

4 Robustness issues: numerical issues, degenerate cases.

4.1 double arithmetic

4.1.1 Small questions

Let us use `double` arithmetic. For each of the following statements, answer whether it is true or false, and justify in less than one line.

$$a > b \Leftrightarrow a - b > 0 \quad (1)$$

$$(a * b) * c = a * (b * c) \quad (2)$$

$$a + b = b + a \quad (3)$$

$$a * (b + c) = a * b + a * c \quad (4)$$

$$x > y \Rightarrow \text{sqrt}(x) > \text{sqrt}(y) \quad (5)$$

$$(\text{for } x, y \geq 0) \quad x * x \geq y * y \Rightarrow x \geq y \quad (6)$$

$$a, b, c \text{ integers in } [-2^{20}, 2^{20}] \Rightarrow (a - b) * (a - c) = a * a + a * (c - b) - b * c \quad (7)$$

4.1.2 A function

What does the following function return when called on a `double` in the open interval $] -2^{50}, 2^{50}[$?

```
double WhoAmI{double x}
{
    double a = 3377699720527872.0;    // 2^50 + 2^51
    double s = x+a;
    double r = s-a;
    return r;
}
```

4.1 Correction:

4.1.1 Small questions

- (1) true. That's what denormalized numbers are for.
- (2) false. Computer multiplication is not associative
- (3) true. Rounding is the same in both cases.
- (4) false. No distributivity.
- (5) false. It is possible that $\text{sqrt}(x) = \text{sqrt}(y)$ even for $x \neq y$.
- (6) true. Rounding preserves order (in the broad sense).
- (7) true. All values are integers $< 2^{54}$ so they are exact with 53 significant digits. (Of course there was a mistake in the question, it should read $(a - b) * (a - c) = a * a - a * (b + c) + b * c$.)

4.1.2 A function

Answer: Rounding to closest half-integer.

Proof: $2^{51} = 2^{51} + 2^{50} - 2^{50} < x + a < 2^{51} + 2^{50} + 2^{50} = 2^{52}$. So, the value of first significant bit of `s` is $2^{52-1} = 2^{51}$, and the value of the 53rd significant bit of `s` is $2^{52-53} = 0.5$. Since the rounding mode is to closest, `s` becomes the half-integer closest to `x+a`. Finally, `r` is the half-integer that is closest to `x`.

4.2 Circle intersection

Let C_1 and C_2 be two circles of respective centers (x_1, y_1) and (x_2, y_2) and respective radii r_1 and r_2 ,

4.2.1 Predicate

Write the predicate testing if C_1 and C_2 intersect as the sign of a polynomial in $x_1, y_1, r_1, x_2, y_2, r_2$.

4.2.2 Precision

Assume that the input data $x_1, y_1, r_1, x_2, y_2, r_2$ are integers in $[-2^b, 2^b]$, and that the computations are performed with `double`. For which values of b is the predicate guaranteed to give the correct result?

(Recall that according to IEEE754 norm, `double` are stored with 53 significant bits.)

4.2 Correction:

4.2.1 Predicate

p_i denotes the center of $C_i, i = 1, 2$.

If $\|p_1 p_2\| > r_1 + r_2$ then disjoint

If $\|p_1 p_2\| = r_1 + r_2$ then tangent

If $\|p_1 p_2\| < \|r_1 - r_2\|$ then disjoint (one circle is contained in the other)

If $\|p_1 p_2\| = \|r_1 - r_2\|$ then tangent

otherwise intersecting.

To put this in a single polynomial:

$s = \text{sign}(((x_1 - x_2)^2 + (y_1 - y_2)^2 - (r_1 + r_2)^2)((x_1 - x_2)^2 + (y_1 - y_2)^2 - (r_1 - r_2)^2))$

If $s < 0$ then intersection; If $s = 0$ then tangent; If $s > 0$ then disjoint.

4.2.2 Precision

The sign of a product is the product of the signs of the factors.

$x_1 - x_2, y_1 - y_2,$ and $r_1 + r_2$ are integers in $[-2^{b+1}, 2^{b+1}]$

$(x_1 - x_2)^2, (y_1 - y_2)^2,$ and $(r_1 + r_2)^2$ are integers in $[0, 2^{2b+2}]$

$(x_1 - x_2)^2 + (y_1 - y_2)^2 \in [0, 2^{2b+3}]$

$(x_1 - x_2)^2 + (y_1 - y_2)^2 - (r_1 + r_2)^2 \in [-2^{2b+2}, 2^{2b+3}]$

For the computations to be exact it is enough that $2b + 3 \leq 53$, i.e. $b \leq 25$. Same for $(x_1 - x_2)^2 + (y_1 - y_2)^2 - (r_1 - r_2)^2$. Remark: there is no risk of overflow.