## Cálculo Numérico - IME/UERJ

## Gabarito - Lista de Exercícios 1 - Aritmética de ponto flutuante

1. (a) 19

(d) 1,5

(g) 12,25

(b) 226

- (e) 1,59375
- (h) 0,328125

(c) 65

- (f) 12,625
- (i) 0,890625

2. (a) 10111

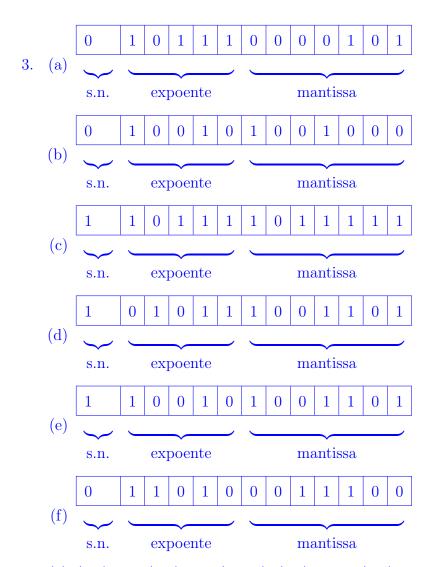
(d) 10,1

(g)  $1010,00\overline{0011}$ 

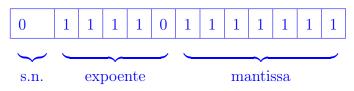
- (b) 11111111
- (e)  $0.0\overline{0011}$
- (h) 111111,11001111010111000011 (com 20 casas)

- (c) 101000110111
- (f)  $11,\overline{1100}$

(i)  $0,\overline{1100}$ 

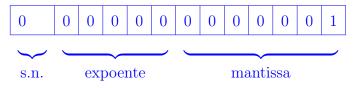


- 4. (a) (3-a) 266; (3-b) 12,5 (exato); (3-c) -446; (3-d) 0,10009765625 (3-e) -12,8125 (3-f) 2496
  - (b) Maior número positivo



$$1,11111111 \times 2^{15} = (65280)_{10}$$

Menor número positivo (o número positivo mais próximo de zero está na forma desnormalizada)



$$0,0000001 \times 2^{-14} = 2^{-21} \approx 4,7684 \times 10^{-7}$$

- (c) 100,5
- (d) 19,875

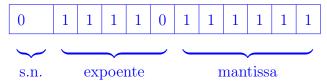
(e) 
$$m = 1,1001100 \times 2^4$$
 (exato -  $E_{abs} = E_{rel} = 0$ );  
 $n = 1,1110001 \times 2^6 = 120,5$  ( $E_{abs} = 0,25$ ,  $E_{rel} \approx 2,0790 \times 10^{-3}$ );  
 $p = 1,0100000 \times 2^1$  (exato -  $E_{abs} = E_{rel} = 0$ );  
 $a = 1,1100110 \times 2^8 = 460$  ( $E_{abs} = 0,25$ ,  $E_{rel} \approx 5,4318 \times 10^{-4}$ );  
 $b = 1,1100010 \times 2^8 = 452$  ( $E_{abs} = 1,25$ ,  $E_{rel} \approx 2,773 \times 10^{-3}$ ).

(a) Menor número positivo (o número positivo mais próximo de zero está na forma desnormalizada)



$$0,000001 \times 2^{-14} = 2^{-20} \approx 9,5367 \times 10^{-7}$$

Maior número positivo



$$1,111111 \times 2^{15} = (2-2^{-6}) \times 2^{15} = 65024$$

(b)

$$(4,25)_{10} = 1,000100 \times 2^{2}$$
  
+  $\frac{e}{4,25+e} = 0,0000000111111111... \times 2^{2}$ 

Note que no arredondamento para 6 dígitos na mantissa:

 $4,25 + e \approx 1,000100 \times 2^2 = 4,25.$ 

Portanto,

 $e = 0,00000001111111111... \times 2^2$ .

Normalizando e:

$$e = 1, 111111111... \times 2^{-8} \times 2^2 = 1, 111111111... \times 2^{-6}$$
.

Usando aproximação com 6 dígitos na mantissa:

$$e \approx 10,000000 \times 2^{-6} = 2^{-5} = 0,03125.$$

(c) Em binário e na forma normalizada,

$$4,25 = 100,01 = 1,000100 \times 2^{2}$$

Então, na forma normalizada, os bits da mantissa são: 000100.

O menor número maior que 4,25 é o número obtido na forma normalizada somando 1 à mantissa da representação de 4,25 na máquina. Então,

$$000100 + 1 = 000101.$$

Logo, este número será em decimal:

$$1,000101 \times 2^2 = (100,0101)_2 = 4,3125.$$

(d) Em binário e na forma normalizada,

$$80 = 1010000 = 1,010000 \times 2^{6}$$
.

Então, na forma normalizada, os bits da mantissa são: 010000.

O maior número menor que 80 é o número obtido na forma normalizada subtraindo 1 da mantissa da representação de 80 na máquina. Então,

$$010000 - 1 = 001111.$$

Logo, este número será em decimal:

$$1,001111 \times 2^6 = (1001111)_2 = 79.$$

(e) 
$$(0,8)_{10} = (0,110011001100...)_2 = 1,100110 \times 2^{-1}$$
  
 $(5)_{10} = (101)_2 = 1,010000 \times 2^2$   
 $0,8 \times 5 = 1,11111111 \times 2^1 \approx 10,000000 \times 2^1 = 1,000000 \times 2^2 = 4,0.$