

Họ và tên: Nguyễn Vũ Đức Hạnh

Mã số sinh viên: 24520453

ASSIGNMENT 1 (+) - FLOATING POINT (*)

Giải thích vì sao có sự khác biệt bên dưới?

$$(1e20 + -1e20) + 3.14 = 3.14$$

$$1e20 + (-1e20 + 3.14) = 0$$

Giải thích:

Số dấu phẩy động chỉ có thể biểu diễn chính xác một tập hợp hạn chế các số. Các số khác được làm tròn để phù hợp với số bit hiện có => khi thực hiện phép tính có thể dẫn đến mất độ chính xác khi thực hiện với số có mũ chênh lệch lớn

$$\text{Ta có: } 3.14 = 1.57 \times 2^1$$

$$1e20 \approx 1.01001101... \times 2^{66}$$

Với **$(1e20 - 1e20) + 3.14 = 3.14$** :

=> $(1e20 - 1e20)$ có cùng mũ và phần giá trị nhưng ngược dấu => Kết quả: 0

$$\Rightarrow 0 + 3.14 = 3.14.$$

Với **$1e20 + (-1e20 + 3.14) = 0$** ;

$-1e20 + 3.14$. Để cộng, phải điều chỉnh số mũ của cả 2 lại cho giống nhau. Mũ của $-1e20$ dưới dạng dấu phẩy động là khoảng 66 (số lớn), trong khi mũ của 3.14 là 1 (số nhỏ). => Dịch phải phần giá trị của 3.14 đi $(66 - 1) = 65$ bit để số mũ giống nhau. Nhưng biểu diễn dấu phẩy động chỉ có 52 bit phần giá trị nên dịch 65 bit vượt giới hạn biểu diễn => 3.14 trở thành 0 sau làm tròn

$$\Rightarrow -1e20 + 0 = -1e20 \text{ (3.14 bị mất hoàn toàn sau khi làm tròn)}$$

$$\Rightarrow 1e20 + (-1e20) = 0.$$