

CS 208

Mon. 9 Jan 2023

Character "code points"

A — 65 = 0x41

α — 0x03B1

U+03B1

Encodings

How do I store a
Codepoint in a File or
in memory?

0x03B1

0000 0011

byte
"high-order"

1011 0001

byte
"low-order"

UTF-16 LE

all codepoints
encoded as
2-byte
sequences
(ie. 16 bits)

↳ "little
endian"
byte
order

19 00 00 00 ← little endian

00 00 00 19 ← big endian

16^1 's 1's

$$1 \times 16 + 9 = 25_{\text{ten}}$$

00 00 0 | 00 ← big

256^1 's
 16^2

16's 1's

$$256 + 0 + 0 = 256$$

Wikipedia (or "tutorial")
Two's Complement
integers