Why Triangular and Trapezoid Membership Functions: A Simple Explanation

Vladik Kreinovich, Olga Kosheleva, and Shahnaz Shabazova

Abstract In principle, in applications of fuzzy techniques, we can have different complex membership functions. In many practical applications, however, it turns out that to get a good quality result – e.g., a good quality control – it is sufficient to consider simple triangular and trapezoid membership functions. There exist explanations for this empirical phenomenon, but the existing explanations are rather mathematically sophisticated and are, thus, not very intuitively clear. In this paper, we provide a simple – and thus, more intuitive – explanation for the ubiquity of triangular and trapezoid membership functions.

1 Ubiquity of Triangular and Trapezoid Membership Functions

Why fuzzy sets and membership functions: reminder. In the traditional 2-valued logic, every property is either true or false. Thus, if we want to formally describe a property like "small" in the traditional logic, then every value will be either small or not small.

This may sound reasonable until one realizes that, as a result, we have a threshold value t separating small values from non-small one:

- every value below t is small, while
- every value above t is not small.

This means that, for any small $\varepsilon > 0$ – e.g., for $\varepsilon = 10^{-10}$:

- the value $t - \varepsilon$ is small while

Vladik Kreinovich and Olga Kosheleva
University of Texas at El Paso, 500 W. University, El Paso, TX 79968, USA
e-mail: vladik@utep.edu, olgak@utep.edu

Shahnaz Shabazova
Azerbaijan Technical University, Baku, Azerbaijan
e-mail: shahbazova@gmail.com

- a practically indistinguishable value $t+\varepsilon$ is *not* small.

This does not seem reasonable at all.

To make a formalization of properties like "small" more reasonable, Lotfi Zadeh proposed:

- instead of deciding which value is small and which is not,
- to assign, to each possible value x of the corresponding quantity, a degree $\mu(x)$ to which, according to the expert, this value is small (or, more generally, to what extend this value satisfies the corresponding property); see, e.g., [1, 2, 5, 7, 8, 9].

This degree can be determined by asking an expert to mark this degree on a scale from 0 to 1. Alternatively, we can ask the expert to mark this degree, e.g., on a scale from 0 to 10, and then divide the resulting marked degree by 10. Thus, we get the values $\mu(x) \in [0,1]$ corresponding to different values x. The function $\mu(x)$ assigning this degree to each possible value is known as a *membership function*, or, alternatively, a *fuzzy set*.

Intuitively, we expect that if two values x and x' are close, then the expert will assign similar degrees to these two values, i.e., that the degrees $\mu(x)$ and $\mu(x')$ will also be close.

Triangular and trapezoid membership functions. In the beginning, practitioners applying fuzzy techniques dutifully followed the definition of a membership function. Namely, for each imprecise property like "small", they asked experts, for many different values of x, what are their degrees $\mu(x)$.

Surprisingly, it turned out that from the application viewpoint, all this activity was wasted: whether we talk about control or planning or whatever other activity, the quality of the result usually does not change if we replace the elicited membership function with a simple piecewise-linear one that has a shape of a triangle or a shape of the trapezoid. Specifically, a triangular membership function has the following form, for some parameters $\widetilde{x}$ and $\Delta > 0$:

- $\mu(x) = 0$ for $x \le \widetilde{x} - \Delta$;
- $\mu(x) = \dfrac{x-(\widetilde{x}-\Delta)}{\Delta}$ for $\widetilde{x}-\Delta \le x \le \widetilde{x}$;
- $\mu(x) = \dfrac{(\widetilde{x}+\Delta)-x}{\Delta}$ for $\widetilde{x} \le x \le \widetilde{x}+\Delta$, and
- $\mu(x) = 0$ for $x \ge \widetilde{x}+\Delta$.

Similarly, a trapezoid membership function has the following form, for some parameters $\widetilde{x}$, δ, and Δ, for which $0 < \delta < \Delta$:

- $\mu(x) = 0$ for $x \le \widetilde{x} - \Delta$;
- $\mu(x) = \dfrac{x-(\widetilde{x}-\Delta)}{\Delta-\delta}$ for $\widetilde{x}-\Delta \le x \le \widetilde{x}-\delta$;
- $\mu(x) = 1$ when $\widetilde{x}-\delta \le x \le \widetilde{x}+\delta$;
- $\mu(x) = \dfrac{(\widetilde{x}+\Delta)-x}{\Delta-\delta}$ for $\widetilde{x}+\delta \le x \le \widetilde{x}+\Delta$, and
- $\mu(x) = 0$ for $x \ge \widetilde{x}+\Delta$.

Why triangular and trapezoid membership functions work so well? To most people, the surprisingly empirical success of triangular and trapezoid membership functions was very unexpected. The only person who was not very surprised was ... Lotfi A. Zadeh himself, since he always had an intuition that in many cases, the simplest – and thus, intuitively clearest methods – work the best:

- when we use simple, intuitively clear methods, we utilize both the formulas and our intuition, while
- when we use complex, difficult-to-intuitively understand methods, we have to rely only on formulas, we cannot use our intuition – and thus, our results are often worse.

On the *qualitative* level, this is a reasonable explanation. However, it is desirable to also have a more convincing, *quantitative* explanation of why triangular and trapezoid membership functions work so well.

There already are explanations for this empirical phenomenon, but they are not very intuitive. In our previous papers [3, 4], we have provided quantitative explanations for the surprising empirical success of triangular and trapezoid membership functions. These explanations are based either on the general ideas of signal processing (and related wavelets) or on the type-2 fuzzy analysis of the problem.

From the mathematical viewpoint, both explanations seem to be reasonable. But, honestly, would Lotfi Zadeh – if he was still alive – be fully happy with these explanations? We do not think so. He would complain that these explanations are too complex and thus, not very intuitive. Would it be possible – he would ask (as he asked in many similar situations) – to come up with simpler, more intuitive explanation, an explanation where we would be able to support the corresponding mathematics by the intuitive commonsense understanding?

What we do in this paper. In this paper, we provide such a simple reasonably intuitive explanation.

Is this a final answer? Probably not. Maybe an even simpler and an even more intuitive explanation is possible. However, the new explanation – motivated by Zadeh's quest for simplicity – is already much simpler that the explanations that we had before, so we decided to submit it for publication.

2 Main Idea

As we have mentioned earlier, one of the main motivations for fuzzy technique was the need to make sure that if the values x and x' are close, then the corresponding membership degrees $\mu(x)$ and $\mu(x')$ should also be close. How can we formalize this idea?

When x and x' are close, i.e., when $x' = x + \Delta x$ for some small Δx, then the difference $\mu(x') - \mu(x) = \mu(x + \Delta x) - \mu(x)$ between the corresponding values of the membership function can be – at least for smooth membership functions – represented as $\mu'(x)\cdot\Delta x + o(\Delta x)$. Therefore, the requirement that this difference is small is equivalent to requiring that the absolute value of the derivative $|\mu'(x)|$ is small.

There are two ways to formalize this requirement:

- we can require that the worst-case value of this derivative is small, or
- we can require that the average – e.g., mean squared – value of this derivative is small.

The smaller the corresponding characteristic, the more the resulting membership function is in line with the original fuzzy idea. Thus, it is reasonable, for both formalizations, to select a membership function for which the value of the corresponding characteristic is the smallest possible.

Let us show that in both cases, this idea leads to triangular and trapezoid membership functions.

3 First Formalization and the First Set of Results

Definition 1. *Let $\underline{x} < \overline{x}$ be two real numbers. For each continuous almost everywhere differentiable function $\mu(x)$ defined on the interval $[\underline{x},\overline{x}]$, let us define its* worst-case non-fuzziness degree $D_w(\mu)$ *as*

$$D_w(\mu) = \max_{x\in[\underline{x},\overline{x}]} |\mu'(x)|.$$

Comment. The requirement that the membership function is almost everywhere differentiable is needed so that we can define the largest value of the derivative. This requirement is not as restrictive as it may seem, since usually, membership functions are piecewise-monotonic, and it is known that all monotonic functions – and thus, all piecewise-monotonic functions – are almost everywhere differentiable.

Proposition 1. *Among all continuous almost everywhere differentiable function $\mu(x)$ defined on the interval $[\underline{x},\overline{x}]$ for which $\mu(\underline{x}) = 0$ and $\mu(\overline{x}) = 1$, the following linear function has the smallest worst-case non-fuzziness degree:*

$$\mu(x) = \frac{x-\underline{x}}{\overline{x}-\underline{x}}.$$

Comment. For reader's convenience, all the proofs are placed in a special Proofs section.

Proposition 2. *Among all continuous almost everywhere differentiable function $\mu(x)$ defined on the interval $[\underline{x},\overline{x}]$ for which $\mu(\underline{x}) = 1$ and $\mu(\overline{x}) = 0$, the following linear function has the smallest worst-case non-fuzziness degree:*

$$\mu(x) = \frac{\overline{x} - x}{\overline{x} - \underline{x}}.$$

Discussion. Thus, if we assume that $\mu(x) = 0$ for all $x \notin [\widetilde{x} - \Delta, \widetilde{x} + \Delta]$ and $\mu(\widetilde{x}) = 1$, then, due to Propositions 1 and 2, the most fuzzy membership function – i.e., the function with the smallest possible worst-case non-fuzziness degree – will be the corresponding triangular function.

Similarly, if we assume that $\mu(x) = 0$ for all $x \notin [\widetilde{x} - \Delta, \widetilde{x} + \Delta]$ and $\mu(x) = 1$ for all $x \in [\widetilde{x} - \delta, \widetilde{x} + \delta]$, then, due to Propositions 1 and 2, the most fuzzy membership function – i.e., the function with the smallest possible worst-case non-fuzziness degree – will be the corresponding trapezoid function.

Thus, we indeed get a reasonably simple explanation for the ubiquity of triangular and trapezoid membership functions.

4 Second Formalization and the Second Set of Results

Definition 3. *Let $\underline{x} < \overline{x}$ be two real numbers. For each continuous almost everywhere differentiable function $\mu(x)$ defined on the interval $[\underline{x}, \overline{x}]$, let us define its* average non-fuzziness degree $D_a(\mu)$ *as*

$$D_a(\mu) = \sqrt{\frac{1}{\overline{x} - \underline{x}} \cdot \int_{\underline{x}}^{\overline{x}} (\mu'(x))^2\,dx}.$$

Proposition 3. *Among all continuous almost everywhere differentiable function $\mu(x)$ defined on the interval $[\underline{x}, \overline{x}]$ for which $\mu(\underline{x}) = 0$ and $\mu(\overline{x}) = 1$, the following linear function has the smallest average non-fuzziness degree:*

$$\mu(x) = \frac{x - \underline{x}}{\overline{x} - \underline{x}}.$$

Proposition 4. *Among all continuous almost everywhere differentiable function $\mu(x)$ defined on the interval $[\underline{x}, \overline{x}]$ for which $\mu(\underline{x}) = 1$ and $\mu(\overline{x}) = 0$, the following linear function has the smallest average non-fuzziness degree:*

$$\mu(x) = \frac{\overline{x} - x}{\overline{x} - \underline{x}}.$$

Discussion. Thus, similarly to the previous section, we can show that the most fuzzy membership functions are triangular and trapezoid ones.

5 Proofs

Proof of Proposition 1. For the linear function $\mu_\ell(x$, we have $\mu_\ell'(x) = K \stackrel{\text{def}}{=} \dfrac{1}{\overline{x} - \underline{x}}$ for all x and thus, $D_w(\mu_\ell) = K$. Let us prove:

- that we cannot have a smaller value of $D_w(\mu)$, and
- that the only function with this value of worst-case degree of non-fuzziness is the linear function.

In other words, let us prove:

- that we cannot have $|\mu'(x)| < K$ for all x, and
- moreover, that we cannot have $|\mu'(x)| \le K$ for all x and $\mu'(x) < K$ for some x.

Indeed, due to the known formula relating integration and differentiation, we have

$$\mu(\overline{x}) - \mu(\underline{x}) = 1 - 0 = \int_{\underline{x}}^{\overline{x}} \mu'(x)\,dx.$$

If we had $|\mu'(x)| \le K$ (hence $\mu'(x) \le K$) for all x and $\mu'(x) < K$ for some x, then we would have

$$1 = \int_{\underline{x}}^{\overline{x}} \mu'(x)\,dx < \int_{\underline{x}}^{\overline{x}} K\,dx = (\overline{x} - \underline{x}) \cdot K = 1,$$

i.e., we would get $1 < 1$, which is a clear contradiction. The proposition is proven.

Proof of Proposition 2 is similar.

Proof of Proposition 3. It is known that for every two numbers a and b, we have

$$\frac{a+b}{2} \le \sqrt{\frac{a^2+b^2}{2}}.$$

Indeed, by squaring both sides, we get an equivalent inequality

$$\frac{a^2+2ab+b^2}{4} \le \frac{a^2+b^2}{2}.$$

Multiplying both sides of this inequality by 4 and moving all the terms into the right-hand side, we get an equivalent inequality $0 \le a^2 - 2ab + b^2 + (a-b)^2$, which is, of course, always true. This argument also shows that the equality is attained only if $a = b$.

Similarly, we can prove that always the arithmetic average of all the values $\mu'(x)$ cannot exceed the mean square average of these values:

$$\frac{1}{\overline{x} - \underline{x}} \cdot \int_{\underline{x}}^{\overline{x}} \mu'(x)\,dx \le \sqrt{\frac{1}{\overline{x} - \underline{x}} \cdot \int (\mu'(x))^2\,dx} = D_a(\mu),$$

and that the equality is only possible when all the values $\mu'(x)$ are equal to each other, i.e., when $\mu(x)$ is a linear function. As we have mentioned in the proof of Proposition 1, the left-hand side of the above inequality is equal to

$$\frac{1}{\overline{x}-\underline{x}} \cdot \int_{\underline{x}}^{\overline{x}} \mu'(x)\,dx = \frac{1}{\overline{x}-\underline{x}} \cdot (\mu(\overline{x}) - \mu(\underline{x})) = \frac{1}{\overline{x}-\underline{x}}.$$

Thus:

- this value $\dfrac{1}{\overline{x}-\underline{x}}$ is indeed the smallest possible value of the average non-fuzziness degree $D_a(\mu)$, and
- the only membership function with this smallest value is indeed the linear function.

The proposition is proven.

Proof of Proposition 4 is similar.

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 7th World Conference on Soft Computing (Baku, Azerbaijan, May 29–31, 2018) for valuable discussions.

References

1. R. Belohlavek, J. W. Dauben, and G. J. Klir, *Fuzzy Logic and Mathematics: A Historical Perspective*, Oxford University Press, New York, 2017.
2. G. Klir and B. Yuan, *Fuzzy Sets and Fuzzy Logic*, Prentice Hall, Upper Saddle River, New Jersey, 1995.
3. O. Kosheleva and V. Kreinovich, “Why triangular membership functions are often efficient in F-transform applications: relation to probabilistic and interval uncertainty and to haar wavelets”, In: J. Medina et al. (eds.), *Proceedings of the 17th International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems IPMU'2018*, Cadiz, Spain, June 11–15, 2018.
4. O. Kosheleva, V. Kreinovich, and S. Shahbazova, “Type-2 fuzzy analysis explains ubiquity of triangular and trapezoid membership functions”, *Proceedings of the 7th World Conference on Soft Computing*, Baku, Azerbaijan, May 29–31, 2018.
5. J. M. Mendel, *Uncertain Rule-Based Fuzzy Systems: Introduction and New Directions*, Springer, Cham, Switzerland, 2017.
6. J. M. Mendel and D. Wu, *Perceptual Computing: Aiding People in Making Subjective Judgments*, IEEE Press and Wiley, New York, 2010.
7. H. T. Nguyen and E. A. Walker, *A First Course in Fuzzy Logic*, Chapman and Hall/CRC, Boca Raton, Florida, 2006.

8. V. Novák, I. Perfilieva, and J. Močkoř, *Mathematical Principles of Fuzzy Logic*, Kluwer, Boston, Dordrecht, 1999.
9. L. A. Zadeh, "Fuzzy sets", *Information and Control*, 1965, Vol. 8, pp. 338–353.