高等学校计算机专业教材

数据库原理及应用教程

陈志泊　李冬梅　王春玲　编

人 民 邮 电 出 版 社

需要全本请访问 www.ertongbook.com

内 容 提 要

本书系统地讲述了数据库技术的基本原理和应用。全书共七章，主要内容包括：数据库系统概述、关系模型、SQL 语言、关系数据库理论、数据库安全与保护、数据库设计和 SQL Server 2000 数据库管理系统。本书除介绍数据库技术的基本原理外，还以 SQL Server 2000 为背景介绍了数据库技术的实现，包括数据库和数据表的维护、查询与统计、视图管理、存储过程和触发器的管理、用户管理、约束和默认管理、数据库的备份和还原、Transact-SQL 程序设计等内容，使读者可以充分利用 SQL Server 2000 平台深刻理解数据库技术的原理，达到理论和实践的紧密结合。

本书内容循序渐进、深入浅出，每一章节都给出了大量的实例。同时，每章后面都附有相应的习题，便于读者巩固所学知识。

本书可作为大学计算机及相关专业的本、专科教材，也可供从事计算机软件工作的科技人员、工程技术人员以及其他有关人员参阅。

图书在版编目（CIP）数据

数据库原理及应用教程 / 陈志泊，李冬梅，王春玲编. —北京：人民邮电出版社，2002.3
高等学校计算机专业教材
ISBN 7–115–10155–8

Ⅰ．数…　Ⅱ．①陈…②李…③王…　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（2002）第 009671 号

出版者的话

为了适应我国大学本科计算机专业教育发展对教材的需要，我社特邀请教育部所属的中国人民大学、中国地质大学、中国农业大学、北京科技大学、北京林业大学、北京语言文化大学（排名不分先后）6 所高等学校的“ 计算机科学与技术系 ”系主任及资深教授组成专家组，规划、编写、审订了本套教材。

读者对象　　本套教材的主要读者对象是普通高等院校计算机科学与技术专业的学生，兼顾信息、自动化及机电类专业学生学习计算机课程的需要。由于这些专业对学生的培养目标是掌握计算机软件与硬件的基本理论和方法，能从事计算机应用、软件研制、技术开发和管理工作的高级技术人才，因此，本套教材的内容既注意计算机高级技术人才所应具有的完整知识结构，又适当侧重主要专业知识。

教材特点　　本套教材参考了美国 IEEE/CS 和 ACM 技术委员会 2001 年推荐的课程 CC2001(Computing Curricula 2001，参见 http://www.csab.org/~csab，http://cs.nju.edu.cn/~gchen/teaching/cc2001/overview-bok.html 2001)及我国几十所高等院校计算机专业的 2001 年教学计划进行规划。教材内容在选择国际上先进的计算机理论、技术的同时，务求符合我国目前教学环境的实际情况，有一定深度，又有较高的实用性。

根据大学教学的特点，本套教材主要包括必修课程、选修课程和辅助课程三类教材。除了每本教材内容自成体系外，还考虑了它在整个教学计划中的安排顺序，适当增加了承上启下的内容。

写作风格　　本套教材根据内容需要，沿着“ 这是什么、有什么用、怎样用、怎样用得更好 ”的思路编写教材，通过讲解具体知识，传授学习方法，使学生达到掌握理论和技术的目的，同时力求文笔流畅，言简意赅。

教材每一章除基本知识外，还有本章要点、小结、思考与练习题。有些教材还附加教学大纲（包括教学重点、难点，讲授知识点的参考学时数），操作性较强的课程还配有实验教材（包括上机应具备的软硬件环境和实验内容及方法）。为了方便教师教学，本套教材提供了演示稿，可到人民邮电出版社网站（http://www.ptpress.com.cn）的“ 教材出版图书出版中心 ”→“ 教材附件 ”→“ 课件 ”下载。

本套教材的作者都具有多年教学经验，教材的草稿也都在各自的授课过程中多次使用。这套教材的出版，为计算机专业的师生提供了新的选择。我们希望这套以易教、易学，朴实、实用为特色的教材在培养信息化建设专业人才方面做出应有的贡献。

欢迎广大读者对本套教材的不足之处提出批评和建议。

2002 年 1 月

编 者 的 话

数据库技术自从 20 世纪 60 年代中期产生以来，无论是理论还是应用方面都已变得相当的重要和成熟，成为计算机科学的重要分支。数据库技术是计算机领域发展最快的学科之一，也是应用很广、实用性很强的一门技术。目前，数据库技术已从第一代的网状、层次数据库系统，第二代的关系数据库系统，发展到以面向对象模型为主要特征的第三代数据库系统。

随着计算机技术飞速发展及其应用领域的扩大，特别是计算机网络和 Internet 的发展，基于计算机网络和数据库技术的信息管理系统、应用系统得到了突飞猛进的发展。如事务处理系统、地理信息系统（GIS）、联机分析系统、决策支持系统、企业资源规划（ERP）、客户关系管理（CRM）、数据仓库和数据挖掘等系统都是以数据库技术作为重要的支撑的。可以说，只要有计算机存在，就存在着数据库技术。因此，数据库技术的基本知识和基本技能已成为计算机及相关专业的必修内容。

本书系统地讲述了数据库技术的基本原理和应用实践。全书共七章，第 1 章是数据库系统概述。主要介绍了数据管理技术发展的三个阶段及各个阶段的特点，数据库管理系统的组成、结构和功能，数据库技术的发展。第 2 章是关系模型。主要介绍了层次模型、网状模型和关系模型的各自特点，关系、关系模式和关系数据库的有关基本概念以及关系代数和关系运算。第 3 章是 SQL 语言，主要介绍了标准 SQL 语言的基本语法以及应用，还列举了大量的实例帮助读者理解和掌握 SQL 语言的使用与特点。第 4 章是关系数据库理论，本章内容是进行数据库设计所必需的理论基础，主要讲解了函数依赖的概念、1NF、2NF、3NF 和 BCNF 的定义及其规范化的方法。第 5 章是数据库安全与保护。本章从安全性控制、完整性控制、并发性控制和数据库恢复四方面讨论了数据库的安全保护功能。第 6 章是数据库设计，介绍了数据库设计的六个阶段：系统需求分析、概念结构设计、逻辑结构设计、物理设计、数据库实施、数据库运行与维护。对于每一阶段，都分别详细讨论了其相应的任务、方法和步骤。第 7 章是 SQL Server 2000 数据库管理系统，主要以 SQL Server 2000 为背景介绍了数据库技术的实现，包括数据库和数据表的维护、查询与统计、视图管理、存储过程和触发器的管理、用户管理、约束和默认管理、数据库的备份和还原、Transact-SQL 程序设计等内容。

本书内容全面，深入浅出，不仅适合于教学，也适合于广大计算机爱好者学习有关数据库的知识。作为教材，建议讲课学时为 40 学时，实验为 20 学时。第 7 章内容可以结合实验内容和其他章节的实践内容同时讲解。另外，除实验学时外，最好安排学生自由上机的时间，加强学生的实际动手能力。

本书由陈志泊主编，全书由陈志泊、李冬梅和王春玲编写。另外，张惠萍、黄金燕等参加了文字编辑、PowerPoint 讲稿的制作工作，马振龙参加了书稿校对方面的工作，在此表示衷心的感谢！

由于时间仓促，加之作者水平有限，书中不当之处在所难免，敬请读者批评指正。

编　者

2002 年 1 月

目　录

第 1 章　数据库系统概述 · · · · · 1

1.1　信息、数据与数据处理 · · · · · 1
1.1.1　数据与信息 · · · · · 1
1.1.2　数据处理 · · · · · 2
1.2　数据库技术的产生、发展与研究领域 · · · · · 2
1.2.1　数据库技术的产生与发展 · · · · · 2
1.2.2　数据库技术的研究领域 · · · · · 6
1.3　数据库系统的组成和结构 · · · · · 6
1.3.1　数据库系统的组成 · · · · · 6
1.3.2　数据库系统的结构 · · · · · 8
1.4　数据库管理系统 · · · · · 11
1.4.1　DBMS 的主要功能 · · · · · 11
1.4.2　DBMS 的组成 · · · · · 12
1.4.3　DBMS 的数据存取的过程 · · · · · 13
1.5　数据模型 · · · · · 13
1.5.1　数据模型的组成要素 · · · · · 13
1.5.2　数据之间的联系 · · · · · 14
1.5.3　数据模型的分类 · · · · · 16
1.6　数据模型与数据库系统的发展 · · · · · 22
1.6.1　第一代数据库系统 · · · · · 22
1.6.2　第二代数据库系统 · · · · · 23
1.6.3　第三代数据库系统 · · · · · 24
1.7　面向对象的数据库技术 · · · · · 25
1.7.1　面向对象的模型的核心概念 · · · · · 25
1.7.2　面向对象的数据库系统的特点 · · · · · 27
1.7.3　面向对象的数据库语言 · · · · · 28
1.7.4　面向对象的数据库的研究内容 · · · · · 28
1.8　数据库技术与其他相关技术的结合 · · · · · 29
1.8.1　分布式数据库 · · · · · 29
1.8.2　主动数据库 · · · · · 31
1.8.3　多媒体数据库 · · · · · 33
1.9　小结 · · · · · 34
习题 · · · · · 34

需要全本请访问 www.ertongbook.com

第 2 章　关系模型 · · · · · 36

2.1　关系模型 · · · · · 36
2.2　关系的形式化定义 · · · · · 37
2.3　关系的性质 · · · · · 39
2.4　关系的键 · · · · · 41
2.4.1　候选键与关系键 · · · · · 41
2.4.2　主属性与非码属性 · · · · · 41
2.4.3　外部关系键 · · · · · 42
2.4.4　关系模型的完整性 · · · · · 42
2.5　关系数据库模式与关系数据库 · · · · · 43
2.5.1　关系模式和关系数据库模式 · · · · · 43
2.5.2　关系数据库 · · · · · 44
2.6　关系代数 · · · · · 44
2.6.1　关系代数的分类及其运算符 · · · · · 44
2.6.2　传统的集合运算 · · · · · 45
2.6.3　专门的关系运算 · · · · · 47
2.7　关系演算 · · · · · 51
2.7.1　元组关系演算语言 · · · · · 52
2.7.2　域关系演算语言 QBE · · · · · 57
2.8　小结 · · · · · 61
习题 · · · · · 61

第 3 章　关系数据库标准语言 SQL · · · · · 63

3.1　SQL 语言的基本概念与特点 · · · · · 63
3.1.1　SQL 语言的发展及标准化 · · · · · 63
3.1.2　SQL 语言的基本概念 · · · · · 64
3.1.3　SQL 语言的主要特点 · · · · · 64
3.2　SQL 数据定义 · · · · · 65
3.2.1　字段数据类型 · · · · · 65
3.2.2　创建、修改和删除数据表 · · · · · 67
3.2.3　设计、创建和维护索引 · · · · · 73
3.3　SQL 数据查询 · · · · · 74
3.3.1　SELECT 命令的格式与基本使用 · · · · · 74
3.3.2　条件查询 · · · · · 76
3.3.3　常用库函数及统计汇总查询 · · · · · 79
3.3.4　分组查询 · · · · · 80
3.3.5　查询的排序 · · · · · 81
3.3.6　数据表连接及连接查询 · · · · · 82

3.3.7 子查询 · · · · · 85
3.3.8 合并查询 · · · · · 90
3.3.9 存储查询结果到表中 · · · · · 90
3.4 SQL 数据更新 · · · · · 90
3.4.1 插入数据记录 · · · · · 90
3.4.2 修改数据记录 · · · · · 92
3.4.3 删除数据记录 · · · · · 93
3.5 视图 · · · · · 93
3.5.1 定义和删除视图 · · · · · 94
3.5.2 查询视图 · · · · · 95
3.5.3 更新视图 · · · · · 95
3.6 SQL 数据控制 · · · · · 96
3.6.1 权限与角色 · · · · · 96
3.6.2 系统权限与角色的授予与收回 · · · · · 97
3.6.3 对象权限与角色的授予与收回 · · · · · 98
3.7 小结 · · · · · 99
习题 · · · · · 99

第 4 章 关系数据库理论 · · · · · 101

4.1 规范化问题的提出 · · · · · 101
4.1.1 规范化理论的主要内容 · · · · · 101
4.1.2 不合理的关系模式存在的存储异常问题 · · · · · 101
4.2 函数依赖 · · · · · 104
4.2.1 函数依赖的定义及性质 · · · · · 104
4.2.2 完全函数依赖与部分函数依赖 · · · · · 106
4.2.3 传递函数依赖 · · · · · 106
4.3 范式 · · · · · 106
4.3.1 第一范式 · · · · · 107
4.3.2 第二范式 · · · · · 108
4.3.3 第三范式 · · · · · 110
4.3.4 BC 范式 · · · · · 112
4.3.5 多值依赖与 4NF · · · · · 115
4.4 关系模式的规范化 · · · · · 118
4.4.1 关系模式规范化的目的和原则 · · · · · 118
4.4.2 关系模式规范化的步骤 · · · · · 118
4.4.3 关系模式规范化的要求 · · · · · 119
4.5 小结 · · · · · 121
习题 · · · · · 122

第 5 章　数据库安全保护 ····· 124

5.1　数据库的安全性 ····· 124
5.1.1　数据库安全性的含义 ····· 124
5.1.2　安全性控制的一般方法 ····· 124
5.2　完整性控制 ····· 129
5.2.1　数据库完整性的含义 ····· 129
5.2.2　完整性规则的组成 ····· 129
5.2.3　完整性约束条件的分类 ····· 130
5.3　并发控制与封锁 ····· 131
5.3.1　数据库并发性的含义 ····· 131
5.3.2　事务（Transaction） ····· 132
5.3.3　并发操作与数据的不一致性 ····· 133
5.3.4　封锁 ····· 135
5.4　数据库的恢复 ····· 141
5.4.1　数据库恢复的含义 ····· 141
5.4.2　数据库恢复的原理及其实现技术 ····· 141
5.4.3　数据库的故障和恢复的策略 ····· 143
5.5　小结 ····· 145
习题 ····· 146

第 6 章　数据库设计 ····· 147

6.1　数据库设计概述 ····· 147
6.1.1　数据库设计的任务、内容和特点 ····· 147
6.1.2　数据库设计方法简述 ····· 148
6.1.3　数据库设计的步骤 ····· 150
6.2　系统需求分析 ····· 153
6.2.1　需求分析的任务 ····· 153
6.2.2　需求分析的方法 ····· 155
6.3　概念结构设计 ····· 157
6.3.1　概念结构设计的必要性 ····· 157
6.3.2　概念模型的特点 ····· 157
6.3.3　概念结构设计的方法与步骤 ····· 158
6.4　逻辑结构设计 ····· 169
6.4.1　逻辑结构设计的任务和步骤 ····· 169
6.4.2　初始关系模式设计 ····· 169
6.4.3　关系模式规范化 ····· 171
6.4.4　模式评价与改进 ····· 171
6.5　数据库物理设计 ····· 173

6.5.1 确定物理结构 ····· 173
6.5.2 评价物理结构 ····· 175
6.6 数据库实施 ····· 175
6.6.1 建立实际数据库结构 ····· 175
6.6.2 装入数据 ····· 175
6.6.3 应用程序编码与调试 ····· 176
6.6.4 数据库试运行 ····· 176
6.6.5 整理文档 ····· 177
6.7 数据库运行和维护 ····· 177
6.7.1 维护数据库的安全性与完整性 ····· 177
6.7.2 监测并改善数据库性能 ····· 177
6.7.3 重新组织和构造数据库 ····· 178
6.8 小结 ····· 178
习题 ····· 179

第 7 章 SQL Server 2000 数据库管理系统 ····· 181

7.1 SQL Server 2000 的新特性 ····· 181
7.1.1 数据库增强 ····· 182
7.1.2 联合数据库服务 ····· 183
7.2 SQL Server 2000 的主要组件 ····· 184
7.3 Transact-SQL 程序设计 ····· 186
7.3.1 变量 ····· 186
7.3.2 流程控制命令 ····· 187
7.3.3 其他命令 ····· 191
7.3.4 常用函数 ····· 194
7.4 企业管理器 ····· 202
7.5 查询分析器 ····· 204
7.6 管理数据库 ····· 205
7.6.1 系统数据库 ····· 205
7.6.2 实例数据库 ····· 206
7.6.3 创建数据库 ····· 206
7.6.4 查看数据库信息 ····· 208
7.6.5 更改数据库 ····· 209
7.6.6 删除数据库 ····· 211
7.6.7 压缩数据库 ····· 211
7.7 管理数据表 ····· 213
7.7.1 用 Enterprise Manager 创建数据库表 ····· 213
7.7.2 修改表 ····· 214

7.7.3 查看表 ····· 215
7.7.4 用 Enterprise Manager 删除表 ····· 217
7.8 用 Enterprise Manager 管理数据 ····· 217
7.8.1 添加数据 ····· 217
7.8.2 删除数据 ····· 217
7.8.3 修改数据 ····· 218
7.9 SQL Server 的账号和存取权限 ····· 218
7.9.1 SQL Server 的验证模式 ····· 218
7.9.2 账号和角色 ····· 220
7.10 索引 ····· 230
7.10.1 创建索引 ····· 230
7.10.2 查看与修改索引 ····· 234
7.10.3 删除索引 ····· 236
7.11 数据完整性 ····· 236
7.11.1 数据完整性概述 ····· 236
7.11.2 规则 ····· 236
7.11.3 默认 ····· 240
7.12 数据查询 ····· 243
7.13 存储过程和触发器 ····· 243
7.13.1 存储过程概述 ····· 243
7.13.2 创建存储过程 ····· 244
7.13.3 管理存储过程 ····· 247
7.13.4 触发器概述 ····· 249
7.13.5 创建触发器 ····· 250
7.13.6 触发器的原理 ····· 253
7.13.7 INSTEAD OF 触发器 ····· 254
7.13.8 管理触发器 ····· 254
7.14 视图 ····· 256
7.14.1 创建视图 ····· 256
7.14.2 管理视图 ····· 257
7.15 备份和还原 ····· 258
7.15.1 备份和还原概述 ····· 259
7.15.2 创建备份设备 ····· 262
7.15.3 数据库备份 ····· 264
7.15.4 还原数据库 ····· 265
7.15.5 备份和还原系统数据库 ····· 266
7.16 小结 ····· 268
习题 ····· 269

参考文献 ····· 274

需要全本请访问 www.ertongbook.com

第 1 章　数据库系统概述

数据库（Database），简单地说就是数据的仓库，即数据存放的地方。我们周围有许多数据库的例子，如通讯录是一个小数据库，图书馆则是一个典型的大型数据库。小数据库尚可用手工管理，而大型数据库必须由计算机进行管理。在计算机三大主要应用领域（科学计算、过程控制和数据处理）中，数据处理所占比例约为 70%。20 世纪 60 年代末，数据库技术作为数据处理的最新技术应运而生。

本章首先回顾数据管理技术的三个发展阶段，然后介绍数据库中有关的主要概念。学习本章后，读者应了解数据库的三个发展阶段及各阶段的主要特点，掌握数据库中的有关的基本概念、数据库系统的组成及各部分的主要功能，重点掌握实体、属性定义和实体型之间的联系类型及特点，了解三种数据模型的特点及区别，为后面各章的学习打下基础。

1.1　信息、数据与数据处理

1.1.1　数据与信息

在数据处理中，我们最常用到的基本概念就是数据和信息，它们有着不同的含义。

1．信息（Information）

（1）信息的定义

在信息社会，信息是一种资源，它与能源、材料一起构成客观世界的三大要素。信息是关于现实世界事物的存在方式或运动状态反映的综合。信息是客观存在的，人类有意识地对信息进行采集并加工、传递，从而形成了各种消息、情报、指令、数据及信号等。

（2）信息的特征

信息的特征如下：

① 信息源于物质和能量。它不可能脱离物质而存在，信息的传递需要物质载体，信息的获取和传递要消耗能量。如信息可以通过报纸、电台、电视、计算机网络进行传递。

② 信息是可以感知的。人类对客观事物的感知，可以通过感觉器官，也可以通过各种仪器仪表和传感器等，不同的信息源有不同的感知形式。如报纸上刊登的信息通过视觉器官感知，电台中广播的信息通过听觉器官感知。

③ 信息是可存储、加工、传递和再生的。人们用大脑存储信息，叫做记忆。计算机存储器、录音、录像等技术的发展，进一步扩大了信息存储的范围。借助计算机，还可对收集到的信息进行整理。

2．数据（Data）

（1）数据的定义

数据是用来记录信息的可识别的符号，是信息的具体表现形式。

（2）数据的表现形式

可用多种不同的数据形式表示同一信息，而信息不随数据形式的不同而改变。如“ 2000 年硕士研究生将扩招 30% ”，其中的数据可改为汉字形式“ 两千年 ”、“ 百分之三十 ”。数据的概念在数据处理领域中已大大地拓宽了，其表现形式不仅包括数字和文字，还包括图形、图像和声音等。这些数据可以记录在纸上，也可记录在各种存储器中。

3．数据与信息的联系

数据是信息的符号表示或载体，信息则是数据的内涵，是对数据的语义解释。如上例中的数据“ 2000 ”、“ 30% ”被赋予了特定的语义，它们就具有了传递信息的功能。

1.1.2 数据处理

数据处理是将数据转换成信息的过程，包括对数据的收集、存储、加工、检索和传输等一系列活动。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，作为决策的依据。

我们可用下式简单地表示出信息、数据与数据处理的关系：

信息＝数据＋处理

数据是原料，是输入，而信息是产出，是输出结果。“ 信息处理 ”的真正含义应该是为了产生信息而处理数据。

1.2 数据库技术的产生、发展与研究领域

1.2.1 数据库技术的产生与发展

数据处理的中心问题是数据管理，数据管理是指对数据进行组织、分类、编码、存储、检索和维护。随着计算机硬件和软件的发展，数据管理经历了人工管理阶段、文件系统阶段和数据库系统阶段三个发展阶段。

1．人工管理阶段

在 20 世纪 50 年代中期以前，计算机主要用于科学计算，硬件中的外存只有卡片、纸带、磁带，没有磁盘等直接存取设备。软件只有汇编语言，没有操作系统和管理数据的软件。数据处理的方式基本上是批处理。

人工管理数据的特点有如下几个。

（1）数据不保存

因为当时计算机主要用于科学计算，对于数据保存的需求尚不迫切。在计算某一课题时将数据输入主存，运算处理后将结果输出，计算机在处理数据过程中不保存程序和数据。

（2）系统没有专用的软件对数据进行管理

每个应用程序都要包括数据的存储结构、存取方法和输入方式等，程序员编写应用程序时，还要安排数据的物理存储，因此，程序员的负担很重。

（3）数据不共享

数据是面向程序的，一组数据只能对应一个程序。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

数据的独立性是指逻辑独立性和物理独立性。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。

在人工管理阶段，程序与数据之间的关系可用图 1-1 表示。

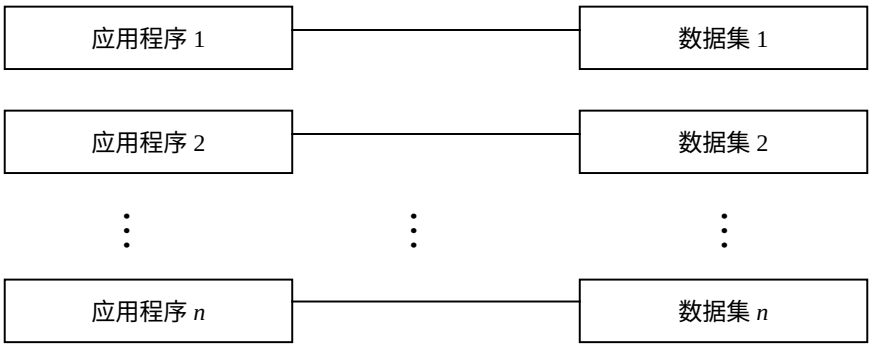


图 1-1　人工管理阶段

2．文件系统阶段

在 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，还大量用于信息管理；对大量的数据进行存储、检索和维护成为紧迫的需求。此时，硬件有了磁盘、磁鼓等直接存取设备；在软件方面，出现了高级语言和操作系统。操作系统中有了专门管理数据的软件，一般称为文件系统。处理方式有批处理，也有联机处理。

（1）特点

文件系统管理数据的特点有如下几个。

① 数据以文件形式长期保存。数据以文件的组织方式，保存在计算机的存储设备上，应用程序可对文件进行查询、修改和增删等处理。

② 文件系统可对数据的存取进行管理。程序员只与存储在存储设备上的文件的文件名打交道，不必关心数据的物理存储，大大减轻了程序员的负担。

③ 文件形式多样化。

文件有顺序文件、倒排文件和索引文件等，因而对文件的记录可顺序访问，也可随机访问，更便于存储和查找数据。

④ 程序与数据间有一定独立性。数据由专门的软件即文件系统进行管理，程序和数据间由软件提供的存取方法进行转换，数据存储发生变化不一定影响程序的运行。

在文件系统阶段，程序与数据之间的关系可用图 1-2 表示。

（2）问题

与人工管理阶段相比，文件系统阶段对数据的管理有了很大的进步，但一些根本性问题仍没有彻底解决，主要表现在以下三方面。

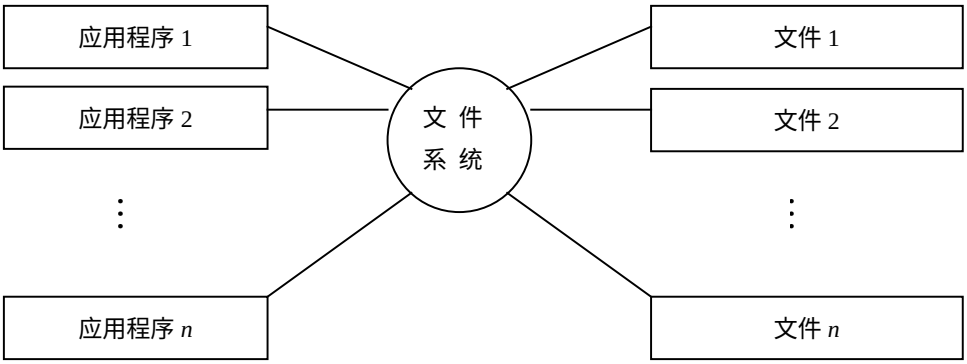


图 1-2　文件系统阶段

① 数据冗余度大。各数据文件之间没有有机的联系，一个文件基本上对应于一个应用程序，即文件仍然是面向应用的。当不同的应用程序所使用的数据具有共同部分时，也必须分别建立自己的数据文件，数据不能共享。

② 数据独立性差。文件是为某一特定的应用服务的，所以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。

③ 数据一致性差。由于相同数据的重复存储、各自管理，在对数据进行更新操作时，不但浪费磁盘空间，同时也容易造成数据的不一致性。

3．数据库系统阶段

从 20 世纪 60 年代后期开始，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。文件系统的数据管理方法已无法适应开发应用系统的的需要。于是为解决多用户、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。

数据库系统管理数据的特点有如下几个。

（1）数据共享性高、冗余低

这是数据库系统阶段的最大改进，数据不再面向某个应用程序而是面向整个系统，当前所有用户可同时存取库中的数据。这样便减少了不必要的数据冗余，节约了存储空间，同时也避免了数据之间的不相容性与不一致性。

（2）数据结构化

在数据库系统中，按照某种数据模型，将应用的各种数据组织到一个结构化的数据库中，系统不仅要考虑某个应用的数据结构，还要考虑整个应用的数据结构，这样，整个应用的数据不是一盘散沙，可表示出数据之间的有机关联。

例如：要建立学生成绩管理系统，系统包含学生（学号、姓名、性别、系别、年龄）、课程（课程号、

课程名)、成绩（学号、课程号、成绩）等数据，分别对应三个文件。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，要想查找某个学生的学号、姓名、所选课程的名称和成绩，必须编写一段比较复杂的程序来实现。而采用数据库方式，由于数据库系统不仅描述数据本身，还描述数据之间的联系，上述查询可以非常容易地联机查到。

（3）数据独立性高

数据的逻辑独立性是指当数据间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，在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时，均可确保数据的物理独立性。

数据的独立性是由系统提供的两个映象的功能保证的，关于二级映象，将在后一节讨论。

（4）有统一的数据控制功能

在数据库系统中，数据由数据库管理系统（Data Base Management System，DBMS）进行统一管理和控制。数据库可为多个用户和应用程序所共享，不同的应用需求，可以从整个数据库中选取所需要的数据子集。另外，对数据库中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四方面的数据控制功能。

① 数据的安全性（Security）控制：防止不合法使用数据造成数据的泄露和破坏，保证数据的安全。例如，系统提供口令检查用户身份或用其他手段来验证用户身份，以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（Integrity）控制：系统通过设置一些完整性规则，确保数据的正确性、有效性和相容性。

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 的正整数表示。

相容性是指表示同一事实的两个数据应相同，否则就不相容，如一个人不能有两个性别。

③ 并发（Concurrency）控制：多个用户同时存取或修改数据库时，防止相互干扰而提供给用户不正确的数据，并使数据库受到破坏。

④ 数据恢复（Recovery）：由于计算机系统的硬件故障、软件故障、操作员的误操作及其他故意的破坏等原因，造成数据库中的数据不正确或数据丢失时，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。

在数据库系统阶段，程序与数据之间的关系可用图 1-3 表示。

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。

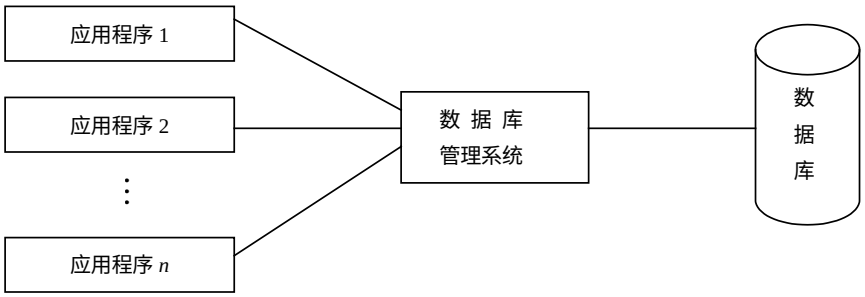


图 1-3　程序与数据之间的关系

1.2.2　数据库技术的研究领域

数据库技术经历了以上三个阶段的发展，现在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，数据库技术仍需不断向前发展。数据库学科的研究范围十分广泛，主要包括以下三个领域。

1．数据库管理系统软件的研制

DBMS 是数据库系统的基础。DBMS 的研制包括研制 DBMS 本身及以 DBMS 为核心的一组相互联系的软件系统，包括工具软件和中间件。研制的目标是提高系统的性能和提高用户的生产率。

2．数据库设计

数据库设计的研究范围包括数据库的设计方法、设计工具和设计理论的研究，数据模型和数据建模的研究，计算机辅助数据库设计及其软件系统的研究，数据库设计规范和标准的研究等。数据库设计的目的是在 DBMS 的支持下，按照应用的要求，为某一部门或组织设计一个结构合理、使用方便、效率较高的数据库及其应用系统。

3．数据库理论

数据库理论的研究主要集中于关系规范化理论、关系数据理论等。近年来，随着人工智能与数据库理论的结合以及并行计算技术的发展，数据库逻辑演绎和知识推理、并行算法等都成为新的研究方向。

随着数据库应用领域的不断扩展，计算机技术的迅猛发展，数据库技术与人工智能技术、网络通信技术、并行计算技术等相互渗透、相互结合，使数据库技术不断涌现新的研究方向。

本书所介绍的数据库系统的基本概念、基本技术和基本知识都是进行上述三个领域研究和开发的基础。

1.3　数据库系统的组成和结构

1.3.1　数据库系统的组成

数据库系统是指在计算机系统中引入数据库后的系统构成。整个数据库系统，主要由数据库、数据库用户、计算机硬件系统、计算机软件系统等几部分组成。

需要全本请访问 www.ertongbook.com