3-2020

Which Are the Correct Membership Functions? Correct "And"- and "Or"- Operations? Correct Defuzzification Procedure?

Olga Kosheleva

Vladik Kreinovich

Shahnaz Shahbazova

Comments:
Technical Report: UTEP-CS-20-22

Which Are the Correct Membership Functions? Correct “And”- and “Or”- Operations? Correct Defuzzification Procedure?

Olga Kosheleva, Vladik Kreinovich, and Shahnaz Shahbazova

Abstract Even in the 1990s, when many successful examples of fuzzy control appeared all the time, many users were somewhat reluctant to use fuzzy control. One of the main reasons for this reluctance was the perceived subjective character of fuzzy techniques – for the same natural-language rules, different experts may select somewhat different membership functions and thus get somewhat different control/recommendation strategies. In this paper, we promote the idea that this selection does not have to be subjective. We can always select the “correct” membership functions, i.e., functions for which, on previously tested case, we got the best possible control. Similarly, we can select the “correct” and- and or-operations, the correct defuzzification procedure, etc.

1 Formulation of the Problem: Is There Such a Thing as the Correct Membership Function?

Need to translate imprecise (fuzzy) expert rules into a precise control strategy. In many application areas, be it medicine or engineering, some professionals are more successful and more skilled than others. It is therefore desirable to somehow incorporate the knowledge and skills of the top experts into a computer-based system that would help other experts in their decisions.

Top experts are usually willing (and even eager) to share their knowledge. Some part of this knowledge they describe in precise numerical terms, such as “if the body temperature is above 100 F, take aspirin”. The problem is that a large part

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

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

of the experts' knowledge is described not in such precise numerical terms, but in terms of imprecise ("fuzzy") words of natural language like "high" or "small". To incorporate this knowledge inside a computer-based system, it is important to translate such knowledge into a precise form.

This was the main motivation of Lofti Zadeh when he invented fuzzy techniques; see, e.g., [1, 4, 9, 13, 14, 18].

Fuzzy techniques: general idea. To explain the problem with which we deal win this paper, let us briefly refresh the main ideas of fuzzy techniques. We will explain these ideas on the example of the most frequent application of fuzzy technique – fuzzy control. Specifically, we will describe the main ideas behind the simplest (and most widely used) approach – Mamdani's approach.

We consider a system whose state, at any given moment of time, is characterized by n numerical values $x_1,\dots,x_n$. For example:

- For a simple moving point-size object, the first three parameters are coordinates and the second three parameters are three components of the velocity.
- For a car, we have two spatial coordinates, two components of velocity, and parameters that describe the orientation of the wheels, the engine's regime, etc.

We want, based on the values x_i describing the system's state, to recommend an appropriate control. The control may also be described by several parameters. Let us concentrate on recommending one of these parameters; we will denote this parameter by u. In these terms, we want, for each tuple $x=(x_1,\dots,x_n)$, to find an appropriate control value u.

To find this value u, we have several (K) if-then rules provided by the experts, i.e., rules of the type "if $A_{k1}(x_1)$ and ... and $A_{kn}(x_n)$, then $B_k(u)$", where $k=1,\dots,K$ is the number of the rule, and A_{ki} and B_k are imprecise properties like "small".

In general, once we have such rules, then the control value u is appropriate for the state x if:

- either the first rule is applicable – i.e., its conditions $A_{1i}(x_i)$ are all satisfied and its conclusion $B_1(u)$ is also satisfied,
- or the second rule is applicable – i.e., its conditions $A_{2i}(x_i)$ are all satisfied and its conclusion $B_2(u)$ is also satisfied, etc.

If we describe the property "u is a reasonable control for the state x" by $R(x,u)$, then the above statement can be described in the following symbolic form:

$$R(x,u)\Leftrightarrow$$

$$(A_{11}(x_1)\,\&A_{12}(x_2)\,\&\,\dots\,\&B_1(u))\vee(A_{21}(x_1)\,\&A_{22}(x_2)\,\&\,\dots\,\&B_2(u))\vee\dots \quad (1)$$

First step: membership functions. The formula (1) may looks like a normal logical formula, but it is not:

- in a usual logical formula, every statement is either true or false, but
- here, a statement like "98 degrees is a high fever" is only true to some extent.

In the computer, true is represented as 1, and false as 0. Thus, to represent statements which are not exactly true and not exactly fully false, Zadeh proposed to use real numbers intermediate between 0 and 1.

Eliciting such numbers for each statement is a usual think in polls, when we are asked, e.g., to estimate, on a scale from 0 to 10, how satisfied we are with the given service. People usually have no problem assigning some numerical value to such statements. The only thing we need to do to get a value from 0 to 1 is to normalize this number – i.e., in our example, to divide it by 10.

For each of the quantities x_i, we can select several possible values $x_{i1},\ldots,x_{ip}$. Then, for each of these values x_{iv} and for each of the corresponding properties A_{ki}, we can ask the expert to estimate the degree to which this property is satisfied by the given value. This degree is usually denoted by $\mu_{ki}(x_{iv})$.

Of course, there are infinitely many possible real numbers x_i, and it is not possible to ask the expert about each of them. So, to find the degree $\mu_{ki}(x_i)$ for all other possible values x_i, we need ro perform some interpolation or extrapolation. As a result, we get a function $\mu_{ki}(x_i)$ that assigns, to each possible value x_i, the degree to which this value satisfies the property A_{ki}. This function is called a *membership function* corresponding to the property A_{ki}.

Similarly, we design a membership function $\mu_k(u)$ describing to what extent the value u satisfies the property B_k.

Second step: "and"- and "or"-operations. Once we know the expert's degrees of confidence $\mu_{ki}(x_i)$ and $\mu_k(u)$ in statement $A_{ki}(x_i)$ and $B_k(u)$, we need to estimate the degree to which the statement $R(x,u)$ is true. For this purpose, fuzzy technique uses special functions $f_\&(a,b)$ and $f_\vee(a,b)$ known as "and"- and "or"-operations (or, for historical reasons, t-norms and t-conorms):

- the function $f_\&(a,b)$ transforms the expert's degree of confidence a and b in statement A and B into an estimate for the expert's degree of confidence in the statement $A\,\&\,B$; and
- the function $f_\vee(a,b)$ transforms the expert's degree of confidence a and b in statement A and B into an estimate for the expert's degree of confidence in the statement $A\vee B$.

By using these operation, we can find, for each u, the degree $\mu_r(x,u)$ to which the control u is reasonable for the input x as

$$\mu_r(x,u) =$$

$$f_\vee(f_\&(\mu_{11}(x_1),\mu_{12}(x_2),\ldots,\mu_1(u)),f_\&(\mu_{21}(x_1),\mu_{22}(x_2),\ldots,\mu_2(u)),\ldots). \quad (2)$$

Final step: defuzzification. If we are designing a recommender system, then the degrees (2) is all we want.

However, if we are designing a system for automatic control, then we need to transform the membership function (2) into an exact value $\bar{u}$. This process is known as *defuzzification.*

The most frequently used defuzzification is based on minimizing the mean squared difference $(\bar{u}-u)^2$ between the estimate and the actual (unknown) value of the optimal control – so that the term corresponding to each u is weighted by the degree $\mu(x,u)$ to which this value u is reasonable. In precise terms, we minimize the expression

$$\int \mu(x,u)\cdot(u-\bar{u})^2\,du.$$

Differentiating this expression with respect to $\bar{u}$ and equating the resulting derivative to 0, we get the following formula

$$\bar{u}=\frac{\int u\cdot\mu(x,u)\,du}{\int \mu(x,u)\,du}.$$

This formula is known as *centroid defuzzification*.

Problem: this is all very subjective. One of the main problems is that all this sounds very subjective. One expert may draw one membership function, another expert may draw a somewhat different one. Because of this subjectivity, many users are reluctant to use fuzzy techniques.

A natural question may be: which membership function is a "correct" one? At present, the usual answer is: there is no correct, objective membership function, it is all subjective.

Our opinion. What we argue in this paper is that while, yes, the existing membership functions are subjective, there *is* such a thing as a correct membership function – as well the correct "and"- and "or"-operations, etc.

Actually, our arguments will be very straightforward and simple. However, they seem to be novel and, in our opinion, are important, so we decided to write them down and to promote them as much as possible.

2 In Each Application Area, There Are the Correct Membership Functions: Our Arguments

Towards an idea. In the control case – where fuzzy techniques are most frequently used – the main objective of the fuzzy technique is to transform the natural-language-based fuzzy rules into a precise control strategy.

If we use different membership functions, we get different control strategies. For each specific system, some strategies are better, some are somewhat worse.

There is usually an objective function that describes what exactly we want. For example, when planning a route for a self-driving car:

- in some situations, e.g., in emergency, we want to find the fastest way from point A to point B;
- in other cases, we want to find a way that – within a given time limit – will lead to the smallest amount of pollution;

- for sending delicate fragile objects, the objective may be to make sure that the ride is as smooth as possible – in some precise sense.

Thus, there exist membership functions for which this objective function attains the largest possible value – these membership functions are the best of this particular situation and are, thus, the *correct* ones.

So what? At first glance, what we just wrote may sound reasonable, but on second thought, this does not sound convincing at all. The only way to find the correct membership function is to know the exact system, and if we know the exact system, then we do not really need the fuzzy techniques: instead of optimizing membership functions, we can simply optimize the control itself.

Still, there are correct membership functions. If we only had one system to control, then, true, the above idea would make no sense at all. In practice, however, we have many different systems to control – and for each of these systems, we may have many different situations to control, with many different objective functions (as we have mentioned above).

In all these situations, the same users provide rules by using largely the same natural-language words. So, a natural idea is as follows:

We consider all the situations $s_1, \ldots, s_a$ for which the users provided the rules *and* for which the system is actually known. For each of these situations, we know the corresponding objective function. Let us denote by $F_i(\mu)$ the value of the objective function which is obtained in the i-th situation if we use a combination

$$\mu = (\mu_{11}, \mu_{12}, \ldots, \mu_1, \mu_{21}, \mu_{22}, \ldots, \mu_2, \ldots)$$

of the membership functions. To each of the situations, we assign a weight w_i describing this situation's importance. In this sense, the overall objective function has the form

$$F(\mu) = w_1 \cdot F_1(\mu) + w_2 \cdot F_2(\mu) + \ldots + w_a \cdot F_a(\mu).$$

Then, we find the combination of membership functions μ for which this overall objective function attains its largest possible value. The membership functions from this combination are thus the best (most effective), i.e., *correct* membership functions.

How are these correct membership functions useful? Why do we need to know which membership functions are the best for known situations? Because usually, there are many *unknown* situations, i.e., situations:

- in which we do not have exact expressions for the system, and
- for which the only thing we know are expert rules.

These are exactly the types of situations for which fuzzy techniques were invented in the first place. Since the "correct" membership functions worked well on previous situations, it is highly possible that they will work well in the new situation as well.

Finding the correct membership functions will make applications of fuzzy techniques easier. If we use the above idea, then, in each application area, after we have found the correct membership function, there is no longer any need to elicit membership degrees etc., no need for any subjectivity at all – all this will be now done "under the hood", inside the system.

All the expert has to do is supply rules, and the system will automatically translate it into a control strategy – by using the stored "correct" membership functions.

Of course, with new applications, we may need to update what we mean by correct membership functions. Of course, for new systems, we may also gain some information about how they actually work, and thus, be able to estimate possible numerical consequences of different combinations of membership functions.

We will then be able to add this new system to the list of situations based on which we search for the correct membership function, and thus, maybe come up with new, slightly different "correct" membership functions.

Caution. We are *not* yet describing *how* to find the "correct" membership functions, we are just explaining why finding such correct membership functions may be very useful. There are many different optimization techniques, so we should not worry about how to optimize – although, of course, it would be nice to see which optimization techniques are more efficient for finding such "correct" membership functions (see our discussion in the next section).

3 Not Only Membership Functions

Not only membership functions. Why limit ourselves to selecting membership functions? We can similarly select the best "and"- and "or"-operations, the best defuzzification procedure, etc.:

- we find the "and"- and "or"-operations and/or a defuzzification procedure that work the best on known cases, and
- we can then use these optimal ("correct") "and"- and "or"-operations and defuzzification in new situations.

This has been partly done. For a medicine-related recommendation system, this was done by the designers of the world's first expert system MYCIN, an expert system for diagnosing rare blood diseases; see, e.g., [2]. To find the best "and"-operation, the designers of MYCIN selected numerous pairs of related statements (A_i, B_i), $i = 1, 2, \ldots$, and asked the experts to estimate their degree of confidence a_i, b_i, and c_i in the statements A_i, B_i, and $A_i \,\&\, B_i$. Then, they found an "and"-operation $f_\&(a,b)$ that provided the best match for all these cases, i.e., for which $f_\&(a_i,b_i) \approx c_i$ for all i. This can be done, e.g., by the usual Least Squares technique, when we minimize the sum

$$(f_\&(a_1,b_1) - c_1)^2 + (f_\&(a_2,b_2) - c_2)^2 + \ldots$$

This selection – as well as the similar selection of the most efficient “or”-operation – led to a very successful expert system.

Interestingly, at first, the MYCIN designers thought that they found general laws of human reasoning. However, when they tried to apply their “and”- and “or”-operations to a different application domain – geosciences – the results were not good. It turned out that in different application domains, different “and”- and “or”-operations work best. This makes perfect sense; for example:

- in medicine, unless we have a real emergency, it is important to be very cautious, so as not to harm the patient; so, the doctors do not recommend a serious procedure like surgery unless they are reasonably sure that it will be successful – and they perform a lot of tests if they are not sure;
- on the other hand, if an oil company is too cautious, its competitors will be the first to exploit the new possibility, so actions need to be bold; empty wells and resulting losses are OK (and inevitable) as long as overall, the company is profitable.

For control, also, depending on different objective functions – whether we want stability or smoothness – different “and”- and “or”-operations turn out to be better; see, e.g., see, e.g., [6, 7, 8, 10, 11, 12, 13, 16, 17].

An interesting possibility. It may turns out that several different combinations of membership functions, “and”- and “or”-operations, and defuzzification lead to exactly the same control and are, thus, equally good.

This happened in physics, where, as the famous mathematician and physicist Henri Poincaré noticed in the early 20 century, we can describe the same gravity phenomena is different geometries – we just need to modify physical equations:

- we can have Einstein’s curved space, then particles move along inertial lines, or
- we can have the flat Minkowski space-time in which gravity is an extra force; see, e.g., [15] (see also [3, 5]).

If such a situation occurs in fuzzy – that several options lead to equally good control (or equally good recommendations) – then out of several equally good options we can select, e.g., the one which is the easiest to compute.

We can go even further. Why limit ourselves to the usual $[0,1]$-based fuzzy sets? We can alternatively consider interval-valued (or, more generally, type-2) interpretations of natural-language words, in which, for each possible value x of the corresponding quantity, instead of a single degree $\mu(x)$, we have an interval $\left[\underline{\mu}(x),\overline{\mu}(x)\right]$ of possible degree values.

It would be interesting to know which technique leads to the best control and/or to the best recommendations. This will convert current philosophical and case-by-case discussions into a practical problem: as the proverb says, the proof of the pudding is its eating. Maybe in some application areas, type-1 will be better, in others type-2?

We can similarly compare Mamdani’s technique with a sometimes used fuzzy logic technique, in which we use fuzzy implication $f_{\to}(a,b)$ to get a more direct description of the expert’s rules:

$$\mu_r(x,u) = f_\&(f_\rightarrow(f_\&(\mu_{11}(x_1),\mu_{12}(x_2),\ldots),\mu_1(u)),$$

$$f_\rightarrow(f_\&(\mu_{21}(x_1),\mu_{22}(x_2),\ldots),\mu_2(u)),\ldots).$$

It would be interesting to know which one leads to better control or better recommendations.

Similarly, we can compare the fuzzy approach with probabilistic approaches proposed for describing natural-language rules.

Opportunities are endless, let us follow them, there is a lot of work ahead!

Acknowledgments

This work was supported in part by the US National Science Foundation grants 1623190 (A Model of Change for Preparing a New Generation for Professional Practice in Computer Science) and HRD-1242122 (Cyber-ShARE Center of Excellence).

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. B. G. Buchanan and E. H. Shortliffe, *Rule Based Expert Systems: The MYCIN Experiments of the Stanford Heuristic Programming Project*, Addison-Wesley, Reading, Massachusetts, 1984.
3. S. N. Gupta, "Einstein's and other theories of gravitation", *Review of Modern Physics*, 1957, Vol. 29, pp. 334–336.
4. G. Klir and B. Yuan, *Fuzzy Sets and Fuzzy Logic*, Prentice Hall, Upper Saddle River, New Jersey, 1995.
5. V. Kreinovich, "Gupta's derivation of Einstein equations", *Soviet Academy of Sciences Doklady*, 1975, Vol. 222, No. 2, pp. 319–321 (in Russian); translated into English in *Soviet Physics Doklady*, 1975, Vol. 20, No. 5, pp. 341–342.
6. V. Kreinovich, "From semi-heuristic fuzzy techniques to optimal fuzzy methods: mathematical foundations and applications", In: B. Papadopoulos and A. Syropoulos (eds.), *Current Trends and Developments in Fuzzy Logic, Proceedings of the First International Workshop*, Thessaloniki, Greece, October 16–20, 1998, pp. 1–62.
7. V. Kreinovich and S. Kumar, "Optimal choice of &- and $\vee$-operations for expert values," *Proceedings of the 3rd University of New Brunswick Artificial Intelligence Workshop*, Fredericton, New Brunswick, Canada, 1990, pp. 169–178.
8. V. Kreinovich, C. Quintana, R. Lea, O. Fuentes, A. Lokshin, S. Kumar, I. Boricheva, and L. Reznik, "What non-linearity to choose? Mathematical foundations of fuzzy control", *Proceedings of the 1992 International Conference on Fuzzy Systems and Intelligent Control*, Louisville, Kentucky, 1992, pp. 349–412.
9. J. M. Mendel, *Uncertain Rule-Based Fuzzy Systems: Introduction and New Directions*, Springer, Cham, Switzerland, 2017.
10. H. T. Nguyen and V. Kreinovich, "Methodology of fuzzy control: an introduction", In: H. T. Nguyen and M. Sugeno (eds.), *Fuzzy Systems: Modeling and Control*, Kluwer, Boston, Massachusetts, 1998, pp. 19–62.

11. H. T. Nguyen, V. Kreinovich, and D. Tolbert, “On robustness of fuzzy logics”, *Proceedings of the 1993 IEEE International Conference on Fuzzy Systems FUZZ-IEEE’93*, San Francisco, California, March 1993, Vol. 1, pp. 543–547.
12. H. T. Nguyen, V. Kreinovich, and D. Tolbert, “A measure of average sensitivity for fuzzy logics”, *International Journal on Uncertainty, Fuzziness, and Knowledge-Based Systems*, 1994, Vol. 2, No. 4, pp. 361–375.
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. H. Poincaré, *Science and Hypothesis*, Bloomsbury Academic, London, 2018.
16. M. H. Smith and V. Kreinovich, “Optimal strategy of switching reasoning methods in fuzzy control”, Chapter 6 in H. T. Nguyen, M. Sugeno, R. Tong, and R. Yager (eds.), *Theoretical Aspects of Fuzzy Control*, J. Wiley, N.Y., 1995, pp. 117–146.
17. D. Tolbert, *Finding “and” and “or” operations that are least sensitive to change in intelligent control*, University of Texas at El Paso, Department of Computer Science, Master’s Thesis, 1994.
18. L. A. Zadeh, “Fuzzy sets”, *Information and Control*, 1965, Vol. 8, pp. 338–353.