Computer Arithmetic

CS/SE 4X03

Ned Nedialkov

McMaster University

September 16, 2021

Cancellations

Consider $x - y$.

Assume no roundoff in the subtraction, i.e. $\mathsf{fl}\,(x-y) = \mathsf{fl}\,(x) - \mathsf{fl}\,(y)$.

The relative error is

$$\begin{aligned}
\left|\frac{\mathsf{fl}\,(x-y)-(x-y)}{x-y}\right| &= \left|\frac{\mathsf{fl}\,(x)-\mathsf{fl}\,(y)-(x-y)}{x-y}\right| \\
&= \left|\frac{(\mathsf{fl}\,(x)-x)-(\mathsf{fl}\,(y)-y)}{x-y}\right| \\
&\leq \frac{|\mathsf{fl}\,(x)-x|+|\mathsf{fl}\,(y)-y|}{|x-y|}
\end{aligned}$$

If $x \approx y$ this ratio can be large.

Example 1. Consider a decimal FP system with $t = 5$ digits. Let $x = 9.23450001$ and $y = 9.23455001$.

Assuming rounding to the nearest, what is the relative error in (a) fl $(x + y)$, (b) fl $(x - y)$?

x and y are represented as fl $(x) = 9.2345$ and fl $(y) = 9.2346$

Unit round of is 5×10^{-5}

(a)

$$\begin{aligned} \text{fl}\,(x+y) &= \text{fl}\big[\text{fl}\,(x) + \text{fl}\,(y)\big] = \text{fl}\,(9.2345 + 9.2346) = \text{fl}\,(1.84691 \times 10) \\ &= 1.8469 \times 10 \end{aligned}$$

$$\left| \frac{\text{fl}\,(x+y) - (x+y)}{x+y} \right| = \left| \frac{1.8469 \times 10 - 1.846905002 \times 10}{1.846905002 \times 10} \right|$$
$$\approx 2.7 \times 10^{-6} < 5 \times 10^{-5}$$

Example 1. cont.

(b)

$$\begin{aligned}\mathsf{fl}(x-y) &= \mathsf{fl}\big[\mathsf{fl}(x) - \mathsf{fl}(y)\big] = \mathsf{fl}(9.2345 - 9.2346) = \mathsf{fl}\left(-1.0000 \times 10^{-4}\right)\\ &= -1.0000 \times 10^{-4}\end{aligned}$$

$$\begin{aligned}\left|\frac{\mathsf{fl}(x-y) - (x-y)}{x-y}\right| &= \left|\frac{-1.0000 \times 10^{-4} - (-5.0000 \times 10^{-5})}{-5.0000 \times 10^{-5}}\right|\\ &= \left|\frac{-10 - (-5)}{-5}\right|\\ &= 1 \gg 5 \times 10^{-5}\end{aligned}$$

Example 2. How to evaluate $\sqrt{x+1} - \sqrt{x}$ to avoid cancellations?

For large x, $\sqrt{x+1} \approx \sqrt{x}$.

$$(\sqrt{x+1} - \sqrt{x})\frac{\sqrt{x+1} + \sqrt{x}}{\sqrt{x+1} + \sqrt{x}} = \frac{1}{\sqrt{x+1} + \sqrt{x}}$$

Evaluate

$$\frac{1}{\sqrt{x+1} + \sqrt{x}}$$

Let $x = 100\,000$. In a 5-digit decima arithmetic,
$x + 1 = 1.0000 \times 10^5 = 100\,001$ rounds to 1.0000×10^5.

Then $\sqrt{x+1} - \sqrt{x}$ gives 0, but

$$\frac{1}{\sqrt{x+1} + \sqrt{x}} = \frac{1}{\sqrt{1.0000 \times 10^5} + \sqrt{1.0000 \times 10^5}} = 1.5811 \times 10^{-3}$$

Example 3. Consider approximating e^{-x} for $x > 0$ by

$$e^{-x} \approx 1 - x + \frac{x^2}{2!} - \frac{x^3}{3!} + \cdots (-1)^k \frac{x^k}{k!}$$

for some k

From $e^{-x} = 1/e^x$, it is better to approximate

$$\boxed{e^x \approx 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots + \frac{x^k}{k!}}$$

✓ taylor series

and then compute $1/e^x$

Solving $ax^2 + bx + c$

Compute the roots of $ax^2 + bx + c = 0$

$$x_{1,2} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

If $b^2 \gg 4|ac|$, there may be cancellations

Example 4.

Consider 4-digit decimal arithmetic and take $a = 1.01$, $b = 98.73$, $c = 4.03$

$$b^2 = 9748, \quad 4ac = 16.28, \quad b^2 - 4ac = 9732$$

$$d = \sqrt{b^2 - 4ac} = 98.65$$

$$-b + d = -98.73 + 98.65 = -0.08, \quad -b - d = -98.73 - 98.71 = -197.4$$

$$x_1 = (-b + d)/(2a) = -3.960 \times 10^{-2}$$

$$x_2 = (-b - d)/(2a) = -97.71$$

Exact roots rounded to 4 digits -4.084×10^{-2}, -97.71

cont.

$d = \sqrt{b^2 - 4ac}$, avoid cancellations in $\pm b + d$

Use $x_1 x_2 = c/a$

Compute using

$d = \sqrt{b^2 - 4ac}$
if $b \geq 0$
  $x_1 = -(b + d)/(2a)$
  $x_2 = c/(ax_1)$
else
  $x_1 = (-b + d)/(2a)$
  $x_2 = c/(ax_1)$

This algorithm gives $x_1 = -97.71$, $x_2 = -4.084 \times 10^{-2}$

Exact roots rounded to 4 digits -97.71, -4.084×10^{-2}