Aula AOC - 24/08

segunda-feira, 24 de agosto de 2020 08:53

# Aula AOC 24/08

- Aritmética de ponto flutuante:
  - ↳ Notação de ponto fixo (= ponto da notação decimal)
  - ↳ Notação de ponto flutuante

- O problema do ponto fixo é que o tamanho da parte inteira e da fracionária fica fixo com relação a seu armazenamento em memória.

- O ponto flutuante se move dinamicamente para uma posição conveniente usando a notação científica.

$$\pm M \cdot B^{\pm E}$$

Sinais — Mantissa — base (binária - 2) — Expoente

| 1 bit | 8 bits | 23 bits |
|---|---|---|
| Bit do sinal | Expoente ↑ Posição do ponto | Significando |

- **A mantissa é armazenada em complemento-dois**
- **O expoente é representado na notação *excess* ou *biased* (*polarizado*)**

  **Exemplo:**
  - **Excess (bias) 128 significa**
  - **Campo expoente de 8 bit**
  - **Faixa de valores puros 0-255**
  - **Subtraia 128 para obter o valor correto**
  - **Nova faixa -128 to +127**

## • Números Fracionários

9.81 → 9 − 1001 . 1100 ...

| | | | | |
|---|---|---|---|---|
| | | ×2 | | se ≥1 = 1 / else = 0 |
| | 0.81 | → | 1,62 | → 1 |
| ↑ | 0,62 | → | 1,24 | → 1 |
| | 0,24 | → | 0,48 | → 0 |
| | 0,48 | → | 0,96 | → 0 |

⋮

| | Sign | Biased exponent | Fraction | Value |
|---|---|---|---|---|
| positive zero | 0 | 0 | 0 | 0 |
| negative zero | 1 | 0 | 0 | −0 |
| plus infinity | 0 | all 1s | 0 | ∞ |
| minus infinity | 1 | all 1s | 0 | −∞ |
| quiet NaN | 0 or 1 | all 1s | ≠ 0; first bit = 1 | qNaN |
| signaling NaN | 0 or 1 | all 1s | ≠ 0; first bit = 0 | sNaN |
| positive normal nonzero | 0 | all 1s | f | $2_{e-16383}(1.f)$ |
| negative normal nonzero | 1 | all 1s | f | $-2_{e-16383}(1.f)$ |
| positive subnormal | 0 | 0 | f ≠ 0 | $2_{e-16383}(0.f)$ |
| negative subnormal | 1 | 0 | f ≠ 0 | $-2_{e-16383}(0.f)$ |

Continuação 25/08

## • Operações aritméticas c/ ponto flutuante

$X = 0,3 \cdot 10^2 = 30$

$Y = 0,2 \cdot 10^3 = 2000$

$X + Y = (0,3 \cdot 10^{2-3} + 0,2) \cdot 10^3 = 0,23 \cdot 10^3 = 230$

$X - Y = (0,3 \cdot 10^{2-3} - 0,2) \cdot 10^3 = -0,17 \cdot 10^3 = -170$

$X \cdot Y = (0,3 \cdot 0,2) \cdot 10^{2+3} = 0,06 \cdot 10^5 = 6000$

$X \div Y = (0,3 \div 0,2) \cdot 10^{2-3} = 1,5 \cdot 10^{-1} = 0,15$

- Fases importantes:
  - Verificar se é zero
  - Ajustar os expoentes
  - Realizar a operação
  - Normalizar o resultado

## • Bits de GUARDA

- Bits dentro das arquiteturas que têm como função principal melhorar as aproximações do processo aritmético de ponto flutuante.

## • Arredondamento

- Round to nearest
- Round towards $+\infty / -\infty$
- Round towards 0

- Infinito
  - O infinito é tratado como o caso limite da aritmética real.

- NaN
  - Quiet - Erro sem aviso
  - Signaling - Erro c/ aviso

Casos QUIET:

Somas e subtrações c/ $\infty$

Multiplicar $0 \cdot \infty$

Dividir $\frac{0}{0}$ ou $\frac{\infty}{\infty}$

Raiz de números negativos