



国家科学技术学术著作出版基金资助出版
国家九五重点图书

计算机应用技术前沿丛书



工程数据库管理系统

董金祥 孙建伶 编著

浙江大学出版社

国家科学技术学术著作出版基金资助出版

工程数据库管理系统

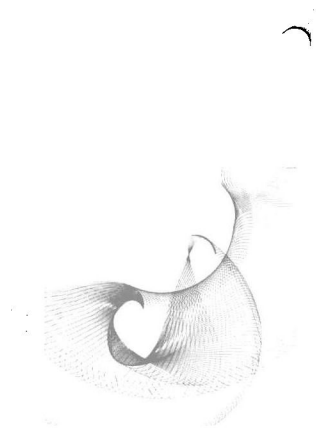
浙江大學出版社

国家九五重点图书

工程数据库管理系统

董金祥 孙建伶 编著

计算机应用技术前沿丛书



图书在版编目 (CIP) 数据

工程数据库管理系统 / 董金祥, 孙建伶编著. —杭州:
浙江大学出版社, 2000.12
ISBN 7-308-02218-8

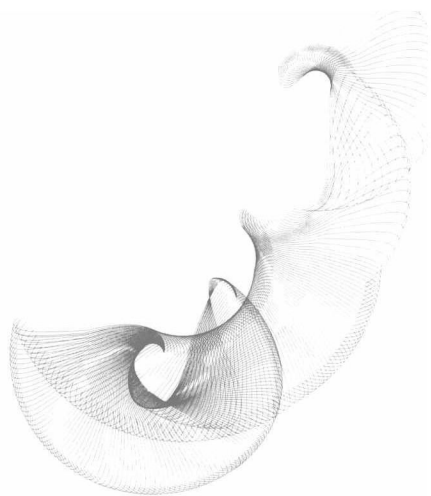
I. 工... II. ①董...②孙... III. 工程数据库-数
据库管理系统 IV. TP392

中国版本图书馆 CIP 数据核字 (2000) 第 12149 号

总 责 编 陈晓嘉 梁 兵
封面设计 俞亚彤
版式设计 孙海荣
责任绘图 张作梅
责任出版 李慧华

出版发行: 浙江大学出版社
(杭州浙大路 38 号 邮政编码 310027)
(E-mail: zupress@mail.hz.zj.cn)
(网址: <http://www.zjupress.com>)

责任编辑: 梁 兵 李玲如
排 版 者: 浙江大学出版社电脑排版中心
印 刷: 浙江印刷集团公司
经 销: 浙江省新华书店
开 本: 889mm×1194mm 1/16
印 张: 18.5
字 数: 472 千
版 印 次: 2000 年 12 月第 1 版第 1 次印刷
书 号: ISBN 7-308-02218-8/TP·190
定 价: 37.00 元



主 编 潘云鹤

副主编 陈 纯

顾 问(以姓氏笔划为序)

王 选 李三立 李国杰 李衍达

沈昌祥 张 钹 张效祥 汪成为

何志均 杨芙清 金怡濂 陈火旺

胡启恒 唐泽圣 戴汝为



序

20 世纪是一个科技、经济空前发展的时代,从世纪初相对论、量子理论的创立到今天以信息产业为龙头的高科技产业成为经济发展的第一支柱,人类社会发生了根本性的变革。而在这场以科学技术为社会发展直接动因的变革中,意义最深远、影响最广泛的就是计算机及其相关技术的发展和运用。

在过去的 50 年里,计算机已从最初的协助人类进行精密和复杂运算的单一功能的运算器发展成为能够模拟人类智慧、感觉,可以适应环境、具有多种功能的设备;计算机及其技术的应用已从尖端科学、国防工业延伸到了人类活动的各个领域。它不仅充实和革新了传统的科学与技术研究手段,带来了新的方法和理念,成为科学技术谋求新发展的沃壤,而且作为信息处理的核心工具和技术支撑起庞大的信息产业,进而改变了全球的产业格局。

目前,一个由遍布全世界的计算机连接起来的跨越国界、巨大而高速的多媒体信息网络正在形成,因此,不仅信息产业本身,而且各行各业都将面临计算机应用技术研究和发展的重大课题。我们针对国内外有关计算机应用技术的最新进展,根据浙江大学及国内其他学者在计算机辅助设计系统和方法、计算机动画、设计思维模拟、自主式智能系统、智能决策系统、软件智能化技术、工程数据库管理系统、图象与声音

压缩技术等方面取得的成果,组织编撰了《计算机应用技术前沿》丛书。我们希望,通过这套书,为从事相关领域研究工作的专家和学者提供一些有参考价值的方法和技术,为普及、推广,进而开拓计算机及其网络的新的应用技术作出一点贡献。

潘云鹤

一九九九年十月



前言

80 年代,关系型数据库系统在商务信息管理领域占据了主导地位。与此同时,人们开始试图将信息管理技术推向更广的应用领域,如机械工程、软件工程、VLSI 设计、建筑、科学计算可视化等。在此推广过程中,关系型数据库的弱点日益显现,它无法很好地支持上述列举的所谓的“非标准”数据库应用,这促使数据库研究者进一步探索支持工程应用的、具有更强功能的新型数据库系统,即工程数据库系统。

工程数据库系统的探索方向大致有两种:

1. 改良型工程数据库系统:在现有的关系数据模型的基础上加以扩充。

2. 变革型工程数据库系统:建立新型数据模型,如面向对象模型。

由于数据库最早是为商务信息管理开发的,因此,许多数据库著作都论述了数据库在商务管理方面的应用与技术。迄今为止,还没有一本书全面论述工程应用领域中的数据库概念、工具、设计方法和实现技术。本书试图就上述这些方面展开讨论。

本书是作者多年来在工程数据库领域教学与科研成果的总结。在论及改良型工程数据库主要内容的同时,重点论述的是面向对象数据库。面向对象数据库从面向对象程序设计的语言中借鉴了面向对象的思想来开发能支持“非标准”应用的数据库系统。面向对象数据库是工程数据库的一种实现方式。

在本书中涉及面向对象数据库管理技术的各章中,均以一個面向对象的数据库系统 OSCAR 作为例子。OSCAR 是浙江大学人工智能研究所投入大量人力物力,经过十年时间自行设计开发的,拥有自主版权。OSCAR 的功能覆盖了工程对象建模、存储管理、版本管理、事务管理、查询处理等数据库的主要功能。OSCAR 支持 STEP 标准,提供了 STEP 的 EXPRESS 语言建模和 SDAI 数据存取。OSCAR 还是一个可以和程序设计语言无缝接口的工程数据库管理系统,从数据建模到数据操纵均不存在阻抗不匹配问题。OSCAR 是目前国内功能最齐全、实用化程度最高的系统之一,已在多家单位投入实际应用。

本书由以下三个部分组成:

第一部分(1~2 章)是概述,从总体角度介绍数据模型从层次模型、网状模型、关系模型到面向对象模型的发展,以及数据库体系结构从集中式系统到客户/服务器系统、分布式系统、多数据库系统、联邦数据库系统的发展;分析关系型数据库应用于工程领域的局限性,工程数据库系统为克服这些缺陷所采用的实现途径,以及面向对象工程数据库管理系统的特征。

第二部分(3~11 章)结合 OSCAR 系统,全面介绍工程数据库管理系统的各种功能及其系统实现技术,包括体系结构、编程界面、查询处理、物理对象管理、并发控制、恢复、安全管理、版本管理、数据交换等各个方面。

第三部分(12 章)介绍了 OSCAR 在 CAD/CAPP/CAM, PDM, CIMS 中的应用。

本书的写作分工如下:董金祥完成 1, 2, 5, 10, 11, 12 章;孙建伶完成 3, 4, 6, 7, 8, 9 章。全书由董金祥统稿。在成书过程中,周波博士提出了许多建设性的意见,寿宇澄、冯柯、顾云苏、姚建中、陈俊等做了大量资料整理和文字工作,在此一并表示感谢。

最后,衷心感谢国家科学技术部设立的国家科学技术出版基金对本书出版的支持。

目 录

1 数据库系统的发展	1
1.1 概述	1
1.2 数据模型的发展	9
1.3 数据库体系结构的发展	25
2 工程数据库管理系统	32
2.1 关系型数据库应用于工程领域的局限性	33
2.2 工程数据库管理系统的实现途径	37
2.3 面向对象的工程数据库管理系统	51
3 工程数据库管理系统的体系结构	60
3.1 工程数据库的客户/服务器体系结构	60
3.2 客户/服务器环境下的缓冲一致性问题及并发控制方法	66
3.3 一种改进的工程数据库管理系统体系结构	68
3.4 OSCAR 的体系结构	72
4 数据库与程序设计语言的紧密结合	83
4.1 数据库与程序设计语言之间的失配	83
4.2 解决失配问题的途径	85
4.3 面向对象数据库程序设计语言的持久性模型	89
4.4 两种文化传统的冲突	90
4.5 面向对象数据库无缝的 C++ 应用编程界面	91
4.6 ODMG C++ 绑定	97

5	面向对象数据库的查询处理	115
5.1	概述	115
5.2	面向对象查询处理的体系结构	116
5.3	面向对象查询优化技术	121
5.4	查询执行器的设计	133
5.5	面向对象数据库中的索引技术	136
6	物理对象管理	145
6.1	概述	145
6.2	对象标识	146
6.3	指针调和	152
6.4	数据聚簇	166
6.5	废料收集	176
7	工程数据库的并发控制	187
7.1	数据库的一致性和常规并发控制技术	187
7.2	新的并发控制需求	188
7.3	基于事务语义的并发控制	189
7.4	基于抽象数据类型语义的并发控制	190
7.5	支持协作过程的并发控制	191
7.6	基于语义的并发控制评价	193
7.7	支持设计过程的事务管理	194
8	数据库恢复技术	201
8.1	数据库系统的故障	201
8.2	基于日志的恢复	202
8.3	ARIES 恢复策略	208
8.4	客户/服务器环境下的恢复	214
8.5	B 树的恢复	215
8.6	OSCAR 的恢复技术	216
9	工程数据库的安全控制	219
9.1	关系数据库的安全控制	219
9.2	面向对象数据库的安全控制	222
9.3	数据库的多级安全控制	225

9.4	OSCAR 的权限管理	228
10	版本管理	232
10.1	版本和配置	232
10.2	版本管理的基本功能	233
10.3	版本管理模型	234
10.4	OSCAR 的版本管理	242
11	工程数据库的数据交换与 STEP 标准	245
11.1	工程领域的的数据交换要求	245
11.2	STEP 概述	248
11.3	STEP 的技术原理和组成结构	250
11.4	EXPRESS 语言	253
11.5	STEP 实现方式	254
11.6	OSCAR 支持下的 STEP 标准实现	256
12	工程数据库管理系统 OSCAR 的应用	263
12.1	OSCAR 特点简介	263
12.2	以 OSCAR 为核心的集成化 CAD/CAPP/CAM 系统	266
12.3	OSCAR 在 PDM 中的应用	273
12.4	OSCAR 在 CIMS 中的应用	277

1 数据库系统的发展

我们的世界正日趋复杂,知识体系的内涵日益丰富。据估计,每过五年,关于技术、科学、商业等各方面的知识总量就翻一番。因此,不要说全部掌握复杂的知识总体,就是仅仅检索相关信息也越来越难。越来越多的系统向数据密集型(data intensive application)方向发展,如管理信息系统、办公信息系统、银行信息系统、民航订票系统、情报检索系统等与日常生活息息相关的系统^[1],这些系统的共同特点是:

- (1)数据量大,需要大容量的数据存储器。
- (2)数据与应用程序相对独立,即在程序运行结束后,数据仍可永久地存在。
- (3)数据可以被共享。

因此,管理这些系统产生的大量、永久、共享的数据,越来越成为计算机应用研究的一个重要方面。在这种需求推动下,数据管理技术应运而生,并成为计算机科学中的一个重要分枝。它的最终目的是要达到数据和程序的分离,使程序设计人员和使用人员不必进行数据管理方面的工作,就可在一个数据安全、一致、共享的环境中进行工作。

1.1 概 述

50年代,在数据管理领域首先出现的是文件系统。直到目前,文件系统仍有着广泛的应用。在文件系统中,数据按其内容、用途、结构等组成若干命名的文件,文件之间通过文件名来相互区别。这样的数据管理体系能在一定程度上解决数据的存储和共享等问题,但随着应用领域对数据管理要求的日益提高,文件系统的缺点也逐步显露出来,主要表现为:

- (1)与应用系统的接口很不方便。应用程序的设计者必须对文件的逻辑和

物理结构有非常清楚和全面的了解;接口操作是打开、关闭、移位、读、写等底层文件操作,而真正对数据进行的处理都是在应用程序中进行的。不同的应用程序对数据的操作不可避免地会有重复,但这些共同的部分无法写成共享的程序代码。

(2) 文件系统造成数据冗余。由于不同文件之间的相同数据无法共享,而应用系统因各自功能的不同,对数据的视角和划分方式也会不同,因而不可避免地造成同一数据必须存放在不同文件中的情况。这种数据冗余给数据更改带来隐患,难以保证更改后的数据仍能保持一致。

(3) 数据与程序难以分离。文件系统中,文件结构的任何一次修改都有可能导致程序的修改。在现代软件越来越复杂的情况下,这将大大加重软件维护的工作量。

(4) 文件系统很难支持并发访问(concurrent access)。文件系统最初产生时的想法就是作为某一应用程序的附属部分,因而没有支持并发访问的机制。而现代的应用越来越强调并发,文件系统已经不能适应这种要求。

(5) 文件系统难以做到规范化。由于很多情况下,命名、编码、结构等是人为定义的,所以难以对文件系统做到统一管理。另外,数据的安全保密也不能得到保证。

从本质上说,文件系统由于不强调数据项的存在,因而出现全体数据“整进整出”的现象,同时,由于文件系统的总体构架中没有把数据和程序同等对待,而是把数据当成程序的附庸,因而无法从全局上对数据进行管理。

针对以上情况,人们逐步转向了数据库系统(database system),它的主要特点是数据统一管理和数据共享。数据库系统的一个首要思路是把数据当作与程序对等的实体来对待,是为所有应用程序服务的共享资源。当然,这种服务必须是有序的,并且具有较高层次的应用程序接口。在数据库系统中,这一功能是由数据库管理系统 DBMS(database management system)完成的。数据库管理系统使得用户不再需要关心具体的数据存储和实现细节,也不需要介入打开、关闭、读、写等低级文件操作,可以在更高层次上访问和观察数据,即从数据本身的内容上来考虑。数据存储结构的修改对应用程序也是透明的,不会要求程序做相应修改。由于数据的独立性,可以支持并发访问机制。同时,由于数据不是单纯存放在以“名”来区分的文件中,因此数据可以按照它们的本来含义进行分类和组织,有利于从全局的高度来管理数据,减少数据冗余。由于有了专门的数据管理系统,因此可在其中实现更多的功能,例如数据安全性(security)、完整性约束(integrity constraint)、数据恢复(recovery)、并发控制(concurrent control)、多种用户界面(multiple user interface)等。从根本上说,由于数据库系统把数据的管理作为一个相对独立的工作,因此有可能在这个领域发展专门的数据管理技术,最终为不同的应用系统服务,真正达到数据与程序分离的目的。

1.1.1 数据库、数据库管理系统和数据库系统

数据库系统的结构如图 1.1 所示。由图中可以看出,数据库是数据的汇集地,是存储数据的地方。数据以一定的组织形式组成数据库。数据库管理系统是管理数据库的软件,它负责实现数据库系统的各种功能。各应用程序通过数据库管理系统访问数据库,而数据库的规划、设计、协调、维护等工作则由数据库管理员 DBA(database administrator)来完成。数据库系统就是应用程序、数据库管理系统、数据库和数据库管理员的总和。

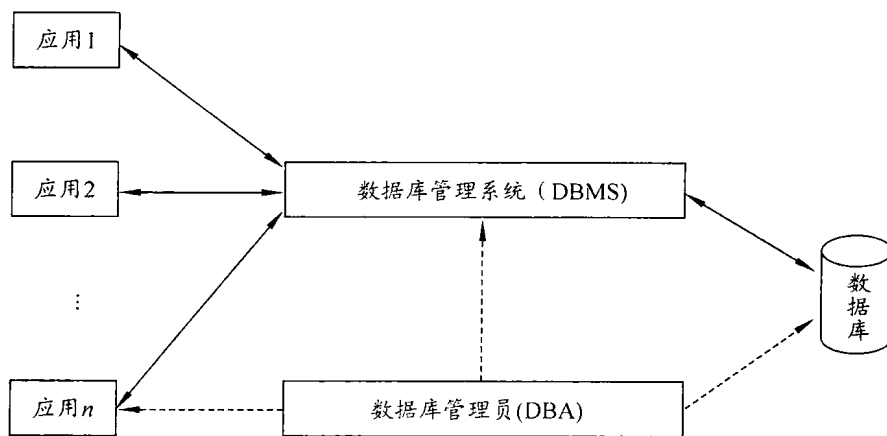


图 1.1 数据库系统

在数据库系统中,DBMS 是核心。整个数据库系统的功能大部分由 DBMS 的特点和功能决定。从应用的角度来研究数据库,主要也是从研究 DBMS 来入手。不同的 DBMS 的功能有共同之处,但实现方式和实现程度是不一样的,因而产生的效果也是不一样的。这样,在为某一应用领域选择数据库系统的时候,关键是考察 DBMS 的特点是否能满足领域的要求。以下方面的内容是几乎所有 DBMS 都覆盖的,但具体功能会有所不同。

高级的用户接口

DBMS 首先保证应用程序与数据库物理结构无关。对用户来说,关心的只是数据中所包含的信息,而不是数据的存放格式,也就是说只关心数据的逻辑结构。DBMS 负责数据的逻辑结构和物理结构之间的映射。

当应用程序访问数据库时,DBMS 负责寻找访问路径和访问策略,并检查访问权限。为此,DBMS 向应用程序和用户提供的非工程化的数据操纵语言 DML (data manipulate language)。应用程序和用户通过 DML 对数据进行操纵。这样,用户看到的是数据的逻辑结构,操纵用非过程化的 DML,保证应用程序不需关心细节,从而实现数据的独立性。

数据字典管理

在数据库中保存永久和共享的数据,因此数据的定义应独立于应用程序,永久性地保留在数据库中。这些信息就是数据字典。数据字典不但保存数据的逻辑属性,还包含数据的存储结构定义和访问、管理数据所必需的信息。

并发控制

越来越多的应用领域要求程序之间的共享。数据库系统作为不附属于单一应用程序的独立数据处理系统,必须提供并发控制能力,以实现高粒度的数据共享。

查询处理和查询优化

数据库数据的应用很多情况下是通过查询进行的。查询包括数据检索、修改和定义新的数据等,是用户对数据库所提出的各种访问请求。

数据库恢复功能

由于应用程序不再关心数据的存储,因此更要求数据库保证所存储的数据在各种情况下都是安全的,包括在发生故障的情况下,数据库能恢复到一致的状态。

完整性约束检查

数据库中的数据是永久和共享的,因此不但需要进行语法检查,还要进行语义检查。

数据在语义方面的约束就是完整性约束。完整性约束对确保数据的正确性,防止错误数据的进入是十分重要的。

访问权限控制

数据库是共享资源,任何应用程序和用户只能在拥有访问权限的前提下才能对数据进行存取。控制用户是否有权对数据进行某种操作的机制就是 DBMS 的访问权限控制。访问权限控制不仅能控制用户访问数据时进行的操作,还可以控制用户的访问范围。

DBMS 是数据库系统的核心,也是研究最多的技术凝聚点。在设计数据库系统时,选择什么样的 DBMS 是关键,但 DBMS 并不是数据库系统的全部。作为一个完整的数据库系统,还需要设计和载入数据、开发基于数据库的应用程序,以及进行数据库的维护、调整、用户管理等工作。数据库系统中的各种因素是相互影响和制约的,有时还需要在各方面之间进行折衷和妥协。DBMS 的设计和选择也要受其他方面的影响,要充分考虑应用的需求。

1.1.2 数据库的应用领域

可以说,几乎所有计算机应用领域都是数据库的应用领域。当前计算机的应用都趋向于大型化和全面化,对数据的管理也提出了较高的要求。在这种情况下,数据库的应用往往随着计算机应用的深入而被自然地引进。从数据库和应用程序的对应关系来看,数据库的应用领域可以分为传统领域和新领域。前者以结构化程序为主要应用程序,后者以新一代程序设计技术,主要是以面向对象技术为主要应用方向。

传统领域

从数据库发展历史上看,数据库的应用领域是随着计算机应用领域的扩展而扩展的。

数据库出现的 60 年代末期是程序设计进入结构化设计的时代:计算机的应用从单纯的科学计算向计算、推理、管理一体化方向发展,计算机越来越多地涉及到管理功能。总的来说,数据库应用的领域基本上也就是结构化程序设计的应用范围。这些领域是以事务处理为特征的。典型的应用领域有:工业数据管理、商用数据管理等。

新领域

随着计算机技术的发展,计算机的应用逐渐扩展到非事务处理领域。以面向对象程序设计为代表的新式程序设计使计算机应用达到一个新的境界。与传统领域相比,新领域的数据库有着数据量大、数据之间联系复杂的特点,因而,对数据库管理也提出了更高的要求。应用于新领域的数据库系统必须具有很强的数据抽象和表达能力、很高的数据操纵效率和灵活的应用程序接口能力。典型的数据库应用新领域有:计算机辅助设计(CAD)、计算机集成制造系统(CIMS)、专家系统、多媒体信息处理等。新的应用领域还对并发控制、数据库完整性约束等有专门的要求。

1.1.3 数据库与相关领域的关系

硬件

在所有的计算机应用中,硬件系统都是基础。所有的软件系统都建立在硬件的基础上,数据库系统也不例外。硬件的发展目前非常迅速,这极大地影响着软件系统的设计。例如,以往的软件设计都非常强调效率,而现在,由于硬件的飞速发展,效率问题的解决更多地依赖于硬件,而软件系统可以更加自由地解决应用