Why Storytelling Is Good for Education? Why Practitioners Stopped Eliciting Membership Functions and “And”- and “Or”-Operations from the Experts?

Christian Servin, Olga Kosheleva, and Vladik Kreinovich

Abstract In this paper, we deal with two questions – which we show to be related: why storytelling is good for education and why practitioners stopped eliciting membership functions and “and”- and “or”-operations from the experts.

1 Formulation of two problems

The first problem. In many disciplines – such as mathematics – we want to teach students rigorous thinking. However, empirical evidence shows that many successful teaching techniques – including techniques for teaching mathematics – use informal storytelling; see, e.g., [3, 4, 10, 11, 12]. A natural question is: why?

The second problem. There are many successful applications of fuzzy techniques (see, e.g., [1, 5, 6, 8, 9, 15]), in particular, many successful applications to control. These applications usually have three stages:

- first, we elicit imprecise (“fuzzy”) rules form the experts,
- then, we use fuzzy techniques to transform these rules into a precise control strategy, and
- finally, if needed, we tune the resulting controller to make it most effective.

Christian Servin
Information Technology Systems Department, El Paso Community College (EPCC)
919 Hunter Dr., El Paso, TX 79915-1908, USA, e-mail: cservin1@epcc.edu

Olga Kosheleva
Department of Teacher Education, University of Texas at El Paso, 500 W. University
El Paso, Texas 79968, USA, e-mail: olgak@utep.edu

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

Interestingly, while the first and the third stages are, in essence, the same now as they were decades ago, in implementing the second stage, there is a drastic different between what was done earlier, and what is done now:

- when fuzzy applications started, designers of the corresponding fuzzy systems spent a lot of times eliciting, from the experts, the exact shapes of the membership functions and, often, the exact "and"- and "or"-operations that best describe the expert's thinking;
- in contrast, at present, this is practically never done: the designers use simple – e.g., triangular – membership functions and simple – min or product – "and"-operations (t-norms).

This change saves the designers a lot of effort. For example, the designers of the world's first effective medical expert system MYCIN spent several years – and several million dollars – finding out which "and"-operation best describes the reasoning of medical doctors; see, e.g., [2].

This does not mean that the designers use only simple membership functions and "and"- and "or"-operations in their designs. The resulting systems often use complex membership functions and complex "and"- and "or"-operations – but this adjustment of membership functions and "and"- and "or"-operations happens only on the third – tuning – stage. This is no longer done by first eliciting this information from the users.

So, now we elicit much less information from the experts than in the past – and still the results are as good. Intuitively, the more information we get form the experts, the better our system should be – but it looks like in this particular case, the corresponding additional information does not improve the quality of the resulting system – since its lack does not make the systems perform any worse. But why? How can we explain this?

What we do in this paper. In this paper, we provide answers to both questions – and we explain the relation between these two answers. Specifically, in Section 2, we provide the answer to the second question, then in Section 3 we provide the answer to the first question, and explain how these two answers are related.

2 Why practitioners stopped eliciting membership functions and "and"- and "or"-operations

Let us first come up with a framework for analyzing this problem. Both membership functions and "and"- and "or"-operations deal with fuzzy degrees, degrees that describe the expert's degree of confidence in different statements. According to the well-known seven-plus-minus-two law (see, e.g., [7, 13]), a person can meaningful distinguish between 5 and 9 different degrees – and for most people, it is 7 degrees. Thus, while, from the purely mathematical viewpoint, fuzzy degrees can take infinitely many values – namely, all possible real values from the interval $[0,1]$ – in reality, most experts can only meaningfully distinguish 7 different degrees. One of these degrees is "absolutely true" – that corresponds to 1, another is "absolutely false", which corresponds to 0. This leave us with 5 degrees corresponding to uncertainty – and thus, located inside the open interval $(0,1)$.

To describe the most appropriate numerical values corresponding to these 5 degrees, we need to select 5 points in this open interval. Since we have no reason to believe that the distance between any pair of neighboring degrees is larger than for any other pair, it makes sense to assume that all these distances are equal, i.e., that each of the 7 degrees (including 0 and 1) is located at the same distance d from the neighboring degrees. Since the first value is 0, the next is thus d, the next after that is $2d$, etc., and the final seventh degree is $6d$. Since we know that the final seventh degree is 1, this means that $6d = 1$, so $d = 1/6$, and the corresponding degrees are 0, 1/6, 1/3, 1/2, 2/3, 5/6, and 1.

To make computations more convenient, let us use the scale $[0,6]$ instead of the usual $[0,1]$. In this scale, the actual degrees are integers: 0, 1, 2, 3, 4, 5, and 6. In this scale, the two simplest "and"-operations – min and product – get the following form:

- minimum and maximum remains minimum and maximum, while
- the product becomes $r/6$ for $r = a \cdot b$.

For integers a and b, the value $v \stackrel{\text{def}}{=} (a \cdot b)/6$ is sometimes not an integer. Since in this scale, only integer degree values make sense, we will round the value into an integer V by using the usual rounding rules:

v	$[0,0.5)$	$[0.5,1.5)$	$[1.5,2.5)$	$[2.5,3.5)$	$[3.5,4.5)$	$[4.5,5.5)$	$[5.5,6]$
V	0	1	2	3	4	5	6

One can easily show that in terms of the integer-valued r, this rounding leads to the following rules:

r	$[0,2]$	$[3,8]$	$[9,14]$	$[15,20]$	$[21,26]$	$[27,33]$	$[33,36]$
V	0	1	2	3	4	5	6

So how will the product look like in this framework? Let us use this framework to see how the product will look like. For example, for $a = 3$ and $b = 5$, the actual product is $r = 15$, so, based on the above rules, we take the degree 3 as the resulting value. By repeating these calculations for all possible pairs (a,b), we get the following table:

	0	1	2	3	4	5	6
0	0	0	0	0	0	0	0
1	0	0	1	1	1	1	1
2	0	1	1	1	2	2	2
3	0	1	1	2	2	3	3
4	0	1	1	2	2	3	4
5	0	1	1	3	3	4	5
6	0	1	2	3	4	5	6

One can see that all these values are either equal to $\min(a,b)$ or differ from $\min(a,b)$ by one – i.e., by the smallest distinguishable difference.

What about other "and"-operations? what about "or"-operations? One of the possible ways to elicit a fuzzy degree is to ask n experts and if m of them say "yes", take the value m/n. In this case, from the purely mathematical viewpoint, the degree can be viewed as a probability – the probability that a randomly selected expect says "yes".

In general, for random events, the probability that both occur is between $\max(a+b-1,0)$ and $\min(a,b)$ – and when the correlation is non-negative, between $a\cdot b$ and $\min(a,b)$; see, e.g., [14]. In fuzzy application, "and" is usually applied to conditions from the same rule provided by the same expert – so it is reasonable to expect these conditions to be non-negatively correlated. Thus, the corresponding "and"-operations are between $a\cdot b$ and min.

Since the results of $a\cdot b$ and $\min(a,b)$ differ by no more than one level, the same is true for any two operations between these two. So, no matter what "and"-operation the expert actually uses, the result is minimally different from min. This is why current practitioners simply use min instead of the time-consuming elicitation.

For any "or"-operation $f_\vee(a,b)$, we can use the same result since, as is well known, its dual $1-f_\vee(1-a,1-b)$ is an "and"-operation – and thus, close to min. Therefore, the original "or"-operation is close to the dual of min, i.e., to max.

What about membership functions? Why do practitioners use piece-wise linear membership functions? Any smooth function can be expanded into Taylor series, in which the next term after linear is quadratic, proportional to x^2. For values $a\in[0,6]$, the original term x^2 corresponds to $a^2/6$ – the diagonal terms in the above table. We can see that these terms are minimally different from a. Since quadratic terms lead to an almost the same result as linear, this explains why practitioners do not consider such terms (and thus, do not consider higher order terms), and only use piecewise linear functions.

All this explains why practitioners nowadays do not elicit membership functions and "and"- and "or"-operations.

3 But why storytelling?

Why storytelling? In precise sciences, learning means moving from a wrong precise statement to a correct precise statement: e.g., from the wrong $3\cdot 3=6$ to the correct $3\cdot 3=9$. This means changing, for each of these statements, our degree of confidence in them from 0 to 1 or from 1 to 0. To many students, such an abrupt change is difficult. To make learning easier, it is desirable to replace a single abrupt change with two or more changes which are smaller than by 1. For each intermediate statement, the degree should thus be inside $(0,1)$.

What does that mean? Degrees inside $(0,1)$ were introduced by Zadeh to describe imprecise ("fuzzy") natural-language statements. So, intermediate statements should be imprecise natural-language statements. This is exactly what storytelling is about – so this explains why storytelling is effective in education.

How is this related to the second problem? In the traditional approach to fuzzy, it is not enough to say that we have degrees between 0 and 1 to be able to apply fuzzy techniques. We also need to select the shapes of the membership functions, to select "and"- and "or"-operations, etc. What we showed is that for the simple purpose of describing expert knowledge, this additional knowledge is, in effect, not needed – different selections lead to almost same (minimally distinguishable) results. From this viewpoint, once we see that it is reasonable to use degrees between 0 and 1, we can already use fuzzy techniques.

For storytelling, this means that we do not need to worry about membership functions or about "and"- and "or"-operations: we can, in effect, use whatever we want, and the results will be good – as indeed they are when we use storytelling in education.

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).

The authors are thankful to the anonymous referees for valuable suggestions.

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. M. F. Dahlstrom, "Using narratives and storytelling to communicate science with nonexpert audiences", *Proceedings of the National Academy of Sciences*, 2014, Vol. 111, Suppl. 4, pp. 13614–13620, doi 10.1073/pnas.1320645111
4. J. Gottschall, *The Storytelling Animal: How Stories Make Us Human*, Houghton Mifflin Harcourt, Boston, Massachusetts, USA, 2012.
5. G. Klir and B. Yuan, *Fuzzy Sets and Fuzzy Logic*, Prentice Hall, Upper Saddle River, New Jersey, 1995.
6. J. M. Mendel, *Explainable Uncertain Rule-Based Fuzzy Systems*, Springer, Cham, Switzerland, 2024.
7. G. A. Miller, "The magical number seven plus or minus two: some limits on our capacity for processing information", *Psychological Review*, 1956, Vol. 63, No. 2, pp. 81–97.
8. H. T. Nguyen, C. L. Walker, and E. A. Walker, *A First Course in Fuzzy Logic*, Chapman and Hall/CRC, Boca Raton, Florida, 2019.

9. V. Novák, I. Perfilieva, and J. Močkoř, *Mathematical Principles of Fuzzy Logic*, Kluwer, Boston, Dordrecht, 1999.
10. R. Olson, *Houston, We Have a Narrative: Why Science Needs Story*, University of Chicago Press, Chicago, Illinois, USA, 2015.
11. S. Papert, *Mindstorms: Children, Computers, and Powerful Ideas*, Basic Books, New York, New York, USA, 1980.
12. W. Quesenbery and K. Brooks, *Storytelling for User Experience: Crafting Stories for Better Design*, Rosenfeld Media, Brooklyn, New York, USA, 2010.
13. S. K. Reed, *Cognition: Theories and Application*, SAGE Publications, Thousand Oaks, California, 2022.
14. D. J. Sheskin, *Handbook of Parametric and Nonparametric Statistical Procedures*, Chapman and Hall/CRC, Boca Raton, Florida, 2011.
15. L. A. Zadeh, "Fuzzy sets", *Information and Control*, 1965, Vol. 8, pp. 338–353.