

高校教材

现代数据库系统

XIANDAI
SHUJUKU XITONG
SHIYONG JIAOCHENG

实用教程

◎ 徐洁磐 张 剡 封 玲 编著

 人民邮电出版社
POSTS & TELECOM PRESS

现代数据库系统实用教程

徐洁磐 张 剌 封 玲 编著

人民邮电出版社

图书在版编目 (CIP) 数据

现代数据库系统实用教程 / 徐洁磐, 张劼, 封玲编著. —北京: 人民邮电出版社, 2006.10
ISBN 7-115-14747-7

I. 现... II. ①徐...②张...③封... III. 数据库系统—教材 IV. TP311.13

中国版本图书馆 CIP 数据核字 (2006) 第 082970 号

内 容 提 要

本书是一部关于现代数据库系统的基本原理与技术的教科书, 本书突出了数据库和新技术的应用以及教学的需要。在编写中注重原理与技能的并重, 传统内容与先进性实用内容并重以及课堂教学与实验并重。学生学后能掌握数据库的基本原理, 熟练掌握一个数据库产品的操作并学会开发、设计与管理数据库的能力。

本书由 16 章构成, 内容包括基本原理、基本操作、数据库设计及应用等, 突出 Web 数据库、数据交换等内容。本书各章均配有丰富的习题以及复习提纲, 同时本书还附有实习指南供数据库实验课之用。

本书结构清晰, 概念准确, 文字叙述简洁明了, 可读性强, 既便于教师课堂讲授, 又便于自学者阅读。

本书可作为普通高校计算机应用类及相关专业的本科“数据库课程”教材和教学参考书, 同时也可以作为数据库应用开发人员的参考书。

现代数据库系统实用教程

-
- ◆ 编 著 徐洁磐 张 劼 封 玲
责任编辑 梁 凝
 - ◆ 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>
河北三河市海波印务有限公司印刷
新华书店总店北京发行所经销
 - ◆ 开本: 787×1092 1/16
印张: 16.75
字数: 398 千字 2006 年 10 月第 1 版
印数: 1 - 4 000 册 2006 年 10 月河北第 1 次印刷

ISBN 7-115-14747-7/TN · 2773

定价: 26.00 元

读者服务热线: (010)67129258 印装质量热线: (010)67129223

前 言

随着计算机科学与技术的进步以及数据库应用的日益广泛,有力地带动了数据库的教学与培训,目前我国高校本科专业中开设“数据库”课程越来越多,它们包括:

- (1) 传统的研究型为主的“计算机”专业;
- (2) 新兴的以应用为主的“计算机应用”专业;
- (3) 近年来出现大量与计算机应用紧密相关的专业,如“软件工程”、“信息管理”、“网络通信”专业;如“自动控制”、“地理信息系统”及“电子无线电”等工科类专业;如“电子商务”、“工程管理”、“企业管理”等管理类专业,它们也大量开设数据库课程;
- (4) 大量的非计算机专业,即与计算机应用无直接关系但须掌握一定的计算机知识的专业也需设置数据库课程。

我们可以将这些“数据库”课程分为三类:

- (1) 第一类为计算机研究型专业类——以原理与理论为主;
- (2) 第二类为计算机应用及相关专业类——原理与实用并重;
- (3) 第三类为非计算机专业类——以实用为主。

这三种类型的“数据库”课程其目标、要求及内容均不相同,对教材要求也完全不同,目前在市场上所能见到的“数据库”教材以第一及第三类为多见,而第二类“数据库”教材则较为少见,但由于其实际应用需求及学生数量均占三类之首,因此编写此类“数据库”教材成为当务之急。其目的是编写一本为教学、实验、复习及习题提供全面材料的教材,该教材以原理与实用并重,并特别注意理论与实用的有机结合与统一。

基于上述之目标,本书的编写要求是:

- (1) 了解数据库系统的基本概念与基本原理/理论;
- (2) 以 SQL Server2000 为背景掌握数据库的使用方法;
- (3) 学会数据库的最新技术,包括数据交换的原理与操作以及了解 Web 数据库的使用;
- (4) 学会数据库设计和管理的原理与方法;
- (5) 了解数据库的传统与最新的应用。

根据上述五个要求编写教材,本教材共 16 章包括如下一些内容:

第 1 章~第 4 章(共四章):介绍数据库的基本概念、原理与理论以及关系数据库的基本原理,并以 SQL Server2000 为背景作说明。

第 5 章~第 7 章(共三章):介绍数据库操作的核心部分,包括数据定义、操纵及控制的基本内容和相关的 SQL 操作,并以 SQL Server2000 为例介绍详细的操作。

第 8 章~第 12 章(共五章):介绍数据交换的原理与操作,并以 SQL Server2000 为例作具体的介绍。

第 13 章~第 15 章(共三章):介绍数据库设计与管理的基本原理与方法。

第 16 章(共一章):介绍数据库的应用领域与新的应用。

最后本教材还附有八个实验指导书供上机实习之用。

本教材分为五部分，其中第一、二部分是最基础性的内容，所有选用本教材的学生均必须掌握，而其余三部分内容可以根据不同专业、不同需要有选择性地讲授。

本教材是一本特色明显的教材，主要表现为如下四个方面：

1. 先进性

(1) 本教材突出数据模型，以它作为数据库学科的核心与基础。

(2) 本教材扩充了传统的以集中式、单机为模式的数据库教学体系，引入了 C/S、B/S，以及 Web 数据库为模式的分布式方式，从而建立了五种数据交换方式的扩展体系。

(3) 本教材着重于数据库标准化的介绍，本书以 SQL'92 为基础作全面介绍，同时也介绍了 SQL'99 及 SQL'03 等内容。

(4) 本教材不仅介绍数据库传统应用，也介绍了非事务处理及数据仓库等新的应用。

(5) 本教材着重于在网络环境下的数据安全与数据管理的介绍。

(6) 本教材摒弃了部分过时及不适当的内容，如层次模型、网状模型、关系演算、嵌入式方式、查询优化等内容以及尚处于研究阶段的分布式数据库等内容。

按上述六点所构作的教材，突出了先进性、摒弃了过时的内容，使整个教材具有新颖、先进的特点，能适应目前数据库的发展与数据库应用开发的要求。

2. 实用性

本教材面向实用，使学生在掌握数据库一般原理的基础上能学会操作、使用一种数据库管理系统产品（如 SQL Server2000），并学会设计与管理数据库，了解数据库的各种应用。

3. 理论与实用的结合

计算机应用及相关专业的学生既需了解数据库的基本原理与理论，又需学会数据库的使用，两者不可缺一，关键是需两者的有机结合，而目前教材中普遍存在理论与实际脱节的倾向。在本教材中以理论引导实用，又以实用支撑理论，两者无缝结合，使学生学完该课程后，既能对数据库基础理论有完整的了解，同时也能操作与开发数据库应用。

4. 适合教学

本教材内容丰富、结构合理、书写精练，适合于教学，每章配有大量习题，并附有复习提要，鉴于数据库课程的实用性，教材还附有八个实验内容。

本书可作为高等院校计算机应用专业及计算机应用相关专业的本科“数据库”课程教材，也可作为数据库应用开发人员的参考材料以及相关应用、培训教材。

值本书付梓之际，作者首先向山东大学董继润教授表示谢意，他为审阅本书付出了艰辛的劳动，并提出了很多宝贵意见，同时感谢南京大学张德富教授、费翔林教授对本书出版的支持。

本书为南京大学金陵学院教材丛书之一，得到了南京大学计算机软件新技术国家重点实验室的支持，在本书编写过程中还得到了陈巧珍老师的具体帮助，在此一并表示感谢。

由于作者水平有限，书中的错误在所难免，恳切希望读者批评指正。

编著者

2006 年端午佳节于南京大学金陵浦苑
南京大学计算机软件新技术国家重点实验室

目 录

第 1 章 数据库系统概述	1
1.1 数据库系统的发展及当前主流	1
1.1.1 数据库系统发展简史	1
1.1.2 当前数据库主流产品介绍	2
1.2 数据库系统的若干基本概念	3
1.3 应用环境与数据库系统的发展	8
1.4 数据库系统的特点	11
习题 1	12
第 1 章复习指导	12
第 2 章 数据模型	14
2.1 数据模型的基本概念	14
2.2 数据模型的四个世界	15
2.3 概念世界与概念模型	15
2.3.1 E-R 模型	15
2.3.2 面向对象模型	20
2.4 信息世界与逻辑模型	22
2.4.1 概述	22
2.4.2 关系模型与关系模型数据库管理系统	23
2.4.3 面向对象模型与面向对象数据库管理系统	28
2.5 计算机世界与物理模型	28
2.5.1 数据库的物理存储介质	28
2.5.2 磁盘存储器及其操作	29
2.5.3 文件系统	32
2.5.4 数据库存储结构	33
习题 2	34
第 2 章复习指导	35
第 3 章 关系数据库系统介绍	37
3.1 关系数据库系统概述	37
*3.2 关系模型的数学基础——关系代数	38
3.2.1 关系的表示	38
3.2.2 关系操纵的表示	39
3.2.3 关系模型与关系代数	41

3.2.4 关系代数中的扩充运算	41
3.2.5 关系代数实例	44
3.2.6 关系代数与关系数据库系统	45
3.3 关系数据库语言 SQL	46
3.3.1 SQL 概貌	46
3.3.2 SQL 的功能	48
习题 3	49
第 3 章复习指导	50
第 4 章 关系数据库管理系统 SQL Server2000	51
4.1 SQL Server2000 系统概貌	51
4.1.1 SQL Server2000 的应用环境与结构模式	51
4.1.2 SQL Server2000 的版本	51
4.1.3 SQL Server2000 的运行平台	51
4.1.4 一个典型的系统架构	53
4.1.5 SQL Server2000 组成	54
4.1.6 SQL 语言	57
4.2 SQL Server2000 的功能与特色	57
4.2.1 SQL Server2000 的特色	57
4.2.2 SQL Server2000 的基本功能	58
4.2.3 SQL Server2000 的扩充功能	60
习题 4	60
第 4 章复习指导	60
第 5 章 关系数据库管理系统中数据定义及相关 SQL 语句	62
5.1 关系数据库管理系统的数据库定义功能介绍	62
5.2 SQL 的数据定义语句	63
5.2.1 SQL 基本数据类型	63
5.2.2 SQL 的模式定义语句	63
5.2.3 SQL 的基表定义语句	64
5.2.4 SQL 的索引定义语句	65
5.3 SQL Server2000 中的数据定义功能	66
习题 5	67
第 5 章复习指导	68
第 6 章 关系数据库管理系统的数据库操纵及相关 SQL 语句	69
6.1 关系数据库管理系统的数据库操纵功能介绍	69
6.2 SQL 的数据操纵语句	70
6.2.1 SQL 的查询语句	70

6.2.2 SQL 的更新语句	77
6.2.3 SQL 的统计、计算及分类语句	79
6.3 关系数据库管理系统的视图功能与相关的 SQL 语句	81
6.3.1 关系数据库管理系统的视图功能	81
6.3.2 SQL 有关视图的语句	81
6.4 SQL Server2000 中数据操纵及视图功能	82
习题 6	83
第 6 章复习指导	85
第 7 章 关系数据库管理系统数据控制功能及相关 SQL 语句	87
7.1 关系数据库管理系统控制功能概述	87
7.2 关系数据库的安全性控制	87
7.2.1 数据库安全的基本概念与内容	88
7.2.2 数据库的安全标准	90
7.2.3 SQL 对数据库安全的支持	92
7.2.4 SQL Server2000 中的数据安全性功能	94
7.3 关系数据库的完整性控制	98
7.3.1 关系数据库完整性控制的功能	98
7.3.2 完整性规则的三个内容	98
7.3.3 完整性约束的设置、检查与处理	99
7.3.4 触发器	101
7.3.5 SQL Server2000 中的数据完整性功能	102
7.4 事务处理	104
7.4.1 事务	104
7.4.2 事务的性质	105
7.4.3 事务活动	106
7.4.4 有关事务的 SQL 语句	106
7.4.5 SQL Server2000 中的事务功能	106
7.5 并发控制技术	107
7.5.1 事务的并发执行	107
7.5.2 封锁	110
7.5.3 封锁协议	110
7.5.4 两阶段封锁协议	112
7.5.5 封锁粒度	112
7.5.6 活锁与死锁	113
7.6 数据库故障恢复技术	114
7.6.1 概述	114
7.6.2 数据库故障分类	114
7.6.3 数据库故障恢复三大技术	115

7.6.4 恢复策略	116
7.6.5 SQL Server2000 中的故障恢复技术	117
习题 7	117
第 7 章复习指导	119
第 8 章 关系数据库系统中的数据交换原理	122
8.1 概述	122
8.1.1 数据交换模型	122
8.1.2 数据交换的五种方式	122
8.2 数据交换的管理	124
8.2.1 会话管理	124
8.2.2 连接管理	125
8.2.3 游标管理	126
8.2.4 诊断管理	126
8.2.5 动态 SQL	127
8.3 数据交换的流程	128
习题 8	129
第 8 章复习指导	129
第 9 章 数据交换之一——人机交互方式	131
9.1 人机交互方式概述	131
9.2 SQL Server2000 中的人机交互方式功能	131
9.3 SQL Server2000 中的人机交互方式的操作	132
9.3.1 企业管理器	132
9.3.2 SQL 查询分析器	136
9.3.3 事件探查器	138
9.4 SQL Server2000 中的人机交互方式的操作例子	139
习题 9	144
第 9 章复习指导	144
第 10 章 数据交换之二——自含式 SQL 及 T-SQL	146
10.1 自含式 SQL 概述	146
10.2 T-SQL 介绍	146
10.2.1 数据类型、变量与表达式	147
10.2.2 核心 SQL 操作	148
10.2.3 数据交换操作	149
10.2.4 算法程序设计语言中的程序流控制及输出语句	151
10.2.5 函数	151
10.2.6 文本、图像操作	152

10.3 T-SQL 的编程	152
习题 10	153
第 10 章复习指导	154
第 11 章 数据交换之三——调用层接口及 ODBC	155
11.1 调用层接口概述	155
11.2 C/S 结构方式	155
11.3 ODBC 介绍	156
11.3.1 ODBC 接口	157
11.3.2 ODBC 工作流程	158
11.3.3 ODBC 函数集	162
11.4 存储过程	163
习题 11	164
第 11 章复习指导	164
第 12 章 数据交换之四——Web 数据库与 ADO 接口	166
12.1 Web 方式概述	166
12.2 因特网与 Web 应用	166
12.3 Web 数据库介绍	167
12.4 B/S 结构方式	168
12.5 ASP 接口方式	169
12.6 ADO 控件	169
习题 12	172
第 12 章复习指导	173
第 13 章 关系数据库的规范化方法	175
13.1 规范化方法的起因	175
13.2 函数依赖	177
13.3 函数依赖与范式	179
13.4 范式间关系	183
13.5 规范化所引起的一些问题	184
13.6 关系数据库规范化的非形式化判别法	185
习题 13	185
第 13 章复习指导	186
第 14 章 数据库设计	188
14.1 数据库设计概述	188
14.2 数据库设计的需求分析	189
14.2.1 需求调查	189

14.2.2 需求分析	189
14.2.3 数据需求分析说明书	191
14.3 数据库概念设计	192
14.3.1 数据库概念设计概述	192
14.3.2 数据库概念设计的过程	193
14.3.3 数据库概念设计说明书	196
14.4 数据库的逻辑设计	196
14.4.1 数据库逻辑设计基本方法	196
14.4.2 关系视图设计	198
14.4.3 数据库逻辑设计说明书	199
14.5 数据库的物理设计	199
14.5.1 存取方法设计	200
14.5.2 存储结构设计	200
14.5.3 数据库物理设计说明书	201
习题 14	202
第 14 章复习指导	202
第 15 章 数据库管理	205
15.1 数据库管理概述	205
15.2 数据库管理的内容	205
15.3 数据库管理员 DBA	209
习题 15	210
第 15 章复习指导	210
第 16 章 数据库系统的应用	212
16.1 概述	212
16.2 数据库在事务处理领域中的应用	212
16.2.1 电子商务	213
16.2.2 企业资源规划	214
16.2.3 客户关系管理	216
16.3 数据库在非事务处理领域中的应用	219
16.3.1 数据库在工程领域中的应用及工程数据库介绍	219
16.3.2 数据库在多媒体领域中的应用及多媒体数据库介绍	221
16.3.3 数据库在 GIS 中的应用及空间数据库介绍	224
16.4 数据库在分析领域中的应用	226
16.4.1 决策支持系统	226
16.4.2 数据仓库的基本原理	228
16.4.3 联机分析处理 OLAP	232
16.4.4 数据挖掘	238

16.4.5 DSS 中的建模与展示.....	239
16.4.6 DSS 整体结构.....	240
习题 16.....	240
第 16 章复习指导	241
附录：“数据库课程”实验指导书	244
参考文献.....	253

第 1 章 数据库系统概述

本章主要介绍数据库系统的基本概况，包括发展简况、基本概念、特点与基本内容，本章对全书内容具有提纲挈领的作用。

1.1 数据库系统的发展及当前主流

1.1.1 数据库系统发展简史

数据库系统发展至今已有近 50 年的历史，随着计算机网络的发展，计算机应用领域的扩大，它的功能与适应范围也愈来愈广，到目前已成为信息系统的基本的、主要的支撑，它的发展经历了下面若干个阶段。

1. 人工管理阶段

20 世纪 40 年代自计算机出现至 50 年代这段时间中，由于当时计算机自身很简单且所含数据量小（如仅有 1024~4096 个单元），因此对数据的管理均是手工操作故称人工管理阶段。

2. 文件系统阶段

文件系统是数据库系统发展的初级阶段，它出现于 20 世纪 50 年代，它提供了简单的数据共享与数据管理能力，但是它无法提供完整统一的数据管理与共享能力。由于它的功能简单，因此它只能附属于操作系统而不能成为独立的软件，目前一般将其看成仅是数据库系统的雏形，而不是真正的数据库系统。

3. 层次数据库与网状数据库系统阶段

20 世纪 60 年代以后，真正的数据库系统：层次数据库与网状数据库开始发展，它们为统一管理与共享数据提供了有力支撑，这个时期数据库系统蓬勃发展形成了有名的“数据库时代”。但是这两种系统也存在不足，主要是它们脱胎于文件系统，受文件的物理影响较大，给数据库使用带来诸多不便，同时，此类系统的数据模式构造繁琐，难以推广使用。

4. 关系数据库系统阶段

关系数据库系统出现于 20 世纪 70 年代，在 80 年代得到蓬勃发展，并逐渐取代前两种系统。关系数据库系统结构简单、使用方便、逻辑性强、物理性少，因此在 80 年代以后一直占据数据库领域的主导地位。由于此系统来源于商业应用，适合于事务处理领域并在事务处理领域发挥主要作用。

5. 关系数据库系统的扩充阶段

20 世纪 90 年代以后数据库应用逐步扩充至非事务领域及分析领域的应用，同时计算机网络与因特网的出现也使传统关系数据库的应用环境受到影响，因此须对关系数据库作必要的扩充与改造，它包括：

- (1) 在传统关系数据库系统扩充非事务处理应用的能力，即所谓的对象关系数据库。
- (2) 使传统关系数据库系统能适应网络与因特网的交互能力，即所谓的 Web 数据库。
- (3) 使传统关系数据库系统扩充分析应用的能力即所谓的数据仓库。

这三种扩充功能目前已成为关系数据库的发展主流。

1.1.2 当前数据库主流产品介绍

当前数据库的主流产品均是关系数据库扩充形式，其规模大致分为大型、中型、小型及桌面式等四种，下面介绍这四种数据库的代表性的产品。

1. 大型数据库产品

大型数据库的代表性产品是 Oracle 与 DB2。

(1) Oracle 数据库管理系统

Oracle 数据库管理系统目前的版本是 Oracle V9i，它在全球的销售量很大，它除具有关系数据库的基本功能外，还具有一定的面向对象功能，它支持 Web 功能，能存储大对象数据（如图像、语音、视频及音频等数据），在它的扩充部分中还有数据仓库的功能，该产品在我国主要用于公安、金融以及大型企业中。

有关此产品的详细介绍可参阅 ORACLE 公司网站：<http://www.oracle.com>。

(2) DB2

DB2 是 IBM 公司的产品，它的前身是关系数据库管理系统的第一代产品 SYSTEM-R。它是一种关系扩充型产品，主要适用于 IBM 的大型机中，具有通用性强并有较好的并行存储与并行计算能力，该产品在我国主要用于金融、气象等大型企业、事业部门。

有关此产品的详细介绍可参阅 IBM 公司网站：<http://www.ibm.com>。

2. 中型数据库产品

中型数据库产品的代表是 Sybase 公司的产品，其主打产品有两种 Sybase Adaptive Server Enterprise 及 Sybase Adaptive Server Anywhere，它是一种关系扩充型产品，它在中国销售良好，主要用户有铁道部门、水利部门及政府系统等。

有关此产品的详细介绍可参阅 Sybase 公司的网站：<http://www.sybase.com>。

3. 小型数据库产品

小型数据库产品的代表是微软公司的 SQL Server 2000，该产品是关系模型产品并有一定的扩充功能。如 Web 功能、数据仓库功能以及大对象数据类型的表示能力，该产品适合于微型机环境并与微软公司的软件环境协调一致（如 Windows 操作系统、ODBC、OLEDB 接口及 ASP、C#、VB、VC 等开发工具），该产品在我国使用广泛，主要用于中小型企业及教育

机构中。

4. 桌面式数据库产品

桌面式数据库产品的代表是微软公司的 Access，该产品是关系模型产品，但能适应 Web 环境，该产品是以数据库为核心，并有多种开发工具相配套的产品，它是微软公司 OFFICE 系列中的一个产品，该产品适用于微型机环境并与微软公司软件环境协调一致，该产品在我国同样也使用广泛，主要用于小型企、事业单位中的简单应用。

有关上面两种微软公司产品的详细介绍可参阅微软公司网站：<http://www.microsoft.com>。

1.2 数据库系统的若干基本概念

本节将对与数据有关的六个基本概念作介绍，它们是数据、数据库、数据库管理系统、数据库管理员、数据库系统及数据库应用系统。

1. 数据 (data)

(1) 什么是数据

数据是现实世界中客体在计算机中的抽象表示，具体地说，它是一种存储于计算机内的符号串。

(2) 数据的特性

数据有下面五个特性：

① 数据表现形式的多样性

从表现形式看，数据表现的形式很多，除常用的数字、文字、时间等表示形式外，还包括图像、图形、语言、视频等多媒体数据以及表示知识、规则、数学符号及推理等的抽象数据，数据表现形式的多样性为数据的应用提供了有力的基础。

② 数据的可构造性

从结构看，数据分为结构化数据 (structured data)、半结构化数据 (semistructred data) 与非结构化数据 (non-structured data)。所谓非结构化数据即表示符号串是不规则结构形式，所谓半结构化数据即表示符号串呈半规则结构形式，如文件中之流式文件，以及因特网中的 Web 结构等均属非结构化及半结构化形式。而在软件中的数据大多是有结构的，它们称结构化数据。首先，结构化数据有型 (type) 与值 (value) 之分，数据的型给出了数据表示的类型，如整型、实型、字符型等，而数据的值给出了符合给定型的数值。随着应用需求的扩大，数据的型有了进一步的扩大，它包括了将多种相关数据以一定结构方式组合构成特定的数据框架称为数据结构 (data structure)，具有统一结构形式的数据结构的具体描述可称为数据模式 (data schema)。

③ 数据的挥发性与持久性

从存储时间看，数据一般分为两部分，其中一部分与程序仅有短时间的交互关系，随着程序的结束而消亡，它们称为临时性数据或挥发性 (transient) 数据。这类数据一般存放于计算机内存中；而另一部分数据则对系统起着长期持久的作用，它们称为持久性 (persistent)

数据，这类数据一般存放于计算机中的次级存储器（如磁盘）内。

④ 数据的私有性与共享性

从其使用对象看，数据可分为私有性与共享性数据两种。为特定应用（程序）服务的数据称私有（private）数据，而为多个应用（程序）服务的数据则称为共享（share）数据。

⑤ 数据的海量性

从其存储数量看，数据可分为小量、大量及海量三种。数据的量是衡量与区别数据的重要标志，这主要是由于数据“量”的变化可能会引起数据“质”的变化。数据量由小变大后，数据就需要进行管理，需要保护与控制。目前数据以海量数据为多见，因此一般数据均需管理、保护与控制。

随着技术的进步与应用的扩大，数据的特性都在发生变化，这些变化主要表现为：

- 数据的量由小量到大量进而到海量；
- 数据的组织由非结构化到结构化；
- 数据的服务范围由私有到共享；
- 数据的存储周期由挥发到持久。

数据的这些变化使得现代数据具有海量的、结构化的、持久的和共享的特点，本书如不作特别说明，所提数据即具此四种特性。

下面讨论数据与其他几个重要概念间的关系，它们是数据与软件及数据与信息间的关系。

(3) 数据与软件（software）

首先讨论软件，软件是计算机科学与技术的一大门类，它是建立在计算机硬件之上的一种运行（或处理）实体。软件一般由程序与数据两部分组成，其中程序给出了运行的过程而数据则给出了运行的对象与结果，程序与数据的关系如图 1.1 所示。

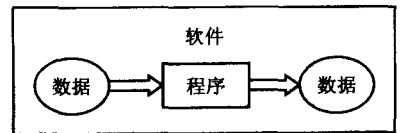


图 1.1 程序与数据关系示意图

在软件中数据（主要指其结构）是其最稳定部分，而程序则是可变部分，因此数据称为软件中的不动点（fixed point），它在软件中起着基础性的作用。

软件发展至今，程序与数据间的不同关系形成了目前流行的两种结构方式。

① 以程序为中心的结构：在此种软件结构中以程序为中心以数据为辅助，即每个程序有若干个数据为其支撑，它们构成了如图 1.2 (a) 所示结构。

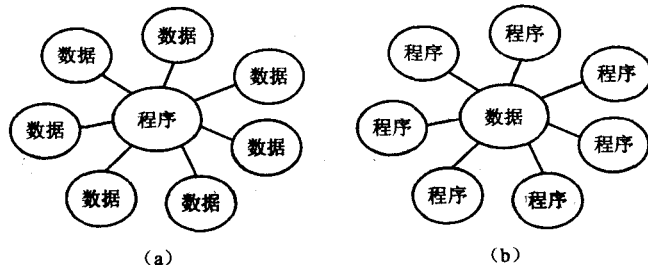


图 1.2 软件两种结构图

② 以数据为中心的结构：在此种软件结构中以数据为中心以程序为辅助，即以一个数据集合为中心围绕它有若干个程序对数据作处理，它们构成了如图 1.2 (b) 所示的结构。

在目前,大多数软件结构采用以数据为中心的结构。

在过去,软件是以程序为中心,而数据则以私有形式从属于程序。在此种系统中,数据是分散、凌乱的,它造成了数据管理的混乱,如数据冗余大、一致性差、结构复杂等多种弊病,但经过若干年的发展,数据在软件中的地位和作用发生了本质的变化,在软件中它已占主体地位,而程序则已退居附属地位,它们构成了以数据为中心的结构。在此种结构中,需要对数据作集中、统一的管理,并使其为多个应用程序共享,它们构成了如图1.3所示的结构图,这种结构方式为数据库系统的产生与发展奠定了基础。

(4) 数据与信息(information)

数据与信息是两种不同的概念,在实际应用中往往引起混淆,数据是指客体在计算机中表示的形式而信息则是客体的语义表现。数据与信息表示了一个事物的两个不同方面的理解,但是数据与信息又有关联,即数据可以表示信息,而信息也在一定程度可以反映数据。

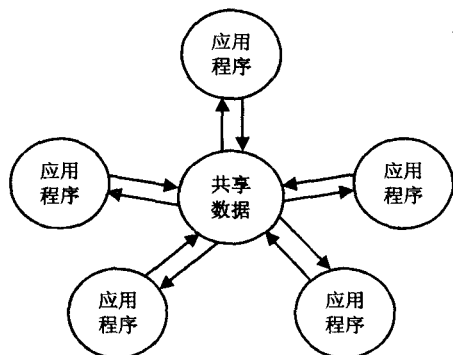


图 1.3 以数据为主体的软件系统

2. 数据库(database, 简称 DB)

数据库是数据的集合,它具有统一的结构形式,存放于统一的存储介质内,并由统一机构管理,它由多种应用数据集成,并可被应用所共享。

数据库存放数据,数据按所提供的数据模式存放,它能构造复杂的数据结构以建立数据间内在联系与复杂关系,从而构成数据的全局结构模式。

数据库中的数据具有“集成”、“共享”的特点,亦即是数据库集中了各种应用的数据,并对其进行统一的构造与存储,而数据可为不同的应用服务与使用。

3. 数据库管理系统(database management system, 简称 DBMS)

数据库管理系统是统一管理数据库的一种软件(属系统软件),它负责:

- 组织数据库中的数据;
- 操纵数据库中的数据;
- 维护数据库中的数据;
- 控制及保护数据不受破坏;
- 数据库中的数据交换;
- 数据库中的数据服务;
- 数据字典。

数据库中的数据是具有海量级的数据,且其结构复杂,因此是需要管理的,它主要有如下几方面功能:

(1) 数据模式定义

数据库管理系统负责为数据库构造模式,亦即是它为数据库构造其统一的数据框架。

(2) 数据存取的物理构造

数据库管理系统负责为数据模式的物理存取构造有效的存取方法与手段,如构造索引