

# **MAC0210: Relatório EP 1**

18 de Setembro de 2016

**Nathan Benedetto Proença - 8941276**  
**Victor Sena Molero - 8941317**

## **Sumário**

## Problem 1

**Questão 1 (3.11):** Suponha que temos um sistema de representação de ponto flutuante com base 2 e,

$$x = \pm S \times 2^E,$$

$$\text{com } S = (0.1b_2b_3b_4 \dots b_{24}),$$

$$\text{i.e., } \frac{1}{2} \leq S < 1$$

onde o expoente  $-128 < E < 127$ .

a) Qual é o maior número de ponto flutuante desse sistema?

*Resposta.*  $2^{126} - 2^{101}$ , basta preencher todos os bits (de  $b_2$  até  $b_{24}$ ) e escolher o maior expoente possível.  $\square$

b) Qual é o menor número de ponto flutuante positivo desse sistema?

*Resposta.*  $2^{-128}$ , basta escolher a menor mantissa possível (0.1) e o menor expoente possível ( $-127$ ).  $\square$

c) Qual é o menor inteiro positivo que não é exatamente representável nesse sistema?

*Resposta.*  $2^{24} + 1$ , basta escolher a menor mantissa não representável (0.10...01) e o menor expoente para o qual ela representa um inteiro (25).  $\square$

## Problem 2

**Questão 2 (5.1):** Qual é a representação do número  $1/10$  no formato IEEE single para cada um dos quatro modos de arredondamento?

*Resposta.* A resposta curta é:

$$x = 0.0001\overline{1}$$

,

$$\text{round\_down}(x) = \text{round\_towards\_zero}(x) = 1.10011001100110011001100 \times 2^{-4}$$

e

$$\text{round\_up}(x) = \text{round\_to\_nearest}(x) = 1.10011001100110011001101 \times 2^{-4}$$

.

Se  $x$  é uma representação binária exata de  $1/10$ ,  $x = 0.0001\overline{1} = 1.\overline{1001} \times 2^{-4}$  (os números abaixo da barra representam uma dízima periódica, se repetem infinitamente).

Calculamos então  $x_- = 1.10011001100110011001100 \times 2^{-4}$  e  $x_+ = 1.10011001100110011001101 \times 2^{-4}$ . Se o modo é *Round Down*, o número será representado por  $x_-$  e se for *Round Up*, será  $x_+$ , ambos pela definição dos modos.

Já que  $x > 0$ , o modo *Round towards zero* também usará  $x_-$ , porém  $x$  é mais próximo de  $x_+$  do que de  $x_-$ , basta perceber que o erro relativo entre  $x$  e  $x_-$  é maior que  $1/2$ , portanto, o modo *Round to nearest* usará a representação  $x_+$ .  $\square$

E para os números  $1 + 2^{-25}$

*Resposta.*  $\square$

e  $2^{130}$ ?

*Resposta.*  $\square$