

CS Basics - Exercises

Encoding of Strings

E. Benoist

Fall Term 2022-23

1 ASCII

For reference, you can use the table provided by Wikipedia: <https://en.wikipedia.org/wiki/ASCII>.

Exercise 1. What are the strings that corresponds to the following list of bytes (encoded in ASCII).

- 48 65 6c 6c 6f 20 57 6f 72 6c 64
- 4c65 2070 7269 7820 6573 7420 3a20 4348
4620 3132 332c 2d0a

2 Latin 1

For reference, you can use the table provided by Wikipedia: https://en.wikipedia.org/wiki/ISO/CEI_8859-1.

Write the strings corresponding to the following bytes written in hexadecimal :

Exercise 2.

44 e9 6a e0 20 61 72 72 69 76 e9 20 e0 20 42 e2
6c 65 21 0a

48 f6 69 20 7a e4 6d 65 2c 20 67 f6 74 73 3f 0a

Encode the following strings using Latin 1 encoding:

- 50 étudiants sont à l'école!
- Die Häuser sind schön und hübsch.

3 UTF8 encoding

For reference, you may use `U0000.pdf`, `U0080.pdf` and `U4E00.pdf`, which can be found in the `utf-8` directory.

Exercise 3. Write the following sentences in UTF-8 encoding as hexadecimal values. Do not forget spaces and other special characters.

1. “a”, “à”, “ä”, “ü”, chinese character U+4E01 (looks like a T).
2. Student
3. Biel/Bienne et/und Bern(e)
4. 50 étudiants sont à l'école!
5. Die Häuser sind schön und hübsch.
6. Die Schüler sind tüchtig. Sie können UTF-8!

Exercise 4. Decode the following UTF-8 encoded sentences (bytes are shown in hexadecimal notation):

1. 4C 27 C3 A9 63 6F 6C 65 20 C3 A0 20 42 69 65 6E
6E 65 20 63 27 65 73 74 20 73 75 70 65 72 21 0A
2. 53 69 65 20 6D C3 BC 73 73 65 6E 20 55 54 46
2D 38 20 6b C3 B6 6E 6E 65 6E 2E 0A