

GH

中等专业学校工科电子类教材

沈美琴 主编

数据库 原理及应用

(修订版)

西安电子科技大学出版社

73.967
SMU
C-1

中等专业学校工科电子类教材

数据库原理及应用

(修订版)

沈美琴 主编

西安电子科技大学出版社

1996

(陕)新登字 010 号

内 容 提 要

本教材共分数据库系统概论、常见的数据模型、FOXBASE+的基本概念、FOXBASE+数据库文件的建立、FOXBASE+的数据处理、FOXBASE+的命令文件、FOXBASE+应用、多用户 FOXBASE+初步、FoxPro 2.5 基础和 FoxPro 2.5 的应用等 10 章。书中详细阐述了关系型数据库系统 FOXBASE+的各类文件的建立和使用,对 FOXBASE+的各类命令的使用作了详尽的描述,对应用软件的设计与调试技巧作了较全面的阐述,并对最新关系型数据库 FoxPro 2.5 的基本概念、特点以及各类命令的使用与程序设计方法作了必要的介绍。

本书既适合中等专业学校计算机专业作为教材用书,同时也适宜于各类大专层次有关专业学生使用。对有关工程技术人员来说,本书也不失为一本实用技术参考书。

中等专业学校工科电子类教材

数据库原理及应用

(修订版)

沈美琴 主编

责任编辑 云立实

西安电子科技大学出版社出版

陕西省高陵县印刷厂印刷

陕西省新华书店发行 各地新华书店经售

开本 787×1092 1/16 印张 21 字数 499 千字

1994 年 5 月第 1 版 1996 年 11 月修订版 1996 年 11 月第 6 次印刷 印数 52 001-62 000

ISBN 7-5606-0458-7/TP·0202(课)

定价: 16.80 元

出版说明

根据国务院关于高等学校教材工作的规定，我部承担了全国高等学校和中等专业学校工科电子类专业教材的编审、出版的组织工作。由于各有关院校及参与编审工作的广大教师共同努力，有关出版社的紧密配合，从1978~1990，已编审、出版了三个轮次教材，及时供给高等学校和中等专业学校教学使用。

为了使工科电子类专业教材能更好地适应“三个面向”的需要，贯彻国家教委《高等教育“八五”期间教材建设规划纲要》的精神，“以全面提高教材质量水平为中心，保证重点教材，保持教材相对稳定，适当扩大教材品种，逐步完善教材配套”，作为“八五”期间工科电子类专业教材建设工作的指导思想，组织我部所属的九个高等学校教材编审委员会和四个中等专业学校专业教学指导委员会，在总结前三轮教材工作的基础上，根据教育形势的发展和教学改革的需要，制订了1991~1995年的“八五”（第四轮）教材编审出版规划。列入规划的，以主要专业主干课程教材及其辅助教材为主的教材约300多种。这批教材的评选推荐和编审工作，由各编委会或教学指导委员会组织进行。

这批教材的书稿，其一是从通过教学实践、师生反映较好的讲义中经院校推荐，由编审委员会（小组）评选择优产生出来的，其二是在认真遴选主编人的条件下进行约编的，其三是经过质量调查在前几轮组织编定出版的教材中修编的。广大编审者、各编审委员会（小组）、教学指导委员会和有关出版社，为保证教材的出版和提高教材的质量，作出了不懈的努力。

限于水平和经验，这批教材的编审、出版工作还可能有缺点和不足之处，希望使用教材的单位，广大教师和同学积极提出批评和建议，共同为提高工科电子类专业教材的质量而努力。

机械电子工业部电子类专业教材办公室

前 言

本教材系按机械电子工业部的工科电子类教材 1991~1995 年编审出版规划,由计算机专业教学指导委员会征稿并推荐出版。责编委曾玉琨。

本教材由上海市轻工业学校沈美琴担任主编,上海第二工业大学何守才担任主审。

本课程的参考学时为 80 学时,其主要内容共分为 10 章。第一章通过对数据库的基本概念、发展趋势及现实世界数据描述的介绍,使读者对数据库系统有一概括的了解;第二章介绍了常见的数据模型:层次数据库、网络数据库及关系数据库的基本概念及代表性系统;第三章到第八章详细介绍了关系数据库系统 FOXBASE+,对 FOXBASE+的各类文件的建立和使用,各种命令的作用均作了大量详尽的描述,并对一个应用程序作了全面的分析,同时详细介绍了应用软件的设计及调试技巧;第九章介绍了 FoxPro 2.5 的基本概念和特点;第十章介绍了 FoxPro 2.5 的各类命令的使用及程序设计方法。

本教材中作了“*”标记的章节可以作为阅读材料处理,也可只作简单介绍,重点是第四章~第七章。

由于编者水平有限,书中难免还存在一些缺点和错误,殷切希望广大读者批评指正。

编 者

目 录

第一章 数据库系统概论	1
第一节 数据管理概述	1
第二节 现实世界的数据库描述	6
第三节 数据库的基本概念	12
第四节 数据库系统	20
第五节 数据库的完整性和安全性	25
习题一	27
第二章 常见的数据模型	28
第一节 层次数据库	28
第二节 网络数据库	36
第三节 关系数据库	47
习题二	77
第三章 FOXBASE+ 的基本概念	81
第一节 FOXBASE+ 的系统环境和主要性能指标	81
第二节 FOXBASE+ 的文件类型和数据类型	84
第三节 FOXBASE+ 的变量、函数和表达式	87
第四节 FOXBASE+ 的句法和操作规则	99
习题三	110
第四章 FOXBASE+ 数据库文件的建立	112
第一节 数据库的建立	112
第二节 数据输入	115
第三节 数据输出	120
第四节 数据库结构的修改	123
习题四	127
第五章 FOXBASE+ 的数据处理	128
第一节 记录的删除	128
第二节 数据的编辑	131
第三节 排序和索引	134
第四节 数据的检索	140
第五节 数据的统计	144
第六节 多工作区操作	149
第七节 文件操作	157
习题五	160
第六章 FOXBASE+ 的命令文件	161
第一节 命令文件的建立和调用	161
第二节 设置命令	163
第三节 程序交互式数据输入	166

第四节	定位格式输出	171
第五节	几种程序语句	174
第六节	过程及调用	182
第七节	模块设计和菜单技术	185
习题六		192
第七章	FOXBASE+应用	194
第一节	应用软件的设计	194
第二节	应用实例	196
第三节	报表处理	250
第四节	标签输出	255
第五节	FOXBASE+与其他高级语言接口	258
习题七		268
第八章	多用户 FOXBASE+初步	271
习题八		281
第九章	FoxPro 2.5 基础	282
第一节	概论	282
第二节	FoxPro 的特点	290
习题九		294
第十章	FoxPro 2.5 的应用	296
第一节	FoxPro 2.5 的菜单系统	296
第二节	FoxPro 2.5 各类命令的使用	298
第三节	FoxPro 2.5 程序设计	305
习题十		310
附录一	FOXBASE+命令菜单	313
附录二	FOXBASE+函数一览	323

第一章 数据库系统概论

数据库是当前计算机系统的一个重要组成部分。数据库技术是在数据管理技术的发展中逐渐形成的,它的概念、原理和方法还在继续发展和变化着。随着科学技术的进步,反映客观世界和人类生产活动、社会活动的的数据量急剧增长,使用数据库对信息进行管理已经成为计算机应用的又一广阔领域。

第一节 数据管理概述

数据管理是数据处理的中心问题,它指的是对数据的组织、编目、定位、存储、检索和维护等内容。

由于社会生产力的高速发展,新技术层出不穷,信息量急剧膨胀,整个人类社会的信息化程度越来越高,人类对信息处理的方法也越来越先进。随着计算机软、硬件的发展,数据管理的内容已从数值计算发展到对文字、声音、图像……各类信息进行处理。

一、数据管理技术的发展

至今为止,数据管理的发展经历了 3 个阶段,即人工管理阶段、文件系统管理阶段和数据库系统管理阶段。

1. 人工管理阶段

50 年代中期以前,计算机主要应用于科技计算。当时计算机的硬件情况是:没有磁盘等直接存取的存储设备,只有磁带、卡片、纸带等。软件的情况是:没有操作系统,没有管理数据的软件;处理数据的方式是批处理。

人工管理阶段的特点:

① 由于主要应用于科技计算,一般不需要长期保存数据,只需在计算时将数据输入计算机,用完后就撤去,因而不保留数据。

② 对数据进行处理时,程序员不仅要规定数据的逻辑结构,还要在程序中设计物理结构(存储结构)及存取方法(输入输出方法)等,程序中存取数据的子程序会随着存储的改变而改变。也就是说,没有软件系统对数据进行管理,数据与程序不具有独立性,数据的存储一旦变化就必须修改程序。

③ 基本上没有文件的概念,即使有,也大多是顺序文件。如果使用时出现非顺序文件,就必须由程序员自行设计组织方式和存取方式,没有进行转换的专用软件。

④ 数据面向应用,一组数据对应一个程序(如图 1-1)。由于各应用程序所处理的数据之间不可能完全没有关系,所以程序之间会有大量的重复数据。

2. 文件系统管理阶段

从 50 年代后期到 60 年代中期,计算机不仅应用于科技计算,还大量应用于管理,计算机的软硬件有了很大变化和发展。硬件情况:外存有磁盘、磁鼓等设备,尤其是磁盘成为联机操作主要的直接存取外存设备。软件情况:操作系统中包含了文件系统(即信息管

理模块),专门用于管理数据;数据处理的方式有计算机文件批处理和实时联机处理。

文件系统管理阶段的特点:

① 数据长期保留在外存上。

由于计算机大量用于数据处理,对文件要反复进行查询、修改、插入和删除等操作,因而需要保留数据。

② 程序和数据有一定的独立性,文件的逻辑结构与物理结构已有所区别。

由管理软件进行数据管理,程序与数据之间由存取方法进行转换,程序库中有共同的查询数据和修改数据的例行程序。

③ 文件的物理结构中除了直接存取方式外,还增加了索引文件和链接文件,对文件中的记录可以顺序地访问,也可以随机地访问。

④ 数据以文件为单位共享(如图 1-2 所示)。数据的物理存储的改变仍需修改用户的应用程序,因此数据仍面向应用。文件不易扩充,修改浪费时间,数据冗余度较大。

⑤ 数据的存取以记录为单位。

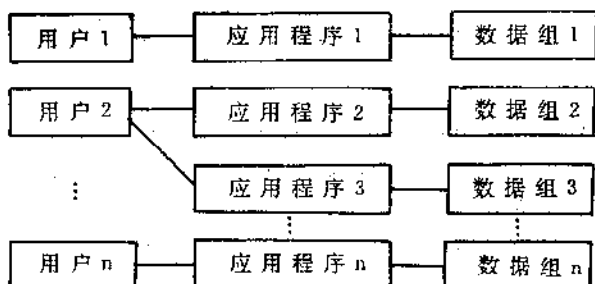


图 1-1 人工管理阶段数据与程序对应关系

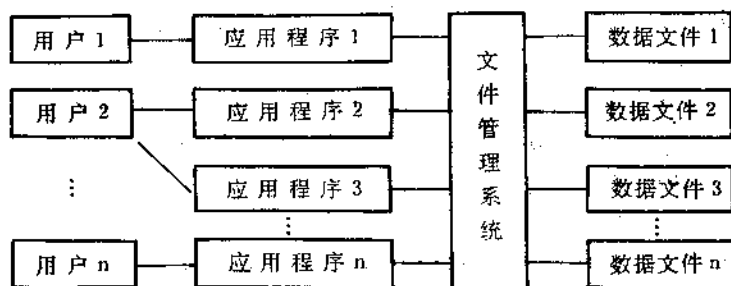


图 1-2 文件系统管理阶段数据与程序对应关系

3. 数据库系统管理阶段

60 年代末期,管理规模越来越大,数据量随之急剧增加,数据共享的要求也更高。此时计算机的硬件情况是:有了大容量的磁盘。硬件价格下降,而软件的销售价格上升。为降低编制和维护系统软件及应用程序的成本,解决多用户共享问题,出现了数据库系统。数据库的数据处理方式中更多地采用了实时联机处理,并开始考虑分布处理。

初期的数据库系统有如下特点:

① 对数据进行严格细致的描述,使文件、记录、数据项等数据单位之间的联系清晰,结构简单。

② 允许用户以记录或数据项作为单位进行访问,也允许关键字检索和文件之间的交叉访问。

③ 解决了数据的应用独立于数据的存储这个问题。数据的物理存储可以很复杂,同样

的物理数据可以导出多个不同的逻辑文件，用户以简单的逻辑结构操作数据时，不需要考虑数据的存储情况；改动数据的物理位置和存储结构时不必修改或重写程序；用户的逻辑数据与它们的物理存储之间的转换由管理软件完成。数据库系统中程序与数据的关系如图 1-3 所示。

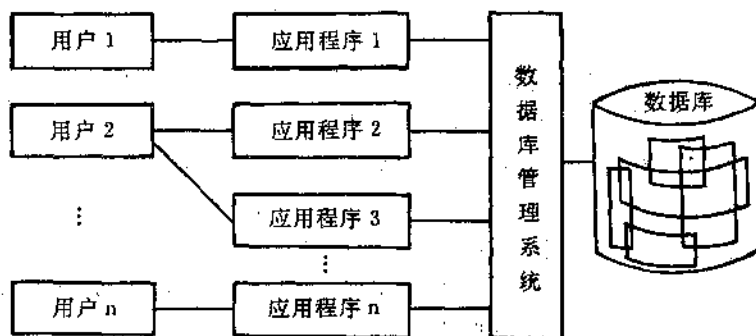


图 1-3 数据库系统中数据与程序的关系

初期的数据库系统中数据整体逻辑仅仅是用户逻辑的简单并集，在用户增多、逻辑文件日渐庞杂的情况下，数据库的组织越来越乱。为了提高效率，减少冗余，增加新数据，常常需要改变数据的整体逻辑结构，这时会导致用户逻辑文件的修改。这种情况促使人们把用户观点的逻辑结构从整体结构中独立出来，形成了现今的数据库系统。它具有数据独立、数据共享和数据结构化等特点，我们将在本章的第三节中详细阐述。

目前的数据库具有三级结构和两级独立性。三级结构是指用户数据逻辑结构、数据的物理存储结构和数据的整体逻辑结构；数据的两级独立性是指数据的物理存储的变化不影响整体的逻辑结构或用户程序，数据整体逻辑结构的改变也不影响用户应用程序。

“物理”和“逻辑”这两个词分别用来说明数据的两种形式。用户或应用程序员使用的数据结构称为逻辑结构；实际存储在磁盘、磁带或其他存储介质上的数据结构称为物理结构。两者之间的区别见图 1-4 所示。

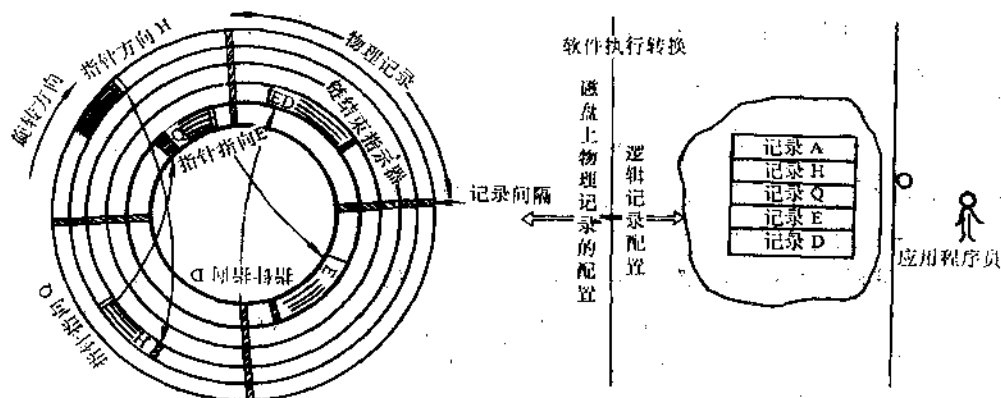


图 1-4 逻辑记录与物理记录的区别

磁盘上的物理记录包含了几个连在一起的逻辑记录,用一个指针把一个逻辑记录链接到另一个逻辑记录上,应用程序员按链的顺序请求访问逻辑记录文件,没有必要知道这个链的具体内涵。软件可以按照各种需要的顺序把逻辑结构传送给不同的应用程序,用户不必考虑物理数据结构。

二、中文信息处理

在我国,绝大多数的数据信息是中文汉字,因而中文信息处理就成了我国计算机系统应用和研究的一个重要领域,它直接关系到计算机能否在我国普及。

中文信息处理的范围很广,包括汉字信息处理、中文编辑排版、中文通信、中文情报检索、中文语言自动处理、中外文互译、中文数据库等。汉字信息处理又包括汉字输入、汉字库、汉字输出三个重要部分,其中汉字输入是其他中文信息处理工作的前提和基础。在计算机系统中普及中文信息的处理,必须解决汉字的输入输出问题。这正是国内外许多科技人员集中力量主攻的汉字信息处理的难题。

近年来,不少计算机系统已装备了汉字输入输出设备或汉字终端,并出现了一批科技成果。目前常见的是在 IBM-PC 及其兼容机上运行的汉字磁盘操作系统、CCDOS 和 APPLE 机上的汉字系统。为了使微机系统能运行中文操作系统,很多厂商生产微机时在维持原来机器的硬件系统(键盘、主机、显示器、打印机)的结构形式和软件功能的基础上,增加了一些专用的外部模块以专门处理汉字,使微机系统具备了处理汉字的良好软硬件环境。

计算机系统处理中文信息,首先要解决汉字进入计算机或转换成计算机代码的问题。汉字在不断演变中,其总数无法精确估计,而且繁体和简体、古今各种书体并存,字数在六万个以上,要全体进入计算机是很困难的,也是没有必要的。经过研究分析,发现常用汉字只是其中一个小小的子集。为了适应中文信息处理的需要,我国在 1981 年制定了“信息交换用汉字编码字符集基本集 GB2312-80”国家标准,规定常用汉字总数为 6 367 个,其中一级字 3 755 个,二级字 3 008 个,并给每个汉字分配了代码,将它们作为汉字信息交换标准代码或机器内码。在这些常用字中,有 103 个汉字使用频率接近 40%,其中“的、一、在、是、了、不、和、有”8 个汉字的使用频率超过 10%。因此对这些使用频率高的汉字,在信息处理中还应有特殊的考虑,以提高汉字检索的速度。下面我们就汉字信息处理的三个重要部分分别作一些介绍。

1. 汉字输入

一般情况下,把汉字输入计算机有 3 种方法,即汉字图像识别、汉字语言识别和汉字编码输入。迄今为止,已提出的汉字编码输入方案已达数百种之多,有些已在实际系统中使用。各类编码的输入方案有许多相似之处,可以大致归纳成以下几类。

① 邮电编码。采用我国邮电部的“标准电码本”的编码作为汉字的输入代码,一字一码,每码的定长是 4 位十进制数字。这种方法的优点是输入方法极简单,可在通用键盘上操作,其缺点是数码与汉字音形特征间没有规律性联系,一般人无法记忆。

② 整字输入。基本上也采用一字一码,在有 3 000~4 000 字的大键盘上,以一定的规则(如按词组联想、部首为序或常用字集中等)排列汉字。经过专门训练的人员每分钟可输入 50~60 个汉字。这种方法的缺点是设备庞大复杂,价格昂贵,非专门人员不易使用。目

前经过改进,使用笔触式输入,缩小了键盘。近来出现了光键盘(汉字一页一页显示在屏幕上),用户用光笔在光键盘上取字,大大提高了汉字输入的速度。

③ 字元编码。字元是组成汉字的基本构件。给每个字元分配一个代码,则每个汉字可由若干代码表示。根据字元位置(上下、左右、内外)顺序将其代码排列得到汉字的代码。由于字元个数明显少于汉字个数,因此可用中等键盘。这种方法的缺点是汉字字元的拆拼方法较难掌握,一般用户使用不方便。

④ 笔形编码。通过汉字笔划分类,给每一类笔划一个代码,按汉字笔法顺序或四角顺序得到汉字代码。这种方法的优点是可以在小键盘上输入,缺点是汉字取笔较难掌握。

⑤ 拼音编码。在通用键盘上像西文一样进行输入。我国的现代汉语拼音已被联合国承认,用若干字母(按拼音规则)表示一个汉字。问题是同音字难以处理(现采用同音异字在屏幕上同时显示由用户选择输入的办法),发音不准或不知道如何发音的汉字无法输入。

⑥ 组合编码。这种方法综合上述各类编码的优点,如最常用字采用一字一键以提高输入速度;以音为主或以形为主的,采用音形结合以克服拼音重码多和笔形拆字难等缺点。目前可行的方法多属组合编码。

汉字输入要在中西文操作系统 CCDOS 支持下进行。在 IBM-PC 机上运行的 CCDOS 有国标码、电报码、音韵码、首尾码、区位码和拼音码等汉字处理输入方法,其中区位码、首尾码和拼音码较为常用。目前崛起的五笔字形汉字输入以其输入速度快、与四通汉字处理机通用的优点成为十分受欢迎的输入方式。

2. 汉字库

汉字库是汉字信息处理的重要部分,也称汉字发生器。汉字库分成活字式、固定文字板式、活动文字板式和点阵式几种。在计算机系统中常用的是点阵式汉字库,其中包括汉字标准代码、点阵字模和索引信息等,一般存放在磁盘等存储设备中。点阵汉字产生的文字质量、文字字体与点阵数目有关。目前使用最广的是宋体字。点阵数目有如下几种:

① 16×16 点阵。它可表示简单汉字,难以表示复杂汉字;可以表示横和竖,难以表示撇和捺。

② 24×24 点阵。它表现宋体汉字可达到实用程度,但表现少数复杂汉字还有一定困难。

③ 32×32 点阵。有较强的表现能力,近似于参考字体,基本上满足实用要求。

④ 128×128 点阵。表现能力等效于参考字段,质量很高,可用于中文编辑排版。其缺点是大量耗费存储空间,无法在存储量不大的系统中使用。

3. 汉字输出

汉字输入的逆过程就是汉字输出,其过程是将汉字代码再转换回汉字。汉字输出分为屏幕显示和打印输出两类。中文信息系统对汉字打印机的要求是:设备适宜,打印速度快,字体质量好,操作噪音小,功能齐全,操作简单,维护方便,价格便宜等。目前汉字打印机的概况在表 1-1 中列出,这些打印机各有优缺点,有的使用较广,有些则较少使用。

表 1-1 各种汉字打印机

打印机种类	工 作 原 理	优 缺 点
活字击打式	用电子机械装置驱动活字或字锤击打色带或纸,把汉字记录在普通纸上	打印的文字质量高,保存方便 机构复杂、设备庞大、噪音大、可靠性差,逐渐被淘汰
点阵击打式	用针头击打色带,由点阵在普通纸上形成汉字	打印机结构简单,设备小重量轻,使用普通纸,记录便于保存,使用方便,价格便宜,在市场上占绝对优势 印字质量比活字差,有噪音
电灼式	在纸上涂覆碳粉精和铝蒸发层,用电极放电灼灼铝层使其蒸发,暴露着色层显现字迹	印字速度快,文字质量高 机器运转费用高,使用特殊纸,有气味
热敏式	在纸基上涂覆着色层,用电流给印字头加热,使记录纸产生着色反应,显现字迹	打印机体积小,重量轻,噪音小,价格便宜 打印速度受限制,记录结果不易保存
喷墨式	使高速带电墨水粒子通过记录信号控制的电场到达记录纸形成字迹	打印机体积小,重量轻,几乎无噪音,用普通纸,印字质量高,记录速度较快,有发展前途
静电式	利用静电现象形成汉字潜像,使潜像经过显影、定影记录在静电纸上	整机结构简单,维护方便,可靠性好,无噪音,印字速度高,质量好 操作复杂,机器运转费用高
电子照相光纤式	用光线使带电鼓面曝光,生成静电潜像,通过显影后转印在普通纸上	印字速度快,分辨率高,字迹清晰,用普通纸,费用低,便于推广,适用于大、中型主机和汉字信息处理系统配套
电子照相激光式	用激光束在感光体上扫描,生成静电潜像,经显影后转印在普通纸上	输出速度快,印字质量好 设备庞大,价格昂贵

第二节 现实世界的数据描述

我们知道,在数据库中存储数据的目的,主要是帮助人们去处理和控制与这些数据相关的事物。因此,数据库往往不是独立存在的,它通常是一个更大的信息控制系统的一部分,如图 1-5 所示。人们通过对客观事物的观测,获得大量的信息,对这些信息进行规范化(即记录、整理和归类)、数据化后送入数据库保存,这时有一部分信息直接送入控制决策机构。为了确定策略,控制决策机构向数据库发出询问,并利用响应后提供的信息及其他有关信息作出决策后,再行控制客观事物。

例如,一个工厂的生产管理系统中,客观事物可指产品的产量、质量、销售情况等;控制决策机构是生产领导部门;策略是指提高产品产量或质量的具体措施。由数据库管理人

员把得到的信息(产量、质量、销售情况)规范化、数据化后送入数据库,生产领导部门可以通过一定的操作询问,向数据库了解产品的各种情况,根据这些情况并参考市场需求信息和对产品质量的反馈信息,提出对产品的产量或质量的修正意见及措施,然后在实际生产中加以实施(控制客观事物)。

下面我们来讲解信息的循环过程。

一、信息和数据

信息(Information)是向人们提供关于现实世界新的事实的知识,数据(Data)是用以载荷信息的物理符号,两者不可分离又有一定区别。确切地说,数据是信息的具体表现,而信息是消化了的数据,但信息是更基本的反映现实的概念。我们可以把数据管理说成信息管理,但信息管理不一定是数据管理。

图 1-5 的信息管理系统中,信息从客观事物出发,规范化、数据化后,流经数据库,通过决策机构最后又回到现实世界,这一循环过程经历了 3 个领域:现实世界、观念世界和数据世界。

现实世界是存在于人脑之外的客观世界,这个世界包括客观事物和它们间的相互联系。

观念世界也叫信息世界,是现实世界在人们头脑中的反映。客观世界中的客观事物在观念世界里称为实体。反映事物间联系的是实体模型。

数据世界也叫机器世界。信息经加工编码(数据化)后成为数据,每一实体的数据称为记录。事物及联系在数据世界中用数据模型来描述。

从现实世界到数据世界,客观事物及其联系经历了两级抽象描述——实体模型和数据

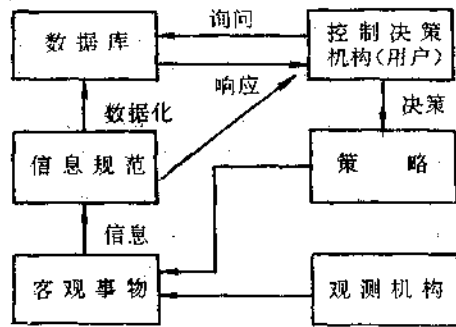


图 1-5 信息控制系统

解释过程

抽象过程

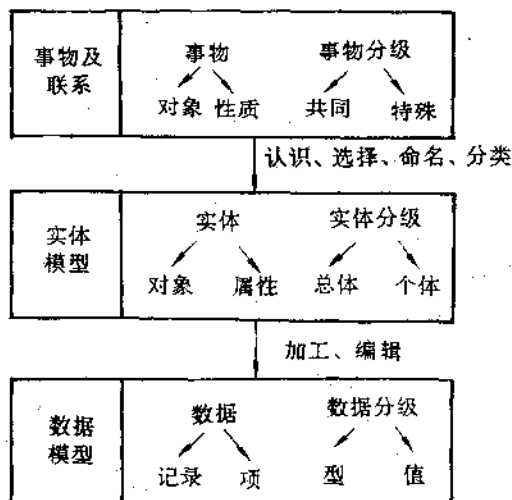
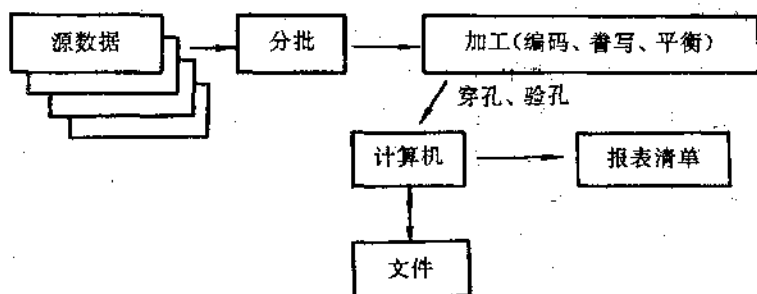


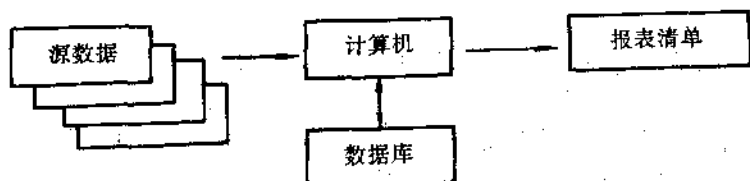
图 1-6 信息的三个领域

模型。数据库的核心是数据模型。为了得到正确的模型，首先要充分了解客观事物。信息循环流经 3 个领域时的内容及其联系可在图 1-6 中清晰地反映出来。

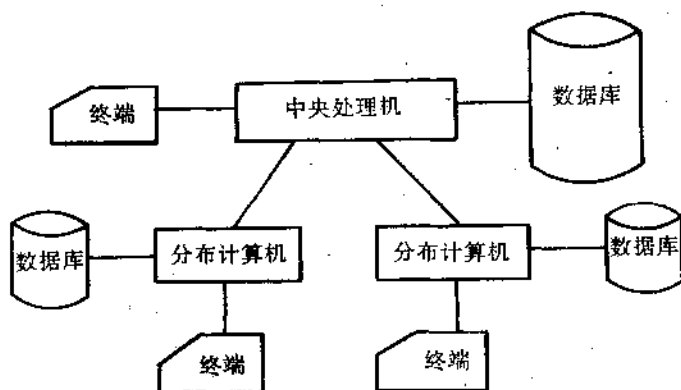
随着人类社会的发展，信息在各种活动中居于越来越重要的地位，它对人类活动，特别是在管理工作中起非常重要的作用。信息系统就是为企业或组织提供所需信息的工具。一个电子计算机化的信息系统称为人-机系统。在人-机系统中可以从两个角度来看待信息系统：从人的角度看，信息系统是研究数据管理的，称为用户观点；从机器的角度看，信息系统是研究数据处理方式的，称为机器观点。目前，数据处理的方式可分为批处理系统、联机集中处理系统和联机分布处理系统(如图 1-7)。



(a) 文件批处理



(b) 联机处理



(c) 分布式处理

图 1-7 数据处理方式

二、实体模型

在观念世界中，我们用实体描述客观事物。术语“实体”没有确切的定义，只能说实体是存在的且可分的，是任何一个我们所关心的“事物”。这事物可以指人，也可以指物；可以指概念性的东西，也可指事物与事物的联系。实体模型则是确定数据库包含哪些信息内容的关键，是设计数据库的先导。如图 1-6 所示，在实体模型中，实体可分为“对象”和“属性”两大类，如人、椅子、工厂、学校、王小红、汽车、人民大学等都是“对象”；“属性”表示“对象”的某种特征，如对象“人”有姓名、年龄、性别、民族、籍贯等特征。属性可分为原子属性和可分属性两类：不能再细分的属性称为原子属性，如性别、颜色等；可以再细分的属性称为可分属性，如工资还可分为应发工资、扣发金额和实发工资等。可分属性由原子属性和可分属性组成。

实体模型的对象和属性是客观事物中对象与性质的抽象描述，它们既有联系又有区别。联系是绝对的：一个对象具有某些属性，若干属性描述某个对象。区别是相对的：一个对象具有某个属性，可能又是另一些属性描述的对象。如“汽车”是对象，它具有型号、颜色、所属单位等属性，其中“所属单位”又可作为对象，被单位名、单位地址、电话号码、……等属性来描述。对象与属性之间具有相对性的原因是描述的事物不同，观察研究问题的角度不同。在构造实体模型时，必须辩证地研究客观事物。

在实体模型中，实体又可分为两级：一级是个体，指的是单个的能互相区别的特定实体，如王小红、人民大学、上海机床厂等；另一级是总体，指的是某类个体组成的集合，如“人”可指王小红、李平、张连生、……等个体组成的集合，“学校”指人民大学、上海机电工业学校、向明中学等个体组成的集合。

在各自包含有若干个体的两个总体之间，有以下 3 种联系方式。

① 一对一联系(见图 1-8(a))。如果总体 A 中的任一个体至多对应总体 B 的一个个体，反之，总体 B 中的任一个体至多对应总体 A 中的一个个体，则称 A 对 B 是一对一的联系。

例如，医院中病房内病人与床位之间，班级与正班长之间都是一对一联系。

② 一对多联系(见图 1-8(b))。如果总体 A 中至少有一个个体对应

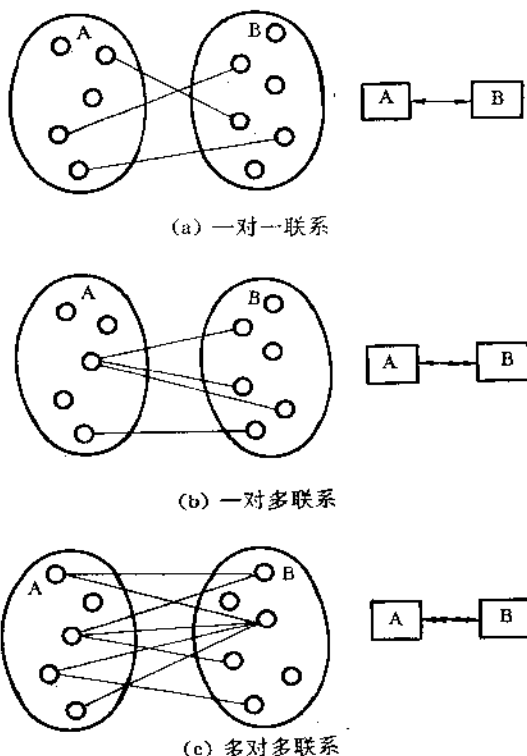


图 1-8 总体之间的联系方式

于总体 B 中一个以上个体,反之,总体 B 中任一个体至多对应于总体 A 中一个个体,则称 A 对 B 是一对多联系。

例如,学校对学生,父亲对子女等都是一对多联系。

③ 多对多联系(见图 1-8(c))。如果总体 A 中至少有一个个体对应总体 B 中一个以上个体,反之,总体 B 中也至少有一个个体对应 A 中的一个以上个体,则称 A 对 B 是多对多联系。

例如,教师与学生,商店与商品等都是多对多联系。

在考察和研究了客观事物与它们之间的联系后,就可以开始对实体模型进行描述。在实体模型中,要对每一个实体加以命名,并描述其间的各种联系。

例 工厂实体模型。

用矩形表示对象,用椭圆表示属性,我们可作出工厂实体模型图如图 1-9。

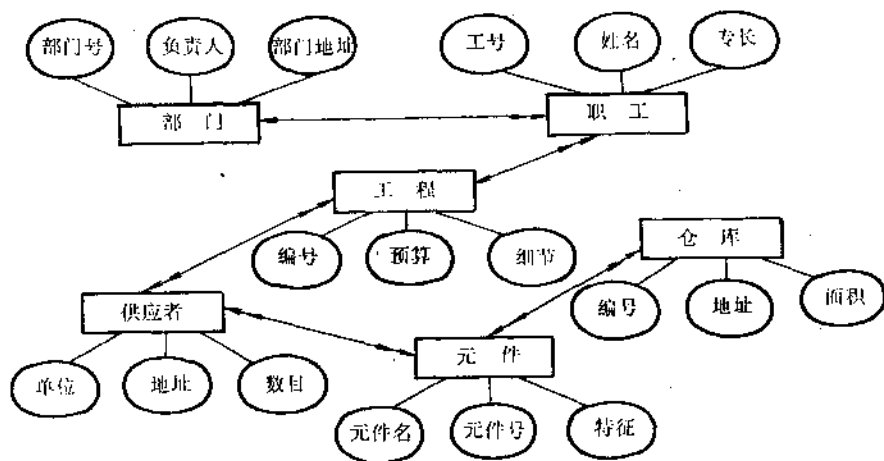


图 1-9 工厂实体模型

在工厂实体模型中,描述了部门、职工、仓库、元件、工程等对象。部门具有属性:部门号、负责人和部门地址;职具有属性:工号、姓名和专长;工程具有属性:编号、预算和细节;……等等。在此模型中,部门与职工为一对多联系,因为一个部门可以有多个职工,而一个职工只能在一个部门供职。同理,工程与职工也是一对多联系。供应者与工程、供应者与元件、仓库和元件是多对多联系,读者可自行解释其原因。

三、数据模型

数据模型是数据化后的实体模型,是对客观事物及其联系的数据描述,是数据库设计的核心之一。

用数据来描述实体时,用以描述对象的数据称为“记录”;用以描述属性的数据称为“项”。一个对象具有若干属性,因此一个记录也由若干项组成;属性分为原子属性和可分属性,项也分为基本项(对应于原子属性,是具有名称的最小逻辑数据单位)和组合项(对应于可分属性,由基本项和组合项组成)。一般情况下,数据项名就用属性名表示,但含义不同。作为属性名时,它表示观念信息;作数据项名时,则表示数据信息,它包含了项的特