

Arvusüsteemid (3)

- Väärtus = numbrite kaalutud summa
 - $173,42 = 1 \cdot 10^2 + 7 \cdot 10^1 + 3 \cdot 10^0 + 4 \cdot 10^{-1} + 2 \cdot 10^{-2}$
 - Täisosa / murdosa
- Üldistatud esitusviis
 - $D = \sum d_i \cdot r^i, \quad i = -n, \dots, p-1$
 - $d_{p-1} d_{p-2} \dots d_2 d_1 d_0, d_{-1} d_{-2} \dots d_{-n} / r$ - baas (radix)
- John F. Wakerly, Digital design: principles and practices. Pearson/Prentice Hall.
- Vt ka Harri Lenseni materjale (Moodle's)

Püsi- ja ujukomaarvud

- Täisarvud: $-2^n \dots 2^{n-1}$ (2^n-2^0)
- Püsikomaarvud: $-2^n \dots 2^{n-2^{-f}}$
- Ujukomaarvud
 - eksponent: täisarv
 - mantiss: normaliseeritud püsikomaarv
- Põhitehted teostatavad täisarvude vahel
 - koma asukoha korrigeerimine:
 - püsikomaarvud – korrutamisel ja jagamisel
 - ujukomaarvud – normaliseerimine

Täis- / püsikoma- / ujukomaarv

- Täisarv
 - 1+15 bitti \approx -32000 kuni +32000, täpsus 1
 - + lihtsad operatsioonid
 - puudub murdosa
- Püsikomaarv
 - 1+5+10 bitti \approx -32 kuni +32, täpsus ~ 0.001 ($\sim 0.03\%$)
 - + lihtsad liitmine ja lahutamine
 - normaliseerimine vajalik korrutamiseks ja jagamiseks (nihutamine)
- Ujukomaarv
 - 1+5+10 bitti \approx -64000 kuni +64000, täpsus $\sim 0.1\%$
 - + paindlik vahemik
 - normaliseerimine vajalik (analüüs + nihutamine)

Püsikoma-arvud – leidmine

- $5,125_{10} \rightarrow 101,001_2$
 - Täisosa: $5 = ((1) \cdot 2 + 0) \cdot 2 + 1 = 101_2$
 - $5 / 2 = 2$ jääk 1
 - $2 / 2 = 1$ jääk 0
 - Murdosa: $0,125 = (((1)/2 + 0)/2 + 0)/2 = 0,001_2$
 - $0,125 \cdot 2 = 0,25$ – täisosa 0; murdosa 0,25
 - $0,25 \cdot 2 = 0,5$ – täisosa 0; murdosa 0,5
 - $0,5 \cdot 2 = 1$ – täisosa 1; murdosa 0
- $101,001_2 \rightarrow 5,125_{10}$
 - $1 \cdot 2^2 + 0 \cdot 2^1 + 1 \cdot 2^0 + 0 \cdot 2^{-1} + 0 \cdot 2^{-2} + 1 \cdot 2^{-3} = 5,125_{10}$
 - Täisosa: $101_2 / 1010_2 = 0_2$ jääk 101_2 😊

Püsikoma-arvud – ümardamine

- 10-arvud – 0,026 \rightarrow 0,03 / 0,024 \rightarrow 0,02
- 2-arvud – piiriks $0,0\dots5_{10}$ asemel $0,0\dots1_2$
- Mõjutab teisenduste täpsust!
 - $2,6_{10} = 10,10011001100110\dots_2$
 - $10,10011001_2 = 2,59765625_{10}$
 - $10,10011010_2 = 2,6015625_{10}$
- Negatiivsed arvud?
 - $-2,6_{10} = 11101,01100110011001\dots_2$
 - $11101,01100110_2 \rightarrow -10,10011010_2 = -2,601562_{10}$
 - $11101,01100111_2 \rightarrow -10,10011001_2 = -2,59765625_{10}$

Näited – 1+3+6 bitti

2,7 [2==10]

$$0,7 \cdot 2 = \underline{1},4$$

$$0,4 \cdot 2 = \underline{0},8$$

$$0,8 \cdot 2 = \underline{1},6$$

$$0,6 \cdot 2 = \underline{1},2$$

$$0,2 \cdot 2 = \underline{0},4$$

$$0,4 \cdot 2 = \underline{0},8$$

$$0,8 \cdot 2 = \underline{1},6$$

0010,1011001

1,9 [1==01]

$$0,9 \cdot 2 = \underline{1},8$$

$$0,8 \cdot 2 = \underline{1},6$$

$$0,6 \cdot 2 = \underline{1},2$$

$$0,2 \cdot 2 = \underline{0},4$$

$$0,4 \cdot 2 = \underline{0},8$$

$$0,8 \cdot 2 = \underline{1},6$$

$$0,6 \cdot 2 = \underline{1},2$$

0001,1110011

1,3 [1==01]

$$0,3 \cdot 2 = \underline{0},6$$

$$0,6 \cdot 2 = \underline{1},2$$

$$0,2 \cdot 2 = \underline{0},4$$

$$0,4 \cdot 2 = \underline{0},8$$

$$0,8 \cdot 2 = \underline{1},6$$

$$0,6 \cdot 2 = \underline{1},2$$

$$0,2 \cdot 2 = \underline{0},4$$

0001,0100110

- Ümardamine & positiivne / negatiivne

0010,101101

1101,010011

0001,111010

1110,000110

0001,010011

1110,101101

Püsikoma-arvud – operatsioonid

- n bitti, m kohta pärast koma $(2^{n-m} \cdot 2^{-m}) \approx N/2^m$
 - $0,625_{10} = 1/2 + 1/8 = 0000,10100000_2 = 160 / 256$
 - $2,7_{10} = 0010,101101_2 = 173/64$
- Liitmine & lahutamine
 - $a+b = (A/2^m) + (B/2^m) = (A+B)/2^m$ – OK
- Korrutamine
 - $a*b = (A/2^m) * (B/2^m) = (A*B)/2^{2m}$ – “liiga väike”
 - n & n bitti $\rightarrow 2n$ bitti $\rightarrow m$ bitti sabast maha...
- Jagamine
 - $a/b = (A/2^m) / (B/2^m) = (A/B)/2^0$ – “liiga suur”

Püsikoma – liitmine & lahutamine

- $a+b = (A/2^m)+(B/2^m) = (A+B)/2^m$ – OK

$$\begin{array}{r} 2,7 \quad 0010,101101 \\ + 1,9 \quad + 0001,111010 \\ = 4,6 \quad = 0100,100111 \quad [4,609375] \end{array}$$

$$\begin{array}{r} 1,3 \quad 0001,010011 \\ - 1,9 \quad + 1110,000110 \\ = -0,6 \quad = 1111,011001 \quad [-0,609375] \end{array}$$

$$\begin{array}{r} 2,7 \quad 0010,101101 \\ - 1,3 \quad + 1110,101101 \\ = 1,4 \quad = 0001,011010 \quad [1,40625] \end{array}$$

Püsikoma – korrutamine

- $a*b = (A/2^m)*(B/2^m) = (A*B)/2^{2m}$ – “liiga väike”
 - n & n bitti \rightarrow $2n$ bitti \rightarrow m bitti sabast maha...

$1,3*1,9=2,47 \rightarrow 83*122=10126 \rightarrow /64=158 \rightarrow 2,46875$

$01010011 * 01111010 = 010011110001110 \rightarrow$

$010011110|001110 \rightarrow 010,011110 == 2,46875$

$1,3*2,7=3,51 \rightarrow 83*173=14359 \rightarrow /64=224 \rightarrow 3,5$

$01010011*010101101 = 011100000010111 \rightarrow$

$011100000|010111 \rightarrow 011,100000 == 3,5$

$1,9*2,7=5,13 \rightarrow 122*173=21106 \rightarrow /64=329 \rightarrow 5,140625$

$01111010*010101101 = 0101001001110010 \rightarrow$

$0101001001|110010 \rightarrow 0101,001001 == 5,140625$

Püsikoma – jagamine

- $a \cdot b = (A/2^m)/(B/2^m) = (A/B)/2^0$ – “liiga suur”
 - lisabitid arvutamisel?!

$2,7/1,3=2,0769\dots \rightarrow 173/83=2,0843\dots \rightarrow$

$\dots *64=133,3975 \rightarrow 133/64=2,078125$

$010101101 / 01010011 = 10[,000100111\dots] \rightarrow 010,000101$

$5,13/2,7=1,9 \quad (5,125/2,703125=1,8959537572)$

$0101.001000/0010.101101 = 0001.111001 = 121/64 = 1,890625$

$328/173 = 1 ???$

$(328*64)/173 = 20992/173 = 121,341\dots \rightarrow 121/64 = 1,890625$

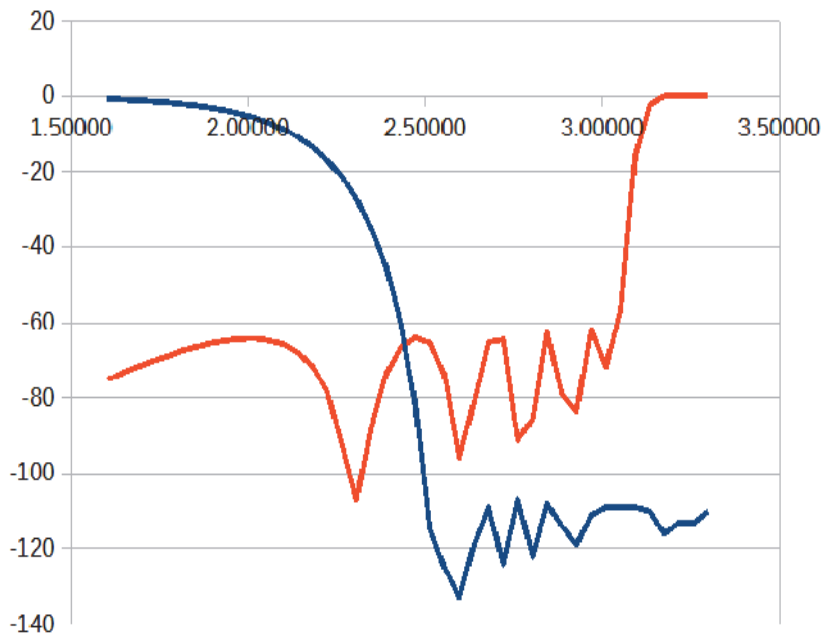
Püsikoma-arvud – täpsus

- Ujukoma (32 bitti) vs püsikoma (N+12 bitti)

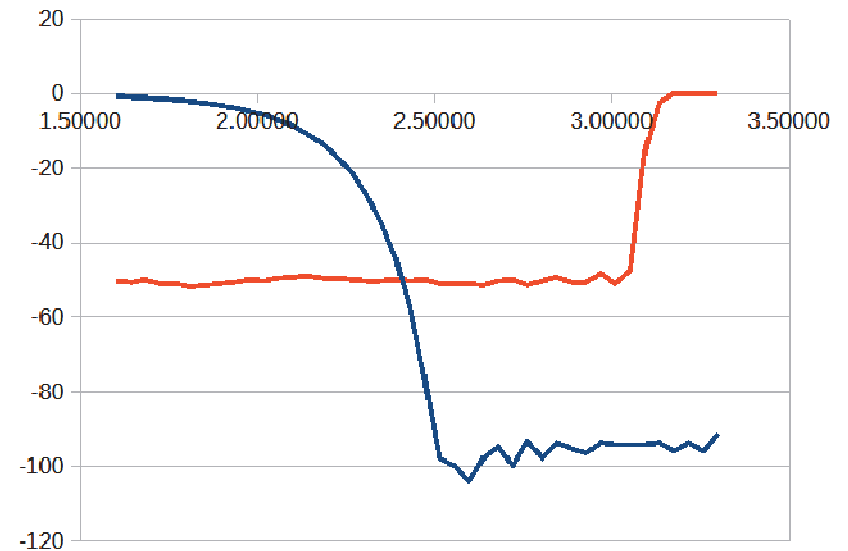
```
for (i=0;i<TAB_SIZE;i++)  
    sum+=coeffs[i]*delayed[i];
```

```
#define FX_FRACT    12  
for (i=0;i<TAB_SIZE;i++)  
    sum+=((coeffs[i]*delayed[i])>>FX_FRACT);
```

Ujukoma



Püsikoma



Ujukoma-arvud – operatsioonid

- Märk 1; eksponent k ja mantiss m bitti
 - $\pm(N/2^m)*2^K$; $1 > N \geq 0.5$ ja $2^{k-1} > K \geq -2^{k-1}$
 - $1,5 = 0,75*2^1 = 0|0001|11000000$
 - Võimalik ka märgiga mantiss ja eksponendi variandid
- Liitmine & lahutamine
 - $a+b = (A*2^{AE})+(B*2^{BE}) = (A*2^{AE})+(B*2^{AE-x}) = (A+B/2^x)*2^{AE}$
 - $AE > BE$ (e. $AE=BE+x$) ja $1 > \text{mantiss} > 0$
 - eel- ja järelkorrektsioonid võivad olla vajalikud!
- Korrutamine & jagamine
 - $a*b = (A*2^{AE})*(B*2^{BE}) = (A*B)*2^{AE+BE}$
- Normaliseerimine enne/pärast tehteid

Ujukoma-arvud – näited

- Märgibitt, astendaja 5, mantiss 10 bitti
- $1,3 = 0,65 \cdot 2^1$
 - $0|00001|1010011001 == 0,64941\dots$
 - $0|00001|1010011010 == 0,65039\dots$
- $1,9 = 0,95 \cdot 2^1$
 - $0|00001|1111001100 == 0,94921\dots$
 - $0|00001|1111001101 == 0,95019\dots$
- $2,7 = 0,675 \cdot 2^2$
 - $0|00010|1010110011 == 0,67480\dots$

Ujukoma-arvud – tehted

- Liitmine / lahutamine

- Astendajad peavad olema võrdsed

- $2,7 - 1,3 \equiv 0,675 \cdot 2^2 - 0,65 \cdot 2^1 =$
 $0,675 \cdot 2^2 - 0,325 \cdot 2^2 = (0,675 - 0,325) \cdot 2^2 =$
 $0,35 \cdot 2^2 = 0,7 \cdot 2^1 \equiv 1,4$

- Korrutamine

- $1,3 \cdot 1,9 \equiv 0,65 \cdot 2^1 \cdot 0,95 \cdot 2^1 = 0,6175 \cdot 2^2 \equiv 2,47$

- Jagamine

- $2,7 / 1,3 \equiv 0,675 \cdot 2^2 / 0,65 \cdot 2^1 =$
 $1,038461... \cdot 2^1 = 0,519230... \cdot 2^2 \equiv 2,076923...$

Ujukoma-arvud – tehted

- Astendaja 5 bitti, mantiss 10 bitti ($1 \leq |M| < 2$)
 - märgibitt == 0, pole näidatud
 - $121 = 1111001 = 1,111001000 * 10^{00110}$
 - $3,1416 = 11,00100100001111... = 1,100100100 * 10^{00001} [3,140625]$
- Liitmine, astendajad peavad olema võrdsed
 - $121 + 3,1416 == 124,1416$
 - $1,111001000 * 10^{00110} = 1,111001000 * 10^{00110}$
 - $1,100100100 * 10^{00001} = 0,000011001 * 10^{00110}$
 - $1,111100001 * 10^{00110} = 124,125$

Ujukoma-arvud – tehted

- Korrutamine

- $121 * 3,1416 == 380,1336$

- $1,111001000 * 10^{00110} * 1,100100100 * 10^{00001} =$
 $10,111110000 * 10^{00111}$ [vajalik normaliseerimine] =
 $1,011111000 * 10^{01000} = 380,0$

- Jagamine

- $121 / 3,1416 == 38,5154061624$

- $1,111001000 * 10^{00110} / 1,100100100 * 10^{00001}$

- $1,001101000 * 10^{00101} == 38,53125$

Ujukoma-arvud – muud formaadid

- Erinevad formaadid ja tehted
 - <https://ati.ttu.ee/~lrv/ujukoma.html>
 - IEEE 754 standard
- Posit
 - IEEE 754 asendus / täiendus
 - <https://www.sigarch.org/posit-a-potential-replacement-for-ieee-754/>
 - [https://en.wikipedia.org/wiki/Unum_\(number_format\)](https://en.wikipedia.org/wiki/Unum_(number_format))
- Ligikaudne arvutamine
 - Approximate computing

IEEE 754 ujukomaformaad

- https://en.wikipedia.org/wiki/IEEE_754
- Märk, astendaja, mantiss
 - Normeeritud mantiss – 1,0...2,0
 - Vanim bitt alati 1 – ei kasutata
 - Astendaja nn kõrvalekaldega (bias)

10-arv	2-arv	$M_2 * 2^{E_2}$	S	$E_2 + 01111_2$	M_2	$[M_{10} * 2^{E_{10}}]$	[10-arv]
2,5	10.1	$1.0100000000 * 2^1$	0	10000	0100000000	$1,25 * 2^1$	2,5
1,1	1.0(0011)	$1.0001100110 * 2^0$	0	01111	0001100110	$1,099609375 * 2^0$	1,099609375
-2,2	-10.(0011)	$-1.0001100110 * 2^1$	1	10000	0001100110	$1,099609375 * 2^1$	-2,1992187500
-0,056	-0.0000111001010110...	$-1.1100101011 * 2^{-101}$	0	01010	1100101011	$-1,7919921875 * 2^{-5}$	-0,055999755859375
1,0	1.0	$1.0 * 2^0$	0	01111	0000000000	$1,0 * 2^0$	1,0
+0,0	0.0		0	00000	0000000000		+0,0
-0,0	-0.0		1	00000	0000000000		-0,0
$+\infty$			0	11111	0000000000		$+\infty$
$-\infty$			1	11111	0000000000		$-\infty$

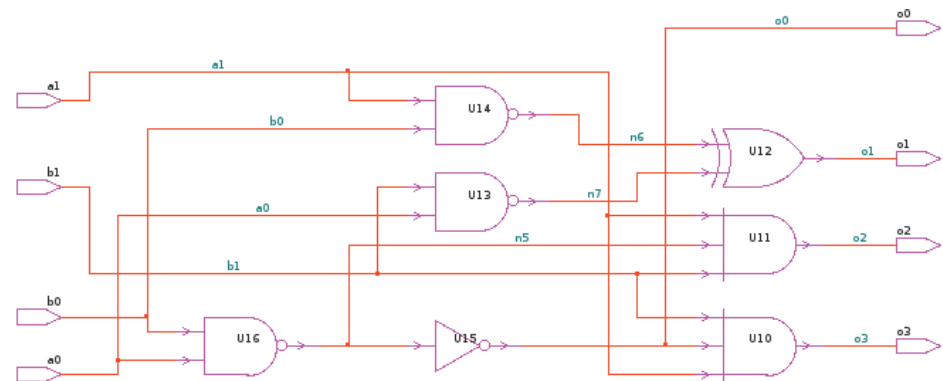
Ligikaudne arvutamine

- Arvutustäpsus polegi nii oluline
 - Närvivõrgud (NN) – 8 bitti täitsa OK 😊
 - Eksponent 3 bitti (-4...+3)
 - Märgiga mantiss 5 bitti ($0,5 \leq |M| < 1$)
 - 2,7 $\sim\sim 0,1011 * 10^010 = 2,75$
 - -2,7 $\sim\sim 1,0101 * 10^010$
 - 5,13 $\sim\sim 0,1010 * 10^011 = 5,0$
 - 0,45 $\sim\sim 0,1110 * 10^111 = 0,4375$
 - <https://en.wikipedia.org/wiki/Minifloat>
 - https://asawicki.info/articles/fp8_tables.php

Ligikaudne arvutamine

- Arvutustäpsus polegi nii oluline
 - Korrutamistel noorimad järgud vähemtähtsad
 - Lihtsustatud korrutaja – 2x2 bitti

A x B	00	01	10	11
00	000	000	000	000
01	000	001	010	011
10	000	010	100	110
11	000	011	110	111



11 eg

vs

8 eg

