Chopping and Rounding

21 January 2026 09:11

Chopping and rounding of numbers

Definition. (Chopped and rounded number)

Let x be a real number given in the floating point representation as

$$x = (-1)^s \times (0.d_1 d_2 d_3 \cdots d_n d_{n+1} \cdots)_\beta \times \beta^e$$

The floating point approximation of x using n-digit chopping is given by

$$fl(x) = (-1)^s \times (0.d_1 d_2 d_3 \cdots d_n)_\beta \times \beta^e$$

The floating point approximation of x using n-digit rounding is given by

$$fl(x) = \begin{cases} (-1)^s \times (0.d_1 d_2 d_3 \cdots d_n)_\beta \times \beta^e & 0 \le d_{n+1} < \beta/2 \\ (-1)^s \times (0.d_1 d_2 d_3 \cdots (d_n+1))_\beta \times \beta^e & \beta/2 \le d_{n+1} < \beta. \end{cases}$$

Examples: $\pi = (-1)^0 \times 0.31415926\cdots \times 10^1$

$$fl_5(\pi) = (-1)^0 \times 0.31415 \times 10^1 \sim \pi \quad \text{(chopping)}$$

$$fl_5(\pi) = (-1)^0 \times 0.3141\underline{6} \times 10^1 \sim \pi \quad \text{(rounding)}$$

Machine epsilon

The machine epsilon of a computer is the smallest positive floating-point number δ such that

$$fl(1+\delta) > 1$$

For any floating point number $\hat{\delta} < \delta$, we have $fl(1+\hat{\delta}) = 1$, and $1+\hat{\delta}$ and 1 are identical within the computer's arithmetic.

Types of errors

Error = True Value − Approximate value
|Error| − Absolute error

Relative error = $\frac{\text{Error}}{\text{True Value}}$

% error = $|RE| \times 100$

x − TV, x_A − appr. value

$E(x_A) = x - x_A$ $E_a(x_A) = |E(x_A)|$

$E_r(x_A) = \frac{E(x_A)}{x}$ $(x \neq 0)$