WHY INTERVALS? WHY FUZZY NUMBERS? TOWARDS A NEW JUSTIFICATION

Vladik KREINOVICH

Department of Computer Science,
University of Texas at El Paso, USA
E-mail: vladik@utep.edu

Abstract

The purpose of this paper is to present a new characterization of the set of all intervals (and of the corresponding set of fuzzy numbers). This characterization is based on several natural properties useful in mathematical modeling; the main of these properties is the necessity to be able to combine (fuse) several pieces of knowledge.

Keywords: interval uncertainty, fuzzy numbers, foundations

1 Interval Uncertainty and Interval Computations: A Brief Reminder

Why intervals: a practical explanation. One of the main source of information about the physical world is measurements; see, e.g., [20]. Measurements are never 100% accurate. As a result, the result $\widetilde{x}$ of the measurement is, in general, different from the (unknown) actual value x of the desired quantity. The difference $\Delta x \stackrel{\text{def}}{=} \widetilde{x} - x$ between the measured and the actual values is usually called a *measurement error*.

The manufacturers of a measuring device usually provide us with an upper bound Δ for the (absolute value of) possible errors, i.e., with a bound Δ for which we guarantee that $|\Delta x| \le \Delta$. The need for such a bound comes from the very nature of a measurement process: if no such bound is provided, this means that the difference between the (unknown) actual value x and the observed value $\widetilde{x}$ can be as large as possible. In other words, if we measure, say, a temperature to be 100, in reality, this temperature could be $> 10^3$ or even $> 10^6$ – such an uncertainty is reasonable for a guess but not for a measurement.

Since the (absolute value of the) measurement error $\Delta x = \tilde{x} - x$ is bounded by the given bound Δ, we can therefore guarantee that the actual (unknown) value of the desired quantity belongs to the interval $[\tilde{x} - \Delta, \tilde{x} + \Delta]$. For example, if the measured value of the temperature is $\widetilde{x} = 100$ and the upper bound on the measurement error is $\Delta = 10$, then we can guarantee that the actual value of the temperature x must be within the interval $[100 - 10, 100 + 10] = [90, 110]$.

Traditional probabilistic approach to describing measurement uncertainty. In many practical situations, we not only know the interval $[-\Delta, \Delta]$ of possible values of the measurement error; we also know the probability of different values Δx within this interval [20, 21, 23]. This knowledge underlies the traditional engineering approach to estimating the error of indirect measurement, in which we assume that we know the probability distributions for measurement errors Δx_i.

In practice, we can determine the desired probabilities of different values of Δx_i by comparing the results of measuring with this instrument with the results of measuring the same quantity by a standard (much more accurate) measuring instrument. Since the standard measuring instrument is much more accurate than the one use, the difference between these two measurement results is practically equal to the measurement error; thus, the empirical distribution of this difference is close to the desired probability distribution for measurement error.

Interval approach to measurement uncertainty. As we have mentioned, in many practical situations, we do know the probabilities of different values of the measurement error. There are two cases, however, when this determination is not done:

- First is the case of cutting-edge measurements, e.g., measurements in fundamental science. When a Hubble telescope detects the light from a distant galaxy, there is no “standard" (much more accurate) telescope floating nearby that we can use to calibrate the Hubble: the Hubble telescope is the best we have.

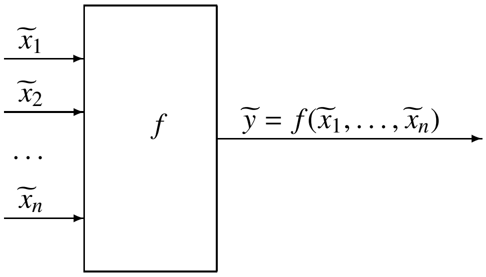
- The second case is the case of measurements on the shop floor. In this case, in principle, every sensor can be thoroughly calibrated, but sensor calibration is so costly – usually costing ten times more than the sensor itself – that manufacturers rarely do it.

In both cases, we have no information about the probabilities of Δx; the only information we have is the upper bound on the measurement error.

In this case, after performing a measurement and getting a measurement result $\widetilde{x}$, the only information that we have about the actual value x of the measured quantity is that it belongs to the interval $\mathbf{x} = [\widetilde{x} - \Delta, \widetilde{x} + \Delta]$.

Why indirect measurements. In the previous text, we considered an idealized situation when we can directly measure the value of the desired quantity.

In many real-life situations, we are interested in the value of a physical quantity y that is difficult or impossible to measure directly. Examples of such quantities are the distance to a star and the amount of oil in a given well. Since we cannot measure y directly, a natural idea is to measure y *indirectly*. Specifically, we find some easier-to-measure quantities $x_1, \ldots, x_n$ which are related to y by a known relation $y = f(x_1, \ldots, x_n)$; this relation may be a simple functional transformation, or complex algorithm (e.g., for the amount of oil, numerical solution to an inverse problem). Then, to estimate y, we first measure the values of the quantities $x_1, \ldots, x_n$, and then we use the results $\widetilde{x}_1, \ldots, \widetilde{x}_n$ of these measurements to to compute an estimate $\widetilde{y}$ for y as $\widetilde{y} = f(\widetilde{x}_1, \ldots, \widetilde{x}_n)$:



For example, to find the resistance R, we measure current I and voltage V, and then use the known relation $R = V/I$ to estimate resistance as $\widetilde{R} = \widetilde{V}/\widetilde{I}$.

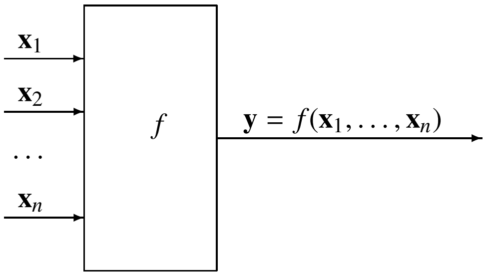
Computing an estimate for y based on the results of direct measurements is called *data processing*; data processing is the main reason why computers were invented in the first place, and data processing is still one of the main uses of computers as number crunching devices.

Comment. In this paper, for simplicity, we consider the case when the relation between x_i and y is known exactly; in some practical situations, we only known an approximate relation between x_i and y.

Why interval computations. As we have mentioned, measurements are never 100% accurate; as a result, the measured values $\widetilde{x}_i$ are, in general, different from the (unknown) actual values x_i of the measured quantities.

In particular, in the case of interval uncertainty, after we performed a measurement and got a measurement result $\widetilde{x}_i$, the only information that we have about the actual value x_i of the measured quantity is that it belongs to the interval $\mathbf{x}_i = [\widetilde{x}_i - \Delta_i, \widetilde{x}_i + \Delta_i]$. In such situations, the only information that we have about the (unknown) actual value of $y = f(x_1, \ldots, x_n)$ is that y belongs to the range $\mathbf{y} = [\underline{y}, \overline{y}]$ of the function f over the box $\mathbf{x}_1 \times \ldots \times \mathbf{x}_n$:

$$\mathbf{y} = [\underline{y}, \overline{y}] = \{f(x_1, \ldots, x_n) \,|\, x_1 \in \mathbf{x}_1, \ldots, x_n \in \mathbf{x}_n\}.$$



The process of computing this interval range based on the input intervals $\mathbf{x}_i$ is called *interval computations*; see, e.g., [5, 6, 17].

Possibility of linearization. In many practical situations, the dependence $y = f(x_1, \ldots, x_n)$ of the desired quantities y on the uncertain parameters x_i is reasonably smooth, and the measurement uncertainty Δx_i is relatively small. In such cases, we can safely linearize the dependence of y on x_i.

Specifically, since the function $f(x_1, \ldots, x_n)$ is reasonably smooth, and the inputs $x_i = \widetilde{x}_i - \Delta x_i$ differ only slightly from the known value Δ_i, we can thus ignore quadratic and higher order terms in the expansion of f and approximate the function f, in the vicinity of the approximate values $(\widetilde{x}_1, \ldots, \widetilde{x}_n)$, by its linear terms:

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

where

$$\Delta y \stackrel{\text{def}}{=} c_1 \cdot \Delta x_1 + \ldots + c_n \cdot \Delta x_n,$$
$$\widetilde{y} \stackrel{\text{def}}{=} f(\widetilde{x}_1, \ldots, \widetilde{x}_n), \quad c_i \stackrel{\text{def}}{=} \frac{\partial f}{\partial x_i}.$$

Linearization: resulting formula. One can easily show that when each of the variables Δx_i takes possible values from the interval $[-\Delta_i, \Delta_i]$, then the largest possible value of the linear combination Δy is

$$\Delta = |c_1| \cdot \Delta_1 + \ldots + |c_n| \cdot \Delta_n,$$

and the smallest possible value of δy is $-\Delta$. Thus, in this approximation, the interval of possible values of Δy is $[-\Delta, \Delta]$, and the desired interval of possible values of y is $[\widetilde{y} - \Delta, \widetilde{y} + \Delta]$.

2 Interval Uncertainty and Interval Computations: Traditional Challenges

2.1 First Challenge: Non-Linearity

In some practically important cases, non-linear terms cannot be ignored. In the previous text, we assumed that the measurement errors are small and therefore, terms quadratic in these errors can be ignored. This assumption is often justified: for example, if a measurement accuracy is 3%, then its square is $0.03^2 \approx 0.1\% \ll 3\%$ and therefore, can indeed be safely ignored.

In some practical situations, however, quadratic and higher order terms can no longer be ignored. For example, if the measurement accuracy is $\Delta x \approx 30\%$, then the square of Δx is $\approx 10\%$ – no longer much smaller than Δx.

Need to take non-linearity into account. It is therefore desirable to design new interval estimation techniques that would take the corresponding quadratic, cubic, etc., terms into account.

Difficulty. In general, the above problem is computationally difficult: even for quadratic functions, in general, computing the exact bound in case of interval uncertainty is an NP-hard problem [13, 22].

Successes. Getting reasonable interval estimates for y for non-linear functions under interval uncertainty is one of the main directions of interval computations [5, 6]. Researchers have designed several useful algorithms, and have successfully used these algorithms in numerous practical applications.

2.2 Second Challenge: Partial Information About Probabilities

Situation. So far, we have described two extreme situations:

- the case where we have a complete information on which values x_i (or Δx_i) are possible, and what the frequencies of different possible values are; in this text, we call this case *probabilistic uncertainty*;
- the case where we only know the range of possible values of x_i (or Δx_i), and we do not have any information about the frequencies at all; we call this case *interval uncertainty*.

In many real-life cases, we have an intermediate situation: we have some (partial) information about the frequencies (probabilities) of different values of x_i (or Δx_i), but we do not have the complete information about these frequencies.

Algorithms and successes. The partial information can be represented as bounds $[\underline{F}(x), \overline{F}(x)]$ on the (unknown) cumulative distribution function $F(x)$ (such bounds are called *p-boxes*), as bounds on moments, etc.

In many such situations, there exist efficient algorithms for processing such uncertainty, and many practical applications of these algorithms; see, e.g., [2, 10, 11, 12, 14, 24] and references therein.

3 Interval Uncertainty and Interval Computations: A New Challenge

Description of a new challenge: first approximation. In this paper, we discuss the following new challenge.

We have mentioned that since the measurement error $\Delta x = \tilde{x} - x$ is bounded by the manufacturer's bound Δ, we can guarantee that possible values of the measured physical quantity belong to the interval $[\tilde{x} - \Delta, \tilde{x} + \Delta]$. But are all values from this interval possible?

In other words, is the set X of all possible values of the desired quantity equal to this interval – or is it a proper subset of this interval?

More accurate formulation of the new challenge. Usually, the literal answer to the above question is "no". The reason for that is as follows: The manufacturer's bound is often an overestimate, because it is difficult to estimate Δ precisely, and so the manufacturer, because of his desire to *guarantee* the accuracy, prefers to give an upper estimate for this bound. Because of that, the actual set X is usually smaller than the interval $[\widetilde{x} - \Delta, \widetilde{x} + \Delta]$.

Assume now that we know the exact upper bound Δ_e for the error. This means that all possible values of x belong to an interval $[\tilde{x} - \Delta_e, \tilde{x} + \Delta_e]$, or, that the set X is a subset of this interval. Then our question is: is X equal to this interval? I.e., *is the set X an interval?*

Computational aspect of the new challenge. In some cases, the set X may not be an interval. Then, we must somehow approximate it. What family of sets should we use for this approximation?

4 What Is Known About This New Challenge: Description and Limitations

There are several results which justify the use of intervals.

4.1 Limit Approach

Description. Some results justify intervals along the same lines as normal distributions are justified in statistics:

- if we have many small independent errors, then, due to the Central Limit Theorem, the distribution for their sum is close to Gaussian;

- similarly, due to a special limit theorem, if the measurement error Δx is a sum of several small errors, each of which independently takes values in some set X_i, then the set of possible values for their sum is close to an interval; see, e.g., [9].

Limitations. This approach works well if we have already eliminated large error components and all remaining components of the measurement error are relatively small.

In many practical situations, we may still have error components which are much larger than others. In this case, the above approach does not work.

4.2 Consistency Approach

Description. Another justification of intervals come from the fact that intervals describe not only measurement uncertainty, but also uncertainty related to expert estimates. In such estimates, it is important to be able to check consistency.

If several experts present their estimates of possible values of the desired quantity x in term of intervals $[\underline{x}_i, \overline{x}_i]$, then checking consistency of this knowledge is easy: it is sufficient to check that all these intervals have a non-empty intersection, i.e., that $\max\limits_i \underline{x}_i \le \min\limits_j \overline{x}_j$. It turns out that intervals are the only sets for which such a feasible algorithm for checking consistency is possible; see, e.g., [15, 16].

Limitations. Expert estimates are sometimes educated guesses. As a result, the corresponding intervals do not necessarily contain the actual value of the desired quantity. Hence, these intervals may be (and often are) inconsistent.

In contrast, intervals coming from measurements are *guaranteed* to contain the actual values. For such intervals, there is always a non-empty intersection – because the actual value belongs to this intersection.

Of course, we may have inconsistent intervals, but this would simply mean that one of the measuring instrument has broken down. Detecting such a breakdown is an important practical problem, but it is such a rare event but we do not want to make it a foundation of our treatment of measurement-related uncertainty.

4.3 Invertibility Approach

Description. In standard arithmetic, if we, e.g., accidentally add a wrong number y to the preliminary result x, we can undo this operation by subtracting y from the result $x + y$. It turns out that a similar possibility to invert (undo) addition holds for intervals (although in case of intervals, we cannot simply undo addition by subtracting y from the sum).

It also turns out that if we add a single set that is not an interval, we lose invertibility. Thus, the invertibility requirement leads to a new characterization of the class of all intervals; see, e.g., [1, 8].

Limitations. For AI-type applications, when we explore possible data processing algorithms, it makes sense to try some algorithm and then "undo" it.

In data processing, the algorithm is usually well know and well established. So, while computer breakdowns do occur, and it is important to be able to recover from them, these breakdowns are rather rare events. Hence – similarly to the consistency approach – we do not want to make these events a foundation of our treatment of measurement-related uncertainty.

4.4 Summary

In short, the existing justifications of intervals are not fully helpful for measurement-related uncertainty. It is therefore desirable to come up with a new more convincing justification.

This is what we will do in this paper.

5 From Set and Interval Uncertainty to Fuzzy Sets and Fuzzy Numbers

Need for fuzzy values. In many real life situations, we cannot directly measure the values $x_1, \ldots, x_n$. Instead, we only have the *expert's estimates* of these values. Instead of operations with real numbers, we thus have to perform operations with these estimates.

Expert estimates are usually formulated in terms of words of natural language (e.g., "x is approximately equal to 1"). In order to apply computer operations to such estimates, we must first describe them in computer-understandable (numerical) terms.

One of the most natural ways to describe the expert's uncertain ("fuzzy") knowledge about a quantity X is to describe, for each real number x, our degree of belief that x is a possible value of the quantity X. This degree of belief is usually denoted by $\mu_X(x)$, and the corresponding description is called a *fuzzy set* (see, e.g., [7, 19]).

If we know the fuzzy sets that correspond to different inputs $X_1, \ldots, X_n$, then we will be able, using the well-known Zadeh's *extension principle* [7, 19], to describe the fuzzy set Y, i.e., in other words, to describe, for each real number y, how possible it is that this number is the actual value of Y. This description is very informative, but in many real-life situations (like in the oil example) we are not so much interested in this "fine structure" of our beliefs as in making a simple decision of what values of Y are possible and what values are not.

To make such a "binary" ("yes-no") decision, we must select some threshold degree of belief $\alpha \in (0, 1]$ and separate all possible real numbers y into two groups:

- For some real numbers y, our degree of belief that y is a possible value of Y exceeds (or is equal to) the threshold α ($\mu_Y(y) \geq \alpha$). We assume that such values y are *possible* for Y.
- For some other values y, our degree of belief that y is a possible value of Y is smaller than the threshold α ($\mu_Y(y) < \alpha$). We assume that such values y are *not possible* for Y.

The set of all y selected as possible is called the $\alpha-$*cut* of the corresponding fuzzy set and denoted by ${}^{\alpha}Y$.

This necessity to make a decision leads to the following natural alternative representation of a fuzzy set: to describe a fuzzy set X, for every $\alpha \in (0, 1]$, we describe the set ${}^{\alpha}X$ of all the values that will be assumed possible if take α as a threshold. This family of sets $\{{}^{\alpha}X\}$ is monotonic (${}^{\alpha}X \subseteq {}^{\beta}X$ if $\alpha \geq \beta$), and completely describes the original fuzzy set.

Operations on fuzzy sets. The above representation of a fuzzy set as a family of its $\alpha-$cuts is a natural background for defining operations on fuzzy sets, in particular, operations that correspond to standard arithmetic operations.

In order to define the result $X \circ Y$ of applying an operation $\circ$ (e.g., addition, subtraction, multiplication, etc.) to fuzzy sets X and Y, let us fix a threshold α and find out what values of $X \circ Y$ are possible for this particular threshold. For this threshold, only values from ${}^{\alpha}X$ are possible values of X, and only values from ${}^{\alpha}Y$ are possible values of Y. By applying the operation $\circ$ to all possible pairs $x \in {}^{\alpha}X$ and $y \in {}^{\alpha}Y$, we get the set of all possible values of $X \circ Y$. In other words, the $\alpha-$cut ${}^{\alpha}(X \circ Y)$ of the desired fuzzy set $X \circ Y$ has the following form ([7], Section 4.4):

$$ {}^{\alpha}(X \circ Y) = \{x \circ y \,|\, x \in {}^{\alpha}X,\ y \in {}^{\alpha}Y\}. $$

In particular, for addition ($\circ = +$), we have

$$ {}^{\alpha}(X + Y) = \{x + y \,|\, x \in {}^{\alpha}X,\ y \in {}^{\alpha}Y\}. $$

Comment. Under certain reasonable conditions, this definition is equivalent to the more standard one, that stems from the extension principle [3, 4, 18].

Fuzzy sets used in data processing. In different situations, different fuzzy sets are possible; for example, we sometimes only know that the value of a certain quantity X is "large". The greater the value x, the greater our degree of belief that this x is large, so the corresponding $\alpha-$cuts are semi-infinite intervals.

Such knowledge is possible; however, for data processing, such vague information is practically useless. Since in this paper, we are only interested in data processing applications, we will therefore restrict ourselves only to the fuzzy sets in which for every α, the $\alpha-$cut is *bounded*.

There is one more property that is natural to assume: if the values $x^{(1)}, x^{(2)}, \ldots, x^{(k)}, \ldots$ are all possible, and the sequence $x^{(k)}$ converges to a certain number x, then no matter how accurately we compute x, we will always find a number $x^{(k)}$ that is indistinguishable from x and possible. Therefore, it is natural to assume that this limit value x is also possible. In other words, it is natural to assume that every $\alpha-$cut contains all its limit points, i.e., that it is a *closed* set.

Combining these two conditions, we arrive at the assumption that each $\alpha-$cut is bounded and closed. On the real line, bounded and closed sets are exactly compact sets, so, we will call the fuzzy sets with such $\alpha-$cuts *compact fuzzy sets*.

Fuzzy numbers. An important particular case of a compact fuzzy set is a *fuzzy number* in which each $\alpha-$cut is a (closed) interval.

In fuzzy data processing, mainly fuzzy numbers are used; however, more general fuzzy sets are also sometimes needed: For example, if for some quantity x, with some degree of belief α, we know that x^2 belongs to the interval $[1, 4]$, and we know nothing about the sign of x, then the corresponding $\alpha-$cut (set of possible values of x) is not an interval, but a union of two disjoint intervals $[-2, -1] \cup [1, 2]$.

Justification of fuzzy numbers. There are some papers which justify fuzzy numbers. For example, in [1], the above invertibility justification of intervals has been extended to a justification of fuzzy numbers. However, this justification has the same limitation as the above justification of intervals. We therefore need a new more convincing justification.

6 Towards a New Justification of Intervals and Fuzzy Numbers

In this section, we provide a new justification of intervals – and thus, of fuzzy numbers as fuzzy sets for whom all α-cuts are intervals.

6.1 Motivations

Reminder: we are looking for a measurement-related justification. As we have mentioned, there are already exist justifications of intervals which are mainly oriented towards expert estimates. In this paper, we are therefore mainly interested in a measurement-related justification.

We are looking for a family of bounded closed sets. In the measurement case, the set of possible values of a quantity is always guarantee to be contained in an interval $[\widetilde{x} - \Delta, \widetilde{x} + \Delta]$. Thus, any set X of possible values of a quantity must be a subset of an interval – i.e., it must be a *bounded* set.

Similarly to the fuzzy case, we can also argue that it is sufficient to consider *closed* sets X. Thus, we are looking for a family of bounded closed sets.

Need to fuse several measurement results. If we are not satisfied with the accuracy of a single measurement, then it is natural to perform additional measurements.

After each measurement, we have a set X of possible values of this quantity which are consistent with the result of this measurement. The (unknown) actual value of the desired quantity must belong to this set X.

After several measurements, we have several set $X^1, X^2, \ldots, X^k$. We know that the desired value x must belong to each of these sets. Thus, the set of possible values x is the *intersection* of all these sets.

In other words, our family of sets must be closed under intersection.

Need for data processing. As we have mentioned, we are often interested not in the values of the directly measured quantities $x_1, \ldots, x_n$, but rather in the value of some quantity $y = f(x_1, \ldots, x_n)$.

In the linearized case, the dependence between $\Delta y = \widetilde{y} - y$ and $\Delta x_i = \widetilde{x_i} - x_i$ take a form $\Delta y = \sum\limits_{i=1}^{n} c_i \cdot \Delta x_i$. Hence, the dependence of y on x_i is also linear:

$$y = y_0 + \sum_{i=1}^{n} c_i \cdot x_i, \quad \text{where } y_0 \stackrel{\text{def}}{=} \widetilde{y_i} - \sum_{i=1}^{n} c_i \cdot \widetilde{x_i}.$$

Once we know the sets $X_1, \ldots, X_n$ of possible values of $x_1, \ldots, x_n$, then the set Y of possible values of Y takes the form

$$Y = \left\{ y_0 + \sum_{i=1}^{n} c_i \cdot x_i : x_1 \in X_1, \ldots, x_n \in X_n \right\}.$$

This set is called a *Minkowski linear combination* of the sets $X_1, \ldots, X_n$ and is denoted by $y_0 + c_1 \cdot X_1 + \ldots + c_n \cdot X_n$.

In other words, our family of sets must be closed under Minkowski linear combination.

Possibility to represent the set of values in a computer. We want to be able to represent these sets inside a computer. Inside a computer, we can represent only finitely many parameters. Thus, it is reasonable to require that our family of sets must be a finite-parametric family $X(a_1, \ldots, a_n)$, i.e., the result of a continuous mapping of a subset of R^n into the class of all sets.

On the class of all bounded closed sets, there is a natural metric – Hausdorff metric $d_H(X, Y)$. This metric is defined as the smallest $\varepsilon > 0$ for which X is contained in the ε-neighborhood of Y and Y is contained in the ε-neighborhood of X, i.e., for which

$$\forall x \in X\, \exists y \in Y\, (d(x, y) \le \varepsilon) \,\&\, \forall y \in Y\, \exists x \in X\, (d(x, y) \le \varepsilon),$$

where $d(x, y) = |x - y|$ is the standard distance between the points on the real line.

Thus, we require that the mapping that describes our family is continuous in terms of the (topology corresponding to) the Hausdorff metric.

Closed family of sets. Similarly to the requirement that each set from the family is closed, the family must also be closed in the sense of the Hausdorff metric.

Now, we are ready to formulate our main result.

6.2 Main Result

Definition. *We say that a family of bounded closed subsets of R is* finite-dimensional *if for some integer n, this family is an image of a subset of R^n under some continuous mapping (continuous in the sense of Hausdorff metric on the set of all closed bounded sets).*

Theorem. *Let $\mathcal{F}$ be a non-empty closed finite-dimensional family of bounded closed sets of R which is closed under intersection and Minkowski linear combination. Then, $\mathcal{F}$ is either the family of all one-point sets or the family of all intervals.*

Comment. Thus, if we exclude the case when all the values are known exactly, we get a new justification for intervals.

Proof.

1°. Let us first consider the case when $\mathcal{F}$ contains only 1-point sets. Let us prove that in this case, the family $\mathcal{F}$ coincides with the family of all 1-point sets.

Indeed, let $X = \{x\} \in \mathcal{F}$ be any set from the family $\mathcal{F}$. Due to closeness under Minkowski linear combination, the family $\mathcal{F}$ contains sets $y_0 + X$ for all $y_0 \in R$. In particular, for every $x' \in R$, we can take $y_0 = x' - x$ and conclude that $\{x'\} \in \mathcal{F}$. Thus, $\mathcal{F}$ is the family of all possible 1-point sets.

2°. Let us consider the case when not all sets from $\mathcal{F}$ are 1-point sets. In this case, $\mathcal{F}$ contains a set X with at least two points. For this set, $\inf X < \sup X$.

Let us prove that the family $\mathcal{F}$ contains a set Y with $\inf Y = 0$ and $\sup Y = 1$.

Indeed, we can take $Y = y_0 + c \cdot X$, with $y_0 = -\inf X$ and

$$c = \frac{1}{\sup X - \inf X}.$$

Since every set from $\mathcal{F}$ is closed, the set Y contains its own inf and sup, so $\{0, 1\} \subseteq Y$.

3°. Let us prove that the family $\mathcal{F}$ contains the interval $[0, 1]$.

Indeed, let Y be the set from Part 2 of this proof. For every m, the family $\mathcal{F}$ contains the set

$$Y_m \stackrel{\text{def}}{=} \frac{1}{m} \cdot Y + \ldots + \frac{1}{m} \cdot Y \ \ (m \text{ times}).$$

Every element of Y_m has the form $(y_1 + \ldots + y_m)/m$, where $y_i \in Y$. Since $\inf Y = 0$ and $\sup Y = 1$, every element from Y is between 0 and 1. Thus, every element from Y_m is also between 0 and 1. By taking values $x_i = 0$ and 1, we conclude that $\{0, 1/m, 2/m, \ldots, 1\} \subseteq Y_n$. Thus, $\{0, 1/m, 2/m, \ldots, 1\} \subseteq Y_m \subseteq [0, 1]$. When $m \to \infty$, we have $\{0, 1/m, 2/m, \ldots, 1\} \to [0, 1]$, thus $Y_m \to [0, 1]$. Since $\mathcal{F}$ is a closed family, we thus conclude that $[0, 1] \in \mathcal{F}$.

4°. Let us prove that the family $\mathcal{F}$ contains an arbitrary interval $[a, b]$, with $a \le b$.

Indeed, due to closeness under Minkoswki linear combination, the family $\mathcal{F}$ contains $a + (b - a) \cdot [0, 1]$, which is exactly $[a, b]$.

5°. We have just proven that the family $\mathcal{F}$ contains all intervals. To complete our proof, it is sufficient to show that it does not contain any set which is not an interval. We will prove this by reduction to a contradiction. Let us assume that the family $\mathcal{F}$ contains a set S which is not an interval.

5.1°. Let us first prove that the family $\mathcal{F}$ contains a 2-point set.

Indeed, since the set S is not an interval, this means that there exists an element $s_0 \in [\inf S, \sup S]$ for which $s \notin S$. Let

$$s^+ \stackrel{\text{def}}{=} \inf\{s \in S : s > s_0\}$$

and

$$s^- \stackrel{\text{def}}{=} \sup\{s \in S : s < s_0\}.$$

Since S is a closed set, it contains both s^- and s^+. By definition of s^- and s^+, the set S cannot contain any elements from the open interval (s^-, s^+). Thus, $S \cap [s^-, s^+] = \{s^-, s^+\}$.

According to Part 4 of our proof, the family $\mathcal{F}$ contains the interval $[s^-, s^+]$. Since the family $\mathcal{F}$ is closed under intersection, it contains $S \cap [s^-, s^+] = \{s^-, s^+\}$.

5.2°. Let us first prove that the family $\mathcal{F}$ contains an arbitrary 2-point set $\{a, b\}$, with $a < b$.

This conclusion follows from the fact that $\{a, b\} = y_0 + c \cdot \{s^-, s^+\}$, where y_0 and $c > 0$ are solutions to the system of linear equations $a = y_0 + c \cdot s^-$ and $b = y_0 + c \cdot s^+$, i.e., $c = (b - a)/(s^+ - s^-)$ and $y_0 = a - c \cdot s^-$.

5.3°. Let us now get the desired contradiction.

Let n be the dimension of the family $\mathcal{F}$. From Part 5.2 of this proof, it follows that for every $i = 0, 1, 2, \ldots$, we have $\{0, 2^i\} \in \mathcal{F}$. Since the family $\mathcal{F}$ is closed under Minkowski linear combination, we conclude that for every $c_0, \ldots, c_n$, we have

$$C(c_0, \ldots, c_n) \stackrel{\text{def}}{=} c_0 \cdot \{0, 2^0\} + c_1 \cdot \{0, 2^1\} + \ldots + c_n \cdot \{0, 2^n\} \in \mathcal{F}.$$

For $c_i \approx 1$, all these sets are different: each coefficient c_i can be determined from the value $c_i \cdot 2^i \in C(c_0, \ldots, c_n)$, and these values (for $c_i \approx 1$) are all different. Thus, we have a subfamily $C(c_0, \ldots, c_n)$ which is determined by $n + 1$ parameters $c_0, \ldots, c_n$.

So, we have a non-degenerate $(n + 1)$-dimensional family of sets within the family $\mathcal{F}$ – which contradicts to our assumption that the family $\mathcal{F}$ is n-dimensional. This contradiction proves that the family $\mathcal{F}$ cannot contain a set which is not an interval and thus, coincides with the family of all intervals.

The theorem is proven.

7 Conclusion

In this paper, after reviewing interval uncertainty, we present computational challenges associated with interval uncertainty. After this, we describe several approaches to fuzzy sets and fuzzy numbers, and explain the intrinsic connection between interval computations, interval uncertainty, fuzzy sets, and fuzzy numbers. This connection is then used to provide a new justification of intervals and fuzzy numbers.

This justification is based on several natural properties useful in mathematical modeling; the main of these properties is the necessity to be able to combine (fuse) several pieces of knowledge.

The first type of fusion is related to the fact that in many practical situations, we have two (or more) different pieces of knowledge about the value of the same quantity. In the crisp case, each piece of knowledge can be described by a set of all the values which are consistent with this knowledge. In this case, the actual (unknown) value of the quantity belongs to the *intersection* of the corresponding sets. Thus, the class $\mathcal{F}$ of all the sets which are used to describe uncertainty must be closed under intersection.

The second type of fusion comes from the fact that in many practical situations, we have partial knowledge about two or more different quantities $x_1, \ldots, x_n$, and we want to describe the resulting knowledge about a quantity

y which is related to $x_1, \ldots, x_n$ by a known dependence $y = f(x_1, \ldots, x_n)$. In the simplest case when the dependence f is linear, we conclude that the class $\mathcal{F}$ must be closed under the corresponding set operation – which is called Minkowski linear combination.

It turns out that the class of all intervals is, in effect, the only class which is closed under intersection and Minkowski linear combination; thus, in the fuzzy case, the class of fuzzy numbers can be similarly justified. So, we provide a new theoretic foundation for computationally merging interval computation and fuzzy logic in computational intelligence.

Acknowledgments

This work was supported in part by NASA under cooperative agreement NCC5-209, NSF grants EAR-0225670 and DMS-0532645, Star Award from the University of Texas System, and Texas Department of Transportation grant No. 0-5453.

The author is thankful to the participants of IEEE Symposium on Foundations of Computational Intelligence FOCI'07 for valuable suggestions.

References

[1] Bouchon-Meunier B., Kosheleva O., Kreinovich V., Nguyen H.T., 1997, *Fuzzy numbers are the only fuzzy sets that keep invertible operations invertible*, Fuzzy Sets and Systems, Vol.91, No.2, pp. 155-163.

[2] Ferson S., 2002, *RAMAS Risk Calc 4.0: Risk Assessment with Uncertain Numbers*, CRC Press, Boca Raton, Florida.

[3] Herencia J.A., 1996, *Graded sets and points: a stratified approach to fuzzy sets and points*, Fuzzy Sets and Systems, Vol.77, pp. 191-202.

[4] Herencia J.A., 1997, *Graded numbers and graded convergence of fuzzy numbers*, Fuzzy Sets and Systems, Vol.88, No.2, pp. 183-194.

[5] Interval computations website http://www.cs.utep.edu/interval-comp

[6] Jaulin L., Keiffer M., Didrit O., Walter E., 2001, *Applied Interval Analysis*, Springer-Verlag, London.

[7] Klir G., Yuan B., 1995, *Fuzzy sets and fuzzy logic: theory and applications*, Prentice Hall, Upper Saddle River, New Jersey.

[8] Kosheleva, O., Kreinovich V., 1999, *Only Intervals Preserve the Invertibility of Arithmetic Operations*, Reliable Computing, Vol.5, No.4, pp. 385-394.

[9] Kreinovich V., 1995, *Why intervals? A simple limit theorem that is similar to limit theorems from statistics*, Reliable Computing, Vol.1, No.1, pp. 33-40.

[10] Kreinovich V., 2004, *Probabilities, Intervals, What Next? Optimization Problems Related to Extension of Interval Computations to Situations with Partial Information about Probabilities*, Journal of Global Optimization, Vol.29, No.3, pp. 265-280.

[11] Kreinovich V., 2006, *Statistical Data Processing under Interval Uncertainty: Algorithms and Computational Complexity*, In: Lawry J., Miranda E., Bugarin A., Li S., Gil M.A., Grzegorzewski P., Hryniewicz O. (Eds.), Soft Methods for Integrated Uncertainty Modeling, Springer-Verlag, pp. 11-26.

[12] Kreinovich V., Ferson S., Ginzburg L., 2003, *Exact Upper Bound on the Mean of the Product of Many Random Variables With Known Expectations*, Reliable Computing, Vol.9, No.6, pp. 441-463.

[13] Kreinovich V., Lakeyev A., Rohn J., Kahl P., 1997, *Computational Complexity and Feasibility of Data Processing and Interval Computations*, Kluwer, Dordrecht.

[14] Kuznetsov V.P., 1991, *Interval statistical models*, Moscow, Radio i Svyaz Publ. (in Russian).

[15] Misane D., Kreinovich V., 1995, *A new characterization of the set of all intervals, based on the necessity to check consistency easily*, *Reliable Computing*, Vol.1, No.3, pp. 285-298.

[16] Misane D., Kreinovich V., 1996, *The necessity to check consistency explains the use of parallelepipeds in describing uncertainty*, Geombinatorics, Vol.5, No.3, pp. 109-120.

[17] Moore R.E., 1959, *Automatic error analysis in digital computation*, Technical Report Space Div. Report LMSD84821, Lockheed Missiles and Space Co.

[18] Nguyen H.T., 1978, *A note on the extension principle for fuzzy sets*, J. Math. Anal. and Appl., Vol.64, pp. 359-380.

[19] Nguyen H.T., and Walker E.A., 2005, *A first course in fuzzy logic*, CRC Press, Boca Raton, Florida.

[20] Rabinovich S., 2005, *Measurement Errors and Uncertainties: Theory and Practice*, Springer-Verlag, New York.

[21] Sheskin D.J., 2004, *Handbook of Parametric and Nonparametric Statistical Procedures*, Chapman & Hall/CRC, Boca Raton, Florida.

[22] Vavasis S.A., 1991, *Nonlinear Optimization: Complexity Issues*, Oxford University Press, New York.

[23] Wadsworth Jr. H.M. (ed.), 1990, *Handbook of statistical methods for engineers and scientists*, McGraw-Hill Publishing Co., New York.

[24] Walley P., 1991, *Statistical reasoning with imprecise probabilities*, Chapman and Hall, New York.