

# 信息组织与信息处理

高等学校教学用书

张耀清 编

XINXIZUZHIYU  
XINXICHULI

中国矿业大学出版社

## 序

我国首批微机的引进是在1980年,在这以前,国产的计算机一般都不配置操作系统。计算机的操作、使用和维护比较复杂,应用计算机来解决各种问题主要依靠经过专门训练的计算机专业人员。

80年代,微机在我国得到了推广和某种程度的普及。非计算机专业的技术人员稍加培训就能够初步地操作和使用计算机。工作在各种岗位上的技术人员,他们熟悉自己业务范围内的各种细节,一旦掌握了计算机技术就如虎添翼,而成为计算机应用领域中的一支强大队伍。他们初步地使用某种计算机语言或dBASE来编写较为简单的程序并不困难,但是要进一步提高,就遇到诸如数据结构、文件组织、程序设计的概念和方法等问题。从事信息处理的人员还会遇到DBMS(数据库管理系统)、MIS(管理信息系统)、DSS(决策支持系统)等问题。在计算机专业,这是许多门课程。但是这些非计算机专业的技术人员往往没有时间,也不太必要去逐门学习这些课程,他们需要有一门课程对上述知识进行综合学习,以便可以花费较少的时间和精力去获得这些知识的精华。本书正是这种课程的一本教材。

孙辑正 于中国矿业大学 1992年1月

# 前 言

在高科技时代,信息处理技术推动着社会前进。以微型计算机为代表的信息技术走进了社会的各行各业,为人们的生产、生活、娱乐和办公等活动大开方便之门,成为人们得心应手的工具。信息处理技术的发展和运用,为人们迎接当代的信息大爆炸提供了重要手段,信息处理技术牵动着当代创业者的心弦。

在信息时代,信息处理技术和处理的结果主宰着人们的各种重大行为,一切立志献身四化大业的专家、教授、厂长、经理、工程师和政府官员们都在更新自己的知识结构,努力学习和应用信息处理技术,以迎接高科技时代的挑战。以经济建设为中心的改革开放,使科学技术、经济建设日新月异,激烈的竞争使每个厂、矿、企业和政府部门的业务活动迅速走向了基于计算机的信息处理轨道。管理信息系统、决策支持系统、办公自动化系统和专家系统等基于计算机的信息处理技术为企业的管理、决策、计划、预测做出重大的贡献,已经成为发展生产、提高技术、增加效益的重要组成部分。大量的事实告诉人们,必须学会信息组织与信息处理的理论、技术和技能,才能驾驭计算机这个高科技的向导,把人类社会引向更加美好的未来。

很多非计算机专业的科技人材,他们学了信息处理技术后,创造出重大业绩。在学校,很多非计算机专业的学生,积极辅修“数据结构”、“数据库”、“信息系统”等课程。种种现象表明,信息处理的理论、技术、技能是科技人员必备的能力。

为了适应当前的社会需求,作者在“信息组织与信息处理”讲义的多次应用基础上,结合科研活动,经过认真修改,并加强了信息处理技术的基本理论、基本技术和基本技能,增加程序示例,充实了系统分析与设计等实践性很强的应用环节,构成本书的基本内容。

全书共有七章,第一章介绍信息处理的基础知识和计算机的系统结构、概念和应用的基本常识;第二章和第六章讲述信息组织与信息处理的基本理论;第三章、第四章和第五章讲解信息组织与信息处理的基本技术和基本技能,包括了大量的程序示例和结构图形;第七章以煤矿基建概预算信息系统开发为例,说明信息组织与信息处理的具体方法和途径。为了便于复习,每章后边都有复习思考题,在书后还附有 dBASE 函数和主要操作命令。

本书可作为大专院校非计算机专业的教材,也可供各类工程技术人员自学信息组织与信息处理的理论与技术。

在本书编写过程中得到中国矿业大学孙辑正教授的指导,并为本书写了“序”,特此致谢!

由于本人水平有限,再加上时间仓促,书中难免有疏漏和错误,敬请广大读者批评指正。

作者于中国矿业大学 1992年1月

# 目 录

<b>第一章 信息处理的基本知识</b> .....	(1)
§ 1-1 信息与数据 .....	(1)
一、信息与数据的概念 .....	(2)
二、数据的层次化结构 .....	(4)
§ 1-2 信息的代码表示 .....	(6)
一、机器内部码 .....	(6)
二、数据代码 .....	(8)
§ 1-3 计算机软件知识 .....	(10)
一、系统软件.....	(11)
二、应用软件.....	(12)
§ 1-4 计算机语言与程序设计 .....	(13)
一、计算机语言体系.....	(13)
二、程序设计的基本方法.....	(17)
§ 1-5 微型计算机操作与应用 .....	(21)
一、微型计算机的优点.....	(21)
二、微型计算机的构成.....	(21)
三、微型计算机主要操作命令.....	(23)
复习思考题 .....	(26)
<b>第二章 数据的基本结构</b> .....	(28)
§ 2-1 线性表 .....	(28)
一、线性表及其存储结构 .....	(28)
二、栈和队列 .....	(31)
三、数组 .....	(36)
§ 2-2 链表 .....	(37)
一、线性链表.....	(37)
二、链栈与链队.....	(39)
三、其它形式的链表.....	(41)
§ 2-3 字符串 .....	(42)
一、字符串的结构和运算.....	(42)
二、文本编辑.....	(43)
§ 2-4 树及其应用 .....	(43)
一、树的构成.....	(44)
二、二叉树及其运算.....	(45)
三、树的应用示例.....	(50)

§ 2-5 图及其应用 .....	(53)
一、图的建立 .....	(53)
二、建立计算机通讯网 .....	(54)
三、拓扑排序 .....	(57)
四、关键路径 .....	(58)
复习思考题 .....	(60)
<b>第三章 数据的排序与查找 .....</b>	<b>(62)</b>
§ 3-1 内排序方法 .....	(62)
一、选择排序 .....	(63)
二、插入排序 .....	(65)
三、冒泡排序 .....	(66)
四、两路归并排序 .....	(68)
五、希尔排序 .....	(69)
六、快速排序 .....	(71)
七、堆排序 .....	(73)
八、基数排序 .....	(75)
九、各种排序方法比较 .....	(77)
§ 3-2 线性表查找 .....	(78)
一、顺序查找 .....	(78)
二、折半查找 .....	(79)
三、插值查找 .....	(81)
四、分块查找 .....	(82)
五、顺序查找算法分析 .....	(83)
§ 3-3 树表查找 .....	(84)
一、二叉排序树查找 .....	(84)
二、二叉排序树的结构分析 .....	(85)
§ 3-4 哈希表查找 .....	(86)
一、哈希表的结构 .....	(86)
二、哈希表查找 .....	(87)
三、哈希查找示例 .....	(87)
复习思考题 .....	(90)
<b>第四章 文件组织 .....</b>	<b>(92)</b>
§ 4-1 文件的基本概念 .....	(92)
一、文件的名词、术语 .....	(92)
二、文件的性质和用途 .....	(93)
§ 4-2 外部存储器 .....	(95)
一、磁带机 .....	(96)
二、磁盘存储设备 .....	(96)
三、光盘存储器 .....	(98)
§ 4-3 顺序文件 .....	(99)

一、顺序文件的结构与特点	103
二、顺序文件的操作	105
§ 4-4 索引文件	103
一、索引文件的组织形式	103
二、索引顺序文件的结构	104
三、索引顺序文件的操作	106
§ 4-5 特种形式的文件	107
一、虚拟存取文件	107
二、直接存取文件	109
三、多关键字文件	112
四、文件的结构与性能综述	113
§ 4-6 随机文件的程序设计示例	115
一、文件处理程序的建立	115
二、程序示例分析	117
复习思考题	120
<b>第五章 数据库管理系统及其应用</b>	122
§ 5-1 数据库的基础知识	123
一、数据库的特征	123
二、数据的三个领域	125
三、实体与模型	127
§ 5-2 数据库管理系统	129
一、数据库管理系统的结构与功能	129
二、DBMS 的种类与应用	131
§ 5-3 dBASE-Ⅲ 系统概述	133
一、dBASE-Ⅲ 的主要功能	133
二、dBASE-Ⅲ 的文件类型	135
三、变量与表达式	135
§ 5-4 dBASE 程序设计框架	138
一、dBASE 程序的构成和特点	138
二、程序模块的建立	140
三、程序控制功能	142
§ 5-5 各类程序模块示例	144
复习思考题	153
<b>第六章 管理信息系统</b>	155
§ 6-1 管理信息系统的结构与功能	156
一、管理信息系统的物理结构	156
二、管理信息系统的功能	157
§ 6-2 管理信息系统模型示例	161
一、社会性的管理信息系统	161

二、企业的管理信息系统 .....	(163)
6-3 管理信息系统开发 .....	(166)
结构系统分析与设计方法 .....	(166)
二、系统分析 .....	(170)
三、系统设计 .....	(176)
§ 6-4 结构化模块设计 .....	(180)
一、模块的结构分析 .....	(180)
二、结构化模块的设计 .....	(182)
§ 6-5 决策支持系统 .....	(186)
一、DSS 的结构与功能 .....	(186)
二、MIS 与 DSS 的比较 .....	(188)
复习思考题 .....	(188)
<b>第七章 煤矿基建概预算信息系统开发</b> .....	(190)
§ 7-1 煤矿基建概预算信息调研 .....	(191)
一、基建概预算的特点和工作内容 .....	(191)
二、基建工程费用的构成 .....	(192)
三、概预算中取费的依据 .....	(193)
§ 7-2 施工费用模型设计 .....	(195)
一、直接定额费的计算 .....	(195)
二、施工总费用生成 .....	(197)
三、单位工程预算示例 .....	(199)
§ 7-3 施工概预算信息系统结构设计 .....	(203)
一、系统逻辑设计 .....	(203)
二、系统物理设计 .....	(204)
三、模块设计 .....	(205)
§ 7-4 系统实施示例 .....	(207)
复习思考题 .....	(213)
<b>附录</b> .....	(214)
附录 I dBASE-Ⅲ 操作命令表 .....	(214)
附录 II ASCII 码表 .....	(219)
附录 III dBASE-Ⅲ 函数表 .....	(221)
附录 IV 文件运算中常用的 BASIC 函数 .....	(222)
<b>主要参考文献</b> .....	(223)

## 第一章 信息处理的基本知识

在当今的信息社会里,人们用信息处理技术解决自己的业务信息处理问题,已司空见惯了。各行各业应用信息处理技术,不仅技术水平大为提高,而且经济效益十分显著,进而促使其迅速发展。信息处理技术发展和应用的特点概括如下:

1. 系统性 信息处理技术都是基于计算机功能,并且由计算机硬件,计算机软件、数据对象、通讯网络,经过严格分析设计组成的系统;

2. 经济性 凡是应用这种技术的单位和行业都收到了明显的经济效益和显著的社会效益;

3. 广泛性 世界上任何一个国家、任何一个行业、任何一个系统、任一单位,任何复杂的理论、工程、技术和难题都可运用信息技术;

4. 通用性 信息处理技术属横向学科,任何有数据的地方都可应用,不受行业限制,操作、维护、使用方法一样,信息格式相同;

5. 辅助性 计算机可以辅助各种业务工作,例如:CAD、CAM、CAP、CAI、CASE等,只要有信息技术辅助,就可使行业技术水平大为提高;

6. 共享性 信息技术中的各种资源,如数据资源、软件资源、硬件资源等,都可在系统通讯网络的支持下,实现资源共享;

7. 使用方便性 当代的信息处理系统都有友好的人—机接口,可帮助用户学习,提示操作方法,为推广应用创造方便条件;

8. 智能性 人工智能的发展,知识工程、专家系统的应用,不仅解决繁杂重复的数据处理,而且能帮助人们演绎推理、思维,应用智能技术和理论自动解决问题,使计算机具有某些人类的智能。

从上述特点可知,信息处理涉及各种专门的学科,如信息和数据的知识、计算机及软件知识,程序和语言的知识,系统的建立和管理的知识和汉字信息知识等。

### § 1-1 信息与数据

数据的概念很具体,只要写个数码都可以称之为数据。人们了解数据不是为了知道一个数码,而是为了解决或记录自己在生产和生活中的各种问题。如何记录,如何表现,如何应用数据解决现实问题是十分重要的。数据必须赋之于物理概念,才是对人们有意义的信息。例如识别职工档案的信息,有如下一段数码:

“1867、1、33、4、07800、791005”

这是代表一个职工的简历代码,职工号:“1867”,唯一标识一个职工的代号;“1”是代表男性;现年“33”岁;“4”是代表助教;工资是“78”元;1979年10月5日参加工作的。借助于数码记载事实,研究事实,用计算机处理数据,就是处理事实。有物理含义的数据,人们称为信息。信



息只对相关的人们有直接利害关系,人们也特别关心对自己有用的信息。

通过抽象的数据来记载,表述具体的事实是我们的宗旨。为此,必须研究抽象数据的概念、分类、性质、处理方法等问题。

### 一、信息与数据的概念

数据(DATA)和信息(INFORMATION)是数据处理中的基本概念。数据是原材料,信息是产品。数据是抽象的表述,信息是赋予物理概念且对人们有直接利害关系的数据;或者说数据是用以载荷信息的物理符号;数据是记录下来的可鉴别的符号,信息则是对数据的解释。总之,这两个概念既不相同,又密切相关,而且如影随形地应用于现实社会里人们各项活动中,因而在习惯上往往混为一谈,或不作严格的区分,可叫做“数据”,也可称为“信息”。

#### 1. 数据

数据是描述客观事物的数值、图形、符号、图像、声音等信息所记录下来事实。这些数据的形式都可以输入计算机进行处理,是计算机加工处理的原材料。记录方式可以有很多种,同一事实用数值、文字、图形、符号、代码都可以记录,由于形式不同,处理方式也不同,对数据必须作具体的分析。

##### 1) 数据的特征

- ① 数据的表达形式 文字、符号、数值、图形、照片、色彩、声音等;
- ② 数据的类型 字母、数字、汉字、逻辑类别和各种运算符号;
- ③ 数据的长度 多少位、多少字节,有无小数位及小数的位数;
- ④ 数据的属性 数据的属性可以是生产、销售、产品、人事、贸易等业务记录。这些业务数据的表达方式可以是数值型、字符型、逻辑型等;
- ⑤ 数据的处理要求 精度、格式、用途、输出形式、处理频度;
- ⑥ 吞吐量 每天、每月、输入—输出数据量有多大,批处理或实时处理;
- ⑦ 数据处理过程形态的转变和识别。

##### 2) 数据的流动情况分析

每个具体问题在处理过程中数据如何变动?来龙去脉?如何采集?如何处理?有何种用场?在处理过程中和其他数据的关系?分析数据的合理流动过程,设计合理的数据流程对数据处理有极大的用处。

##### 3) 数据字典

在数据处理过程中,要把数据记录下来,如数据的名称、属性、容量、结构、频度、用途、表达形式等,作为存贮、处理和分析数据的依据。这些内容称为数据字典。

##### 4) 数据文件或数据库

以文件形式组织起来的同一种结构的大量数据,存入计算机,提供计算或处理,常称之为数据文件,经过规范化处理的大量数据文件称为数据库。

总之,关于数据的各种特征和表达形式搞清楚之后,才能更好的组织和处理数据。

#### 2. 信息

信息是具有特定含义的数据,是描述、分析、认识客观事物的依据。信息是经过加工而获得的有用的数据。在社会活动中,人们接触到信息的概念很多,理解也不尽相同。

- ① 信息是新内容、新知识的消息;
- ② 信息是事先不知道结果的消息;

- ③ 信息是使不确定因素减少的有用知识；
- ④ 信息是对客观事物观察而得到的有用知识。

总之,信息来源于实践,反映事物特征和规律,是一种有用的知识。信息是有物理意义的具体数据。

当前,流行说法:“信息就是金钱”,或“信息就是生命”,说明在信息社会,人们对信息的迫切愿望。从这句话可看出信息对人们有切身利益关系。因此,各种“信息中心”、“信息公司”、“信息站”应运而生,力图为人们提供有用的信息,这就是人们对信息的理解和应用的根本原因所在。

#### 1) 信息的性质

信息的价值体现在它的时间性、准确性、适用性和相关性。对一个决策者来说,失去其中任何一条都会导致决策的失误,甚至造成巨大的损失。

#### 2) 信息的分类

通常可分成自然信息和社会信息。

① 自然信息是由于自然环境的变化而发送的信息,各种生物接受自然信息而求生存发展。例如,冬天雁南飞。

② 社会信息是人类社会发生的一切信息的总称,包括:政治信息、经济信息、军事信息、科技信息、市场信息和各种新闻。在改革开放的今天,经济、科技、生产信息更为人们所关注。

#### 3. 数据分类

为了得到有用的信息,计算机需处理各种数据,数据大体上可分为如下几种类型:

##### 1) 按数据处理的任务分类

① 标识数据 把属性相同的数据归并排列,用以说明其它数据,起标识作用。例如,职工的工作证号,标识一个职工,同样有学生证号、准考证号、企业编号、工程编号等;

② 数值数据 表示量的属性数据。如产值、产量、工资收入等;

③ 控制数据 用以驱动并控制设备运行的数据。如运算符、指令、开关命令或一段程序。

##### 2) 按数据的表示方法分类

① 字母数据 由 ASCII 码组成的各种字符串;

② 数字数据 由数码表示的数值数据;

③ 字母数字数据 各种数字和字母组成的字符组合。

##### 3) 按数据的格式分类

① 格式化数据 具有约定的固定格式的数据,格式化表示可使机器处理更为简便易行;

② 非格式化数据 没有固定格式的数据,书写不受任何约束;

③ 文字数据 非数值数据处理中,经常会遇到一段文字说明,一篇文章、一个文件,都是按某种语言格式输入、编辑和存贮的。在一些程序中也可以增加注释语句,说明程序的功能或使用中注意事项。

#### 4. 信息的级

社会应用实践,常把信息分成四个级。

① 战略级 决策人员使用的简明扼要的信息报告,是发展趋势分析和估计的依据,是

极为重要的信息,是高级领导关注的信息;

② 战术级 是经理们所关心的,是作出战略决策所需要的一系列具体目标。如逐年的经营状况和预算等周期性的或随机的信息报告;

③ 操作级 是操作人员日常处理的信息。如提供性能报告和直接操作反馈所要求的报告;

④ 事务级 是大量事务处理的信息,是一般的业务人员从事的工作内容。如帐务往来、各种票据、各类生产数据等信息。

## 二、数据的层次化结构

数据是客观事实的记录。由于事实的多样性、复杂性,相应的数据表示也有多样和复杂的构造。数据本身的形式和结构直接影响它的处理方法和表述。为此,必须研究数据的构成方式,由于计算机只能处理已存贮的数据,应研究数据构成和存贮的关系。

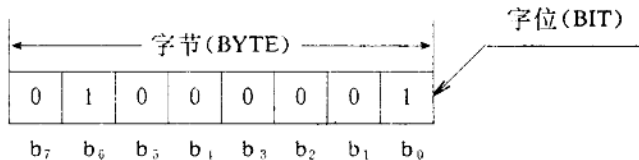
按数据的构成层次,可以有从小到大,从简到繁的七种结构关系:位、字节、数据项、数据对象、记录、文件和数据库,形成了明显的层次化结构。

### 1. 位(BIT)

计算机处理的一个数位,是一个电脉冲,用 0、1 表示,是二进制数的组成单元,也是数据记载中的最小单位。表达形式因载体不同而不同,穿孔卡片上的一个孔,或一个孔位,磁性介质上的一小片磁化体,或未磁化的字位。计算机术语中的“位”(BIT),如八位机、十六位机、三十二位机,其中的位就是说计算机能同时处理由几个脉冲组成的数据。位(BIT)是计算机系统中最小的数据单元。

### 2. 字节(BYTE)

一个 ASCII 码字符要用八位(BIT)来表示,这八位就称为一个字节,由于这个原因,字节就成为计算机处理数据的基本单位。中央处理机表示符号的位组是 8 个 BIT,计算机处理的字长,都是 8 的整数倍。表示一个字节的形成如下:



其中;最高位  $b_7$  是校验位,进行奇偶校验用,以检查数据的正确性。

高三位是表示 ASCII 码的高位数。例如在图中  $b_6b_5b_4$ ,表示成 '100' 是二进制数,是十六进制数的 '4'。

低四位:  $b_3b_2b_1b_0$  是表示 ASCII 码的低位数, '0001' 是二进制数,也是十六进制数的 '1'。

高三位和低四位一起表示字母 'A' 的 ASCII 码的十六进制数 41。在数据处理中,每个字母或符号的 ASCII 码值都占有一个字节,是计算机中统计数据量的基本单位,如一张软磁盘容量是 360KB,或者称为 360 千字节,其中“B”就是字节。“K”是  $K=2^{10}=1024$ 。PC 机内存是 512KB,一个 BASIC 解释程序是 25KB。K=1024 是一个常用单位,是计算机数据计量的进位标识符之一。

ASCII 码字符表,见附录 I。

### 3. 数据项(DATA ITEM)

数据项是构成数据记录的最小单位,它有独立的物理含意。数据项有时也称字段或场,在组织或采集数据时必须从数据项的名称、性质、长度、构成形式等方面予以具体描述。数据项是描述实体属性的基本单元。

### 4. 数据对象(DATA OBJECT)

数据对象是被研究的实体属性的数据描述,是性质相同元素的集合。例如十进制数的数据对象集合是: $N=\{0,1,2,3,\dots,9\}$ ,英文字母的数据对象集合: $C=\{A,B,C,\dots,Z\}$ 。

### 5. 数据记录(DATA RECORD)

数据记录是所研究的数据对象中各个数据项的集合,是文件的基本组成单位。例如研究学生的学习成绩时,数据对象就是学生及功课,数据项就是每门功课,每个学生就是一个记录。如图 1-1 所示。

姓名	学号	语文	数学	外语	总分
丁一	0388125	70	80	70	220
王文	0488072	80	70	80	230
张明	0388062	90	80	90	260
李立	0488035	60	70	60	190
田红	0488093	50	60	60	170
⋮	⋮	⋮	⋮	⋮	⋮

图 1-1 学生成绩登记表

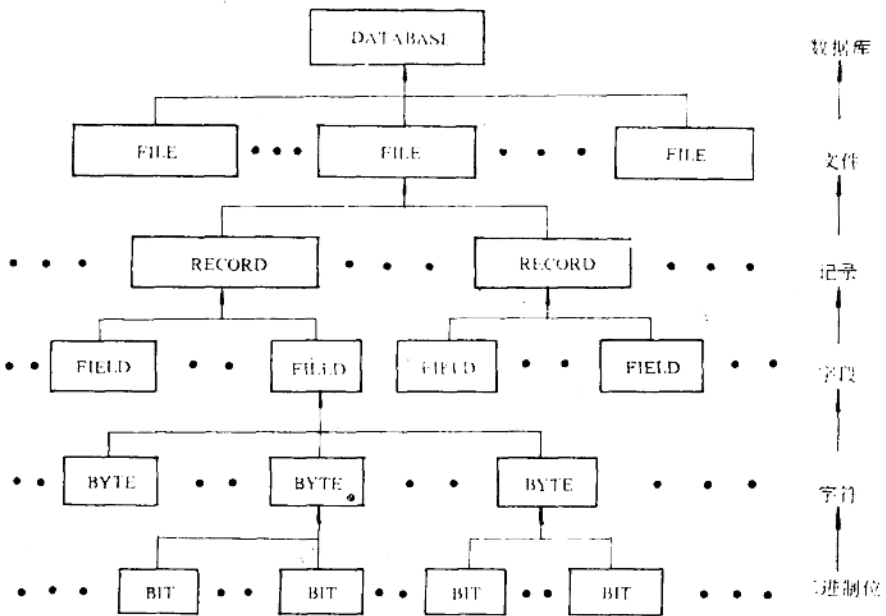


图 1-2 数据的层次化结构图

## 6. 文件(FILE)

文件是性质相同的记录集合。如上述的学生成绩登记表就可视为一个文件。不同性质的数据可构成不同的文件。不同文件中的数据项是不相同的。文件中记录的多少无足轻重,关键是相同的记录可以用同样的程序处理。

## 7. 数据库(DATA BASE)

许多具有特定联系的文件组成数据库,经过规范化,消除数据冗余和依赖关系。一般来说文件都要经过重新组织,消除大量的重复数据,而成为一个合格的数据库。

上述七种基本概念,在计算机数据处理中经常用到,它们之间有层次化结构,如图 1-2 所示。

# § 1-2 信息的代码表示

在信息处理过程中,要对大量的信息(数据或资料)进行计算机加工处理,以提供决策所需的信息。为此,首先要将信息存贮在计算机可以直接读写的介质上,例如文件、数据库所依附的磁盘、磁带。处理完成后又要输出、打印或显示成人们可识别的信息。在输入、处理与输出的过程中,信息的表达形式要进行多次转换,每次转换必须保持数据的一致性和唯一性。为了适应这种要求,必须有一套简便的表述信息的方法,即所谓信息的代码表示。

### 一、机器内部码

把人们能识别的信息,变成计算机能识别的信息,即机内码。

#### 1. 数据的转换过程

1) 原始数据的转换 人们把原始数据经键盘输入计算机,存在磁盘上,或用穿孔机把数据变成纸带上的信息孔位,再用光电机或读孔机读入计算机,存在磁盘上或磁带上,用鼠标器或光笔把图形或符号写入计算机并存盘。凡此种种,都是把人们能识别的信息变成计算机可以识别的信息。

2) 各种数据载体上(磁盘、磁带、纸带、卡片等)的信息,计算机都可以直接或间接的读取,并进行各种处理。这些信息在中央处理机 CPU—CENTRAL PROCESSING UNIT 进行处理时都是以“0”、“1”表示的电脉冲。

3) CPU 处理之后的信息,又以二进制码形式存入磁盘、磁带或穿孔。输出这些计算的结果时,还要把磁盘上的二进制码转换成人们可以识别的数值、字符或汉字信息,并以表格样式输出打印,或者绘制成某种图形。

这些转换的过程都是由计算机、外部设备及有关软件自动完成的。

#### 2. 常用的数制

表示数值的方法称为数制。常用的数制有如下几种。

1) 二进制数 以“0”和“1”两个码表示数,如“010”代表 2,“011”代表 3,是 CPU 可以识别和处理的数据。

2) 八进制数 用“0,1,2,……,7”八个码表示的数,逢 8 进 1,如(13)<sub>8</sub>表示十进制数的(11)<sub>10</sub>。

3) 十进制数 日常生活中用的数制,以 0,1,……,9 来标识。

4) 十六进制数 以“0,1,2,……,9,A,B,……,F”十六个数码表示的数,逢 16 进 1。计

计算机内的存储地址码或计算结果值都用十六进制数表示,如 $(1A)_{16}$ 表示十进制数的 $(26)_{10}$ 。如地址码“FFF”就是相关数值的最高地址码值。

以上的各种数制可以互相转换,如 $(32)_{10}$ 可表示成

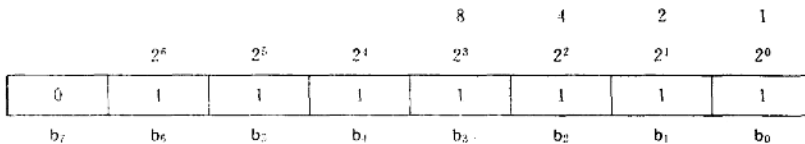
$$(100000)_2 = (10)_8 = (32)_{10} = (20)_{16}$$

它们的对应关系如下表所示,转换数据有专门的算法。

表 1-1 常用数制对应关系表

十进制数	二进制数	8421BCD 码	八进制数	十六进制数
0	0	0	0	0
1	1	1	1	1
2	10	2	2	2
3	11	3	3	3
4	100	4	4	4
5	101	5	5	5
6	110	6	6	6
7	111	7	7	7
8	1000	8	10	8
9	1001	9	11	9
10	1010		12	A
11	1011		13	B
12	1100		14	C
13	1101		15	D
14	1110		16	E
15	1111		17	F

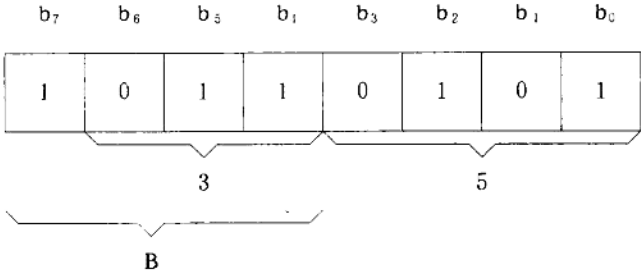
5) 8421BCD 码 它是用二进制码表示十进制数的方法,表达形式如下图。用 $b_3b_2b_1b_0$ 四位二进制码可表示 $0\sim 9$ 中任何一个码,实现十进制数的二进制表示。对于多位数,一个字节的右边四位; $b_3b_2b_1b_0$ 表示一个数码, $b_6b_5b_4$ 是高三位也表示一个数码,例如十进制数 12,可写成 0010010,而非二进制的 1100。



### 3. 字母数字码

若把 $A\sim Z, 0\sim 9$ 同时用二进制码来表示,则称为字母数字码。最常用的字母数字码是美国信息交换标准代码,统称 ASCII 码,是英语 AMERICAN STANDARD CODE FOR INFOR

MATION INTERCHANGE 的字头缩写。用一个字节可表示任何一个字母、数码或符号,例如“5”的 ASCII 码 B5 可表示成为:



其中:低四位表示 5,高三位表示 3,如果采用奇校验,则在 b<sub>7</sub> 中置“1”,使字节中“1”的个数为奇数个。这样,高位就可按四位计,变成十六进制的“B”,所以“5”的 ASCII 码值为 (B5)<sub>16</sub>。每个字母、数字、符号都有相应的 ASCII 码值(见附录 I)。

**二、数据代码**

用计算机进行数据处理时,经常用数字、符号或字母来代替某些信息进行运算。编码的功能可归纳为:

其一:为数据项、记录、词组或文件提供一个表达简洁、概念清楚的标识符,既可节省存储空间,又便于查找和运算。

其二:标准化的代码便于数据的交流和共享。

因此,在计算机数据处理行业中广泛应用各类数据代码。

表 1-2 所示为“煤炭工业机电安装工程预算定额”手册中的一页,每项工程定额都有专用的代码。例如表中所示的 0.022~0.035t 液压支架代码是“0014”,只要取出其中各项数据,就可见液压支架的运行指标。如果按代码进行查询计算,将是十分方便的。各行各业都有统一的编码资料,登记、注册,全国通用。

**1. 编码的原则**

各行各业的编码都是有组织、有计划进行的,一般遵守如下原则:

- 1) 选用最小值,或最短的代码值;
- 2) 代码设计要逻辑性强、直观、形象、便于记忆、便于读写;
- 3) 尽可能用通用代码标准,如国标、部标、行业标准代码;
- 4) 尽可能利用计算机设备的特征,以便于输入、输出和计算;
- 5) 代码要符合习惯用法,用英语缩略词,以便于记忆、便于推广;
- 6) 代码值要便于扩充,利于修改。

常用的编码和分类方法,如图 1-3 所示。

表 1-2

采区支架(柱)安装

电算 编号	定额编号		0014	0015	0016	0017	0018		
	项 目		液压支架		金属支架		铰接顶梁		
			0.022~0.035t	0.038~0.06t	0.014~0.035t	0.042~0.062t	0.017~0.031t		
	定额单位		30架						
	基价(元)		25.45	38.84	24.44	39.02	21.89		
其中	人工费(元)		22.04	34.73	20.62	34.52	21.16		
	材料费(元)		3.41	4.11	3.81	4.50	0.73		
	机械费(元)		—	—	—	—	—		
4003 4004 4005 8002 8001	名 称	单 位	单 价	数 量					
	综合工日	工日	5.47	4.03	6.35	3.77	6.31	3.86	
	材 料	机械油	kg	1.07	0.3	0.4	0.4	0.5	0.4
		黄干油	kg	1.74	0.4	0.2	0.2	0.25	0.4
		煤 油	kg	0.61	1	1.2	1	1.2	
		破 布	kg	2.01	0.4	0.5	0.5	0.6	
		砂 布	张	0.20	2	2	2	2	

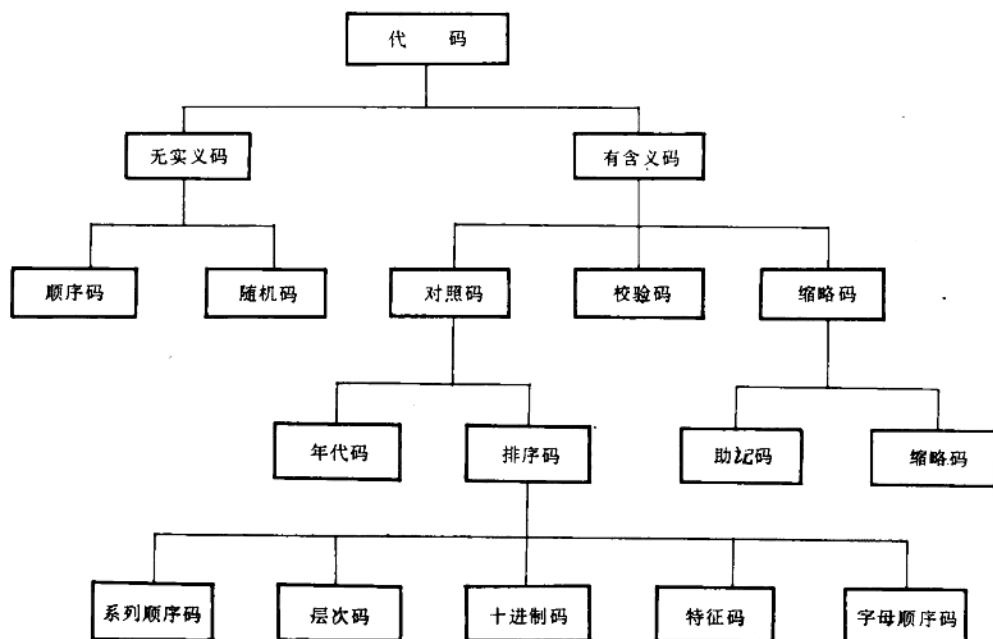


图 1-3 编码方法分类树



## 2. 常用的编码方法

数据编码方法主要有图 1-3 介绍的几种,这里仅就其中的几种予以说明。

1) 顺序码 顺序码就是按自然顺序编写的代码,如书的页码,时间顺序码,职工证代码,电话码等。这种方法简便易行、易于理解,便于推广。其码值一般没有物理意义,不反应事物的特征,是应用十分广泛的代码形式。如表 1-2 的“定额编号”、“电算编号”都是顺序码。

2) 区间码 区间码是把数域分成若干组,每个区间表示一组数,码中的数字和位序相互对应,且代表某种意义。区间码有如下几种:

☆层次码 按层次区间进行分类编码,其结构关系如图 1-3 所示,层次清晰、概念分明、易于识别、易于扩充;

☆十进制编码 每位十进制数代表一个类别。如图书馆编码:

500——自然科学;510——数学;520——天文学;……

每个数据区间又可有自己的细化编号,扩充十分方便。

3) 助记码 将编码对象的名称、规格等特征作为编码的一部分,例如:

TV—B—12——12 英寸黑白电视机;

TV—C—20——20 英寸彩色电视机。

其中 TV、B、C 都是英语词头,且与人们习惯一致。

4) 缩略码 在编码对象的英语名称中,提取 n 个关键的字母做代码。例如:

Amt——总额(Amount);

Cont——合同(Contract);

INV·NO——发票号(Invoice Number)。

5) 校验码 为了保证数据输入、输出和运算结果的正确性,编码结构中都设校验位,计算机在接受数据或运算时都按某种规则自动进行校验,确保数据的正确性。

## § 1-3 计算机软件知识

计算机软件一词有如下三层含义:

1. 个体含义 软件是指用计算机语言编写的指令序列,及其相应的说明性文件。指令序列就是程序,是用计算机语言对运算对象及处理规则的描述,说明性文件是指运行软件所必须的资料说明。

2. 整体含义 软件是指在特定的计算机系统中所有个体软件的总和,亦即,在计算机系统中除硬件以外的所有成分的集合体。

3. 科学含义 软件是指在研究、开发、应用、维护上述含义下的软件所涉及的理论、方法、技术所构成的科学。在这种含义下的软件,也称为软件学。

软件在计算机科学及其应用中的作用可归纳如下几个方面:

1. 软件是用户和硬件之间的界面 人们只有通过软件才能使用计算机硬件,为了用计算机解决自己的问题,首先要编制或购买软件,架起“人一机”接口的桥梁。

2. 软件是计算机系统的司令部 如操作系统可以管理整个计算机系统硬件、外设、软件、文档资料,分配运算任务,确定工作内容,保证系统的自动化运行。没有操作系统,计算机将无法工作。