7-1-2023

# How to Propagate Interval (and Fuzzy) Uncertainty: Optimism-Pessimism Approach

Vinícius F. Wasques
*São Paulo State University*, vwasques@outlook.com

Olga Kosheleva
*The University of Texas at El Paso*, olgak@utep.edu

Vladik Kreinovich
*The University of Texas at El Paso*, vladik@utep.edu



Comments:
Technical Report: UTEP-CS-23-28
To appear in *Proceedings of the 20th World Congress of the International Fuzzy Systems Association IFSA'2023*, Daegu, South Korea, August 20-24, 2023.

# How to Propagate Interval (and Fuzzy) Uncertainty: Optimism-Pessimism Approach

**Vinícius F. Wasques**[a], **Olga Kosheleva**[b], and **Vladik Kreinovich**[c]
[a]Ilum School of Science, São Paulo State University UNESP, São Paulo, Brazil, vwasques@outlook.com
[b]Teacher Education, Univ. of Texas at El Paso, El Paso, TX 79968, USA, olgak@utep.edu
[c]Computer Science, Univ. of Texas at El Paso, El Paso, TX 79968, USA, vladik@utep.edu

## Abstract

In many practical situations, inputs to a data processing algorithm are known with interval uncertainty, and we need to propagate this uncertainty through the algorithm, i.e., estimate the uncertainty of the result of data processing. Traditional interval computation techniques provide guaranteed estimates, but from the practical viewpoint, these bounds are too pessimistic: they take into account highly improbable worst-case situations when all the measurement and estimation errors happen to be strongly correlated. In this paper, we show that a natural idea of having more realistic estimates leads to the use of so-called interactive addition of intervals, techniques that has already been successful used to process interval uncertainty. Thus, we provide a new justification for these techniques. If we use a known interpretation of a fuzzy set as a nested family of intervals – its alpha-cuts – then we can naturally extend our results to the case is fuzzy uncertainty.

**Keywords:** interval uncertainty, fuzzy uncertainty, uncertainty propagation, interactive addition of intervals

## 1 Formulation of the Problem

**Need for data processing.** One of the main objectives of science is to predict the future state of the world, i.e., the future values of all the quantities that characterize this state. One of the main objectives of engineering is to come up with designs and controls that lead to a better future state.

How do we predict the future state of the world? We measure or estimate the current state and we use the known relation between the current and the future state of the world to make these predictions. In other words, to predict the future value $y$ of one of the desired quantities, we find the current quantities $x_1,\dots,x_n$ that are related to $y$ by a known dependence $y = f(x_1,\dots,x_n)$, and then we use the estimates $\widetilde{x}_1,\dots,\widetilde{x}_n$ of the current quantities – coming either from measurements or from expert estimates – to compute the estimate $\widetilde{y} = f(\widetilde{x}_1,\dots,\widetilde{x}_n)$ for the desired quantity $y$. This computation is an important particular case of *data processing*.

**Need for uncertainty propagation.** Measurement are never absolutely accurate. Similarly, expert estimates are never absolutely accurate. In both cases, each estimate $\widetilde{x}_i$ is, in general, somewhat different from the actual (unknown) value $x_i$ of the corresponding quantity. The difference $\Delta x_i \stackrel{\text{def}}{=} \widetilde{x}_i - x_i$ is known as the *approximation error*, or, in case of measurements, a *measurement error*.

Because of these differences, the value $\widetilde{y}$ that is obtained by applying the algorithm $f$ to the measurement results is, in general, different from the value $y$ that we would have obtained if we knew the actual values $x_i$. It is important to know how accurate is our estimate i.e., how big is the difference $\Delta y \stackrel{\text{ref}}{=} \widetilde{y} - y$ between the estimate $\widetilde{y}$ and the actual value $y$. Estimating this difference – i.e., the uncertainty with which we know $y$ – based on the uncertainty with which we know the inputs $x_i$ is known as *uncertainty propagation*.

**Interval uncertainty.** To perform uncertainty propagation, we need to know what information we have about the uncertainty in $x_i$. Let us start with the case of measurements. Often, the only information that we have about the measurement error is the range $[\underline{\Delta},\overline{\Delta}]$ of its possible values of measurement error (provided by the manufacturer of the measuring instrument). For example, if the upper bound on the absolute value of the measurement error is 0.1, this means that the measurement error lies in the interval $[\underline{\Delta},\overline{\Delta}] = [-0.1, 0.1]$.

In such cases, once we know the measurement result $\widetilde{x}$, the only conclusion we can make about the actual (unknown) value $x$ of the corresponding quantity is that this value $x$ is somewhere in the interval $[\underline{x},\overline{x}] \stackrel{\text{def}}{=} [\widetilde{x}-\overline{\Delta},\widetilde{x}-\underline{\Delta}]$. For example, if we have an instrument with measurement error not exceeding 0.1, and the measurement result is 1.0, this means that the actual value is somewhere in the interval $[1.0-0.1,1.0+0.1] = [0.9,1.1]$. This situation is known as *interval uncertainty*; see, e.g., [4].

**Fuzzy uncertainty.** Interval uncertainty means that we are absolutely sure that the measurement error is in the interval $[\underline{\Delta},\overline{\Delta}]$. As we have mentioned, this interval is provided by the manufacturer, and the manufacturer wants to make sure that it is not sued if its instrument is used in critical situations and underestimation of the measurement error leads to a disaster. Because of this fear, manufacturers usually overestimate their bounds on measurement error. Designers of measuring instruments usually say that, e.g., while the only interval they can absolutely guarantee is $[-0.1,0.1]$, they are very confident – e.g., with degree of confidence 0.99 – that it will actually be within a narrower interval, e.g., $[-0.05,0.05]$; with somewhat smaller degree of confidence – that it will be even within a narrower interval, etc.

In such situations, in addition to the wide interval corresponding to 100% certainty (and, correspondingly, 0% uncertainty), we have narrower intervals corresponding to intermediate levels of uncertainty $\alpha$ – all the way to 1. These nested intervals form, in effect, what is called a fuzzy set (see, e.g., [2]) – so that an interval corresponding to degree $\alpha$ is the $\alpha$-cuts of this set.

Similar things happens if instead of measurement results, we use expert estimates – we have wider intervals of possible values that, with low uncertainty, contain the actual value, and we have narrower intervals corresponding to higher uncertainty.

From this viewpoint, a fuzzy set is nothing else but a nested family of intervals, so processing fuzzy information simply means processing several interval cases. Because of this, in the following text, we will concentrate on interval uncertainty.

**Propagation of interval uncertainty: usual approach.** If all we know about each input $x_i$ is that its value is located somewhere in the interval $\mathbf{x}_i = [\underline{x}_i,\overline{x}_i]$, then the only thing we can conclude about the value $y = f(x_1,\ldots,x_n)$ is that this value is somewhere in the set $\mathbf{y} = f(\mathbf{x}_1,\ldots,\mathbf{x}_n) \stackrel{\text{def}}{=} \{f(x_1,\ldots,x_n) : x_1 \in \mathbf{x}_1,\ldots,x_n \in \mathbf{x}_n)\}$. The task of computing this set is known as *interval computations*; see, e.g., [4].

**Monotonic case.** In many practical situations, the function $f(x_1,\ldots,x_n)$ is monotonic with respect to each of its variables. For example, it can be non-strictly increasing with respect to each input – the simple example of such function is addition $f(x_1,x_2) = x_1+x_2$.

In this monotonic case, the largest possible value of this function is attained when all the inputs $x_i$ attain their largest value $x_i = \overline{x}_i$, and the smallest possible value of this function is attained when all the inputs $x_i$ attain their smallest value $x_i = \overline{x}_i$. Thus, in this case, we have $f([\underline{x}_1,\overline{x}_1],\ldots,[\underline{x}_n,\overline{x}_n]) = [f(\underline{x}_1,\ldots,\underline{x}_n), f(\overline{x}_1,\ldots,\overline{x}_n)]$. For example, for addition, we get $[\underline{x}_1,\overline{x}_1]+[\underline{x}_2,\overline{x}_2] = [\underline{x}_1+\underline{x}_2,\overline{x}_1+\overline{x}_2]$.

**Case of relatively small measurement errors.** Another important case is the practically important case if when the measurement errors are relatively small. In general, by definition of the measurement error $\Delta y = \widetilde{y}-y$, we have $\Delta y = f(\widetilde{x}_1,\ldots,\widetilde{x}_n) - f(x_1,\ldots,x_n)$. By definition of the measurement error, we have $\Delta x_i = \widetilde{x}_i - x_i$, so $x_i = \widetilde{x}_i - \Delta x_i$. Substituting these expressions for $x_i$ into the above formula, we get

$$\Delta y = f(\widetilde{x}_1,\ldots,\widetilde{x}_n) - f(\widetilde{x}_1-\Delta x_1,\ldots,\widetilde{x}_n-\Delta x_n). \quad (1)$$

Since the measurement errors $\Delta x_i$ are small, terms which are quadratic (or of higher order) with respect to $\Delta x_i$ can be safely ignored in comparison with the linear terms. For example, if $\Delta x_i \approx 10\%$, then $(\Delta x_i)^2 \approx 1\%$ – which is indeed much smaller. Thus, we can expand the right-hand side of the expression (1) in Taylor series and ignore quadratic and higher order terms in the corresponding expansion. As a result, we get the following expression: $\Delta y = c_1\cdot\Delta x_1+\ldots+c_n\cdot\Delta x_n$, where $c_i \stackrel{\text{def}}{=} \dfrac{\partial f}{\partial x_i}_{|x_1=\widetilde{x}_1,\ldots,x_n=\widetilde{x}_n}$. Each term $c_i\cdot\Delta x_i$ is monotonic in $\Delta x_i$: increasing if $c_i \geq 0$ and decreasing if $c_i < 0$. So, when $\Delta x_i$ is in the interval $[\underline{\Delta}_i,\overline{\Delta}_i]$, the largest value $\overline{a}_i$ of each term $c_i\cdot\Delta x_i$ is equal to $c_i\cdot\overline{\Delta}_i$ if $c_i \geq 0$, and to $c_i\cdot\underline{\Delta}_i$ if $c_i < 0$. The largest possible value $\overline{\Delta}$ of $\Delta y$ can be obtained by adding these $n$ largest values: $\overline{\Delta} = \overline{a}_1+\ldots+\overline{a}_n$.

Similarly, the smallest value $\underline{\Delta}$ of each term $c_i\cdot\Delta x_i$ is equal to $c_i\cdot\underline{\Delta}_i$ if $c_i \geq 0$, and to $c_i\cdot\overline{\Delta}_i$ if $c_i < 0$. The smallest possible value $\underline{\Delta}$ of $\Delta y$ can be obtained by adding these $n$ smallest values: $\underline{\Delta} = \underline{a}_1+\ldots+\underline{a}_n$.

**Problem with the usual approach.** The usual approach corresponds to the most pessimistic case, when we consider the worst possible scenario. For example, in the case of addition, we consider the case when each input $x_i$ attain its largest possible value – i.e., that the correpsonding measurement error attains its smallest possible value. We have already mentioned that even for a single variable, it is not very probable that the measurement error will attain its smallest possible

value. That two such rare events will both happen is highly improbable. It is therefore desirable to come up with more realistic estimates of the results of data processing. This is what we will be doing in this paper.

## 2 From Pessimistic to Optimistic Approach

**Idea.** We consider situations in which all we know for each input $x_i$ is that its values are located on the corresponding interval $[\underline{x}_i, \overline{x}_i]$. We do not have any information about the relation between different inputs. If they change independently from each other, then yes, it is possible that both inputs $x_i$ attain their largest values, and it is possible that they both will attain their smallest values, so we get the *widest* possible interval – which we called pessimistic.

But it may happen that the values $x_i$ are related, so that when $x_1$ is the largest, $x_2$ must be the smallest and vice versa. In this case, we get the narrowest possible interval. It is therefore reasonable to consider the *optimistic* approach – i.e., to consider the *narrowest* possible interval. Let us describe this idea in precise terms.

**Towards the precise formulation of the optimistic approach.** We know that for each input $x_i$, both endpoints $\underline{x}_i$ and $\overline{x}_i$ are possible – if one or both of these values was not possible, we would have a narrower interval of possible values of $x_i$. We have no such confidence about the possibility of intermediate values. So, to minimize the set of all possible values of $y = f(x_1, \ldots, x_n)$, let us consider the case when only these two values are possible for each input $x_i$.

Then, the set $X$ of possible tuples $x = (x_1, \ldots, x_n)$ must have the property that for each $i$, at least one tuple must contain $\underline{x}_i$, and at least one tuple must contain $\overline{x}_i$. For each set with this property, we can consider the interval hulls generated by all the corresponding values $f(x)$. The intersection of all these interval is thus the narrowest possible interval for $y$. Let us describe it in terms of a precise definition.

**Definition 1.** *Let $\mathbf{x}_1 = [\underline{x}_1, \overline{x}_1]$, …, $\mathbf{x}_n = [\underline{x}_n, \overline{x}_n]$ and let $f(x_1, \ldots, x_n)$ be a real-valued function of n real variables. We say that a set X of tuples $x = (x_1, \ldots, x_n)$ is* possible *if the following three conditions are satisfied:*

- *for each tuple $x \in X$ and for each i, the i-th element for every i is equal either to $\underline{x}_i$ or to $\overline{x}_i$;*
- *for each i, there exists a tuple $x \in X$ for which $x_i = \underline{x}_i$; and*
- *for each i, there exists a tuple $x \in X$ for which $x_i = \overline{x}_i$.*

*For each possible set X, we define the interval $f(X)$ as $\left[\min_{x \in X} f(x), \max_{x \in X} f(x)\right]$. By the* optimistic range *$f_o(\mathbf{x}_1, \ldots, \mathbf{x}_n)$ of the function $f(x_1, \ldots, x_n)$ on the intervals $\mathbf{x}_1, \ldots, \mathbf{x}_n$, we mean the intersection of the intervals $f(X)$ for all possible tuples X.*

**Algorithms for computing the optimistic range: general comment.** There are finitely many possible tuples – namely, $2^n$ of them. Thus, there exist finitely many sets of such tuples. In principle, we can therefore test all possible sets of such tuples, and thus, find all the intervals $f(X)$ and their intersection. So, in principle, the optimistic range is computable. Of course, this computation requires unrealistic exponential time. However, we will show that in many practical cases, there are feasible algorithms for computing the optimistic range. Let us describe some of these algorithms.

**Algorithm for computing the optimistic range: case of $n = 2$.** For $n = 2$, we get the following result.

**Proposition 1.** *For $n = 2$, $f_o([\underline{x}_1, \overline{x}_1], [\underline{x}_2, \overline{x}_2]) = [\underline{\Delta}, \overline{\Delta}]$, where $\underline{\Delta} = \max(\min(f(\underline{x}_1, \underline{x}_2), f(\overline{x}_1, \overline{x}_2)), \min(f(\underline{x}_1, \overline{x}_2), f(\overline{x}_1, \underline{x}_2)))$ and $\overline{\Delta} = \min(\max(f(\underline{x}_1, \underline{x}_2), f(\overline{x}_1, \overline{x}_2)), \max(f(\underline{x}_1, \overline{x}_2), f(\overline{x}_1, \underline{x}_2)))$.*

**Proof.** If a possible set $X$ is a proper subset of a possible set $X'$, then clearly $f(X)$ is a subinterval of the interval $f(X')$. Thus, to find the desired intersection, it is sufficient to consider minimal possible sets, i.e., possible sets for which no proper subset is possible.

For $n = 2$, for $i = 1$, in a possible set, there must be a tuple containing $\underline{x}_1$ and there must be a tuple containing $\overline{x}_1$. If these two tuples contain different endpoints of the 2nd interval, then we get two possible sets used in the formulation of the proposition: the sets $X_1 = \{(\underline{x}_1, \underline{x}_2), (\overline{x}_1, \overline{x}_2)\}$ and $X_2 = \{(\underline{x}_1, \overline{x}_1), (\overline{x}_1, \underline{x}_2)\}$.

Let us now consider the remaining cases when the two above-mentioned tuples contain the same endpoint of the 2nd interval. If this endpoint is $\underline{x}_2$, this means that we must have at least one more tuple containing $\overline{x}_2$. If this third tuple contains $\underline{x}_1$, then $X$ contains $(\overline{x}_1, \underline{x}_2)$ and $(\underline{x}_1, \overline{x}_2)$, i.e., contains the set $X_2$ – and is, thus, not minimal. If this third tuple contains $\overline{x}_1$, then $X$ contains $(\underline{x}_1, \underline{x}_2)$ and $(\overline{x}_1, \overline{x}_2)$, i.e., contains the set $X_1$ – and is, thus, not minimal.

Similarly, if the common endpoint is $\overline{x}_2$, this means that we must have at least one more tuple containing $\underline{x}_2$. If this third tuple contains $\underline{x}_1$, then $X$ contains $(\overline{x}_1, \overline{x}_2)$ and $(\underline{x}_1, \underline{x}_2)$, i.e., contains the set $X_1$ – and is, thus, not minimal. If this third tuple contains $\overline{x}_1$, then $X$ contains $(\underline{x}_1, \overline{x}_2)$ and $(\overline{x}_1, \underline{x}_2)$, i.e., contains the set $X_2$ – and is, thus, not minimal.

So, the only minimal possible sets are $X_1$ and $X_2$. One can check that the intersection of two corresponding intervals is exactly the expression from the formulation of Proposition 1. The proposition is proven.

**Case of a monotonic function.** For functions that are (non-strictly) increasing in both variables, we get a simpler result.

**Proposition 2.** *If a function $f(x_1,x_2)$ is (non-strictly) increasing in each of its variables, then $f_o([\underline{x}_1,\overline{x}_1],[\underline{x}_2,\overline{x}_2]) = [\underline{\Delta},\overline{\Delta}]$, where $\underline{\Delta} = \min(f(\underline{x}_1,\overline{x}_2),f(\overline{x}_1,\underline{x}_2))$ and $\overline{\Delta} = \max(f(\underline{x}_1,\overline{x}_2),f(\overline{x}_1,\underline{x}_2))$. This interval is equal to $f(X)$, for $X = \{(\underline{x}_1,\overline{x}_2),(\overline{x}_1,\underline{x}_2)\}$.*

**Proof.** This result follows from Proposition 1. Indeed, for an increasing function, we have $f(\underline{x}_1,\underline{x}_2) \le f(\overline{x}_1,\overline{x}_2)$, thus

$$\min(f(\underline{x}_1,\underline{x}_2),f(\overline{x}_1,\overline{x}_2)) = f(\underline{x}_1,\underline{x}_2).$$

Here, due to monotonicity, $f(\underline{x}_1,\underline{x}_2) \le f(\underline{x}_1,\overline{x}_2)$ and $f(\underline{x}_1,\underline{x}_2) \le f(\overline{x}_1,\underline{x}_2)$, thus

$$f(\underline{x}_1,\underline{x}_2) \le \min(f(\underline{x}_1,\overline{x}_2),f(\overline{x}_1,\underline{x}_2))$$

and so indeed

$$\underline{\Delta} = \min(f(\underline{x}_1,\overline{x}_2),f(\overline{x}_1,\underline{x}_2)).$$

The formula for $\overline{\Delta}$ is proven similarly.

Since addition is clearly increasing in each of the variables, we have the following corollary:

**Corollary.** $[\underline{x}_1,\overline{x}_1] +_o [\underline{x}_1,\overline{x}_2] = [\underline{\Delta},\overline{\Delta}]$, *where* $\underline{\Delta} = \min(\underline{x}_1+\overline{x}_2,\overline{x}_1+\underline{x}_2)$ *and* $\overline{\Delta} = \max(\underline{x}_1+\overline{x}_2,\overline{x}_1+\underline{x}_2)$.

*Comment.* These operations were first proposed by Kaucher and are thus known as *Kaucher arithmetic*; see, e.g., [3].

**Case of general $n$.** What happened with addition of intervals for the generic $n$ – which corresponds to the linearization case of uncertainty propagation? In this case, we have the following result.

**Definition 2.** *By a* midpoint *of an interval $[\underline{x}_i,\overline{x}_i]$, we mean the value $m_i = (\underline{x}_i+\overline{x}_i)/2$.*

*By a* half-width *(*radius*) of an interval $[\underline{x}_i,\overline{x}_i]$, we mean the value $r_i = (\overline{x}_i-\underline{x}_i)/2$.*

*Comments.* Once we know the midpoint and the radius, then we can reconstruct the original interval as $[\underline{x}_i,\overline{x}_i] = [m_i-r_i,m_i+r_i]$. Vice versa, when the interval already has the form $[\widetilde{x}_i-\Delta_i,\widetilde{x}_i+\Delta_i]$, where $\widetilde{x}_i$ is the measurement result and $\Delta_i$ is the bound on the absolute value of the measurement error, then the midpoint is exactly the measurement result and the radius $r_i$ is exactly the given error bound.

**Proposition 3.** *For addition $f(x_1,\ldots,x_n) = x_1+\ldots+x_n$ of intervals $[\underline{x}_i,\overline{x}_i]$, the optimistic range has the following form:*

- *If $2\max\limits_i r_i \ge \sum\limits_j r_j$, then $\sum\limits_o [\underline{x}_i,\overline{x}_i] = [m-r,m+r]$, where $m = m_1+\ldots+m_n$ and $r = 2\max\limits_i r_i - \sum\limits_j r_j$.*
- *If $2\max\limits_i r_i < \sum\limits_j r_j$, then $\sum\limits_o [\underline{x}_i,\overline{x}_i] = \emptyset$.*

**Proof.**

1°. Let us first consider the first case, when the inequality $2\max\limits_i r_i \ge \sum\limits_j r_j$ holds. Let $i_0$ denote the index of the interval with the largest radius: $r_{i_0} = \max\limits_i r_i$. Subtracting $r_{i_0}$ from both side of the above inequality, we get $r_{i_0} \ge \sum\limits_{j\ne i_0} r_j$, and we also get $r = 2\max\limits_i r_i - \sum\limits_j r_j = r_{i_0} - \sum\limits_{j\ne i_0} r_j$.

By definition, a possible set $X$ must include a tuple for which $x_{i_0} = \overline{x}_{i_0} = m_{i_0}+r_{i_0}$. For this tuple, for each $j \ne i_0$, we have $x_j \ge \underline{x}_j = m_j - r_j$, thus

$$x_1+\ldots+x_n \ge m_{i_0}+r_{i_0}+\sum_{j\ne i_0}(m_j-r_j) =$$

$$\sum_i m_i + \left(r_{i_0} - \sum_{j\ne i_0} r_j\right) = m-r.$$

Similarly, we conclude that a possible set $X$ must contain a tuple for which $\sum\limits_i x_i \ge m+r$. Thus, for each possible set $X$, we have $f(X) \supseteq [m-r,m+r]$. On the other hand, for

$$X = \{(\underline{x}_1,\ldots,\underline{x}_{i_0-1},\overline{x}_{i_0},\underline{x}_{i_0+1},\ldots,\underline{x}_n),$$

$$(\overline{x}_1,\ldots,\overline{x}_{i_0-1},\underline{x}_{i_0},\overline{x}_{i_0+1},\ldots,\overline{x}_n)\},$$

we have $f(X) = [m-r,m+r]$. Thus, indeed, the intersection of all the sets $f(X)$ is indeed the interval $[m-r,m+r]$.

2°. Let us now consider the second case, when $2\max\limits_i r_i < \sum\limits_j r_j$. Since each $r_i$ is smaller ot equalo than the maximum, this implies that $2r_i < \sum\limits_j r_j$. Subtracting $r_i$ from both sides of this inequality, we conclude that $r_i < \sum\limits_{j\ne i} r_j$.

Let us now consider the set $X_1$ consisting of all the tuples in which one of the values $x_i$ is equal to $\overline{x}_i$ and all other values $x_j$ are equal to $\underline{x}_j$. For each of these tuples, we have

$$x_1+\ldots+x_n = \overline{x}_i + \sum_{j\ne i}\underline{x}_j = m_i+r_i+\sum_{j\ne i}(m_j-r_j) =$$

$$\sum m_i + \left( r_i - \sum_{j \neq i} r_j \right) < \sum m_i = m.$$

Thus, for this set, the upper bound of $f(X_1)$ is smaller than $m$.

On the other hand, we can similarly prove that for the set $X_2$ consisting of all the tuples in which one of the values $x_i$ is equal to $\underline{x}_i$ and all other values $x_j$ are equal to $\overline{x}_j$, the upper bound of $f(X_2)$ is larger than $m$. Thus, the intersection of the intervals $f(X_1)$ and $f(X_2)$ is empty – and hence, the intersection of all the intervals $f(X)$ corresponding to possible sets $X$ is also empty. The proposition is proven.

**Limitations of the optimistic approach.** For two intervals, optimistic approach means that when $x_1$ attains its largest possible value $\overline{x}_1$, the quantity $x_2$ attains its smallest possible value $\underline{x}_2$. This is as improbable as the corresponding assumption in the pessimistic case.

Thus, both pessimistic and optimistic approaches are not realistic: the pessimistic approach is too pessimistic, it consider too many possible values of $\Delta y$, while the optimistic approach is too optimistic, it consider too few possible values of $\Delta y$. So, to get a realistic estimate, we need to combine the pessimistic and optimistic results into a single optimistic-pessimistic approach.

## 3 From Pessimistic and Optimistic to Realistic Approach: Case When the Optimistic Approach Leads to a Nonempty Interval

**Idea.** Let us first consider the case when the optimistic approach leads to a non-empty interval. In this case, we have two intervals: the pessimistic interval $A$ and the optimistic intervals $B$. We need to combine them into a single interval.

This problem is similar to the general problem that we started with: if we knew how to combine two numerical estimates $a$ and $b$ into a single numerical estimate $c(a,b)$, then a reasonable idea would be to consider the set of all possible values $c(a,b)$ when $a \in A$ and $b \in B$: $c(A,B) = \{c(a,b) : a \in A, b \in B\}$. To utilize this idea, we need to find reasonable ways to combine two numerical estimates.

**How to combine two numerical estimates.** What are the reasonable properties of the combination function $c(a,b)$? First, since we have two numerical estimates of the same quantity, a reasonable requirement is that the combined estimate should be within the interval generated by these two estimates:

$$\min(a,b) \leq c(a,b) \leq \max(a,b). \qquad (2)$$

The second reasonable property comes from the fact that often, what we estimate is a combination of two of more parts: e.g., we may be estimating the population of a state, which is equal to the sum of populations of all its counties. In particular, for the case of two parts, we have estimates $a_1$ and $b_1$ for the first part, and estimates $a_2$ and $b_2$ for the second part.

There are two possible ways to get a combined estimates for the sum of the two values We can first combine estimates for each part, into two combined estimates $c(a_1,b_1)$ and $c(a_2,b_2)$, and then add the resulting combined estimates into a single value $c(a_1,b_1) + c(a_2,b_2)$. Alternatively, we can first add the estimates for both parts, resulting in $a = a_1 + a_2$ and $b = b_1 + b_2$, and then combine these two estimates, resulting in $c(a,b) = c(a_1 + a_2, b_1 + b_2)$. It is reasonable to require that these two ways lead to the same value, i.e., that

$$c(a_1,b_1) + c(a_2,b_2) = c(a_1 + a_2, b_1 + b_2). \qquad (3)$$

**Definition 3.** *We say that a function $c(a,b)$ is a* combination rule *if it satisfies the conditions (2) and (3).*

**Proposition 4.** *A function is a combination rule if and only if it has the form $c(a,b) = \gamma \cdot a + (1-\gamma) \cdot b$ for some $\gamma \in [0,1]$.*

**Proof.** It is known – see, e.g., [1] – that every bounded function that satisfies the condition (3) is linear, i.e., has the form $c(a,b) = \gamma \cdot a + \delta \cdot b$ for some $\gamma$ and $\delta$. The condition (2) for $a = b$ implies that $\delta = 1 - \gamma$, and the same condition for $a \neq b$ implies that $\gamma \in [0,1]$.

*Comment.* If we additionally require that the combined value does not depend on the order of the estimates, i.e., that $c(a,b) = c(b,a)$, then, as one can easily check, the only possible value is $\gamma = 0.5$.

**So how to combine two intervals.** The above-described function is non-strictly increasing in both $a$ and $b$, thus we can use the formula for computing range of a monotonic function – that we described in Section 1 – and come up with an explicit expression for the combined interval:

$$c([\underline{a},\overline{a}],[\underline{b},\overline{b}]) = [\gamma \cdot \underline{a} + (1-\gamma) \cdot \underline{b}, \gamma \cdot \overline{a} + (1-\gamma) \cdot \overline{b}].$$

Interestingly, if we take the pessimistic interval as $[\underline{a},\overline{a}]$ and the optimistic interval as $[\underline{b},\overline{b}]$, we get the formulas for so-called *interactive addition for intervals*, formulas that have been successfully used in many applications; see, e.g., [5]. Thus, our analysis provides a new justification for these formulas.

## 4 From Pessimistic and Optimistic to Realistic Approach: Case When the Optimistic Approach Leads to the Empty Set

**What are the reasonable transformations: discussion.** But what can we do if the optimistic approach leads to the empty set? In this case, we only have one interval – the pessimistic one. So, to get a narrower realistic interval, we can only use this given interval. In other words, we need an operation $F$ that takes an interval $[\underline{a},\overline{a}]$ and returns its subinterval $F([\underline{a},\overline{a}]) \subseteq [\underline{a},\overline{a}]$.

To analyze what are the natural properties of this transformation, let us take into account that we are processing numerical values of the corresponding quantities, but these numerical values depends on the selection of the measuring unit, on the selection of the starting point, and – sometimes – on the choice of the sign (e.g., what is positive electric charge and what is negative is just a question of choice).

If we change the measuring unit to a new one which is $\lambda$ times smaller, then all numerical values are multiplies by $\lambda$: $x \mapsto \lambda \cdot x$; e.g., 2 meters becomes 200 centimeters. If we change the direction, then all values are multiplied by $-1$: $x \mapsto -x$. If we change the starting point to a new one which is $x_0$ units earlier, then we need to add $x_0$ to all numerical values: $x \mapsto x + x_0$. For example, Year $x = 2$ in the French revolution calendar – that started in 1789 – is year $2 + 1789 = 1791$ in the regular calendar. It general, if we make all these changes, then we get a linear transformation $x \mapsto \lambda \cdot x + x_0$ for all possible real numbers $\lambda \neq 0$ and $x_0$.

It is reasonable to require that the selection of a subinterval should not depend on these choices, i.e., that it should lead to the same subinterval no matter what scale we use.

**Definition 4.** *We say that a function $F$ that maps each interval into its subinterval is* scale-invariant *if for every interval $[\underline{a},\overline{a}]$ and for all possible real numbers $\lambda \neq 0$ and $x_0$, we have the following property:*

*if $F([\underline{a},\overline{a}]) = [\underline{b},\overline{b}]$, then we should have $F([\underline{A},\overline{A}]) = [\underline{B},\overline{B}]$, where $[\underline{A},\overline{A}] \stackrel{\text{def}}{=} \lambda \cdot [\underline{a},\overline{a}] + x_0$ and $[\underline{B},\overline{B}] \stackrel{\text{def}}{=} \lambda \cdot [\underline{b},\overline{b}] + x_0$.*

**Proposition 5.** *A function a function $F$ that maps each interval into its subinterval is scale-invariant if and only if it has the form $F([m-r,m+r] = [m-\gamma\cdot r, m+\gamma\cdot r]$ for some $\gamma \in [0,1]$.*

**Proof.** One can easily check that the above formula is indeed scale-invariant. Vice versa, let us assume that we have a scale-invariant transformation $F$, and let us proved that it is described by the above formula.

For degenerate intervals, i.e., intervals of the type $[a,a]$ with $r = 0$, the requirement that the result is a subinterval leads to $F([a,a]) = [a,a]$, exactly as the formula predicts. So, to complete the proof, it is sufficient to consider non-degenerate intervals $[m-r,m+r]$ with $r > 0$.

For the interval $[\underline{a},\overline{a}] = [-1,1]$, the requirement that the result $[\underline{b},\overline{b}] \stackrel{\text{def}}{=} F([-1,1]$ should not change under the transformation $x \mapsto -x$ implies that should have $-[\underline{b},\overline{b}] = [\underline{b},\overline{b}]$, i.e., that $\underline{b} = -\overline{b}$. If we denote $\overline{b}$ by $\gamma$, we thus get $F([-1,1]) = [-\gamma,\gamma]$. The requirement that the resulting interval be a subinterval of the original interval $[-1,1]$ leads to $\gamma \in [-1,1]$.

Now, every non-degenerate interval $[m-r,m+r]$ can be obtained from the interval $[-1,1]$ if we apply the transformation $x \mapsto r \cdot x + m$. Since we have $F([-1,1]) = [-\gamma,\gamma]$, scale-invariance implies that $F([r\cdot(-1)+m, r\cdot 1+m]) = [r\cdot(-\gamma)+m, r\cdot\gamma+m]$, i.e., exactly to the desired formula $F([m-r,m+r] = [m-\gamma\cdot r, m+\gamma\cdot r]$. The proposition is proven.

**Acknowledgments.** This work was supported in part by the National Science Foundation grant EAR-2225395, and by the AT&T Fellowship in Information Technology.

## References

[1] J. Aczél, J. Dhombres, *Functional Equations in Several Variables*, Cambridge University Press, 2008.

[2] R. Belohlavek, J. W. Dauben, G. J. Klir, *Fuzzy Logic and Mathematics: A Historical Perspective*, Oxford University Press, New York, 2017.

[3] E. Kaucher, "Uber Eigenschaften und Anwendungsmöglichkeiten der erweiterten Intervallrechnung und des hyperbolische Fastkörpers über R", *Computing*, 1977, Suppl. 1, pp. 81–94.

[4] G. Mayer, *Interval Analysis and Automatic Result Verification*, de Gruyter, Berlin, 2017.

[5] V. Wasques, A. Andrade, P. Zanineli, "Associative property of interactive addition for intervals: application in the malthusian model", *Proceedings of the 2023 Annual Conference of the North American Fuzzy Information Processing Society NAFIPS 2023*, Cincinnati, Ohio, May 31–June 2, 2023.