Arvusüsteemid (3)

- Väärtus = numbrite kaalutud summa
  - $173{,}42 = 1\cdot10^{2}+7\cdot10^{1}+3\cdot10^{0}+4\cdot10^{-1}+2\cdot10^{-2}$
  - Täisosa / murdosa
- Üldistatud esitusviis
  - $D = \Sigma\, d_i \cdot r^i, \quad i = -n, \ldots, p-1$
    - $d_{p-1}\, d_{p-2} \ldots d_2\, d_1\, d_0\, ,\, d_{-1}\, d_{-2} \ldots d_{-n}$ / 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 ... 2^n-1 (2^n-2^0)
- Püsikomaarvud: -2^n ... 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 ≈ -32000 kuni +32000, täpsus 1
  + lihtsad operatsioonid
  – puudub murdosa
- Püsikomaarv
  - 1+5+10 bitti ≈ -32 kuni +32, täpsus ~0.001 (~0.03%)
  + lihtsad liitmine ja lahutamine
  – normaliseerimine vajalik korrutamiseks ja jagamiseks (nihutamine)
- Ujukomaarv
  - 1+5+10 bitti ≈ -64000 kuni +64000, täpsus ~0.1%
  + paindlik vahemik
  – normaliseerimine vajalik (analüüs + nihutamine)

Püsikoma-arvud – leidmine

- $5{,}125_{10}$ –> $101{,}001_2$
  - 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
- $101{,}001_2$ –> $5{,}125_{10}$
  - $1\cdot 2^2 + 0\cdot 2^1 + 1\cdot 2^0 + 0\cdot 2^{-1} + 0\cdot 2^{-2} + 1\cdot 2^{-3} = 5{,}125_{10}$
  - Täisosa: 101_2 / 1010_2 = 0_2 jääk 101_2 ☺

Püsikoma-arvud – ümardamine

- 10-arvud – 0,026 -> 0,03 / 0,024 -> 0,02
- 2-arvud – piiriks $0,0\ldots5_{10}$ asemel $0,0\ldots1_{2}$
- Mõjutab teisenduste täpsust!
  - $2,6_{10} = 10,10011001100110\ldots_{2}$
  - $10,10011001_{2} = 2,59765625_{10}$
  - $10,10011010_{2} = 2,6015625_{10}$
- Negatiivsed arvud?
  - $-2,6_{10} = 11101,01100110011001\ldots_{2}$
  - $11101,01100110_{2} \rightarrow -10,10011010_{2} = -2,601562_{10}$
  - $11101,01100111_{2} \rightarrow -10,10011001_{2} = -2,59765625_{10}$

Näited – 1+3+6 bitti

2,7 [2==10]	1,9 [1==01]	1,3 [1==01]
0,7·2 = 1,4	0,9·2 = 1,8	0,3·2 = 0,6
0,4·2 = 0,8	0,8·2 = 1,6	0,6·2 = 1,2
0,8·2 = 1,6	0,6·2 = 1,2	0,2·2 = 0,4
0,6·2 = 1,2	0,2·2 = 0,4	0,4·2 = 0,8
0,2·2 = 0,4	0,4·2 = 0,8	0,8·2 = 1,6
0,4·2 = 0,8	0,8·2 = 1,6	0,6·2 = 1,2
0,8·2 = 1,6	0,6·2 = 1,2	0,2·2 = 0,4
0010,101100**1**	0001,111001**1**	0001,010011**0**

- Ümardamine & positiivne / negatiivne

0010,101101	0001,111010	0001,010011
1101,010011	1110,000110	1110,101101

Püsikoma-arvud – operatsioonid

- n bitti, m kohta pärast koma ($2^{n-m}-2^{-m}$) ≈ $N/2^m$
  - $0{,}625_{10}$ = 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 → 2n bitti → 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

```
   2,7      0010,101101
+  1,9    + 0001,111010
=  4,6    = 0100,100111    [4,609375]

   1,3      0001,010011
-  1,9    + 1110,000110
= -0,6    = 1111,011001    [-0,609375]

   2,7      0010,101101
-  1,3    + 1110,101101
=  1,4    = 0001,011010    [1,40625]
```

Püsikoma – korrutamine & jagamine

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

1,3*1,9=2,47 → 83*122=10126 → /64=158 → 2,46875

01010011 * 01111010 = 010011110001110 →

010011110|001110 → 010,011110 == 2,46875

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

2,7/1,3=2,0769... → 173/83=2,0843... →

...*64=133~~,3975~~ → 133/64=2,078125

010101101 / 01010011 =

10[,000100111...] → 010,000101

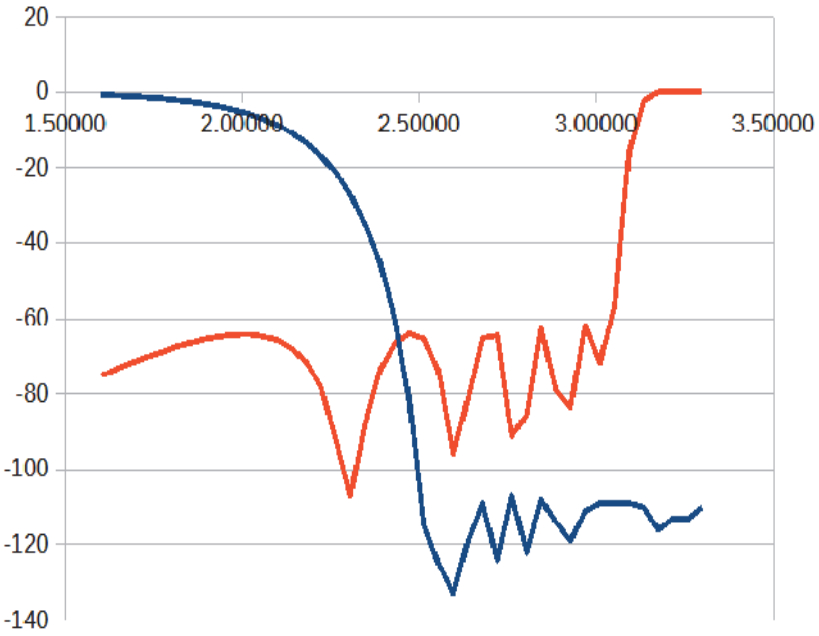
Püsikoma-arvud – täpsus

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

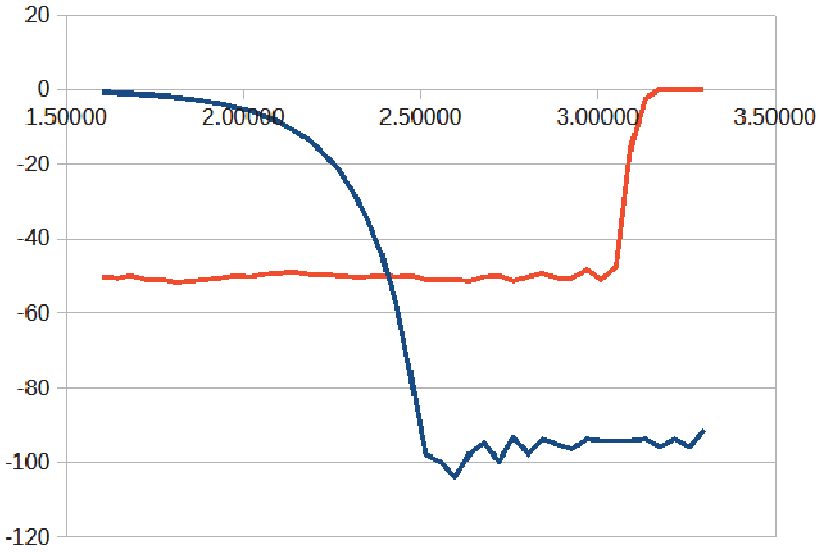
```
for(i=0;i<TAB_SIZE;i++)
  sum+=coeffs[i]*delayed[i];
```

```
#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 > 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...
  - 0|00001|1010011010 == 0,65039...
- $1{,}9 = 0{,}95 \cdot 2^1$
  - 0|00001|1111001100 == 0,94921...
  - 0|00001|1111001101 == 0,95019...
- $2{,}7 = 0{,}675 \cdot 2^2$
  - 0|00010|1010110011 == 0,67480...

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

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)
- Ligikaudne arvutamine
  - Approximate computing

IEEE 754 ujukomaformaat

- 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
+∞			0	11111	0000000000		+∞
-∞			1	11111	0000000000		-∞

Ligikaudne arvutamine

- Arvutustäpsus polegi nii oluline
  - Närvivõrgud (NN) – 8 bitti täitsa OK ☺
  - Korrutamistel noorimad järgud vähemtähtsad

  - Lihtustatud 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**