Ćwiczenia

Jakub Kowal

November 10, 2025

1 Lista 1

1.5

33-bitowe słowo $x = sm2^c$
cecha 8 bitów ze znakiem
mantysa 24 bity $\in [\frac{1}{2}, 1)$

1.5.a)

$x_{max} = (0.11...1)_2 * 2^{127} = (1 - 2^{-24}) * 2^{127} = 1,7 * 10^{38}$
$x_{min} = (0.10...0)_2 * 2^{-127} = 2^{-1} * 2^{-127} = 2^{-128}$
$[-x_{max}, -x_{min}] \cup [x_{min}, x_{max}]$

1.5.b)

$(-x_{min}, x_{min})$

1.5.c)

$\epsilon = \frac{1}{2} * \beta^{1-t} = 2^{-1} * 2^{-23} = 2^{-24}$

Single

1.5.a)

$x_{max} = (2 * 2^{-23}) * 2^{127} \approx 3,4 * 10^{38}$
$x_{min} = 1 * 2{-}126 = 2^{-126}$
$x_{minsub} = (0.00.....01)_2 * 2^{-126} = 2^{-23} * 2^{-126} = 2^{-149}$
$[-x_{max}, -x_{min}] \cup \{0\} \cup [x_{min}, x_{max}]$

1.5.b)

$(-x_{minsub}, x_{minsub})$

1.5.c)

$\epsilon = \frac{1}{2}\beta^{1-t} = 2^{-1} * 2^{-23} = 2^{-24}$

1.6

1.6.a)

x $= 1 + 2^{-24}$
$\epsilon = 2^{-23}$
$x^{-} = 1$
$x^{+} = 1 + 2^{-23}$

1.6.b)

$x \bigoplus 1 = 1$
$x \in (0, 2^{-24} + 2^{-25}]$

1.6.c)

$x \bigoplus = x$
$x < \frac{x}{2^{23}}$

1.7

$A(a_1, ..., a_n) = (...(a_1 \oplus a_2) \oplus a_3) \oplus ...) \oplus a_n = (...(a_1 + a_2)(1 + \delta_1) + a_3)(1 + \delta_2) + ...) + a_n)(1 + \delta_{n-1}) = a_1 + a_1 * E_1 + a_2 + a_2 * E_1 + a_3 + a_3 * E_2 ... + a_n + a_n * E_{n-1} = (a_1 + a_2 + ... + a_n) + (a_1 * E_1 + a_2 * E_1 + ... + a_n * E_{n-1}) = S + (a_1 * E_1 + a_2 * E_1 + ... + a_n * E_{n-1})$
$|\delta| \leq \epsilon$
$1 + E_k = \prod_{i=1}^{k}(1 + \delta_k) \leftarrow wprowadzamy$
$(a_1 * E_1 + a_2 * E_1 + ... + a_n * E_{n-1}) = E_{max} = \prod_{i=1}^{n-1}(1 + |\delta_i|) - 1 \leq \prod_{i=1}^{n-1}(1 + \epsilon) - 1$
$a_1 * E_1 + a_2 * E_1 + ... + a_n * E_{n-1} \leq (a_1 + a_2 + ... + a_n) * E_{max} \leq ((a_1 + a_2 + ... + a_n) * \prod_{i=1}^{n-1}(1 + \epsilon)) - 1 = S(1 + \epsilon)^{n-1} - 1$
$\frac{|\tilde{S} - S|}{|S|} \leq \frac{\not{S} + S * (1 + s)}{|S|} \leftarrow$ Zmazał za szybko
$\tilde{S} \leq S + S * (1 + \epsilon)^{n-1} - 1$

1.8

P(n-1)=$\prod_{i=0}^{n-1} q_i$ — dokładny wynik
Q(n-1)=fl(P(n-1)) —wynik w fl
Q(n-1)=$P(n-1)(1+\sum_{i=1}^{n-1} q_i), \forall i |\delta_i| \leq \epsilon$ n=1 Q(1)=$a_0 \odot a_1 = a_0 a_1 (1+\delta_1) = P(1)(1+\sum_{i=1}^{1} q_i)$
Zachodzi dla 1
Q(n+1)=$a_0 \odot a_1 \odot \cdots \odot a_{n+1}$=Q(n)$\odot a_{n+1} = (Q(n)a_{n+1})(1+\delta_n+1) = P(n)(1+\sum_{i=1}^{n})a_{n+1}(1+\delta_{n+1}) = P(n+1)(1+\sum_{i=1}^{n}\delta_1)(1+\delta_{n+1}) \approx P(n+1)(1+\delta_{n+1}+\sum_{i=1}^{n}\delta_i) = P(n+1)(1+\sum_{i=1}^{n+1}\delta_i)$

|P(n)-Q(n)|=$|P(n)\sum_{i=1}^{n}\delta_i|$

1.9

$A1(a,b) = a^2 - b^2$
$fl(a^2-b^2) = fl(fl(a^2) - fl(b^2)) = (a^2(1+\delta_1) - b^2(1+\delta_2))(1+\delta_3) \approx a^2 - b^2 + (a^2-b^2)\delta_3 + a^2\delta_1 - b^2\delta_2$
Błąd względny:
$\frac{|a^2-b^2+(a^2-b^2)\delta_3+a^2\delta_1-b^2\delta_2-(a^2-b^2)|}{|a^2-b^2|} = \frac{|(a^2-b^2)\delta_3+a^2\delta_1-b^2\delta_2|}{|a^2-b^2|} \leq |\delta_3| + \frac{|a^2\delta_1-b^2\delta_2|}{|a^2-b^2|} \leq |\delta_3| + \frac{a^2|\delta_1|-b^2|\delta_2|}{|a^2-b^2|} \leq \epsilon + \frac{a^2\epsilon+b^2\epsilon}{|a^2-b^2|} = \epsilon(1+\frac{a^2+b^2}{|a^2-b^2|})$
$A2(a,b) = (a-b)(a+b)$
$((a-b)(1+\delta_1)(a+b)(1+\delta_2))(1+\delta_3) = (a^2-b^2)(1+\delta_1)(1+\delta_2)(1+\delta_3) \approx (a^2-b^2)(1+\delta_1+\delta_2+\delta_3)$
Błąd względny:
$\frac{|(a^2-b^2)(1+\delta_1+\delta_2+\delta_3)-(a^2-b^2)|}{|a^2-b^2|} = |\delta_1+\delta_2+\delta_3| \leq |\delta_1|+|\delta_2|+|\delta_3| \leq 3\epsilon$

2 Lista 2

2.1

$y = \sqrt{x^2+1} - 1$
TW z wykładu:
Jeśli x,y - dodatnie liczby w dwójkowej arytmetyce float, takie że:
$x > y, 2^{-q} \leq 1 - \frac{y}{x}$
to przy odejmowaniu tracimy najwyżej q bitów q = 2
$2^{-2} \leq 1 - \frac{1}{\sqrt{x^2+1}}$
$\frac{1}{4} = 1 - \frac{1}{\sqrt{x^2+1}}$
$\frac{1}{\sqrt{x^2+1}} \leq \frac{3}{4}$
$\sqrt{x^2+1} >= \frac{3}{4}$

2.2

Wskazówka:
Twierdzenie z 1 zadania: $2^{-q} <= 1 - \frac{a}{b} <= 2^{-P} \wedge b - a, b > a$
$x = \frac{1}{2}$
$\cos \frac{1}{2} = 0,8775825$
$2^{-q} <= 1 - \frac{0,8775825}{1} <= 2^{-P}$
$2^{-q} <= 0,1224175 <= 2^{-P}$
$q = 4 \wedge P = 3$

2.3

f(x)=$x^{-1}(1 - \cos x)$
a) $\lim_{x \to 0} \frac{1-\cos x}{x} (De'Hospital) = \lim_{x \to 0} \frac{\sin x}{1} = 0$
b) $\cos x \approx 1$
$x \approx 2\pi k, k \in Z$
$1 - \cos x = 2 \sin^2 \frac{x}{2}$
f(x)=$\frac{2 \sin^2 \frac{x}{2}}{x} = (t = \frac{x}{2}) = \frac{\cancel{2} \sin^2 t}{\cancel{2} t} = t (\frac{\sin t}{t})^2 = \frac{x}{2} (\frac{\sin \frac{x}{2}}{\frac{x}{2}})^2$
$\lim_{x \to 0} \frac{\sin t}{t} = 1$ dla $x \to 0 (t \to 0)$ mamy $f(t) = t 1^2 = t = \frac{x}{2}$

2.4

$f(x) = \sqrt{x+2} - \sqrt{x}$
Problem dla dużych x : $\sqrt{x} \approx \sqrt{x+2}$
$\sqrt{x+2} - \sqrt{x} / * \frac{\sqrt{x+2} + \sqrt{x}}{\sqrt{x+2} + \sqrt{x}}$
$\frac{x+2-x}{\sqrt{x+2} + \sqrt{x}} = \frac{2}{\sqrt{x+2} + \sqrt{x}}$
$a^2 - b^2 = (a - b)(a + b)$

2.6

$4ac \to 0$
x=$\frac{-b \mp \sqrt{b^2 - 4ac}}{2a}$
$x_1 x_2 = \frac{c}{a}$ $q = \frac{1}{2a}(-b - sqm?(b)\sqrt{b^2 - 4ac})$
$x_1 = \frac{-b + \sqrt{b^2 - 4ac}}{2a}$
$x_2 = \frac{-b - \sqrt{b^2 - 4ac}}{2a} = \frac{-(b + \sqrt{b^2 - 4ac})}{2a}$
$x_1 = \frac{c}{a x_2}$

$$b = 0 a x^2 + c = 0$$
$$\Delta = -4ac$$

$b < 0$
$x_1 = \frac{|b| + \sqrt{b^2 - 4ac}}{2a}$
$x_2 = \frac{c}{2a}$

2.8

$\frac{|f(\hat{x})-f(x)|}{|f(x)|} = \frac{|f(x+x\delta)-f(x)|}{|f(x)|} \approx \frac{|f`(x)x\delta|}{|f(x)|} = cond(x)|\delta|$
$x^\alpha (x^\alpha)` = \alpha x^{\alpha-1} \frac{|(\alpha x^{alpha-1})x\delta|}{|x^\alpha|} = \frac{|\alpha x^\alpha \delta|}{|x^\alpha|} = \alpha\delta$
$\sin x (\sin x)` = \cos x \frac{|(\cos x)x\delta|}{|\sin x|} = \frac{x\delta}{\tan x}$
$e^x (e^x)` = e^x$
$x^{-1}e^x (x^{-1}e^x)` = x^{-1}e^x + (-1)x^{-2}e^x = \frac{e^x}{x} - \frac{e^x}{x^2}$
$\arcsin x$