Problem Sheet #6

Problem 6.1: *unicode and utf-8 encoding*   (2+2 = 4 points)

The content of a file containing UTF-8 Unicode encoded text is given by the following sequence of bytes in hexadecimal notation:

```
f0 9f 98 8a 20 2b 20 f0 9f 95 b6 20 3d 20 f0 9f 98 8e 0a
```

a) Write each byte in binary notation.

b) Identify the unicode code points of the characters. What is the text stored in the file?

Problem 6.2: *valid utf-8 encodings and unicode code points*   (1+1+1+1 = 4 points)

Given the following byte sequences, determine whether they are valid UTF-8 encodings. Explain your reasoning.

a) 0xC1 0x8F

b) 0xE4 0x68 0xA2

c) 0xF4 0x8F 0xBF 0xBF

d) 0xF5 0x80 0x80 0x80

Problem 6.3: *time zones*   (1+1 = 2 points)

The ISO 8601 format for date and time values supports numeric time zone offsets. The world, however, is a bit more complex since we have divided the world into regions using a common time zone. In other words, a time zone is a region using the same local time. Hoever, both the region and the numeric offset to UTC time are not static and vary over time. A simple example are European countries using summer and hence the numeric offset to UTC time changes twice per year.

We meanwhile often identify a time zone by prominent locations in a certain area. For example, you might set your computer clock to be consistent with `Europe/Stockholm` if you live in Sweden. The Internet Assignet Numbers Authority (IANA) maintains a database with time zone information, which essentially consists of complex rules capturing the history of time zones of the world (and the comments in the files make for an interesting read).

RFC 9557 introduces an extended date and time format allowing us to append a time zone name to an ISO 8601 date and time format. This enables programs to correctly calculate times for events in the past according to the time zone offset that was effective in the past.

a) Sort the following date and time values into ascending order.

- `2024-10-11T15:38:29+02:00`
- `2024-10-11T15:38:29-01:15`
- `2024-10-11T01:38:29+12:00`
- `2024-10-11T14:38:29-02:00`

b) What are the extended RFC 9557 date and time values including time zone names for the following events?

- Time in Los Angeles when the Super Bowl LVI was kicked off.
- Time in Stockholm when Neil Amstrong stepped on the moon.
- Time in New York when Lady Diana had a car accident in Paris.
- Time in Beijing when the distributed denial of service attack on Dyn started in 2016.