Symmetry approach explains the computational advantage of intuitionistic fuzzy sets over their interval-valued reformulation

Olga Kosheleva[1] and Vladik Kreinovich[2]
Departments of [1]Teacher Education and [2]Computer Science
University of Texas at El Paso
500 W. University
El Paso, Texas 79968, USA
olgak@utep.edu, vladik@utep.edu

Abstract

We explain the main ideas behind the symmetry approach – which is one of the main tools of modern physics, and we show that this approach can explain the computational advantage of intuitionistic fuzzy logic in comparison with its interval-valued reformulation.

1 Why symmetries?

In this paper, we will use symmetry approach – one of the main tools of modern physics. Since many people in the computing community are not well familiar with this approach, let us briefly explain its main ideas.

Why symmetries. One of the main objectives of science and engineering is to predict future events – and to make sure that these future events are as beneficial to the humankind as possible. In general, what is the reason why we can make predictions in the first place?

For example, why are we sure that if we drop a pen (or any other object), it will start falling down with the acceleration of 9.81 m/sec^2? Because the same experiment was performed at different locations on the Earth, with the experimenters facing different directions, and in all these cases, we observed the same phenomenon. In other words, we can move to a different location, we can rotate – and this phenomenon will not change. In geometric terms, this phenomenon is symmetric with respect to shifts and rotations.

Such situations are ubiquitous. In all these situations, we have some transformations that do not change the corresponding physical process. In physics, such transformations are known as *symmetries*. In ordinary speech, we most apply this word to invariance under geometric transformations such as rotations or reflections in a plane, but in physics, this term is applied to any transformations, not necessarily geometric. For example, electromagnetic interactions will not change if we replace all positive charges with negative ones and vice versa – in physics terms, this is symmetry.

The idea of symmetry can be (and is) applied to even more complex situations than these. For example, how do we know that Ohm's law will hold when students do experiments in our lab? Because this law was followed on many different locations on Earth – and even on many locations outside the Earth surface.

Symmetries are one of the main tools in modern physics. Physicists fully understood the importance of symmetries around the 1960s, and this understanding revolutionized physics; see, e.g., [3, 15]. Even the way physical theories are proposed has changed. In the past, starting from Newton, new physical theories were proposed in terms of differential equations. Nowadays, starting with quarks, new theories are described by listing the corresponding symmetries – and differential equations can be then derived from these symmetries. This is not only about new theories: it turned out that most fundamental physical theories that were proposed before this symmetries revolution – such as Maxwell's electrodynamics, quantum physics, special relativity, general relativity, etc. – can be reformulated in terms of the corresponding symmetries, symmetries that uniquely determine the corresponding differential equations; see, e.g., [4].

Symmetries have been useful in computing as well. The same symmetries ideas have been applied to many application areas beyond physics. For example, it turned out that many empirically successful selection of intelligent techniques – be it selection of membershipo functions and "and"- and "or"-operations in fuzzy logic (see, e.g., [2, 6, 10, 13, 14, 17]), selecting activation functions in neural networks, and appropriate functions in genetic algorithms and simulated annealing – all this can be naturally explained by the corresponding symmetries; see, e.g., [12] and references therein. As for modern deep neural networks with their spectacular successes – most of their empirically successes features can also be explained by symmetries; see, e.g., [7].

2 Symmetries and optimality

In the previous section, we mentioned that symmetries can help explain the empirically successful algorithmic (and other) choices. Let us explain how exactly the symmetries approach can help find optimal solutions to practical problems. To explain this, we first need to describe, in precise terms, what we mean by transformations and what we mean by optimality.

What we mean by symmetries. If transformations g and g' both preserve the corresponding phenomenon, then if we apply them one after another – i.e., if we form what is known as a *composition* of two transformation, then this composition also preserves the phenomenon. For example, if we first go 10 meters forward – which preserves the above-described pen-dropping behavior – and then go 10 meters to the right – which also preserves this behavior – then the resulting diagonal shift also preserves this behavior. Similarly, if we perform the inverse transformation g^{-1} – e.g., go back 10 meters – the behavior with not change.

Thus, the set of phenomenon-preserving transformations must be closed under composition and under taking the inverse. Such sets of transformations are known as *transformation groups.*

What we mean by optimal. There may be different optimality criteria, but they are all based on the possibility to compare two alternatives a and b:

- Sometimes, we can conclude that a is better; we will denote it by $a > b$.
- Sometimes, we can conclude that b is better, i.e., that $b > a$.
- Sometimes, we can conclude that a and b are, from the decision maker's viewpoint, of equal value: we will denote it by $a \sim b$.
- And sometimes, we cannot meaningfully compare the two alternatives.

In other words, to describe what is optimal, we need to have two binary relations $>$ and $\sim$. Of course, these relations must be consistent: e.g., if a is better than b and b is better than c, then a should be better than c. In these terms, an alternative a is optimal if it is either better than or of the same quality as all other alternatives b, i.e., if for every b, we have either $a > b$ or $a \sim b$.

What if there are several different optimal alternatives which are equally good with respect to the given optimality criterion? This means that we can now use some other criterion to optimize something else. For example, if we are looking for the most healthy diet, and we came up with several equally healthy diets, then a natural idea is to select the one which is the tastiest. If we have several algorithms that lead to the same optimal accuracy, we can select the one which is, e.g., the fastest.

In general, if an optimality criterion does not lead to a unique optimal alternative, this means that this criterion is not final: we can use this non-uniqueness to optimize something else. So, it makes sense to consider *final* optimality criteria, for which there is exactly one optimal alternative.

The optimality criterion should be invariant. If there is a transformation that preserves the corresponding phenomenon, then which alternative is better should not change if we apply this transformation. For example, since the laws of electrodynamics do not change under spatial shift, we expect that the relative quality of two electric devices will remain the same if we simply move both devices the same distance in the same direction.

Definitions and the result. Now we are ready to formulate all this in precise terms, and to finally explain how symmetries lead to optimality.

Definition 1. *Let A be a set. We will call its elements* alternatives*. By an* optimality criterion *on the set A, we mean a pair of relations $(<, \sim)$ for which the following conditions are satisfied for all a, b, and c:*

- *if $a > b$ and $b > c$, then $a > c$;*
- *if $a > b$ and $b \sim c$, then $a > c$;*
- *if $a \sim b$ and $b > c$, then $a > c$;*
- *if $a \sim b$ and $b \sim c$, then $a \sim c$;*
- *if $a \sim b$, then $b \sim a$;*
- *we always have $a \sim a$;*
- *we cannot have $a > b$ and $a \sim b$.*

We say that an alternative $a_{\rm opt}$ is optimal *if for every $a \in A$, we have either $a_{\rm opt} > a$ or $a_{\rm opt} \sim a$. We say that the optimality criterion is* final *if for this criterion, there is exactly one optimal alternative.*

Definition 2. *Let G be a transformation group. We say that the optimality criterion is G-*invariant *if for all alternatives $a, b \in A$ and for all transformations $g \in G$, we following two conditions are satisfied:*

- *if $a > b$, then $g(a) > g(b)$; and*
- *if $a \sim b$, then $g(a) \sim g(b)$.*

Proposition. *If an alternative $a_{\rm opt}$ is optimal with respect to some final G-invariant optimality criterion, then this alternative is also G-invariant, i.e., $g(a_{\rm opt}) = a_{\rm opt}$ for all $g \in G$.*

Comment. Thus, to find the optimal alternative, we only need to consider G-invariant (*symmetric*) alternatives – and if there is only one such invariant alternative, then this is exactly the optimal one.

Proof of the Proposition. Let us show that for every g, we indeed have $g(a_{\rm opt}) = a_{\rm opt}$.

Indeed, by definition of optimality, for every alternative a, we have either $a_{\rm opt} > a$ or $a_{\rm opt} \sim a$. In particular, this is true for each alternative $g^{-1}(a)$: for every a, we have either $a_{\rm opt} > g^{-1}(a)$ or $a_{\rm opt} \sim g^{-1}(a)$. Since the optimality criterion is G-invariant, we can conclude that for every a, we have either $g(a_{\rm opt}) > g(g^{-1}(a)) = a$ or $g(a_{\rm opt}) \sim g(g^{-1}(a)) = a$. Thus, for every a, have either $g(a_{\rm opt}) > a$ or $g(a_{\rm opt}) \sim a$. By definition of the optimal alternative, this means that the alternative $g(a_{\rm opt})$ is optimal. But since our optimality criterion is final, this means that it has only optimal alternatives, thus indeed $g(a_{\rm opt}) = a_{\rm opt}$.

The statement is proven.

What we do in this paper. In this paper, we apply the symmetry approach to intuitionistic and interval-values fuzzy logics. So, let us briefly recall these two ideas.

3 Intuitionistic and interval-values fuzzy logics: a brief reminder

Traditional fuzzy logic: a brief reminder. In the early 1960s, Lotfi Zadeh, who was one of the leading specialists in control, noticed that in many cases, skilled human controllers control systems – like chemical plants – more effectively than the supposedly optimal automatic controllers. Of course, one cannot be better than the optimal, so this simply meant that the supposedly optimal controllers were not really optimal. To be more precise, they were optimal with respect to the model used in their design – but this model was not always an adequate description of the actual physical systems. In other words, skilled controllers have some additional knowledge that was not incorporated into the existing models.

Skilled controllers were willing to share this knowledge – but unfortunately they could only describe it by using imprecise ("fuzzy") words from natural language such as "small". To describe this knowledge, Zadeh designed special techniques – that he called *fuzzy*. Specifically, to describe a word like "small", he suggested to ask the experts to mark, for each possible value x of the corresponding quantity, the degree $m(x)$ – on the scale from 0 to 1 – to which the expert believes that this value satisfies the given property (e.g., to which x is small). This technique has indeed led to many practical successes; see, e.g., [2, 6, 10, 13, 14, 17].

In practice, we also use logical combinations of statements – i.e., complex statement obtained by applying logical connectives like "and", "or", and "not". To estimate our degree of confidence in such statements, we need to extend the usual logical operations from the set $\{0,1\} = \{\text{"false"}, \text{"true"}\}$ to the whole interval $[0,1]$. Such extensions are known as, correspondingly, "and"-operations $f_\&(a,b)$ (also called *t-norms*), "or"-operations $f_\vee(a,b)$ (also called *t-conorms*), and negation operations $f_\neg(a)$. The simplest such operations are $f_\&(a,b) = \min(a,b)$, $f_\vee(a,b) = \max(a,b)$, and $f_\neg(a) = 1 - a$.

Limitation of the traditional fuzzy techniques. One of the limitations of the traditional fuzzy technique is that it cannot distinguish between two completely different situation:

- when we know nothing about a statement, and
- when we have a large number of arguments for the statement and an equal number of arguments against it.

In both cases, we have equal number of arguments for and against, so it makes sense to describe each of these situations by a number which is equidistant from both 1 ("definitely yes") and 0 ("definitely no"), i.e., by the number 0.5.

To overcome this limitation, two different techniques were developed: intuitionistic fuzzy logic and interval-valued fuzzy logic.

Intuitionistic fuzzy logic. In the above situation, a natural idea is to explicitly take into account how many arguments we have for and again this statement:

- in the first case, we have no arguments for each side, so we can say that $m_+(x) = m_-(x) = 0$, while
- in the second case, we have many arguments for both side, so we can take $m_+(x) = m_-(x) = 0.5$.

In general, instead of a single value $m(x)$, we have two values:

- the value $m_+(x)$ describes to what extend the expert is confident that the given statement is true, and
- the value $m_-(x)$ describes to what extend the expert is confident that the given statement is false.

This idea – first proposed by Krassimir Atanassov – is known as *intuitionistic fuzzy* approach; see, e.g., [1, 16]. Its name came from intuitonistic logic, in which some statements are neither true nor false – this makes sense, e.g., in constructive interpretation, where $\forall x A(x)$ would mean that we have an algorithm that proves $A(x)$ for each x, and its negation would mean that we can algorithmically construct an object x for which this property does not hold – and such algorithms do not always exist.

Logical operations can be naturally extended to this case:

$$f_\&((a_+, a_-), (a_+, a_-)) = (\min(a_+, b_+), \max(a_-, b_-)),$$

$$f_\vee((a_+, a_-), (a_+, a_-)) = (\max(a_+, b_+), \min(a_-, b_-)), \text{ and}$$

$$f_\neg((a_+, a_-)) = (a_-, a_+).$$

An alternative approach: interval-valued fuzzy logic. Another approach to overcoming the above-described limitation of the traditional fuzzy logic is to take into account that in the above two situations, we may eventually end up with different degrees:

- in the first case, after we gain more information about the statement, we may end up in any degree of confidence: from 0 when all future evidence will be against the statement, to 1, when all future evidence will be for this statement; so, the set of all possible values of the future expert's degree of confidence is the whole interval $[0, 1]$;
- in the second case, only the value 0.5 is possible – which corresponds to the degenerate interval $[0.5, 0.5]$.

In general, we can have all possible intervals $[\underline{a}, \overline{a}] \subseteq [0, 1]$. This idea was proposed by Zadeh himself and then actively developed by Jerry Mendel and his group; see, e.g., [10].

Logical operations can be naturally extended to such case: for each operation $f(a, b)$, we can take (see, e.g., [5, 8, 9, 11]):

$$f([\underline{a}, \overline{a}], [\underline{b}, \overline{b}]) \stackrel{\text{def}}{=} \{f(a, b) : a \in [\underline{a}, \overline{a}], b \in [\underline{b}, \overline{b}]\}.$$

As a result, we get the following operations:

$$f_\&([\underline{a},\overline{a}],[\underline{b},\overline{b}]) = [\min(\underline{a},\underline{b}),\min(\overline{a},\overline{b})];$$

$$f_\vee([\underline{a},\overline{a}],[\underline{b},\overline{b}]) = [\max(\underline{a},\underline{b}),\max(\overline{a},\overline{b})];$$

$$f_\neg([\underline{a},\overline{a}]) = [1-\overline{a},1-\underline{a}].$$

4 Relation between intuitionistic and interval-valued fuzzy techniques and how symmetries can explain this relation

Interestingly, from the purely mathematical viewpoint, these two approaches are equivalent. Let us start with the well-known fact: that while the ideas behind these two approaches are different, from the purely mathematical viewpoint, these two approaches are equivalent – in the sense that we can transform them into one another while preserving all three logical operations. This can be attained when we:

- assign, to each intuitionistic degree (a_+,a_-), the interval $[a_+,1-a_-]$ and, vice versa,
- assign, to each interval $[\underline{a},\overline{a}]$, the intuitionistic degree $(\underline{a},1-\overline{a})$.

However, from the computational viewpoint, the intuitionistic approach is more efficient. Indeed, in the intuitionsitic approach, all we need is min and max, while in the interval-valued approach, we also need the subtraction $1-a$.

To perform a substraction of an n-bit number a, we need to flip all n bit – so, we need n bit operations, which for 64-bit words means 64 operations. In contrast, to compute min or max, we only need to decide which of the two numbers to use – i.e., to decide which of the two numbers are larger. In half of the cases, the first (largest) bits are different, so we get our answer already after 1 bit comparison. In situations in which their first bits are the same, in half of these cases – i.e., in 1/4 of overall cases – the second bits are different, so we get the answer after 2 bit operations, etc. In general, with probability 2^{-k}, we need k bit operations, so the overall expected number of bit operations is equal to

$$s = 2^{-1}\cdot 1 + 2^{-2}\cdot 2 + \ldots + 2^{-k}\cdot k + \ldots$$

To compute this sum, let us divide all the terms by 2, then we get

$$\frac{s}{2} = 2^{-2}\cdot 1 + 2^{-3}\cdot 2 + \ldots + 2^{-k}\cdot (k-1) + 2^{-(k+1)}\cdot k + \ldots$$

If we subtract the second sum from the first one, for each k, we get the term

$$2^{-k}\cdot k - 2^{-k}\cdot (k-1) = 2^{-k},$$

so the overall difference takes the following form:

$$\frac{s}{2} = 2^{-1} + 2^{-2} + \ldots + 2^{-k} + \ldots$$

The right-hand side is a well-known geometric progression whose sum is 1. So, $s/2 = 1$ and $s = 2$.

Thus, each min or max operation requires, on average, 2 bit operations – which is 32 times fewer than 64 bit operations needed to compute $1 - a$, Thus, intuionistic fuzzy logic is indeed computationally much more efficient than its interval-valued form.

How can we explain it in terms of symmetries? When we do not known whether the statement S is true or its negation $\neg S$ is true, we can view both S or $\neg S$ as the main statement. In other words, we have a natural symmetry: we can swap a statement and its negation.

Intuitionistic fuzzy logic is invariant with respect to this swap: all we have to do to perform this operation is to swap the two degrees, i.e., to go from (a_+, a_-) to (a_-, a_+): no arithmetic operations are needed. In contrast, in the interval-valued approach, a similar swap requires two negations.

Our general result says that for any invariant optimality criterion, the optimal alternative is also invariant. Clearly, comparing computational complexity – i.e., average running time – of the logical operations should not depend on whether we consider the statement or its negation as basic. So, this criterion is invariant with respect to the swap – and the only swap-invariant representation is the intuitionstic fuzzy one.

So, symmetries indeed explain why the intuionistic fuzzy techniques are computationally more efficient than their interval-values reformulation.

Acknowledgments

This work was supported in part by the National Science Foundation grants 1623190 (A Model of Change for Preparing a New Generation for Professional Practice in Computer Science), HRD-1834620 and HRD-2034030 (CAHSI Includes), EAR-2225395 (Center for Collective Impact in Earthquake Science C-CIES), and by the AT&T Fellowship in Information Technology.

It was also supported by a grant from the Hungarian National Research, Development and Innovation Office (NRDI), by the Institute for Risk and Reliability, Leibniz Universitaet Hannover, Germany, and by the European Union under the project ROBOPROX (No. CZ.02.01.01/00/22 008/0004590).

References

[1] K. Atanassov, *Intuitionistic Fuzzy Sets: Theory and Applications*, Springer-Verlag, Berlin, Heidelberg, 1999.

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

[3] R. Feynman, R. Leighton, and M. Sands, *The Feynman Lectures on Physics*, Addison Wesley, Boston, Massachusetts, 2005.

[4] A. M. Finkelstein, V. Kreinovich, and R. R. Zapatrin. "Fundamental physical equations uniquely determined by their symmetry groups," *Lecture Notes in Mathematics*, Springer- Verlag, Berlin-Heidelberg-N.Y., Vol. 1214, 1986, pp. 159–170.

[5] L. Jaulin, M. Kiefer, O. Didrit, and E. Walter, *Applied Interval Analysis, with Examples in Parameter and State Estimation, Robust Control, and Robotics*, Springer, London, 2012.

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

[7] V. Kreinovich and O. Kosheleva, "Optimization under uncertainty explains empirical success of deep learning heuristics", In: P. Pardalos, V. Rasskazova, and M. N. Vrahatis (eds.), *Black Box Optimization, Machine Learning and No-Free Lunch Theorems*, Springer, Cham, Switzerland, 2021, pp. 195–220.

[8] B. J. Kubica, *Interval Methods for Solving Nonlinear Constraint Satisfaction, Optimization, and Similar Problems: from Inequalities Systems to Game Solutions*, Springer, Cham, Switzerland, 2019.

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

[10] J. M. Mendel, *Explainable Uncertain Rule-Based Fuzzy Systems*, Springer, Cham, Switzerland, 2024.

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

[12] H. T. Nguyen and V. Kreinovich, *Applications of Continuous Mathematics to Computer Science*, Kluwer, Dordrecht, 1997.

[13] H. T. Nguyen, C. L. Walker, and E. A. Walker, *A First Course in Fuzzy Logic*, Chapman and Hall/CRC, Boca Raton, Florida, 2019.

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

[15] K. S. Thorne and R. D. Blandford, *Modern Classical Physics: Optics, Fluids, Plasmas, Elasticity, Relativity, and Statistical Physics*, Princeton University Press, Princeton, New Jersey, 2021.

[16] P. M. Vassilev and K. T. Atanassov, *Modifications and Extensions of Intuitionistic Fuzzy Sets*, "Prof. Marin Drinov" Academic Publishing House, Sofia, Bulgaria, 2019.

[17] L. A. Zadeh, "Fuzzy sets", *Information and Control*, 1965, Vol. 8, pp. 338–353.