



Please solve the following exercises at home prior to the tutorial:

Exercise 1

Write the following numbers as 32-bit floating point numbers according to IEEE specification. Provide the result in binary and hexadecimal format (i.e., the 32 bits as stored in memory):

- a) 0
- b) -4278
- c) $-42.15625 \cdot 10^3$

Exercise 2

Find out, how a denormalized number is stored according to IEEE specification. What is a use case for denormalized numbers?

Exercise 3

This exercise uses **decimal** floating point numbers for simplification with 4-digit precision.

- a) Transform the numbers $a' = 1234$ and $b' = 0.0001$ into the normalized notation a and b (decimal).
- b) Calculate the sum $a+b$. What steps are necessary?
- c) Calculate $a \cdot b$. What are the necessary steps for multiplication?

The following exercises will be done during the tutorial:

Exercise 4

Given are two IEEE single-precision floating point numbers $a = 3D800000_{16}$ and $b = 40000000_{16}$.

- a) Calculate the sum $a + b$
- b) Calculate the product $a \cdot b$

Give results in binary representation.

Exercise 5

- a) Calculate $(1690 + 3) + 4$ using **decimal** floating point arithmetic with 3-digit precision.
- b) Calculate $(3 + 4) + 1690$ using **decimal** floating point arithmetic with 3-digit precision.

Note: According to IEEE specification, numbers are rounded, not truncated.