



Profesor: Gonzalo Hernández.
Auxiliar: Constanza Maturana, Tomás Gonzalez
Fecha: 19 de Marzo

Auxiliar 1: Representación Numérica y Errores

Resumen Materia

- Representación de números enteros:** $p = \sum_{i=0}^n \alpha_i b^i$ donde $b = 2, 8, 10$ y $\alpha_i \in \{0, \dots, b-1\}$
- Representación de números reales:** $x = \pm m \cdot b^z$ donde $b \in \mathbb{N}$, $z \in \mathbb{Z}$, $m \in [0, 1] \cap \mathbb{R}$
- Representación de punto flotante:** $x = \pm 0.m_1 m_2 \dots m_t \cdot b^z$ donde $b \in \mathbb{N}$, $t \in \mathbb{N} \setminus \{\infty\}$, $z \in \mathbb{Z}$, $m_i \in \{0, \dots, b-1\}$
- Representación de números reales en $[0, 1]$:** $p = \sum_{i=0}^n \alpha_i 2^{-i}$ donde $\alpha_i \in \{0, 1\}$
- Representación de punto flotante actual:** $x = (-1)^s 2^{(z-c)} (1 + m)$ donde $s \in \{0, 1\}$, $c = 1023$, $z \in \{0, 1, \dots, 2^{11} - 1\}$, $m = m_1 m_2 \dots m_{52}$ con $m_i \in \{0, 1\}$
- Error Absoluto:** $E_A(x, x^*) = |x - x^*|$
- Error Relativo:** $E_R(x, x^*) = \frac{|x - x^*|}{|x|}$
- Relación entre punto flotante y un real:** $fl(x) = x(1 + \delta)$ donde δ puede ser cero, negativo o positivo. δ es el error relativo y se cumple que $|\delta| \leq 5 \times 10^{-t}$ donde t es la cantidad de dígitos significativos. El valor $eps = 5 \times 10^{-t}$ es la precisión de la máquina.
- Propagación de errores:** $fl(x \otimes y) = (x \otimes y)(1 + \varepsilon)$ para $\otimes = +, -, *, /$
 - Error suma: $\varepsilon_s = \frac{x}{x+y} \varepsilon_x + \frac{y}{x+y} \varepsilon_y$
 - Error resta: $\varepsilon_r = \frac{x}{x-y} \varepsilon_x - \frac{y}{x-y} \varepsilon_y$
 - Error multiplicación: $\varepsilon_m = \varepsilon_x + \varepsilon_y$
 - Error división: $\varepsilon_d = \varepsilon_x - \varepsilon_y$
 - Error general: $\varepsilon_\Phi = \frac{x}{\Phi(x,y)} \frac{\partial \Phi}{\partial x} \varepsilon_x + \frac{y}{\Phi(x,y)} \frac{\partial \Phi}{\partial y} \varepsilon_y$
- Estabilidad:** si los llamados números de condicionamiento $\mu_x = \frac{x}{\Phi(x,y)} \frac{\partial \Phi}{\partial x}$ y $\mu_y = \frac{y}{\Phi(x,y)} \frac{\partial \Phi}{\partial y}$ se pueden acotar por constantes, entonces la secuencia de operaciones matemáticas es estable. En caso contrario, es inestable.

Ejercicios

- Considere una modificación de la codificación floating point inicial de 20 bits con los siguientes parámetros: 1 bit signo, 1 bit signo exponente, 7 bits exponente y 11 bits para la mantisa.
 - De acuerdo a esta codificación floating point determine el número real representado según: 01101010111110001111
 - De acuerdo a esta codificación floating point determine la codificación del número real: 56.715
 - De acuerdo a esta codificación floating point determine el número real máximo, mínimo y más cercano a cero
- Realice un análisis de primer orden para determinar la propagación de errores generados por las siguientes operaciones matemáticas aplicadas a dos reales x e y :
 - $\phi(x, y) = e^{xy}$
 - $\phi(x, y) = \ln(x^2 + y^2)$
 - $\phi(x, y) = x^2 y^2$

(d) $\phi(x, y) = \sqrt{\frac{x}{y}}$

Respuestas

1. (a)

Signo	Mantisa	Signo	Exponente	Exponente	Mantisa
0		1		1010101	11110001111

Signo Mantisa: +
 Signo Exponente: -
 Exponente: $1 \times 2^6 + 0 \times 2^5 + 1 \times 2^4 + 0 \times 2^3 + 1 \times 2^2 + 0 \times 2^1 + 1 \times 2^0 = 85$
 Mantisa: $1 \times 2^{-1} + 1 \times 2^{-2} + 1 \times 2^{-3} + 1 \times 2^{-4} + 0 \times 2^{-5} + 0 \times 2^{-6} + 0 \times 2^{-7} + 1 \times 2^{-8} + 1 \times 2^{-9} + 1 \times 2^{-10} + 1 \times 2^{-11} = 0.94482421875$
 $\Rightarrow$ Número: $+0.94482421875 \times 2^{-85} = 2.4423133625642525496 \times 10^{-26}$

(b) $56.715 = 0.56715 \times 10^2 = 0.56715 \times 10^2 \times \frac{2^6}{2^6} = 0.56715 \times \frac{10^2}{2^6} \times 2^6 = 0.56715 \times 1.5625 \times 2^6 = 0.886171875 \times 2^6$
 Signo Mantisa: +
 Signo Exponente: +

		resto		
Exponente:	6 : 2 = 3	0	↑	$\Rightarrow 0000110$
	3 : 2 = 1	1		
	1 : 2 = 0	1		

Mantisa:
 $m = 0.886171875 \Rightarrow m \cdot 2^{11} \approx 1814$

		Resto		
1814 : 2 = 907	0			$\uparrow \Rightarrow 11100010110$
907 : 2 = 453	1			
453 : 2 = 226	1			
226 : 2 = 113	0			
113 : 2 = 56	1			
56 : 2 = 28	0			
28 : 2 = 14	0			
14 : 2 = 7	0			
7 : 2 = 3	1			
3 : 2 = 1	1			
1 : 2 = 0	1			

$\Rightarrow$

Signo	Mantisa	Signo	Exponente	Exponente	Mantisa
0		0		0000110	11100010110

(c) i. Real Máximo: $0\ 0\ 1111111\ 11111111111 = \frac{2^{11}-1}{2^{11}} \times 2^{127} = 0.99951171875 \times 2^{127}$
 ii. Real Mínimo: $1\ 0\ 1111111\ 11111111111 = -\frac{2^{11}-1}{2^{11}} \times 2^{127} = -0.99951171875 \times 2^{127}$
 iii. Real Más cercano a cero positivo: $0\ 1\ 1111111\ 00000000001 = 2^{-11} \times 2^{-127} = 2^{-138} = 2.869859254937225361 \times 10^{-42}$
 iv. Real Más cercano a cero negativo: $1\ 1\ 1111111\ 00000000001 = -2^{-11} \times 2^{-127} = -2^{-138} = -2.8698592549372253613 \times 10^{-42}$

 2. (a) $\varepsilon_\Phi = \frac{x}{\Phi(x,y)} \frac{\partial \Phi}{\partial x} \varepsilon_x + \frac{y}{\Phi(x,y)} \frac{\partial \Phi}{\partial y} \varepsilon_y = \frac{x}{e^{xy}} y e^{xy} \varepsilon_x + \frac{y}{e^{xy}} x e^{xy} \varepsilon_y = xy (\varepsilon_x + \varepsilon_y)$ inestable
 - (b) $\varepsilon_\Phi = \frac{x}{\ln(x^2+y^2)} \frac{2x}{x^2+y^2} \varepsilon_x + \frac{y}{\ln(x^2+y^2)} \frac{2y}{x^2+y^2} \varepsilon_y = 2 \frac{x^2 \varepsilon_x + y^2 \varepsilon_y}{(x^2+y^2) \ln(x^2+y^2)}$ inestable
 - (c) $\varepsilon_\Phi = \frac{x}{x^2 y^2} 2xy^2 \varepsilon_x + \frac{y}{x^2 y^2} 2x^2 y \varepsilon_y = 2 (\varepsilon_x + \varepsilon_y)$ estable
 - (d) $\varepsilon_\Phi = \frac{x}{\sqrt{\frac{x}{y}}} \frac{1}{2y\sqrt{\frac{x}{y}}} \varepsilon_x + \frac{y}{\sqrt{\frac{x}{y}}} \left(-\frac{1}{2} \frac{x}{y^2 \sqrt{\frac{x}{y}}} \right) \varepsilon_y = \frac{1}{2} (\varepsilon_x - \varepsilon_y)$ estable