

O Padrão IEEE 754

por Gabriel da Cunha Dertoni
nº usp: 11795717

O Padrão para Aritmética com Pontos Flutuantes IEEE (IEEE 754) é um padrão estabelecido em 1985 pelo Instituto de Engenheiros Eletricistas e Eletrônicos^[1] (IEEE). Esse padrão se tornou amplamente adotado e é utilizado até os dias atuais como uma convenção de como a aritmética com pontos flutuantes (números reais) deve funcionar.^[2]

Na época em que foi estabelecido, cada fabricante utilizava seu próprio padrão de aritmética com pontos flutuante^[3]. Por consequência, havia grande incompatibilidade entre as máquinas. Dessa necessidade surgiu o padrão IEEE 754, como uma colaboração entre pesquisadores e projetistas de microprocessadores liderada pelo IEEE.

O padrão IEEE 754 descreve os números numa forma de notação científica, que em base binária pode ser descrita a partir da forma:

$$a \times 2^e$$

Para tal, se secciona os bits fornecidos em três partes distintas: o sinal, o expoente e a mantissa. O sinal, consiste de apenas um bit e representa se o número é positivo ou negativo (0 positivo, 1 negativo). A mantissa (a) representa a parte fracionária do número. Para qualquer valor da mantissa, exceto 0, sempre haverá algum 1 mais a esquerda representando o bit mais significativo do número. Como esse bit quase sempre se encontra presente, ele é redundante e, portanto pode ser ignorado na mantissa. Para tal é necessário que a mantissa comece a partir do segundo bit mais significativo. O expoente (e) representa o número de casas que precisam ser deslocadas para que o segundo bit mais significativo da mantissa se encontre na primeira posição. Dessa forma a mantissa apenas armazena os bits após a vírgula^[1].

Por exemplo, o número -83,5. Primeiro se configura o bit de sinal, no caso 1, já que se trata de um número negativo. Em seguida, se calcula o número em base binária $83,5_{10} = 1010011,1_2$, para tornar o segundo bit mais significativo como primeiro bit após a vírgula, é necessário deslocar 6 casas, portanto $a = 0100111_2$ e $e = 6_{10} = 110_2$, os bits restantes são preenchidos com 0. Entretanto, há uma exceção, ao tentar representar o número 1, na primeira etapa, a mantissa seria apenas zeros e o expoente também, uma vez que não é necessário deslocar casa alguma e não haveria forma de distinguir entre 0 e 1. Por conta disso, ao armazenar o expoente, se adiciona $127_{10} = 1111111_2$ a ele. Dessa forma, o número 1 pode ser representado por $0\ 01111111\ 000000000000000000000000$ (seções separadas por espaço, 32 bits). Por fim, o exemplo de -83,5 poderá ser armazenado como:

1 1000101 010011100000000000000000 (32 bits)



Figura 1: Distribuição e posicionamento dos bits para 32 bits totais.

Essa forma de representar números reais, entretanto, possui desvantagens. Por conta da forma como ela é formulada, números que não se encaixam no padrão $x / 2^k$, onde x é o número que se deseja representar, devem ser arredondados. Portanto, números racionais como $\frac{1}{5}$ ou $\frac{1}{3}$ podem apenas ser aproximados no IEEE 754. Para tal, se pode utilizar de arredondamento para cima, para baixo, ou para o número mais próximo^[2].

Para realizar aritmética com números no padrão IEEE 754, é necessário primeiro separar a representação novamente em mantissa, expoente e sinal. Ao realizar a adição, a primeira etapa é igualar os expoentes dos dois números sendo adicionados dividindo um deles por 2 repetidas vezes (aumentar e) até que os expoentes se igualem. Depois, pode se somar as mantissas e depois normalizar o resultado novamente ao padrão. Na multiplicação basta multiplicar as mantissas e somar os expoentes^[3].

Referências:

[1] IEEE 754. Disponível em: <https://pt.wikipedia.org/wiki/IEEE_754>. Acesso em 23 mar. 2020.

[2] Representação em Ponto Flutuante. Disponível em: <http://www.ic.uff.br/~simone/scminter/contaulas/6_FLOAT.pdf>. Acesso em 23 mar. 2020.

[3] Padrão IEEE 754-2008. Disponível em: <<https://ufsj.edu.br/portal2-repositorio/File/nepomuceno/ca/03a-compieee.pdf>>. Acesso em 23 mar. 2020.

[4] Floating Point Numbers (Part2: Fp Addition) - Computerphile. Disponível em: <https://www.youtube.com/watch?v=782QWNOD_Z0&t=66s>. Acesso em 23 mar. 2020.