# Flyttal

mantissa

$$x = \pm \left( d_0 + \frac{d_1}{\beta} + \frac{d_2}{\beta^2} + \dots + \frac{d_{t-1}}{\beta^{t-1}} \right) \cdot \beta^e$$

bas, bas, bas, bas → exponent

$0 \le d_k \le \beta - 1, \quad L \le e \le U$

exponentområde

$d_k, k = 0, \dots, t-1$ – mantissa (heltal)

precision

| bas | $t$ | $L$ | $U$ | precision |
|---|---|---|---|---|
| 2 | 24 | -126 | 127 | 32 bitar (enkel) |
| 2 | 53 | -1022 | 1023 | 64 bitar (dubbel) |

**Enkel precision:** $\pm \; e_1 e_2 \dots e_8 \; d_0 d_1 \dots d_{22}$

1 bit tecknet, 8 bitar exponent, 23 bitar mantissa

$1 + 8 + 23 = 32$ bitar

**Dubbel precision:** $\pm \; e_1 \dots e_{11} \; d_0 \dots d_{51}$

1 bit, 11 bitar exponent, 52 bitar

$1 + 11 + 52 = 64$ bitar

Antalet olika tal som räknas:

$$(*)\quad 2(\beta-1)\beta^{t-1}(U-L+1)+1$$

| Bas | t | L | U | precision |
|---|---|---|---|---|
| 2 | 24 | -126 | 127 | 32 (enkel) |
| 2 | 53 | -1022 | 1023 | 64 (dubbel) |

Exempel: floatgui. m

$\beta = 2$, $p = t = 3$, $L = -1$, $U = 1$

Antalet flyttal (se *):

$$2(2-1)2^{3-1}(1-(-1)+1)+1 =$$

$$= 2 \cdot 1 \cdot 4 \cdot 3 + 1 = 25$$

Exempel: $1 = [1+0] \cdot 2^0$

Exponenten: $0 + 1023 = 1024 - 1 =$

$= 1 \cdot 2^{10} - 1 \cdot 2^0 =$

$$\begin{array}{r} 0\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1 \\ -\ 1\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0 \\ \quad 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 1 \\ \hline 0\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1 \end{array}$$

---

Exempel: $0.5 = 1 \cdot 2^{-1} = [1+0] \cdot 2^{-1}$

Exponenten $(-1)$: $-1 + 1023 = 1024 - 2$

$$\begin{array}{r} 0\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 0 \\ -\ 1\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0 \\ \quad 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 1\ 0 \\ \hline 0\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 0 \end{array}$$

---

Exempel: $0.1 = [1.6] \cdot 2^{-4} = [1+0.6] 2^{-4}$

Exponenten $-4$ lagras som:

$-4 + 1023 = 1019 = 1024 - 5 =$

$= 1 \cdot 2^{10} - (1 \cdot 2^2 + 1 \cdot 2^0) =$

$$\begin{array}{r} 0\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1 \\ -\ 1\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0 \\ \quad 0\ 0\ 0\ 0\ 0\ 0\ 0\ 1\ 0\ 1 \\ \hline 0\ 1\ 1\ 1\ 1\ 1\ 1\ 0\ 1\ 1 \end{array}$$

Exempel:

$1 = [1] \cdot 2^0$

I binär form:

4 bitar 4 4 4 4

0 01111111111 0 ... 0

tecken 11 bitar exp. mantissa 52 bitar

I hexadecimalt form:
grupperar i 4 bitar:

0011 1111 1111 0...0 ... 0...0

1) 2) 3) alla 0

$0011 = 1 \cdot 2^0 + 1 \cdot 2^1 + 0 \cdot 2^2 + 0 \cdot 2^3 =$

$= 1 + 2 = 3$

$1111 = 1 \cdot 2^0 + 1 \cdot 2^1 + 1 \cdot 2^2 + 1 \cdot 2^3 = 1 + 2 +$

$+ 4 + 8 = 15 = F$

$1111 = 15 = F$

0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15

det samma A B C D E F

1 i hexad. form är:

3FF0000000000000

Konvertera från binär till hexadecimalt (bas 16) form:

1010 0111 1011

Svar: grupperar i 4 bitar:

1010 0111 1011

a) b) c)

a) $1010 = 0\cdot 2^0 + 1\cdot 2^1 + 0\cdot 2^2 + 1\cdot 2^3 =$

$1\,0\,1\,0$ under $2^3\ 2^2\ 2^1\ 2^0$

$= 10 = A$

b) $0111 = 1\cdot 2^0 + 1\cdot 2^1 + 1\cdot 2^2 + 0\cdot 2^3 =$

$0\,1\,1\,1$ under $2^3\ 2^2\ 2^1\ 2^0$

$= 7$

c) $1011 = 1\cdot 2^0 + 1\cdot 2^1 + 0\cdot 2^2 + 1\cdot 2^3 =$

$= 11 = B$

**Svar: A 7 B**

Hex.: 0; 1; 2; 3; 4; 5; 6; 7; 8; 9; 10; 11;

det samma A B

12; 13; 14; 15

C D E F

$-3.25 = (-1) \cdot 3.25 = (-1) \cdot [1.625] \cdot 2^{1}$

exponent: $1 + 1023 = 1024 = 2^{10}$

$= (-1) \cdot [1 + 0.625] \cdot 2^{1}$

1 presenteras inte i datorn

$0.625 = \frac{1}{2} x_1 + \frac{1}{4} x_2 + \frac{1}{8} x_3 + \frac{1}{16} x_4 + \ldots =$

$= \frac{1}{2} \cdot 1 + \frac{1}{4} \cdot 0 + \frac{1}{8} \cdot 1 + 0$

Steg1. $0.625 - 0.5 = 0.125$

0.5 < 0.625? YES → SKRIV 1

NO → SKRIV 0

Steg2. 0.25 > 0.125? YES → LÄT SKRIVA 0

NO → STOP

Steg3. $0.125 = \frac{1}{8} = 0.125$?

YES

SKRIV 1

STOP.

$-3.25$ Presenteras så här i datorn:

| 1 | 10000000000 | 10100...0 |
|---|---|---|

tecken "−" — exponent 11 bitar — mantissa 52 bitar

$$-9.28 = (-1) \cdot [1.16] \cdot 2^3 =$$

$$= (-1) \cdot [1 + 0.16] \cdot 2^{(3)}$$

(3) ← exponent

$3 + 1023 = 1026 = 2^{10} + 1 \cdot 2^1 + 0 \cdot 2^0$

← Presenteras inte i dator

$$0.16 = \frac{1}{2} x_1 + \frac{1}{4} x_2 + \frac{1}{8} x_3 + \frac{1}{16} x_4 + \ldots$$

$$= \frac{1}{2} \cdot 0 \cdot 2^0 + 0 \cdot \frac{1}{4} \cdot 2^1 + 1 \cdot \frac{1}{8} \cdot 2^2 + 0 \cdot \frac{1}{16} \cdot 2^3 + \ldots$$

Steg 1. $0.5 > 0.16$ → NO → 0

Steg 2. $0.25 > 0.16$ → YES → 0

Steg 3. $0.125 > 0.16$ → NO

$0.16 - 0.125 = 0.035$

Steg 4. $0.0625 > 0.035$ → YES → 0

Steg 5 ⋮

I dator lagrar $-9.28$ presenteras så i binär form i dator:

| 1 | 10000000010 | 0010 ... |
|---|---|---|

1 bit: tecken "−"

11 bitar: exponent

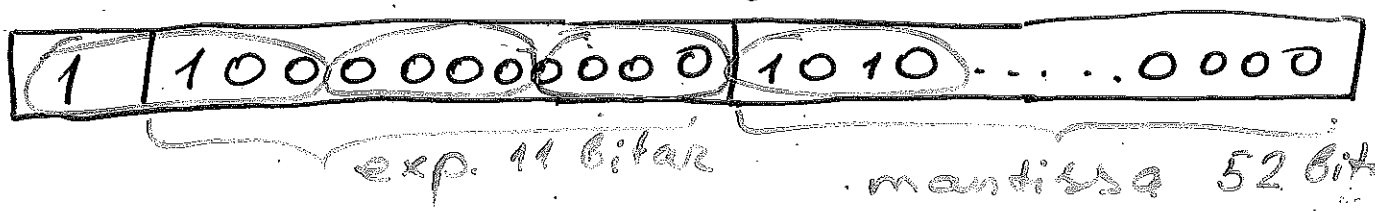
52 bitar: mantissa

Hexadecimalt (bas 16)

vi grupperar fyra bitar av binär form.

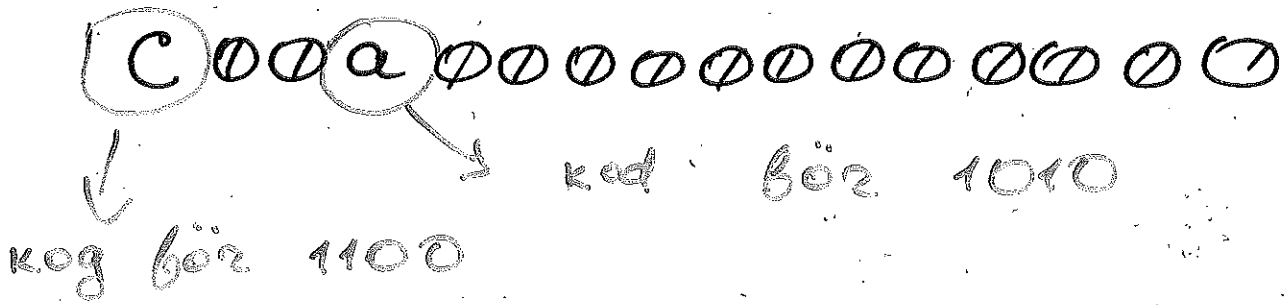
Exempel:

- 3.25 i binär form lagras som




Nu grupperar vi om i 4 bitar:

- 3.25 :




Bas $\beta = 16$ :

0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15

som vanligt

a b c d e f

I exempel:

$1100 = 1 \cdot 2^3 + 1 \cdot 2^2 + 0 \cdot 2^1 + 0 \cdot 2^0 = 12 = c$

$1010 = 1 \cdot 2^3 + 0 \cdot 2^2 + 1 \cdot 2^1 + 0 \cdot 2^0 = 10 = a$