

DATA BASE SYSTEM
IN MICROCOMPUTER

微机数据库—— FoxBASE 基础教程



主编 白旭红 副主编 李宁 编著 白旭红 李宁 王利 郑小玲



DATA BASE
SYSTEM IN
MICROCOMPUTER

首都经济贸易大学出版社

内 容 简 介

本书是以全国计算机等级考试二级——FoxBASE 数据库管理系统考试大纲为主要依据编写的教科书,是适用于高等学校、中等专业学校以及成人教育类非计算机专业的计算机基础课程。内容包括数据库基础知识、关系型数据库管理语言 FoxBASE 的基本命令、操作方法及程序设计基本方法。

本书在编写过程中力求深入浅出,通俗易懂,亦可作为在职人员自学微机数据库知识或参加等级考试的参考教材。

前 言

随着人类社会的不断进步,随着科学技术的不断发展,以计算机技术为标志的现代信息技术已成为当代人类文明的重要组成部分。因此,计算机知识也就成为人们参与社会生活所必备的知识。近些年来,越来越多的人意识到计算机的重要性,积极参加有关的学习和进修,从而加快了我国计算机技术的普及,促进了应用水平的提高。为适应计算机技术发展的需要,国家教委考试中心于1994年推出了面向社会的“全国计算机等级考试”。这一举措不仅为人才的选拔提供了一个统一、公正、客观的标准,也为计算机基础教育在培养目标和层次方面提供了参考依据。

为满足社会对计算机基础课程和应试教材的需求,我们编写了计算机基础课系列教材,包括本书《微机数据库》与《计算机基础教程》。这套系列教材较全面、系统地介绍了微型机应用的基础知识,是高等学校、中等专业学校以及成人教育类非计算机专业计算机基础课程的理想教材,也可供在职人员自学使用。

本书在编写过程中参考了“全国计算机等级考试”二级——FoxBASE的考试大纲,主要内容参请见目录。

在FoxBASE命令中,多数命令是通过主句与有关子句的选用和组合,把各种功能综合在一起,形成一个功能较强的命令,要求初学者一下子了解这种命令的完整功能,显然比较困难。因此,本书在前七章中所介绍的命令均从横向和纵向这两个不同角度将一个命令的各项功能拆开来介绍,首先从横向介绍主句所体现的基本功能;这就使读者比较容易入门。然后从纵向介绍子句所体现的附加功能,在第七章末给出前七章所介绍过的命令全格式。特别值得一提的是用于记录筛选的<范围>子句、FOR子句、WHILE子句及用于字段筛选的FIELDS子句,它们可以用于多种库文件的操作命令中,而且子句的功能不会因为在不同的命令中而发生改变,因此,学习了一个子句就等于弄清楚一批命令的某种附加功能,事半功倍。

为适应课堂教学或自学需要,每章后面均附有习题,大部分章节还配有上机题目。

参加本教材编写的人员,均为首都经济贸易大学经济信息管理系的教师。其中第一章由王利撰写,第二、四、九章由郑小玲撰写,第三章、第五至第八章、第十四章、第十五章由李宁撰写,第十章至第十三章由白旭红撰写。全书由白旭红主编,并由白旭红与李宁合作总纂,盛定宇审定。

由于本书编写时间比较仓促,有不妥之处恳请读者指正。

编者

1996年10月

目 录

第一章 数据库系统基本概念	(1)
第一节 数据、信息与数据处理	(1)
第二节 计算机数据管理的发展	(4)
第三节 数据库系统的组成	(8)
第四节 关系数据模型	(11)
第二章 FoxBASE 概述	(17)
第一节 FoxBASE 功能特点	(17)
第二节 使用 FoxBASE 的准备	(18)
第三节 FoxBASE 技术指标和文件类型	(19)
第三章 数据库文件及其基本操作	(23)
第一节 数据库文件的基本概念	(23)
第二节 数据库文件的建立	(26)
第三节 数据库文件的打开与关闭	(28)
第四节 数据库文件结构的维护	(28)
第五节 记录的追加、修改和输出	(29)
第六节 当前记录与记录定位	(31)
第七节 插入记录	(33)
第八节 字段的替换修改	(33)
第九节 记录的删除	(34)
第十节 建立数据库文件的其他方法	(36)
第十一节 FoxBASE 的命令格式	(37)
第四章 FoxBASE 表达式	(37)
第一节 常量和变量	(43)
第二节 表达式	(46)
第三节 函数	(51)
第五章 库文件的索引与排序	(67)
第一节 记录顺序的概念	(67)
第二节 索引的建立	(67)
第三节 索引的打开与关闭	(68)
第四节 当前顺序的改变	(69)
第五节 重建索引	(71)

第六节 降序索引和多重顺序索引	(71)
第七节 排序	(72)
第六章 记录的定位	(77)
第一节 物理定位	(77)
第二节 逻辑定位	(77)
第三节 定位是否成功的判别	(83)
第四节 软定位	(84)
第七章 记录和字段的筛选	(88)
第一节 记录筛选和字段筛选的概念	(88)
第二节 记录过滤器筛选	(89)
第三节 FOR 子句筛选	(90)
第四节 WHILE 子句筛选	(91)
第五节 范围控制与<范围>子句	(91)
第六节 记录筛选的优化	(92)
第七节 记录筛选与记录定位的关系	(93)
第八节 字段筛选	(94)
第九节 筛选结果的保存与装入	(96)
第十节 库文件操作命令的筛选功能	(97)
第八章 数据库文件的统计	(105)
第一节 计数、求和与求平均值	(105)
第二节 分类汇总	(107)
第三节 分级汇总	(108)
第九章 内存变量的操作	(116)
第一节 内存变量的数据输入	(116)
第二节 内存变量的管理	(118)
第三节 数组	(122)
第十章 FoxBASE 程序设计基础	(131)
第一节 算法及算法的描述	(131)
第二节 程序设计方法与步骤	(134)
第三节 命令文件的建立与运行	(136)
第四节 简单的顺序控制结构程序举例	(138)
第五节 使用选择控制结构的程序设计	(139)
第六节 使用重复控制结构的程序设计	(145)
第七节 常用的典型算法	(149)
第十一章 FoxBASE 的过程及其调用	(160)
第一节 过程与过程调用	(160)
第二节 过程和嵌套调用	(166)
第三节 参数传递	(167)
第四节 自定义函数	(169)
第十二章 数据的格式输入与输出程序设计	(175)

第一节	屏幕坐标·····	(175)
第二节	清除屏幕及在屏幕上画框·····	(175)
第三节	数据的格式输入和输出命令·····	(178)
第四节	用@···SAY···PICTURE 设计报表程序·····	(184)
第五节	用@···GET···PICTURE 设计屏幕格式输入程序·····	(191)
第六节	数据的打印输出·····	(196)
第十三章	菜单程序块的设计·····	(202)
第一节	菜单程序块的概念·····	(202)
第二节	菜单程序的设计·····	(204)
第十四章	多文件操作·····	(216)
第一节	如何同时打开多个文件·····	(216)
第二节	库文件的批量更新·····	(219)
第三节	文件的联动·····	(222)
第四节	文件的联合·····	(225)
第十五章	实例研究·····	(232)
第一节	库文件批量更新的通用算法·····	(232)
第二节	期刊流通管理系统·····	(236)
附录 A	标签的制作·····	(246)
附录 B	命令、函数及术语索引·····	(249)

第一章 数据库系统基本概念

数据库(Database)是存储在一起的通用化相关数据的集合,它不仅包括描述事物的数据本身,而且包括有关事物之间的联系。数据库系统是实现有组织地、动态地存储大量的相关数据,方便多用户访问的计算机软、硬件资源组成的系统。

数据库技术产生于60年代末,70年代初。在计算机的三大主要应用领域(科学计算、数据处理与过程控制)中,数据处理约占70%左右的比重。计算机作为信息处理的工具,为适应数据处理需求的迅速提高,满足各类信息系统对数据管理的要求,在文件系统的基础上发展起了数据库系统。数据库技术主要研究如何存储、使用和管理数据,是计算机数据管理技术发展的最新阶段。30多年来,数据库在理论上、实现技术上均得到很大的发展,研制出许多商用数据库管理系统,使得计算机应用渗透到各行各业的各类管理工作中。管理信息系统、办公自动化系统、决策支持系统等都是使用了数据库管理系统或数据库技术的计算机应用系统。

现在,数据库技术已成为计算机领域中最重要技术之一,是软件学科的一个独立分支。数据库方法原本是针对事务处理中大量数据管理需求的,但是它的应用范围不断扩大,不仅应用于事务处理,并且进一步应用到情报检索、人工智能、专家系统、计算机辅助设计等,涉及到非数值计算各方面的应用。应用范围的扩大又进一步促进了数据库技术的深入发展,可以说数据库系统已成为当代计算机系统的重要组成部分。

我国自从第一次全国人口普查使用了数据库技术以来,随着微机数据库系统的推广,数据库应用逐渐普及,深受广大用户欢迎。计算机应用人员只有了解数据库系统的基本原理,掌握数据库设计的基础理论,熟悉数据库管理系统特点,才能开发出适用的数据库应用系统。

第一节 数据、信息与数据处理

人类的一切活动都离不开数据,离不开信息。但是在不同的领域里,信息的含义有所不同。一般认为信息(Information)是数据、消息中所包含的意义。因此,信息是人类维持正常活动不可缺少的资源。数据和信息有时可以混用,例如,数据处理也称为信息处理;有时则必须分清,例如,不能把信息系统称为数据系统。

一、数据

所谓数据,通常指用符号记录下来的可加以鉴别的信息。数据的概念包括两个方面:其一,数据内容是事物特性的反映或描述;其二,数据是符号的集合。

由于描述事物特性必须借助一定的符号,这些符号就是数据形式。例如,某人的出生日期是“一九七八年六月二十五日”,当然也可以将以上的汉字形式改为用“06/25/78”来表示。

数据的概念在数据处理领域中比在科学计算领域中已经大大地拓宽了。所谓“符号”,不仅仅指数字、字母、文字和其他特殊字符,而且还包括图形、图像、声音等多媒体数据;所谓“记录

下来”也不仅是指印在纸上,而且包括记录在磁介质、光介质上及半导体存储器里。

数据在空间上的传递称为通信,在时间上的传递称为存储。

二、信息

信息是关于现实世界事物的存在方式或运动形态反映的综合,是人们进行各种活动所需要的知识。数据与信息既有联系又有区别。数据是载荷信息的物理符号或称为载体。数据能表示信息,但并非任何数据都能表示信息,正如人们常说的“如果计算机输入的是垃圾,输出的也会是垃圾”。同一数据也可能有不同的解释,因此,信息只是人们消化理解了的数据。信息是抽象的,不随数据设备所决定的数据形式而改变;而数据的表示方式却具有可选择性。

信息是反映客观现实世界的知识,用不同的数据形式可以表示同样的信息。例如,同样一条新闻在报纸上以文字的形式刊登,在电台以声音的形式广播,在电视上以图像的形式放映以及在计算机网络上以通信形式传播,其信息内容可以相同。当然,由于信息载体不同、喜闻乐见的程度不同、接收对象不同,所产生的效果或作用各异又将另当别论。

三、数据处理

数据处理是指将数据转换成信息的过程。广义地讲,它包括对数据的收集、存储、加工、分类、检索、传播等一系列活动。狭义地讲,它是指对所输入的数据进行加工整理。其基本目的是从大量的、已知的数据出发,根据事物之间的固有联系和运动规律,通过分析归纳、演绎推导等手段,提取出对人们有价值、有意义的信息,作为决策的依据。由此可见,信息是一种被加工成特定形式的数据,这种数据形式对于数据接收者来说是有意义的。对数据的加工可以相对比较简单也可以相当复杂。简单加工包括组织、编码、分类、排序等;复杂加工可以复杂到使用统计学方法、数学模型等对数据进行深层次的加工。

我们可以用下面的式子简单地表示出信息与数据之间的关系:

信息 = 数据 + 处理

在这里,原始数据是原料,是输入,而信息是对原始数据的综合推导加工的产品,是数据处理的输出结果。当两个或两个以上数据处理过程前后相继时,前一过程称为预处理。预处理的输出作为二次数据,成为后面处理过程的输入,此时信息和数据的概念就产生了交叉,表现出相对性。如图 1.1.1 所示。人们有时说“信息处理”,其真正含义应该是为了产生信息而处理数据。

例如,一个人的“出生日期”是有生以来不可改变的基本特征之一,属于原始数据,而“年龄”是用现年与出生日期相减而得到的数字,具有相对性,可视为二次数据。同样道理,



图 1.1.1 数据与信息的关系

“生产日期”、“购置日期”是产品和设备的原始数据,“失效日期”和“资产折旧”是经过简单计算得出的结果。

又如,用手工或计算机填写的发货单,对于发货部门的工作人员来说即为照单发货的信息,但对于仓储部门的管理者来说,它只是核算盘点库存量的原始数据。由于数据与信息之间

存在着这种关系,这两个词有时被交替使用。其根本区别在于信息对当前或将来的行动或决策有价值。

四、信息系统

信息已经受到社会的广泛重视,被看作社会和科学技术发展的三大支柱之一(材料、能源、信息)。在一个组织里,信息已作为人力、物力、财力之外的第四种资源,占有重要地位。当然,信息是一种非物质的,可以管理和控制的,有别于基本资源的新形式的资源。当今社会已进入信息社会的时代,管理人员在工作中多半时间花费在记录、检索和获取信息上。

1. 信息系统

信息系统是指为了某些明确的目的而建立的,由人员、设备、程序和数据集合构成的统一整体。信息系统的主要功能是提供信息,以支持一个组织机构的运行、管理和决策。更确切地说,信息系统将不适用的数据形式加工成可利用的形式。一个信息系统的质量取决于它是否能及时地为用户提供所需要的信息。在一个组织机构中,不同阶层的管理人员因其管理的目标不同,所需要信息亦不相同。信息系统针对各个层次的需求,通过计算机实现信息支持,达到辅助管理的目的。

2. 信息系统分类

一般可将信息系统分为三类:

- . 数据处理系统;
- . 管理信息系统;
- . 决策支持系统。

数据处理系统 EDP(Electronic Data Processing)。用计算机代替繁杂的手工事务处理工作,其目的是提高数据处理的准确性、及时性,节约人力、提高工作效率。例如,计算机运行会计核算软件,对会计的“簿记”事务进行常规处理、提供数据查询、会计报表等功能,使会计部门的日常工作自动化。

管理信息系统 MIS(Management Information System)。由若干子系统构成的一个集成的人机系统,从组织的全局出发,实现数据共享,提供分析、计划、预测、控制等方面的综合信息。其主要目的是发挥系统的综合效益,提高管理水平。例如,某企业管理信息系统由技术管理子系统、人事管理子系统、财务管理子系统、物资管理子系统、生产管理子系统、设备管理子系统、销售管理子系统组成。实现计算机管理能够迅速、准确地提供有关信息,不仅有力地支持各个职能部门的组织管理,并且通过信息共享加强了各子系统之间的协同,使得整个系统有机地联系在一起,避免了重复搜集、统计、存储,既可以保证数据的准确性,又能够大大提高信息利用率,对领导层制定计划、确定经营目标、指挥生产提供有力的信息支持,从而大大提高管理综合效益,增强企业的市场竞争能力。

应该指出,管理信息系统的结构和组织机构的影响是可逆的。现存的组织结构可以影响管理信息系统的设计,管理信息系统也可以影响未来组织机构的调整。由于计算机的引进使工作和管理环境为之更新,原来行之有效的制度、方法和监督机制在新的环境可能不再适用了。必须依据新环境下的特点来制定新制度、采用新方法、以至实施新的业务流程。

决策支持系统 DSS(Decision Support Systyem)。为决策过程提供有效的信息和辅助决策手段的人机系统,其主要目的是帮助决策者提高决策的科学性及有效性。

计算机辅助决策必须积累大量的数据、案例、方法、模型,更进一步还可以利用知识库系统,专家系统。决策支持系统的服务对象是面向某种决策问题的管理人员,它协助决策者在求

解问题的过程中方便地检索出相关数据,对多种可选方案进行测试比较,然后做出决定。这里需要强调指出,决策支持系统只能对决策提供支持,并不能由计算机代替人,自动化地做出决定,人是决策行动的主体。例如,不同的管理人员运行同一套决策支持系统软件,结果可能做出不同的决策来。

第二节 计算机数据管理的发展

通过上一节对信息系统的简单介绍可以了解到,各级各类信息系统都需要大量的数据作为基础,数据处理的中心问题是数据管理。数据管理是指对数据的组织、分类、编码、存储、检索和维护。

像任何其他技术的发展一样,计算机数据管理也经历了由低级到高级的发展过程。计算机数据管理随着计算机硬件(主要是外存储器)、软件技术和计算机应用范围的发展而不断发展,多年来大致经历了如下四个阶段:

- 人工管理阶段;
- 文件系统阶段;
- 数据库系统阶段;
- 分布式数据库系统阶段。

一、人工管理阶段

50年代中期以前,计算机主要用于科学计算。当时在硬件方面,外存储器只有卡片、纸带、磁带,没有象磁盘这样的可以随机访问、直接存取的外部存储设备。软件方面,没有专门管理数据的软件,数据由计算或处理它的程序自行携带。数据处理方式基本是批处理。数据与应用程序之间的关系如图 1.2.1 所示。

这一时期数据管理的特点是:

1. 数据与程序不具有独立性

一组数据对应一组程序。这就使得程序依赖于数据,如果数据的类型、格式或者数据量、存取方法、输入输出方式等改变了,程序必须作相应的修改。

2. 数据不长期保存

由于数据是面向应用程序的,在一个程序中定义的数据,无法被其他程序利用,因此程序与程序之间存在大量的重复数据。

3. 没有对数据进行管理的软件

数据管理任务,包括存储结构、存取方法、输入输出方式等完全由程序设计人员自负其责,这就给应用程序设计人员增加了沉重的负担。

二、文件系统阶段

50年代后期至60年代中后期,计算机开始大量地用于管理中的数据处理工作,大量的数据存储、检索和维护成为紧迫的需求。在硬件方面,可以直接存取的磁鼓、磁盘成为联机的主要外存。在软件方面,出现了高级言语和操作系统。操作系统中的文件系统(有的也称为信息处

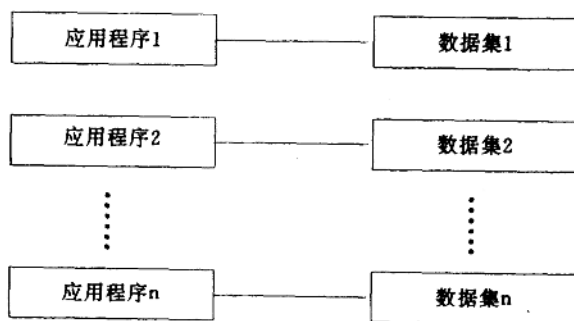


图 1.2.1 人工管理阶段数据与程序的关系

理模块)是专门管理外存的数据管理软件。数据处理方式有批处理,也有联机实时处理。

在这一阶段,程序与数据有了一定的独立性,程序和数据分开存储,有了程序文件和数据文件的区别。数据文件可以长期保存在外存储器上多次存取,方便进行查询、修改、插入、删除等操作。数据的存取以记录为基本单位,并出现了多种文件组织形式,如顺序文件、索引文件、随机文件等。

在文件系统的支持下,数据的逻辑结构与物理结构之间可以有一定的差别,逻辑结构与物理结构之间的转换由文件系统的存取方法来实现。数据与程序之间有了设备独立性,程序只需用文件名访问数据,不必关心数据的物理位置。这样,程序员可以集中精力在数据处理的算法上,而不必考虑数据存储的具体细节。

数据的逻辑结构是指呈现在用户面前的数据结构。数据的物理结构是指数据在物理设备上的实际存储结构。例如,用户看到的记录是按照记录号顺序排列的,而实际上这些记录可能是分散存储在磁盘的不同扇区里,用链接方式组织在一起的。用户访问文件时只需给出文件名、逻辑记录号,而不必关心记录在存储器上的地址和内、外存交换数据的具体过程。此阶段数据与程序的关系如图 1.2.2 所示。

文件系统阶段对数据的管理虽然有了长足的进步,但一些根本性问题仍然没有得到彻底解决,主要表现在以下三个方面:

1. 数据冗余度大

数据冗余是指不必要的重复存储,同一数据项重复出现在多个文件中。在文件系统

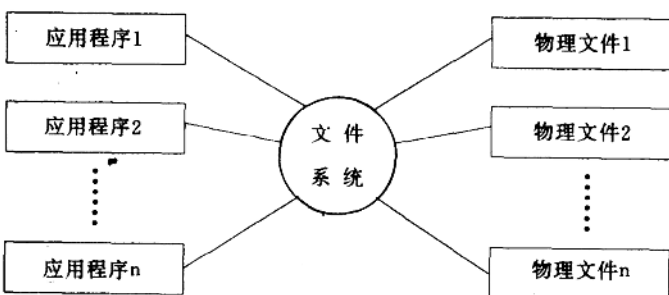


图 1.2.2 文件系统中数据与程序的关系

下,数据文件基本上与各自的应用程序相对应,数据不能以记录和数据项为单位共享。对于不同的应用,即使有很多数据相同,只要逻辑结构不同,用户必须各自建立自己的文件。这不仅浪费存储空间,增加更新开销,更严重的是,由于不能统一修改,容易造成数据的不一致性。

2. 缺乏数据独立性

文件系统中的数据文件是为了满足特定业务领域某部门的专门需要而设计的,服务于某一特定应用程序。数据和程序相互依赖,如果改变数据的逻辑结构或文件的组织方法,必须修改相应的应用程序。同样道理,如果修改应用程序,如改用另一种程序设计语言来编写程序,也将影响数据文件的结构,需要进行转换或重新组织。

3. 数据无集中管理

除了对记录的存取由文件系统承担以外,文件没有统一的管理机构,其安全性与完整性无法保障。数据的维护任务仍然由应用程序来承担。

4. 无弹性、无结构的数据集合

所谓无弹性是指,由于记录的内部结构由应用程序自己定义,而不是由系统来统一管理,因此对现有数据文件的应用不易扩充,不易移植,亦难以增删数据项以适用新的应用要求;无结构是指各个数据文件之间是独立的,缺乏联系,不能反映现实世界事物之间的联系。

这些问题阻碍了数据处理技术的发展,不能满足日益增长的信息需求,这正是数据库技术产生的原动力,也是数据库系统产生的背景。应用需求和计算机技术的发展促使人们研究一种

新的数据管理技术——数据库技术。

三、数据库系统阶段

从 60 年代后期开始,计算机应用于管理的规模更加庞大,需要计算机管理的数据量急剧增长,并且对数据共享的需求日益增强。大容量磁盘(数百兆字节以上)系统的采用,使计算机联机存取大量数据成为可能;软件价格上升,硬件价格相对下降,使独立开发系统维护软件的成本增加。文件系统的数据管理方法已无法适应开发应用系统的需要。为解决数据的独立性问题,实现数据的统一管理,达到数据共享的目的,发展了数据库技术。

数据库(Database)是通用化的相关数据集合,它不仅包括数据本身,而且包括关于数据之间的联系。数据库中的数据不只面向某一项特定应用,而是面向多种应用,可以被多个用户、多个应用程序所共享。例如,某个企业、组织或行业所涉及的全部数据的汇集,其数据结构独立于使用数据的程序,对于数据的增加、删除、修改和检索由系统进行统一的控制,而且数据模型也有利于将来应用的扩展。

为了让多种应用程序并发地使用数据库中具有最小冗余度的共享数据,必须使数据与程序具有较高的独立性。这就需要有一个软件系统对数据实行专门管理,提供安全性和完整性等统一控制机制,方便用户以交互命令或程序方式对数据库进行操作。

为数据库的建立、使用和维护而配置的软件称为数据库管理系统 DBMS(Data Base Management System),它是在操作系统支持下运行的。目前较流行的数据库管理系统包括 Oracle、Informix、Sybase,以上也有微机版本。微机上的数据库管理系统有 dBASE、FoxBASE、FoxPro、FoxPro for Windows、ACCESS、CLIPPER、WINBASE 等。

现在,数据库已成为各类信息系统的核心基础。在数据库管理系统支持下数据与程序的关系如图 1.2.3 所示。

数据库系统的主要特点是:

1. 实现数据共享,减少数据冗余

在数据库系统中,对数据的定义和描述已经与应用程序中分离开来,通过数据库管理系统来统一管理。数据的最小访问单位是数据项,既可以按数据项的名称存取库中某一个或某一组数据项,也可以存取一条记录或一组记录。

建立数据库时,应当以面向全局的观点组织库中的数据,而不能象文件系统那样仅仅考虑某一部门的局部应用。数据库群中存放全组织(如企业)通用化、综合性的数据,某一类应用通常仅使用总体数据的子集,这样才能发挥数据共享的优势。

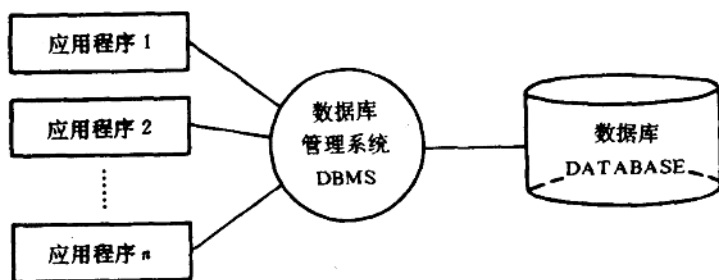


图 1.2.3 数据库系统数据与程序的关系

2. 采用特定的数据模型

整个组织的数据不是一盘散沙,必须表示出数据之间所存在的有机的关联,才能反映现实世界事物之间的联系。也就是说,数据库中的数据是有结构的,这种结构由数据模型表示出来。

文件系统只表示记录内部的联系,类似于描述事物的属性之间的联系,而不涉及不同文件记录之间的联系。要想在不同文件中查询相关的数据,必须编写一个程序。

例如有三个文件:图书(总编号,分类号,书名,作者,出版单位,单价);读者(借书证号,姓名,性别,单位,职称,地址);借阅(借书证号,总编号,借阅日期,备注)。要想查找某人所借图书的书名,此书的出版社及借阅者的职称,则必须编写一段不很简单的程序来实现。

数据库系统不仅表示描述一个事物的各个属性之间的联系,而且还能表示不同事物之间的联系。只要定义好数据模型,上述询问可以非常容易地联机查到。关于数据模型我们将在第四节中详细介绍。

3. 具有较高的数据独立性

在数据库系统中,DBMS 提供映象的功能,确保应用程序对数据结构和存取方法有较高的独立性。数据的物理存储结构与用户看到的逻辑结构可以有很大差别。用户只以简单的逻辑结构来操作数据,无需考虑数据在存储器上的物理位置与结构。

4. 有统一的数据控制功能

数据库作为多个用户和应用程序的共享资源,对数据的存取往往是并发的,即多个用户同时使用同一个数据库。数据库管理系统必须提供并发控制功能、数据的安全性控制功能和数据的完整性控制功能。

四、分布式数据库系统阶段

70年代后期之前,数据库系统多数是集中式的。分布式数据库系统是数据库技术和计算机网络技术相结合的产物,在80年代中期已有商品化产品问世。分布式数据库是一个逻辑上统一、地域上分布的数据集合,是计算机网络环境中各个结点局部数据库的逻辑集合,同时受分布式数据库管理系统的控制和管理,如图1.2.4所示。

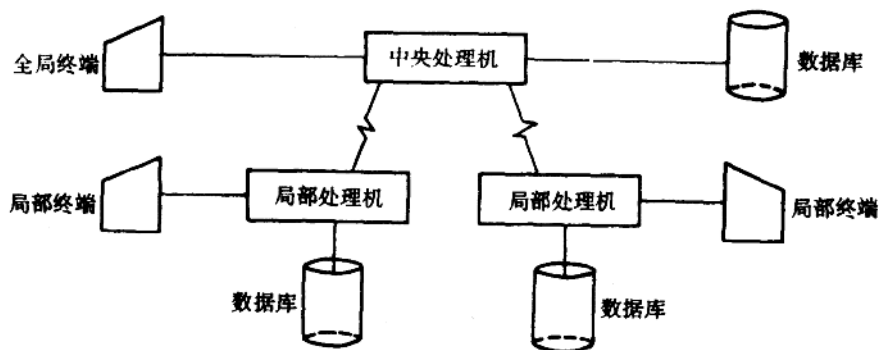


图 1.2.4 分布式数据库系统

分布式数据库在逻辑上象一个集中式数据库系统,实际上数据存储在处于不同地点的计算机网络的各个结点上。每个结点有自己的局部数据库管理系统,有很高的独立性。用户可以由分布式数据库管理系统(网络数据库管理系统),通过网络通信相互传输数据。分布式数据库系统有高度透明性,每台计算机上的用户并不需要了解他所访问的数据究竟在什么地方,就象在使用集中式数据库一样。其主要优点有:

1. 局部自主

网络上每个结点的数据库系统都具有独立处理本地事务的能力(大量的),而且各局部结点之间也能够互相访问、有效地配合处理更复杂的事务。因此,分布式数据库系统特别适合各

个部门的地理位置分散的组织机构。例如,银行业务、飞机订票、多公司企业管理等。

2. 可靠性和可用性

分布式系统比集中式系统有更高的可靠性,在个别结点或个别通信链路发生故障的情况下可以继续工作。一个局部系统发生故障不至于导致整个系统停顿或破坏,只要有一个结点上的数据备份可用,数据则是可用的。可见,支持一定程度的数据冗余是充分发挥分布式数据库系统优点的先决条件之一。

3. 效率和灵活性

分布式系统分散了工作负荷,缓解了单机容量的压力。数据可以存储在邻近的常用结点中;如果本结点的数据子集包含了要查询的全部内容,显然比集中式数据库在全集上查找节省时间。

系统易于实现扩展。例如,一个单位要增加新的机构,分布式数据库系统能够在对现有系统影响较小的情况下实现扩充。由此,扩大系统规模比集中式系统更加方便、经济、灵活。

第三节 数据库系统的组成

数据库系统是指引进数据库技术后的计算机系统,例如,一个以数据库为基础的管理信息系统。数据库系统实现有组织地、动态地存储大量的相关数据,方便多用户数据共享,从而减少数据冗余。

一、数据库系统的组成

数据库系统由五部分组成:硬件系统、数据库集合、数据库管理系统及相关软件、数据库管理员和用户,如图 1.3.1 所示。数据库管理系统 DBMS 在层次结构中的位置如图 1.3.2 所示。

1. 硬件系统

运行数据库系统的计算机需要有足够大的内存、足够大容量的磁盘等联机直接存取设备和较高的通道能力以支持对外存的频繁访问。还需要足够数量的脱机存储介质,如软盘、磁带存放数据库备份。

2. 数据库集合

系统包括若干个设计合理、满足应用需要的数据库。

3. 系统软件

数据库管理系统 (DBMS) 是为数据库的建立、使用和维护而配置的软件,它是数据库系统的核心组成部分,当然也离不开支持其运行的操作系统。如果不仅用数据库管理系统自含的语言开发应用系统,还需要使用其他程序设计语言及工具软件。

4. 数据库管理员

对于较大规模的数据库系统必须有人全面负责建立、维护和管理数据库系统。承担此任务

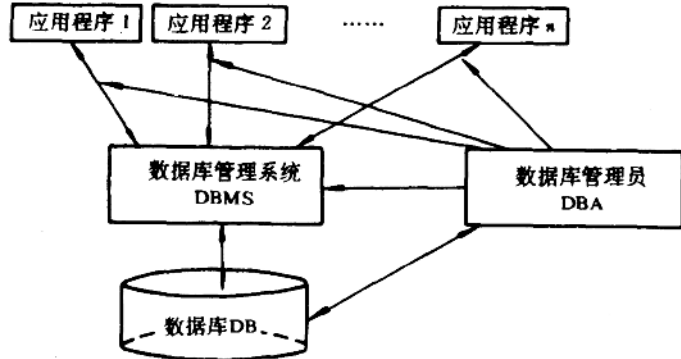


图 1.3.1 DBA 地位示意图

的人员称为数据库管理员 DBA (Database Administrator)。数据库管理员的职责包括:定义并存储数据库的内容,监督和控制数据库的使用,负责数据库的日常维护,必要时重新组织和改进数据库。

5. 用户

数据库系统的用户

分为两类:一类是最终

用户,主要对数据库进行联机查询或通过数据库应用系统提供的界面来使用数据库,这些界面包括菜单、表格、图形和报表;另一类是专业用户,即应用程序员,他们负责设计应用系统的程序模块,对数据库进行操作。

二、DBMS 的主要功能

数据库管理系统作为数据库系统的核心软件,其主要目标是使数据成为方便用户使用的资源,易于为各种用户所共享,并增进数据的安全性、完整性和可用性。这些是通过 DBMS 所支持的三级结构和两级独立性特点来实现的。

数据库具有三级结构,也称为三级模式。即在用户数据逻辑结构与数据的物理存储结构之间加入了数据的整体逻辑结构,整体逻辑结构涉及到所有用户的数据定义,是全局的数据视图。这三级模式之间可以有很大差别,为了实现这三级结构之间的转换,DBMS 提供相邻二级结构之间的映象。用户只需关心自己的局部逻辑结构就可以了,而不必关心数据在系统内的表示与存储。

数据库管理系统保证了数据和程序之间的物理独立性和逻辑独立性。所谓数据的物理独立性是指当数据的存储结构改变时,由系统提供数据的物理结构与逻辑结构之间的映象或转换功能,保持数据的逻辑结构不变,从而应用程序不需要修改。

数据的逻辑独立性是指,由系统提供数据的总体逻辑结构和面向某个具体应用的局部逻辑结构之间的映象或转换功能,当数据总体逻辑结构改变时,通过映象保持局部逻辑结构不变,从而应用程序也不需要修改。数据库的两级独立性和三级结构特点如图 1.3.3 所示。

在数据库系统中,数据是多个用户和应用程序的共享资源,已经从应用程序中完全独立出来,由 DBMS 来统一管理。DBMS 应该提供以下几个方面的功能:

1. 数据库的定义功能

提供数据定义语言 DDL (Data Description Language) 或者操作命令,以便对各级数据模式进行精确的描述。由此,系统必须包含 DDL 的编译或解释程序。这些数据模式并不是数据本身,而是具体 DBMS 所支持的数据模型的框架结构。用 DDL 所作的定义将被系统保留在数据字典中,以便在进行数据操纵和控制时使用。用户可以查阅数据定义以便共享数据库中的数据。

2. 数据操纵功能

为了对数据库中的数据进行追加、插入、修改、删除、检索等操作,DBMS 提供语言或者命

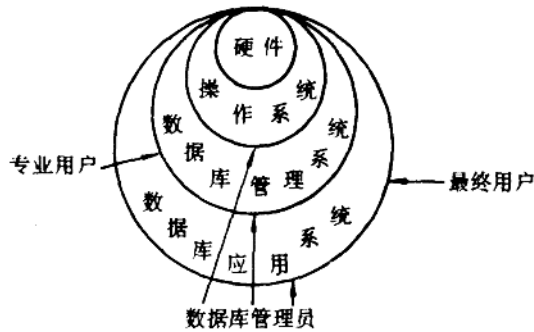


图 1.3.2 数据库系统层次示意图

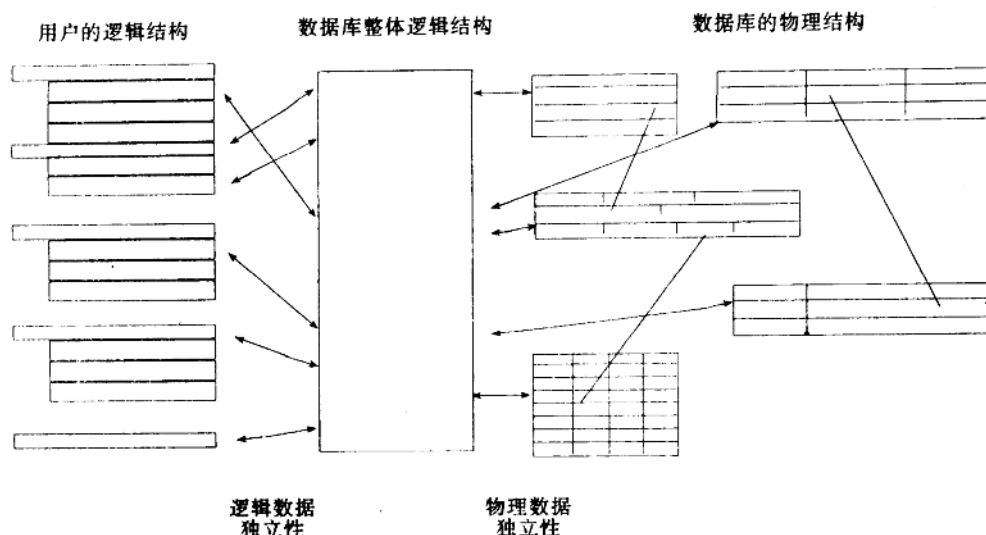


图 1. 3. 3数据库系统的结构

令,称为数据操纵语言 DML(Data Manipulation Language)。不同的 DBMS 语言的语法格式也不相同,以其实现方法而言,可分为两个类型:一类 DML 可以独立交互式使用,不依赖于任何程序设计语言,称为自含型或自主型语言;另一类是宿主型 DML,嵌入到宿主语言中使用,如嵌入 FORTRAN, COBOL, C 等程序设计语言中。在使用高级语言编写的应用程序中,需要调用数据库中的数据时,则要用宿主型 DML 语句来操纵数据。因此, DBMS 必须包含 DML 的编译或解释程序。

3. 数据库运行控制功能

数据库中的数据是提供给多个用户共享的,用户对数据的存取可能是并发的,即多个用户同时使用同一个数据库。DBMS 必须提供以下三方面的数据控制功能:

• 并发控制功能

对多用户并发操作加以控制、协调。例如,当某个用户正在修改某些数据项时,其他用户同时存取,就可能导致错误结果。如果两个用户同时修改同一数据时,先存储的修改就会丢失。数据库管理系统应对要修改的记录采取一定的措施,如锁定,暂时不让其他用户访问,待完成修改存盘之后再开锁。

• 数据的安全性控制

数据安全性控制是对数据库采用的一种保护措施,防止非授权用户存取造成数据泄密或破坏。例如,设置口令、确定用户访问密级和数据存取权限,系统审查通过后才执行允许的操作。

• 数据的完整性控制

数据完整性是数据的准确性和一致性的测度。系统应采取一定的措施确保数据有效、与数据库的定义一致。例如,当输入或修改数据时,不符合建立数据库时的定义或范围等规定的数据库系统不预接受。然而系统并不能保证所有输入数据绝对准确。例如,将日期 12/08/96 误录成 13/08/96 时,由于没有 13 月份,系统不予接受,必须重新录入。但是如果误录成 11/08/96

系统则无法控制。另外,当突然停电、出现硬件故障、软件失效或严重误操作时,系统应提供恢复数据库的功能,如定期转储、恢复备份等,使系统有能力将数据库恢复到损坏之前的某一个状态。

4. 数据字典

数据字典 DD(Data Dictionary)中存放着对实际数据库各级模式所作的定义,即对数据库结构的描述。这些数据是数据库系统中有关数据的数据,称之为元数据(metadata)。DD 提供了对数据库数据描述的集中管理手段,对数据库的使用和操作都要通过查阅数据字典来进行。数据字典经历了人工字典、计算机文件、专用数据字典系统和数据库管理系统与数据字典一体化四个发展阶段。专用的数据字典在系统设计、实现、运行和扩充各个阶段是管理和控制数据库的有力信息工具。

上述几方面是一般的 DBMS 所具备的功能。通常在大、中型计算机上实现的 DBMS 功能比较齐全,而在小型机,尤其是在微机上实现的 DBMS 功能相应有不同程度的减弱。

第四节 关系数据模型

前面多次提到,数据库中的数据是有结构的,这种结构反映出事物及事物之间的联系。任何一个数据库管理系统都是基于某种数据模型的,它不仅管理具体数据,而且要按照模型管理数据间的联系。一个实际数据模型应当反映全组织数据之间的整体逻辑关系。

数据模型由三部分组成,即模型结构、数据操作和完整性规则。其中模型结构是数据模型最基本的部分,它将确定数据库的逻辑结构,是对系统静态特性的描述。数据操作提供对数据库的操纵手段,主要有检索和更新两大类操作。数据操作是对系统动态特性的描述。完整性规则是对数据库有效状态的约束。

我们先介绍模型结构,数据库管理系统所支持的数据模型分为四种:层次模型、网状模型、关系模型、面向对象模型。传统的说法,有三种数据模型,即指前三种。

70 年代是数据库蓬勃发展的年代。层次系统和网状系统占据了整个商用市场,而关系系统仅处于实验阶段。80 年代关系系统逐步代替网状和层次系统而占领了市场。关系模型对数据库的理论和实践产生很大的影响,成为当今最流行的数据库模型。

一、关系模型

在计算机数据管理的历史上,出现过两次飞跃。第一次是数据库技术的出现,它使得数据管理技术步入了一个新的时代。第二次是关系数据模型的诞生,它标志着数据库技术走向成熟。

关系模型是用二维表的形式来表示事物和事物间联系的数据模型。从用户观点来看,关系的逻辑结构是一个二维表,在磁盘上以文件形式存储。例如,表 1.4.1 是一个职工基本情况表,表 1.4.2 是工资表。在关系数据库中,这两张表就是两个关系。在 FoxBASE 中,每一个表均作为一个数据库文件来存储。亦即 FoxBASE 的每个数据库文件就是一个关系。

因为关系模型的用户界面简单,易于为用户接受,并且有严格的理论基础,已成为最主要的数据模型。自 80 年代以来,新推出的数据库管理系统几乎都支持关系模型。早期的许多层次和网状模型系统的产品也加上了关系接口。我国微机上运行的关系数据库管理系统应用很广泛,这对促进我国数据库的应用有极大的推动作用。本书介绍的 FoxBASE 就是一个微机关系型数据库管理系统。