



数据库系统教程

(第二版)

陈有祺 主编

袁晓洁 徐雷 编著

计算机
大专
教材
系列

计算机大专教材系列

[新高职、高自考可选用]

南开大学出版社



数据库系统教程

(第二版)

陈有祺 主编
袁晓洁 徐 雷 编著

南开大学出版社

·天津·

内 容 提 要

本书系统、完整地讲述了当前数据库技术由基本原理到应用实践的主要内容。全书分为两大部分:第一部分是数据库系统的基本原理,由第1章到第7章组成,内容包括:数据模型、数据库系统构成、数据库管理系统、关系数据库、SQL语言、关系数据理论、关系系统的查询优化、数据库环境和数据库设计;第二部分是FoxBASE⁺和FoxPro for Windows程序设计语言,由第8、9章组成,介绍了它们的基本命令和程序设计技术。

本书是在总结了成功的数学方法和经验的基础上,综合数据库领域里重要的基础知识,并理论联系实际,将基本原理融汇贯通于数据库实用技术中编写而成的。它特别适合作为大学、大专、中专及成人教育等计算机专业的数据库教材,同时也可作为非计算机专业的工程技术人员和管理人员学习计算机基础知识的教科书。

图书在版编目(CIP)数据

数据库系统教程 / 陈有祺主编; 袁晓洁, 徐雷编著.
2版(修订版). — 天津: 南开大学出版社, 2000. 11
计算机大专教材系列
ISBN 7-310-00889-8

I. 数... II. ①陈... ②袁... ③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(2000)第 46592 号

出版发行 南开大学出版社

地址: 天津市南开区卫津路 94 号

邮编: 300071 电话: (022) 23508542

出版人 肖占鹏

承 印 南开大学印刷厂印刷

经 销 全国各地新华书店

版 次 2000 年 11 月第 2 版

印 次 2000 年 11 月第 5 次印刷

开 本 787mm×1092mm 1/16

印 张 15

字 数 368 千字

印 数 29001—34000

定 价 30.00 元

“计算机大专教材系列”编委会

主 编
副主编
编 委

陈有祺	吴功宜	王家骅	
朱瑞香	于春凡	孙桂茹	李 信
朱耀庭	周玉龙	辛运帏	刘 军
袁晓洁	李正明	何志红	裴志明
伍颖文			
张 蓓			

第二版出版说明

自本套教材系列出版以来,承蒙全国广大用户和读者的厚爱,至今已发行了数十万册,大部分教材已作了第三次印刷。在此期间,我们收到了广大读者许多宝贵的意见和建议,为我们修改、补充教材内容提供了重要的依据,借此机会,我们表示衷心的感谢!

近几年来,计算机技术的发展日新月异。特别是国际互联网的蓬勃发展,上网用户逐年呈几何级数增加;多媒体技术的普及应用,使我们的工作和生活更加绚丽多彩;图文并茂的 Windows 系统,已代替传统的 DOS 系统展示在初学者的面前。有鉴于此,我们将这套教材作了较大的更新调整,除对原有各教材都作了不同程度的修改、补充外(例如,原来的《计算机应用基础》经大幅度增删修改后以新面貌《计算机基础与应用》面世;原《单片机及其应用》改为《单片微机原理及应用技术》等等),还在原系列中增加了《Internet 应用基础》、《多媒体技术应用基础》和《办公自动化基础教程》等新教材,以适应形势发展的需要。

另一方面,我国的高等教育事业也有了很大的发展。近年来,高等学校自学考试和高等职业技术教育吸引了众多的青年学子。在这些新型的教育体系中,教材的缺乏成为最紧迫的问题之一。根据这种社会需求,我们在这套教材中更加突出了深入浅出、学以致用原则,使得这套教材不仅适合于在校大专学生的需要,也适合高自考和新高职广大学生的需要。尽管如此,但由于我们的水平和经验都有很大的局限性,因而这套教材仍然可能存在许多缺点和不足,敬请同行专家和广大读者继续批评指正。

第一版出版说明

随着计算机应用的日益深入、普及,目前我国正在兴起学习计算机专业知识的高潮,各种有关计算机的书籍如雨后春笋般涌现出来,使广大读者大有应接不暇之势。但是,已经出版的这些书籍中,有的偏深偏专,取材偏多偏全,适合有一定基础的计算机专业人员阅读参考;有的则是普及性读物,只适合急于入门的计算机爱好者使用;在为数不多的教材中,大都是为计算机专业本科生使用而编写的,不适合成人教育和大专类学生的需要。鉴于这种形势,我们决定编写一套适合于计算机类各专业大专学生和成人教育使用的教材。这套教材共有十种,虽然它还不能完全覆盖上述办学层次教学计划中的所有课程,但是它包含了培养一个计算机类专科生的主要教学内容。其中入门的教材有《计算机应用基础》和《C 语言程序设计》;属于专业基础的教材有《16 位微型计算机原理与接口》,《汇编语言程序设计》,《数据结构》和《操作系统》;应用性较强的《单片机及其应用》,《数据库系统教程》,《计算机网络基础》和《软件工程引论》。

这套教材贯彻了理论联系实际、学以致用原则。在取材方面,不追求包罗万象、面面俱到,而着力保证把最基本、最实用的部分包含进来。在叙述方面,力求做到深入浅出,尽量用实例来说明基本概念和基本方法。我们希望这套教材不仅能适合课堂讲授的需要,也便于广大读者自学。这套教材由南开大学计算机与系统科学系的教师们编写而成,他们之中既有教学经验十分丰富的教授、副教授,也有活跃在计算机应用最前沿的青年教师。这些教师不仅具有教本科生、研究生的教学经验,也具有教大专生和成人教育的教学经验,这就使这套教材的质量有了基本的保证。但是由于我们初次编写这类教材,尚未经过实践的检验,缺点和不足之处在所难免,敬希同行专家和广大使用者批评指正。

第一版前言

目前,计算机应用已普及到各个领域,掌握计算机的基础知识,将计算机运用于各个行业,尽最大可能地发挥其潜在能力,已是一种大趋势。在计算机领域里,应用最广、实用性最强的一种技术就是数据库技术,它已渗透到许许多多的应用行业中,成为人们掌握计算机知识所必修的一门课程。为此我们编写了这本《数据库系统教程》,作为人们学习这门技术所必需的一本教材。

到目前为止,有关数据库方面的教材已有许多种,这些书从内容结构上可分为两类:一类是讲数据库理论知识的,并不涉及某个具体的实用数据库系统;另一类是讲述某一个或几个实用数据库系统的技术及使用方法,而不过多地阐述数据库的基础理论。因此,使用上述教材,很难使读者理论联系实际,也很难全面理解和掌握数据库的完整知识和概念。鉴于此,作者在多年的数据库教学实践中,吸取其他多种教材的优点,总结自己成功的教学经验,并结合目前数据库系统的特点,将理论知识和实用技术综合起来编写成这本教材。本教材涉及面广,难度适中,结构编排合理,适用性良好。本书共分8章。第1章以介绍数据库基本概念和思想为主,是全书的基础。

第2章到第5章重点介绍目前最流行的关系数据库的有关理论知识,包括关系数据库基本概念、SQL语言、关系代数、关系数据理论和关系系统的查询优化。

第6章和第7章介绍数据库的环境和数据库的设计知识。

第8章介绍目前应用最广的关系数据库管理系统FoxBASE⁺。它以应用为主线,在介绍基础知识和基本方法的同时,强化了对实用技术的介绍。

本书第1~7章由袁晓洁执笔编写,第8章由徐雷执笔编写,并由袁晓洁对全书进行了统编和审阅。在本书的编写过程中得到了南开大学计算机与系统科学系的陈有祺教授、朱瑞香教授和韩维桓教授的指导与帮助,他们为本书提出了许多宝贵意见,对此我们表示深深的谢意。另外南开大学出版社对该书的编辑、校对、出版等工作付出了大量艰辛的劳动,给予了很大的支持和帮助,在此一并表示感谢。

由于作者水平所限,编写的时间仓促,本书错误之处在所难免,敬请专家和读者批评指正。

作者

第二版前言

该书自 1995 年正式出版以来,深受广大读者特别是大专院校师生的欢迎,被许多院校选为计算机学科或相关专业的授课教材,我们为自己多年教学成果的结晶得到大家的广泛认可而感到欣慰。五年来,不少专家和读者为我们写来了热情洋溢的信,提出了许多宝贵的意见,也提出了殷切的希望,对此我们深表谢意。为了满足广大教师和读者的要求,也为了跟上数据库技术飞速发展的潮流,我们对该书进行了修改和扩充,使其能够更好地适应当前数据库教学方面的需求。

主要改动如下:

(1)结合当前数据库技术的发展,修改了部分已显陈旧的内容。

(2)鉴于目前 Windows 下数据库开发工具的增多,也为了保持与原书程序设计部分的一致性,增加了第 9 章 FoxPro for Windows 内容。

(3)为帮助读者更加形象、深刻地理解该书内容,我们制作了多媒体教学光盘,并随此次再版图书赠送。该光盘包括两个多媒体教学课件:“关系数据库 SQL 标准语言教程”和“缤纷数据库直通车”。教学型智能多媒体课件“关系数据库 SQL 标准语言教程”以 SQL 语言为背景,生成了一个智能导航式教学系统。该系统具有开放性,教师可自行输入教学内容和测试题目,进行导航式教学,并可同时记录学生的进度和学习状况。该课件获南开大学多媒体创作大赛一等奖。自学型智能多媒体课件“缤纷数据库直通车”是本书的多媒体教材,它将《数据库系统教程》的全部内容囊括其中,采用图、文、动画、语音及视频技术相结合的方式,讲解重点内容,配有上百道题目,并可由计算机对答题结果进行自动评判,使读者能够对自学的效果进行检验,帮助读者尽快掌握所学知识。该课件采用浏览器方式导航,结构清晰,较好地再现和扩充了原有的教材的内容,十分适合自学数据库系统课程。该课件在天津市教委组织的多媒体课件评比中获一等奖。

本书第 9 章由康义楠执笔编写,袁晓洁对全书进行了统一修改和审阅。多媒体课件由袁晓洁、徐雷、康义楠、王祥瑞、冯学虎、李昭、曹奇、罗岚、董震、刘涛、林晓民制作。再次感谢广大读者对本书的支持。

作者

2000.1.8

目 录

第 1 章 数据库技术概念

1.1 文件技术与数据库技术	(1)
1.1.1 文件技术的局限	(1)
1.1.2 数据库技术的发展	(3)
1.1.3 数据库技术的特点	(5)
1.1.4 数据库基本概念	(7)
1.2 数据模型	(8)
1.2.1 概念模型	(9)
1.2.2 数据模型三要素	(12)
1.2.3 层次模型	(13)
1.2.4 网状模型	(17)
1.3 数据库系统结构	(21)
1.3.1 数据库的三级模式结构	(21)
1.3.2 数据库系统组成	(23)
1.4 数据库管理系统	(25)
1.4.1 DBMS 基本功能	(26)
1.4.2 DBMS 的组成	(27)
1.4.3 DBMS 的评价与选择	(28)
习题	(30)

第 2 章 关系数据库

2.1 关系模型的基本概念	(31)
2.1.1 关系的通俗解释	(31)
2.1.2 关系的数学定义	(32)
2.1.3 关系模型	(33)
2.1.4 关系数据库管理系统	(35)
2.2 关系代数	(35)
2.2.1 传统的集合运算	(36)
2.2.2 专门的关系运算	(37)
2.2.3 关系代数表达式	(38)
习题	(40)

第3章 关系数据库标准语言—SQL

3.1 SQL 特点	(41)
3.2 SQL 数据定义功能	(42)
3.2.1 建立基本表	(42)
3.2.2 扩充修改基本表	(43)
3.2.3 撤销基本表	(44)
3.3 SQL 数据查询功能	(44)
3.3.1 简单查询	(45)
3.3.2 连接查询	(47)
3.3.3 嵌套查询	(48)
3.3.4 组函数	(49)
3.4 SQL 数据更新功能	(51)
3.4.1 插入数据	(51)
3.4.2 修改数据	(51)
3.4.3 删除数据	(52)
3.5 建立和使用视图	(52)
3.5.1 视图的建立与删除	(53)
3.5.2 视图查询	(54)
3.5.3 视图的更新	(54)
3.6 索引的建立与删除	(55)
3.6.1 索引的建立	(55)
3.6.2 索引的删除	(55)
3.7 SQL 数据控制功能	(56)
3.7.1 使用权的授予	(56)
3.7.2 使用权的收回	(57)
习题	(57)

第4章 关系数据库设计理论

4.1 关系模式的存储异常与数据依赖	(59)
4.2 函数依赖的基本概念	(60)
4.3 关系的规范化	(62)
4.3.1 第一范式	(62)
4.3.2 第二范式	(63)
4.3.3 第三范式	(64)
4.3.4 BCNF 范式	(64)
4.3.5 多值依赖	(66)
4.3.6 第四范式	(68)
4.4 模式分解的优劣	(69)
习题	(70)

第5章 数据库查询的实现与优化

5.1 概述	(72)
--------------	------

5.2 优化的一般策略.....	(75)
5.2.1 逻辑层优化的一般策略	(75)
5.2.2 物理层优化的一般方法	(76)
习题	(76)

第6章 数据库环境

6.1 关系完整性及安全性.....	(77)
6.1.1 完全性的一般方法	(77)
6.1.2 SQL 系统的安全性	(79)
6.1.3 完整性约束条件	(81)
6.1.4 ORACLE 系统的完整性控制	(81)
6.2 事务处理.....	(83)
6.2.1 事务的基本概念	(83)
6.2.2 三个并行问题	(84)
6.2.3 并发控制	(84)
6.2.4 恢复策略.....	(85)
习题	(86)

第7章 数据库设计概念

7.1 规划.....	(87)
7.2 需求分析.....	(88)
7.3 概念结构设计.....	(90)
7.4 逻辑结构设计.....	(91)
7.5 数据库物理设计.....	(91)
7.6 数据库实现.....	(94)
7.7 数据库运行和维护.....	(94)
习题	(95)

第8章 FoxBASE+数据库管理系统

8.1 FoxBASE+ 概述	(96)
8.1.1 FoxBASE+ 文件的结构	(97)
8.1.2 FoxBASE+ 命令语句规则	(98)
8.1.3 FoxBASE+ 的技术指标	(99)
8.2 FoxBASE+ 数据库操作	(100)
8.2.1 对数据库结构的操作	(101)
8.2.2 对数据库记录的操作	(104)
8.2.3 对数据库运算的操作	(116)

8.2.4 数据库之间的操作	(117)
8.2.5 对数据库文件的操作	(122)
8.3 数据类型、常量、变量、表达式和函数	(124)
8.3.1 数据类型	(124)
8.3.2 常量和变量	(125)
8.3.3 表达式	(126)
8.3.4 有关内存变量的操作	(128)
8.3.5 对内存变量数组与数据库变量之间的操作	(131)
8.3.6 对函数的操作	(133)
8.4 FoxBASE+ 程序设计	(142)
8.4.1 命令文件的建立与运行	(143)
8.4.2 交互式命令	(146)
8.4.3 屏幕显示命令	(147)
8.4.4 选择结构命令	(152)
8.4.5 循环结构命令	(154)
8.4.6 事件处理命令	(156)
8.4.7 过程文件	(158)
8.4.8 内存环境	(160)
8.4.9 菜单程序的设计	(163)
8.4.10 工作方式和状态的设置	(169)
习题	(170)

第 9 章 FoxPro 2.5 for Windows 应用

9.1 FoxPro 2.5 系统	(174)
9.1.1 系统概述	(174)
9.1.2 FoxPro 2.5 的配置与运行环境	(175)
9.1.3 FoxPro 的菜单系统与命令窗口	(175)
9.2 FoxPro 数据库操作	(177)
9.2.1 对数据库结构的操作	(177)
9.2.2 对数据库记录的操作	(180)
9.2.3 数据查询与统计	(188)
9.2.4 View 窗口与多重数据库操作	(194)
9.3 FoxPro 开发工具	(198)
9.3.1 屏幕生成器	(198)
9.3.2 报表生成器	(203)
9.3.3 菜单生成器	(208)
9.3.4 编写应用程序	(211)
9.3.5 编制应用程序举例	(214)
习题	(219)

第1 章

数据库技术概论

电子计算机在刚刚问世的前 10 年中,主要是用于解决科学研究和工程设计中的数值计算问题。这个时期它的技术与应用只被少数科学家所拥有。随着计算机自身不断地发展,计算机的应用由科学研究部门扩展到各种企业和行政部门。因此,在计算机的三大主要应用领域,即科学计算、数据处理和过程控制中,数据处理迅速上升为计算机应用的主要方面,数据库技术就是作为数据处理中的一门新技术而发展起来的。

数据处理是指对各种形式的数据进行收集、储存、加工和传播的一系列活动的总和。其目的之一是从大量的、原始的数据中抽取、推导出对人们有价值的信息以作为行动和决策的依据;目的之二是为了借助计算机科学地保存和管理复杂的大量的数据,以便人们能方便而充分地利用这些宝贵的信息资源。

数据库技术所研究的问题就是如何科学地组织和储存数据,如何高效地获取和处理数据。数据库技术作为数据管理的最新技术目前已广泛应用于各个领域,数据库系统已成为当代计算机系统的重要组成部分。掌握设计和开发数据库及其应用技术是非常重要的。

本章主要介绍数据库的基本概念与技术,它们包括:数据模型、数据库体系结构和数据库管理系统等。这些内容是学习后面几章的基础。

1.1 文件技术与数据库技术

在 50 年代到 60 年代之间,计算机不仅用于科学计算,而且已经大量用于数据处理方面,此时计算机外存储器有了磁鼓、磁盘等直接存取的存储设备,为计算机进行事务管理奠定了硬件基础;另外软件技术主要是数据结构设计和数据管理技术的研究也得到了迅速的发展,从而出现了专门的管理数据的软件,这就是所谓的“文件技术”阶段。

1.1.1 文件技术的局限

在文件技术阶段,数据管理的主要特点可概括为如下几点:

1. 外存储器成为计算机系统不可缺少的组成部分,用户可以随时通过程序对文件进行查询、修改和增删等处理;
2. 文件组织形式日益多样化,索引文件、直接存取文件、链接文件纷纷出现,既可以满足批处理应用的需要,又能有效地实现记录的随机存取;
3. 出现了专用数据管理系统的软件,即文件系统,它能对驻留在外存储器上的数据文件实施统一管理。这种专用数据管理软件构成了操作系统的一个重要组成部分。由于应用程序不再需要了解数据在存储介质上的实际地址,因而大大减少了程序设计的工作量;

4. 数据不再仅仅属于某个特定的程序,而可以重复使用。

文件技术的上述特点使得这项技术在 20 世纪 60 年代中得到了充分的发展,把计算机应用推向了一个新的高潮。但由于文件结构的设计仍然是基于特定用途的,程序仍然是基于文件的特定物理结构和存取方法编制的,因此,数据结构与程序之间的依赖关系并未根本改变,因而限制了它的进一步发展。

60 年代中期以后,计算机在数据处理领域的应用迅速发展,由个别部门的应用逐步发展成多个部门的普遍应用,由简单孤立的单项应用发展为彼此相关的复杂应用,从而使管理的规模更加庞大,数据量急剧增长,共享性也更强了。这就带来了数据管理上的一些新问题。现举一例说明。

某学校的学生处、教务处和卫生院均要使用计算机对学生的有关信息进行管理,但其各自处理的内容不同,如用文件系统实现,可按如下方式进行组织。

学生处要处理的信息包括:

学号,姓名,系名,年级,专业,年龄,性别,籍贯,政治面目,家庭住址,个人履历,社会关系
.....

为此,学生处的应用程序员必须定义一个文件 F1,该文件结构中的记录应包括上述几个数据项。

教务处要处理的信息包括:

学号,姓名,系名,年级,专业,课名,成绩,学分.....

显然,教务处的应用程序员需定义一个文件 F2,该文件结构中的每一记录包括以上几个数据项。

类似地,当卫生院要记录和处理学生的有关健康情况信息时,其创建的文件 F3 应包括下列数据项:

学号,姓名,系名,年级,专业,年龄,性别,身高,体重,健康状况.....

这样当上述三个部门共用某台计算机时,该计算机的外存中要同时保存 F1,F2,F3 三种文件,可这三种文件中均有学生的学号、姓名、系名、年级和专业等信息,因此重复的数据项达到了 1/3 以上,严重冗余。

数据冗余不仅浪费存储空间,更严重的是带来潜在的不一致性。由于数据存在多个副本,所以当发生数据更新时,就很可能发生某些副本被修改而另一些副本被遗漏的情况,从而使数据发生不一致,影响数据的正确性和可靠性。比如,某学生由于特殊原因需要转系,他从数学系转到计算机系,当学生处得到该信息后,将该学生所属的系名由数学系改为计算机系,因而 F1 文件中保存了正确的信息。但如果教务处和卫生院没有得到该学生转系的信息,或者没有及时更改 F2 和 F3 文件,这就造成了不一致,使得一个学生属于两个系,人们也就无法确认哪一个是正确的信息。由于数据的使用价值在很大程度上依赖其可靠性,所以,这种不一致的后果是不可忽视的。可以想象,当这种情况发生在军事、航天、金融等行业部门时,其后果是非常严重的。

另一方面,在传统文件技术阶段,文件是为某一特定应用服务的,应用程序和文件是一对一的(见图 1-1),这就造成了应用程序与数据结构过分地互相依赖,而且系统很难扩充。一旦数据的逻辑结构改变,就必须修改应用程序和文件结构的定义;反之,应用程序的改变,也就影响文件的数据结构的改变。比如,学生处在管理学生信息时发现,只登记学生年龄是不够的,还应记录下学生的出生年月,因此需在 F1 的文件中加一个数据项,这就必须修改 F1 的文件结

构,同时还要修改涉及 F1 文件的应用程序;另一方面,若用户原先是用 Pascal 语言编制的访问 F1 文件的应用程序,现想改用 C 语言,那么,F1 文件的数据结构也必须随之改变。由此可看出在文件技术阶段,扩充已有的系统是很困难的,因此,用户迫切希望应用程序能够不受数据结构变化的影响。

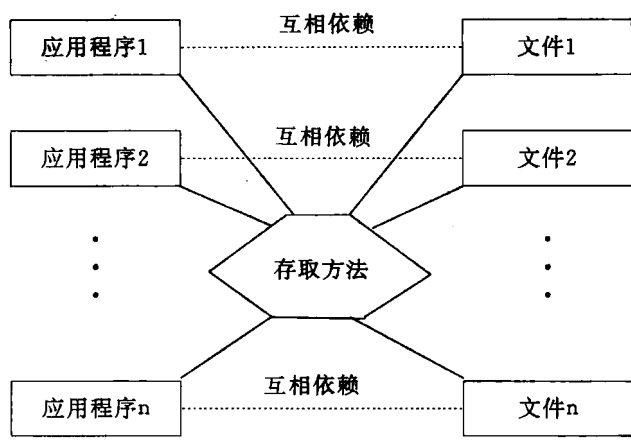


图1-1 应用程序与文件之间一对一关系

此外,文件系统缺乏对数据操作进行控制的方法,对于数据的安全性、保密性和正确性等方面的控制,完全要用户程序自己负责,这使得应用程序的编制相当繁琐。

综上所述,传统的文件技术有许多缺点,不能满足人们的要求,因此迫切需要新的数据管理技术来实现对数据的共享,实现数据与程序的独立性,并提供安全性和完整性,也就是说在操作系统之上必须有一个软件系统——数据库管理系统 DBMS(Data Base Management System),在建立、运用和维护时,对数据库进行统一控制,这就是数据库技术。

1. 1. 2 数据库技术的发展

在 50 年代初,数据处理进入文件处理阶段,实际上是数据库系统的前奏,到了 60 年代,数据库技术开始成形,从四个方面可反映出数据库技术的形成过程。

1. CODASYL 的建立与发展

1959 年 5 月,美国国防部召集了有计算机厂家、用户和政府其它部门等的专家 40 余人,讨论开发事务处理通用语言,并组成了数据系统语言协会(The Conference On Data Systems Languages 简称 CODASYL)。该协会在随后的十年里做了大量的事务处理软件开发和论证工作,其中主要的成果有:

- (1)1960 年,设计并发表了新的语言 COBOL(Common Business Oriented Language)。
- (2)1967 年,CODASYL 的 COBOL 协会专门成立了相应的组织,定名为数据库任务组 DBTG(Data Base Task Group),对语言问题进行研究。DBTG 小组于 1968 年发表了第一个工作报告:有数据库操作功能的 COBOL 的扩展。
- (3)1969 年,CODASYL 委员会公布其研究成果 DBTG 报告,该报告确定并建立了数据库系统的许多概念、方法和技术。
- (4)1971 年,CODASYL 之下又成立了新的组织——数据定义语言(DDL)委员会,并提出了数据定义的语言报告。随后 DBTG 对子模式定义语言和数据操纵语言作为 COBOL 语言扩

展内容作了研究并予以完成。

DBTG 所提议的方法是基于网状结构的,它是数据库网状模型的基础和典型代表。

2. 商用数据库管理系统的发展

60 年代初,许多公司开发了用于商用管理的软件系统,这些系统虽然还称不上是数据库系统,但却具有了现代数据库的特色。这些系统有:GE 公司的 IDS 商用系统;IBM Informatics 公司在 IBM System/360 上进行的 MARK IV 系统;MITRE 公司接受美国空军的委托开发的实验系统 ADAM(Advanced Data Management System);SCD 公司开发的 TDMS(Time — Shared Data Management System)系统等等。

与此同时,IBM 公司却迟迟没有对此项问题进行研究,直至 1965 年才开始提出 GIS (Generalized Information System)。后来又于 1969 年发表了 IMS(Information Management System),它是一个 DB/DC(Data Base/Data Communication)的庞大系统。IBM 与洛克菲尔公司共同开发提出了基于自己的数据库模型:一种以层次数据结构为基础的数据模型及与其相适应的语言 DL/1(Data Language/1)。该产品有极大的竞争力。人们把 IMS 作为当前流行的三大数据库模型之一——层次模型数据库的代表。

3. 关系方法的发展

1970 年,IBM 公司 San Jose 研究所的 E. F. Codd 在美国计算机学会会刊“Communication of the ACM”上发表了题为“大型共享数据库数据的关系模型”(“A Relational Model of Data for Shared Data Banks”)的著名论文,为关系模型的讨论奠定了基础,并将数据库系统的研究推向新的历史时期,即从只以应用为主转向理论研究发展的新时期。

继 1971 年后,Codd 又继续发表了多篇关于关系模型的文章,把数学方法引入到数据库系统,使数据库系统有了坚实的理论基础。Codd 本人作为关系数据库的创始人和奠基人,获得了 1981 年 ACM 图灵奖。

自 Codd 提出关系模型数据库方法以后,1974 年 IBM 就提出了 SQL(Structured Query Language)语言,这是一种基于关系方法实现对数据库存取的语言。大约从 1975 年到 1979 年的五年期间,IBM 公司 San Jose 实验室成功地研制了一个实现 SQL 语言的关系数据库系统原型 System R,该系统在 IBM 370 上运行。之后,IBM 又将 SQL 语言引入到 DB2(IBM Data Base 2)中,配置在 MVS 上运行,并于 1983 年推出 DB2 产品,表明了关系模型数据库系统终于被人们所接受,并得到最终的肯定。

1979 年,一个叫 Ashto-Tate 的小公司引入了微机产品 dBASE II,并称之为关系 DBMS,然而,该产品的成功混淆了数据库处理的主题,事实上,它只是带有通用文件处理能力的编程语言。今天,微机市场已经变得更加成熟和完善,dBASE III+,dBASE IV 才是一个真正的关系 DBMS。

dBASE 开创了微机上数据库技术应用的先河,使得其他厂商也开始把他们的产品从主机向微机移植。Oracle、Focus 和 Ingres 是三个向下移植到微机上的例子,它们是真正的 DBMS 程序。把数据库技术移植到微机上的一个影响就是 DBMS 用户界面的巨大改进。今天,DBMS 产品采用图形用户界面,如 Microsoft Windows,提供了丰富而又强大的用户界面,Microsoft 的 Access 和 Borland 的 Paradox for Windows 就是这样的两个产品。

微机、关系模型和大大改进的用户界面的结合使得数据库技术从组织范畴转移到个人计算机范畴。当这种转移发生时,使用数据库技术的数目激增。在 1980 年,美国有 1 万个使用 DBMS 的地方,而目前有 500 多万个这样的地方。

4. 客户服务器及分布式数据库应用

80 年代中期到后期,终端用户开始使用一种称作局域网(LAN)的新型计算机通信技术将独立的计算机连接起来。这些网络使得计算机能够以过去无法想像的速度从一台计算机向另一台计算机发送数据。这种技术的第一个应用是共享外设,如大容量高速磁盘、昂贵的打印机和绘图机以及通过电子函件方便计算机间通信。

基于 LAN 的体系结构大大不同于基于主机和小型机数据库的体系结构。采用主机或小型机时,只有一个 CPU 参与数据库应用处理,而采用 LAN 时,则有许多 CPU 同时参与。由于这种情形既有利又存在很多问题,因而导致了一种新型的多用户数据处理,称作客户机服务器数据库结构(client server database architecture)。这种结构是今天大多数工作组数据库处理的基础。

分布式数据库是把组织的所有数据分布在很多计算机上,如微机、LAN 服务器和主机,使得进行数据库处理时,需要相互通信。分布式数据库的目标是使得对于每一个用户而言,好像他是组织数据的唯一用户,并提供没有别人使用时他所应该有的一致性、精确性和实用性。

1.1.3 数据库技术的特点

与文件系统相比,数据库技术提供了对数据的更高级更有效的管理,它有如下主要特点:

1. 数据共享性。数据共享的意义是多种应用、多种语言互相覆盖的共享数据集合。在传统的文件组织中,每个数据文件是特定的应用所私有的,如图 1-1 所示。而数据库是从整体观点来看待和描述数据的,数据不再是面向某一应用,而是面向整个系统,如图 1-2 所示。这可以大大减少数据冗余度,既节约存储空间,减少存取时间,又可避免数据之间的不相容性和不一致性。前面所举文件系统例子,若从数据库角度进行组织,则如图 1-3 所示。

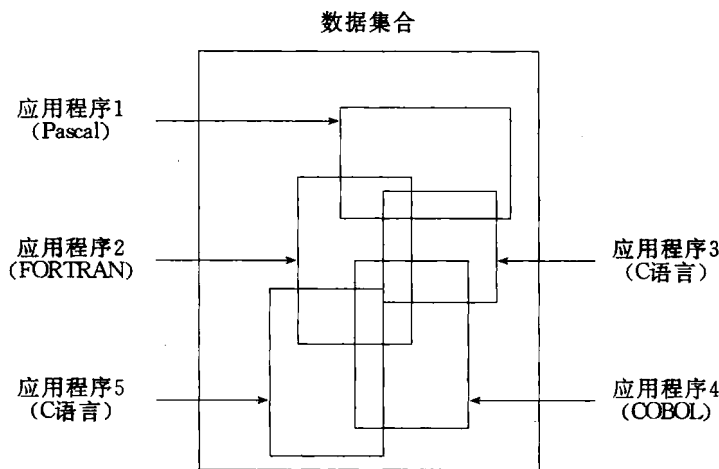


图1-2 应用程序共享数据集合

由该结构可看出,数据不再分属于各个应用程序,而是集中存储在一起。对一个特定的组织而言,除了由于需要保密而强加的某些限制以外,这些数据是为该组织的各个下属部门和各个方面的应用所共享的,这样就大大提高了数据的使用价值。在图 1-3 中,学生记录是为教务处、学生处和卫生院所共享的,若某个学生需要转系,则只要修改学生记录中的系名属性即可,这样就不会出现不一致性的情况。另外,除了共享的数据以外,各部门还可以有自己的私有数