系统。

#### 2) 数据库管理系统

数据库管理系统(Database Management System,DBMS)是管理数据库的系统软件,是位于用户与操作系统之间的一层数据管理软件,是数据库系统的核心组成部分。用户在