Interval Computations as Applied Constructive Mathematics: from Shanin to Wiener and Beyond

Vladik Kreinovich
Department of Computer Science
University of Texas at El Paso
El Paso, Texas 79968, USA
vladik@utep.edu
1. General Problems of Science and Engineering

- The main objective of science is to understand the current state of the world and to predict its future state.
- The main objective of engineering is to find controls and strategies that lead to a better future.
- The state of the world is usually described in terms of real numbers – values of physical quantities.
- Some quantities we can measure directly: e.g., distance from here to our hotel.
- Other quantities \( y \) we cannot measure directly: e.g., distance from here to a nearby star.
2. Indirect Measurements

- Since we cannot measure the quantity of interest $y$ directly, we measure it indirectly.
- Namely, we measure related easier-to-measure quantities $x_1, \ldots, x_n$ and get values $\tilde{x}_i$.
- Then, we use the known relation $y = f(x_1, \ldots, x_n)$ and known (approximate) values of $x_i$ to estimate $y$ as

$$\tilde{y} = f(\tilde{x}_1, \ldots, \tilde{x}_n).$$
3. How Constructive Mathematics Can Help

- Ideally, we want to be able to estimate $y$ with any given accuracy $\varepsilon$.

- For this purpose, we need to have:
  - an algorithm that, given $\varepsilon$, computes the accuracy $\delta$ with we should measure the inputs, and
  - an algorithm $\tilde{f}$ that, when applied to the measurement results $\tilde{x}_i$ to get the desired estimate $\tilde{y}$:
    $$|\tilde{x}_i - x_i| \leq \delta \Rightarrow |\tilde{f}(\tilde{x}_1, \ldots, \tilde{x}_n) - f(x_1, \ldots, x_n)| \leq \varepsilon.$$

- In a nutshell, this is what constructive mathematics is about – when limited to real numbers.
4. Applied Constructive Mathematics

- In practice, our ability to measure accurately is limited.
- So, we have measurement results $\tilde{x}_i$ with some accuracies $\delta_i$: $|\tilde{x}_i - x_i| \leq \delta_i$.
- The only information that we have about the actual value $x_i$ is that $x_i \in [\underline{x}_i, \overline{x}_i]$ def $= [\tilde{x}_i - \delta_i, \tilde{x}_i + \delta_i]$.
- What can we say about $y = f(x_1, \ldots, x_n)$? We can only conclude that $y \in [\underline{y}, \overline{y}]$ def $= \{f(x_1, \ldots, x_n) : x_i \in [\underline{x}_i, \overline{x}_i]\}$.
- Computing $[\underline{y}, \overline{y}]$ is called interval computations.
- Yuri Matiyasevich called it applied constructive mathematics.
5. Why Intervals?

• Usually, we do not just know the upper bound $\delta_i$ on the measurement error $\Delta x_i \stackrel{\text{def}}{=} \bar{x}_i - x_i$: $|\Delta x_i| \leq \delta_i$.

• We also know the probabilities of different values $\Delta x_i$.

• These probabilities come from comparing measurement results with a standard (more accurate) instrument.

• There are two situations when this is not possible:
  
  – state-of-the-art measurement, when we use the most accurate instrument; and
  
  – measurements on the shop floor, where we could calibrate everything, but it would cost too much.

• Then, all we have is an upper bound $\delta_i$ on $|\Delta x_i|$.
6. **Why Wiener? A Brief History of Interval Computations**

- **Origins:** Archimedes (Ancient Greece), N. Wiener (1914)
- **Modern pioneers:** Mieczyslaw Warmus (Poland), Teruo Sunaga (Japan), Ramon Moore (USA), 1956–59
- **First boom:** early 1960s.
- **First challenge:** taking interval uncertainty into account when planning spaceflights to the Moon.
- **Current applications** (sample):
  - design of elementary particle colliders: Martin Berz, Kyoko Makino (USA)
  - will a comet hit the Earth: Martin Berz, Ramon Moore (USA)
  - robotics: L. Jaulin (France), A. Neumaier (Austria)
  - chemical engineering: M. Stadtherr (USA)
7. Interval Computations – How? First Idea

- In a computer, every computation is a sequence of elementary arithmetic operations.
- In mathematical terms, this means that we consider compositions of simple arithmetic functions.
- So, a natural idea – known as straightforward interval computations – is to:
  - find interval analogues of simple arithmetic functions, and then
  - in the original algorithm, replace each arithmetic operation with the corresponding interval one.
8. Interval Analogues of Simple Arithmetic Functions

- When $x_1 \in x_1 = [x_1, \bar{x}_1]$ and $x_2 \in x_2 = [x_2, \bar{x}_2]$, then:
  - The range $x_1 + x_2$ for $x_1 + x_2$ is $[x_1 + x_2, \bar{x}_1 + \bar{x}_2]$.
  - The range $x_1 - x_2$ for $x_1 - x_2$ is $[x_1 - \bar{x}_2, \bar{x}_1 - x_2]$.
  - The range $x_1 \cdot x_2$ for $x_1 \cdot x_2$ is $[y, \bar{y}]$, where
    \[ y = \min(x_1 \cdot x_2, x_1 \cdot \bar{x}_2, \bar{x}_1 \cdot x_2, \bar{x}_1 \cdot \bar{x}_2); \]
    \[ \bar{y} = \max(x_1 \cdot x_2, x_1 \cdot \bar{x}_2, \bar{x}_1 \cdot x_2, \bar{x}_1 \cdot \bar{x}_2). \]

- The range $1/x_1$ for $1/x_1$ is $[1/\bar{x}_1, 1/x_1]$ (if $0 \not\in x_1$).

- These operations are known as interval arithmetic.
9. **Straightforward Interval Computations: Example and Limitations**

- *Example:* \( f(x) = (x - 2) \cdot (x + 2), \ x \in [1, 2]. \)

- How will the computer compute it?
  - \( r_1 := x - 2; \)
  - \( r_2 := x + 2; \)
  - \( r_3 := r_1 \cdot r_2. \)

- *Main idea:* perform the same operations, but with *intervals* instead of *numbers*:
  - \( r_1 := [1, 2] - [2, 2] = [-1, 0]; \)
  - \( r_2 := [1, 2] + [2, 2] = [3, 4]; \)
  - \( r_3 := [-1, 0] \cdot [3, 4] = [-4, 0]. \)

- *Actual range:* \( f(x) = [-3, 0] \subset [-4, 0]. \)

- *Comment:* excess width (4 vs. 3) is unavoidable, since interval computations is NP-hard.
10. What Can We Do?

- Representing an algorithm as a composition of elementary arithmetic functions often does not work.
- Idea: represent it as a composition of some other functions.
- What is the class of functions closed under composition?
- It is reasonable to require that this class is also closed under inversion.
- So, we are looking for group of transformations of $\mathbb{R}^n$.
- The simplest such group is the group of all linear transformations.
- What are other such groups?
11. Wiener Again

- N. Wiener is mostly known as the father of cybernetics, a general theory of biological and engineering systems.
- His interest started when a physiologist noticed that his design resembled the actual neural structure.
- Wiener noticed that when approach a faraway object, we go through five phases.
- At first, we notice a blur – corresponding to all possible transformations.
- Then, we get a shape modulo projective transformations.
- Then, affine, then homotheties, and finally, we identify the object exactly.
12. Wiener (cont-d)

- Wiener mentioned that we are a product of billion years of improving evolution.
- So, if there were other groups, we would have used them.
- So, he conjectured that there are no other groups.
- The only transformation groups containing all linear one are all projective ones and all transformation.
- Surprisingly, this was indeed proven in the 1960s by V. M. Guillemin, I. M. Singer, and S. Sternberg.
13. For the Resulting Transformations, Interval Computations Are Feasible

- The most general case is all possible transformations.
- Locally – in the vicinity of id – they are monotonic.
- Computing the range of monotonic $f(x_1, \ldots, x_n)$ is easy.
- For example, if $f$ increases in all $x_i$, the range is
  \[
  [f(x_1, \ldots, x_n), f(\bar{x}_1, \ldots, \bar{x}_n)].
  \]
- The range of a linear $f(x_1, \ldots, x_n) = a_0 + \sum_{i=1}^{n} a_i \cdot x_i$ on $[\tilde{x}_i - \delta_i, \tilde{x}_i + \delta_i]$ is $[\tilde{y} - \delta, \tilde{y} + \delta]$, where:
  \[
  \tilde{y} = f(\tilde{x}_1, \ldots, \tilde{x}_n) \text{ and } \delta = \sum_{i=1}^{n} |a_i| \cdot \delta_i.
  \]
- Feasible algorithms are known for fractional-linear $f$. 
14. Towards Standard Interval Computations

- If a function is linear or fractional-linear, we apply the known algorithms.
- If not, we check whether the function is monotonic.
- We know that $f \uparrow x_i$ if $\frac{\partial f}{\partial x_i} \geq 0$.
- So, to check, we estimate the range $[d_i, \bar{d}_i]$ of this partial derivative – e.g., by straightforward interval comp.
- If $d_i \geq 0$, we can use monotonicity-based formulas.
- If the function is not monotonic, we try to approximate it by one of the feasible-for-intervals functions.
- Approximations by monotonic functions is a new idea, currently being tested.
- Approximation by fractional-linear functions is a raw idea, no algorithm is know.
15. Standard Interval Computations (cont-d)

- Approximation by linear functions – 1st terms in Taylor series – is well known: for some $\eta \in [x_1, \bar{x}_1] \times \ldots$

\[
f(x_1, \ldots, x_n) = f(\tilde{x}_1, \ldots, \tilde{x}_n) + \sum_{i=1}^{n} \left( \frac{\partial f}{\partial x_i} \bigg|_{\eta} \right) \cdot \Delta x_i.
\]

- Thus, we get a centered form estimate

\[
[y, \bar{y}] \subseteq \tilde{y} + \sum_{i=1}^{n} [d_i, \bar{d}_i] \cdot [-\delta_i, \delta_i].
\]

- This formula is obtained by ignoring second order terms, so its accuracy is $O(\delta_i^2)$.

- To increase its accuracy, we can decrease $\delta_i$ by bisection.
16. Bisection: How?

- If we bisect all the intervals, we get $2^n$ subboxes – too many for large $n$.
- So, we need to decided which interval to bisect, based on the values $|d_i|$ and $\delta_i$.
- The resulting criterion $f(|d_i|, \delta_i)$ should not change if we change the units for measuring $x_i$ or $y$:

  $$f(|d_i|, \delta_i) > f(|c_i|, \gamma_i) \iff f(\mu \cdot \lambda^{-1} \cdot |d_i|, \lambda \cdot \delta_i) > f(\mu \cdot \lambda^{-1} \cdot |c_i|, \lambda \cdot \gamma_i).$$

- This implies bisecting where $|d_i| \cdot \delta_i \to \text{max}$. 
17. Resulting Algorithm

- We need to estimate the range of $f(x_1, \ldots, x_n)$ on intervals $[x_i, \bar{x}_i] = [\tilde{x}_i - \delta_i, \tilde{x}_i + \delta_i]$.

- First, we use straightforward interval computations to find the range $[d_i, \bar{d}_i]$ of each partial derivative $\frac{\partial f}{\partial x_i}$.

- If $d_i \geq 0$ or $\bar{d}_i \leq 0$, monotonicity reduces the problem to a problem with $n - 1$ variables.

- If the result is non-monotonic, we use the centered form estimate: $[y, \bar{y}] \subseteq \tilde{y} + \sum_{i=1}^{n} [d_i, \bar{d}_i] \cdot [-\delta_i, \delta_i]$.

- To get a more accurate estimate, we bisect the interval with the largest product $|d_i| \cdot \delta_i$, and repeat.
18. Monotonicity: Example

• **Idea:** if the range $[r_i, \bar{r}_i]$ of each $\frac{\partial f}{\partial x_i}$ on $x_i$ has $r_i \geq 0$, then

\[
f(x_1, \ldots, x_n) = [f(x_1, \ldots, x_n), f(\bar{x}_1, \ldots, \bar{x}_n)].
\]

• **Example:** $f(x) = (x - 2) \cdot (x + 2)$, $x = [1, 2]$.

• **Case $n = 1$:** if the range $[r, \bar{r}]$ of $\frac{df}{dx}$ on $x$ has $r \geq 0$, then

\[
f(x) = [f(x), f(\bar{x})].
\]

• **AD:** $\frac{df}{dx} = 1 \cdot (x + 2) + (x - 2) \cdot 1 = 2x$.

• **Checking:** $[r, \bar{r}] = [2, 4]$, with $2 \geq 0$.

• **Result:** $f([1, 2]) = [f(1), f(2)] = [-3, 0]$.

• **Comparison:** this is the exact range.
19. Centered Form: Example

- **General formula:**
  \[ Y = f(\tilde{x}_1, \ldots, \tilde{x}_n) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(x_1, \ldots, x_n) \cdot [-\Delta_i, \Delta_i]. \]

- **Example:** \( f(x) = x \cdot (1 - x), \ x = [0, 1]. \)

- Here, \( x = [\tilde{x} - \Delta, \tilde{x} + \Delta], \) with \( \tilde{x} = 0.5 \) and \( \Delta = 0.5. \)

- **Case \( n = 1: \)** \( Y = f(\tilde{x}) + \frac{df}{dx}(x) \cdot [-\Delta, \Delta]. \)

- **AD:** \( \frac{df}{dx} = 1 \cdot (1 - x) + x \cdot (-1) = 1 - 2x. \)

- **Estimation:** we have \( \frac{df}{dx}(x) = 1 - 2 \cdot [0, 1] = [-1, 1]. \)

- **Result:** \( Y = 0.5 \cdot (1 - 0.5) + [-1, 1] \cdot [-0.5, 0.5] = 0.25 + [-0.5, 0.5] = [-0.25, 0.75]. \)

- **Comparison:** actual range \([0, 0.25]\), straightforward \([0, 1].\)
20. Centered Form and Bisection: Example

- **Known:** accuracy $O(\Delta_i^2)$ of first order formula

  $$f(x_1, \ldots, x_n) = f(\tilde{x}_1, \ldots, \tilde{x}_n) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\chi) \cdot (x_i - \tilde{x}_i).$$

- **Idea:** if the intervals are too wide, we:
  - split one of them in half ($\Delta_i^2 \rightarrow \Delta_i^2 / 4$); and
  - take the union of the resulting ranges.

- **Example:** $f(x) = x \cdot (1 - x)$, where $x \in x = [0, 1]$.

- **Split:** take $x' = [0, 0.5]$ and $x'' = [0.5, 1]$.

- **1st range:** $1 - 2 \cdot x = 1 - 2 \cdot [0, 0.5] = [0, 1]$, so $f \uparrow$ and $f(x') = [f(0), f(0.5)] = [0, 0.25]$.

- **2nd range:** $1 - 2 \cdot x = 1 - 2 \cdot [0.5, 1] = [-1, 0]$, so $f \downarrow$ and $f(x'') = [f(1), f(0.5)] = [0, 0.25]$.

- **Result:** $f(x') \cup f(x'') = [0, 0.25]$ – exact.
21. Wiener and Constructive Mathematics Yet Again

- As we have mentioned, the general problem of interval computations is NP-hard.
- It is NP-hard even for computing the range of sample variance \( \frac{1}{n} \cdot \sum_{i=1}^{n} (x_i - a)^2 \), where \( a = \frac{1}{n} \cdot \sum_{i=1}^{n} x_i \).
- So (unless P = NP), the worst-case complexity of interval computations problems is exponential.
- A natural question: what about average computational complexity?
- Here, we need a probability measure on the set of all functions.
- In this problem, Norbert Wiener was also a pioneer with his Wiener measure.
22. Wiener Yet Again (cont-d)

- The original formulas of Wiener were not algorithmic.
- In my early papers, it was shown that constructivization is possible.
- This was inspired by Shanin and constructive mathematics.
- That result was for Wiener’s measure – and real distribution may be different.
- So, recently, we extended these algorithmic results to general probability measures over metric spaces.
- This includes function spaces as particular cases.
23. Beyond Intervals

- Instead of boxes, we can have other sets describing uncertainty – e.g., ellipsoids or zonotopes.
- Instead of numbers, we can have similar uncertainty about more complex objects – e.g., functions.
- This becomes applications of beyond-numbers constructive mathematics.
- All these problems can be naturally reformulated in terms of modal logic.
- Indeed, \( x_i \in [\underline{x}_i, \overline{x}_i] \) means that all values from this interval are possible.
- We want to find when \( y \) is a possible value of \( f(x_1, \ldots, x_n) \).
- Thus, it is also applied modal logic.
- Can Yu. Gurevich’s modal constructive logic help?
24. Interval Computations: Everyone is Welcome

- Interval computations is extremely important for practice.
- There are many theoretical and practical open problems.
- We have regular biannual conferences SCAN’XX, the next one will be in Hungary in September 2020.
- There are annual European SWIM workshops.
- We have a journal *Reliable Computing* (formerly *Interval Computations*), founded by Yu. Matiyasevich.
- Our website is [http://www.cs.utep.edu/interval-comp](http://www.cs.utep.edu/interval-comp)
- Everyone is welcome to visit — or even to join our community!