TALLER PUNTO FLOTANTE

Criterios de evaluación

- Reúnanse por parejas para resolver el problema que le han asignado
- Entregue un informe en Word, utilizando el editor de ecuaciones (nunca a mano)
- Devuelva su solución en la fecha acordada en horas de clase.
- Tenga en cuenta los valores de cada numeral para su resultado.

Ejemplo: sume y multiplique los números 0.5 y –0.4375 en binario, suponiendo 4 bits de precisión y escriba la respuesta en decimal con 4 cifras significativas.

Solución

Para sumar, primero obtendremos la versión binaria normalizada, suponiendo 4 bits de precisión:

$0.5 = 1/2 = 1/2^1 = 0.1 = 0.1 \times 2^0 = 1.000 \times 2^{-1}$

$- 0.4375 = - 7/16 = - 7/2^4 = - 0.0111 = 0.0111 \times 2^0 = - 1.110 \times 2^{-2}$

Ahora siguiendo el algoritmo de la suma:

1) Se desplaza al número con exponente más pequeño (a la derecha), hasta alinearlo con el exponente mayor: $- 1.110 \times 2^{-2} = - 0.111 \times 2^{-1}$

2) Se suman las mantisas: $1.000_2 \times 2^{-1} + (- 0.111_2 \times 2^{-1}) = 0.001_2 \times 2^{-1}$

3) Se normaliza la suma, verificando si existe sobre flujo o bajo flujo:

$0.001_2 \times 2^{-1} = 0.010_2 \times 2^{-2} = 0.100_2 \times 2^{-3} = 1.000_2 \times 2^{-4}$

Puesto que $-126 < - 4 < 127$, no hay sobre flujo ni bajo flujo (El exponente desplazado sería $– 4 + 127 = 123$, y está entre 1 y 254).

4) Redondeo de la suma: $1.000_2 \times 2^{-4}$

Ya está redondeada, por lo que el resultado es:

$\mathbf{1.000_2 \times 2^{-4}} = 0.0001000_2 = 1/2^4 = 1/16 = \mathbf{0.0625 = 6.250 \times 10^{-2}}$

Para multiplicar esos números en binario, suponiendo los 4 bits de precisión, se procederá como sigue:

Primero obtendremos la versión binaria normalizada, suponiendo 4 bits de precisión en cada número: $0.5 = 0.1_2 = 1.000_2 \times 2^{-1}$ y $- 0.4375 = - 0.0111_2 = - 1.110_2 \times 2^{-2}$.

Siguiendo el algoritmo de la multiplicación se tiene:

1) Se suman los exponentes: $-1 + (-2) = -3$

2) Se multiplican las mantisas: $1.000_2 \times 1.110_2 = 1110000_2$. El producto es $1.110000_2 \times 2^{-3}$, pero necesitamos mantenerlo en 4 bits por lo que obtenemos: $1.110_2 \times 2^{-3}$

3) El producto está normalizado y no hay error de sobre flujo (overflow) o bajo flujo (underflow).

4) El redondeo no cambia el producto.

5) Los signos de los operandos son diferentes, de manera que el resultado es:

$$\mathbf{- 1.110_2 \times 2^{-3} = 2.1875 \times 10^{-1}}$$

Problema 1:

1. **(valor 100/3 puntos)** Exprese en decimal el número $E643E9E0_H$ del estándar IEEE 754 de precisión simple, con 8 cifras significativas.
2. **(valor 100/3 puntos)** Se tiene un número positivo, que transformado en formato estándar IEEE 754 de precisión simple es $1.01110010101111111010100_2 x 2^{128}$. Determine el número con 7 cifras significativas, si existe.
3. **(valor 100/3 puntos)** sume los números: $-8800.390625 x 10^{-2}$ y $0.67002009375 x 10^{2}$ dado que se pueden almacenar 6 dígitos en la mantisa y hasta 2 dígitos en el exponente. Determine el signo, el exponente, la mantisa y el valor en base 10, del resultado según la tabla del formato estándar IEEE 754 de precisión simple.

Problema 2:

1. **(valor 100/3 puntos)** Exprese en decimal el número

   $$10111101011001010111111101010000_2$$

   del estándar IEEE 754 de precisión simple, con 7 cifras significativas.
2. **(valor 100/3 puntos)** Si un número es negativo y luego de transformarse en formato estándar IEEE 754 con precisión doble es

   $$1.1011100101000000101010101010100110100100010001011111_2 x 2^{1024},$$

   determine el número con de 7 cifras significativas, si existe.

3. **(valor 100/3 puntos)** Multiplique los números: $-8800.390625 x 10^{-2}$ y $0.67002009375 x 10^{2}$ dado que se pueden almacenar 6 dígitos en la mantisa y hasta 2 dígitos en el exponente. Determine el signo, el exponente, la mantisa y el valor en base 10, del resultado según la tabla del formato estándar IEEE 754 de precisión simple.