Virgola fissa (VF) vs virgola mobile (VM)

Esempio di VF e VM con 6 bit: $iii.fff$ vs $i.fff*10^{ee}$ 6 bit per

Valore minimo (Vmin) e massimo (Vmax)

	Vmin	Vmax
VF	000.000	999.999 ≈ 10^3
VM	0.000* 10^{00}	9.999*10^99≈ 10^{100}

Risoluzione: più piccola differenza descrivibile

	VF	VM
Risoluzione	000.001 = 10^{-3}	1.001* 10^{ee} -1.000* $10^{ee}=¿$ 0.001* 10^{ee} = $10^{-3}*10^{ee}$

Errore assoluto: differenza tra il numero x e la sua rappresentazione e_A (x) = x – rappresentazione(x)

Errore relativo: differenza percentuale tra il numero e la sua rappresentazione e_R (x) = e_A (x)/x

	VF	VM
e_A (x)	≤ 10^{-3}	≤ $10^{-3}*10^{ee}$
$e_R(x)=\frac{e_A(x)}{x}$	$\left[\frac{10^{-3}}{10^3},\frac{10^{-3}}{0}\right)$	$e_R(x)=\frac{e_A(x)}{x}\leq\frac{max\, e_A(x)}{min\, x}$ Ricordando che se x è un numero normalizzato allora $10^{ee}\leq\lvert x\rvert<10^{ee+1}$ $e_R\leq\frac{10^{-3}*10^{ee}}{10^{ee}}=10^{-3}$

Standard IEEE 754

Esempio 1: convertiamo $(17.375)_{10}$, dobbiamo determinare s, m, E

- Il numero è positivo $\rightarrow s = 0$
- Converto la parte intera in binario: $(17)_{10} = (10001)_2$
- Converto la parte frazionaria $(.375)_{10}$ $(= \frac{3}{2^3})$

$$\begin{array}{rll|l} 0.375 \cdot 2 = 0.75 & \text{parte intera} = 0 & \text{MSD} & \\ 0.75 \cdot 2 = 1.5 & \text{parte intera} = 1 & & (.375)_{10} = (.011)_2 \\ 0.5 \cdot 2 = 1.0 & \text{parte intera} = 1 & & \end{array}$$

- Unisco i risultati: $(10001.011)_2$
- Normalizzazione: $(1.0001011)_2 \cdot (10)_2^4$
- Mantissa: $m = 0001\ 0110\ 0000\ 0000\ 0000\ 000$
- Esponente: $E = e + 127 = (4)_{10} + (127)_{10} = (131)_{10} = (10000011)_2$

Esempio 2: convertiamo $(-0.8)_{10}$, dobbiamo determinare s, m, E

- Il numero è negativo $\rightarrow s = 1$
- Converto la parte intera in binario: $(0)_{10} = (0)_2$
- Converto la parte frazionaria $(.8)_{10}$

$$\begin{array}{lll} \mathbf{0.8} \cdot 2 = 1.6 & \text{parte intera} = 1 & \text{MSD} \\ 0.6 \cdot 2 = 1.2 & \text{parte intera} = 1 & \\ 0.2 \cdot 2 = 0.4 & \text{parte intera} = 0 & \\ 0.4 \cdot 2 = \mathbf{0.8} & \text{parte intera} = 0 & \\ 0.8 \cdot 2 = \ldots & \text{la sequenza 1100 si ripete} & \end{array}$$

- Unisco i risultati: $(0.8)_{10} = (0.\overline{1100})_2$
- Normalizzazione: $(0.\overline{1100})_2 = (0.1100\ \overline{1100})_2 = (1.100\ \overline{1100})_2 \cdot (10)_2^{-1}$
- Mantissa: $m = 100\ 1100\ 1100\ 1100\ 1100\ 1100$
- Esponente:
  $E = e + 127 = (-1)_{10} + (127)_{10} = (126)_{10} = (1111110)_2 \rightarrow (01111110)_2$

Numero subnormalizzato

Esempio 1: rappresentare in formato IEEE 754 il numero
$-(0.125)_{10} \cdot 2^{-125}$

- Il numero è negativo $\rightarrow s = 1$
- Il numero è esprimibile come un subnormalizzato? Per esserlo devo poterlo scrivere come $(0.m) \cdot 2^{-126}$ (m ha 23 bit)
- $0.125 \cdot 2^{-125} = 0.125 \cdot 2^{1} \cdot 2^{-126} = 0.25 \cdot 2^{-126}$, quindi posso scriverlo come subnormalizzato
- converto 0.25 $\left(\frac{1}{2^2}\right)$ in base 2: $(0.25)_{10} = (0.01)_2$
- Mantissa $m = 01\ 0 \ldots 0$
- Esponente $E = 0000\ 0000$

Convertire da base 10 a base 2: 33.1001_{10}

Parte intera: 100001

Parte frazionaria:

0.1001*2=0.2002 → 0

0.2002*2=0.4004 → 0

0.4004*2=0.8008 → 0

0.8008*2=1.6016 → 1

0.6016*2=1.2032 → 1

0.2032*2=0.4064 → 0

…

Convertire da base 2 a base 10: 110001.1001101_2

Parte intera: 1+16+32 = 49

Parte frazionaria: 1/2 + 1/16 + 1/32 + 1/128 = 0.6016

Risultato: 49.6016

Convertire il numero -24.511 da base 10 a base 2 in virgola fissa con 4 bit per la parte frazionaria. Identificare il minimo numero di bit per la parte intera

Numero di bit per la parte intera: $-32 \leq -24 \leq 31$

Ci sono 64 livelli quindi 2^6 → n=6

Parte intera.

Conversione di -24 in binario con n=6 e complemento a 2

$24_{10} = 011000_2$

100111+1 = 101000 ← -24

Parte frazionaria su n=4

0.511*2=1.022 → 1

0.022*2=0.044 → 0

0.044*2=0.088 → 0

0.088*2=0.176 → 0

$0.511_{10}=1000_2$

Risultato: 101000.1000

Convertiamo il numero dell'esercizio precedente da virgola fissa a standard IEEE-754

Numero: -24.511

s=1

Parte intera:

$24_{10}=11000_2$

Parte frazionaria:

$0.511_{10}=1000_2$

Numero: 11000.1000 → 1.10001000 * $(10)_2^4$

M = 10001000000000000000000

E = 4+127 = 131_{10} = 10000011_2

Risultato: 1 10000011 10001000000000000000000