

数据库技术原理 与应用教程



徐洁磐 常本勤 编著

本书为教师配有电子教案



机械工业出版社
China Machine Press

TP311.13/314

2008

普通高等院校计算机课程规划教材

数据库技术原理 与应用教程



徐洁磐 常本勤 编著



机械工业出版社
China Machine Press

本书由基础篇、操作篇和应用篇三部分组成，内容涵盖数据库系统的基本概念和理论、常用数据库系统 SQL Server 的相关操作以及数据库技术的相关应用。各章后均有习题及相关的内容小结，附录中还提供了实验指导，帮助读者巩固所学知识。

本书结构合理、内容先进、注重理论与实践相结合，适合作为高等院校计算机及相关专业的学生的数据库课程教材，也可供相关技术人员参考。

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

数据库技术原理与应用教程/徐洁磐等编著. —北京: 机械工业出版社, 2008. 1
(普通高等院校计算机课程规划教材)

ISBN 978-7-111-22945-2

I. 数… II. 徐… 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(2007)第 186615 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: 朱 劼

北京牛山世兴印刷厂印刷·新华书店北京发行所发行

2008 年 1 月第 1 版第 1 次印刷

184mm × 260mm · 18.25 印张

标准书号: ISBN 978-7-111-22945-2

定价: 29.00 元

凡购本书, 如有缺页、倒页、脱页, 由本社发行部调换
本社购书电话: (010)68326294

前 言

近年来,我国在计算机本科系科中开设“数据库”课程的高校越来越多,其中涉及以下三种类型的专业:

- 1)以研究为主的“计算机相关专业”。
- 2)以应用为主的“计算机应用专业”及“计算机应用相关专业”。
- 3)需掌握一定计算机知识的“非计算机专业”。

这三种专业的“数据库”课程的教学目标、要求与教学内容不尽相同,因此需要采用不同类型的教材。但是,目前市场上的数据库教材多面向第一类与第三类专业,面向第二类专业的教材则较为少见,而由于实际应用需要,此类学生数量又占三类专业学生数量之首,因此编写面向此类专业的数据库教材成为当务之急。

本书就是这样一本面向应用型专业的数据库课程教材。本书的编写目标是:以应用为核心,以基础与操作为支撑,注重理论与实际的结合,学生学完本书后既具有数据库的基本理论知识,又能进行数据库操作,而且能从事数据库领域的实际工作。

下面介绍本书编写的具体内容:

1. 数据库应用

本书是计算机应用型教材,因此力求培养学生的以下能力:

- 1)从事数据库应用系统开发的能力。
- 2)初步的数据库设计能力。
- 3)初步从事数据库管理的能力。

2. 数据库基础知识

- 1)能掌握数据库系统与关系数据库系统的一般性原理与基础理论知识。
- 2)在数据库基础知识的介绍中坚持先进性与实用性,淘汰传统教材中落后、过时的内容(如关系演算、查询优化、层次模型、网状模型、嵌入式 SQL 以及传统的分布式数据库系统等),增加先进与实用的内容(如数据交换、面向对象模型与对象-关系模型以及 XML 数据库与 Web 数据库)。

3)在介绍系统一般性原理的同时,以一个具体的系统(SQL Server 2000)为蓝本进行介绍与分析。

3. 数据库操作

- 1)以介绍 SQL 语言作为数据库操作的主要内容。
- 2)SQL 语言的介绍以 ISO SQL 为标准,以 SQL'92 为主要内容并兼顾 SQL'99 及 SQL'03,特别要介绍数据交换中的 SQL 内容,包括人机交互、自含式、调用层接口、Web 数据库与 XML 数据库等 SQL 扩充结构的内容。
- 3)同时介绍 SQL Server 2000 中的 SQL 语句,注重标准化与实用性相结合。

本书由基础篇、操作篇与开发应用篇三部分构成。各部分的具体内容如下：

(1) 基础篇

基础篇由第1章~第6章组成，主要介绍数据库系统与关系数据库系统的一般性原理与基础理论，其中第1章~第3章介绍数据库系统的一般性原理与基本理论，第4章~第6章介绍关系数据库系统的基础理论与原理，并以SQL Server 2000为例说明相关概念与原理。

(2) 操作篇

操作篇由第7章~第12章组成，主要介绍ISO SQL以及SQL Server 2000中的SQL语句及使用方法，内容涉及SQL核心部分及扩展部分等8部分。其中，数据定义、数据操纵、数据交换以及数据控制属核心部分，而人机交互方式、自含式方式、调用层接口方式及Web方式属扩展部分。

(3) 开发应用篇

应用篇由第13章~第16章组成，主要介绍前面所述的应用的三个领域以及应用的三个领域，即传统事务处理领域、非传统事务处理领域和分析领域。

本书是一本特色明显的教材，主要表现在如下几个方面：

• 定位准确

本书面向应用类专业学生需要，书中既有数据库基本原理与基本操作等理论性内容，也有实用性内容，学生学完本书后能掌握数据库的基础知识与基本技能，并能进行实际应用，同时也可作为学习后续课程及进一步研究打下基础。

• 结构合理

本书以实用性为目标，将整个数据库技术内容组织成三大部分，改变了以往复杂、繁琐的结构体系，更适合目标读者的需求。

• 内容先进

本书重点介绍国内外先进成熟技术，抛弃了一些过时陈旧的内容，因而具有明显的时代特征。

• 实用性以及理论与实际结合

本书注重实用，教材内容与数据库实际应用紧密结合。而且，本书注重理论与实际的结合，基础理论能指导实际应用，同时实际应用又能支撑理论，对能力的培养也大有助益。

• 适合教学

本书针对教学需要合理安排结构体系，同时配有大量实验与应用性习题并为授课教师提供丰富成熟的教学课件，因此特别适合于教学。

本书可作为普通高校计算机应用类专业本科生“数据库”课程的教材，也可作为数据库应用开发人员的参考材料以及培训教材。

值本书付梓之际，首先向山东大学董维润教授表示感谢，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与费翔林教授对本书的支持。此外，本书也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，在此一并表示感谢。

由于作者水平有限，错误之处在所难免，恳请读者指正。

编 者

2007年5月于

南京大学计算机软件新技术国家重点实验室

目 录

前言

第一篇 基础篇

第1章 数据、数据管理与数据处理 ... 2

- 1.1 概述 2
- 1.2 数据及其特性 3
- 1.3 数据与数据库 3
- 1.4 数据管理 4
- 1.5 数据管理的变迁 5
- 1.6 数据管理工具与产品 7
- 1.7 数据处理 7
 - 1.7.1 数据处理的环境 8
 - 1.7.2 应用开发 8
 - 1.7.3 数据处理的应用领域 8
- 本章小结 9
- 习题1 10

第2章 数据库的基础知识 11

- 2.1 数据库中的基本概念 11
- 2.2 数据库内部结构体系 16
 - 2.2.1 数据库的三级模式 16
 - 2.2.2 数据库的两级映射 17
- 2.3 数据库应用环境与数据交换 17
- 2.4 数据库系统的特点 19
- 本章小结 21
- 习题2 22

第3章 数据管理中的数据模型 23

- 3.1 数据模型的基本概念 23
- 3.2 数据模型的四个世界 23
- 3.3 概念世界与概念模型 24
 - 3.3.1 E-R 模型 24
 - 3.3.2 面向对象模型 29
- 3.4 信息世界与逻辑模型 31
 - 3.4.1 概述 31

- 3.4.2 关系模型 32

- 3.4.3 面向对象逻辑模型和对象关系模型 36

3.5 计算机世界与物理模型 38

- 3.5.1 数据库的物理存储介质 38
- 3.5.2 磁盘存储器及其操作 38
- 3.5.3 文件系统 39
- 3.5.4 数据库的物理结构 40

本章小结 41

习题3 42

第4章 关系模型的基本理论 43

- 4.1 关系模型的基本理论概述 43
- 4.2 关系代数 43
 - 4.2.1 关系的表示 43
 - 4.2.2 关系操作的表示 44
 - 4.2.3 关系模型与关系代数 47
- 4.3 关系数据库的规范化方法 49
 - 4.3.1 规范化方法的起因 49
 - 4.3.2 函数依赖 51
 - 4.3.3 函数依赖与范式 54
 - 4.3.4 模式分解 56
 - 4.3.5 范式间的关系 59
 - 4.3.6 关系数据库规范化的非形式化判别法 59

本章小结 60

习题4 61

第5章 关系数据库管理系统的组成 及其标准语言 63

- 5.1 关系数据库管理系统概述 63
 - 5.1.1 简介 63
 - 5.1.2 关系数据库管理系统的组成 64
 - 5.1.3 SQL 语言 65
- 5.2 数据构造功能 65
 - 5.2.1 关系数据库 65
 - 5.2.2 基表 65

5.2.3 视图	66	7.2 SQL 的数据定义语句	111
5.2.4 物理数据库	66	7.2.1 SQL 的基本数据类型	111
5.3 数据操纵功能	66	7.2.2 SQL 的模式定义语句	111
5.4 数据控制功能	67	7.2.3 SQL 的表定义语句	112
5.4.1 安全性控制	68	7.2.4 SQL 的索引定义语句	113
5.4.2 完整性控制	72	7.3 SQL Server 2000 中的数据定义 语句	113
5.4.3 事务处理	74	7.4 SQL 的数据操纵功能	115
5.4.4 并发控制	76	7.5 SQL 的数据操纵语句	115
5.4.5 故障恢复	82	7.5.1 SQL 的查询语句	115
5.5 数据交换功能	84	7.5.2 SQL 的更新语句	123
5.5.1 概述	85	7.5.3 SQL 的统计、计算及分类 语句	124
5.5.2 数据交换的管理	86	7.6 SQL 中的视图语句	126
5.5.3 数据交换的流程	90	7.7 SQL Server 2000 中的数据操纵及 视图功能	127
5.6 数据服务及数据字典	91	本章小结	128
5.7 关系数据库管理系统的扩充功能	91	习题 7	129
5.8 关系数据库管理系统的标准 语言 SQL	93	第 8 章 SQL 的数据控制语句	132
5.8.1 SQL 的概貌	93	8.1 SQL 的数据控制功能	132
5.8.2 SQL 的功能	95	8.2 SQL 的安全性控制语句	132
5.8.3 SQL 的三种标准	96	8.2.1 SQL 对数据库安全的支持	132
本章小结	96	8.2.2 SQL Server 2000 中的数据安全性 功能	135
习题 5	97	8.3 SQL 的完整性控制	138
第 6 章 关系数据库管理系统 SQL Server 2000	99	8.3.1 SQL 的完整性控制语句	138
6.1 SQL Server 2000 概貌	99	8.3.2 触发器语句	140
6.1.1 SQL Server 2000 的应用环境与 结构模式	99	8.3.3 SQL Server 2000 中的数据完整性 功能	142
6.1.2 SQL Server 2000 的版本	99	8.4 SQL 的事务语句	145
6.1.3 SQL Server 2000 的运行平台	99	8.4.1 有关事务的 SQL 语句	145
6.1.4 SQL Server 2000 典型的系统 架构	100	8.4.2 SQL Server 2000 中的事务 功能	145
6.1.5 SQL Server 2000 的组成	101	8.5 SQL 的故障恢复功能	146
6.1.6 SQL 语言	104	8.5.1 故障恢复三大功能	146
6.2 SQL Server 2000 的特色	104	8.5.2 SQL Server 2000 中的故障恢复 功能	146
6.3 SQL Server 2000 的基本功能	105	本章小结	147
本章小结	107	习题 8	148
习题 6	108	第 9 章 SQL 的数据交换语句	149
第二篇 操作篇		9.1 会话管理语句	149
第 7 章 SQL 的数据定义与操纵 语句	110		
7.1 SQL 的数据定义功能	110		

9.2 连接管理语句	149	12.2 C/S 结构方式	182
9.3 游标管理语句	150	12.3 SQL/CLI 概述	183
9.4 诊断管理语句	151	12.4 ODBC 介绍	185
9.5 动态 SQL 管理	151	12.4.1 ODBC 接口	185
本章小结	153	12.4.2 ODBC 的工作流程	185
习题 9	153	12.4.3 ODBC 函数集	192
第 10 章 SQL 中数据交换之人机交互		本章小结	192
方式	154	习题 12	193
10.1 人机交互方式概述	154	第 13 章 SQL 中数据交换之 Web	
10.2 SQL Server 2000 中的人机交互		方式	194
方式	154	13.1 Web 应用	194
10.3 SQL Server 2000 中的人机交互		13.1.1 互联网与 Web 应用	194
方式的操作	155	13.1.2 XML 简介	195
10.3.1 企业管理器	155	13.2 XML 数据库	197
10.3.2 SQL 查询分析器	160	13.2.1 XML 数据库的基本原理	197
10.3.3 事件探查器	162	13.2.2 SQL/XML	198
10.4 SQL Server 2000 中的人机交互方式		13.2.3 SQL Server 2000 中的 XML 数	
操作实例	163	据库	199
本章小结	169	13.3 Web 数据库	200
习题 10	169	13.3.1 B/S 结构方式	200
第 11 章 SQL 中数据交换之自含式		13.3.2 Web 数据库基本原理	201
方式	170	本章小结	205
11.1 自含式 SQL 概述	170	习题 13	206
11.2 SQL/PSM 介绍	170	第三篇 开发应用篇	
11.2.1 SQL/PSM 语句	170	第 14 章 数据库应用系统开发	210
11.2.2 SQL/PSM 中模块的建立	171	14.1 数据库应用系统概述	210
11.3 T-SQL 介绍	171	14.1.1 数据库应用系统的组成	210
11.3.1 数据类型、变量与表达式	172	14.1.2 数据库应用系统的开发	211
11.3.2 基本 SQL 操作	173	14.2 数据库应用系统的平台	211
11.3.3 数据交换操作	173	14.2.1 网络结构平台	211
11.3.4 算法程序设计语言中的程序流		14.2.2 中间件	212
控制及输出语句	175	14.3 数据库应用系统中的数据层	217
11.3.5 函数	175	14.3.1 数据库应用系统中数据层的	
11.3.6 文本、图像操作	175	组成	217
11.4 存储过程	176	14.3.2 数据库应用系统中数据层的	
11.5 T-SQL 编程	177	开发	218
本章小结	180	14.4 数据库应用系统中的应用层	218
习题 11	181	14.4.1 数据库应用系统中应用层的	
第 12 章 SQL 中数据交换之调用层		组成	218
接口方式	182	14.4.2 数据库应用系统中应用层的	
12.1 调用层接口概述	182	开发	219

14.5 数据库应用系统中的界面层	219	16.4 数据库管理员	243
14.6 典型的数据库应用系统组成	219	本章小结	244
本章小结	221	习题 16	245
习题 14	222		
第 15 章 数据库设计	223	第 17 章 数据库的应用领域	246
15.1 数据库设计概述	223	17.1 概述	246
15.2 数据库设计的需求分析	224	17.2 数据库在传统事务处理领域中的 应用	246
15.2.1 需求调查	224	17.2.1 电子商务	247
15.2.2 需求分析	224	17.2.2 ERP	248
15.2.3 数据需求分析说明书	225	17.2.3 CRM	250
15.3 数据库的概念设计	226	17.3 数据库在非传统事务处理领域中 的应用	253
15.3.1 数据库的概念设计概述	226	17.3.1 数据库在工程领域中的应用及 工程应用数据库	253
15.3.2 数据库概念设计的过程	226	17.3.2 数据库在多媒体领域中的应用 及多媒体数据库	254
15.3.3 数据库概念设计说明书	229	17.3.3 数据库在 GIS 中的应用及空间 数据库	257
15.4 数据库的逻辑设计	230	17.4 数据库在分析领域中的应用	258
15.4.1 数据库的逻辑设计基本方法	230	17.4.1 决策支持系统	258
15.4.2 关系视图设计	232	17.4.2 数据仓库的基本原理	260
15.4.3 数据库逻辑设计说明书	232	17.4.3 联机分析处理	263
15.5 数据库的物理设计	233	17.4.4 数据挖掘	270
15.5.1 存取方法设计	233	17.4.5 DSS 中的建模与展示	270
15.5.2 存储结构设计	234	17.4.6 DSS 整体结构	271
15.5.3 数据库物理设计说明书	235	本章小结	271
本章小结	236	习题 17	273
习题 15	237	附录 实验指导	274
第 16 章 数据库管理	238	参考文献	283
16.1 数据库管理概述	238		
16.2 数据库管理的内容	238		
16.3 数据管理工具	243		

第一篇 基础篇

数据库技术是计算机学科中的一门重要分支，它已有五十余年历史并已成为一门完整的学科，其主要内容包括基础理论、基本操作及应用等。

数据库技术的基础理论部分是构成该学科的基石，它给出了该学科的抽象的、全局的研究结果并对整个学科起指导性作用。

在本书中，基础部分由两方面内容组成，它们是数据库技术的一般性理论和关系数据库技术的理论。

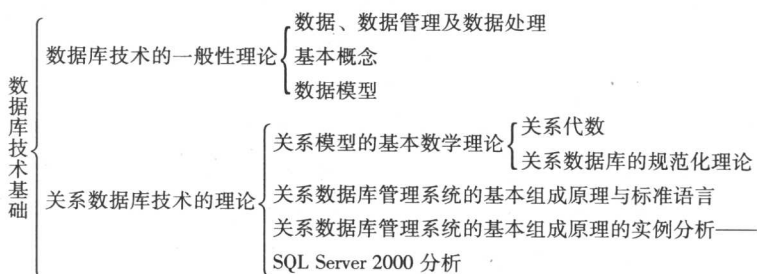
1. 数据库技术的一般性理论

第1章~第3章介绍数据库技术的一般性理论。其中第1章介绍有关数据、数据管理及数据处理的一般性概念，第2章介绍数据库技术中的一些基本概念，第3章介绍数据库系统的核心部分，即数据模型。这三章内容刻画了数据库技术中的基本理论体系。

2. 关系数据库技术的理论

在数据库技术中，目前最为流行的是关系数据库系统，因此本篇中将重点介绍关系数据库技术的理论，它由第4章~第6章组成。其中，第4章介绍关系模型的两种数学理论，它们是关系代数与关系数据库规范化理论；第5章介绍关系数据库管理系统的组成原理及其标准语言 SQL 的概貌；第6章是第5章的一个具体体现，即用第5章的原理对 SQL Server 2000 进行系统分析与介绍。这三章内容概述了关系数据库技术的数学基础与关系数据库管理系统的组成原理。

本篇的组织结构可用下图表示。



数据库技术基础的内容图

第1章 数据、数据管理与数据处理

数据、数据管理与数据处理是全书的讨论主题，在本章中先对它们进行概要介绍，使读者对这些概念有一个宏观、全局的认识。

1.1 概述

当今社会是一个信息社会，“信息”二字使用频率极高。此外，大家还时常听到“数码”、“数字”等新名词，它们在计算机领域中都是“数据”的不同表现形式。一般而言，客观世界的自然界与人类社会中的各种现象与事物都可以抽象为计算机中的数据，我们常说的“数字化城市”、“数字化地球”、“数字化工厂”正是这种抽象的一个体现。

经过这种抽象后，客观世界中的事物与现象均可转化成为计算机中的“数据”，由于数据在现实世界中的重要性，因此必须对它作深入研究，故而出现了数据库技术，它是以“数据”作为其研究对象的一门学科。

随着应用的发展，数据的“量”逐渐增大，由“大量”到“海量”，因此有必要将数据按应用领域“集成”于一起，这就构成了数据库。因此，本书中所说的“数据”指的是以数据库为组成形式的数据。

其次，为方便使用数据，必须对数据加以管理。数据管理一般由一组软件实现，它们称为数据库管理系统，必要时还可由一组人员协助管理数据库，这些人员称为数据库管理员。数据管理是数据库技术的主要研究内容。

在对数据进行抽象后，我们对“客观世界事物与现象”的研究可以转化为对计算机中“数据”的研究，这称为数据处理。数据处理是指对数据进行加工、转换、传输、存取、采集及发布等处理，它是一种新的研究方法 with 思想，可以借助数据处理对客观世界的事物与现象进行研究。正因为如此，数据处理目前已成为世界上林林总总各门学科研究的基本方法与工具。

目前，数据处理有两个主要方向：

1) 数据的事务处理：数据的事务处理是一种由“数据”到“数据”的处理，也就是将一组“数据”经加工而转换成另一种“数据”。数据的事务处理是目前最为常用的一种手段。

2) 数据的分析处理：数据的分析处理是一种由“数据”到“规则”的处理，即将一组“数据”经加工而转换成一组“规则”。数据的分析处理是目前新兴的一种手段，具有创造性，其发展潜力很大。

经数据处理后所获得的“数据”及“规则”在客观世界中可以得到一定的语义解释并可成为客观世界中的研究成果。

上面所述的内容可以用图 1-1 表示。

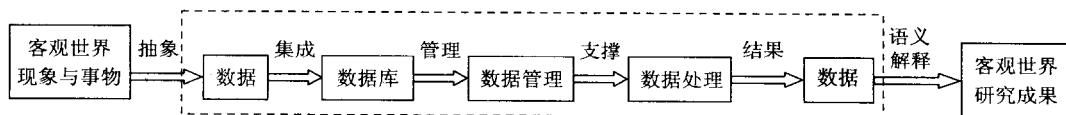


图 1-1 数据处理方法示意图

从图 1-1 中可以看出, 客观世界现象与事物经抽象成为计算机中的数据, 再经数据库与数据管理并以它们作支撑进行数据处理后获得新的数据与规则, 将这些数据与规则进行语义解释后即成为客观世界中的研究成果。在此流程中可以看出, 对客观世界的研究可转换成为数据处理的流程, 而数据库技术即是以研究该部分作为其主要内容, 其中:

- 1) 数据(包括数据库)是数据库技术研究的目标。
- 2) 数据管理是数据库技术研究的内容。
- 3) 数据处理是数据库技术的应用方向。

在这三部分中, 数据与数据管理是数据库技术的主要部分, 而数据处理是以数据与数据管理作为其基本支撑。

1.2 数据及其特性

数据是客观世界中的现象与事物在计算机中的抽象, 是数据库技术研究的主要目标。数据有很多特性, 主要体现在以下几个方面:

1. 数据表示的广泛性

数据是客观世界事物的抽象表示, 现实世界中的客体都能用数据表示。所以我们说, 数据反映了客观世界, 它是现实世界在计算机中的一种模拟。例如, 数据可以表示抽象的符号、推理; 可以表示二维、三维及多维空间结构; 可以表示时刻、日期、年代等时间形式; 也可以表示图形、图像、声音、视频、音频等多媒体形式等。

数据表示的广泛性表明它与现实世界紧密相关, 与人类社会息息相通, 因此我们可以利用数据的这个特性来讨论与研究客观世界与人类社会。

2. 数据是一种重要的信息资源

当今社会, 人们不仅拥有充足的物质财富, 还拥有海量的数据, 这是一笔丰富的信息财富, 可以用它为社会与国民经济建设服务。

3. 数据可以创造财富、创造文明

利用物质资源可以创造财富, 创造文明, 同样利用信息资源也可以创造财富与创造文明, 因为可以通过归纳、整理与分析大量的数据而获得创造性的规则, 从而为人类服务。

1.3 数据与数据库

在计算机技术日益发达的今天, 计算机的存储容量不断扩大, 处理能力不断加快, 数据大都集中存放于计算机或计算机网络中并以数据库(database)的形式出现。因此, 在讨论数据时, 我们主要讨论计算机环境中的数据, 即讨论数据库。本书中所指的数据就是在数据库中数据。

对于数据库, 我们主要研究与讨论的问题是:

- 1) 数据类型与数据结构: 这是一种关于数据的结构形式的研究。
- 2) 数据模式: 这是一种关于数据库中全局统一的数据结构的形式研究。

3) 数据模型: 这是一种关于数据抽象性质表示的研究, 它包括数据的结构, 建立在数据结构上的数据操纵以及数据间内在的语法、语义约束。它反映了数据的静态特性、动态行为以及内部制约关系。

1.4 数据管理

人类社会有着巨大的数据资源，为有效地使用它们，必须对它们进行管理。数据管理是数据库技术研究的核心，它涉及如下几个方面：

1. 数据组织

为便于数据管理，必须对大量或海量的数据进行有序与有机的组织，使其能存储在一个统一的组织结构下，这是数据管理的首要工作。这种数据组织就构成了数据库中数据的实体。

2. 数据定位与查找

在浩如烟海的数据中如何找到所需的数据是数据管理的重要任务，这种查找的难度可用“大海捞针”来形容。查找的关键是数据的定位，即找到数据的位置，只有定位后数据查找才成为可能。因此，数据定位与查找是数据管理的一项艰巨任务。此外，它还包括对数据的修改、删除与增添等工作。

3. 数据保护

数据是一种资源，其中大部分是不可再生资源，因此必须对它们加以保护以防止丢失与破坏。数据保护一般包括以下几个部分：

1) 数据语法与语义正确性保护：数据库中的数据是受一定语法、语义约束的，如职工年龄一般在 18~60 岁之间，职工工资一般在 1000~8000 元之间等，而且职工的工资与其工龄、职务均有一定语义关联，任何违反约束的数据必为不正确数据。因此，必须保护其语法、语义的正确性。

2) 数据访问正确性保护：数据库中的数据是共享的，而共享是受限的，过分的共享会带来安全隐患，如职工对其工资只有读权限而无写与改权限。因此，数据访问权限是受限的，而正确访问权限是受到保护的。

3) 数据动态正确性保护：在多个用户访问同一数据时会相互间产生干扰，从而造成数据的不正确，因此要防止此种现象产生，这称为数据动态正确性保护。

4) 最后，是数据动态正确性保护的另一种现象，即在执行数据操作时要防止因外界干扰而产生的故障。

4. 数据接口

为方便使用数据，必须为不同应用环境的用户提供不同接口，其中包括传统的人机交互环境接口、单机环境接口、网络环境接口、互联网环境接口等。

5. 数据服务与元数据

在数据管理中，还提供大量的服务功能，这称为数据服务(data service)。数据服务一般包括操作服务与信息服务，其中操作服务主要为用户提供操作上的方便，而信息服务则为用户使用数据库提供信息，特别是数据结构信息、数据控制信息，这种信息是有关数据的数据，因此称为元数据(metadata)。元数据是一种特殊的数据服务，由于它的重要性，本书中将对它单独命名并加以介绍。

上述五种数据管理功能在数据库中分两个层次进行管理，第一个层次是低层次管理，它

负责数据管理中简单、常规的管理,相关工作由系统软件——数据库管理系统完成;第二个层次是高层次管理,它由人——数据库管理员进行管理,主要完成数据管理中复杂、智能性管理工作。

1.5 数据管理的变迁

数据管理是数据库技术的核心,在数据库的发展历史中,它经历了多个阶段。

1. 人工管理阶段

自20世纪40年代计算机出现至50年代这段时间中,由于当时计算机结构简单,应用面狭窄且存储单元少,对计算机内的数据的管理非常简单,主要由应用程序编制人员管理各自的数据,此阶段称为人工管理阶段。

2. 文件管理阶段

文件系统是数据库系统发展的初级阶段,它出现于上世纪50年代中期,此时计算机中已有磁鼓、磁盘等大规模存储设备,计算机应用面也逐步拓宽,此时计算机内的数据已开始有专门的软件管理,这就是文件系统(file system)。

文件系统能对数据进行初步的组织,并能对数据进行简单查找及更新操作,但是文件对数据的保护能力差,同时由于当时应用环境简单,因此接口能力差。由于文件系统的管理能力简单,因此它只能附属于操作系统而不能成为独立部分,目前一般将其看成是数据库系统的雏形,而不是真正的数据库系统。

文件系统主要有以下两点不足:

(1) 文件系统的共享性差

在文件系统中,每个文件均为为特定应用程序服务的。在一个计算机中,如果有多个应用,则必须建立多个为应用服务的独立、分散的文件,它们的冗余性差,一致性低,极大地浪费了存储空间且容易造成数据管理的混乱。这些都是文件系统缺乏数据的共享性所带来的弊病。

(2) 文件间缺少内在逻辑联系

由于文件依附于应用程序,不同应用的文件间是彼此隔离的,而且相同应用中的文件也依附于不同的应用需求,它们间也是孤立的。因此,整个文件系统内各文件间是彼此孤立的,是一个无弹性、无结构的数据集合体。这反映了文件系统内在结构上的缺陷,会对数据管理中的数据组织与数据查找更新的能力产生影响,更有甚者,它无法反映数据间内在的逻辑联系,人为制造了“信息孤岛”。

文件系统的这种不足带来了结构上的弊端。这种结构方式一般称为以程序为中心的结构方式,它可用图1-2表示,从图中可以看出,以程序为核心,数据依附于程序,而数据间则彼此隔离与孤立。

3. 数据库管理阶段

自20世纪60年代起,数据管理进入了数据库管理阶段。由于计算机规模日渐庞大,应用日趋广泛,计算机存储设备已出现大容量磁盘与磁盘组,且数据量已由大量跃至海量,传统的文件系统已无法满足新的数据管理要求,因此数据管理职能由附属于操作系统的文件系统而脱离成独立的数据管理机构,即成为数据库管理系统。

数据库管理系统克服了文件系统的不足，特别是在共享性以及数据间逻辑联系方面的不足，使数据库系统成为能适应当代计算机应用发展的数据管理机构。其主要特点是：在数据库中，每个数据不再像文件系统那样仅针对某个应用，而是根据应用全面组织数据，做到数据对所有应用共享，同时根据数据内在关联建立起数据全局、整体的结构化组织。数据库系统的这种结构方式称为以数据为中心的结构方式，它可用图 1-3 表示。在该图中可看到，以整体、全局数据为核心，围绕它的是若干个程序对数据进行处理。

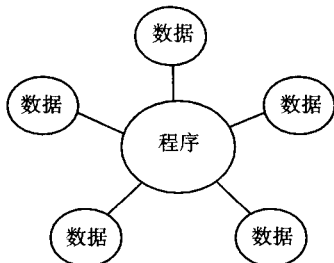


图 1-2 以程序为中心

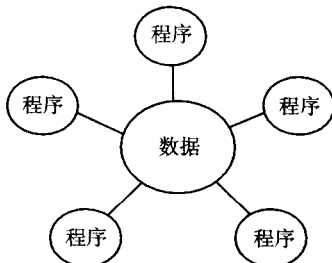


图 1-3 以数据为中心

数据库管理系统阶段因不同的数据结构组织而分为三代，它们是：

(1) 第一代数据库管理——层次与网状数据库管理时代

20 世纪 60 年代以后所出现的数据库管理系统是层次数据库与网状数据库，它们具有真正的数据库管理系统特色。但是，由于它们脱胎于文件系统，受文件的物理影响大，因此给数据库使用带来诸多不便。

(2) 第二代数据库管理——关系数据库管理时代

关系数据库管理系统出现于 20 世纪 70 年代，在 20 世纪 80 年代得到了蓬勃的发展并逐步取代前两种系统。关系数据库管理系统结构简单、使用方便、逻辑性强、物理性少，因此一直占据数据库领域的主导地位。关系数据库管理系统起源于商业应用，它适合于事务处理领域并在该领域内发挥主要作用。

(3) 第三代数据库管理——后关系数据库管理时代

20 世纪 90 年代以后，数据库逐步扩充至非事务处理领域与数据分析领域。此外，网络与互联网的出现也使传统关系数据库应用受到影响，此时需对关系数据库管理系统实行必要的改造与扩充，内容包括：

1) 引入面向对象概念建立对象关系数据库管理系统以适应非事务处理领域的应用。

2) 扩充数据交换能力以适应数据库在网络及互联网环境中的应用。

3) 引入联机分析处理概念建立数据仓库以适应数据分析处理领域的应用。

这三种扩充功能目前已成为关系数据库系统发展趋势。

数据管理变迁的全貌可用图 1-4 表示。

本书将主要介绍数据库管理，重点介绍关系数据库管理，同时对后关系数据库管理也给予适当的

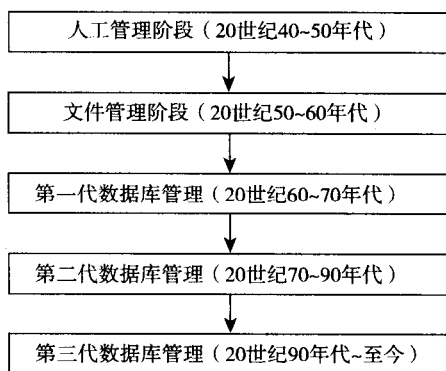


图 1-4 数据管理变迁示意图

关注与介绍。

1.6 数据管理工具与产品

当前数据库管理的主流产品均是后关系数据库管理产品(亦称关系扩充型产品),按其规模大致可分为大型、中型、小型及桌面式四种。下面介绍这四种数据库管理的代表性产品。

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

Oracle 数据库管理系统最新的版本是 Oracle9i, 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, 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、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 Sybase Adaptive Server Enterprise 及 Sybase Adaptive Server Anywhere, 它是一种关系扩充型产品, 在我国主要应用于铁道部门、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, 它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(如 Windows 操作系统、ODBC、OLE DB 接口及 ASP、C#、VB、VC 等开发工具)协调一致。该产品在我国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 Office 产品套件中的 Access, 它是关系模型产品, 但能适应 Web 环境。该产品以数据库为核心, 并有多种配套开发工具, 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>。

1.7 数据处理

当今, 由于数据与人类社会的密切关系, 数据处理已成为计算机应用的重要内容。在本书中, 数据处理主要指的是数据库中数据的应用。

1.7.1 数据处理的环境

在数据处理中,数据存放于计算机中,用户应用数据是通过访问数据库而实现的。而这种访问是在一定环境下进行的,随着计算机技术的发展,数据应用环境也不断变化,迄今为止一共有四种不同的环境,它们是:

1)人机直接交互式环境:在这种环境下,用户为操作员,由操作员直接访问数据库中的数据,这是一种最为原始与简单的访问方式,在数据库发展的初期就采用此种方式。

2)单机、集中式环境:在这种环境下,用户为应用程序,应用程序在机器内(单机)访问数据库中数据,这种访问方式在20世纪70~80年代较为流行,也是一种简单的访问方式。

3)网络分布式环境:在计算机网络出现后,数据访问方式出现了变化。在此种环境中,数据与用户(应用程序)可分处网络不同结点,用户使用数据时可通过接口调用的方式,这种方式目前应用广泛。

4)互联网环境:在当前互联网时代,用户是以互联网中的XML为代表,而数据访问方式则是XML与数据库间的接口调用方式。这种方式也是目前广泛应用的方式。

目前,这四种数据应用环境及访问方式都普遍存在,它为数据应用提供了多种应用手段。

1.7.2 应用开发

数据应用是需要开发的,数据应用开发可分为三个部分,它们是:

1)数据组织的设计:数据应用开发的首要任务是设计一个适合应用需要的数据结构以存储应用使用的数据。

2)数据库应用系统的开发:为方便数据应用,必须开发一个系统,该系统是在一定应用环境下,采用一种合理的数据结构并且与一定的硬件平台、基础软件平台及数据管理软件相结合,具有大量结构化数据与应用处理程序并且有一个友好的可视化界面。这种系统就是数据库应用系统,它可为特定的数据应用提供全面的服务。

3)数据应用的管理:数据库应用系统是数据应用的在计算机中的一种表现形式。数据库应用系统中的数据需要进行管理并不断进行维护与改造,使之能更好为应用服务,这就是数据应用的管理。

数据应用的管理一般由一组人完成,它们称为数据库管理员。

1.7.3 数据处理的应用领域

数据处理领域的应用范围很广,但是一般集中在如下三个方面:

(1)传统的事务处理

事务处理是数据应用的主要领域,也是最传统的领域,它以数据处理为主并具有数据结构简单、事务短、数据操作类型少的特点,目前主要用于电子商务、客户关系管理、企业资源规划以及管理信息系统等应用中。这种应用主要以关系数据库为支撑。

(2)非传统事务处理

非传统事务处理是近期的数据应用领域,它也以数据处理为主,但数据结构复杂,无法