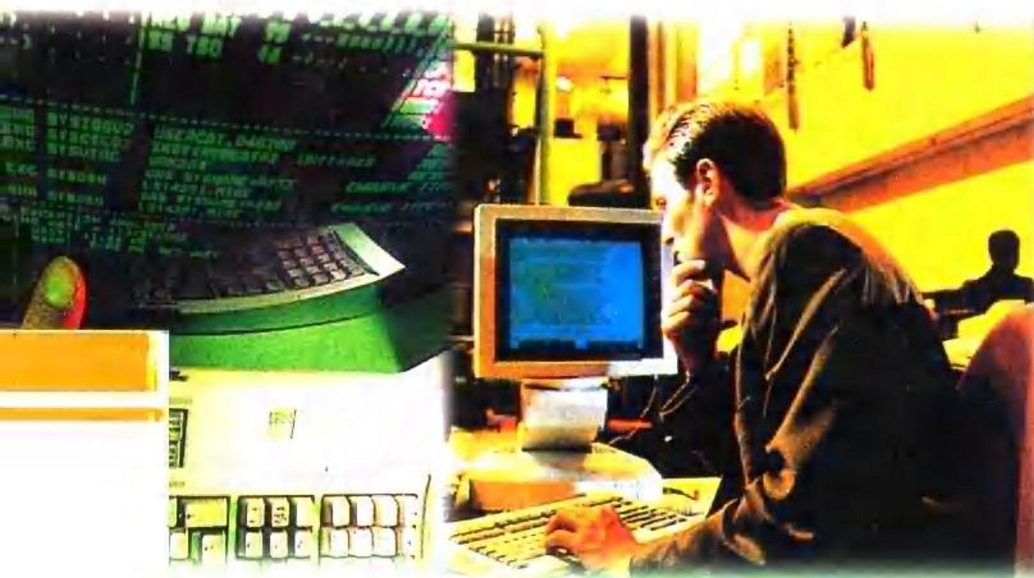




上海“九五”重点图书出版规划项目  
全国名校计算机专业核心课程教材

# 数据库系统原理

徐洁磐 编著



上海科学技术文献出版社

00010170

上海“九五”重点图书出版规划项目  
全国名校计算机专业核心课程教材

# 数据库系统原理

徐洁磐 编著

上海科学技术文献出版社

## 图书在版编目(CIP)数据

数据库系统原理/徐洁磐编著. -上海:上海科学技术文献出版社,1999.8

ISBN 7-5439-1438-7

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(1999)第 30125 号

### 数据库系统原理

徐洁磐 编著

\*

上海科学技术文献出版社出版发行

(上海市武康路 2 号 邮政编码 200031)

全国新华书店经销

上海科技文献出版社昆山联营厂印刷

\*

开本 850×1168 1/32 印张 11 字数 305 000

1999 年 7 月第 1 版 1999 年 7 月第 1 次印刷

印数: 1—2 100

ISBN 7-5439-1438-7/T·553

定价: 20.00 元

上海“九五”重点图书出版规划项目

## 全国名校计算机专业核心课程教材

### 编辑委员会

|    |     |         |          |
|----|-----|---------|----------|
| 顾问 | 陈火旺 | 中国工程院院士 | 国防科学技术大学 |
|    | 盛焕烨 | 教授      | 上海交通大学   |
|    | 顾冠群 | 中国工程院院士 | 东南大学     |

|    |           |    |        |
|----|-----------|----|--------|
| 主任 | 左孝凌       | 教授 | 上海交通大学 |
| 编委 | (按姓氏笔画排序) |    |        |

|     |    |        |
|-----|----|--------|
| 朱洪  | 教授 | 复旦大学   |
| 刘瑞挺 | 教授 | 南开大学   |
| 何炎祥 | 教授 | 武汉大学   |
| 招兆铿 | 教授 | 复旦大学   |
| 金廷赞 | 教授 | 浙江大学   |
| 徐洁磐 | 教授 | 南京大学   |
| 谢康林 | 教授 | 上海交通大学 |

|    |     |     |             |
|----|-----|-----|-------------|
| 策划 | 方金善 | 副编审 | 上海科学技术文献出版社 |
|    | 项暑烽 | 编审  | 上海科学技术文献出版社 |
|    | 廖恬  | 高工  | 上海科学技术文献出版社 |

# 总 序

近 20 年来,计算机学科的发展,促使现代科学迅速崛起。当前计算机技术,已经成为实时控制、信息处理、通讯、企事业管理,以及社会生活各个方面无所不在,无所不用的必不可少的实用工具。计算机技术的应用,冲破了传统学科的分类,例如,经济、艺术、法律、农林、医学等各种学科,都需依赖于计算机的应用,除了各自领域的专业实施外,应用计算机已是各个专业提高效率,发挥潜能,促进发展,与专业息息相关的有效手段。同时,计算机应用的拓广,使计算机科学与现代科学技术相互结合,千丝万缕。当前,计算机专业教育中面临的最大问题,是教学内容远远滞后于应用实践。当前,个人电脑已深入到各行各业,网络、多媒体等在各方面应用普遍,但是在计算机专业教育中,这些内容还难成主流,而一些基础课程,如数据结构、数据库,以及计算机原理与组成等都还不能涉及前沿,有些课程内容,长期滞留在七八十年代的水平。此外,计算机发展,已渗透到很多相关学科,如通讯、视像、激光、生物、化学、管理等各种相关领域,计算机教学的内容颇具知识爆炸之势,因此如何去芜存精,深厉浅揭是当前计算机教育改革的当务之急。本着上述观点,近年来,国内一些院校,曾对部分课程作了分析、实验、论证。大家认为当前世界计算机发展过程中,存在四大流行趋势,即是,面向对象、并行与分布处理、多媒体结构以及网

络运行。如何在教学内容中,反映这四方面的应用是刻不容缓,势所必然之事。

为此,我们经过筹划,组织了这套“全国名校计算机专业核心课程教材”(简称“名校教材”)。“名校教材”,并非都是名家,名作,它只是反映了国内一些名校在教学改革中的一些思想、举措。例如,当前计算机高级语言是大学的最基础先行课程。自C语言以后,很多异军突起,先有C++面向对象,又有VB,VC方便简捷,各领一方;网络流行之后,又有JAVA风靡一时,面对基础语言课程,观点各异,仁智相见,难执一词,难成共识。为此有人建议,语言应自概论讲起,分析语法结构,掌握语言构成规律,读通语言文本,那么任何计算机高级语言,都可举一反三,触类旁通。这种以结构规律来学会应用的方法,就是以不变应万变,任其千变万化,万变不离其宗。这种抓住本质,适应瞬息变化的拓展学科方法,在计算机的专业教育中是极具典型和富于启迪的。本丛书中,还有些教材,力图反映当前计算机学科中的最新进展,如数据库、操作系统两书都在详述基础概念基础上分类介绍了面向对象,并行与分布技术,以及网络、多媒体等实施的内容,使读者既能建立深厚基础,又能高屋建瓴,接近最新领域。

“名校教材”,目前已列为上海“九五”重点图书出版规划项目,并得到国内著名院校的教授、专家支持。

这套丛书,立意是着重基础,反映导向,注重实践。希望每种教材,能有创意,能具共鸣,能被接受,能予推广。但是另一方面,我们也意识到,由于各校情况各异,作者观感不一,理解角度有所不同,所以对教材的选用和编

著,还不易一致认同,我们只希望这套教材不落窠臼,在反映当前学科动向,促进学以致用等方面,能起到推波助澜作用。希望有关院校能根据本身条件,积极使用,参与讨论,以使各书能够不断修改,日臻完善。

最后我们感谢中国高等学校计算机教育研究会对于“名校教材”策划的各种支持,感谢上海科学技术文献出版社为本丛书的立题、报审、出版等所付出的艰辛和努力。

左孝凌

1997.7

# 前 言

数据库技术在我国日渐普及,数据库应用也日益拓宽,据不完全统计,它在计算机应用中占有的百分比已超过 65%,数据库不仅在传统的商业领域中、事务处理领域中发挥着主要作用而且在非传统应用中也起到越来越大作用,如在工程领域中、在数据集成领域中、在统计领域中以及在图形、图像、声音等多媒体领域中都起着重要的数据支撑作用。

随着应用的需要以及技术的成熟,数据库本身也起着显著的变化,关系数据库技术已成为当前数据库发展的主流,层次与网状数据库已退出舞台,成为历史产物,与此同时,分布式数据库、面向对象数据库以及知识库等新型数据库系统正在兴起。

鉴于以上情况,有必要编写一本与数据库新技术有关材料以适应数据库应用的发展,以适应面向 21 世纪的计算机应用的需要的图书,本书正是为此目的而编写,它包括数据库系统的最新内容,其主要有如下几点:

(1) 突出以关系数据库技术为主要内容,删减与压缩层次与网状数据库内容。

(2) 增加分布式数据库、面向对象数据库以及知识库系统内容。

(3) 增加数据库设计内容。

以上述三点为基础,再通过结构上的调整从而构成



一本具有能适应最新数据库技术发展的书稿。

本书可作为高等学校数据库课程教材,也可适用于数据库应用开发人员参考。本书重点在于阐述数据库系统的基本原理与基本技术而并不是某种数据库软件的操作手册。

本书内容叙述深入浅出,文字浅显易懂,组织结构严密,这是本书的一大特色。

本书在编写过程中得到王银根副教授、毛正年老师的帮助,在此特表示感谢。

由于作者水平有限,书中难免有不少缺点和错误,殷切希望读者批评指正。

作者 1998.10 南京

# 目 录

## 第一章 概 论

|                     |       |
|---------------------|-------|
| 1.1 数据与数据库系统.....   | ( 1 ) |
| 1.2 数据库系统的发展.....   | ( 3 ) |
| 1.3 数据库系统的基本特点..... | ( 4 ) |
| 1.4 数据库系统的基本内容..... | ( 8 ) |
| 1.5 数据库系统的工作过程..... | (10)  |
| 习 题 .....           | (11)  |

## 第二章 数据模型

|                     |      |
|---------------------|------|
| 2.1 数据模型的基本概念.....  | (12) |
| 2.2 数据模式.....       | (13) |
| 2.3 数据模型的四个世界.....  | (16) |
| 2.4 概念世界与概念模型.....  | (17) |
| 2.5 信息世界与逻辑模型.....  | (34) |
| 2.6 计算机世界与物理模型..... | (50) |
| 习 题 .....           | (52) |

## 第三章 关系数据库系统

|                           |      |
|---------------------------|------|
| 3.1 关系数据库系统概述.....        | (54) |
| 3.2 关系数据库系统的衡量准则.....     | (56) |
| 3.3 关系代数——关系模型数学理论之一..... | (58) |
| 3.4 关系演算——关系模型数学理论之二..... | (71) |
| 习 题 .....                 | (75) |

## 第四章 关系数据库系统标准语言 SQL

|                      |      |
|----------------------|------|
| 4.1 SQL 概貌 .....     | (77) |
| 4.2 SQL 数据定义功能 ..... | (79) |
| 4.3 SQL 数据操纵功能 ..... | (81) |
| 4.4 视图 .....         | (93) |
| 4.5 嵌入式 SQL .....    | (96) |
| 4.6 动态 SQL .....     | (98) |
| 4.7 SQL 的特点 .....    | (99) |
| 习 题 .....            | (99) |

## 第五章 数据库的物理组织

|                     |       |
|---------------------|-------|
| 5.1 概 论 .....       | (101) |
| 5.2 数据库中数据分类 .....  | (101) |
| 5.3 数据库的物理结构 .....  | (102) |
| 5.4 数据库存储空间组织 ..... | (105) |
| 习 题 .....           | (106) |

## 第六章 关系数据库的查询优化

|                      |       |
|----------------------|-------|
| 6.1 概 述 .....        | (107) |
| 6.2 查询优化的可能性 .....   | (108) |
| 6.3 关系代数等价变换规则 ..... | (110) |
| 6.4 查询优化策略与算法 .....  | (112) |
| 习 题 .....            | (114) |

## 第七章 数据库保护

|                      |       |
|----------------------|-------|
| 7.1 数据库的安全性保护 .....  | (116) |
| 7.2 数据库的完整性保护 .....  | (123) |
| 7.3 数据库的并发控制技术 ..... | (126) |
| 7.4 数据库的恢复技术 .....   | (130) |

|           |       |
|-----------|-------|
| 习 题 ..... | (132) |
|-----------|-------|

## 第八章 关系数据库规范化理论

|                       |       |
|-----------------------|-------|
| 8.1 概 述 .....         | (133) |
| 8.2 规范化理论 .....       | (137) |
| 8.3 规范化所引起的一些问题 ..... | (153) |
| 习 题 .....             | (154) |

## 第九章 数据库设计与数据库管理

|                      |       |
|----------------------|-------|
| 9.1 数据库设计概述 .....    | (155) |
| 9.2 数据库设计的需求分析 ..... | (156) |
| 9.3 数据库概念设计 .....    | (158) |
| 9.4 数据库的逻辑设计 .....   | (166) |
| 9.5 数据库的物理设计 .....   | (172) |
| 9.6 数据库管理 .....      | (174) |
| 9.7 DBA .....        | (176) |
| 习 题 .....            | (177) |

## 第十章 分布式数据库系统

|                                 |       |
|---------------------------------|-------|
| 10.1 概 述 .....                  | (178) |
| 10.2 分布式数据库系统的分类与结构 .....       | (182) |
| 10.3 数据分布策略 .....               | (193) |
| 10.4 查询分解和优化 .....              | (197) |
| 10.5 SYBASE 数据库管理系统的分布式功能 ..... | (211) |
| 10.6 一个典型的分布式数据库系统例子 .....      | (217) |
| 10.7 国际互联网上的数据库分布方式 .....       | (219) |
| 习 题 .....                       | (220) |

## 第十一章 面向对象数据库系统

|                         |       |
|-------------------------|-------|
| 11.1 面向对象概念与数据库系统 ..... | (221) |
|-------------------------|-------|

|      |                      |       |
|------|----------------------|-------|
| 11.2 | 面向对象方法·····          | (223) |
| 11.3 | 面向对象方法介绍·····        | (226) |
| 11.4 | 面向对象数据模型·····        | (233) |
| 11.5 | 面向对象数据库管理系统·····     | (234) |
| 11.6 | 面向对象数据库的应用·····      | (250) |
| 11.7 | 面向对象数据库管理系统产品介绍····· | (262) |
|      | 习 题·····             | (266) |

## 第十二章 知识库系统

|      |                 |       |
|------|-----------------|-------|
| 12.1 | 概 述·····        | (267) |
| 12.2 | 知识库系统原理·····    | (277) |
| 12.3 | 几个典型的知识库系统····· | (294) |
| 12.4 | 知识库系统的应用·····   | (298) |
|      | 习 题·····        | (310) |

## 第十三章 多媒体数据库、工程数据库及数据仓库

|      |                |       |
|------|----------------|-------|
| 13.1 | 多媒体数据库·····    | (311) |
| 13.2 | 工程数据库·····     | (318) |
| 13.3 | 数据分析与数据仓库····· | (323) |
|      | 习 题·····       | (329) |

|      |       |       |
|------|-------|-------|
| 参考文献 | ····· | (331) |
|------|-------|-------|

# 第一章 概 论

## 1.1 数据与数据库系统

计算技术的发展,计算机应用的进一步深入与拓展,使得数据在计算机应用中的地位与作用日益重要,它在商业中、事务处理中占有主导地位,近年来在统计领域,在图形、图像、声音等多媒体领域以及在智能领域等的地位与作用也变得十分重要,因此,数据已成为构成一个计算机系统的重要支持。数据同时也成为一种财富与信息资源,它具有独立地向社会提供信息的能力。

数据(data)是计算机领域中的专用术语,它属软件范畴,它是软件中程序加工的客体,在计算机中一个软件只有有了程序与数据才构成完整的运行实体。数据的表现形式很多,除常用的数字、文字、日期等外,还包括图像数据、图形数据、声音数据等形式。数据一般有型(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,数据的值给出了符合给定型的数值。在软件系统中为表示上的需要,经常须将多个数据组合在一起,它们构成特定的数据框架称为数据结构(data structure),在特定条件下也可称为数据模式(data schema)。数据存放于计算机中的存储体内,少量的、常用的数据一般存放于计算机的内存储器中,这些数据由于受到内存物理因素的影响不能长期存放,只能作短时、临时的储存,故称为临时性数据或挥发性数据,而大量的、非经常性使用的数据一般存放于计算机中次级存储器(目前一般为磁盘或光盘)中,这些数据可在磁盘、光盘中长期保存,故称为持久性数据。数据在计算机软件系统中的作用和地位是随着计算机应用的扩大而逐渐发生变化的。首先是数据量的变化,在计算机应用中对数据量的需求由少量到大量进一步到海量;其次是数据结构的变化由简单到复杂;最后是地位的变化,过去,计算机软件系统是以

程序为主体,数据是程序的附属品,数据在系统内是分散、凌乱的,这也造成了数据管理上的混乱,如数据冗余高、数据的不一致以及数据不安全等多种弊病。近十余年来,数据在软件系统中的地位产生了改变,系统中大多以数据为主体而程序已退居于附属地位。在此种情况下,需要对数据作集中统一的管理,并以程序对数据的共享为目标,这就出现了数据库系统(database system)。数据库系统是以海量的、复杂结构的、持久的、共享的数据的统一管理为目标的计算机应用软件整体,它由下面几个部分组成:

1. 数据库(database,简称 DB)

数据库是数据的集合,它具有一定的组织形式并存放于统一的存贮介质上,它是多种应用数据集成,并可被应用所共享。

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

数据库管理系统是管理数据库的机构,它是一种软件,它负责:

- (1) 数据库中数据的组织;
- (2) 数据库中数据的维护以及保护数据不受破坏;
- (3) 数据库中数据操作。

3. 数据库应用(database application)

数据库提供了数据共享的功能,因此多个应用程序可以使用数据库,应用程序使用数据库是通过 DBMS 实现的。

4. 数据库管理员(database administrator,简称 DBA)

由于数据库具有共享性,因此对数据设计、规划、协调需专职人员负责进行此项工作,这些人员或集体统称为数据库管理员。

5. 系统平台(system platform)

数据库、数据管理系统及数据库应用都是建立在一定计算机平台上的,这种平台包括计算机及操作系统两部分,近期还包括计算机网络。

上述五个部分构成了一个完整的数据库系统,有时为方便起见也可简称为数据库。图 1-1 给出了数据库系统五个部分的

关系。

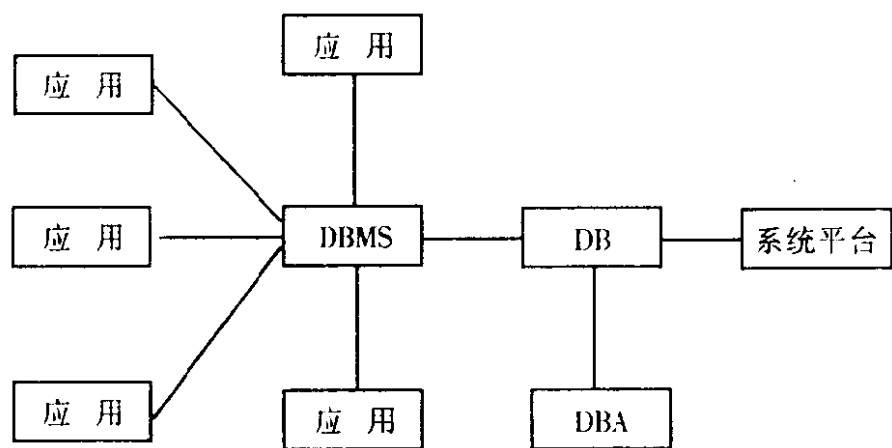


图 1-1 数据库系统示意图

## 1.2 数据库系统的发展

数据库系统的发展分为若干个阶段：

### 1. 文件系统阶段

文件系统是数据库系统的初级阶段，它出现于 20 世纪的 50 年代至 60 年代中，它提供了简单的数据共享与数据管理功能，但它无法提供完整统一管理与数据共享能力，由于它的功能较为简单，一般它均附属于操作系统而并不成为独立的软件。目前，一般已将它看成为仅是数据库系统雏形而并不是真正的数据库系统。

### 2. 初级数据库系统阶段

自 20 世纪的 60 年代起各种数据库系统相继问世，它们是层次模型与网状模型的数据库系统，它们为统一管理与共享数据提供了有力支撑。这个时期由于数据库系统的蓬勃发展，而形成了著名的“数据库时代”。当然，这两种类型的数据库系统也有其一定的不足，最主要的是它们均脱胎于文件系统，因此，受文件中的物理结构影响较大，用户在使用数据库时需要对数据的物理结构有详细的了解，这对使用数据库带来了诸多的麻烦。同时，数据库中表示数据模式的结构方式过于繁琐，这也影响了数据库应用中越来越多的对数据结构的复杂要求的实现。



### 3. 关系数据库系统阶段

关系数据库系统的理论出现于 20 世纪 70 年代初期,其系统形成于 70 年代后期并在 80 年代得到了充分的发展,它具有简单的结构方式与较少的物理表示,使用与操作又极为方便。因此在 80 年代它逐步取代层次型和网状型数据库成为占主导地位的数据库,到目前为止,关系型数据库系统仍占有数据库应用的主要地位。

### 4. 高级数据库系统阶段

由于计算机应用的发展,计算机已从传统的科学计算、事务处理等领域逐步扩展到工程设计、统计、人工智能、多媒体、分布式等领域。这些新的领域需要有新的数据库支撑,而传统关系数据库系统是以商业应用、事务处理为背景而发展起来的。它并不完全适用于新领域的应用,因此,需要有新的数据库系统以满足不同领域的要求。自 20 世纪 80 年代中期开始,各种适应不同领域的新型数据库系统不断涌现,如工程数据库、多媒体数据库、CAD 数据库、图形数据库、图像数据库、智能数据库、分布式数据库以及面向对象数据库等,其中特别是面向对象数据库系统由于其通用性强,适应面广而受到青睐。预计 21 世纪初将会形成多种数据库系统共同支撑应用的局面。当然在那时,关系数据库系统仍将是其中一支强有力的主力军。

在我国,数据库的发展也经历了上述的类似的四个阶段,但是略有不同。由于我国计算机应用发展较技术先进国家略晚一步,因此,文件系统阶段较为短暂,而初级数据库系统阶段基本没有经历,自 80 年代中期起即进入关系型数据库系统阶段,特别是以 dBASE-III 与 Foxbase 为代表的初级关系型数据库系统阶段和近年来以 ORACLE, Sybase 为代表的成熟的关系型数据库系统阶段。目前,我国也已逐渐向高级数据库系统方向靠拢。

## 1.3 数据库系统的基本特点

下面介绍数据库系统的几个基本特点,同时也给出了数据库