北京市高等教育精品教材立项项目

HZ BOOKS
华章教育

普通高等教育精品教材·省级

数据库基础教程

王月海 何丽 孟丹 张艳苏 编著

Fundamentals of Database Systems

机械工业出版社
China Machine Press

北京市高等教育精品教材立项项目

普通高等教育精品教材·省级

数据库基础教程

王月海 何丽 孟丹 张艳苏 编著

Fundamentals of Database Systems

机械工业出版社
China Machine Press

本书以培养应用型软件人才为目标，结合编者多年的教学经验和项目开发经验而编写。

本书分为三部分，第一部分是数据处理和数据库基础理论，主要介绍数据处理技术发展、数据库模型、标准SQL语言、规范化理论、数据库设计、数据库安全与保护；第二部分是数据库应用，主要介绍在SQL Server 2008环境下如何完成数据库的创建、数据库的操作、存储过程与触发器、数据库连接技术、综合数据库应用系统开发案例；第三部分是数据库技术的发展及展望，主要介绍面向对象数据库系统、分布式数据库系统、多媒体数据库系统、移动数据库系统以及数据仓库与数据挖掘。

本书取材新颖，将数据库原理知识与实际数据库开发工具结合在一起，旨在培养读者的综合实践与创新能力。

本书可作为高等院校计算机应用专业以及信息管理等相关专业的教材或参考书，也可作为相关人员学习数据库知识的参考书。

封底无防伪标均为盗版
版权所有，侵权必究
本书法律顾问　北京市展达律师事务所

图书在版编目（CIP）数据

数据库基础教程／王月海等编著．—北京：机械工业出版社，2011.8
（北京市高等教育精品教材立项项目）

ISBN 978-7-111-35592-2

Ⅰ．数…　Ⅱ．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（2011）第157073号

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：李　荣
北京瑞德印刷有限公司印刷
2011年8月第1版第1次印刷
185mm×260mm·17.5印张
标准书号：ISBN 978-7-111-35592-2
定价：30.00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
购书热线：（010）68326294；88379649；68995259
投稿热线：（010）88379604
读者信箱：hzjsj@hzbook.com

前言

数据库技术主要研究如何存储、使用和管理数据。它自20世纪60年代产生至今已有40多年的历史，出现了C. W. Bachman、E. F. Codd和James Gray三位图灵奖获得者，带动了一个巨大的软件产业的发展。数据库技术一直是最活跃、发展速度最快、应用最广的IT技术之一。

在过去20年里，数据库技术的应用有了巨大的增长，几乎每个行业都要用数据库来存储、操纵、检索数据。例如，在商业、医疗保健、教育、政府组织、图书馆、军事、工业控制等领域都有数据库的应用，它已成为信息管理、电子商务、网络服务等应用系统的核心技术和重要基础。从某种意义上说，数据库技术已成为计算机、控制、信息等相关专业的工程技术人员所必须具备的专业知识。

全书共12章，分为三部分。

第一部分是数据处理和数据库基础理论，包括第1～7章，主要介绍数据处理技术发展、数据库模型、标准SQL语言、规范化理论、数据库设计步骤、数据库安全与保护措施；第二部分是数据库应用，包括第8～11章，主要介绍在SQL Server 2008环境下如何完成数据库的创建及对数据库的操作、T-SQL语言、存储过程与触发器、常用数据库访问技术，在第11章中以开发一个学生公寓管理系统为例，介绍如何进行需求分析、概念结构设计、逻辑结构设计、物理结构设计及数据库应用系统的开发等数据库设计步骤，并用C#+SQL Server 2008实现这个实际的系统；第三部分是数据库技术的发展及展望，主要介绍面向对象数据库系统、分布式数据库系统、多媒体数据库系统、移动数据库系统以及数据仓库与数据挖掘。

本书采用理论与应用相结合的写作思路，自始至终贯彻案例教学的思想，使读者清晰地认识到理论和应用各自解决的问题，在内容编排上由理论到实践、从技术基础到具体开发应用。本书取材新颖，书中的案例采用流行的C#开发技术和最新的SQL Server 2008数据库管理系统，将数据库原理知识与实际数据库开发工具结合在一起，旨在培养读者的综合实践与创新能力。

本书在编写过程中，参考了大量的相关技术资料和程序开发源码资料，在此向资料的作者深表谢意，书中的全部程序都已经上机调试通过。由于编者水平和时间有限，书中难免有错误及不妥之处，敬请各位同行和读者不吝赐教，批评指正。

编　者

2011年6月

教学建议

教学内容	教学目标及教学要求参考	课时安排参考	可选实验参考
第1章 数据处理概述	• 熟悉数据、信息等基本概念 • 理解数据处理与数据管理的区别 • 了解数据处理发展的三个阶段 • 掌握数据逻辑结构及数据物理结构	2	
第2章 数据库基础知识	• 了解数据库技术的发展史 • 掌握数据库、数据库管理系统、数据库系统等基本概念 • 理解常用的数据模型 • 掌握数据库系统的三级模式结构 • 理解 DBMS 的工作过程	4	
第3章 关系模型	• 掌握关系模型的三要素 • 掌握关系模型的完整性约束 • 掌握关系代数表达式的运用 • 理解关系代数的查询优化方法	4	安装某个数据库管理系统（SQL Server 2008）。了解 RDBMS 的系统架构、系统服务停止及启用
第4章 关系数据库标准语言——SQL	• 了解 SQL 语言的发展及标准化 • 理解并掌握数据定义、数据查询、数据更新等 SQL 语言的使用 • 掌握视图的定义及使用方法 • 理解数据控制中权限及角色的授予及收回	4～6	在 RDBMS 中建立一个数据库，在查询窗口中进行数据定义、数据查询及各类更新的 SQL 操作
第5章 关系数据库规范化理论	• 理解非规范化的关系模式所存在的异常 • 理解并掌握函数依赖的基本概念 • 掌握 1NF、2NF、3NF 的判定标准 • 理解模式分解的一般方法	4	
第6章 关系数据库设计	• 掌握数据库设计的基本步骤 • 了解需求分析的任务及方法 • 掌握概念结构设计中 E-R 图的画法 • 掌握 E-R 图向关系模型转化的方法 • 理解物理结构设计阶段的任务	4～6	根据一个给定的数据库应用系统的需求，完成数据库设计各个阶段的详细设计报告
第7章 数据库的安全与保护	• 熟悉数据库安全性控制的一般方法 • 掌握 SQL Server 安全性控制的方法 • 理解数据库完整性的意义 • 掌握事务基本概念 • 理解实现并发控制的基本方法 • 掌握 SQL Server 中数据库备份与恢复的方法	4	在 RDBMS 中演练登录用户和数据库用户角色的分配、数据库备份及恢复的基本操作
第8章 SQL Server 2008 数据库系统基础	• 了解 SQL Server 的不同版本 • 熟悉 SQL Server 的主要管理工具 • 掌握在对象资源管理器中进行数据库操作	2	编写存储过程、触发器等 SQL 程序

试读结束，需要全本PDF请购买 www.ertongbook.com

（续）

教学内容	教学目标及教学要求参考	课时安排参考	可选实验参考
第9章 SQL Server 2008 高级应用	• 掌握 T-SQL 的基本语法、流程控制语句 • 掌握存储过程的定义及使用 • 掌握触发器的定义及使用	4～6	
第10章 SQL Server 数据库访问技术	• 熟悉常用的几种数据库访问技术 • 掌握 ODBC 数据源的配置方法 • 掌握 ADO.NET 对象模型 • 熟悉 JDBC 访问数据库的基本过程	4	编写通过 ADO.NET 连接数据库的 C/S 结构的应用程序
第11章 SQL Server 数据库应用系统开发	• 熟悉 .NET 开发环境 • 掌握数据库应用系统开发的各个步骤 • 掌握在 .NET 平台下用 C#实现通用数据库连接类的编写 • 了解水晶报表的使用	2	编写通过 ADO.NET 连接数据库的 B/S 结构的应用程序
第12章 数据库技术的发展	• 了解并熟悉面向对象数据库、分布式数据库、多媒体数据库、移动数据库、数据仓库与数据挖掘等概念	2	
教学总学时建议		40～46	16～20

相关图书

书　名	书号（ISBN）	作　者	译者	出版日期	价格
数据库基础案例教程与实验指导	978-7-111-32156-9	张巨俭		2011	35.00
Access数据库应用教程	978-7-111-33023-3	朱翠娥		2011	29.80
分布式数据库系统原理与应用	978-7-111-34524-4	申德荣 等		2011	35.00
数据库原理与应用教程 第3版	7-111-31204-8	何玉洁		2010	29.80
数据库原理与应用	7-111-19871-9	何玉洁		2006	32.00
数据库技术原理与应用教程	7-111-22945-2	徐洁磐 常本勤		2008	29.00
数据库技术原理与应用教程学习与实验指导	7-111-29126-8	常本勤 徐洁磐		2010	19.80
数据库系统实现 第2版	978-7-111-30287-2	Hector Garcia-Molina	杨冬青 等	2010	59.00
数据库系统导论（原书第8版）	7-111-21333-8	C. J. Date	孟小峰 王珊 等	2007	75.00
数据库系统概念（原书第5版·本科教学版）	7-111-23422-7	Abraham Silberschatz ; Henry F. Korth ; S. Sudarshan	杨冬青 马秀莉 等	2008	45.00
数据库系统概念、设计及应用	7-111-27958-7	S. K Singh	何玉洁	2009	89.00
数据库系统基础教程（英文版·第3版）	7-111-24733-3	Jeffrey D. Ullman; Jennifer Wdom		2008	45.00
数据库系统基础教程（原书第三版）	7-111-26828-4	Jeffrey D. Ullman; Jennifer Wdom	岳丽华 金培权 等	2009	45.00
数据库系统实现 （英文版 第2版）	7-111-28860-2	（美） Jeffrey D Ullman		2009	55.00
数据库应用技术 SQL Server 2005提高篇	7-111-23518-7	张蒲生		2008	29.00
数据库应用技术SQL Server 2005基础篇	7-111-22791-5	张蒲生		2008	26.00
网络数据库技术应用	7-111-24609-1	周玲艳 张希		2008	25.00
SQL Server2005数据库管理与开发实用教程	7-111-286684	李丹		2009	29.00
数据库技术应用教程	7-111-20741-2	何宁 黄文斌 熊建强		2007	29.00
数据库技术应用实验教程	7-111-20953-9	何宁 滕冲 熊素萍		2007	19.00
Visual FoxPro数据库管理系统教程	7-111-22967-4	程玮　陆晶		2008	26.00
Visual FoxPro数据库与程序设计教程	7-111-20561-6	张莹		2007	28.00

目 录

前言

教学建议

第一部分
数据处理和数据库基础理论

第1章 数据处理概述 …… 2

1.1 信息与数据 …… 2

1.1.1 什么是信息 …… 2

1.1.2 什么是数据 …… 3

1.1.3 数据与信息的联系及区别 …… 3

1.2 数据处理 …… 4

1.2.1 数据为什么需要处理 …… 4

1.2.2 数据处理的方法及过程 …… 4

1.2.3 数据处理要解决哪些问题 …… 5

1.2.4 数据处理与数据管理 …… 5

1.3 数据处理技术的发展 …… 5

1.3.1 人工管理阶段 …… 5

1.3.2 文件系统阶段 …… 6

1.3.3 数据库系统阶段 …… 7

1.4 用数据库技术管理数据的优点 …… 9

1.5 小结 …… 10

习题 …… 10

第2章 数据库基础知识 …… 11

2.1 数据库的定义 …… 11

2.2 数据库技术的发展史 …… 12

2.3 数据库系统结构 …… 14

2.3.1 数据库系统的三级模式结构 …… 14

2.3.2 两级模式映射及数据独立性 …… 15

2.3.3 数据库系统的外部体系结构 …… 16

2.4 数据库系统的组成 …… 18

2.5 数据模型 …… 19

2.5.1 数据模型的三要素 …… 20

2.5.2 概念模型 …… 21

2.5.3 常用的逻辑数据模型 …… 23

2.5.4 层次模型 …… 24

2.5.5 网状模型 …… 27

2.5.6 关系模型 …… 29

2.6 数据库系统的核心——DBMS …… 30

2.6.1 什么是DBMS …… 31

2.6.2 DBMS的主要功能 …… 31

2.6.3 DBMS的工作过程 …… 31

2.7 小结 …… 32

习题 …… 33

第3章 关系模型 …… 34

3.1 关系模型的基本概念 …… 34

3.1.1 关系的通俗解释 …… 34

3.1.2 关系的形式化定义 …… 34

3.1.3 关系模式 …… 36

3.1.4 关系数据库 …… 37

3.2 关系模型的完整性 …… 38

3.2.1 实体完整性 …… 38

3.2.2 参照完整性 …… 38

3.2.3 用户自定义完整性 …… 39

3.3 关系操作 …… 40

3.3.1 基本关系操作 …… 40

3.3.2 关系数据语言的分类 …… 40

3.3.3 关系代数概述 …… 41

3.3.4 传统的集合运算 …… 41

3.3.5 专门的关系运算 …… 43

3.3.6 关系运算表达式应用实例 …… 47

3.4 关系数据库的查询优化 …… 48
3.4.1 查询优化问题的提出 …… 48
3.4.2 查询优化的必要性 …… 49
3.4.3 查询优化的一般策略 …… 50
3.4.4 关系代数表达式的等价变换规则 … 51
3.4.5 关系代数表达式的优化算法 …… 52
3.5 小结 …… 54
习题 …… 55
第4章 关系数据库标准语言——SQL … 56
4.1 SQL的基本概念及特点 …… 56
4.1.1 SQL的产生及标准化 …… 56
4.1.2 SQL语言的基本概念 …… 56
4.1.3 SQL语言的特点 …… 57
4.1.4 SQL语言的组成 …… 58
4.2 SQL的数据定义 …… 58
4.2.1 基本表的创建、修改和删除 …… 58
4.2.2 索引的创建和删除 …… 61
4.3 SQL的数据查询 …… 62
4.3.1 单表查询 …… 64
4.3.2 多表连接查询 …… 69
4.3.3 嵌套查询 …… 72
4.3.4 集合查询 …… 77
4.4 SQL的数据更新 …… 78
4.4.1 数据插入 …… 78
4.4.2 数据修改 …… 79
4.4.3 数据删除 …… 80
4.5 视图 …… 80
4.5.1 视图的创建和撤销 …… 81
4.5.2 视图的数据操作 …… 82
4.5.3 视图的优点 …… 83
4.6 SQL的数据控制 …… 84
4.6.1 数据控制简介 …… 84
4.6.2 权限与角色 …… 84
4.6.3 系统权限与角色的授予和收回 … 84
4.6.4 对象权限与角色的授予和收回 … 85
4.7 小结 …… 86
习题 …… 86
第5章 关系数据库规范化理论 …… 88
5.1 数据依赖 …… 88
5.1.1 问题的提出 …… 88
5.1.2 函数依赖的基本概念 …… 89
5.1.3 候选码 …… 91
5.2 关系模式的规范化 …… 91
5.2.1 关系与范式 …… 91
5.2.2 第一范式（1NF） …… 92
5.2.3 第二范式（2NF） …… 92
5.2.4 第三范式（3NF） …… 93
5.2.5 BCNF …… 94
5.2.6 规范化小结 …… 95
5.3 模式分解 …… 95
5.4 小结 …… 96
习题 …… 97
第6章 关系数据库设计 …… 98
6.1 数据库设计概述 …… 98
6.1.1 数据库设计的任务、内容和特点 … 98
6.1.2 数据库设计的方法 …… 99
6.1.3 数据库设计的步骤 …… 99
6.2 系统需求分析 …… 101
6.2.1 需求分析的任务 …… 101
6.2.2 需求分析的方法 …… 102
6.2.3 数据流图和数据字典 …… 102
6.3 概念结构设计 …… 105
6.3.1 概念结构设计的特点 …… 105
6.3.2 概念模型设计的方法与步骤 …… 105
6.3.3 数据抽象与局部E-R图设计 …… 106
6.3.4 视图的集成 …… 109
6.4 逻辑结构设计 …… 110
6.4.1 E-R图向关系模型的转化 …… 111
6.4.2 数据模型的优化 …… 113
6.4.3 设计用户子模式 …… 113
6.5 数据库物理设计 …… 114
6.5.1 确定物理结构 …… 114
6.5.2 评价物理结构 …… 116
6.6 数据库实施 …… 116

6.7 数据库的运行与维护 …… 118
6.8 小结 …… 119
习题 …… 119
第7章 数据库的安全与保护 …… 120
7.1 数据库的安全性 …… 120
7.1.1 数据库安全性概述 …… 120
7.1.2 安全性控制的一般方法 …… 120
7.1.3 SQL Server 的安全性控制 …… 123
7.2 数据库的完整性 …… 127
7.2.1 数据库完整性的含义 …… 127
7.2.2 完整性约束条件 …… 128
7.2.3 完整性控制 …… 129
7.2.4 SQL Server 的完整性控制 …… 130
7.3 数据库的并发控制 …… 131
7.3.1 事务 …… 131
7.3.2 并发操作与数据的不一致性 …… 132
7.3.3 封锁 …… 134
7.3.4 活锁和死锁 …… 137
7.3.5 SQL Server 中的并发控制技术 … 138
7.4 数据库的备份及恢复 …… 141
7.4.1 数据库恢复概述 …… 141
7.4.2 数据库恢复的基本原理及其实现技术 …… 141
7.4.3 数据库的故障及恢复策略 …… 144
7.4.4 SQL Server 数据库备份与恢复技术 …… 146
7.5 小结 …… 149
习题 …… 150

第二部分 数据库应用

第8章 SQL Server 2008 数据库系统基础 …… 152
8.1 SQL Server 2008 版本分类及安装要求 …… 152
8.1.1 SQL Server 2008 的不同版本 …… 152
8.1.2 SQL Server 2008 的安装要求 …… 153
8.2 SQL Server 2008 体系结构 …… 153
8.3 SQL Server 2008 主要管理工具 …… 155
8.3.1 SQL Server 集成管理器 …… 155
8.3.2 SQL Server 配置管理器 …… 157
8.3.3 分析服务 …… 157
8.3.4 数据库引擎优化顾问 …… 158
8.3.5 业务智能开发工具 …… 158
8.3.6 事件探查器 …… 158
8.3.7 SQL Server 文档和教程 …… 159
8.4 SQL Server 2008 数据库管理 …… 159
8.4.1 SQL Server 2008 的系统数据库 … 159
8.4.2 示例数据库 …… 160
8.4.3 数据库的文件与文件组 …… 160
8.4.4 数据库操作 …… 161
8.4.5 数据表操作 …… 165
8.4.6 数据操纵 …… 167
8.5 小结 …… 168
习题 …… 168
第9章 SQL Server 2008 高级应用 …… 169
9.1 T-SQL 语言基础 …… 169
9.1.1 T-SQL 语法约定 …… 169
9.1.2 T-SQL 数据类型 …… 170
9.1.3 变量 …… 172
9.1.4 运算符 …… 173
9.1.5 批处理 …… 174
9.1.6 流程控制语句 …… 175
9.1.7 函数 …… 180
9.2 存储过程 …… 184
9.2.1 存储过程概述 …… 184
9.2.2 存储过程的分类 …… 185
9.2.3 创建存储过程 …… 185
9.2.4 查看存储过程 …… 188
9.2.5 删除存储过程 …… 188
9.2.6 执行存储过程 …… 189
9.2.7 修改存储过程 …… 189
9.3 触发器 …… 190
9.3.1 触发器概述 …… 190
9.3.2 触发器的工作原理 …… 191

9.3.3 创建触发器 …… 193
9.3.4 查看触发器 …… 195
9.3.5 修改触发器 …… 196
9.3.6 删除触发器 …… 197
9.4 小结 …… 197
习题 …… 198
第 10 章 SQL Server 数据库访问技术 … 199
10.1 数据库访问技术概述 …… 199
10.2 ODBC 技术 …… 200
10.2.1 ODBC 概述 …… 201
10.2.2 ODBC 体系结构 …… 201
10.2.3 配置 ODBC 数据源 …… 202
10.3 ADO 和 ADO.NET …… 204
10.3.1 OLEDB …… 204
10.3.2 ADO …… 205
10.3.3 ADO.NET …… 206
10.4 JDBC 技术 …… 209
10.4.1 JDBC 简介 …… 209
10.4.2 JDBC 的基本结构 …… 210
10.4.3 使用 JDBC 访问数据库 …… 211
10.5 小结 …… 212
习题 …… 212
第 11 章 SQL Server 数据库应用系统开发 …… 213
11.1 系统需求 …… 213
11.2 系统功能设计 …… 214
11.3 数据库设计 …… 214
11.3.1 概念结构设计 …… 215
11.3.2 逻辑结构设计 …… 215
11.3.3 物理结构设计 …… 215
11.3.4 创建数据库 …… 217
11.4 系统实现 …… 217
11.4.1 C#语言 …… 217
11.4.2 创建项目 …… 217
11.4.3 通用连接数据库技术的实现 … 218
11.4.4 主窗体界面设计 …… 219
11.4.5 用户登录模块 …… 220
11.4.6 权限管理模块 …… 220
11.4.7 公寓管理员管理模块 …… 222
11.4.8 公寓基本信息管理模块 …… 225
11.4.9 来访人员管理模块 …… 227
11.4.10 查询模块 …… 229
11.4.11 报表打印模块 …… 232
11.5 小结 …… 234
习题 …… 234

第三部分
数据库技术的发展及展望

第 12 章 数据库技术的发展 …… 236
12.1 数据库技术概述 …… 236
12.1.1 新一代数据库系统的特点 …… 236
12.1.2 第三代数据库系统应具备的三个基本特征 …… 237
12.2 面向对象数据库系统 …… 237
12.2.1 面向对象数据库系统概述 …… 237
12.2.2 面向对象数据库系统的功能要求 …… 238
12.2.3 面向对象的基本概念 …… 238
12.2.4 面向对象数据库系统的应用 … 239
12.2.5 对象关系数据库系统 …… 240
12.2.6 RDBMS、ORDBMS 和 OODBMS 的比较 …… 241
12.3 分布式数据库系统 …… 241
12.3.1 分布式数据库系统概述 …… 241
12.3.2 分布式数据库系统的概念及特点 …… 242
12.3.3 分布式数据库系统的体系结构 … 243
12.3.4 分布式事务处理 …… 245
12.4 多媒体数据库系统 …… 247
12.4.1 多媒体数据库系统概述 …… 247
12.4.2 多媒体数据库系统的实现方法 … 248
12.4.3 多媒体数据库系统的主要技术 … 249
12.4.4 多媒体数据库的发展 …… 249
12.5 移动数据库系统 …… 250

试读结束，需要全本PDF请购买 www.ertongbook.com

12.5.1 移动数据库系统概述 …… 250
12.5.2 移动数据库系统的特点及体系结构 …… 250
12.5.3 移动数据库系统的关键技术 … 252
12.5.4 移动事务处理 …… 253
12.5.5 移动数据库的发展 …… 255
12.6 数据仓库与数据挖掘技术 …… 256
12.6.1 数据仓库的产生 …… 256
12.6.2 数据仓库的概念及体系结构 … 257
12.6.3 联机分析处理（OLAP） …… 260
12.6.4 数据挖掘技术 …… 263
12.7 小结 …… 265
习题 …… 266
参考文献 …… 267

第一部分

数据处理和数据库基础理论

- 第1章　数据处理概述
- 第2章　数据库基础知识
- 第3章　关系模型
- 第4章　关系数据库标准语言——SQL
- 第5章　关系数据库规范化理论
- 第6章　关系数据库设计
- 第7章　数据库的安全与保护

第1章 Chapter

数据处理概述

没有人会否认我们正处在一个“信息化”的时代。随着信息化在全球的快速进展，世界对信息的需求快速增长，信息资源已经成为社会各行各业的重要资源和财富。信息技术（Information Technology，IT）已成为支撑当今经济活动和社会生活的基石。

实际上，信息系统和信息处理从人类文明产生开始就已存在，直到电子计算机问世、信息技术实现飞跃以及现代社会对信息需求日益增长，才迅速发展起来。从第一台电子计算机于1946年问世，信息系统经历了由单机到网络，由低级到高级，由电子数据处理到管理信息系统、再到决策支持系统，由数据处理到智能处理的过程。

本章主要介绍信息与数据、数据处理的一些基础知识和数据处理技术发展的几个主要阶段。读者从中可以了解到为什么要使用数据库技术以及数据库技术的重要性。

1.1 信息与数据

1.1.1 什么是信息

信息是一种重要的战略资源，现在各行各业的人们都在越来越多地谈论各种信息。信息已经被列入了社会发展的三大科学支柱（材料、能源、信息）之一，成为现代社会文明和科学发展的重要标志，因此需要通过各种手段充分挖掘其潜在价值。那么，信息的概念到底是什么呢？

信息是一个抽象概念，对于它的定义有很多种。从计算机管理理论的角度来说，可以作如下定义：信息是人们的头脑中对现实世界各种事物的抽象反映，它是反映客观世界里各种事务特征和变化的知识。例如，一年中天气的阴晴雨雪的总数，火车开行的车次、车速，空间卫星的运行轨迹及环绕周期等，都可以称作信息。

按照不同的应用领域，信息可以划分为经济信息、社会信息、科技信息、军事信息等不同的种类。各类信息都具有相同的特点，如下所示：

1）信息是可以识别的。这种识别既可以是由人的感官实现的直接识别，也可以是通过各种探测手段和工具进行的间接识别。例如，对于物体的大小、形状的认识仅通过肉眼就能实现，但是河水的流量、金属的硬度等信息的获取就必须通过特殊的手段间接完成。

2）信息是可以转换的。信息可以以多种不同的形式存在，并且能够在不同的形式之间进行转换。例如，具体的物质信息可以转换成语言、文字、图像等形式，而语言文字也可以转换成广播电视中的电信号或计算机里的各种编码。

3）信息是可以存储的。人对事物的记忆就是大脑对信息的存储。而计算机可以把大量的信息存储在磁盘、磁带、光盘或半导体芯片等多种存储介质中，长期保存，反复使用。这些都是人脑无法实现的。

4）信息是可以处理的。人类的思维活动就是人脑对于信息进行处理的过程。而计算机则要通过由人编制的各种系统软件来完成对信息的处理。

5）信息是可以传递的。人们在日常生活中通过语言、动作、表情进行的相互交流就是信息的传递。在现代电子世界中，信息可以通过代码的形式在电报、电话、光纤、计算机等网络中快捷地传递和交换，实现信息资源的充分共享。

6）信息是可以再生的。存在于人脑中的信息可以通过语言、文字、图表等形式再生成。同样，存储在计算机里的信息也可以通过显示、打印等方式实现再生。

1.1.2　什么是数据

任何具体事物，都要通过信息来反映和认识。但信息作为一种抽象的反映，不可能直接被计算机所接受。处理任何信息，都要通过数据来完成。数据可以用来表示信息，而事物的客观状态又可以由信息来反映。

因此，数据（Data，又称资料）是对客观事物的性质、状态以及相互关系等进行记载的物理符号或这些物理符号的组合。它是可识别的、抽象的符号。这些符号不仅指数字，而且包括字符、文字、图形等。其表现形式如表 1-1 所示。

表 1-1　数据的类型及表现形式

数据类型	表现形式
数值数据	数、字母和其他符号
图形数据	图形或图片
声音数据	声音、噪音或音调
视觉数据	动画或图片
模糊数据	高、矮、胖、瘦等

从表 1-1 中可以看出：数值数据使得客观世界严谨有序，而其他类型的数据使得客观世界丰富多彩。

为了满足计算机处理的需要，人们使用人为规定的数字来表示形形色色的信息。可以这么说，世界上的一切信息都可以用数字来表示，正因为有了数字的表示，计算机才有可能处理现实世界的各种事物。事实上，所有数据在计算机中都是以若干个二进制位代码的形式表示的。但是用数字表示信息并不是因为计算机应用的需要才出现的，早在计算机问世之前，邮电通信中的电报密码就已经使信息数字化了，因为在绝大多数情况下，只有用数字表示信息才是最准确的。例如，我们说某大学很大或很小，并不会给人留下很深的印象，可是如果说某大学共有在校学生 10 000 余人或 800 人，就容易给人留下较深的印象。

1.1.3　数据与信息的联系及区别

信息和数据之间存在着内容和形式上的联系，数据是用来负载信息的物理符号（包括数字、字母和符号），是表达和传播、交换信息所必需的工具；而信息则是对数据的解释，信息要通过数据的形式来表现出来，它只能依靠数据而存在，不可能独立出来。二者之间也有比较明显的不同，数据本身只是一些可以识别的符号，并不具有任何实际意义，只有对数据赋予了某些具体含义之后，数据才能成为信息，即信息是经过加工处理后对客观世界产生影响的数据。另外，信息是更基本的直接反映现实的概念，而数据则是信息的具体表现。同样的信息可以由不同的数据形式表现出来，但是它们所表示的信息内容却不会改变。所以信息不随着负载它的物理设备的改变而改变，而数据则不同，它在计算机化的系统中的表示往往和计算机相关。例如，“今天是星期五。”和“Today is Friday.”就是用两种不同的数据形式表现出来的同样的信息。

1.2 数据处理

1.2.1 数据为什么需要处理

在日常生活中，每天都会有大量的关于不同问题的数据产生出来。由于这些数据涉及不同的领域，而数据之间也没有必要的联系，所以显得毫无规律，杂乱无章。从这些表面化的数据中很难提炼出有价值的信息，而且信息的获取过程也没有特定的规律，这对需要从海量数据中挖掘潜在的有价值信息的工作造成极大的挑战。

为了从获得的大量数据中找出对我们有价值的信息并加以利用，在对具体的数据进行收集、汇总之后，需要通过必要的手段对某些数据按照某种规律进行转化和必要的计算，使其更能反映事物的本质特征，最后再通过对这些数据的分析就可以得出有用的数据。这种通过对具体数据的收集、转化、汇总、分析、计算等处理过程，将大量的表面数据进行简化，从中提炼出能够反映事物本质和内在联系的有价值的数据的过程就是数据处理过程。这些提炼后的数据能够为人类的生产、经营等社会活动提供强有力的参考。

例如，对于购物中心等商业企业，每天都要面对商品、销售、收入、支出等各种信息，产生成千上万的具体数据。通过对这些具体数据进行收集整理，分析计算，统计输出等处理，就可以得到准确反映企业经营、财务、商品情况的各种报表。经过处理后的总销售额、某种商品的销售数据及库存量、总的销售收入和销售利润、税金等数据，将为企业制定销售策略，决定各种商品的进货渠道和数量，根据当前的财务状况确定财务计划等各种经营活动提供准确可靠的依据。

1.2.2 数据处理的方法及过程

人类社会产生时，就出现了原始的数据处理。直到现在，数据处理已经经历了手工、机械和电子化数据处理三个阶段。现在绝大多数数据都是通过计算机来处理的。一般的电子化数据处理包括以下几个过程：

1）数据的收集：根据用户的需要，及时地从产生数据的各个环节收集所有原始数据。有时还要对收集到的数据进行必要的分类和校验，以保证原始数据的正确性、完整性和时效性。

2）数据的转换：为了使各种原始数据能够被计算机承认并接受，需要用格式变换或代码化的处理方法进行数据转换。对于某些与计算机所要求的格式有差异的数据，可以通过对其数据格式进行转换来达到要求。而对于一些文字信息则可以用比较简单的代码来表示。例如，可以用“T”来代表老师，而用“S”代表学生等。

3）数据的组织：对数据按照其内部联系和计算机软件的要求进行整理和编排，使数据在计算机中占用尽量少的内存和其他资源，并能够较快地被软件调用。

4）数据的输入：将经过数据转换并已组织好的数据，按照设定的格式输入到计算机里。

5）数据的处理：利用各种计算机应用软件，对输入的数据进行各种处理，包括索引、排序、统计、计算、修改、更新等操作，从而得到用户需要的结果。

6）数据的输出：将数据处理的结果，按照用户要求的格式从计算机中输出，提供给用户。输出的结果可以是简单的数据，也可以是报表、图形等其他的形式。

7）数据的存储：所有的相关数据，包括输入的原始数据、处理后的结果数据以及处理过程中的中间数据都会被计算机记录到磁盘、磁带等存储介质上，以备今后继续使用。

1.2.3 数据处理要解决哪些问题

通过以上分析可以看到，通过数据处理可以解决下列几个方面的问题：

首先，把收集到的各种原始数据经过分类整理和格式转换变换成为易于观察、分析，并且可以进行进一步处理的有规律的数据。

其次，把大量的具体数据经过加工变为可以反映某种事物本质的、比较精炼的数据。只有这种数据才能够对人类的决策和行动产生影响，因而这也是数据处理的关键。

第三，要把已有的数据进行存储，以备今后继续使用。

1.2.4 数据处理与数据管理

数据处理和数据管理是两个密切联系但又相互区别的概念。如上所述，所谓数据处理就是从已有数据出发，经过适当加工处理得到新的所需要的数据的过程。数据加工处理一般分为数据计算和数据管理两部分。数据计算相对简单，但数据管理却比较复杂，是数据处理过程的主要内容与核心部分，因而数据处理在本质上就可以看做是数据管理。一般认为，数据管理主要是指数据的收集、整理、组织、存储、维护、检索和传送等操作，这些操作都是数据处理业务中重要的和必不可少的基本环节。为此，数据管理是指对数据进行分类、组织、编码、存储、检索和维护的管理活动的总称。就用计算机来管理数据而言，数据管理是指数据在计算机内的一系列活动的总和。在不少文献和著作中，数据处理和数据管理是两个可以替代使用的概念。

1.3 数据处理技术的发展

在计算机的三大主要应用领域（科学计算、数据处理和辅助设计）中，数据处理是计算机应用的主要方面之一。

我们把通过计算机进行的数据处理称为电子数据处理（Electronic Data Processing，EDP）。

在利用计算机进行数据处理的发展历程中，EDP 技术经历了人工管理、文件系统和数据库系统三个阶段。

1.3.1 人工管理阶段

人工管理阶段是从20世纪40年代中期电子计算机问世到20世纪50年代中期，这一阶段计算机主要用于科学计算。从硬件上看，外存只有磁带、卡片、纸带，速度低、内存小，没有磁盘等直接存取的存储设备；从软件上看，没有操作系统，没有管理数据的软件，数据处理方式是批处理。

在人工管理阶段，数据管理的特点是：

1）数据不保存在机器中。因为计算机主要应用于科学计算，一般不需要将数据长期保存。在计算时将数据输入，计算完毕将数据输出。

2）没有软件系统对数据进行管理。程序员不仅要规定数据的逻辑结构，而且还要在程序中设计物理结构，包括存储结构、存取方法、输入/输出方式等。因此程序中存取数据的

试读结束，需要全本PDF请购买 www.ertongbook.com