Titolo nota 24/02/2021

Lezione 24/02 Aritmetica di macchina

$x \in \mathbb{R} \quad x \neq 0 \qquad x = \pm B^p \cdot \sum_{i=1}^{+\infty} d_i \cdot B^{-i}$

$d_1 \neq 0 \qquad d_i \neq B-1$

$\pm$	p	d_i

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

F ha cardinalità finita $\quad x = \pm B^{p} \cdot \sum_{i=1}^{t} d_i \cdot B^{-i}$

$$\# F = 2(m + M + 1)(B-1)\, B^{t-1} + 1$$

$$\omega = B^{-m} \cdot B^{-1}$$

$$\Omega = B^{M} \cdot (B-1) \sum_{i=1}^{t} B^{-i}$$

$-\Omega \quad -\omega \quad 0 \quad \omega \quad \Omega$

$$dist(x, y) = B^{p} \cdot \left(\sum_{i=1}^{t} d_i \cdot B^{-i} + B^{-t} - \sum_{i=1}^{t} d_i \cdot B^{-i} \right) = B^{p-t}$$

$x \in R^+$

$|x| > \Omega$

$x > \Omega$

<u>overflow</u>

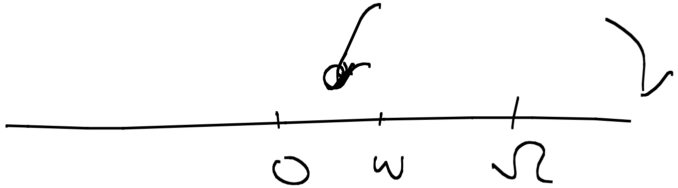
$x \rightsquigarrow +\inf$

$x \in R^-$

$\underline{0 < x < \omega}$

<u>underflow</u>

$x \rightsquigarrow 0$



$x \in \mathbb{R}$

$x \in [\omega, \Omega]$

$\rightsquigarrow$

$\tilde{x} \in F$

$$f : [\omega, \Omega] \longrightarrow F$$

$$\mathbb{R} \ni x = B^p \cdot \sum_{i=1}^{+\infty} d_i \cdot B^{-i} \qquad d_1 \neq 0 \qquad d_i \neq B-1$$

$$p \in [-m, M]$$

f = troncamento

$$F \ni \tilde{x} = B^p \cdot \sum_{i=1}^{t} d_i \cdot B^{-i} = \text{più grande numero di macchina più piccolo di } x$$

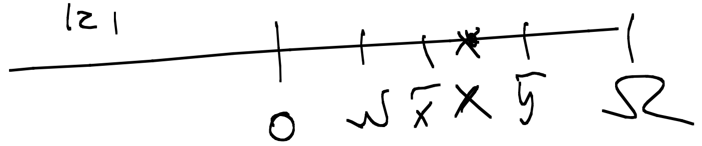
$\omega \quad \tilde{x} \quad x \quad \tilde{y} \quad \Omega$

$-\Omega \quad \tilde{y} \quad x \quad \tilde{x} \quad -\omega$

$$x \in \mathbb{R} \rightsquigarrow \tilde{x} = \mathrm{trn}(x) = f(x)$$

errore di rappresentazione

$$x = \beta^{p} \cdot \sum_{i \geq 1}^{t} d_i \cdot \beta^{-i} \qquad d_1 \neq 0$$



$$\epsilon_x = \frac{\hat{x} - x}{x} \qquad (x \neq 0)$$

$$|\epsilon_x| = \left| \frac{\hat{x} - x}{x} \right|$$

$$|\epsilon_x| = \left| \frac{\hat{x} - x}{x} \right| \leq \frac{|\tilde{y} - \tilde{x}|}{|x|} \leq \frac{\beta^{p-t}}{|x|} \leq \frac{\beta^{p-t}}{\beta^{p-1}} = \beta^{1-t}$$

u ϵ

$$\epsilon_x = \frac{\hat{x} - x}{x} \qquad |\epsilon_x| \le \epsilon(u) \text{ preciznost mašine}$$

$$\hat{x} = x + \epsilon_x \cdot x = x \cdot (1 + \epsilon_x) \qquad |\epsilon_x| \le \epsilon, u$$

$$\hat{x} \in \hat{F}, \quad \tilde{y} \in \hat{F}, \quad \underline{\tilde{x} + \tilde{y}} \notin \hat{F}$$

$$\tilde{x} \oplus \tilde{y} \qquad \tilde{x} \ominus \tilde{y} \qquad \tilde{x} \otimes \tilde{y} \qquad \tilde{x} \oslash \tilde{y}$$

$$\tilde{x} \oplus \tilde{y} = \text{trn.}(\tilde{x} + \tilde{y}) = (\tilde{x} + \tilde{y})(1 + \epsilon_1) \qquad |\epsilon_1| \le u$$

ϵ_1 errori locali delle operazioni

ϵ_x errore di rappresentazione

$|\epsilon_1|, \; |\epsilon_x| \leq u$

$$f(x) = x^2 - 1 = (x-1)\cdot(x+1)$$

$$x \rightsquigarrow \tilde{x} = x \cdot (1 + \epsilon_x)$$

$$y = \tilde{x} * \tilde{x} \rightsquigarrow \tilde{x} \otimes \tilde{x} = \tilde{x}^2 \cdot (1 + \epsilon_1)$$

$$y - 1 \rightsquigarrow y \ominus 1 = (\tilde{x} \otimes \tilde{x}) \ominus 1$$

$$(\tilde{x} \otimes \tilde{x}) \ominus 1 = \left(\tilde{x}^2 \cdot (1+\epsilon_1) - 1\right)(1+\epsilon_2)$$

$$= \left[x^2 (1+\epsilon_x)^2 (1+\epsilon_1) - 1\right](1+\epsilon_2)$$

$$\doteq x^2 \cdot (1+2\epsilon_x)(1+\epsilon_1)\cdot(1+\epsilon_2) - (1+\epsilon_2)$$

$$\doteq x^2 \cdot (1+2\epsilon_x+\epsilon_1+\epsilon_2) - (1+\epsilon_2)$$

$$\doteq (x^2-1) + 2x^2\epsilon_x + x^2\cdot\epsilon_1 + \epsilon_2\cdot(x^2-1)$$

$$(\tilde{x} \otimes \tilde{x}) \ominus 1 \doteq (x^2 - 1) + 2\epsilon_x x^2 + x^2 \cdot \epsilon_2 + \epsilon_2 \cdot (x^2 - 1)$$

$$\frac{(\tilde{x} \otimes \tilde{x}) \ominus 1 - (x^2 - 1)}{(x^2 - 1)} \doteq \frac{2x^2}{x^2 - 1}\epsilon_x + \frac{x^2}{x^2 - 1}\epsilon_2 + \epsilon_2$$

$$x \rightsquigarrow \tilde{x} \rightsquigarrow (\tilde{x} \ominus 1) \otimes (\tilde{x} \oplus 1)$$

$$= \left(x \cdot (1 + \epsilon_x) - 1\right)(1 + \delta_1) \cdot \left(x(1 + \epsilon_x) + 1\right)(1 + \delta_2)(1 + \delta_3)$$

$$= \left(x(1+\epsilon_x)-1\right)\left(x(1+\epsilon_x)+1\right)(1+\delta_1)(1+\delta_2)(1+\delta_3)$$

$$\doteq \left[x^2\cdot(1+\epsilon_x)^2-1\right]\cdot(1+\delta_1+\delta_2+\delta_3)$$

$$\doteq \left[x^2\cdot(1+2\epsilon_x)-1\right](1+\delta_1+\delta_2+\delta_3)$$

$$\doteq (x^2-1)+(x^2-1)(\delta_1+\delta_2+\delta_3)+2x^2\epsilon_x$$

$$\frac{(\tilde{x}\ominus 1)\otimes(\tilde{x}\oplus 1)-(x^2-1)}{x^2-1} = \frac{2x^2}{x^2-1}\epsilon_x+\delta_1+\delta_2+\delta_3$$

$$\frac{(\tilde{x} \otimes \tilde{x}) \ominus 1 - (x^2 - 1)}{x^2 - 1} = \frac{2x}{x^2-1}\epsilon_x + \frac{x^2}{x^2-1}\epsilon_1 + \epsilon_2$$

$$\frac{(\tilde{x} \ominus 1) \otimes (\tilde{x} \oplus 1) - (x^2 - 1)}{x^2 - 1} = \frac{2x}{x^2-1}\epsilon_x + \delta_1 + \delta_c + \delta_3$$

$$|\delta_1 + \delta_2 + \delta_3| \leq |\delta_1| + |\delta_1| + |\delta_3| \leq 3u$$

$$\left|\frac{x^2}{x^2-1}\epsilon_1 + \epsilon_2\right| \leq \frac{x^2}{|x^2-1|}|\epsilon_1| + |\epsilon_2| \leq u + u \cdot \frac{x^2}{|x^2-1|}$$

$f : \mathbb{R} \to \mathbb{R}$ $\qquad$ f razionale

$f(x) = e^x - 1$

$x \rightsquigarrow f(x)$

$\tilde{x} \longrightarrow f(\tilde{x})$

$$\varepsilon_{in} = \frac{f(\tilde{x}) - f(x)}{f(x)}$$

<u>errore inerente</u> $\quad (f(x) \neq 0)$

$$\varepsilon_{alg} = \frac{g(\tilde{x}) - f(\tilde{x})}{f(\tilde{x})}$$

<u>errore algoritmico</u>

$$\epsilon_{tot} = \frac{g(\tilde{x}) - f(x)}{f(x)}$$

$$\underline{\epsilon_{tot}} \doteq \epsilon_m + \epsilon_{alg}$$

$$\frac{g(\tilde{x}) - f(x)}{f(x)} = \frac{g(\tilde{x}) - f(\tilde{x}) + f(\tilde{x}) - f(x)}{f(x)} =$$

$$\frac{g(\tilde{x}) - f(\tilde{x}) + f(\tilde{x}) - f(x)}{f(x)} =$$

$$\frac{g(\tilde{x}) - f(\tilde{x})}{f(x)} + \frac{f(\tilde{x}) - f(x)}{f(x)} \quad \epsilon_{m'}$$

$$\frac{f(\tilde{x})}{f(x)} - 1 = \epsilon_m$$

$$= \left(\frac{g(\tilde{x}) - f(\tilde{x})}{f(\tilde{x})} \cdot \frac{f(\tilde{x})}{f(x)}\right) + \epsilon_{m'}$$

$= c \log (c n + 1) + c n$

$= c \log + c n$

Stabilito dall'algoritmo

Componente del calcolo di f e parte del problema ricorsivo

$$f(x) = \frac{x-1}{x} = 1 - \frac{1}{x}$$