CALCOLO NUMERICO → Sviluppo di algoritmi numerici per risolvere problemi matematici

ARITMETICA di MACCHINA

Costruirsi un modello ad alto livello per capire quale sia l'algoritmo più conveniente

NUMERO

esempio: $\frac{1}{3}$ → rapporto tra due cifre

$0,3333 \cdots 333$

Indichiamo con "B" la base, con $B \in \mathbb{N}$, $B \geq 2$

TEOREMA di RAPPRESENTAZIONE IN BASE (2.1.1.)

Sia $x \in \mathbb{R}$, $x \neq 0$. Allora esistono e sono univocamente determinati

① $\{d_i\}_{i \geq 1}$, con $d_1 \neq 0$, $0 \leq d_i \leq B-1$ e d_i non definitivamente uguali a $B-1$, dette "cifre" della rappresentazione

② Un intero $p \in \mathbb{Z}$ (intero negativo) detto "esponente" della rappresentazione

tali per cui si ha:

esponente

mantissa

$$x = \text{sign}(x) \cdot B^{p} \sum_{i=1}^{+\infty} d_i B^{-i}$$

RAPPRESENTAZIONE NORMALIZZATA IN VIRGOLA MOBILE

↓

primo numero dopo la virgola NON è zero

esempio

$B = 10$, numero: $0,01$

$$\Rightarrow \quad x = +10^{p}\left(d_1 \cdot 10^{-1} + d_2 \cdot 10^{-2} + \cdots\right)$$

$$+d_1 \cdot 10^{p-1} + d_2 \cdot 10^{p-2}$$

$\downarrow$

È R primo numero e deve essere necessariamente $\neq 0$

Quindi vado a vedere la prima cifra "non zero" e in questo caso si trova nella posizione 10^{-2}

$$\Rightarrow \quad p-1 = -2 \longrightarrow p = -1$$

$$x = +10^{-1}\left(d_1 \cdot 10^{-2}\right)$$

esempio

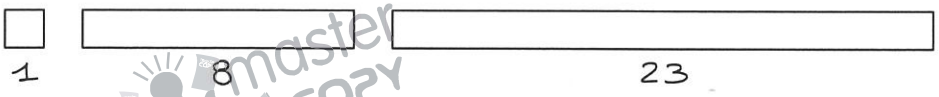
$$0,\overline{9} = 9\cdot 10^{-1} + 9\cdot 10^{-2} + 9\cdot 10^{-3} + \cdots = \sum_{i=1}^{+\infty} 9\cdot 10^{-i} = 9\sum_{i=1}^{+\infty} 10^{-i} =$$

$$= 9\cdot \sum_{i=1}^{+\infty}\left(\frac{1}{10}\right)^i = \frac{9}{10}\cdot \sum_{i=0}^{+\infty}\left(\frac{1}{10}\right)^i$$

$$\lim_{k\to+\infty} \sum_{i=0}^{k} \alpha^i = \lim_{k\to} \frac{\alpha^{k+1}-1}{\alpha-1} = \frac{1}{1-\alpha} \qquad \left(\alpha = \frac{1}{10}\right)$$

$$\Rightarrow \frac{9}{10}\cdot\frac{1}{1-\frac{1}{10}} = \frac{9}{10}\cdot\frac{10}{9} = 1$$

Dal teorema di rappresentazione in base, segue che la rappresentazione di un numero reale nel calcolatore può avvenire assegnando delle posizioni di memoria per: SEGNO, ESPONENTE e CIFRE DELLA RAPPRESENTAZIONE



$$X = \pm 2^{p} \sum_{i=1}^{23} d_i \cdot 2^{-i} \qquad d_1 \neq 0,\ p \in [-127, 127]$$

Definizione (2.1.1.)

Definiamo "INSIEME DEI NUMERI di MACCHINA" con t cifre, base B e range $(-m, M)$ l'insieme dei numeri reali:

$$F(B,t,m,M) = \{0\} \cup \left\{ x \in \mathbb{R} : x = \mathrm{sgn}(x) B^{p} \sum_{i=1}^{t} d_i \cdot B^{-i},\ 0 \leq d_i \leq B-1,\ d_1 \neq 0,\ -m \leq p \leq M \right\}$$

base ← RANGE
cifre
unisco lo zero perché non ha rappresentazione normalizzata

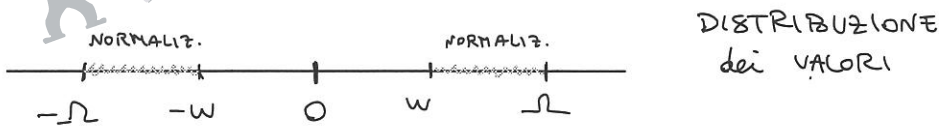
OSSERVAZIONI

- L'insieme dei numeri di macchina $F(B,t,m,M)$ ha cardinalità finita $N = 2B^{t-1}(B-1)(M+m+1) + 1$
- $x \in F(B,t,m,M)$ e $x \neq 0 \Rightarrow w = B^{-m-1} \leq |x| \leq B^{M}(1 - B^{-t}) = \Omega$

w: Più piccolo valore positivo di macchina
Ω: Più grande valore positivo di macchina

→ se $p = M$, si introducono rappresentazioni speciali per i simboli $\pm\infty$ e "NaN" (NOT A NUMBER)

- L'insieme dei numeri di macchina $\mathbb{F}(B, t, m, M)$ è simmetrico rispetto all'origine



- Posto $x = (-1)^s B^p \alpha \in \mathbb{F}(B, t, m, M)$ allora il successivo numero di macchina risulta essere $y = (-1)^s B^p (\alpha + B^{-t})$ per cui in un sistema a virgola mobile la distanza $|y - x| = B^{p-t}$ → se aumenta p, ottengo "spazi" più grandi

DISTANZA TRA DUE VALORI

esempio

±	$0 \le p \le 255$	$d_1 \cdots d_{23}$

Che numero associo a questa configurazione di registro?

- Se $1 \le p \le 254$ $\quad x = \pm 2^{p-126} \left(1 \cdot 2^{-1} + \sum_{i=1}^{23} d_i \cdot 2^{-i-1}\right)$
- Se $p = 0$ $\quad x = \pm 2^{-126} \left(\sum_{i=1}^{23} d_i \cdot 2^{-i}\right)$
- Se $p = 255$ $\quad \begin{cases} +\text{inf.} \\ -\text{inf.} \\ \text{NaN} \end{cases}$ → seq. di cifre non tutte nulle con meno davanti

↳ Tutti 1 e tutti zero

ARITMETICA di MACCHINA

Rappresentare un numero reale non nullo $x \in \mathbb{R}, x \neq 0$, in macchina significa APPROSSIMARE x con un numero $\tilde{x} \in \mathcal{F}(B, t, m, M)$ commettendo un ERRORE RELATIVO di rappresentazione:

$$\varepsilon_x = \frac{\tilde{x} - x}{x} = \frac{\eta_x}{x}, \quad x \neq 0$$

ERRORE RELATIVO di RAPPRESENTAZIONE → Non dipende dalla grandezza del numero, bensì dall'aritmetica della macchina

con $\eta_x = \tilde{x} - x$ "ERRORE ASSOLUTO di RAPPRESENTAZIONE"

N.B. L'errore relativo è importante per una valutazione qualitativa del fenomeno studiato!

Dato $x \in \mathbb{R}, x \neq 0$, distinguiamo 2 casi:

1. $|x| < w$ (UNDERFLOW) oppure $|x| > \Omega$ (OVER FLOW)
2. $w \leq |x| \leq \Omega$

↳ Tecniche di approssimazione:

- ARROTONDAMENTO: il numero x viene approssimato con il numero rappresentabile $\tilde{x}$ più vicino
- TRONCAMENTO: il numero x viene approssimato con il più grande numero rappresentabile $\tilde{x}$ il cui valore assoluto risulti minore o uguale al valore assoluto di x
- ROUND TOWARD PLUS/MINUS INFINITY: x viene approssimato al più piccolo/grande numero rappresent. maggiore/minore del dato