

电子计算机应用系列教材

# 微型计算机数据库技术 及其应用

张淑奇 胡绪英 编著  
陈莘萌 樊成



科学出版社

电子计算机应用系列教材

# 微型计算机数据库技术及其应用

张淑奇 胡绪英 编著  
陈莘萌 樊成

科学出版社

1993

(京) 新登字 092 号

## 内 容 简 介

本书为电子计算机应用系列教材之一,讲述微机的数据库技术。全书共十一章。第一章扼要地介绍了数据库系统的概貌、由来和发展。第二至第四章讲述了关系方法、层次方法和网状方法。第五章介绍数据库系统设计的理论、方法和步骤,重点介绍了关系数据库的设计。第六章描述数据库的保护。第七、九、十一章分别讨论了 INGRES, dBASE III 和 dBASE III PLUS 三个数据库管理系统。第八、十两章提供了两个数据库应用系统设计与实现的实例。

本书较好地体现了理论与实践的结合,不仅有较系统的理论,而且有丰富的例子,且文字表达简洁、流畅、通俗易懂。

本书适宜作为计算机培训班的教材,可供微机数据库应用系统的开发者、设计者学习、参考。

35008 / 17 16

## 电子计算机应用系列教材 微型计算机数据库技术及其应用

张淑奇 胡绪英 编著  
陈莘萌 樊成  
责任编辑 那莉莉

科学出版社出版

北京东黄城根北街16号

邮政编码: 100707

北京市华星计算机公司激光照排

国防科工委印刷厂印刷

新华书店北京发行所发行 各地新华书店经售

\*

1993年1月第一版 开本: 787×1092 1/16

1993年1月第一次印刷 印张: 18½

印数: 0001—5300 字数: 424000

ISBN 7-03-001345-X/TP·83

定价: 10.50元

# “电子计算机应用系列教材”主持、组织编著单位

## 主持编著单位：

国务院电子信息系统推广应用办公室

## 组织编著单位：

广东、广西、上海、山东、山西、天津、云南、内蒙古、  
四川、辽宁、北京、江苏、甘肃、宁夏、江西、安徽、  
河北、河南、贵州、浙江、湖北、湖南、黑龙江、福建、  
新疆、广州、大连、宁波、西安、沈阳、武汉、青岛、  
重庆、哈尔滨、南京等35省、市、自治区、计划单列市

电子振兴  
计算机领导小组办公室  
科技工作

# “电子计算机应用系列教材”联合编审委员会名单

(以姓氏笔划为序)

## 主编审委员：

王长胤\* 苏世生 何守才 陈有祺 陈莘萌\* 邹海明\* 郑天健  
殷志鹤 童 颖 赖翔飞 (有“\*”者为常务主编)

## 常务编审委员：

于占涛 王一良 冯锡祺 刘大昕 朱维华 陈火旺 陈洪陶 余 俊  
李 祥 苏锦祥 佟震亚 张广华 张少润 张吉生 张志浩 张建荣  
钟伯刚 胡秉光 高树森 徐洁盘 曹大铸 谢玉光 谢育先 韩兆轩  
韩培尧 董继润 程慧霞

## 编审委员：

王升亮 王伦津 王树人 王振宇 王继青 王翰虎 毛培法 叶以丰  
冯鉴生 刘开瑛 刘尚威 刘国靖 刘晓融 刘德镇 孙令举 孙其梅  
孙耕田 朱泳岭 许震宇 何文兴 陈凤枝 陈兴业 陈启泉 陈时锦  
邱玉辉 吴宇尧 吴意生 李克洪 李迪义 李忠民 迟忠先 沈林兴  
肖金声 苏松基 杨润生 吕福德 张志弘 张银明 张 勤 张福源  
张翼鹏 郑玉林 郑 重 郑桂林 孟昭光 林俊伯 林钧海 周俊林  
赵振玉 赵惠溥 姚卿达 段银田 钟维明 袁玉馨 唐肖光 唐楷全  
徐国平 徐拾义 康继昌 高登芳 黄友谦 黄 侃 程锦松 楼朝城  
潘正运 潘庆荣

## 秘书组：

秘 书 长：胡茂生  
副秘书长：何兴能 林茂荃 易 勤 黄雄才

## 序

当代新技术革命的蓬勃发展,带来社会生产力新的飞跃,引起整个社会的巨大变革.电子计算机技术是新技术革命中最活跃的核心技术,在工农业生产、流通领域、国防建设和科学研究方面得到越来越广泛的应用.

党的十一届三中全会以来,我国计算机应用事业的发展是相当迅速的.到目前为止,全国装机量已突破三十万台,十六位以下微型计算机开始形成产业和市场规模,全国从事计算机科研、开发、生产、应用、经营、服务和教学的科技人员已达十多万人,与1980年相比,增长了近八倍.他们在工业、农业、商业、城建、金融、科技、文教、卫生、公安等广阔的领域中积极开发应用计算机技术,取得了优异的成绩,创造了显著的经济效益和社会效益,为开拓计算机应用的新局面作出了重要贡献.实践证明,人才是计算机开发应用的中心环节.我们必须把计算机应用人才的开发与培养放在计算机应用事业的首位,要坚持不懈地抓住人才培养这个关键.

从目前来看,我国计算机应用人才队伍虽然有了很大的发展,但是这支队伍的数量和质量还远不适应计算机应用事业发展的客观需要,复合型人才的培养与教育还没有走上规范化、制度化轨道,教材建设仍显薄弱,培训质量不高.因此,在国务院电子信息系统推广应用办公室领导、支持下,全国35个省、市、自治区、计划单列市计算机应用主管部门共同组织118所大学和科研单位的400多位专家、教授编写了全国第一部《电子计算机应用人才培训大纲》以及与之配套使用的电子计算机应用系列教材,在人才培训和开发方面做了一件很有意义的工作,对实现培训工作规范化、制度化将起到很好的推动作用.

《电子计算机应用人才培训大纲》和“电子计算机应用系列教材”贯穿了从应用出发、为应用服务,大力培养高质量、多层次、复合型应用人才这样一条主线.培训大纲总结了近几年各地计算机技术培训正反两方面的经验,提出了计算机应用人才的层次结构、不同层次人才的素质要求和培训途径,制定了一套必须遵循的层次化培训办学规范,编制了适应办学规范的“课程教学大纲”.这部培训大纲为各地方、各部门、各单位制定人才培训规划和工作计划提供了原则依据,为科技人员、管理人员以及其他人员学习计算机技术指出了努力方向和步聚,为社会提供了考核计算机应用人才的客观尺度.“电子计算机应用系列教材”是培训大纲在教学内容上的展开与体现,是我国目前规模最大的一套计算机应用教材.教材的体系为树型结构,模块化与系统性、连贯性、完整性相兼容,教学内容注重实用性、工程性、科学性,并具有简明清晰、通俗易懂、方便教学、易于自学等特点,是一套很好的系列教材.

这部培训大纲和系列教材的诞生是各方面团结合作、群策群力的结果,它的公开出版和发行,对计算机应用人才的培训工作将起到积极的推动作用.希望全国各地、各部门、各单位广泛运用这套系列教材,发挥它应有的作用,并在实践中检验、修改、补

充和完善它。

通过培训教材的建设，把培训工作与贯彻国家既定的成人教育、函授教育、电视教育和科技人员继续工程教育等制度相结合，逐步把计算机应用人才的培训工作引向规范化、制度化轨道，为培训和造就大批高素质、多层次、复合型计算机应用人才而努力奋斗，更好地推动计算机应用事业向深度和广度发展。

**李祥林**

1988年10月17日

# 前 言

60年代开始发展的数据库技术在70年代已趋成熟，并已在发达国家中普遍使用。在我国，70年代末才接触这一软件技术。80年代初，计算机事业的发展使得微型计算机逐步普及，国民经济的各个领域都开始开发自己的信息管理系统。与此同时，微型计算机数据库管理系统也相应地进入了我国计算机软件市场，作为一种开发信息管理系统的工具被使用。近几年来，国内的软件工作者相继开发了一些微型机数据库管理系统。人们明显地感觉到，用数据库管理系统开发一个数据库应用系统，比用其它高级程序设计语言效率要高，开发周期要短，且简单易学。这也是微型计算机数据库管理系统得以广泛应用的原因之一。随着计算机应用的发展，微型计算机数据库技术的应用也将愈来愈广泛。

然而，目前我国计算机应用人才的队伍形成的时间不长，其中相当一部分人虽然在从事计算机应用的工作，但却没有系统地学习过计算机专业的课程，同时随着计算机应用事业的发展，需要培养和造就一大批高质量的、深通计算机技术的人才。因此，尽快满足社会的这一要求，就显得尤为重要了。本书正是在这种形势下，根据我们多年的工作和实践而编写的。

在本书的编写中，我们既注重了数据库技术原理的讲述，又注重了理论与应用的结合。在应用方面，我们着重于数据库的设计。为了使读者能根据数据库设计理论和自身的工作环境设计出性能良好的数据库应用系统，书中提供了两个数据库应用系统设计与实现的实例，详细地描述了数据库的设计过程并附有完整的应用程序。考虑到读者的实习环境，本书向读者介绍了两个关系模型的数据库管理系统，即在 M68000 微型机上运行的 INGRES 和被誉称为“大众数据库”的 dBASE III。IBM-PC 系列微型计算机的普遍使用将使得 dBASE 系列软件具有广阔的应用前景。为了适应信息管理的发展，本书的最后一章向读者介绍了可在局部网上运行的 dBASE III PLUS。

本书适用于计算机专业本科以外的各专业学生，也适用于各种类型的计算机培训班，同时还可作为计算机水平考试的参考书。

使用本书时应注意，第三、第四两章可根据学员的程度选讲或者不讲。第七、第九两章和第八、第十两章作为两组，可根据实习条件取其中一组。

由于水平所限，书中难免存在欠妥之处，恳请读者批评指正。

# 目 录

第一章 数据库系统概貌 .....	1
1.1 信息、数据及数据处理 .....	1
1.2 数据库的产生 .....	5
1.3 数据库系统的体系结构 .....	9
1.4 数据库系统的组成 .....	10
1.5 数据库管理系统 .....	11
1.6 用户存取数据的过程 .....	14
1.7 数据模型 .....	14
习题 .....	18
第二章 关系方法 .....	19
2.1 关系模型的数学理论基础 .....	19
2.2 关系代数 .....	24
2.3 关系演算 .....	29
2.4 关系数据库的数据语言 .....	32
习题 .....	38
第三章 层次方法 .....	40
3.1 IMS 系统简介 .....	40
3.2 数据模型 .....	40
3.3 数据子模型 .....	45
3.4 数据操作语言 .....	48
习题 .....	51
第四章 网状方法 .....	52
4.1 DBTG 系统结构 .....	52
4.2 DBTG 数据结构 .....	53
4.3 DBTG 的模式描述 .....	59
4.4 DBTG 的子模式 .....	62
习题 .....	63
第五章 数据库设计 .....	64
5.1 关系数据库设计理论 .....	64
5.2 数据库设计的主要目标 .....	81
5.3 数据库设计的基本步骤 .....	83
习题 .....	89
第六章 数据库的保护 .....	91
6.1 问题的提出 .....	91
6.2 数据库保护的基本内容 .....	91
6.3 安全性保护 .....	91



6.4	完整性保护 .....	93
6.5	并发控制 .....	94
6.6	数据库的后援与恢复 .....	95
	习题 .....	96
<b>第七章</b>	<b>INGRES 关系数据库系统 .....</b>	<b>97</b>
7.1	概论 .....	97
7.2	体系结构 .....	98
7.3	数据结构 .....	104
7.4	数据操作 .....	107
7.5	数据库的管理与维护 .....	119
7.6	嵌套的 QUEL 语言——EQUEL/C 简介 .....	121
	习题 .....	127
<b>第八章</b>	<b>人事、工资综合管理系统的设计与实现 .....</b>	<b>128</b>
8.1	人事、工资综合管理数据库的设计 .....	129
8.2	数据库的建立 .....	135
8.3	人事、工资综合管理系统应用程序设计 .....	140
8.4	人事、工资综合管理系统的应用程序 .....	157
	习题 .....	175
<b>第九章</b>	<b>dBASE III .....</b>	<b>176</b>
9.1	dBASE 概述 .....	176
9.2	体系结构 .....	176
9.3	数据结构 .....	177
9.4	数据操作语言 .....	181
9.5	程序设计语言及应用程序 .....	204
9.6	dBASE III 数据库系统的保护 .....	216
9.7	dBASE III 与程序设计语言的通讯 .....	218
	习题 .....	226
<b>第十章</b>	<b>图书管理系统的设计与实现 .....</b>	<b>227</b>
10.1	图书管理数据库的设计 .....	227
10.2	图书管理数据库的建立 .....	234
10.3	命令文件的设计 .....	236
10.4	图书管理系统的命令文件 .....	237
<b>第十一章</b>	<b>dBASE III PLUS 简介 .....</b>	<b>253</b>
11.1	单用户 dBASE III PLUS 系统 .....	253
11.2	dBASE III PLUS 的网络功能 .....	255
<b>附录 A</b>	<b>INGRES 数据语言 QUEL .....</b>	<b>261</b>
<b>附录 B</b>	<b>dBASE III 命令一览表 .....</b>	<b>272</b>
<b>参考文献</b>	<b>.....</b>	<b>286</b>

# 第一章 数据库系统概貌

## 1.1 信息、数据及数据处理

自人类社会开始的那天起,就有两类性质的工作,一类是物质生产,另一类是对社会的管理.随着社会的发展,对社会的管理日趋复杂.在现代社会中,对社会的管理变得更为广泛、复杂.例如,国家对各部、委的管理,各部对部属企业、事业单位的管理,各企业、事业单位对本单位的管理.一个单位自身的管理又包括生产管理(业务管理)、供销管理、设备管理、财务管理、人事管理等等.所有这些管理,归结到一点,就是对各种信息的管理.

在现代社会里,信息在各种社会活动中居越来越重要的地位.由于信息量的急剧增加,信息的形式多样、结构复杂以及人们对信息要求的及时性,使得信息处理成为一个极其重要的问题.理论上,信息论已发展为信息科学,被视为现代科学技术的三大支柱之一;实践上,电子计算机所引起的重大的技术革命为信息化社会的形成提供了强有力的现代化信息处理的工具.从现在起,本书所提到的信息管理,都是以电子计算机为工具所进行的信息管理.

信息处理有时又称为数据处理.信息与数据在概念上有无区别呢?严格地讲,是有区别的,实际上,它们是不同的领域中的概念.

### 1.1.1 现实世界

人们通过对信息的管理来实现对人类社会的管理.本节所说的现实世界也只能是人类社会的同义词.它既包括人类的精神世界,也包括人类的物质世界.现实世界就是被管理的对象.

#### 1. 实体集与实体

在现实世界中,有无穷无尽的客观存在着的事物:医院中有病人、医生、护士、医疗器械、药品等;交通工具具有火车、汽车、飞机、轮船;旅社中有房间、床位、旅客;银行中有储户帐号等;仓库中有物资、设备.

注意,管理的对象不是某个事物,而是某类事物.例如,一个旅社里所管理的不是某一个房间,而是所有的房间;某次班机机票的预订是指该次班机的全部机票,而不是某一张;一个仓库管理员所管理的不仅是某一种物资,而是该仓库的所有物资.一个单位的人事部门所管理的是该单位的全体职工,而不是某一个职工.

我们把现实世界中某类事物的全体称做实体集,把该类事物中能相互区分的每个个体称为实体.

这里所说的某类事物的“全体”,是有一定范围的.例如,房间是一实体集,但并不是指

地球上的所有房间.它可以是某一旅社的所有房间,因为这个旅社的管理员只关心他的旅社中的房间;也可以是某机关单位中的房间,该单位的房管部门只对此感兴趣.一句话,实体集是对有限的管理范围来说的.

此外,实体集总是个有穷集合,而不可能是无穷集合.

## 2. 属性

每一实体集都有一定的特性.例如,某一单位的职工,都有姓名、性别、年龄、文化程度、政治面貌等特性;某一仓库的物资都有品名、编号、制造厂家、进库日期、数量等特性;某次班机的机票都有日期、班次、座位、起始地点、终止地点、票价等特性;…….

我们把能够表示实体集的固有特征的那些特性称为该实体集的属性.

一个实体集究竟有多少属性,是随管理业务的需要而异的.例如,对瓷器这一实体集进行管理.在某瓷厂对本厂生产的所有瓷器进行管理时,就不必有制造厂家这一属性,而在一个瓷器商店中,因为有一些不同厂家生产的瓷器,因此在管理中,就需要有这个属性了.

## 3. 实体标识

在一个实体集中,任何两个实体都是可以相互区别的,也就是说,它们所具有的某个或某些属性是不同的.我们能够将一个实体与其它实体区别开来的属性或属性集合称为该实体集的标识属性.

例如,一次班机的机票,座位号码这一属性能将一张机票同其它机票唯一地区别开来.因此,座位号码这一属性就是机票这一实体集的标识属性.

再例如,某一仓库所库存的零件,同一品名,但制造厂家不同,即同一品名的零件,由几个厂家同时生产,同一厂家又制造几种不同的零件.因此要标识一个零件,只用品名这一属性不行,只用制造厂家这一属性也不行,但用品名和制造厂家这两个属性的组合便可以.那么这两个属性的组合就是零件这一实体集的标识属性.

### 1.1.2 信息世界

所谓信息管理,实际上是社会上的各级管理机关对它们所辖的众实体集所具有的各属性和实体集之间的联系的管理.

#### 1. 记录型

一个中学要管理下列三个实体集:

全体教师

全体学生

所有班级

对于教师这一实体集,有下列属性:

姓名

性别

年龄

文化程度

专业

所教课程

家庭住址

对于学生这一实体集,有下列属性:

学号

姓名

性别

年龄

所在班级

对于班级这一实体集,有下列属性:

班次

班主任名

班长名

人数

我们把用来描述某个实体集的属性的有序排列称为记录型。

由上可以看出,一个实体集和一个记录型是一一对应的关系。

实体集:教师

记录型:

姓名	性别	年龄	文化程度	专业	所任课程	家庭住址
----	----	----	------	----	------	------

实体集:学生

记录型:

姓名	性别	年龄	所在班级
----	----	----	------

实体集:班级

记录型:

班次	班主任名	班长名	人数
----	------	-----	----

还可以举出很多例子。

## 2. 属性值和记录值

对于一个实体集,根据管理业务的需要构造一个记录型。那么,对于该实体集中的每个实体,对应于这个记录型中的每一个属性都有特定的内容。例如,班级这个实体集,它对应的记录型包括四个属性。对于这个实体集中的每个实体,这四个属性都有特定的内容。当“班次”这个属性的内容为“初一(1)班”时,“班主任名”的内容为“张波”,“班长名”的内容为“李红”,“人数”的内容为54。我们把“初一(1)班”、“张波”、“李红”、“54”分别叫做这四个属性的属性值。一般地说,对于一个实体,每个属性的内容都称为该属性的一个属性值。一个实体集有多少个实体,那么每个属性就应有多少个属性值。

属性→	班次	班主任名	班长名	人数
属性值 {	初一(1)	张波	李红	54
	初一(2)	赵祥	王军	50
	⋮	⋮	⋮	⋮
	初三(5)	蒋平	刘芳	53

以上,我们是纵向看问题的. 下面让我们横向看一下:

记录型 {	班次	班主任名	班长名	人数
记录值 {	初一(1)	张波	李红	54
	初一(2)	赵祥	王军	50
	⋮	⋮	⋮	⋮
	初三(5)	蒋平	刘芳	53

我们把用来描述每一个实体的各属性值的有序排列称为一个记录,或者称为相应于该记录型的一个记录值.

一个实体集中有多少个实体,那么这个记录型就有多少个记录值. 如果上述学校共有 25 个班,那么该记录型就有 25 个记录值.

### 3. 关键字

我们将一实体集中的标识属性称为关键字.

### 4. 文件

一个记录型连同它的所有记录值的集合,称为一个文件.

## 1.1.3 计算机世界

当完成了从现实世界到信息世界的抽象之后,是写在纸上,存放于文件柜内,还是在计算机的外存贮器上,这似乎没有本质的差别,看起来只不过是存贮介质的改变而已. 然而,就是这一改变,使得信息管理从低级的、低效率的手工操作一跃而进入高速的、现代化的管理阶段.

在计算机世界中,有它的专用名词和术语,下面我们扼要地介绍一下.

### 1. 二进制位

计算机存贮器(包括内存贮器和外存贮器)存贮数据的最小单位是二进制位. 它所表示的数据要么是二进制的 0,要么是二进制的 1,别无其它.

### 2. 字节

字节是内存贮器编址的最小单位. 一个字节含有 8 个二进制位. 对于计算机的 I/O 编码,一个字符均占一个字节. 例如 ASCII 编码中的字符 A 在计算机内部表示为 01000001,相当于八进制的 101,十六进制的 41,十进制的 65.

### 3. 卷

卷是外存贮器的单位。一盘磁带，一个活动的磁盘组，一个温盘或者一片软磁盘都称为一卷。

### 4. 磁道

一个磁盘面被分为若干个同心圆，每个同心圆都是用来记载数据的。这每一个同心圆被称为一个磁道。

在 IBM-PC/XT 及其兼容机上用的软盘均为  $5\frac{1}{4}$  英寸<sup>①</sup> 的。有双面双密度的软盘，即磁盘的两面都可以存贮信息。这种盘每面有 40 个磁道，每个磁道上又被分为 8 等份，每份称为一个扇区。一个扇区有 512 个字节。因此双面双密度的  $5\frac{1}{4}$  英寸的软盘的存贮容量为 327680 字节，约为 0.3 兆字节，如图 1.1 所示。另外，还有高密度的  $5\frac{1}{4}$  英寸的软盘，它每面有 80 个磁道，每个磁道也被分为 8 个扇区。因此存贮容量成倍地增长。

### 5. 逻辑记录

在一个文件中，一个记录值称为一个逻辑记录。

### 6. 物理记录

物理记录是计算机的内存贮器和外存贮器交换数据的单位，又称为块。通常，一个物理记录包含一个或多个逻辑记录。

现在我们就可以说清楚信息与数据之间的区别了。

当人们把信息存入计算机的存贮器后，便表示信息的所有字符都转化成了二进制代码，这些二进制代码被称为数据。

一则信息可以用不同的数据来表示。这种不同在计算机化了的信息系统中，往往与计算机系统有关。

因此我们说数据是计算机化了的信息。对信息的处理是通过对数据的处理来实现的。用计算机进行信息处理的过程如下：

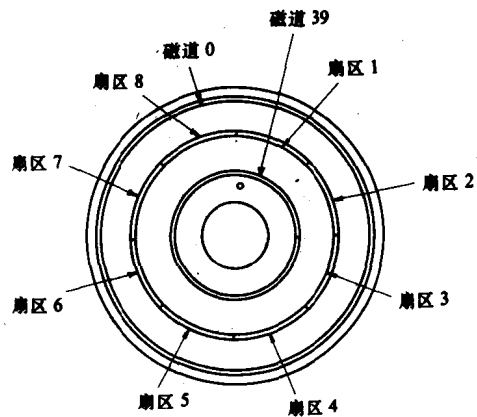
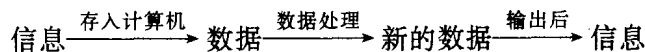


图 1.1

## 1.2 数据库的产生

数据库技术的产生并不是偶然的，是数据管理发展的必然结果，并将随着数据管理技

<sup>①</sup> 1 英寸=0.0254 米。

术的进一步发展日趋完善。

### 1.2.1 数据管理的发展

自 1946 年世界上第一台计算机问世,到数据库系统出现之前,随着计算机软件 and 硬件的发展,数据管理经历了两个阶段。

第一阶段:(1946~50 年代中期)

这一阶段中,发达国家中面临着大量的科学计算.这些科学上的计算用当时的计算工具是无法实现的.科学要发展,计算机便应运而生.因而,在它问世后,首先被应用于科学计算.也正因为刚刚问世,所以还很不完善.硬件方面,外存贮器只有磁带、纸带、卡片,软件方面,没有操作系统和高级语言,没有对数据进行管理的软件.用户只能用机器语言或汇编语言编程,且用户必须和计算机硬件打交道.程序员不仅要规定数据的逻辑结构,还要规定其存贮结构、存取方法、输入输出的方式及其数据转换.因此,数据和其相应的程序是不可分的,即彼此是互相依赖的,亦即数据对于程序来说不具有独立性.

第二阶段:(50 年代后期~60 年代中期)

随着科学的发展,计算机也在不断地发展.到了第二阶段,计算机不仅用于科学计算,同时也用于数据处理.大量的数据需要长期地保存在外存贮器中,并需要反复地进行查询、修改、插入、删除等处理.这就极大地刺激了硬件和软件的发展.硬件方面,出现了大存贮量的磁盘、磁鼓等直接存贮设备.软件方面,出现了高级语言、编译系统、操作系统以及在操作系统支持下的数据管理软件——文件管理系统.处理方式也有了改进,既可以批量处理,又可以联机实时处理.

由于有了文件管理系统,数据以文件的形式存贮在外存贮器上,程序员只需规定文件的逻辑结构.至于文件的存贮结构、存取方法,输入输出方式及其数据的转换均由文件管理系统自行处理.文件的逻辑结构成为程序与数据的一个界面.文件存贮结构的变化不会引起逻辑结构的变化,因而相应的程序亦可不变.这就使得程序对于数据有一定的独立性.

由于有了直接存取的存贮设备,文件的组织已多样化,如索引文件、链接文件、倒排文件、直接存取文件等.它们都不同程度地提高了更新数据的速度.

总之,文件管理系统的出现使得数据管理比第一阶段有了很大的改进.

但是,这一阶段的数据管理仍存在着较大的缺点.首先,文件的逻辑结构是面向应用的,即一个文件基本上对应于一个或几个应用.如果两个应用程序对应的两个文件有一部分相同,那么这个相同的部分在两个文件中都有存贮,这样就浪费外存空间.如果我们把同一数据的多次存贮称为数据的冗余,那么这样的数据管理方法会使得数据冗余度很大.其次,由于数据的冗余度大,就很容易造成数据的不一致性(或者叫做数据的不相容性).例如,同一个职工,在一个文件中的年龄是 30 岁,而在另外一个文件中却是 25 岁,这样的数据只能提供不正确的信息,且应用不易扩充.这些缺点只能在数据库系统中才能得以克服.

### 1.2.2 数据库的定义

数据库技术是数据管理的最新技术,数据库管理系统是当代计算机系统的—一个重要

组成部分. 然而要用几句话严格地、简明地、准确地给数据库下个定义, 却不是一件容易的事. 事实上, 现在已有的关于数据库的定义也是不尽相同的. 这是因为人们从不同的角度用不同的观点来看待数据库的结果. 这里, 我们也只能对数据库给出一个轮廓性的描述.

数据库是一个集成化的有穷的数据集合. 它按照数据间的自然联系以最佳的方式组织、存贮数据, 使之具有最小冗余度. 它能提供给多个用户共享, 有的系统允许几个用户并发地使用数据库. 它具有较高的数据独立性, 确保数据的安全性与完整性.

根据上述, 数据库是**集成的**. 它把一个企业或者一个部门的若干种性质不同的数据及其相互之间的关系根据某些约定, 经规范化后, 再进行存贮.

例如, 我们为某个部门设计了各个子部门都要用到的职工档案, 其中包括人事档案、工资档案、业务档案、住房档案等以及它们之间的联系.

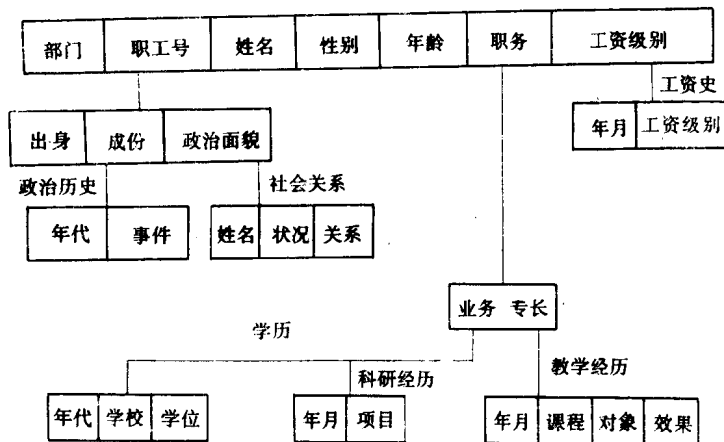
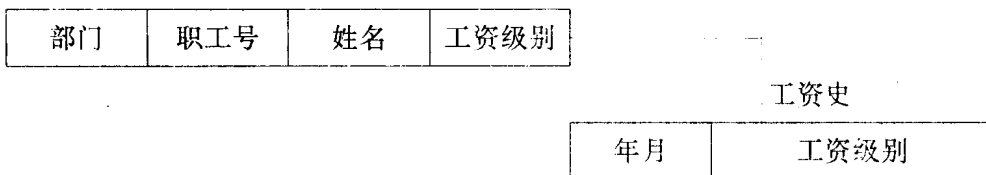


图 1.2

从图 1.2 可知, 数据库的数据是从全局出发来组织的 (不再是面向应用, 而是面向数据的). 这样的结构称为数据库的全局逻辑结构, 它反映了数据库中全部数据间的逻辑关系.

数据库中的数据是可**共享的**. 所谓共享, 指的是数据库的数据可以提供给多个用户使用. 每个用户感兴趣的不是数据库中的全部数据, 而是它的某个子集. 数据库系统允许这些子集以许多不同的方式重叠.

例如, 对于图 1.2 中的全部数据, 劳资科所关心的是:



主管审干的科室感兴趣的是:



部门	职工号	姓名	性别	年龄
----	-----	----	----	----

人事

出身成分	政治面貌
------	------

政治历史

年代	事件
----	----

社会关系

姓名	状况	与职工关系
----	----	-------

技术科需要掌握的是：

部门	职工号	姓名	性别	年龄	职务	业务专长
----	-----	----	----	----	----	------

学历

年代	学校	学位
----	----	----

科研经历

年 月	项 目
-----	-----

教育科所关心的是：

部门	职工号	姓名	职务	业务特长
----	-----	----	----	------

教学经历

年月	课程	授课对象	效果
----	----	------	----

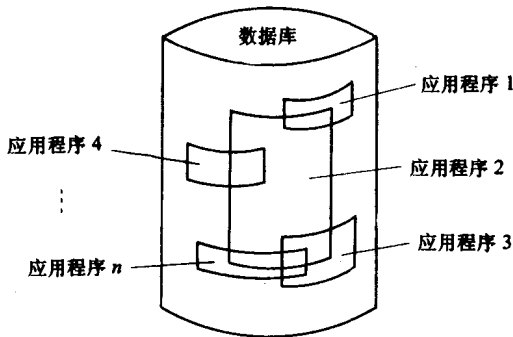


图 1.3

为多用户数据库系统。

读者还可以举出更多的例子。上面所列举的这些表示单个用户的数据逻辑关系，叫做数据的局部逻辑结构。

从上述这些数据的局部逻辑结构可以看出，“共享”的意思是说，数据库的独立的一片数据可以被它的几个不同的用户共同使用，即每个用户都可以存取同一片数据而应用于不同的目的。图 1.3 可以形象地表示数据库这一特征。

关于共享，还有另外一层意思，那就是并发共享，即几个用户可以在同一时间内存取数据库，甚至有可能存取同一数据。一个数据库系统支持这种共享形式，就称之为多用户数据库系统。