

List 1 report

Albert Kołodziejski

October 21, 2023

Exercise 1

Results:

	Float16	Float32	Float64
eps()	0.000977	1.1920929e-7	2.220446049250313e-16
my_eps()	0.000977	1.1920929e-7	2.220446049250313e-16
float.h	-	1.19209e-07	2.22045e-16
nextfloat()	6.0e-8	1.0e-45	5.0e-324
my_eta()	6.0e-8	1.0e-45	5.0e-324
floatmax()	6.55e4	3.4028235e38	1.7976931348623157e308
my_max()	6.55e4	3.4028235e38	1.7976931348623157e308
float.h	-	3.40282e+38	1.79769e+308

QA:

How macheps relate to precision of arithmetic?

Precision of arithmetic is a upper bound of realative error,

$\text{eps} = 2^{-t}$.

Macheps is distance to next bigger number representet in that arytmetic.

$\text{macheps} = 2^{-(t-1)}$

$$\text{macheps} = 2^{-(t-1)} = 2^{-t+1} = 2^{-t} * 2 = \text{eps} * 2$$

What is the relationship between the number eta and the number MIN_{sub} ?

MIN_{sub} is smallest subnormal number that can be represented. Subnormal means that it mantis starts with 0 instead of 1. Eta is next number after zero. Both are the same. In my results it is hard to see becouse julia rounds those number, if we would take bits of both number we would see that they are the same.

What does the function `floatmin()` return and what is the relationship of with MIN_{nor} ?

`floatmin()` returns minimal normalaized nummber, so it is equal to MIN_{nor} .

Exercise 2

Results:

	Float16	Float32	Float64
experiment()	-0.000977	1.1920929e-7	-2.220446049250313e-16
eps()	0.000977	1.1920929e-7	2.220446049250313e-16

Conclusions:

We can get epsilon from this formula, but if numbers of bit used for mantis is odd we will get negativ value.

Exercise 3

Results:

[illegible]

Adding 1 to the end of mantis should create all numbers between 1 and 2

QA:

How numbers are distributed in $[0.5, 1]$ and how can be represented?

They are distributed evenly with step = 2^{-53} , they can be represented as $x = 0.5 + k \cdot \text{step}$, where $k = 1, 2, \dots, 2^{51} - 1$

How numbers are distributed in $[2, 4]$ and how can be represented?

They are distributed evenly with step = 2^{51} , they can be represented as $x = 2 + k \cdot \text{step}$, where $k = 1, 2, \dots, 2^{51} - 1$