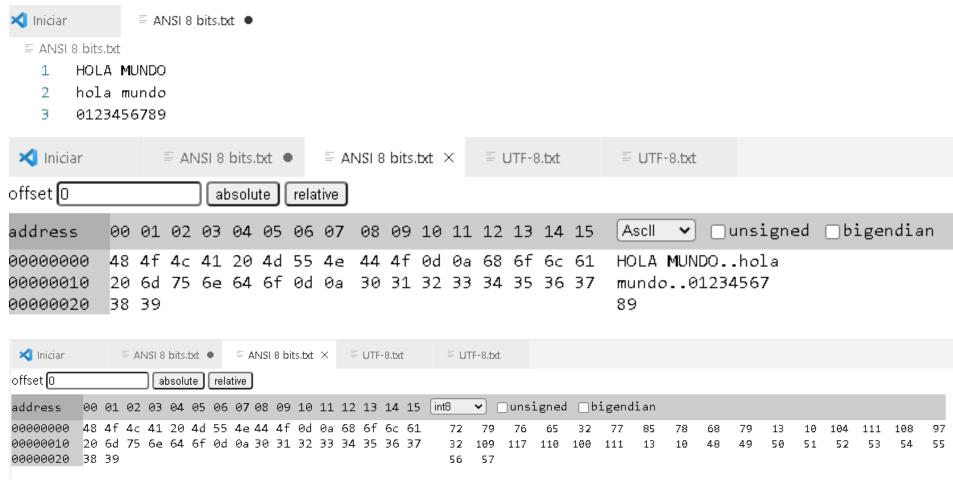
## Codificación y decodificación de caracteres

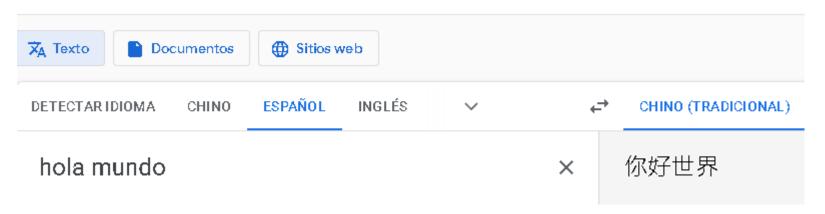
ANSI (8 bits) que sustituye a ASCII (7 bits)



## UTF-8 (Plano básico multilingüe)

https://www.charset.org/utf-8

## Google Traductor

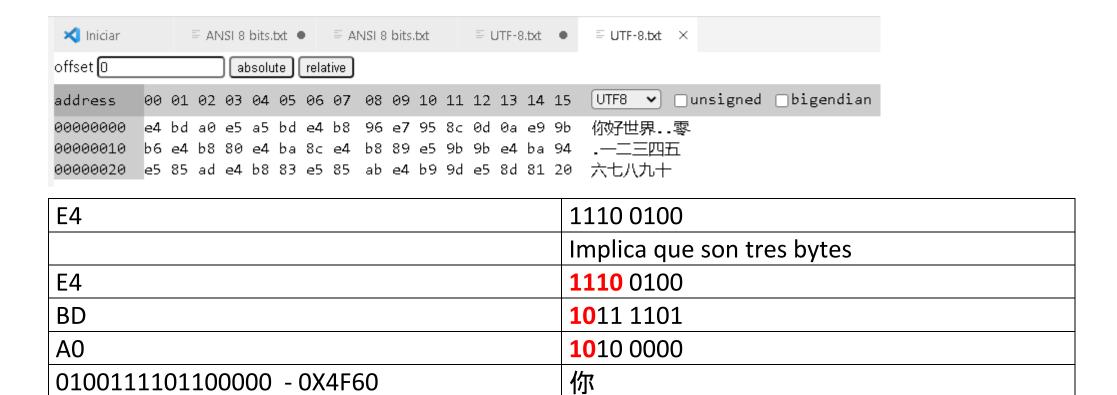


## Los números en chino del 1 al 10

Número	Chino	Pinyin
0	零/○	líng
1	_	уī
2	=	èr
3	三	sān
4	四	sì
5	五	₩ŭ
6	六	liù
7	t	qī
8	Д	bā
9	九	jiŭ
10	+	shí



- UTF-8.txt
  - 1 你好世界
  - 2 零一二三四五六七八九十



E5	<b>111</b> 0 0101
	Implica que son tres bytes
E5	<b>1110</b> 0101
A5	<b>10</b> 10 0101
BD	<b>10</b> 11 1101
0101 1001 0111 1101 - 0X597D	拨

E4	1110 0100
	Implica que son tres bytes
E4	<b>1110</b> 0100
B8	<b>10</b> 11 1000
96	<b>10</b> 01 0110
0100 1110 0001 0110 - 0X4E16	世

E7	1110 0111
	Implica que son tres bytes
E7	<b>1110</b> 0111
95	<b>10</b> 01 0101
86	<b>10</b> 00 1100
0111 0101 0100 1100- 0X754C	界

Unicode	Carácter	Byte 1	Byte 2	Byte 3	Byte 4
U+4F60	你	<b>1110</b> 0100	<b>10</b> 11 1101	<b>10</b> 10 0000	
U+597D	䥽	<b>1110</b> 0101	<b>10</b> 10 0101	<b>10</b> 11 1101	
U+4E16	世	<b>1110</b> 0100	<b>10</b> 11 1000	<b>10</b> 01 0110	
U+754C	界	<b>1110</b> 0111	<b>10</b> 01 0101	<b>10</b> 00 1100	

UTF-8 Resto de planos (pares subrogados) - 010000-10FFFF

Las direcciones van entre 0x 0000

Dec	Hex	UTF-8 Hex	Char	Unicode description
74773	U+12415	F0 92 90 95	7	<b>Cuneiform Numeric Sign One Gesh2</b>
74767	U+1240F	F0 92 90 8F	<b>*</b>	Cuneiform Numeric Sign Four U

F0	<b>1111 0</b> 000
	Implica que son cuatro bytes
FO	<b>1111 0</b> 000
92	<b>10</b> 01 0010
90	<b>10</b> 01 0000
95	<b>10</b> 01 0101
0 0001 0010 0100 0001 0101 0x 12415	Cuneiform Numeric Sign One Gesh2

F0	<b>1111 0</b> 000
92	<b>10</b> 01 0010
90	<b>10</b> 01 0000
8f	<b>10</b> 00 1111
0 0001 0010 0100 0000 1111 0x 1240F	<b>Cuneiform Numeric Sign Four U</b>