

第三节 信息编码

本节主要内容

- 一. 数值在计算机中的表示形式
- 二. 字符编码

一、数值在计算机中的表示形式

1、什么是数制

数制是人们使用符号来记数的方法。按进位的方法进行计数，称为进位计数制。常用的数制有十进制、二进制、十六进制、八进制等。

- 十进制数用D (Decimal) 来表示
- 二进制数用B (Binary) 来表示
- 八进制数用O (Octal) 来表示
- 十六进制用H (Hexadecimal) 来表示
- 计算机中识别的是二进制

(1) 十进制

数码：0, 1, 2, 3, 4, 5, 6, 7, 8, 9

基数：10

位权： 10^i ($i = \dots, -2, -1, 0, 1, 2, \dots$)

原则：逢10进1

例如： $523791.24D = 5 \times 10^5 + 2 \times 10^4 + 3 \times 10^3 + 7 \times 10^2$
 $+ 9 \times 10^1 + 1 \times 10^0 + 2 \times 10^{-1} + 4 \times 10^{-2}$

(2) 二进制

数码：0, 1

基数：2

位权： 2^i ($i = \dots, -2, -1, 0, 1, 2, \dots$)

原则：逢2进1

例如： $110.101B = 1 \times 2^2 + 1 \times 2^1 + 0 \times 2^0 + 1 \times 2^{-1}$
 $+ 0 \times 2^{-2} + 1 \times 2^{-3}$

(3) 八进制

数码：0, 1, 2, 3, 4, 5, 6, 7

基数：8

位权： 8^i ($i = \dots, -2, -1, 0, 1, 2, \dots$)

原则：逢8进1

例如： $6734.620 = 6 \times 8^3 + 7 \times 8^2 + 3 \times 8^1 + 4 \times 8^0 + 6 \times 8^{-1} + 2 \times 8^{-2}$

(4) 十六进制

数码：0, 1, 2, 3, 4, 5, 6, 7, 8, 9, A, B, C, D, E, F

基数：16

位权： 16^i ($i = \dots, -2, -1, 0, 1, 2, \dots$)

原则：逢16进1

例如： $3AF.2H = 3 \times 16^2 + 10 \times 16^1 + 15 \times 16^0 + 2 \times 16^{-1}$

(5) 机制之间的转换

- 二进制、八进制、十六进制转换成十进制数
 - 方法：按权展开，并求和；
 - $1011B=?$
- 十进制转换成二进制、八进制、十六进制
 - 方法：用2/8/16连续去除十进制数，直到商等于0为止，整数部分逆序取余、小数部分正序去整。

举例说明：

- $98.6875D = (1100010.1011) B$
- $98.6875D = (142.54) O$
- $98.681875D = (62.AC) H$

- 二进制整数和八进制以及十六进制之间的快速转化
 - 由于二进制数中的8个数 $(000)_2, (001)_2, (010)_2, \dots, (111)_2$ 分别与八进制数中的8个数码 $0, 1, 2, \dots, 6, 7$ 对应，可根据这一对应关系完成二进制数到八进制数的转换。
 - 从最低位开始向高位进行，每3位二进制数用1位八进制数表示，高位或低位不足一组的补0；
 - 将二进制数 $(11101110.00101011)_2$ 转换成八进制数
 - 结果为356.126

- 由于二进制数中的16个数 $(0000)_2, (0001)_2, (0010)_2, \dots, (1111)_2$ 分别与十六进制数中的16个数码 $0, 1, 2, \dots, 9, A, B, C, D, E, F$ 对应，可根据这一对应关系完成二进制数到十六进制数的转换。
- 从最低位开始向高位进行，每4位二进制数用1位十六进制数表示，高位或低位不足一组的补0；
- 将二进制数 $(10111100101.00011001101)_2$ 转换成八进制数
- 结果为5E5.19A

二、字符编码

编码是指对输入到计算机中的某种非数值类型数据用二进制数表示的转换规则。

美国标准信息交换码 (American Standard Code for Information Interchange , **ASCII**) 是目前计算机中普遍采用的标准编码形式。

字符编码

- ASCII码是用7位二进制编码表示1个字符，用一个字节存放共表示 $2^7 = 128$ 个字符。
- ASCII码最高位为0，真正使用的是低7位。
- 共包括33种控制字符，普通字符95个，包括52个英文大小写字母，10个数字以及其它字符。

汉字编码

是指汉字在计算机中表示的形式。我国国家标准采用连续的两个字节表示。且规定每个字节的最高位为1，以与ASCII码最高位为0之间进行区分。

ASCII码表——常考考点出处

ASCII 字符代码表 一

高四位		ASCII非打印控制字符										ASCII 打印字符											
		0000					0001					0010	0011	0100	0101	0110	0111						
		0					1					2	3	4	5	6	7						
低四位		+进制	字符	ctrl	代码	字符解释	+进制	字符	ctrl	代码	字符解释	+进制	字符	+进制	字符	+进制	字符	+进制	字符	+进制	字符	ctrl	
0000	0	0	BLANK NULL	^@	NUL	空	16	▶	^P	DLE	数据链路转意	32		48	0	64	@	80	P	96	`	112	p
0001	1	1	☺	^A	SOH	头标开始	17	◀	^Q	DC1	设备控制 1	33	!	49	1	65	A	81	Q	97	a	113	q
0010	2	2	☺	^B	STX	正文开始	18	↕	^R	DC2	设备控制 2	34	"	50	2	66	B	82	R	98	b	114	r
0011	3	3	♥	^C	ETX	正文结束	19	!!	^S	DC3	设备控制 3	35	#	51	3	67	C	83	S	99	c	115	s
0100	4	4	◆	^D	EOT	传输结束	20	¶	^T	DC4	设备控制 4	36	\$	52	4	68	D	84	T	100	d	116	t
0101	5	5	♣	^E	ENQ	查询	21	♫	^U	NAK	反确认	37	%	53	5	69	E	85	U	101	e	117	u
0110	6	6	♠	^F	ACK	确认	22	■	^V	SYN	同步空闲	38	&	54	6	70	F	86	V	102	f	118	v
0111	7	7	●	^G	BEL	震铃	23	↕	^W	ETB	传输块结束	39	'	55	7	71	G	87	w	103	g	119	w
1000	8	8	◻	^H	BS	退格	24	↑	^X	CAN	取消	40	(56	8	72	H	88	X	104	h	120	x
1001	9	9	○	^I	TAB	水平制表符	25	↓	^Y	EM	媒体结束	41)	57	9	73	I	89	Y	105	i	121	y
1010	A	10	◻	^J	LF	换行/新行	26	→	^Z	SUB	替换	42	*	58	:	74	J	90	Z	106	j	122	z
1011	B	11	♂	^K	VT	竖直制表符	27	←	^[ESC	转意	43	+	59	;	75	K	91	[107	k	123	{
1100	C	12	♀	^L	FF	换页/新页	28	└	^\ FS	文件分隔符	44	,	60	<	76	L	92	\	108	l	124		
1101	D	13	♪	^M	CR	回车	29	↔	^] GS	组分隔符	45	-	61	=	77	M	93]	109	m	125	}	
1110	E	14	🎵	^N	SO	移出	30	▲	^_	RS	记录分隔符	46	.	62	>	78	N	94	^	110	n	126	~
1111	F	15	☼	^O	SI	移入	31	▼	^-	US	单元分隔符	47	/	63	?	79	O	95	_	111	o	127	␣ ^Back space

注：表中的ASCII字符可以用:ALT + “小键盘上的数字键”输入

本节习题

1、十进制数向二进制数转换时，十进制数91相当于二进制数（ ）

- A. 1101011
- B. 1101111
- C. 1110001
- D. 1011011

【答案】 D

本节习题

2、“美国信息交换标准代码”的缩写是（ ）

A . EBCDIC

B . ASCII

C . GB2310 80

D . BCD

【答案】 B

本节习题

3、已知英文字母m的ASCII码值为109，那么英文字母p的ASCII码值为（ ）

- A. 111
- B. 112
- C. 113
- D. 114

【答案】 B