

Flyttal komplettering (32-bitars)

Format: s / c / f
Antal bitar: 1 8 23

s är teckenbit: 0 motsvarar + och 1 motsvarar −.

c är karakteristika: $c = \exp + 127$.

f kallas fraction och är normaliserade mantissan utan inledande 1.

Högsta värde (255) och lägsta värde (0) på karakteristikan är reserverade för ∞ resp. 0.

$\pm \infty$: s / 255 / 0

± 0 : s / 0 / 0 (Obs! i detta fall sätts inledande mantissabiten = 0.)

Talområde för exponenten: $-126 \leq \exp \leq 127$

Största belopp blir: $1.111\ 1111\ 1111\ 1111\ 1111\ 1111 \cdot 2^{+127} \approx 2^{+128}$

Minsta belopp (förutom 0) blir: $1.000\ 0000\ 0000\ \dots 0 \cdot 2^{-126} = 2^{-126}$

För att kunna representera tal med ännu mindre belopp än minsta beloppet ovan har man infört undantaget nedan, som dock ger sämre upplösning än normalfallet.

Undantag:

Man kan i uttrycket för "0" ovan välja $f \neq 0$.

Största belopp skulle i detta fall bli:

$$0.111\dots 1 \cdot 2^{-127} = (1 - 2^{-23}) \cdot 2^{-127}$$

Det uppstår då ett "glapp" till minsta beloppet (2^{-126}) i "normalfallet" ovan. För att undvika "glappet" bestämmer man att exponenten skall sättas till -126 istället för -127 då $c = 0$.

Största beloppet i undantagsfallet blir då: $(1 - 2^{-23}) \cdot 2^{-126}$