

Architettura degli Elaboratori A.A. 2018/2019

Rappresentazione dei Numeri Razionali

Esercizi

Dott. Mirko Staderini – mirko.staderini@unifi.it

DiMaI – Università degli Studi di Firenze

Viale Morgagni, 65 – 50134 Firenze

Esercizio 1 – Conversione da IEEE754 a decimale

Eeguire la decodifica in decimale del seguente numero rappresentato in formato IEEE754 in singola precisione:

X =

0	1	0	0	0	0	1	1	0	1	0	0	1	0	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---

Soluzione

- Segno (1 bit):

$$S = 0 \text{ (numero positivo)}$$

- Esponente (8 bit):

L'esponente è memorizzato in forma polarizzata, al solito sottraiamo la costante di polarizzazione

$$E = EXP - BIAS = (10001110)_2 - (2^7 - 1) = 2^1 + 2^2 + 2^7 - (2^7 - 1) = 7$$

- Mantissa (23 bit):

$$F = 10010101111$$

Ricordando che:

$$v(X) = (-1)^S \cdot 1, F \cdot 2^E$$

otteniamo:

$$X = (1,10010101111)_2 \cdot 2^7 = (11001010,1111)_2$$

Dove moltiplicare per 2^7 significa spostare la virgola a destra di 7 posizioni

$$\text{Parte intera: } (11001010)_2 = 2^1 + 2^3 + 2^6 + 2^7 = 202$$

$$\text{Parte decimale: } (0,1111)_2 = 2^{-1} + 2^{-2} + 2^{-3} + 2^{-4} = 0.5 + 0.25 + 0.125 + 0.0625 = 0.9375$$

$$X = 202,9375$$

Esercizio svolto – Somma in Floating Point

Siano

$$X = 202,9375$$

$$Y = -54,8125$$

Calcolare la loro somma eseguita in codifica IEEE754

Soluzione

Sappiamo dagli Esercizi precedenti che

$$X_{IEEE754} = 0 \ 10000110 \ 100101011110000000000000$$

$$X_2 = 1,10010101111 \cdot 2^7$$

$$Y_{IEEE754} = 1 \ 10000100 \ 101101101000000000000000$$

$$Y_2 = -1,101101101 \cdot 2^5 = -0,01101101101 \cdot 2^7$$

N.B. Y è negativo. In questi casi si sottrae il valore più piccolo (in valore assoluto) da quello più grande.

$$\begin{array}{r} |Z| = |X| - |Y| \quad \begin{array}{cccccccccccc} 1 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 1 & 1 & 1 & - \\ 0 & 0 & 1 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 0 & 1 & = \\ \hline 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & \end{array} \end{array}$$

$$Z = 1,00101000010 \cdot 2^7$$

Segno positivo: $S = 0$ (numero positivo)

Mantissa: $F = 00101000010$

Esponente: $EXP = E + BIAS = 7 + 127 = 134 = 10000110$

$$Z_{IEEE754} = 0 \ 10000110 \ 001010000100000000000000$$

Possiamo verificare il risultato

Segno positivo: 0

Esponente:

$$E = EXP - BIAS = (10000110)_2 - (2^7 - 1) = 2^1 + 2^2 + 2^7 - (2^7 - 1) = 7$$

Mantissa:

$$F = 0010100001$$

Ricordando che:

$$v(Z) = (-1)^S \cdot 1, F \cdot 2^E$$

otteniamo:

$$Z = (1,0010100001)_2 \cdot 2^7 = (10010100,001)_2$$

Dove moltiplicare per 2^7 significa spostare la virgola a destra di 7 posizioni

Parte intera: $(10010100)_2 = 2^2 + 2^4 + 2^7 = 4 + 16 + 128 = 148$

Parte decimale: $(0,001)_2 = 2^{-3} = 0.125$

$$Z = 148,125$$

Infatti

X	202,9375	-
Y	54,8125	=
Z	148,1250	

Esercizio svolto – Errori di approssimazione

Dato il numero $A = 250,90625$:

- 1) rappresentare A in virgola mobile secondo il seguente formato:
 - 1 bit per il segno (bit contrassegnato con “0”);
 - 4 bit per l’esponente (da bit “1” a bit “4”) in forma polarizzata;
 - 7 bit per la mantissa (da bit “5” a bit “11”).

0	1	2	3	4	5	6	7	8	9	10	11

- 2) Data la rappresentazione di A trovata al punto precedente, calcolare il valore B effettivamente rappresentato. A quanto ammonta l’errore di approssimazione?
- 3) Quanti bit aggiuntivi servirebbero per rappresentare il numero senza errori di approssimazione?

Soluzione

- 1) Segno: 0 (positivo)
Parte intera: 11111010

N	N mod 2
250	0
125	1
62	0
31	1
15	1
7	1
3	1
1	1
0	

Parte decimale: $0,90625 = 0,11101$

N	2*N	Trunc(2*N)
0,90625	1,8125	1
0,8125	1,625	1
0,625	1,25	1
0,25	0,5	0
0,5	1	1
0		

Quindi: 11111010,11101

Normalizzando: $1,111101011101 \cdot 2^7$

Esponente: $7 + \text{BIAS} = 7 + (2^4 - 1) = 14 = 8 + 2 + 4 = 1110$

$A' =$

0	1	2	3	4	5	6	7	8	9	10	11
0	1	1	1	0	1	1	1	1	0	1	0

2) $B = 1,1111010 \cdot 2^7 =$
 $= 11111010 = (2^1 + 2^3 + 2^4 + 2^5 + 2^6 + 2^7) = 2 + 8 + 16 + 32 + 64 + 128 = 250$

$B = 250$

3) L'errore di approssimazione è $|A - B| = |250,90625 - 250| = 0,90625$

4) Sarebbero necessari 5 ulteriori bit di mantissa, per un totale di 12

0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
0	1	1	1	0	1	1	1	1	0	1	0	1	1	1	0	1