

计算机技术工人水平考试辅导丛书

微型计算机 数据录入人员水平考试辅导

张茂芝 主编
清华大学出版社

微型计算机

数据录入人员水平考试辅导

张德成 主编
清华大学出版社

计算机技术工人水平考试辅导丛书

微型计算机 数据录入人员水平考试辅导

张茂芝 主编

清华大学出版社

内容简介

本书为数据录入人员水平考试的复习辅导材料。内容包括:数据录入基础知识,数据录入设备基础知识,操作基本知识,印刷出版的一些常识,PC DOS和CC DOS的命令与操作,汉字录入技术,数据录入实用软件EDLIN、英文和汉字字处理软件WordStar,关系数据库系统dBASE的使用,数据录入的生产过程与经验,书末附有数据录入常用英语词语、数据录入人员级水平考试大纲和1989年数据录入人员水平考试笔试试题及答案。

本书可供从事数据录入的技术和管理人员、办公室文秘人员、初级程序员、系统操作员以及想用计算机录入并处理数据的其他人员学习、参考。

计算机技术工人水平考试辅导丛书 微型计算机数据录入人员水平考试辅导

张 茂 芝 主编

清华大学出版社出版

北京清华园

北京西城华海印刷厂印刷

新华书店总店科技发行所发行

开本:787×1092 1/16 印张:7 1/2 字数:178千字

1991年7月第1版 1991年7月第1次印刷

印数:0001~10000

ISBN 7-302-00941-4 /TP·344

定价:4.50元

前 言

本书是根据《数据录入人员级水平考试大纲(试行)》为数据录入人员水平考试辅导而编写的。随着计算机特别是微型计算机的广泛使用,目前专门从事录入的操作员、技术和管理人员以及把数据录入作为一种方便的工具的人数急剧增加,他们都迫切希望有一本比较系统的专门讲述数据录入的小册子,以便学习和参考。正是基于这个目的,编写了此书。希望本书能有益于应考人员的系统复习,也有益于其他人员的学习和参考。

本书第一章简单地介绍了数据录入的概念、对数据录入人员的基本要求并简单地介绍了数制和码制。要求录入人员掌握一些二进制和十六制的知识是必要的,也是必考内容之一。第二、第三章简单地介绍了微型计算机的基本知识和操作基本知识。第四章讲了PC DOS操作系统的命令及其应用。第五、第六、第八及第九这四章则是总结数据录入实用软件的使用,读者可根据实际工作情况选择阅读;对录入人员来讲,至少应掌握一至二种录入软件的使用。第六和第八章分别讲英文字处理软件WordStar和汉化的WordStar。英文录入人员可以第六章为主,汉字录入人员可以第八章为主,两章可互为参考。英文和汉字WordStar的功能和操作命令在许多方面是相同的或相似的。第九章从数据录入的角度讲了dBASE关系数据库,而数据处理方面的内容则涉及不多。第七章讲汉字录入技术。这一章没有讲多少汉字输入的具体方法,这不是说不重要,只是考虑到那样做势必篇幅太大,且必要性不大。第十章简单地总结了目前数据录入的作业流程和技术管理方面的一些经验。学习和使用计算机必须要勤于实践,只要读者按照正确的途径进行练习,必定会有满意的效果。

在这里我真诚地感谢北京软件行业协会沈林兴高级工程师、机电部计算机技术培训中心邵祖英教授、北京信息工程学院陈一凡教授、北京数据转换专业协会秘书长孙华顺以及清华大学出版社、北京水平考试实施办和北京电脑天地学校的同志们对本书的大力支持和关心。本书成稿时间较为仓促,加之本人水平有限,缺点错误在所难免,恳请广大读者批评指正。

张茂芝

一九九〇年五月于清华大学

目 录

第一章	数据录入基础知识	1
1.1	数据录入的基本概念以及数据录入人员的基本要求	1
1.1.1	数据录入的基本概念	1
1.1.2	数据录入人员的基本要求	1
1.2	数制、字符和汉字编码	2
1.2.1	数制	2
1.2.2	数制间的转换	4
1.2.3	字符和汉字编码	6
第二章	数据录入设备基本知识	11
2.1	计算机常识	11
2.1.1	计算机发展简史	11
2.1.2	计算机的基本组成部分	11
2.1.3	计算机操作系统	13
2.1.4	计算机软件常识	13
2.1.5	磁盘和磁带	14
2.2	IBM PC 及其兼容机	16
2.2.1	硬件基本配置	16
2.2.2	软件基本配置	18
2.2.3	国产兼容机	19
第三章	操作基本知识和出版方面的一些常识	20
3.1	开关机	20
3.2	计算机键盘及常用功能键和编辑键简介	20
3.3	键盘录入练习	21
3.3.1	操作姿势和要领	21
3.3.2	指法练习	22
3.4	标点符号的处理及自由行尾	22
3.5	表格的分类和结构	23
3.6	常用印刷字体与字号	24
3.7	繁体字和非正规简化字	26
第四章	PC DOS命令及其使用方法	28
4.1	引言	28
4.2	PC DOS 的启动过程	28
4.3	PC DOS 常用控制键及简单命令	29
4.4	PC DOS 命令的语法格式	30
4.5	PC DOS 命令行编辑	31
4.6	PC DOS 命令种类	31
4.7	文件	31
4.7.1	文件名	31

4.7.2	设备文件	32
4.7.3	文件目录和路径	32
4.8	PC DOS 提供的常用命令	33
第五章	行编辑软件EDLIN	42
5.1	打开文件	42
5.2	行编辑命令	43
第六章	英文字处理软件WordStar	47
6.1	功能菜单和命令种类	47
6.2	启动WordStar	47
6.3	开启功能菜单	48
6.4	主功能菜单	49
6.5	辅助功能菜单	49
6.6	块功能菜单	50
6.6.1	文件保存	50
6.6.2	块操作	50
6.6.3	两个文件之间的块操作	51
6.7	屏幕功能菜单	52
6.7.1	屏幕的有关术语	52
6.7.2	编排的应用	54
6.8	快速功能菜单	54
6.8.1	“快”字的含意	55
6.8.2	查找、查找—替换命令	55
6.8.3	查找串中的通配符	56
6.8.4	重复插入一个短语	56
6.9	打印功能菜单	57
6.9.1	打印效果	57
6.9.2	页面布局	58
6.9.3	垂直页面控制	59
6.9.4	编页命令	59
6.9.5	MailMerge和SpellStar	60
第七章	汉字录入技术	70
7.1	汉字的特征信息和输入编码方案	70
7.1.1	汉字特征信息	70
7.1.2	汉字输入编码方案	71
7.2	计算机汉字输出	73
7.2.1	汉字库	73
7.2.2	显示和打印	73
7.3	汉字操作系统CC DOS	74
7.3.1	CC DOS简介	74
7.3.2	启动CC DOS	74

第八章	汉字文字编辑软件Wordstar	76
8.1	启动Wordstar	76
8.2	D 进入编辑	76
8.3	基本编辑方法	77
8.4	编辑技巧	78
8.5	查找与替换操作	78
8.6	排版	79
8.7	文章页面设计	79
8.8	退出编辑	80
8.9	P 打印文件/中断	81
8.10	R 运行程序	82
8.11	N 编辑非文书文件	82
8.12	M 合并打印	82
8.13	E 更换文件名	84
8.14	O 拷贝文件	84
8.15	Y 删除文件	84
8.16	X 退出	84
第九章	用dBASE III进行数据录入	85
9.1	引言	85
9.2	进入和退出	85
9.3	建库命令	86
9.3.1	库文件的结构	86
9.3.2	用Create 命令建立库文件的结构	86
9.4	输入数据	87
9.5	显示输入内容	88
9.6	增删记录	88
9.7	全屏幕操作常用键	89
9.8	其它的一些常用命令	90
第十章	数据录入的全过程	93
10.1	明确录入要求	93
10.2	选择合适的录入程序	94
10.3	键盘输入	94
10.4	校核	95
10.5	数据处理	97
10.6	输出成品	98
附录一	数据录入常用英语词语	99
附录二	数据录入员级水平考试大纲(试行)	102
附录三	一九八九年计算机软件人员水平考试试题(数据录入员级)及参考答案	103

第一章 数据录入基础知识

1.1 数据录入的基本概念以及数据录入人员的基本要求

1.1.1 数据录入的概念

数据录入的全称是计算机辅助数据录入(CADE——Computer Aided Data Entry),国际上是从六十年代开始发展起来的,并逐步形成一个行业。作为一个行业,我国大陆则是从八十年代初起步的。数据录入也叫数据转换(data conversion)。数据是信息的载体,常见的数据是文字、数字、图表、符号等。数据转换就是将它们转换成机器可读形式(machine readable form)。数据录入与传统的打字不完全相同,它必须按照一定的要求,遵循某种格式,常常需要加上各种控制符,以便于数据库系统管理、调用、或者适应编辑、排版的需要。数据录入随着计算机技术的发展而发展,是迅速发展的办公室自动化(OA——Office Automation)的一个重要组成部分。

数据录入目前主要用于:建立和更新数据库(data base),办公室自动化管理,出版和印刷等方面。

用于数据录入的设备主要有:数据录入专用设备、各种微机(IBM PC及其兼容机等)、各种中英文电子打字机、计算机终端等。还有光电扫描输入设备,由于其成本高和精度差的原因,尚未广泛采用。

1.1.2 数据录入人员的基本要求

使用数据录入设备进行数据录入最基本的操作是从键盘上输入各种数据,所以按照要求既快又准的键盘输入是录入人员的基本功。

数据录入需要“打字”,但与传统的打字不同。数据录入整个过程从头至尾都要使用计算机,借助于计算机进行录入、处理和输出成品。就电子打字机而言,它与传统的打字机也有着根本的不同。它是配有电脑的打字机,是与计算机密切相关的。掌握计算机的基本知识对提高工作效率,保证质量都有很大的帮助。

数据录入设备种类繁多,使用的环境和处理的对象也不尽相同,每个录入人员必须熟练地使用数据录入设备并能对其进行日常维护。

由于计算机大量地使用各种英文提示信息,还有大量的英文数据及其它语种的数据,因此录入人员应有一定的外语(主要是英语)阅读能力,能阅读英文屏幕信息,至少会使用一种英文字处理软件,掌握字处理软件的主要功能,还要能识别英文手写体字符。

综上所述,数据录入人员的基本要求是:要有相当准确度和速度的键盘录入能力;要有计算

机的基本知识;要熟悉并熟练地使用数据录入设备;要有一定的英文阅读能力;要有紧张、细致、认真、负责的工作态度。

1.2 数制、字符和汉字编码

1.2.1 数制(numeric system)

人们在生产与生活中需要计数。在生活中用得最多的是十进制数。总结了十进制数的特点,就很容易理解和掌握其它数制的特点。我们主要考虑正整数。

1.十进制(Decimal)

十进制数的特点是:

(1)有10个数字:0,1,2,3,4,5,6,7,8,9。

这10个数字表示了10种状态。十进制的基数是10。基数就是计数制中数字符号状态的个数。

(2)逢十进一。

除了这两个基本特点外,还可以看出十进制数的一些特点:

(3)每一个数字符号所代表的数值既与它本身有关,又与它在数中的位置有关。正数的数位越往左,数位越高,同样一个数字所代表的数值就越大。

某一数位上的一个单位表示的数的大小称做权或权重(weight)。

十进制数某一数位上的权是它右边一位上的权的十倍,个位上的权为1,十位上的权为10,百位上的权为100,等等。一个十进制数,百位上的1表示一个100,百位上的5表示五个100,百位上的5所表示数的大小是由数字5乘以百位上的权得到的,即 $100 \times 5 = 500$,百位上的5代表500。

(4)一个十进制数的数值是:①将各位上的数字乘以相应位上的权;②再取其积的和。

例如:十进制数 $128 = 1 \times 10^2 + 2 \times 10^1 + 8 \times 10^0 = 100 + 20 + 8$

有时为了避免混淆,将十进制数128写成 $(128)_{10}$,下标10表示十进制数。

一个m位十进制正整数D表示的数值为:

$$(D)_{10} = D_{m-1} \times 10^{m-1} + D_{m-2} \times 10^{m-2} + \dots + D_1 \times 10^1 + D_0 \times 10^0$$

这个十进制数,数字从第(m-1)位(最高位)到第0位(即个位,最低位)共m位。最高位的数字是 D_{m-1} ,其权为基数10的(m-1)次幂: 10^{m-1} ,最低位上的数字为 D_0 ,其权为10的零次幂: $10^0 = 1$ 。

(如果是负数,它所表示的大小是等式右边的整个式子前加上负号。如果是小数,则小数部分的权的幂次取负值。)

除了我们所习惯的十进制外,生活中还有十二进制(12支铅笔是一打,钟表的时针12小时走一圈等等)和其它数制。在计算机中,机器数则广泛使用二进制数。几乎所有的数字电子计算机都是用二进制进行运算的。

2.二进制(Binary)

电子计算机中使用二进制数的理由很简单,就是因为二进制数只有两种状态,电子线路容易表示,容易实现。比如,可以用电子开关的接通表示1,用开关的断开表示0;可以用二进制数1表示有,二进制数0表示无。(当然逻辑上反过来也可以,但这样规定对初学者可能有点不习惯。)

二进制数的两个基本特点是:

- (1)有两个数字符号:0和1。二进制数字符号状态的个数为2,即基数是2。
- (2)逢二进一。

二进制数的规律与十进制数的规律有很多共同之处。

用二进制数来表示十进制正整数:

表1.1列出对应十进制数0到8的9个二进制数。

表1.1

十进制数	二进制数
0	0
1	1
2	10
3	11
4	100
5	101
6	110
7	111
8	1000

二进制数100表示十进制数6,可以写成 $(100)_2 = (6)_{10}$ 。

一位二进制数可以表示 $2^1 = 2$ 个十进制数:0和1;二位二进制数可以表示 $2^2 = 4$ 个十进制数:0到3;三位二进制数可以表示 $2^3 = 8$ 个十进制数:0到7;n位二进制数所能表示的数的个数为 2^n 个,它们对应十进制数0到 $2^n - 1$ 。如果用二进制数表示字符,每一个数表示一个字符,则n位二进制数所能表示的字符个数也为 2^n 个。

英文字母大小写共52个,数字10个,加上空格,符号和一些控制字符,不超过128个, $128 = 2^7$,所以可以用7位二进制代码来表示。若用8位,则可表示 $2^8 = 256$ 个字符。

汉语中单字的个数,包括常用字、次常用字和稀用字,远远大于256,所以不能用7位或8位二进制代码来表示。如果用14位二进制数来表示, $2^{14} = 16384$,则最多可表示一万六千多个汉字和符号。

二进制数中的一位叫做1比特(bit),8个比特构成一个字节(byte)。一个英文字符可以用1个字节来表示,一个汉字可以用2个字节来表示。

通常,文件的大小是以字节为单位来计量的。

数在机器中的表示形式称为机器数。机器数用二进制数表示。(注:机器数的字长取决于构成寄存器的触发器的个数。)

3.十六进制数(Hexadecimal)

二进制数简单,但写起来太长,容易出错。计算机中还常常使用十六进制数和八进制数。

十六进制数的特点是:

- (1)有16个数字:0、1、2、3、4、5、6、7、8、9、A、B、C、D、E、F,基数是16。
- (2)逢十六进一。

(读者会很容易地总结出八进制数的两个基本特点。)

十六进制数的16个数字是在十进制数字的基础上扩充增加了6个数字:A、B、C、D、E、F。(在这里把通常的字母看成数字。)

下表列出对应十进制数0—16的17个二进制数和十六进制数。

表1.2

十进制数	二进制数	十六进制数
0	0	0
1	1	1
2	10	2
3	11	3
4	100	4
5	101	5
6	110	6
7	111	7
8	1000	8
9	1001	9
10	1010	A
11	1011	B
12	1100	C
13	1101	D
14	1110	E
15	1111	F
16	10000	10

1.2.2 数制间的转换

数制间的转换实质上是基数间的转换。

1.二进制和十六进制正整数转换成十进制数
方法:用各位数字乘以相应的权,再将其积相加取和。

例1. $(1011)_2=1\times2^3+0\times2^2+1\times2^1+1\times2^0$
 $=8+0+2+1=(11)_{10}$

例2. $(3AF)_{16}=3\times16^2+10\times16^1+15\times16^0$
 $=3\times256+10\times16+15\times1=(943)_{10}$

2.十进制正整数转换成二进制和十六进制数

方法:将十进制数转换成二进制数,采用“除以2取余”的方法。同样地,将十进制数转换成十六进制数采用“除以16取余”的方法。

例1 将十进制数107转换成二进制数。

2	107	
2	53	1
2	26	1
2	13	0
2	6	1
2	3	0
2	1	1
	0	1

$$(107)_{10} = (1101011)_2$$

步骤:将107除以2,商53,余数1。第一个余数对应所转换的二进制数的最低位;将53除以2,商26,余1;将26除以2,商13,余0;依此类推,最后一步,将上一个商1除以2,得0,余1。注意最后的余数1对应所转换的二进制数的最高位。

小结:

(1)用2不断地去除要转换的十进制数。如果余数为1,则相应的二进制位上的数字为1。如果余数为0,则相应的二进制位上的数字为0。

(2)做第二步和以后各步除法时,被除数是上一次除法运算所得到的商,除数始终为2。

(3)反复用2除,直到最后一步除法所得的商为0时结束。

(4)第一步除法的余数对应最低位上的数字,最后一步除法的余数为最高位上的数字,写结果二进制数时要由下往上写。注意例子中箭头所指的方向是由下而上的。这样直接写出的二进制数即为所求。

(注:这里的107是作为十进制数来转换成二进制数的,要与后面讲的字符编码把107看成三个字符相区别。)

例2 将十进制数107转换成十六进制数。

16	107	
16	6	B(11)
	0	6

$$(107)_{10} = (6B)_{16}$$

例3 将十进制数212转换成十六进制数。

16	212	
16	13	4
	0	D(13)

$$(212)_{10} = (D4)_{16}$$

例4 将十进制数835转换成十六进制数。

16	8	3	5	
16	5	2		3
16	3			4
	0			3



$$(835)_{10} = (343)_{16}$$

3. 二进制数与十六进制数的互相转换

因为 $2^4=16$ ，四位二进制码与一位十六进制码相对应，所以十六进制数与二进制数之间的互相转换很容易。对于十六进制数转换成二进制数，只需将每一位十六进制数字变换成四位二进制数就可以了。对于二进制数转换成十六进制数，只需从最低位起，将每四位二进制数字分成一组，再分别将每一组转换成十六进制数字即可。

例1 将十六进制数6B转换成二进制数。

6对应的四位二进制码为0110，B对应的四位二进制码为1011。所以 $(6B)_{16} = (01101011)_2$ 。(其中最前面的0可略去不写。)

例2 将二进制数1101011转换成十六进制数。

从后(最低位)往前(高位)的第一组四位二进制码为1011，对应十六进制码B，剩下的一组为110，对应十六进制码为6。所以 $(1101011)_2 = (6B)_{16}$ 。

1.2.3 字符编码

最常见的符号信息为文字符号，文字符号的最小单位称为字符(character)。对于一些非文字符号(如一些控制符)也可以称它们为字符。计算机系统不能直接处理文字符号，只能处理表示文字信息或符号信息的数码，称为代码(code)。用计算机处理数据，首先必须把这些数据按照某种规律和约定，编成号，即一组数码，这就是字符编码。

最常见的字符包括数字、字母、汉字和各种符号，也就是在数据录入业中我们通常所说的“数据”一词的具体内容。

字符编码要选择对哪些字符进行编码，以及如何对所选择的字符集进行编码，并确定每一个字符代码的值。字符代码具有权威性，需由权威机构批准。一旦制定，有关方面必须遵守。只有这样，大家才能有章可循，才能实现不同机器之间的信息交换。在特殊情况下，在局部范围内，一种字符编码的内容(包括所选字符和代码取值)可能与其它的字符集有些不同；但在一个系统内部，字符编码必须统一。我们的任务不是去研究字符如何编码，而是使用现成的字符代码。数据录入中常用的字符代码有西文ASCII和EBCDIC字符代码、我国汉字国家标准GB2312-80字符代码等。

1. ASCII代码字符集

ASCII代码是一种使用很广的字符编码，其全称是美国标准信息交换代码。ASCII是英文American Standard Code for Information Interchange的字首缩写。IBM-PC及兼容机就使用了ASCII码。标准ASCII码是7位(bit)长，取值范围是十进制数值0到127(对应十

六进制数值00到7F),共128个。前32个和最后一个退格符,我们把它们看成控制字符,第33个为空格(也称硬空格,在键盘上敲空格键可以输入),其余的是可显示图形符号的字符,共94个。这94个字符中,包括26个大写英文字母、26个小写英文字母、10个数字,以及加减乘除号、百分号、大于号、小于号、标点符号等32个符号。扩展的ASCII码是8位,取值范围是十进制数值0到255(对应十六进制数值00到FF),共256个。

表1.3 ASCII码字符集速查表

DECIMAL VALUE		0	16	32	48	64	80	96	112	128	144	160	176	192	208	224	240
HEXA DECIMAL VALUE		0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
0	0	BLANK (NULL)	▶	BLANK (SPACE)	0	@	P	'	p	Ç	É	á				∞	≡
1	1	☺	◀	!	1	A	Q	a	q	ü	æ	í				β	±
2	2	☺	↕	"	2	B	R	b	r	é	Æ	ó				Γ	≥
3	3	♥	!!	#	3	C	S	c	s	â	ô	ú				π	≤
4	4	♦	¶	\$	4	D	T	d	t	ä	ö	ñ				Σ	∫
5	5	♣	§	%	5	E	U	e	u	à	ò	Ñ				σ	∫
6	6	♠	=	&	6	F	V	f	v	å	û	ä				μ	÷
7	7	•	↓	'	7	G	W	g	w	ç	ù	ó				τ	≈
8	8	•	↑	(	8	H	X	h	x	ê	ÿ	ï				ø	°
9	9	○	↓	)	9	I	Y	i	y	ë	Ö	Γ				θ	•
10	A	○	→	*	:	J	Z	j	z	è	Ü	Γ				Ω	•
11	B	♂	←	+	;	K	I	k	{	ï	ç	½				δ	√
12	C	♀	└	,	<	L	\	l	!	î	£	¼				∞	n
13	D	♪	↔	-	=	M	J	m	}	ì	¥	ì				φ	²
14	E	♪	▲	.	>	N	^	n	~	Ä	ß	«				€	■
15	F	☼	▼	/	?	O	_	o	Δ	Å	ƒ	»				∩	BLANK IF

工作中需要查阅ASCII字符集和有关的其它字符集以及汉字编码表。要学会使用。在表1.3 ASCII码表中,上面第一行和左边第一列为十进制数值,上面第二行和左边第二列为十六进制数值。表中字符的十进制编码值是对应的上行值与对应的左列值之和。对应的十六进制编码值是二位数,上行值为高位数字,左列值为低位数字。

例如,由字符查找其十进制编码值和十六进制编码值:大写字母A,其对应的上方第一行

十进制数值为64, 左边第一列十进制值为1, 则A的ASCII码十进制值为64+1=65。大写字母A的十六进制编码上行值为4, 左列值为1, 则A的十六进制编码值为41。同一字符的十进制值与十六进制值当然是相等的 $(65)_{10}=(41)_{16}$ 。有时把 $(41)_{16}$ 写成41H, 意义是相同的。

如果知道一个字符的十六进制编码, 也很容易查出其对应的字符。例如十六进制编码7E, 先在第二行中找到十六进制高位码7, 再在左二列找到十六进制低位码E, 根据7所在的列与E所在行的交叉点, 可确定对应的字符为~。如果知道一个字符的十进制编码值, 也不难确定其对应的字符。例如要查找对应十进制值85的字符, 先将85减去最接近85的16的整倍数值80, 其差为5。找出80所处的列与5所处的行的交叉点, 即可确定其对应的字符为大写字母U。由表可以看出, 数字0到9, 大写字母A到Z, 小写字母a到z, 它们的编码值分别是连续递增的, 依次加1。如果知道小写字母i的十六进制编码值为69, 则可推出小写字母j的十六进制编码值为6A。在ASCII字符集中, 左右双引号是用同一个字符, 减号和英文连字符也用同一个字符, 不加区分。

2.EBCDIC代码字符集

在数据录入业, 尤其是国外的一些大数据库, 广泛使用这种字符编码, 它的全称是扩展的二——十进制交换码(Extended Binary——Coded Decimal Interchange Code)。各个公司和机构使用的EBCDIC字符集与编码主要部分或大部分相同, 但有一些差别, 使用时要注意各个公司的不同要求, 必要时要进行查表。表1.4给出一张EBCDIC字符集代码表。

表1.4 EBCDIC 字符集

字符	十六进制码	字符	十六进制码	字符	十六进制码	字符	十六进制码
Space	40	a	81	A	C1	0	F0
¢	4A	b	82	B	C2	1	F1
.	4B	c	83	C	C3	2	F2
<	4C	d	84	D	C4	3	F3
(	4D	e	85	E	C5	4	F4
+	4E	f	86	F	C6	5	F5
	4F	g	87	G	C7	6	F6
&	50	h	88	H	C8	7	F7
!	5A	i	89	I	C9	8	F8
\$	5B	j	8B	J	D1	9	F9
*	5C	l	8C	K	D2		
)	5D	j	91	L	D3		
:	5E	k	92	M	D4		
┌	5F	l	93	N	D5		
=	60	m	94	O	D6		
/	61	n	95	P	D7		
,	6B	o	96	Q	D8		
%	6C	p	97	R	D9		
└	6D	q	98	\	E0		
>	6E	r	99	S	E2		
?	6F	ı	9C	T	E3		
:	7A	~	A0	U	E4		
#	7B	s	A2	V	E5		
@	7C	t	A3	W	E6		
┐	7D	u	A4	X	E7		
,	7E	v	A5	Y	E8		
"	7F	w	A6	Z	E9		
		x	A7				
		y	A8				
		z	A9				
		{	AB				
		[	AC				
		^	BC				

在EBCDIC码表里,数字0到9是连续的,对应十六进制码Hex.F0到Hex.F9(经常用Hex.表示十六进制码),但26个大小写字母都分成了三段,小写字母a(Hex.81)到i(Hex.89)、j(Hex.91)到r(Hex.99)、s(Hex.A2)到z(Hex.A9)各成一段,大写字母A(Hex.C1)到I(Hex.C9)、J(Hex.D1)到R(Hex.D9)、S(Hex.E2)到Z(Hex.E9)各成一段。在排列上,与ASCII码相比,EBCDIC不连续,EBCDIC码小写字母排在前,大写字母排在后。由于编码方式不一样,ASCII与EBCDIC字符的代码值当然不相同,空格在ASCII中的十六进制代码为20,在EBCDIC中的十六进制代码为40。

3. 汉字标准交换码(Chinese standard exchange code)

ASCII和EBCDIC是西文字符信息交换码,我国也制定了信息交换汉字编码标准。我国是个多民族国家,大多数人属于汉族,使用的文字主要是汉字。但要区分“汉字”和“中文”的提法,不能混同,中文应包括少数民族的文字。全国各民族的文字中除汉字外都是拼音文字。

制定汉字编码国家标准时要考虑汉字的一些属性(attribute)。汉字是表意文字,或称象形文字(ideographic character),它的每个字有其特有的形状和构造。这是不同于各种拼音文字的一大特色。汉语单音词特别多,组词能力特别强。从这一点上来讲,在信息时代,汉字有其特殊的优越性,可以很方便地构成新的词汇。汉字的突出缺点是单字字量特别多。我国汉字的字量多达五、六万。对于不同的汉字,其使用频度的差别是很大的。同一汉字在不同专业领域中使用时,其频度也有差异。对不同专业字频分别进行统计,在若干有代表性的专业领域中统计出各自所用汉字的频度后,取其平均值得到一种综合频度。在实用处理上,根据汉字的使用频度可以把汉字分为常用字、次常用字、稀用字、罕用字或一级字、二级字、三级字等几个等级。等级的名称和范围在不同场合下是不完全相同的。1981年我国颁布了《信息交换用汉字编码字符集——基本集》,即国家标准GB2312—80。这个标准用两个7位二进制代码代表一个汉字或非汉字图形字符,用以作为计算机系统之间汉字信息交换用的标准代码。在国家标准GB2312中除682个非汉字图形符号外,规定了作为进行一般汉字信息处理交换用的一级汉字3755个,我们称为常用字;二级汉字3008个,我们称为次常用字;共计6763个汉字。根据需要今后尚需制定扩充集(除基本集外还有第一辅助集和第二辅助集)。

根据我国1975年进行的查频统计,基本集内6763个汉字复盖率可达99.99%左右,其中3755个一级常用汉字的复盖率可达99.9%左右。

为便于人们检索汉字代码,字符集采用了特殊的排序方法。由于一级汉字都是“常用字”,一般都知道读音,故按汉语拼音字母顺序排列。(多音字取它的常用音,同音调字以起笔笔形横、竖、撇、捺、点、折为序。若起笔相同,则按第二笔,依次类推。)二级汉字大部分较生僻,不容易掌握读音,所以按部首排列较易检索。(部首与一般通用字典相同,略有合并,部首顺序按笔画数排列,变形部首排在正部首之后,同部首字按除去部首以外的画数排列,同画数的字也按起笔笔形顺序排列。)

国家标准GB2312《信息交换用汉字编码字符集》是汉字信息处理系统中的基础性代码标准之一。它广泛用于我国通用汉字系统的信息交换及硬、软件设计工作中。例如汉字字模库、汉字整个输入键盘盘面文字的选择、汉字输入码的转换、汉字输出设备的汉字地址码、处理汉字的高级语言、汉字数据库系统等都采用GB2312或以GB2312为基础进行设计。

GB2312代码表纵向分成94个区(section),横向将每个区分成94个位置(position)。在