Cancellation I

- Catastrophic cancellation and benign cancellation
- Catastrophic cancellation :

  $b = 3.34, a = 1.22, c = 2.28,$

  true answer: $b^2 - 4ac = 0.0292$

  $b^2 \approx 11.2, 4ac \approx 11.1 \Rightarrow$ computed answer $= 0.1$

  error $= 0.1 - 0.0292 = 0.0708$

  computed answer $= 0.1 = 1.00 \times 10^{-1}$

  ulps $= 0.01 \times 10^{-1} = 10^{-3}$

  $0.0708 \approx 70$ ulps

Cancellation II

- A large error happens when subtracting two close numbers
- Benign cancellation: subtracting exactly known numbers, by guard digits
- $\Rightarrow$ small relative error
- In the example, b^2 and $4ac$ already contain errors

Avoid Catastrophic Cancellation I

- By rearranging the formula
- Example

$$\frac{-b + \sqrt{b^2 - 4ac}}{2a} \tag{1}$$

- If $b^2 \gg 4ac \Rightarrow$ no cancellation when calculating $b^2 - 4ac$ and $\sqrt{b^2 - 4ac} \approx |b|$
  Then $-b + \sqrt{b^2 - 4ac}$ has a catastrophic cancellation if $b > 0$

Avoid Catastrophic Cancellation II

- Multiplying $-b - \sqrt{b^2 - 4ac}$, if $b > 0$

$$\frac{2c}{-b - \sqrt{b^2 - 4ac}} \tag{2}$$

- Use (1) if $b < 0$, (2) if $b > 0$
- Difficult to remove all catastrophic cancellations, but possible to remove some by reformulations
- Another example: $x^2 - y^2$

  Assume $x \approx y$

  $(x - y)(x + y)$ is better than $x^2 - y^2$

Avoid Catastrophic Cancellation III

x^2, y^2 may be rounded $\Rightarrow x^2 - y^2$ may be a catastrophic cancellation

$x - y$ by guard digit

- A catastrophic cancellation is replaced by a benign cancellation

  Of course x, y may have been rounded and $x - y$ is still a catastrophic cancellation.

  Again, difficult to remove all catastrophic cancellations, but possible to remove some

Avoid Catastrophic Cancellation IV

- Calculating area of a triangle

$$A = \sqrt{s(s-a)(s-b)(s-c)}, s = \frac{a+b+c}{2} \tag{3}$$

a, b, c: length of three edges

If $a \approx b + c$, then

$$s = (a+b+c)/2 \approx a,$$

and $s - a$ may have a catastrophic error

Example: $a = 9.00, b = c = 4.53$

$s = 9.03, A = 2.342$

Avoid Catastrophic Cancellation V

Computed solution: $A = 3.04$, error ≈ 0.7

ulps $= 0.01$, error $= 70$ ulps

- A new formulation by Kahan [1986] , $a \geq b \geq c$

$$A = \frac{\sqrt{(a+(b+c))(c-(a-b))(c+(a-b))(a+(b-c))}}{4} \tag{4}$$

$A \approx 2.35$, close to 2.342

- Conclusion: sometimes a formula can be rewritten to have higher accuracy using benign cancellation

Avoid Catastrophic Cancellation VI

- Only works if guard digit is used; most computers use guard digits now

Exactly Rounded Operations I

- Round then calculate $\Rightarrow$ may not be very accurate
- Exactly rounded: compute exactly then rounded to the nearest $\Rightarrow$ usually more accurate
- The definition of rounding
- 12.5 $\Rightarrow$ 12 or 13 ?
- Rounding up: 0, 1, 2, 3, 4 $\Rightarrow$ down, 5, 6, 7, 8, 9 $\Rightarrow$ up

  Why called "rounding up"? Always up for 5
- Rounding even:

  the closest value with even least significant digit

Exactly Rounded Operations II

50% probability up, 50% down

example: $12.5 \Rightarrow 12$; $11.5 \Rightarrow 12$

- Reiser and Knuth [1975] show rounding even may be better

Exactly Rounded Operations III

Theorem 1

Let

$$x_0 = x, x_1 = (x_0 \ominus y) \oplus y, \ldots, x_n = (x_{n-1} \ominus y) \oplus y$$

. *If* $\oplus$ *and* $\ominus$ *are exactly rounded using rounding even, then*

$$x_n = x, \forall n \text{ or } x_n = x_1, \forall n \geq 1$$

.

$x \oplus y, x \ominus y$: computed solution

- Consider rounding up,

Exactly Rounded Operations IV

$\beta = 10, p = 3, x = 1.00, y = -0.555$

$x - y = 1.555, x \ominus y = 1.56, (x \ominus y) + y = 1.56 - 0.555 = 1.005, x_1 = (x \ominus y) \oplus y = 1.01$

$x_1 - y = 1.565, x_1 \ominus y = 1.57, (x_1 \ominus y) + y = 1.57 - 0.555 = 1.015, x_2 = (x_1 \ominus y) \oplus y = 1.02$

Increased by 0.01 until $x_n = 9.45$

- Rounding even:

$x - y = 1.555, x \ominus y = 1.56, (x \ominus y) + y = 1.56 - 0.555 = 1.005, x_1 = (x \ominus y) \oplus y = 1.00$

$x_1 - y = 1.555, x_1 \ominus y = 1.56, (x_1 \ominus y) + y = 1.56 - 0.555 = 1.005, x_2 = (x_1 \ominus y) \oplus y = 1.00$

Exactly Rounded Operations V

- How to implement "exactly rounded operations"?
  We can use an array of words or floating-points
  But you don't have an infinite amount of spaces
- Goldberg [1990] showed that using two guard digits and one sticky bit the result is the same as using exactly rounded operations (details not discussed)