Inf250 - Aula de Ponto Flutuante

Depto de Informática - UFV

Ricardo Ferreira
ricardo@ufv.br

2024

Plano de Aula

- Float IEEE 754 com 32 bits
- Float de "7 bits"didático...
  - Conversão
  - Adição
  - Multiplicação
  - Verilog

Plano de Aula

- Float IEEE 754 com 32 bits

- Float de "7 bits"didático...
  - Conversão
  - Adição
  - Multiplicação
  - Verilog

Plano de Aula

- Float IEEE 754 com 32 bits

- Float de "7 bits"didático...
  - Conversão
  - Adição
  - Multiplicação
  - Verilog

Plano de Aula

- Float IEEE 754 com 32 bits
- Float de "7 bits"didático. . .
  - Conversão
  - Adição
  - Multiplicação
  - Verilog

Plano de Aula

- Float IEEE 754 com 32 bits

- Float de "7 bits"didático...
  - Conversão
  - Adição
  - Multiplicação
  - Verilog

Plano de Aula

- Float IEEE 754 com 32 bits

- Float de "7 bits"didático...
  - Conversão
  - Adição
  - Multiplicação
  - Verilog

Representação de Números em C

- **unsigned int x = 5;**:
  - Binário (32 bits):

    $$00000000\ 00000000\ 00000000\ 00000101_2$$

  - Hexadecimal: 0x00000005
- **int x = -5;**:
  - Binário (32 bits - complemento de dois):

    $$11111111\ 11111111\ 11111111\ 11111011_2$$

  - Hexadecimal: 0xFFFFFFFB
- **float x=5.0;**:
  - Representação IEEE 754 (32 bits):

    $$0\ \boxed{1000000\ 1}\ 0100000\ 00000000\ 00000000_2$$

  - Hexadecimal:

    $$0x40A00000$$

Representação de Números em C

- **unsigned int x = 5;**:
  - Binário (32 bits):

$$00000000\ 00000000\ 00000000\ 00000101_2$$

  - Hexadecimal: 0x00000005
- **int x = -5;**:
  - Binário (32 bits - complemento de dois):

$$11111111\ 11111111\ 11111111\ 11111011_2$$

  - Hexadecimal: 0xFFFFFFFB
- **float x=5.0;**:
  - Representação IEEE 754 (32 bits):

$$0\ \boxed{1000000\ 1}\ 0100000\ 00000000\ 00000000_2$$

  - Hexadecimal:

0x40A00000

Representação de Números em C

- **unsigned int x = 5;**:
  - Binário (32 bits):

    $$00000000\ 00000000\ 00000000\ 00000101_2$$

  - Hexadecimal: 0x00000005
- **int x = -5;**:
  - Binário (32 bits - complemento de dois):

    $$11111111\ 11111111\ 11111111\ 11111011_2$$

  - Hexadecimal: 0xFFFFFFFB
- **float x=5.0;**:
  - Representação IEEE 754 (32 bits):

    $$0\ \boxed{1000000\ 1}\ 0100000\ 00000000\ 00000000_2$$

  - Hexadecimal:

    $$0x40A00000$$

Representação de Números em C

- **unsigned int x = 5;**:
  - Binário (32 bits):

    $$00000000\ 00000000\ 00000000\ 00000101_2$$

  - Hexadecimal: 0x00000005
- **int x = -5;**:
  - Binário (32 bits - complemento de dois):

    $$11111111\ 11111111\ 11111111\ 11111011_2$$

  - Hexadecimal: 0xFFFFFFFB
- **float x=5.0;**:
  - Representação IEEE 754 (32 bits):

    $$0\ \boxed{1000000\ 1}\ 0100000\ 00000000\ 00000000_2$$

  - Hexadecimal:

    0x40A00000

Representação de Números em C

- **unsigned int x = 5;**:
  - Binário (32 bits):

$$00000000\ 00000000\ 00000000\ 00000101_2$$

  - Hexadecimal: 0x00000005
- **int x = -5;**:
  - Binário (32 bits - complemento de dois):

$$11111111\ 11111111\ 11111111\ 11111011_2$$

  - Hexadecimal: 0xFFFFFFFB
- **float x=5.0;**:
  - Representação IEEE 754 (32 bits):

$$0\ \boxed{1000000\ 1}\ 0100000\ 00000000\ 00000000_2$$

  - Hexadecimal:

$$0x40A00000$$

Representação de Números em C

- **unsigned int x = 5;**:
  - Binário (32 bits):

$$00000000\ 00000000\ 00000000\ 00000101_2$$

  - Hexadecimal: 0x00000005
- **int x = -5;**:
  - Binário (32 bits - complemento de dois):

$$11111111\ 11111111\ 11111111\ 11111011_2$$

  - Hexadecimal: 0xFFFFFFFB
- **float x=5.0;**:
  - Representação IEEE 754 (32 bits):

$$0\ \boxed{1000000\ 1}\ 0100000\ 00000000\ 00000000_2$$

  - Hexadecimal:

$$0x40A00000$$

Representação de Números em C

- **unsigned int x = 5;**:
  - Binário (32 bits):

    $$00000000\ 00000000\ 00000000\ 00000101_2$$

  - Hexadecimal: 0x00000005
- **int x = -5;**:
  - Binário (32 bits - complemento de dois):

    $$11111111\ 11111111\ 11111111\ 11111011_2$$

  - Hexadecimal: 0xFFFFFFFB
- **float x=5.0;**:
  - Representação IEEE 754 (32 bits):

    $$0\ \boxed{1000000\ 1}\ 0100000\ 00000000\ 00000000_2$$

  - Hexadecimal:

    $$0x40A00000$$

Representação de Números em C

- **unsigned int x = 5;**:
  - Binário (32 bits):

$$00000000\ 00000000\ 00000000\ 00000101_2$$

  - Hexadecimal: 0x00000005
- **int x = -5;**:
  - Binário (32 bits - complemento de dois):

$$11111111\ 11111111\ 11111111\ 11111011_2$$

  - Hexadecimal: 0xFFFFFFFB
- **float x=5.0;**:
  - Representação IEEE 754 (32 bits):

$$0\boxed{1000000\ 1}0100000\ 00000000\ 00000000_2$$

  - Hexadecimal:

$$0x40A00000$$

Representação de Números em C

- **unsigned int x = 5;**:
  - Binário (32 bits):

$$00000000\ 00000000\ 00000000\ 00000101_2$$

  - Hexadecimal: 0x00000005
- **int x = -5;**:
  - Binário (32 bits - complemento de dois):

$$11111111\ 11111111\ 11111111\ 11111011_2$$

  - Hexadecimal: 0xFFFFFFFB
- **float x=5.0;**:
  - Representação IEEE 754 (32 bits):

$$0\boxed{1000000\ 1}0100000\ 00000000\ 00000000_2$$

  - Hexadecimal:

$$0x40A00000$$

Formato IEEE 754 (32 bits)

Representação de um número em ponto flutuante:

- **Estrutura:** $\underbrace{S}_{\text{Sinal}} \; \underbrace{E_7 E_6 E_5 E_4 E_3 E_2 E_1 E_0}_{\text{Expoente (8 bits)}} \; \underbrace{M_{22} M_{21} \ldots M_0}_{\text{Mantissa (23 bits)}}$
- **Fórmula para codificação:**

$$(-1)^S \times 2^{(E-127)} \times (1 + M)$$

Onde:

  - S: Bit de sinal (0 para positivo, 1 para negativo)
  - E: Valor do expoente com viés (bias) de 127
  - M: Mantissa (fração normalizada)

- **Exemplo - Número 5.0 em IEEE 754 (32 bits):**

$$\underbrace{0}_{S} \; \underbrace{10000001}_{E} \; \underbrace{01000000000000000000000}_{M}$$

Formato IEEE 754 (32 bits)

Representação de um número em ponto flutuante:

- **Estrutura:** $\underbrace{S}_{\text{Sinal}} \; \underbrace{E_7 E_6 E_5 E_4 E_3 E_2 E_1 E_0}_{\text{Expoente (8 bits)}} \; \underbrace{M_{22} M_{21} \dots M_0}_{\text{Mantissa (23 bits)}}$
- **Fórmula para codificação:**

$$(-1)^S \times 2^{(E-127)} \times (1 + M)$$

Onde:

- S: Bit de sinal (0 para positivo, 1 para negativo)
- E: Valor do expoente com viés (bias) de 127
- M: Mantissa (fração normalizada)

- **Exemplo - Número 5.0 em IEEE 754 (32 bits):**

$$\underbrace{0}_{S} \; \underbrace{10000001}_{E} \; \underbrace{01000000000000000000000}_{M}$$

Formato IEEE 754 (32 bits)

Representação de um número em ponto flutuante:

- **Estrutura:** $\underbrace{S}_{\text{Sinal}}\ \underbrace{E_7 E_6 E_5 E_4 E_3 E_2 E_1 E_0}_{\text{Expoente (8 bits)}}\ \underbrace{M_{22} M_{21} \ldots M_0}_{\text{Mantissa (23 bits)}}$
- **Fórmula para codificação:**

$$(-1)^S \times 2^{(E-127)} \times (1 + M)$$

Onde:

- S: Bit de sinal (0 para positivo, 1 para negativo)
- E: Valor do expoente com viés (bias) de 127
- M: Mantissa (fração normalizada)

- **Exemplo - Número 5.0 em IEEE 754 (32 bits):**

$$\underbrace{0}_{S}\ \underbrace{10000001}_{E}\ \underbrace{01000000000000000000000}_{M}$$

Formato IEEE 754 (32 bits)

Representação de um número em ponto flutuante:

- **Estrutura:** $\underbrace{S}_{\text{Sinal}} \; \underbrace{E_7 E_6 E_5 E_4 E_3 E_2 E_1 E_0}_{\text{Expoente (8 bits)}} \; \underbrace{M_{22} M_{21} \ldots M_0}_{\text{Mantissa (23 bits)}}$
- **Fórmula para codificação:**

$$(-1)^S \times 2^{(E-127)} \times (1+M)$$

  Onde:
  - S: Bit de sinal (0 para positivo, 1 para negativo)
  - E: Valor do expoente com viés (bias) de 127
  - M: Mantissa (fração normalizada)
- **Exemplo - Número 5.0 em IEEE 754 (32 bits):**

$$\underbrace{0}_{S} \; \underbrace{10000001}_{E} \; \underbrace{01000000000000000000000}_{M}$$

Formato IEEE 754 (32 bits)

Representação de um número em ponto flutuante:

- **Estrutura:** $\underbrace{S}_{\text{Sinal}} \; \underbrace{E_7 E_6 E_5 E_4 E_3 E_2 E_1 E_0}_{\text{Expoente (8 bits)}} \; \underbrace{M_{22} M_{21} \ldots M_0}_{\text{Mantissa (23 bits)}}$
- **Fórmula para codificação:**

$$(-1)^S \times 2^{(E-127)} \times (1 + M)$$

Onde:

- S: Bit de sinal (0 para positivo, 1 para negativo)
- E: Valor do expoente com viés (bias) de 127
- M: Mantissa (fração normalizada)

- **Exemplo - Número 5.0 em IEEE 754 (32 bits):**

$$\underbrace{0}_{S} \; \underbrace{10000001}_{E} \; \underbrace{01000000000000000000000}_{M}$$

Formato IEEE 754 (32 bits)

Representação de um número em ponto flutuante:

- **Estrutura:** $\underbrace{S}_{\text{Sinal}}\ \underbrace{E_7 E_6 E_5 E_4 E_3 E_2 E_1 E_0}_{\text{Expoente (8 bits)}}\ \underbrace{M_{22} M_{21} \dots M_0}_{\text{Mantissa (23 bits)}}$
- **Fórmula para codificação:**

$$(-1)^S \times 2^{(E-127)} \times (1 + M)$$

Onde:

- S: Bit de sinal (0 para positivo, 1 para negativo)
- E: Valor do expoente com viés (bias) de 127
- M: Mantissa (fração normalizada)

- **Exemplo - Número 5.0 em IEEE 754 (32 bits):**

$$\underbrace{0}_{S}\ \underbrace{10000001}_{E}\ \underbrace{01000000000000000000000}_{M}$$

Exceções no Formato IEEE 754 (32 bits)

Exceções Especiais:

- **NaN (Not a Number):** $S = 0 \quad E = 11111111 \quad M \neq 0$ Indica um valor indefinido ou inválido, como o resultado de $0/0$ ou $\sqrt{-1}$.
- **Infinito (∞):** $S = 0 \quad E = 11111111 \quad M = 0$ Representa um valor muito grande, como o resultado de $1/0$ (infinito positivo) ou $-1/0$ (infinito negativo).
- **Zero (0):** $S = 0 \quad E = 00000000 \quad M = 0$ Valor zero, pode ser tanto $+0$ (positivo) quanto -0 (negativo), dependendo do bit de sinal.
- **Zero Subnormal:**

$$S = 0 \quad E = 00000000 \quad M \neq 0$$

  Representa números muito próximos de zero, mas que ainda possuem precisão limitada.

Exceções no Formato IEEE 754 (32 bits)

Exceções Especiais:

- **NaN (Not a Number):** $S = 0 \quad E = 11111111 \quad M \neq 0$ Indica um valor indefinido ou inválido, como o resultado de $0/0$ ou $\sqrt{-1}$.
- **Infinito (∞):** $S = 0 \quad E = 11111111 \quad M = 0$ Representa um valor muito grande, como o resultado de $1/0$ (infinito positivo) ou $-1/0$ (infinito negativo).
- **Zero (0):** $S = 0 \quad E = 00000000 \quad M = 0$ Valor zero, pode ser tanto $+0$ (positivo) quanto -0 (negativo), dependendo do bit de sinal.
- **Zero Subnormal:**

$$S = 0 \quad E = 00000000 \quad M \neq 0$$

Representa números muito próximos de zero, mas que ainda possuem precisão limitada.

Exceções no Formato IEEE 754 (32 bits)

Exceções Especiais:

- **NaN (Not a Number):** $S = 0 \quad E = 11111111 \quad M \neq 0$ Indica um valor indefinido ou inválido, como o resultado de $0/0$ ou $\sqrt{-1}$.
- **Infinito (∞):** $S = 0 \quad E = 11111111 \quad M = 0$ Representa um valor muito grande, como o resultado de $1/0$ (infinito positivo) ou $-1/0$ (infinito negativo).
- **Zero (0):** $S = 0 \quad E = 00000000 \quad M = 0$ Valor zero, pode ser tanto $+0$ (positivo) quanto -0 (negativo), dependendo do bit de sinal.
- **Zero Subnormal:**

$$S = 0 \quad E = 00000000 \quad M \neq 0$$

Representa números muito próximos de zero, mas que ainda possuem precisão limitada.

Exceções no Formato IEEE 754 (32 bits)

Exceções Especiais:

- **NaN (Not a Number):** $S = 0 \quad E = 11111111 \quad M \neq 0$ Indica um valor indefinido ou inválido, como o resultado de $0/0$ ou $\sqrt{-1}$.
- **Infinito (∞):** $S = 0 \quad E = 11111111 \quad M = 0$ Representa um valor muito grande, como o resultado de $1/0$ (infinito positivo) ou $-1/0$ (infinito negativo).
- **Zero (0):** $S = 0 \quad E = 00000000 \quad M = 0$ Valor zero, pode ser tanto $+0$ (positivo) quanto -0 (negativo), dependendo do bit de sinal.
- **Zero Subnormal:**

$$S = 0 \quad E = 00000000 \quad M \neq 0$$

Representa números muito próximos de zero, mas que ainda possuem precisão limitada.

Representação Simplificada IEEE 754

Formato Simplificado:

- Considere uma representação com **3 bits de expoente** e **4 bits de mantissa**.
- A fórmula geral para representar o número n é:

$$n = \boxed{e_2}\boxed{e_1}\boxed{e_0}\boxed{M_3}\boxed{M_2}\boxed{M_1}\boxed{M_0}$$

- Onde:

$$n = 2^{e-3} \times (1 + M)$$

- e_2, e_1, e_0 representam o expoente (e) com um **bias** de 3.
- M_3, M_2, M_1, M_0 representam a mantissa $M = \frac{1}{2}M_3 + \frac{1}{4}M_2 + \frac{1}{8}M_1 + \frac{1}{16}M_0$.

Representação Simplificada IEEE 754

Formato Simplificado:

- Considere uma representação com **3 bits de expoente** e **4 bits de mantissa**.
- A fórmula geral para representar o número n é:

$$n = \boxed{e_2}\boxed{e_1}\boxed{e_0}\boxed{M_3}\boxed{M_2}\boxed{M_1}\boxed{M_0}$$

- Onde:

$$n = 2^{e-3} \times (1 + M)$$

- e_2, e_1, e_0 representam o expoente (e) com um **bias** de 3.
- M_3, M_2, M_1, M_0 representam a mantissa $M = \frac{1}{2}M_3 + \frac{1}{4}M_2 + \frac{1}{8}M_1 + \frac{1}{16}M_0$.

Representação Simplificada IEEE 754

Formato Simplificado:

- Considere uma representação com **3 bits de expoente** e **4 bits de mantissa**.
- A fórmula geral para representar o número n é:

$$n = \boxed{e_2}\boxed{e_1}\boxed{e_0}\boxed{M_3}\boxed{M_2}\boxed{M_1}\boxed{M_0}$$

- Onde:

$$n = 2^{e-3} \times (1 + M)$$

- e_2, e_1, e_0 representam o expoente (e) com um **bias** de 3.
- M_3, M_2, M_1, M_0 representam a mantissa $M = \frac{1}{2}M_3 + \frac{1}{4}M_2 + \frac{1}{8}M_1 + \frac{1}{16}M_0$.

Representação Simplificada IEEE 754

Formato Simplificado:

- Considere uma representação com **3 bits de expoente** e **4 bits de mantissa**.
- A fórmula geral para representar o número n é:

$$n = \boxed{e_2}\boxed{e_1}\boxed{e_0}\boxed{M_3}\boxed{M_2}\boxed{M_1}\boxed{M_0}$$

- Onde:

$$n = 2^{e-3} \times (1 + M)$$

- e_2, e_1, e_0 representam o expoente (e) com um **bias** de 3.
- M_3, M_2, M_1, M_0 representam a mantissa $M = \frac{1}{2}M_3 + \frac{1}{4}M_2 + \frac{1}{8}M_1 + \frac{1}{16}M_0$.

Representação Simplificada IEEE 754

Formato Simplificado:

- Considere uma representação com **3 bits de expoente** e **4 bits de mantissa**.
- A fórmula geral para representar o número n é:

$$n = \boxed{e_2}\boxed{e_1}\boxed{e_0}\boxed{M_3}\boxed{M_2}\boxed{M_1}\boxed{M_0}$$

- Onde:

$$n = 2^{e-3} \times (1 + M)$$

- e_2, e_1, e_0 representam o expoente (e) com um **bias** de 3.
- M_3, M_2, M_1, M_0 representam a mantissa $M = \frac{1}{2}M_3 + \frac{1}{4}M_2 + \frac{1}{8}M_1 + \frac{1}{16}M_0$.

Exemplo de Representação Simplificada: 1.6

Representação do número 1.6 (aproximado para 1.625):

- O número 1.6 está entre 1 e 2, então o expoente $e = 0$, portanto $2^0 = 2^{3-3}$
- Para a mantissa, utilizamos a fração binária 1/2 e 1/8, aproximando para 1.625: $1.6 \approx 1 + \frac{1}{2} + \frac{1}{8} = 1.625$
- A mantissa M é representada como:

$$M_3 M_2 M_1 M_0 = 1010$$

(onde $M_3 = 1/2$ e $M_1 = 1/8$)

- O número 1.6 em nossa representação simplificada:

$$n = \boxed{011}\boxed{1010}$$

$$n = 2^{3-3} \times (1 + 0.625) = 1.625$$

Exemplo de Representação Simplificada: 1.6

Representação do número 1.6 (aproximado para 1.625):

- O número 1.6 está entre 1 e 2, então o expoente $e = 0$, portanto $2^0 = 2^{3-3}$
- Para a mantissa, utilizamos a fração binária 1/2 e 1/8, aproximando para 1.625: $1.6 \approx 1 + \frac{1}{2} + \frac{1}{8} = 1.625$
- A mantissa M é representada como:

$$M_3 M_2 M_1 M_0 = 1010$$

(onde $M_3 = 1/2$ e $M_1 = 1/8$)

- O número 1.6 em nossa representação simplificada:

$$n = \boxed{011}\boxed{1010}$$

$$n = 2^{3-3} \times (1 + 0.625) = 1.625$$

Exemplo de Representação Simplificada: 1.6

Representação do número 1.6 (aproximado para 1.625):

- O número 1.6 está entre 1 e 2, então o expoente $e = 0$, portanto $2^0 = 2^{3-3}$
- Para a mantissa, utilizamos a fração binária 1/2 e 1/8, aproximando para 1.625: $1.6 \approx 1 + \frac{1}{2} + \frac{1}{8} = 1.625$
- A mantissa M é representada como:

$$M_3 M_2 M_1 M_0 = 1010$$

(onde $M_3 = 1/2$ e $M_1 = 1/8$)

- O número 1.6 em nossa representação simplificada:

$$n = \boxed{011}\boxed{1010}$$

$$n = 2^{3-3} \times (1 + 0.625) = 1.625$$

Exemplo de Representação Simplificada: 1.6

Representação do número 1.6 (aproximado para 1.625):

- O número 1.6 está entre 1 e 2, então o expoente $e = 0$, portanto $2^0 = 2^{3-3}$
- Para a mantissa, utilizamos a fração binária $1/2$ e $1/8$, aproximando para 1.625: $1.6 \approx 1 + \frac{1}{2} + \frac{1}{8} = 1.625$
- A mantissa M é representada como:

$$M_3 M_2 M_1 M_0 = 1010$$

(onde $M_3 = 1/2$ e $M_1 = 1/8$)

- O número 1.6 em nossa representação simplificada:

$$n = \boxed{011}\boxed{1010}$$

$$n = 2^{3-3} \times (1 + 0.625) = 1.625$$

Exemplo de Representação Simplificada: 5

Representação do número 5:

- O número 5 é maior que 2, então ajustamos o expoente:

$$2^0 \times 5 = 2^1 \times 2.5 = 2^2 \times 1.25$$

- Logo, o expoente é 2, ou $2^{5-3} = 2^2$.
- Para a mantissa, representamos 1.25 como:

$$1.25 = 1 + \frac{1}{4} = 1 + 0.25$$

- A mantissa M é representada como: $M_3 M_2 M_1 M_0 = 0100$ (onde $M_2 = 1/4$)
- O número 5 em nossa representação simplificada:

$$n = \boxed{101}\boxed{0100}$$

$$n = 2^{5-3} \times (1 + 0.25) = 4 \times (1 + \frac{1}{4}) = 4 \times \frac{5}{4} = 5$$

Exemplo de Representação Simplificada: 5

Representação do número 5:

- O número 5 é maior que 2, então ajustamos o expoente:

$$2^0 \times 5 = 2^1 \times 2.5 = 2^2 \times 1.25$$

- Logo, o expoente é 2, ou $2^{5-3} = 2^2$.
- Para a mantissa, representamos 1.25 como:

$$1.25 = 1 + \frac{1}{4} = 1 + 0.25$$

- A mantissa M é representada como: $M_3 M_2 M_1 M_0 = 0100$ (onde $M_2 = 1/4$)
- O número 5 em nossa representação simplificada:

$$n = \boxed{101}\boxed{0100}$$

$$n = 2^{5-3} \times (1 + 0.25) = 4 \times (1 + \frac{1}{4}) = 4 \times \frac{5}{4} = 5$$

Exemplo de Representação Simplificada: 5

Representação do número 5:

- O número 5 é maior que 2, então ajustamos o expoente:

$$2^0 \times 5 = 2^1 \times 2.5 = 2^2 \times 1.25$$

- Logo, o expoente é 2, ou $2^{5-3} = 2^2$.
- Para a mantissa, representamos 1.25 como:

$$1.25 = 1 + \frac{1}{4} = 1 + 0.25$$

- A mantissa M é representada como: $M_3 M_2 M_1 M_0 = 0100$ (onde $M_2 = 1/4$)
- O número 5 em nossa representação simplificada:

$$n = \boxed{101}\boxed{0100}$$

$$n = 2^{5-3} \times (1 + 0.25) = 4 \times (1 + \frac{1}{4}) = 4 \times \frac{5}{4} = 5$$

Exemplo de Representação Simplificada: 5

Representação do número 5:

- O número 5 é maior que 2, então ajustamos o expoente:

$$2^0 \times 5 = 2^1 \times 2.5 = 2^2 \times 1.25$$

- Logo, o expoente é 2, ou $2^{5-3} = 2^2$.
- Para a mantissa, representamos 1.25 como:

$$1.25 = 1 + \frac{1}{4} = 1 + 0.25$$

- A mantissa M é representada como: $M_3 M_2 M_1 M_0 = 0100$ (onde $M_2 = 1/4$)
- O número 5 em nossa representação simplificada:

$$n = \boxed{101}\boxed{0100}$$

$$n = 2^{5-3} \times (1 + 0.25) = 4 \times (1 + \frac{1}{4}) = 4 \times \frac{5}{4} = 5$$

Exemplo de Representação Simplificada: 5

Representação do número 5:

- O número 5 é maior que 2, então ajustamos o expoente:

$$2^0 \times 5 = 2^1 \times 2.5 = 2^2 \times 1.25$$

- Logo, o expoente é 2, ou $2^{5-3} = 2^2$.
- Para a mantissa, representamos 1.25 como:

$$1.25 = 1 + \frac{1}{4} = 1 + 0.25$$

- A mantissa M é representada como: $M_3 M_2 M_1 M_0 = 0100$ (onde $M_2 = 1/4$)
- O número 5 em nossa representação simplificada:

$$n = \boxed{101}\boxed{0100}$$

$$n = 2^{5-3} \times (1 + 0.25) = 4 \times (1 + \frac{1}{4}) = 4 \times \frac{5}{4} = 5$$

Exemplo de Representação Simplificada: 0.4

Representação do número 0.4:

- O número 0.4 é menor que 1, então ajustamos o expoente:

$$2^0 \times 0.4 = 2^{-1} \times 0.8 = 2^{-2} \times 1.6$$

- Logo, o expoente é -2, ou $2^{1-3} = 2^{-2}$.
- Para a mantissa, representamos 1.6 como:

$$1.6 \approx 1 + \frac{1}{2} + \frac{1}{8} = 1.625$$

- A mantissa M é representada como: $M_3 M_2 M_1 M_0 = 1010$ (onde $M_3 = 1/2$ e $M_1 = 1/8$).
- O número 0.4 em nossa representação simplificada (aproximado para 0.40625):

$$n = \boxed{001}\boxed{1010}$$

$$n \approx 2^{1-3} \times (1 + 0.625) = \frac{1}{4} \times \frac{8+4+1}{8} = \frac{13}{32} = 0.40625$$

Exemplo de Representação Simplificada: 0.4

Representação do número 0.4:

- O número 0.4 é menor que 1, então ajustamos o expoente:

$$2^0 \times 0.4 = 2^{-1} \times 0.8 = 2^{-2} \times 1.6$$

- Logo, o expoente é -2, ou $2^{1-3} = 2^{-2}$.
- Para a mantissa, representamos 1.6 como:

$$1.6 \approx 1 + \frac{1}{2} + \frac{1}{8} = 1.625$$

- A mantissa M é representada como: $M_3 M_2 M_1 M_0 = 1010$ (onde $M_3 = 1/2$ e $M_1 = 1/8$).
- O número 0.4 em nossa representação simplificada (aproximado para 0.40625):

$$n = \boxed{001}\boxed{1010}$$

$$n \approx 2^{1-3} \times (1 + 0.625) = \frac{1}{4} \times \frac{8+4+1}{8} = \frac{13}{32} = 0.40625$$

Exemplo de Representação Simplificada: 0.4

Representação do número 0.4:

- O número 0.4 é menor que 1, então ajustamos o expoente:

$$2^0 \times 0.4 = 2^{-1} \times 0.8 = 2^{-2} \times 1.6$$

- Logo, o expoente é -2, ou $2^{1-3} = 2^{-2}$.
- Para a mantissa, representamos 1.6 como:

$$1.6 \approx 1 + \frac{1}{2} + \frac{1}{8} = 1.625$$

- A mantissa M é representada como: $M_3 M_2 M_1 M_0 = 1010$ (onde $M_3 = 1/2$ e $M_1 = 1/8$).
- O número 0.4 em nossa representação simplificada (aproximado para 0.40625):

$$n = \boxed{001 \mid 1010}$$

$$n \approx 2^{1-3} \times (1 + 0.625) = \frac{1}{4} \times \frac{8+4+1}{8} = \frac{13}{32} = 0.40625$$

Exemplo de Representação Simplificada: 0.4

Representação do número 0.4:

- O número 0.4 é menor que 1, então ajustamos o expoente:

$$2^0 \times 0.4 = 2^{-1} \times 0.8 = 2^{-2} \times 1.6$$

- Logo, o expoente é -2, ou $2^{1-3} = 2^{-2}$.
- Para a mantissa, representamos 1.6 como:

$$1.6 \approx 1 + \frac{1}{2} + \frac{1}{8} = 1.625$$

- A mantissa M é representada como: $M_3 M_2 M_1 M_0 = 1010$ (onde $M_3 = 1/2$ e $M_1 = 1/8$).
- O número 0.4 em nossa representação simplificada (aproximado para 0.40625):

$$n = \boxed{001}\boxed{1010}$$

$$n \approx 2^{1-3} \times (1 + 0.625) = \frac{1}{4} \times \frac{8+4+1}{8} = \frac{13}{32} = 0.40625$$

Exemplo de Representação Simplificada: 0.4

Representação do número 0.4:

- O número 0.4 é menor que 1, então ajustamos o expoente:

$$2^0 \times 0.4 = 2^{-1} \times 0.8 = 2^{-2} \times 1.6$$

- Logo, o expoente é -2, ou $2^{1-3} = 2^{-2}$.
- Para a mantissa, representamos 1.6 como:

$$1.6 \approx 1 + \frac{1}{2} + \frac{1}{8} = 1.625$$

- A mantissa M é representada como: $M_3 M_2 M_1 M_0 = 1010$ (onde $M_3 = 1/2$ e $M_1 = 1/8$).
- O número 0.4 em nossa representação simplificada (aproximado para 0.40625):

$$n = \boxed{001}\boxed{1010}$$

$$n \approx 2^{1-3} \times (1 + 0.625) = \frac{1}{4} \times \frac{8+4+1}{8} = \frac{13}{32} = 0.40625$$