1、一个二进制是一位

2、8个位对应一个字节，常用16进制表示

3、7位的字符编码，即每个字节的最高位第8位为0，其余7位的某个值对应着某个字符。

1. ASCII字符集共27=128个字符 = 33个控制字符 + 95个可见字符

编码：就是将二进制转化为字符串

unicode到utf8转化规则：

0000 - 007f //0xxxxxxx

0080 - 07ff //110xxxxx 10xxxxxx

0800 - ffff //1110xxxx 10xxxxxx 10xxxxxx

实例：严

Unicode:4e 25

gbk:d1 cf

uft8:e4 b8 a5

说明：4e 25的二进制是1001110 00100101

4e 25 处于 0800 - ffff之间，

所以套用对应的模版1110xxxx 10xxxxxx 10xxxxxx

从最后一位开始填入到x（不足补0），得到

utf8的二进制 11100100 101110 00 10100101

得到e4 b8 a5