Otto-von-Guericke-University Magdeburg
Max Planck Institute for Dynamics of Complex Technical Systems
Computational Methods in Systems and Control Theory

Dr. Jens Saak, Dipl.-Math. Martin Köhler
Website: http://www.mpi-magdeburg.mpg.de/2802004/sc1/

Winterterm 2014/2015

Scientific Computing 1
3rd Homework

Handout: 11/12/2014 **Return:** 11/19/2014

Make sure you follow the basic rule:

"When reading the code in about six months and asking yourself: who wrote this crap? The answer should not be: YOU!"

Basically that means:

- Try to always use meaningful names for functions, variables, ...
- Write documentation wherever necessary.
- Use indentation to increase readability of the code.
- Add a short statement describing its purpose and basic behavior to each function.
- ...

Exercise 1: (4 Points)

a.) Convert $(1011.101)_2$ and $(0.011111\ldots)_2$ to the decimal system.

b.) Convert $(1CBA)_{16}$ and $(C2D2.E3)_{16}$ into the binary and the decimal system.

c.) Convert $(131)_{10}$ and $(0.3)_{10}$ into the hexadecimal system.

d.) Convert $(763)_{10}$ and $(101101)_2$ into the octal system (base 8).

Exercise 2: (3 Points)

Proof that the grouping of 4 digits in the binary-system to one digit in the hexadecimal-system is correct.

Exercise 3: (6 Points)

a.) Draw all positive numbers that can be expressed using $\mathbb{M}(2, 3, -1, 3)$ on a linearly scaled number ray and on a logarithmically scaled number ray. What difference between the linearly and the logarithmically scaled plot do you recognize?

b.) Given is an arbitrary machine number set

$$\mathbb{M} := \mathbb{M}(p, t, e_{\min}, e_{\max}) := \{ \pm 0.\alpha_1\alpha_2 \dots \alpha_t \cdot p^b \quad | \alpha_i \in \{0, \dots, p-1\}, \alpha_1 \neq 0, \\ e_{\min} \leq b \leq e_{\max}\} \cup \{0\}\,.$$

Proof that for two neighboring powers p^b and p^{b+1} ($e_{min} \leq b < b+1 \leq e_{max}$) the number of representable elements in $\mathbb{M} \cap [p^b, p^{b+1}]$ is the same independent of the choice of b. What can you say about the relative length of two neighboring intervals of this type?

Exercise 4: (5 Points)

Write a C program which prints all numbers that are contained in a given $\mathbb{M}(p, t, e_{min}, e_{max})$. Use the output to plot the members of $\mathbb{M}(2, 4, -2, 4)$ and $\mathbb{M}(3, 2, -1, 2)$ on the number ray. (This can be done with an arbitrary tool, e.g., MATLAB®, gnuplot, TikZ, or by hand.)

Exercise 5: (5 Points)

Write a C function which determines the machine epsilon in double precision. Compile the program using the following compiler options:

- without any extra option (`gcc -o outfile -i input.c`),
- using *SSE2* optimizations (`gcc -mfpmath=sse -msse2 -o output input.c`),
- using the "fast-math" option (`gcc -ffast-math -o output input.c`),
- using the "float-store" option (`gcc -ffloat-store -o output input.c`),
- and second level optimizations (`gcc -O2 -o output input.c`).

What do you recognize? What are possible reasons for this behavior? If you do not use the virtual machine for this exercise please denote some details about your machine (32/64bit, Operating System, Compiler).
Extra (2 Points): Can you design an algorithm which is invariant under such compiler optimizations?

Exercise 6: (2 Points)

You will get a C program from the previous exercise via e-mail. Take a look at it and comment it. Think about:

- Is the code readable or well formed?
- Is the purpose obvious?
- Are unclear statements documented?
- Are function and variable names meaningful?
- Are there parts which can be implemented better or more efficiently?
- ...

Overall Points: 25

In the Tutorial

The following exercises are not part of the homework and will be solved in the tutorial. However, if you solve them successfully during the homework, you can earn some extra points.

Exercise 7:

Determine the absolute and the relative error of 0.5403023059 and π in

a.) $\mathbb{M}(10, 3, -2, 2)$

b.) $\mathbb{M}(2, 3, -2, 3)$

c.) $\mathbb{M}(2, 5, -2, 2)$

Exercise 8:

Compute the smaller one of the two solutions to the quadratic equation

$$x^2 - 1.8x + 0.0001, \quad (a = 0.9,\, b = 0.0001)$$

using

a.) $x_1 = a - \sqrt{a^2 - b}$,

b.) $x_2 = a + \sqrt{a^2 - b}, \quad x_1 = \frac{b}{x_2}$.

Solve the equation in $\mathbb{M}(10, 3, -\infty, \infty)$ and $\mathbb{M}(10, 4, -\infty, \infty)$ and determine the relative error.