Interval Uncertainty as the Basis for a General Description of Uncertainty: A Position Paper

Vladik Kreinovich
Department of Computer Science
University of Texas at El Paso
500 W. University
El Paso, TX 79968, USA
vladik@utep.edu

Abstract

Uncertainty is ubiquitous. Depending on what information we have, we get different types of uncertainty. For each type of uncertainty, techniques have been developed for efficient representation and processing of this uncertainty. However, the plethora of different uncertainty techniques is often confusing for practitioners. The situation is especially difficult in frequent situations when we need to gauge the uncertainty of the result of complex multi-stage data processing, and different data inputs are known with different types of uncertainty. To avoid this problem, it is necessary to develop and implement a general approach to representing and processing different types of uncertainty. In this paper, we argue that the most appropriate foundation for this general approach is interval uncertainty.

Uncertainty is ubiquitous. All the data comes either from measurements or from expert estimates. Neither measurements nor expert estimates are absolutely accurate, so we always have to deal with uncertainty; see, e.g., [8].

The situation is especially critical for dynamical and/or spatial data:

- For dynamical data, we not only have uncertainty about the corresponding values, we also have *temporal uncertainty*, i.e., uncertainty about the moment of time.
- For spatial data, we not only have uncertainty about the corresponding values, we also have *spatial uncertainty*, i.e., uncertainty about the spatial location.

Different types of uncertainty. Depending on what information we have, we have different types of uncertainty.

Probabilistic uncertainty. In the traditional approach, we assume that we know the probability of each possible value of the measurement error; in this case, we have *probabilistic* uncertainty [8].

Interval uncertainty. In many practical situations, we only know the upper bound Δ on the measurement error. In this case, based on the measurement result $\widetilde{x}$, the only information that we have about the actual (unknown) value x of the corresponding quantity is that x belongs to the interval $[\widetilde{x}-\Delta, \widetilde{x}+\Delta]$. This situation is known as *interval* uncertainty; see, e.g., [2, 4].

Case of imprecise probabilities. In some cases, in addition to the interval, we also have partial information about the probabilities; this case is handled by imprecise probability techniques.

Cased of imprecise ("fuzzy") uncertainty. Sometimes, we only have an expert estimate whose accuracy is described by imprecise ("fuzzy") words from natural language such as "approximate", "small", etc.; to describe such uncertainty, we can use *fuzzy techniques* specifically designed to handle this type of uncertainty; see, e.g., [3, 7].

Different types of uncertainty lead to a practical challenge. For each type of uncertainty, techniques have been developed for efficient representation and processing of this uncertainty. The plethora of different uncertainty techniques is often confusing for practitioners.

The situation is especially difficult in frequent situations when we need to gauge the uncertainty of the result of complex multi-stage data processing, and different data inputs are known with different types of uncertainty.

For example, in biomedical spatial data processing, we often need to combine measurement results – which are usually known with probabilistic or interval uncertainty – with the imprecise expert estimates of severity of different symptoms. This difficulty inhibits our ability to process such data, since at presence a service consuming two data sets and producing a third one may not know how to represent the uncertainty of its output dataset because the spatial uncertainty of the input datasets is of different type – and thus, have different representations.

What is needed to confront this challenge. To facilitate processing of uncertainty, especially dynamical and spatial uncertainty, it is necessary to develop and implement a general approach to representing and processing different types of uncertainty.

Our proposal. We believe – and we will try to convince the readers – that the most appropriate foundation for this general approach is interval uncertainty.

Interval uncertainty is the most fundamental type of uncertainty. Interval uncertainty is the most fundamental one: indeed, in measurements, no matter how partial our knowledge, we always know the upper bound on the measurement error – otherwise, if we cannot even guarantee an upper bound, this is not a measurement.

Once we know the upper bound, we know the interval.

Interval uncertainty can be used as a basis for all other types of uncertainty. From the computational viewpoint, interval uncertainty can be used as a basis for all other types of uncertainty.

For example, often, in addition to the interval, we have partial information about the probabilities. The *exact* information about the probability distribution means that we know the exact values of the probability density function, cumulative distribution function (cdf) $F(x) \stackrel{\text{def}}{=} \text{Prob}(X \le x)$, moments, etc. *Partial* information means that instead of the exact values of these statistical characteristics c, we only know *bounds* $\underline{c}$ and $\overline{c}$ on these values, i.e., we only know *intervals* $[\underline{c}, \overline{c}]$ of possible values of these characteristics.

For example, the frequently used technique for describing imprecise probabilities is the technique of p-boxes (probability boxes), in which for every x, we only know the interval $[\underline{F}(x), \overline{F}(x)]$ of possible values of cdf $F(x)$; see, e.g. [1, 6].

Similarly, it is known that fuzzy knowledge can be equivalently described by listing intervals of possible values (known as "alpha-cuts") corresponding to different degree of certainty; see, e.g., [3, 5, 7].

Proposal. We therefore propose to use interval-based representations as a uniform way of representing uncertainty, especially dynamical and spatial uncertainty.

The main benefit of this uniform representation it will facilitate aggregation of different types of uncertainty and thus, help with the propagation of uncertainty through data processing algorithms.

In computer terms, this uniform representation will facilitate communication between services producing data products – and between these services and the users of these services.

Acknowledgments. This work was supported in part by the National Science Foundation grants HRD-0734825 and HRD-1242122 (Cyber-ShARE Center of Excellence) and DUE-0926721, by Grant 1 T36 GM078000-01 from the National Institutes of Health, and by a grant on F-transforms from the Office of Naval Research.

The author is thankful to Paulo Pinheiro da Silva for valuable discussions.

References

[1] S. Ferson, V. Kreinovich, J. Hajagos, W. Oberkampf, and L. Ginzburg, *Experimental Uncertainty Estimation and Statistics for Data Having Interval Uncertainty*, Sandia National Laboratories, 2007, Publ. 2007-0939.

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

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

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

[5] H. T. Nguyen and V. Kreinovich, "Nested intervals and sets: concepts, relations to fuzzy sets, and applications", In: R. B. Kearfott et al (eds.), *Applications of Interval Computations*, Kluwer, Dordrecht, 1996, pp. 245–290.

[6] H. T. Nguyen, V. Kreinovich, B. Wu, and G. Xiang, *Computing Statistics under Interval and Fuzzy Uncertainty*, Springer Verlag, 2012.

[7] H. T. Nguyen and E. A. Walker, *First Course In Fuzzy Logic*, CRC Press, Boca Raton, Florida, 2006.

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