

Question 1

Code point to binary : U+1F974

0x1 = 0001

F = 1111

9 = 1001

7 = 0111

4 = 0100

→ 0001 1111 1001 0111 0100

∴ 1 1111 1001 0111 0100

Since U+1F974 > 0x10000, it uses 4 bytes (UTF-8 pattern)

Pattern : 11110xxx 10xxxxxx 10xxxxxx 10xxxxxx

To make it to 21 bits, add 4 zeros in front:

→ 0000 1111110 0101110 100

Group to 3, 6, 6, 6 bits format:

→ 000 011111 100101 110100

Final conversion

Byte 1 : 11110 000 → 0xF0

Byte 2 : 10 011111 → 0x9F

Byte 3 : 10 100101 → 0xA5

Byte 4 : 10 110100 → 0xB4

∴ F0 9F A5 B4