

CS208

Wed, 27 Sep 2023

UTF-8 example

U+00E9 é

2nd row of UTF-8 chart

U+0080 - U+07FF	110xxxxx	10xxxxxx
-----------------	----------	----------

U+00E9 0000 0000 1110 1001



α = U+03B1

α , $u+03B1$ / Row 2 of A
Chart

0000 0011 1011 0001

110 01110

CE

10 110001

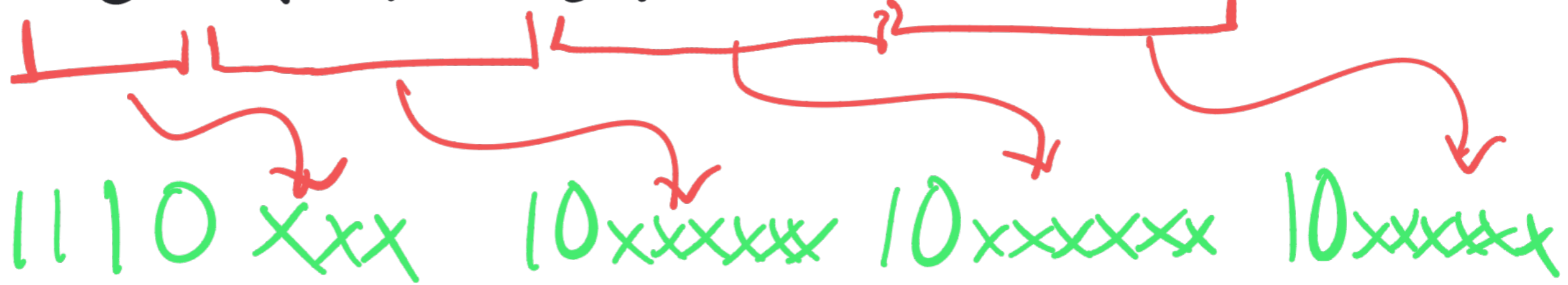
BI

Utf 1F600

Row 4
of chart

Row 4

0001 1111 0110 0000 0000



11110000

1001111

10011000

10000001

0xF0

9F

98

80

to_utf(buff, 0x0351);

buff[0] = 0xC0;

110 000000

buff[0] = buff[0] |

middle_bits(0, 0, 0)

buff[1] = 0x80;

110 01110

10 000000

buff[1] = buff[1] | middle_bits(0, 0, 0);

Printing bytes

in C

in hex

