

## Códigos basados en binario

En UTF-16 cada punto de código entre U+0000 y U+FFFF se codifica, sin cambios, utilizando 16 bits

Sus características principales son:

- Es capaz de representar cualquier carácter Unicode.
- Utiliza símbolos de longitud variable: 1 o 2 palabras de 16 bits por carácter Unicode (2 o 4 bytes). La unidad de información es la palabra de 16 bits.
- Está optimizado para representar caracteres del plano básico multilingüe (BMP) y caracteres del rango U+0000 a U+FFFF. El BMP contiene la gran mayoría de caracteres y sistemas de escritura en uso en la actualidad. Cuando se limita al plano básico multilingüe, UTF-16 puede ser considerado una forma de codificación con símbolos de tamaño fijo (16 bits).
- No superposición: Los símbolos de 1 palabra (16 bits) utilizan un subconjunto de valores que no puede utilizarse en símbolos de 2 palabras (32 bits).

ASCII es un código de 7 bits con 128 caracteres (2<sup>7</sup>) definidos, pero además cuenta con 33 caracteres no imprimibles y 95 imprimibles y comprende tanto letras, signos de puntuación y números como caracteres de control.

Normalmente el código ASCII se extiende a 8 bits (1 byte) añadiendo un bit de control, llamado bit de paridad.