



[Cod: CM-334 Curso: Análisis Numérico I]

[Tema: Aritmética del computador]

Laboratorio N° 1

1. Realice las siguientes operaciones con una mantisa de 4 bits, considerando la siguiente tabla y de la respuesta en base decimal y binaria.

Mantisa	$n = -3$	$n = -2$	$n = -1$	$n = 0$	$n = 1$	$n = 2$	$n = 3$	$n = 4$
0.1000 ₍₂₎	0.0625	0.125	0.25	0.5	1	2	4	8
0.1001 ₍₂₎	0.0703125	0.140625	0.28125	0.5625	1.125	2.25	4.5	9
0.1010 ₍₂₎	0.078125	0.15625	0.3125	0.625	1.25	2.5	5	10
0.1011 ₍₂₎	0.0859375	0.171875	0.34375	0.6875	1.375	2.75	5.5	11
0.1100 ₍₂₎	0.09375	0.1875	0.375	0.75	1.5	3	6	12
0.1101 ₍₂₎	0.1015625	0.203125	0.40625	0.8125	1.625	3.25	6.5	13
0.1110 ₍₂₎	0.109375	0.21875	0.4375	0.875	1.75	3.5	7	14
0.1111 ₍₂₎	0.1171875	0.234375	0.46875	0.9375	1.875	3.75	7.5	15

(a) $\left(\frac{1}{3} + \frac{1}{5}\right) + \frac{1}{6}$

(b) $\left(\frac{1}{10} + \frac{1}{3}\right) + \frac{1}{5}$

(c) $\left(\frac{3}{17} + \frac{1}{9}\right) + \frac{1}{7}$

2. Suponga que x y y son vectores de punto flotante y el producto $f(x, y) = \sum_{i=1}^n x_i y_i$ es calculado usando aritmética de punto flotante. Muestre que el error en el cálculo realizado es

$$|\bar{f}(x, y) - f(x, y)| \leq n\epsilon|x|^T|y| + O(\epsilon^2)$$

3. Evalúe la función

$$y_1(x) = \frac{\log(1+x)}{x}$$

para $x \approx 0$ usando doble precisión. Grafique la función en el intervalo $[-10^{-15}, 10^{-15}]$. Repita el experimento usando

$$y_2(x) = \begin{cases} \frac{\log(1+x)}{(1+x)-1} & \text{si } 1+x \neq 1 \\ 1 & \text{si } 1+x = 1 \end{cases}$$

4. En una máquina binaria ¿cuál es el error de redondeo unitario para mantisas de 48 bits?
5. Calcule un valor aproximado del epsilon de la máquina usando el algoritmo 1
6. Demuestre que $4/5$ no se puede representar de manera exacta como número de máquina. ¿Cuál será el número de máquina más cercano?. ¿Cuál será el error de redondeo relativo que se produce cuando se almacena internamente este número?

Algorithm 1 Epsilon de máquina

```
1: desde  $k \leftarrow 1$  hasta 100 hacer
2:    $s \leftarrow 0.5 * s$ 
3:    $t \leftarrow s + 1.0$ 
4:   si  $t \leq 1.0$  entonces
5:      $s \leftarrow 2.0 * s$ 
6:     Escribir  $k - 1, s$ 
7:     Salir
8:   fin si
9: fin desde
```

7. Muestre ejemplos de que es posible que $\text{fl}[\text{fl}(xy)z] \neq \text{fl}[x\text{fl}(yz)]$ para números de máquina x, y, z .
8. Demuestre que si x, y son números de máquina de 32 bits y $|y| \leq |x|2^{-25}$ entonces $\text{fl}(x + y) = x$.
9. Encuentre el número de máquina de 32 bits que está a la derecha de $1/9$.
10. Determine una cota para el error relativo que surge al calcular $(a + b)/(c + d)$ para números de máquina a, b, c, d .
11. ¿Cuántos bits de precisión se pierden en una computadora cuando efectuamos la sustracción $x - \sin x$ para $x = \frac{1}{2}$?
12. ¿Cuántos bits de precisión se pierden en una computadora cuando efectuamos la sustracción $1 - \cos x$ para $x = \frac{1}{4}$?
13. Escriba un programa calcular

$$f(x) = \sqrt{x^2 + 1} - 1$$

$$g(x) = x^2/(\sqrt{x^2 + 1} + 1)$$

para una sucesión de valores de x como: $8^{-1}, 8^{-2}, 8^{-3}, \dots$ ¿Los resultados son iguales?

14. Analice el problema de calcular $\sinh x$ a partir de su definición: $\sinh x \equiv \frac{1}{2}(e^x - e^{-x})$.
15. Si $x_0 > -1$ entonces la sucesión

$$x_{n+1} = 2^{n+1} \left[\sqrt{1 + 2^{-n}x_n} - 1 \right]$$

converge a $\ln(x_0 + 1)$. Modifique esta fórmula de tal manera que no haya pérdida de dígitos significativos.

16. Diseñe un programa que imprima los valores de las siguientes funciones

$$f(x) = x^8 - 8x^7 + 28x^6 - 56x^5 + 70x^4 - 56x^3 + 28x^2 - 8x + 1$$

$$g(x) = ((((((x - 8)x + 28)x - 56)x + 70)x - 56)x + 28)x - 8)x + 1$$

$$h(x) = (x - 1)^8$$

en 101 puntos igualmente espaciados cubriendo el intervalo $[0.99, 1.01]$. Analice los resultados.

Uni, 21 de marzo de 2019*

*Hecho en L^AT_EX