

Vírgula Flutuante

Trabalho para Casa: TPC2

Alberto José Proença

Metodologia

Leia as folhas do enunciado, e responda obrigatoriamente às questões colocadas na folha fornecida para o efeito, as restantes de acordo com as suas expectativas de graus de exigência.

Relembra-se que o objetivo dos TPC's é fomentar o estudo individual e contínuo, complementado por trabalho em grupo, sendo contabilizado o esforço de se tentar chegar ao resultado (que deverá ser fundamentado na aula) em detrimento da correção do mesmo. A correção dos trabalhos far-se-á na aula da semana em que o trabalho é entregue.

A **penalização por fraude** tem como primeira consequência uma avaliação negativa.

Prazos

Entrega **impreterível** até à hora de início da sessão TP seguinte, com a presença do estudante durante a sessão TP. Não serão aceites trabalhos entregues depois deste prazo.

Introdução

A lista de exercícios que se apresenta segue diretamente o material apresentado na aula teórica sobre representação de números em vírgula flutuante (ver sumário e sugestões de leituras), podendo requerer conceitos básicos adquiridos anteriormente.

Enunciado dos exercícios

Representação de valores em vírgula flutuante

Considere 2 novos formatos de vírgula flutuante, representados com 8-bits, baseados na norma IEEE:

- formato PEQUENO1:
 - o bit mais significativo contém o bit do sinal
 - os 4 bits seguintes formam o expoente (em excesso de 7)
 - os últimos 3 bits representam a mantissa
- formato PEQUENO2:
 - o bit mais significativo contém o bit do sinal
 - os 3 bits seguintes formam o expoente (em excesso de 3)
 - os últimos 4 bits representam a mantissa

Para todos os restantes casos, as regras são as mesmas que as da norma IEEE (valor normalizado, subnormal/desnormalizado, representação do 0, \pm infinito, NaN).

1. (A) Complete a expressão que, a partir dos campos em binário, permite calcular o valor em decimal para cada um dos formatos normalizados: $V = (-1)^S * 1.F * 2^{??}$
2. (A) Para ambos os formatos, apresente os seguintes valores em decimal:
 - a) O maior número finito positivo
 - b) O número negativo normalizado mais próximo de zero
 - c) O maior número positivo subnormal/desnormalizado
 - d) O número positivo subnormal/desnormalizado mais próximo de zero
3. (A) Calcule os valores (número real, \pm infinito, NaN) correspondentes aos seguintes padrões de bits no formato PEQUENO1:
 - a) 10110011
 - b) 01111010
 - c) 10010001
 - d) 00000011
 - e) 11000001
4. (R) Codifique os seguintes valores como números de vírgula flutuante no formato PEQUENO1:
 - a) -111.01_3
 - b) 1/8 Ki (e.g., para representar a dimensão de um ficheiro em bytes)
 - c) $-0x18C$
 - d) 110.01
 - e) 0.005_8
5. (R/B) Converta os seguintes números PEQUENO1 em números PEQUENO2. *Overflow* deve ser representado por \pm infinito, *underflow* por ± 0 e arredondamentos deverão ser para o valor par mais próximo.
 - a) 00110011
 - b) 11101001
 - c) 00010000
 - d) 11001110
 - e) 10000010
6. (B) Considere o desenvolvimento de código científico em C para execução num *notebook* atual, cuja especificação impõe que os números reais sejam representados com pelo menos 8 algarismos significativos. **Indique, justificando**, se consegue representar essas variáveis como `float` ou se tem de as representar como `double`.
7. (B) Um valor do tipo real (*float*) vem representado na norma IEEE 754 por $V = (-1)^S * 1.F * 2^{(Exp-127)}$, se estiver normalizado. **Indique, explicitando** os cálculos, qual o maior inteiro ímpar que é possível representar exatamente, neste formato.

