2° semestre – 2018/2019

master COPY

Calcolo numerico

Gemignani Luca
Poloni Federico

6 CFU

master COPY


master
COPY

4/03/2019

L'ARITMETICA DEL CALCOLATORE

→ RAPPRESENTAZIONE IN BASE E NUMERI DI MACCHINA

Un numero può essere rappresentato in modi diversi:

$\frac{1}{3} = 0,333\ldots = 0,\overline{3}$ → per essere inserito in macchina è necessaria un'unica rappresentazione.

TEOREMA DI RAPPRESENTAZIONE IN BASE

$B \in \mathbb{N}, B > 1$ → B è la base di numerazione

Dato $x \in \mathbb{R}, x \neq 0$, esistono e sono univocamente determinati:

1) $p \in \mathbb{Z}$ → ESPONENTE della rappresentazione;

2) una successione di numeri naturali $\{d_i\}_{i \geq 1}$ con $d_1 \neq 0$, $0 \leq d_i \leq B-1$ e di non definitivamente uguali a $B-1$,

$\{d_i\}_{i \geq 1}$ sono le CIFRE DELLA RAPPRESENTAZIONE,

tali che

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

detta **RAPPRESENTAZIONE NORMALIZZATA IN VIRGOLA MOBILE (floating point)**

La rappresentazione si dice NORMALIZZATA poiché $d_1 \neq 0$, e in virgola mobile poiché è necessario aggiustare p per avere la parte intera nulla e d_1 non nulla.

Le condizioni $d_1 \neq 0$ e di non definitivamente uguali a $B-1$ sono state introdotte per GARANTIRE L'UNICITÀ DELLA RAPPRESENTAZIONE.

Esempio 1

$B = 10 \qquad x = 0,01 = 0,1 \cdot 10^{-1}$

$x = +10^{0}\ (0,01) \qquad \text{sign}(x)\, B^{-p} \sum_{i=1}^{+\infty} d_i B^{-i}$

$d_1 = 0$ allora non è normalizzata

$x = +10^{-1}\ (0,1) = +10^{1} \cdot (0 \cdot 10^{0} + 1 \cdot 10^{-1}) = 10^{0}$ OK

$d_1 \neq 0$ OK

SERIE GEOMETRICA

$\sum_{i=0}^{n} \vartheta^{i} = \frac{1-\vartheta^{n+1}}{1-\vartheta}$ Se pongo $\vartheta = \sum_{i=1}^{n} \left(\frac{1}{10}\right)^{n} = 1$

$\underbrace{0,\overline{9}}_{\vartheta}$

Esempio 2

$x = 1 = 0,1 \cdot 10^{1}$

In macchina → $x = +10^{1}(0,1) = +10^{1}(0 \cdot 10^{0} + 1 \cdot 10^{-1}) = 10^{0} = 1$

Però anche $0,\overline{9} = 1$ (vedi serie geometrica)

In forma normalizzata $y = 0,\overline{9} \rightarrow y = +10^{0} \sum_{i=1}^{+\infty} 9 \cdot 10^{-i}$ → ma tutte le cifre sono definitivamente uguali a 9, allora non è accettabile → la RAPPRESENTAZIONE È UNICA.

- $x = 0$ non ammette una rappresentazione normalizzata.
- I numeri complessi sono rappresentati come coppia di numeri reali.

I numeri vengono trasferiti in macchina in un REGISTRO = sequenza FINITA di celle di memoria in cui vengono collocati dei numeri che saranno memorizzati. Per cui l'INSIEME DEI NUMERI DI MACCHINA (quelli rappresentabili in macchina) ha CARDINALITÀ FINITA → da questo ne segue che spesso vi saranno delle approssimazioni.

Rappresentazione di un numero nel calcolatore

SEGNO	ESPONENTE	MANTISSA
±	p	cifre della rappresentazione

Es. se ho a disposizione 32 bit, avrò 1 cella per il segno, 8 per l'esponente, e 23 per la mantissa.

↓

anche l'esponente avrà un segno, che occuperà una cella

$\pm(1\cdot 2^6 + 1\cdot 2^5 + \dots + 1\cdot 2^0)$

→ $p \in [-127, 127]$ → sarebbe $2^7 - 1 = 128 - 1 = 127$

$x = \pm 2^p \cdot \sum_{i=1}^{23} d_i 2^{-i}$



INSIEME DEI NUMERI DI MACCHINA

Def L'insieme dei numeri di macchina in rappresentazione floating point con t-cifre, base B e range $(-m, M)$ è l'insieme dei numeri reali

$$F(B,t,m,M) = \{0\} \cup \{x \in \mathbb{R} \mid x = \text{sign}(x) B^p \sum_{i=1}^{t} d_i B^{-i},\ 0 \le d_i \le B-1,\ d_1 \neq 0,\ -m \le p \le M\}$$

Osservazioni:

→ $F(B,t,m,M)$ ha cardinalità finita $N = 2(B^{t-1}(B-1)(M-m+1)+1)$;

→ Lo zero è un numero di macchina e ha due rappresentazioni (±)

→ Posso fare delle semplificazioni per liberare spazio:

- $d_0 = 1$ non viene rappresentata
- per l'esponente vengono eliminate le configurazioni tutti zero e tutti uno.
- Non rappresento il segno dell'esponente → $p \in (0, 254)$

Allora $x = \pm 2^{(p+126)} \sum_{i=1}^{23} d_i \cdot 2^{-i-1}$

↳ in questo modo risparmio la cella per il segno

→ Il numero più piccolo rappresentabile è $w = +B^{-M-1}$

Il numero più grande rappresentabile è $\Omega = B^M(1-B^{-t})$

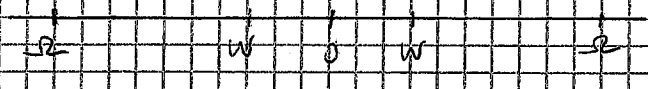
Allora se $x \in F(B,t,m,M)$ e $x \neq 0$, $B^{-M-1} \leq |x| \leq B^M(1-B^{-t})$

Numeri inferiori di w o superiori a Ω non sono rappresentabili → allora è stata introdotta una RAPPRESENTAZIONE DENORMALIZZATA, in cui d₁ può essere uguale a zero.

$$x = \pm 2^{p-1\leq} \sum_{i=1}^{t} a_i 2^{-i} \quad \text{se } p=0$$

se $p = M$ vengono introdotte delle rappresentazioni speciali per i simboli $\pm\infty$ < NaN → not a number

→ $F(B,t,m,M)$ è simmetrico rispetto all'origine



→ DISTANZA TRA DUE NUMERI CONSECUTIVI

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

$|y-x| = B^{p-t}$ varia con p (ordine di grandezza dei numeri considerati)

Da questo ne deriva che i numeri di macchina saranno fitti intorno a w e radi intorno a Ω.

A noi interessano le distanze relative < errori piccoli su distanze piccole, errori grandi su distanze grandi

→ = percentuale dell'errore

Motivo per il quale sono stati inseriti i numeri denormalizzati (per ridurre i vuoti ed evitare ecc. l'alta percentuale dell'errore).

→ $x \in \mathbb{R}$, $x \neq 0$

- $|x| \in [w, \Omega]$ → è garantita un'ACCURATEZZA UNIFORME della rappr.
- $|x| \notin [w, \Omega]$ < OVERFLOW / UNDERFLOW → non si garantisce accuratezza

↳ a questi numeri possono essere attribuiti $w, \Omega, \pm\infty, 0$ ...