UTF16包含1个字符, UTF8包含2个字符.

如何判断: 首先读取第一个字节并将其转换为二进制表示. 若最高位是0, 则该字符使用1个字节编码; 若从最高位顺序是110, 则该字符使用2个字节编码; 若从最高位顺序是1110, 则该字符使用3个字节编码; 若从最高位顺序是11110, 则该字符使用4个字节编码.

UTF16使用双字节或者4字节编码, 不能表示UCS中所有字符.