Journal of Uncertain Systems
Vol.12, No.3, pp.164-170, 2018
Online at: www.jus.org.uk

Towards Foundations of Interval and Fuzzy Uncertainty

Mahdokht Afravi[1], Kehinde Akinola[1], Fredrick Ayivor[1], Ramon Bustamante[1]
Erick Duarte[1], Ahnaf Farhan[1], Martha Garcia[1], Govinda K.C.[1]
Jeffrey Hope[1], Olga Kosheleva[1], Vladik Kreinovich[1,*], Jose Perez[1]
Francisco Rodriguez[1], Christian Servin[2], Eric Torres[1], Jesus Tovar[1]
[1] *University of Texas at El Paso, El Paso, TX 79968, USA*
[2] *El Paso Community College, 919 Hunter, El Paso, TX 79915 USA*

Received 1 December 2017; Revised 31 March 2018

Abstract

In this paper, we provide a theoretical explanation for many aspects of interval and fuzzy uncertainty: Why boxes for multi-D uncertainty? What if we only know Hurwicz's optimism-pessimism parameter with interval uncertainty? Why swarms of agents are better than clouds? Which confidence set is the most robust? Why μ^p in fuzzy clustering? How do degrees of confidence change with time? What is a natural interpretation of Pythagorean and fuzzy degrees of confidence?

Keywords: interval uncertainty, fuzzy uncertainty, foundations of uncertainty techniques

1 Formulation of the Problem

Uncertainty is ubiquitous. When we process data, this data comes either from measurement, or from expert estimates. Measurements are never 100% accurate, so there is always a difference $\Delta x = \widetilde{x} - x$ between the measurement result $\widetilde{x}$ and the actual (unknown) value x of the measured quantity; see, e.g., [17].

Simplest case: interval uncertainty. For measuring instruments, manufacturers usually provide an upper bound Δ on the absolute value of the measurement error Δx: $|\Delta x| \le \Delta$. If this is the only information that we have about the measurement uncertainty, then, based on the measurement result $\widetilde{x}$, the only conclusion that we can make about the actual (unknown) value x of the measured quantity is that x belongs to the interval $[\underline{x}, \overline{x}] \stackrel{\text{def}}{=} [\widetilde{x} - \Delta, \widetilde{x} + \Delta]$. Because of this fact, such uncertainty is known as *interval uncertainty*; see, e.g., [7, 11, 13].

Probabilistic and fuzzy uncertainty. In addition to the upper bound on the measurement error Δx, we often also have additional information about different possible values of Δx.

- Sometimes, we know the frequencies (probabilities) of different values Δx; this is the case of *probabilistic uncertainty*.
- Sometimes, we do not know the actual frequencies, but we have the expert's estimates of how possible it is to have different values. This case is known as *fuzzy uncertainty*; see, e.g., [1, 8, 12, 15, 16, 20].

What we do in this paper. In this paper, we deal with foundational aspects of interval, probabilistic, and fuzzy uncertainty.

*Corresponding author.
Emails: mafravi@miners.utep.edu (M. Afravi), kaakinola@miners.utep.edu (K. Akinola), fayivor@miners.utep.edu (F. Ayivor), rabustamante2@miners.utep.edu (R. Bustamante), ejduarte2@miners.utep.edu (E. Duarte), afarhan@miners.utep.edu (A. Farhan), mgarcia75@miners.utep.edu (M. Garcia), gbkc@miners.utep.edu (K.C. Govinda), jshope2@miners.utep.edu (J. Hope), olgak@utep.edu (O. Kosheleva), vladik@utep.edu (V. Kreinovich), jmperez6@miners.utep.edu (J. Perez), farodriguez-cabral@miners.utep.edu (Francisco Rodriguez), emtorres6@miners.utep.edu (C. Servin), jmtovar2@miners.utep.edu (E. Torres), cservin@gmail.com (J. Tovar).

What are the foundational problems of interval, probabilistic, and fuzzy uncertainty? The ultimate goal of all data processing, of all the attempts to understand better how the world functions, is to make the world a better place. Based on the results of analyzing the data, we make an appropriate decision – to select an engineering design, to better control a vehicle, to make a more productive financial investment. From this viewpoint, it is important not just to process uncertain data, but also to make decisions under the corresponding uncertainty.

If we know the exact consequences of each action, then we can estimate the *utility* $u(a)$ of each alternative, and thus, we can select the alternative with the largest value of this utility; see, e.g., [4, 5, 10, 14, 18]. In case of uncertainty, we know, at best, an interval $[\underline{u}(a), \overline{u}(a)]$ of possible values of utility. How can we make decisions in this situations? In such case, decision theory recommends to maximize the value $\alpha \cdot \overline{u}(a) + (1-\alpha) \cdot \underline{u}(a)$, where $\alpha \in (0,1)$ reflects the degree of the decision maker's optimism or pessimism; see, e.g., [6, 9, 10].

To use this idea, we must know the exact value of α. However, in situations of interval uncertainty, we can only expect to know α also with interval uncertainty. How can we then make a decision? This is the problem that we consider in Section 2.

Interval uncertainty means that for each quantity, we have an interval of possible values. For data processing, we are interested not only in possible values of each quantity x_i, but also in possible values of the whole tuple $x = (x_1, \cdots, x_n)$. What n-dimensional sets should we consider? In interval computations, practitioners usually consider *boxes* $[\underline{x}_1, \overline{x}_1] \times \cdots \times [\underline{x}_n, \overline{x}_n]$, since boxes often lead to efficient results. But why? This is a problem that we consider in Section 3.

In Sections 4 and 5, we consider situations when, in addition to intervals, we have information about frequencies – i.e., the case of probabilistic uncertainty. In Section 4, we consider the problem of decision making under probabilistic uncertainty. Specifically, our analysis of decision making under probabilistic uncertainty explains why the swarm of agents performs better than the cloud of agents. In Section 5, we analyze how to translate probabilistic uncertainty – which often comes, as with normal distributions, without explicit bounds – into such bounds, i.e., into interval and set uncertainty. The corresponding sets are known as *confidence sets*. In this chapter, we analyze how to select confidence sets so that they will be the most robust, i.e., the least dependent on the corresponding uncertainty.

Finally, in Sections 6–8, we consider the case of fuzzy uncertainty. Many fuzzy techniques have been thoroughly analyzed, and most of them have a solid theoretical foundations, but some things still remain heuristic. One of the main such things is the empirical success of functions μ^p in fuzzy clustering. In Section 6, we describe a theoretical foundation for this empirical success.

In the traditional fuzzy logic, we describe the expert's degree of confidence in each statement by a single number – a number that does not change with time. Of course, in real life, people are reluctant to describe their fuzzy ideas in terms of precise numbers. A more adequate description involves using an interval of possible values such degrees; this is analyzed in Section 6. And, of course, people's opinions change; this is analyzed in Section 7.

2 What if We Only Know Hurwicz's Optimism-Pessimism Parameter with Interval Uncertainty?

Formulation of the problem. In many practical situations, we do not know the exact consequences of each possible action. As a result, instead of a single utility value u, we can characterize each possible action by the interval $[\underline{u}, \overline{u}]$ of possible utility values. In such situations, decision theory recommends to select an alternative for which the following combination is the largest: $\alpha \cdot \overline{u} + (1-\alpha) \cdot \underline{u} = \underline{u} + \alpha \cdot (\overline{u} - \underline{u})$, where the parameter α – known as Hurwicz's *optimism-pessimism* parameter – may be different for different people. The value $\alpha = 1$ corresponds to absolute optimists, the value $\alpha = 0$ describes a complete pessimist, and values between 0 and 1 describe reasonable decision makers.

The parameter α needs to be determined based on a person's preferences and decisions. Often, however, in different situations, the decisions of the same person correspond to somewhat different values α. As a result, instead of a single value α, we have the whole range $[\underline{\alpha}, \overline{\alpha}]$ of possible values. In this case, how should we make decisions?

Our solution. For each $\alpha \in [\underline{\alpha}, \overline{\alpha}]$, the interval $[\underline{u}, \overline{u}]$ is equivalent to $\underline{u} + \alpha \cdot (\overline{u} - \underline{u})$. This expression is monotonic in α, so in general, the original interval $[\underline{u}, \overline{u}]$ is equivalent to the following range of possible

values $[\underline{u}_1, \overline{u}_1] = [\underline{u} + \underline{\alpha} \cdot (\overline{u} - \underline{u}), \underline{u} + \overline{\alpha} \cdot (\overline{u} - \underline{u})]$. Similarly, the new interval is equivalent to $[\underline{u}_2, \overline{u}_2] = [\underline{u}_1 + \underline{\alpha} \cdot (\overline{u}_1 - \underline{u}_1), \underline{u}_1 + \overline{\alpha} \cdot (\overline{u}_1 - \underline{u}_1)]$, etc. At each step, the width of the original intervals decreases by the factor $\overline{\alpha} - \underline{\alpha}$; thus, $\overline{u}_k - \underline{u}_k = (\overline{\alpha} - \underline{\alpha})^k \cdot (\overline{u} - \underline{u})$. In the limit, the intervals $[\underline{u}_k, \overline{u}_k]$ tend to a single point

$$u = \underline{u} + \underline{\alpha} \cdot (\overline{u} - \underline{u}) + \underline{\alpha} \cdot (\overline{u} - \underline{u}) \cdot (\overline{\alpha} - \underline{\alpha}) + \underline{\alpha} \cdot (\overline{u} - \underline{u}) \cdot (\overline{\alpha} - \underline{\alpha})^2 + \cdots$$

The corresponding geometric progression adds to $\underline{u} + \alpha \cdot (\overline{u} - \underline{u})$ for $\alpha = \underline{\alpha}/(1 - (\overline{\alpha} - \underline{\alpha}))$. This is the desired equivalent value of α for the case when we know α with interval uncertainty – and this is how we should make decisions in this case.

3 Why Boxes for Multi-D Uncertainty?

Formulation of the problem. In practice, we often do not have a complete information about the values of several quantities $x_1, \ldots, x_n$: several possible tuples $x = (x_1, \cdots, x_n)$ are consistent with our knowledge. How can we describe the set of possible values of x? In many practical applications, approximating this set by a box $[\underline{x}_1, \overline{x}_1] \times \cdots \times [\underline{x}_n, \overline{x}_n]$ leads to reasonable estimates. Why boxes? Why not other families of sets?

What are reasonable properties of the corresponding family F of approximating sets? Usually, we know some bounds on all x_i, so each set S should be bounded.

It is reasonable to require that each set $S \in F$ is closed: indeed, if $a = \lim a_n$ for $a_n \in S$, then for any measurement accuracy, a is indistinguishable from some possible $a_n \in S$ and thus, we will never be able to tell that a is not possible.

Similarly, the family F should be closed.

Also, often, the observed tuple consists of two independent components $x_i = y_i + z_i$; so, if the set Y of all y's and the set Z of all z's are possible, then the *Minkowski sum* $Y + Z \stackrel{\text{def}}{=} \{y + z : y \in Y, z \in Z\}$ is also possible. So, F should be closed under Minkowski sum.

Finally, the numerical values of each quantity x_i change if we change the starting point and/or measuring unit: $x_i \to a_i + b_i \cdot x_i$. These changes do not change what is possible and what is not, so the family F should be closed under these transformations.

Our result: Every family F that satisfies these properties contains all the boxes.

Discussion. Thus, the boxes form the simplest possible approximating family.

Idea of the proof. For each set S, we can take $S_1 = S$, $S_{k+1} = 0.5 \cdot (S_k + S_k)$, etc. In the limit, we get a convex hull of S.

By appropriate re-scaling $x_i \to b_i \cdot x_i$, we can shrink this set S in all directions except one i_0. In the limit, we get an interval parallel to the i_0-th axis. By shifting and re-scaling, we get all possible intervals parallel to this axis.

For n intervals $[\underline{x}_i, \overline{x}_i]$ each of which is parallel to the i-th axis and has all other coordinates 0s, their Minkowski sum is the box $[\underline{x}_1, \overline{x}_1] \times \ldots \times [\underline{x}_n, \overline{x}_n]$. Thus, each family F satisfying the above properties contains all the boxes. (Also, it is easy to prove that the class of all boxes satisfies these properties.)

4 Why Swarms of Agents are Better than Clouds?

Formulation of the problem. Expert estimates are often very crude and approximate. It is known, however, that if we ask several independent experts to estimate a quantity, and then average the resulting estimates, we often get a reasonably accurate estimate. This is the main idea behind *crowd intelligence*.

It was recently discovered that if we first divide the experts into small groups ("swarms"), let each group combine their estimates into a single value, and only average the group results, we get a much more accurate estimate of the quantity of interest; see, e.g., [19]. But why does this more complicated *swarm intelligence* lead to better results than the crowd intelligence?

Our explanation. Let $x_1, \ldots, x_n$ be estimates of different experts, and let $\rho(x)$ denote the probability density function that describes the probabilities of different values of the expert's approximation error. Then, ideally, we should select an estimate a with the largest possible likelihood $L \stackrel{\text{def}}{=} \prod_{i=1}^{n} \rho(x_i - a)$.

It is known that when $\rho(x)$ is a normal distribution, maximum likelihood estimate is indeed the arithmetic average $a = (\sum_{i=1}^{n} x_i)/n$. However, for other distributions, the maximum likelihood estimate is different from the arithmetic average – which explains why the crowd intelligence results are not perfect.

When we allow agents from each swarm s to come to an agreement between themselves, it is reasonable to assume that within each swarm, the agents will come up with a maximum likelihood (ML) estimate $a(s)$ for which the value $\prod_{i \in s} \rho(x_i - a(s))$ is the largest. It is known that when the number of combined estimates is reasonably large, the ML estimates $a(s)$ are approximately normally distributed. Thus, to combine values corresponding to different swarms, we can use the method which is optimal for normal distributions – i.e., taking the arithmetic average.

So, the estimate provided by swarm intelligence is asymptotically equal to the maximum likelihood one and is, thus, (asymptotically) optimal.

5 Which Confidence Set is the Most Robust?

Formulation of the problem. Once we have a probability distribution, then for each confidence level α, we can form a *confidence set* S, i.e., a set S for which $P(S) = \alpha$. There are many such sets, which one should we choose?

It is reasonable to select S to be connected. For each connected confidence set, we can define the degree of robustness by finding out how much the probability changes if we change the set slightly around some point x on the border ∂S of the set S. If we add or subtract a small piece of volume ΔV, then the change in probability ΔP can be found from the definition of the probability density function (pdf) $\rho(x) \stackrel{\text{def}}{=} \lim_{\Delta V \to 0} \frac{\Delta P}{\Delta V}$, as $\Delta P = \rho(x) \cdot \Delta V + o(\Delta V)$. Thus, for fixed (and small) ΔV, this change is proportional to $\rho(x)$.

The value of $\rho(x)$ is, in general, different for different points $x \in \partial S$. As a measure of robustness of the given set S, it is reasonable to select the worst-case value $r(S) \stackrel{\text{def}}{=} \max_{x \in \partial S} \rho(x)$. How can we select the confidence set with the smallest possible value of $r(S)$?

Solution. We show that the smallest possible value is attained when $\rho(x)$ is constant on ∂S. This means that for usual distributions like Gaussian, the corresponding confidence set is $S = \{x : \rho(x) \geq \rho_0\}$ for some ρ_0.

Discussion. For Gaussian distributions, such an ellipsoid is indeed usually selected as a confidence set.

Idea of the proof. If the function $\rho(x)$ is not constant, then we can expand S slightly in the vicinity of points x_0 for which $\rho(x_0) \approx \max_x \rho(x)$ – thus making the border value of $\rho(x)$ smaller, and at the same time shrink S slightly near all other points $x \in \partial S$, to compensate for the increase in $P(S)$ caused by expanding. After this shrinking, the border values of $\rho(x)$ at the corresponding points x will increase – but if we shrink a little bit, these values will still be smaller than $\max_x \rho(x)$. This way, we will get a new confidence set for which $\max_x \rho(x)$ is slightly smaller.

Thus, for the set S at which $r(S)$ is the smallest, the values of pdf cannot be non-constant on the border. So, for the optimal (most robust) confidence set, we indeed have $\rho(x) = \text{const}$ for all $x \in \partial S$.

6 Why μ^p in Fuzzy Clustering?

Formulation of the problem. One of the main algorithms for clustering n given d-dimensional points selects K "typical" values c_k and assignments $k(i)$ for each i from 1 to n so as to minimize the sum $\sum_i (x_i - c_{k(i)})^2$. This minimization is usually done iteratively.

- First, we pick c_k and assign each point x_i to the cluster k whose representative c_k is the closest to x_i.
- Then, we freeze $k(i)$ and select new typical representatives c_k by minimizing the objective function. This leads to c_k being an average of all the points x_i assigned to the k-th cluster.
- Then, the procedure repeats again and again – until the process converges.

In practice, for some objects, we cannot definitely assign them to a single cluster. In such cases, it is reasonable to assign, to each object i, degrees μ_{ik} of belongs to different clusters k, so that $\sum\limits_k \mu_{ik} = 1$. In this case, it seems reasonable to take each term $(x_i - c_k)^2$ with the weight μ_{ik}, i.e., to find the values μ_{ik} and c_k by minimizing the expression $\sum\limits_{i,k} \mu_{ik} \cdot (x_i - c_k)^2$. However, this expression is linear in μ_{ik}, and the minimum of a linear function under linear constraints is always at a vertex, i.e., when one value μ_{ik} is 1 and the rest are 0s. To come up with truly *fuzzy* clustering, with $0 < \mu_{ik} < 1$, we need to replace the factor μ_{ik} with a non-linear expression $f(\mu_{ik})$, i.e., to minimize $\sum\limits_{i,k} f(\mu_{ik}) \cdot (x_i - c_k)^2$. In practice, the functions $f(\mu) = \mu^p$ works the best; see, e.g., [2]. Why?

Our explanation. The weights μ_{ik} are normalized so that their sum is 1. So, if we delete some clusters or add more clusters, we need to re-normalize these values. A usual way to do it is to multiply them by a normalization constant c. It is therefore reasonable to require that the relative quality of different clustering ideas not change is we simply re-scale. This implies, e.g., that if $f(\mu_1) \cdot v_1 = f(\mu_2) \cdot v_2$, then after re-scaling $\mu_i \to c \cdot \mu_i$, we should have $f(c \cdot \mu_1) \cdot v_1 = f(c \cdot \mu_2) \cdot v_2$. We show that this condition implies that $f(\mu) = \mu^p$.

Indeed,

$$\frac{f(c \cdot \mu_2)}{f(c \cdot \mu_1)} = \frac{v_1}{v_2} = \frac{f(\mu_2)}{f(\mu_1)},$$

thus

$$r \stackrel{\text{def}}{=} \frac{f(c \cdot \mu_1)}{f(\mu_1)} = \frac{f(c \cdot \mu_2)}{f(\mu_2)}$$

for all μ_1 and μ_2. Thus, the ratio r does not depend on μ: $r = r(c)$, and $f(c \cdot \mu) = r(c) \cdot f(\mu)$. It is known that the only continuous solutions of this functional equations are $f(\mu) = C \cdot \mu^p$.

Why? We can rather easily solve the equation $f(c \cdot \mu) = r(c) \cdot f(\mu)$ when $f(\mu)$ is differentiable. Indeed, since $f(\mu)$ is differentiable, the ratio $r(c) = f(c \cdot \mu)/f(\mu)$ is also differentiable. If we differentiate both sides of the equation with respect to c, we get $\mu \cdot f'(c \cdot \mu) = r'(c) \cdot f(\mu)$. For $c = 1$, we get $\mu \cdot \dfrac{df}{d\mu} = p \cdot f$, where $p \stackrel{\text{def}}{=} r'(1)$. If we move all the terms containing f to one side and all others to another, we get $\dfrac{df}{f} = p \cdot \dfrac{d\mu}{\mu}$. Integrating, we get $\ln(f) = p \cdot \ln(\mu) + c_1$. If we apply exp to both sides, we get $f(\mu) = C \cdot \mu^p$, where $C = \exp(c_1)$.

Minimization is not affected if we divide the objective function by C and get $f(\mu) = \mu^p$.

7 How do Degrees of Confidence Change with Time?

Formulation of the problem. Situations change. As a result, our degrees of confidence in statements based on past experience decrease with time. How can we describe this decrease? If our original degree of confidence was a, what will be our degree of confidence $d_t(a)$ after time t?

It is clear that $d_t(a)$ should be increasing in a and decreasing in t, but there are many such functions.

Our idea. Let $f_{\&}(a, b)$ be an "and"-operation, i.e., a function that transforms degrees of confidence a and b in statements A and B into an estimate for our degree of confidence in $A \,\&\, B$. There are two ways to estimate our degree of confidence in $A \,\&\, B$ after time t:

- we can apply the function d_t to both a and b, and then combine them into $f_{\&}(d_t(a), d_t(b))$,
- or we can apply d_t directly to $f_{\&}(a, b)$, resulting in $d_t(f_{\&}(a, b))$.

It is reasonable to require that these two expressions coincide: $f_{\&}(d_t(a), d_t(b)) = d_t(f_{\&}(a, b))$.

Simplest case. If $f_{\&}(a, b) = a \cdot b$, then the above equality becomes $d_t(a \cdot b) = d_t(a) \cdot d_t(b)$. It is known that all monotonic solutions to this equation have the form $d_t(a) = a^{p(t)}$ for some $p(t)$.

The dependence on t can be found if we take into account that the decrease during time $t = t_1 + t_2$ can also be computed in two ways:

- directly, as $a^{p(t_1+t_2)}$,

- or by first considering decrease during t_1 ($a \to a' = a^{p(t_1)}$), and then a decrease during time t_2: $a' \to (a')^{p(t_2)} = \left(a^{p(t_1)}\right)^{p(t_2)} = a^{p(t_1)\cdot p(t_2)}$.

It is reasonable to require that these two expressions coincide, i.e., that $p(t_1 + t_2) = p(t_1) \cdot p(t_2)$. The only monotonic solutions to this equation are $p(t) = \exp(\alpha \cdot t)$, so we get $d_t(p) = a^{\exp(\alpha\cdot t)}$.

Comment. It is worth mentioning that for small t, we get $d_t(p) \approx a + \alpha \cdot t \cdot a \cdot \ln(a)$ and is thus related to entropy $-\sum a_i \cdot \ln(a_i)$.

General case. It is known that every "and"-operation can be approximated, with any accuracy, by a Archimedean one, i.e., by an operation of the type $f_{\&}(a.b) = g^{-1}(g(a) \cdot g(b))$. Thus, for re-scaled values $a' = g(a)$, we have $f_{\&}(a', b') = a' \cdot b'$, hence $d_t(a') = (a')^{\exp(\alpha\cdot t)}$ and, in the original scale,

$$d_t(a) = g^{-1}\left((g(a))^{\exp(\alpha\cdot t)}\right).$$

8 Towards a Natural Interval Interpretation of Pythagorean and Complex Degrees of Confidence

Formulation of the problem. Often, we only know the expert's degrees of confidence $a, b \in [0, 1]$ in statements A and B, and we need to estimate the expert's degree of confidence in $A \,\&\, B$. The algorithm $f_{\&}(a, b)$ providing the corresponding estimate is known as an *"and"-operation*, or a *t-norm.* One of the most frequently used "and"-operation is $\min(a, b)$. Similarly, one of the most frequently used "or"-operation is $\max(a, b)$.

Often, it is difficult for an expert to describe his/her degree of certainty by a single number a, an expert is more comfortable describing it by range (interval) $[\underline{a}, \overline{a}]$ of possible values. (An alternative way of describing this is as an *intuitionistic fuzzy degree*, i.e., a pair of values $\underline{a}$ and $1 - \overline{a}$.) If we know intervals $[\underline{a}, \overline{a}]$ and $[\underline{b}, \overline{b}]$ corresponding to a and b, then the range of possible degree of confidence in $A \,\&\, B$ is formed by values $\min(a, b)$ corresponding to all $a \in [\underline{a}, \overline{a}]$ and $b \in [\underline{b}, \overline{b}]$. Since $\min(a, b)$ is monotonic, this range has the form $[\min(\underline{a}, \underline{b}), \min(\overline{a}, \overline{b})]$. Similarly, the range for $A \vee B$ is $[\max(\underline{a}, \underline{b}), \max(\overline{a}, \overline{b})]$.

A recent paper [3] describes extensions $[\mathrm{absmin}(\underline{a}, \underline{b}), \mathrm{absmin}(\overline{a}, \overline{b})]$ and $[\mathrm{absmax}(\underline{a}, \underline{b}), \mathrm{absmax}(\overline{a}, \overline{b})]$ of the above definitions from $a, b \in [0, 1]$ to $a, b \in [-1, 1]$, where:

- $\mathrm{absmin}(a, b)$ is a if $|a| < |b|$, b is $|a| > |b|$, and $-|a|$ if $|a| = |b|$ and $a \neq b$,
- $\mathrm{absmax}(a, b)$ is a if $|a| > |b|$, b if $|a| < |b|$, and $|a|$ if $|a| = |b|$ and $a \neq b$.

These operations have nice properties, but what is their meaning?

Our explanation. In addition to *closed* intervals, let us consider *open* and *semi-open* ones. An open end will be then denoted by the negative number: e.g., $(0.3, 0.5]$ is denoted as $[-0.3, 0.5]$, and $(0.3, 0.5)$ as $[-0.3, -0.5]$.

By considering all possible cases, one can show that for two intervals $A = [\underline{a}, \overline{a}]$ and $B = [\underline{b}, \overline{b}]$, the range of possible values $\{\min(a, b) : a \in A, b \in B\}$ is indeed $[\mathrm{abs}\min(\underline{a}, \underline{b}), \mathrm{abs}\min(\overline{a}, \overline{b})]$. Similarly, the range of possible values of $\max(a, b)$ forms the interval $[\mathrm{abs}\max(\underline{a}, \underline{b}), \mathrm{abs}\max(\overline{a}, \overline{b})]$.

Thus, we get a natural interval explanation of the extended operations.

Acknowledgments

This work was supported in part by the National Science Foundation grant HRD-1242122 (Cyber-ShARE Center of Excellence).

The authors are thankful to all the participants of the 2017 UTEP/NMSU Workshop on Mathematics, Computer Science, El Paso, Texas, November 4, 2017, and Computation Science for valuable discussions.

References

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

[2] Bezdek, J.C., *A Primer on Cluster Analysis: 4 Basic Methods That (Usually) Work*, 1st Edition, Design Publishing, Sarasota, Florida, 2017.

[3] Dick, S., Yager, R., and O. Yazdanbakhsh, On Pythagorean and complex fuzzy set operations, *IEEE Transactions on Fuzzy Systems*, vol.24, no.5, pp.1009–1021, 2016.

[4] Fishburn, P.C., *Utility Theory for Decision Making*, John Wiley & Sons Inc., New York, 1969.

[5] Fishburn, P.C., *Nonlinear Preference and Utility Theory*, The John Hopkins Press, Baltimore, Maryland, 1988.

[6] Hurwicz, L., *Optimality Criteria for Decision Making Under Ignorance*, Cowles Commission Discussion Paper, Statistics, no.370, 1951.

[7] Jaulin, L., Kiefer, M., Didrit, O., and E. Walter, *Applied Interval Analysis, with Examples in Parameter and State Estimation, Robust Control, and Robotics*, Springer, London, 2001.

[8] Klir, G., and B. Yuan, *Fuzzy Sets and Fuzzy Logic*, Prentice Hall, Upper Saddle River, New Jersey, 1995.

[9] Kreinovich, V., Decision making under interval uncertainty (and beyond), *Human-Centric Decision-Making Models for Social Sciences*, edited by Guo, P., and W. Pedrycz, Springer Verlag, pp.163–193, 2014.

[10] Luce, R.D., and R. Raiffa, *Games and Decisions: Introduction and Critical Survey*, Dover, New York, 1989.

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

[12] Mendel, J.M., *Uncertain Rule-Based Fuzzy Systems: Introduction and New Directions*, Springer, Cham, Switzerland, 2017.

[13] Moore, R.E., Kearfott, R.B., and M.J. Cloud, *Introduction to Interval Analysis*, SIAM, Philadelphia, 2009.

[14] Nguyen, H.T., Kosheleva, O., and V. Kreinovich, Decision making beyond Arrows 'impossibility theorem', with the analysis of effects of collusion and mutual attraction, *International Journal of Intelligent Systems*, vol.24, no.1, pp.27–47, 2009.

[15] Nguyen, H.T., and E.A. Walker, *A First Course in Fuzzy Logic*, Chapman and Hall/CRC, Boca Raton, Florida, 2006.

[16] Novák, V., Perfilieva, I., and J. Močkoř, *Mathematical Principles of Fuzzy Logic*, Kluwer, Boston, Dordrecht, 1999.

[17] Rabinovich, S.G., *Measurement Errors and Uncertainty: Theory and Practice*, Springer Verlag, Berlin, 2005.

[18] Raiffa, H., *Decision Analysis*, Addison-Wesley, Reading, Massachusetts, 1970.

[19] Rosenberg, L., Baltaxe, D., and N. Pescetelli, Crowds vs. smarms, a comparison of intelligence, *Proceedings of the IEEE Swarm/Human Blended Intelligence Workshop SHBI'2016*, Cleveland, Ohio, October 21–23, 2016.

[20] Zadeh, L.A., Fuzzy sets, *Information and Control*, vol.8, pp.338–353, 1965.