J3/97-171

Date: May 13, 1997
To: J3
From: William (Bill) Walster
Subject: Additional thoughts on interval arithmetic

Introduction to Interval Arithmetic

Abstract

Most people have not heard of interval arithmetic. The purpose of this article is to provide a simple non-technical introduction to the subject. In so doing, it is hoped that the paradigm shift from *point* to *interval* arithmetic will be accelerated.

1 History

Forty years ago Ramon E. Moore had his first interval thoughts while an employee of Lockheed Missiles and Space Co. Inc. in Sunnyvale CA. At the April 19, 1997 kick-off meeting of Sun Microsystems' university R & D program in interval arithmetic, Ray explained his thinking as follows. In 1957 at LMSC, he was considering how scientists and engineers represent measurements and computed values as $x \pm \epsilon$, where x is the measurement or result and ϵ is a bound on its accuracy.

Example 1. Measure a room with a tape measure and observe that it is 10 feet $\pm$ 1/4 inch wide, by 20 feet $\pm$ 1/4 inch long. From these two observations, the area of the room can be calculated to be: $(10' \pm \frac{1}{4}'') \times (20' \pm \frac{1}{4}'')$ or $200\text{ft}^2 \pm$ some number of square feet. But what is the error in the estimated area of the room? It is not obvious!

If propagating observation errors through the computation in this simple example is not obvious, think about the problem of bounding errors in more complex scientific and engineering calculations. While computers make it easy to perform ever more computations with ever increasing speed in every conceivable technical and commercial field, *point* (as opposed to *interval*) arithmetic provides no obvious mechanical way to bound errors in

final results. Consequently, more and more computing is being done with no knowledge of the end results' accuracy. Striking evidence of this fact exists in the dearth of formal error analyses in scientific and technical computing. In most cases the effort is so daunting that error analyses are not performed.

Ray had a better idea. He reasoned that since $x \pm \epsilon$ consists of two numbers x and ϵ, why not use two *different* numbers to represent exactly the same information contained in $x \pm \epsilon$? That is, instead of $x \pm \epsilon$, use $x - \epsilon$ and $x + \epsilon$, which define the endpoints of an *interval* containing the true quantity in question, x. It was this simple idea that gave birth to *interval arithmetic* and the branch of applied mathematics known as *interval analysis*. With points there is a disconnect between mathematics and computing. With intervals there is not. It is this fact and the results that flow from it that leads those in the interval community to believe that a paradigm shift in computing from point to interval arithmetic has begun.

2 Interval Arithmetic

Arithmetic on intervals is defined so that any interval arithmetic operation produces a new interval that must contain the set of all possible results.

Definition. Let X and Y stand for the intervals $[a,\ b]$ and $[c,\ d]$; where a and c are the lower bounds on the intervals X and Y,and b and d are the upper bounds on the intervals X and Y,respectively. Interval addition of the two intervals X and Y is defined as follows:

$$X + Y = [a,\ b] + [c,\ d] = [a + c,\ b + d]. \tag{1}$$

That is, the lower bound of the sum of two intervals is the sum of their lower bounds, and the upper bound of the interval sum is the sum of their upper bounds.

If needed to guarantee containment when interval arithmetic is implemented on a computer, directed rounding must be used to round lower bounds down (toward $-\infty$) and upper bounds up (toward $+\infty$). Interval arithmetic must always produce intervals that contain the set of all possible results.

Subtraction, multiplication and division are defined analogously to addition. Division by an interval containing zero is well defined and is used in the development of important algorithms that have no counterpart in point arithmetic.

With this modest introduction to interval arithmetic, it is possible to continue with the example of computing the area of the room.

Example 1, continued. Using intervals the width and length of the room can be represented using 5 decimal digits of accuracy:

$$W = [9' \; 11\frac{3}{4}'' , \; 10' \; \frac{1}{4}'' \;] = [9 + \frac{11.75}{12}, \; 10 + \frac{.25}{12}]\text{ft} = [9.9791, \; 10.021]\text{ft}; \text{ and} \quad (2)$$

$$L = [19' \; 11\frac{3}{4}'' , \; 20' \; \frac{1}{4}'' \;] = [19 + \frac{11.75}{12}, \; 20 + \frac{.25}{12}]\text{ft} = [19.979, \; 20.021]\text{ft}. \quad (3)$$

The area of the room can be calculated using 5 decimal digit interval arithmetic:

$$\begin{aligned} A &= W \times L = [9.9791, \; 10.021] \times [19.979, \; 20.021]\text{ft}^2 && (4)\\ &= 9.9791 \times 19.979, \; 10.021 \times 20.021]\text{ft}^2 && (5)\\ &= [199.37, \; 200.64]\text{ft}^2. && (6) \end{aligned}$$

Using point notation, the area of the room is:

$$\begin{aligned} a &= \left(\frac{199.37 + 200.64}{2} \pm \frac{200.64 - 199.37}{2}\right)\text{ft}^2 && (7)\\ &= (200.005 \pm 0.635)\,\text{ft}^2. && (8) \end{aligned}$$

But with only 5 decimal digits of accuracy, it is necessary to round 200.005 to 200.00 or 200.01, with a corresponding increase in the error tolerance from .635 to .64 to cover the endpoints of the interval.

Without interval arithmetic it is not obvious how the point result should be obtained.. It is even less clear how to automate computing point results with correct error tolerances..

Example 2. Now consider a slightly more interesting problem. Suppose after making the above measurements of W and L, it is desired to determine if the room is square. From the Pythagorean:

$$\begin{aligned} D^2 &= W^2 + L^2 = \left([9.9791,\ 10.021]^2 + [19.979,\ 20.021]^2\right) \text{ft}^2 && (9) \\ &= [9.9791^2 + 19.979^2,\ 10.021^2 + 20.021^2]\text{ft}^2 && (10) \\ &= [498.74,\ 501.27]\text{ft}^2; \text{ where} && (11) \end{aligned}$$

D is the diagonal of the room. Suppose a direct measurement of D is made and the following value is obtained:

$$\begin{aligned} D &= [22' \ 4\frac{3}{4}'' \ ,\ 22' \ 5\frac{1}{4}'' \ ] = [22.395,\ 22.438]', \text{ or} && (12) \\ D^2 &= [501.53,\ 503.47]\text{ft}^2. && (13) \end{aligned}$$

The two interval results for computing D^2 do not overlap! The two intervals are disjoint. If the accuracy of the measurements can be trusted, this result implies the room is not square or the Pythagorean Theorem is false. To verify this conclusion, a prudent investigator will replicate the measurements, the calculations and also measure the other diagonal. This simple example illustrates an important property of interval computations: Interval arithmetic can be used to perform numerical proofs. Numerical proofs are an automatic by-product of interval arithmetic.

3 The Interval Paradigm

The simple examples above, illustrate some of the features of interval arithmetic that are not present in point arithmetic. Among others, the interval paradigm has the following characteristics:

1. Fallible data can be represented using intervals and any arithmetic computation can be performed using interval arithmetic to obtain guaranteed bounds on the set of all possible results. Doing the equivalent error analysis using points is so complex that such error analyses are rarely conducted.

2. Interval Arithmetic provides an automatic way to bound errors from all sources, including:

    - observation or measurement errors;
    - modeling errors;
    - machine rounding errors; and
    - interactions among all of the above.

3. Using interval arithmetic, it is possible to solve problems that are thought to be in principle unsolvable by many mathematicians, scientists and engineers. The two most noteworthy exemplar problems are:

    - nonlinear systems of equations; and
    - nonlinear programing (or nonlinear global optimization).

Evidence can be found in the March 1997 issue of SIAM News on page 8 of the degree to which it is commonly, but mistakenly believed that nonlinear optimization problems may *never* be solved. For additional information on interval arithmetic and how it has been used to compute guaranteed bounds on otherwise unsolvable problems, see: http://cs.utep.edu/interval-comp/main.html.