

1.3 数据的采集与编码 1.3.2 数字化 1.3.4 编码

【课时目标】

1. 理解编码的概念与具体应用，掌握常见的字符编码。
2. 掌握声音和图像数字化的原理并能计算声音与图像的存储容量。

【知识梳理-课前预习】

1. 字符编码

编码(Encoding)是信息从一种形式按照某种规则或格式，从一种形式转换为另一种形式的过程。**解码**是编码的逆过程。常见的字符编码有**ASCII**、Unicode 及各种汉字编码。

(1) ASCII 码

ASCII 主要用于表示英语字符。基本的 ASCII 码共有 **128** 个，由 **7** 位的二进制组成。bb 在计算机中存储时最高位加 **0**，占用 1B。二进制范围为 **0000 0000** 到 **0111 1111**，即十六进制的 **00** ~ **7F**。

(2) 汉字编码

汉字编码分为外码、**交换码**、**机内码** 和字形码。

外码(又称输入码)：用来将汉字输入到计算机中的一组键盘符号常见的输入码：拼音码、五笔字形码等。

交换码(又称区位码)：用于各种数字系统之间的汉字信息的通信交换。区位码也称为《信息交换用汉字编码及字符集》，简称 GB2312 编码。

处理码(又称汉字机内码)：是计算机内部用于信息处理的汉字代码。机内码也称为内码，不同的字符集有不同的表示方式，GB2312 采用 **2** 个字节表示一个汉字。

2. UltraEdit 查看字符内码

1 个英文字符(ASCII 码)占用 **1** 个字节存储空间，1 个中文字符需要占用 **2** 个字节的存储空间；左边的显示内容是字符对应的十六进制格式的内码，1 个英文字符需要 **2** 位 16 进制位表示，1 个中文字符需要 **4** 位 16 进制位表示。例如用 UltraEdit 软件观察字符“AF 相机”的内码，结果如下图所示：

0	1	2	3	4	5	6	7	8	9	a	b	c	d	e	f
00000000h: 41 46 CF E0 BB FA ; AF 相机															

A (41H), F (46H), 相 (CFE0H), 机 (BBFAH)

4. 条形码

条形码(barcode)是将宽度不等的多个黑条和白条，按照一定的编码规则排列，用以表达一组信息的图形标识符。我国普遍采用的 EAN13 条形码，由 13 位数字组成，前 3 位数字表示国家代码，最后一位是校验码。



5. 二维码

二维条码/二维码(2-dimensional bar code)是用某种特定的几何图形按一定规律在平面上(二维方向)分布的黑白相间的图形记录数据符号信息。

【课堂练习】

1. 大写字母 H 的 ASCII 码为 72, 则大写字母 J 的 ASCII 值的十六进制值是 **D**
 A. 50H B. 49H C. 48H D. 4AH
2. ASCII 码表中的大写字母 Z 后有 6 个其他字符, 接着便是小写字母。现在已知: 字母 Y 的 ASCII 码为 $(1011001)_2$, 则字母 a 的 ASCII 码用十六进制表示是 **A**
 A. 61H B. 62H C. 63H D. 64H
 (Handwritten: Y 2 00 00 00 00 a 59 5A 60 61)
3. 用 UltraEdit 软件观察到“学”字的内码为 D1A7H, 其对应的二进制编码是 **D**
 A. 11010001 B. 010000110100010 C. 11000110 D. 1101000110100111⁷
4. 用 UltraEdit 软件观察“Book”这几个字符的内码, 如图所示, 则“Look”这几个字符的编码是 **C** **B-42H**

0	1	2	3	4	5	6	7	8	9	a	b	c	d	e	f
00000000h: 42 6F 6F 6B										; Book					

- A. 42 6F 6F 6B B. 4B 6F 6F 6B **B C D E F G H I J K L**
 C. 4C 6F 6F 6B D. 4A 6F 6F 6B **42** **4C**
5. 图中内码表示的对象可能是 **B**

Offset	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
00000000	C6	BB	B9	FB	61	70	70	6C	65							

- A. 9 个 ASCII 字符 **汉字** B. 2 个中文字符, 5 个 ASCII 字符
 C. 5 个中文字符 D. 4 个中文字符, 1 个 ASCII 字符
6. 下列说法正确的是 **D**

Offset	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
00000000	A3	C1	BC	C6	BB	AE										
	A计划															

- A. “A” 占用了 1 个字节 B. “计” 占用了 1 个字节 **2字节**
 C. “A” 的内码为 A3 **A3 C1** D. “计” 的内码为 BC C6
7. 使用 UltraEdit 软件观察字符内码, 部分界面如图所示。下列说法正确的是 **B**

00000000h:	50	61	6E	64	61	73	CA	FD	BE	DD	B4	F3	CA	A6		
	; Pandas数据大师															

- A. “数” 的内码为 “FD BE” **CA FD** B. “a” 的内码为 “61”
 C. “大” 占用 2bit **2B** D. “P” 占用 2Byte **1B**
8. 某图书馆系统计划为其藏书采用新的条形码系统, 每本书的条形码能够唯一标识, 现至少要标识 10000 本书。条形码由一系列黑条和白条组成, 其中每一条都可以编码为 0 或 1。该条形码至少需要的二进制编码位数是 **C** **$2^4 > 10000$**
 A. 10 位 B. 12 位 C. 14 位 D. 16 位
9. 下列说法不正确的是 **B**
- A. 我国普遍采用 EAN13 条形码, 这种条形码由 13 位数字组成
 B. 条形码是汉字编码方式之一 **汉字编码 GB2312、GBK、GB18030 和 Unicode**
 C. 二维码可以存储包括网址、名片、文本、特定代码在内的各种信息
 D. 二维码中可能包含病毒信息, 因此不能盲目扫描来历不明的二维码

【课后作业】

1.使用 UltraEdit 软件观察字符，下列说法正确的是 **D** !



- A.字符“!”的内码占一个字节 **2B** B.字符“J”的十六进制内码为 50H **G-47,J-4A**
 C.字符“10”的十六进制内码为 **3AH** D.图中共有 6 个 ASCII 编码, 4 个汉字编码

(50)3

2.下列关于数制和信息编码的说法, 正确的是 **D**

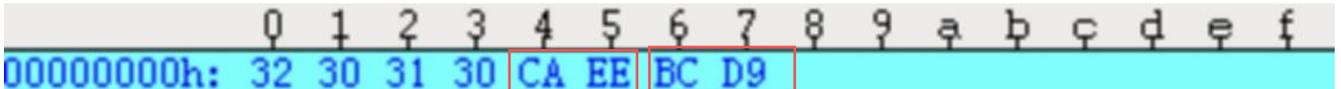
- A. 末位为 0 的 **任意进制** 正整数, 转换为十进制数后一定为 **偶数** **奇数进制, 奇偶不定**
 B.若字符“Y”的内码用十六进制表示为 59H, 则“Z”的内码用十六进制表示为 60H **5AH**
 C.英文字符在计算机内部使用 **ASCII 码** 进行存储和处理 **二进制**
 D.用 x 位 k 进制数给 y 种不同的信息进行编码, 必须满足 $y \leq k^x$

3.小王用 UltraEdit 软件观察字符的内码, 如图所示, 下列说法中正确的是 **B**

CD	AC	D2	BB	B8	F6	CA	C0	BD	E7	: 同一个世界
CD	AC	D2	BB	B8	F6	C3	CE	CF	EB	: 同一个梦想
4F	6E	65	20	64	72	65	61	6D		: One dream
4F	6E	65	20	57	6F	72	6C	64		: One World

- A. “同一个梦想”字符在计算机内存处理和使用时使用的是 **交换码** **二进制**
 B. “One dream”字符在计算机内存存储时, 每个字符占 1B
 C. 字母“O”的 ASCII 码是 01000111B **× O-4F**
 D. “世界”内码的十六进制表示是 BBB8H **CA C0 H**

4. 用 UltraEdit 软件观察字符内码, 结果如下图所示, 则图中内码表示的对象可能是 **C**



- A. 8 个英文字符 B. 4 个汉字字符
 C. 4 个英文字符, 2 个汉字字符 D. 2 个英文字符, 3 个汉字字符

5. 某一数字显示器由 7 段发光二极管组成, 通过各二极管的亮灭来显示数字。若用“1”表示亮, “0”表示灭, 按图 2 所示的字母先后顺序对各二极管的亮灭状态编码, 按此编码方法, 图 1 中数字“9”所对应的编码是 **B**



图 1

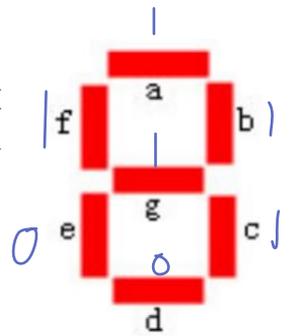


图 2

abcdefg
1110011

- A. 1111011 B.1110011 C.1111100 D.1111110

6. 电子表中常用由七段灯管组成的“**8**”来显示信息, 其中每段灯管有“亮”和“暗”两种状态, 则该符号能显示的最多信息数为 **C** **2⁷=128**

- A. 8 B. 16 C. 128 D. 256

7.有一种利用打孔透光原理设计的简易身份识别卡: 每张卡在规定位置上有一排顶打孔位, 读卡器根据透光检测判断哪些孔位已打孔, 哪些未打孔, 从而识别出卡的编码。如果要设计一种供 257 人使用的身份卡, 则卡上的预打孔位至少需要 **B**

- A、7 个 B、9 个 C、10 个 D、12 个 **2⁹>257**

8. 某无人便利店的商品便签上有二维码, 扫描后可以查看具体的商品价格、实时数量、图片等信息, 则下列说法不正确的是 **C**

A.该商品的二维码属于数据

B. 一个 24×24 由黑白像素组成的二维码至少需要 72 字节的存储空间

24*24/8B=72B

- C. 扫描二维码后可以查看商品的图片，这是一个**编码**过程 **解码**
 D. 不同时间扫描二维码后看到的商品实时数量不同，说明信息具有时效性

9. 使用 Word 软件进行文字处理时，伴随着“录入→保存→打印”的过程，其中所涉及的汉字编码分别为 **D**

- A. 外码→处理码→交换码 B. 外码→交换码→字形码
 C. 输入码→交换码→机内码 D. 输入码→处理码→字形码

10. 使用 UltraEdit 软件观察字符“A-Z a-z 0-9”的内码，部分界面如图所示。

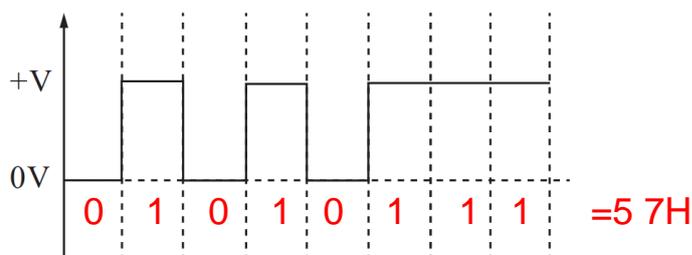
	0	1	2	3	4	5	6	7	8	9	a	b	c	d	e	f
00000000h:	41	2D	5A	20	61	2D	7A	20	30	2D	39					

; A-Z a-z 0-9

下列说法正确的是 **D**

- A. 字符“F”、“d”、“6”的内码值由小变大 **6 F d 由小到大**
 B. 字符“a”与“A”内码值之差的十进制表示为 **20** **32D=20H**
 C. 字符“2”内码值的二进制表示为“00000010” **32H ← 2**
 D. 图中所有字符内码值的十进制表示均小于 128 **0-127**

11. 单极性不归零码是传输二进制数据时采用的一种编码方式，用正电压表示二进制数 1，用零电压表示二进制数 0（按从左往右的顺序编码）。某 8 位二进制数的传输图像如图所示，下列选项能表示该数的是 **D**



- A. 10101000B B. 11101010B C. 53H D. 57H

12. 使用 UltraEdit 软件观察字符内码，部分界面如图所示。

	0	1	2	3	4	5	6	7	8	9	a	b	c	d	e	f
00000000h:	63	3D	28	6D	2B	31	29	2A	36	30	2F	6D				

; c=(m+1)*60/m

则“a=6*n+1”对应的字符内码的界面为 **B**

- A.

00000000h:	63	3D	36	2A	6E	2B	31
------------	----	----	----	----	----	----	----

 B.

00000000h:	61	3D	36	2A	6E	2B	31
------------	----	----	----	----	----	----	----

 C.

00000000h:	61	3D	36	2A	6D	2B	31
------------	----	----	----	----	----	----	----

 D.

00000000h:	61	3D	36	30	2A	6E	2B	31
------------	----	----	----	----	----	----	----	----

13. 下列关于信息编码的说法，正确的是 **D**

- A. 用手机扫描支付二维码是信息编码的过程 **解码**
 B. 只有数字化的数据才能被保存和处理 **← 在计算机中**
 C. 已知大写字母“I”的 ASCII 值是 49H，则小写字母“i”的 ASCII 值是 **6AH** **69H**
 D. n 进制表达式 $(6n+(1)n)$ 的计算结果为两位 n 进制数，则 n 一定为 7 **大小写相差 20H**