Why Interval-Valued (and Type-2) Fuzzy Methods Are Often More Effective

Christian Servin[1], Olga Kosheleva[2], and Vladik Kreinovich[3]
[1]Computer Science and Information Technology Systems Department,
El Paso Community College, 919 Hunter, El Paso, Texas 79915, USA,
cservin1@epcc.edu
[2,3]Departments of [2]Teacher Education and [3]Computer Science
University of Texas at El Paso,
500 W. University, El Paso, Texas 79968, USA
olgak@utep.edu, vladik@utep.edu

1. Formulation of the problem

- In this talk, our objective is to explain why:
  - interval-valued and type-2 fuzzy techniques are often more effective
  - than the more traditional (type-1) fuzzy techniques.
- In order to formulate the phenomenon that we are trying to explain, we need to briefly recall:
  - why fuzzy techniques were invented in the first place,
  - why interval-valued and type-2 techniques appeared, and
  - in what exactly sense these techniques are more effective.

2. Why fuzzy in the first place: a brief reminder

- In the 1960s, Lotfi Zadeh was:
  - one of the world's main specialists in automated control and
  - one of the most popular textbook on this topic.
- He noticed that in many practical situations:
  - the control that was optimal (based on the existing models of the corresponding systems)
  - was often less effective than control by expert controllers.
- A natural explanation is that the models on which optimization was based were approximate.
- These models clearly did not fully adequately reflect the expert knowledge.
- And the experts were often indeed able to point out some knowledge that the existing models missed.

3. Why fuzzy in the first place: a brief reminder (cont-d)

- Some of this additional knowledge was formulated in precise mathematical terms.
- It was relatively easy to incorporate it into the models used to find the control.
- The challenge was that:
  - most of this missing knowledge was formulated in terms of imprecise ("fuzzy") words from natural language,
  - like "small" or "about 1".
- For example, many people know how to drive, and they can describe their driving strategy.
- For example, you can ask an experienced driver what to do when :
  - he car is on the freeway going 100 km/h and
  - the car in front at a distance 10 m slows down to 95.

4. Why fuzzy in the first place: a brief reminder (cont-d)

- Most people will correctly reply "brake a little bit".
- But an automatic controller cannot understand this recommendation.
- It needs to know with what force – and for how many milliseconds – to press the brakes.
- We therefore need techniques for translating such imprecise knowledge into precise computer-understandable terms.
- Zadeh called such techniques *fuzzy* – and came up with several ideas on how to perform this translation.

5. Traditional (type-1) fuzzy techniques: description and limitations

- In the historically first fuzzy techniques provided by Zadeh, each term like "small" was represented by a function $m(x)$ that assigns:
  - to each possible value x of the corresponding quantity,
  - a degree to which this value x satisfies the desired property,
  - e.g., a degree to which x is small.
- This function is known as a *membership function* or, alternatively, a *fuzzy set*.

6. Traditional (type-1) fuzzy techniques: description and limitations (cont-d)

- The idea is to elicit these degrees from the expert, i.e., to ask the expert to mark this degree, e.g., on the interval $[0, 1]$, so that:
  - to values x which are definitely not small, we assign the degree $m(x) = 0$,
  - to values x which are definitely small, we assign the degree $m(x) = 1$, and
  - to values x that are somewhat small are assigned degrees $m(x)$ between 0 and 1.
- The main limitation of this approach is that:
  - similarly to how an expert cannot precisely describe what exactly control value to apply
  - all he/she can say is that this control can be small.,
  - the same expert has similar trouble describing his/her exact degree of confidence that a given control value x is small.

7. Traditional (type-1) fuzzy techniques: description and limitations (cont-d)

- For example, to many people, 25 C is comfortable while 28 C may be somewhat too hot (especially when there is humidity).
- But is the degree of hotness 0.8? 0.81? 0.79?
- This is difficult to distinguish.
- To be able to more adequately describe expert knowledge, Zadeh himself suggested alternatives to the type-1 techniques.
- For example:
  - instead of asking the expert to assign a single number $m(x)$,
  - we can ask the expert to assign an *interval* $[\underline{m}(x), \overline{m}(x)]$ of possible values.
- The resulting techniques are known as *interval-valued* fuzzy techniques.

8. Traditional (type-1) fuzzy techniques: description and limitations (cont-d)

- Alternatively, we can allow the expert to mark his/her degree of confidence that x is small by words from natural language.
- Then, we can use fuzzy techniques to translate this word into precise terms.
- In this case, the resulting degree $m(x)$ is not a single number, but a fuzzy subset of the interval $[0, 1]$.
- The resulting techniques are known as *type-2* fuzzy techniques.
- These techniques indeed lead to a more adequate representation of expert knowledge.

9. Fuzzy techniques in situations without expert knowledge

- Originally, fuzzy techniques were invented to describe expert knowledge.
- Later, it turned out that these techniques are also effective in situations when there is no expert knowledge at all.
- Namely, when we want to fit the data, it is often effective:
  - to find the fuzzy rules that best describe this data, and then
  - to use these rules to predict how the system will react to different inputs.
- This effectiveness makes perfect sense: this is how we learn to do things.
- We use the data to come up with informal (fuzzy) rules.
- And since we humans are a product of billions of years of improving evolution, this must be an effective strategy.

10. In situations without expert knowledge, interval and type-2 methods are more effective, but why?

- Interesting, it turned out that in situations without expert knowledge, interval-valued and type-2 fuzzy techniques often are more effective.
- Namely, when we compare type-2 and interval or type-2 methods with the same number of parameters:
  - interval and type-2 methods provide a more accurate description of the phenomenon and/or
  - a better quality control.
- But why?
- It is clear why interval-valued and type-2 techniques are more effective when the objective is to represent expert knowledge.
- But why are they more effective in situations when there is no expert knowledge?
- In this talk, we provide a theoretical explanation for this phemenon.

11. General idea: the more options, the better the approximation

- In order to explain the above puzzling phenomenon, let us briefly recall why, in general:
  - some approximations are more accurate and
  - some are less accurate.
- In general, there are many ways to approximate a dependence, there are many possible families of approximating functions.
- For each family, there is a certain number of different approximating options.
- Usually, the more options we have, the more accurate the approximation.
- We can illustrate this natural idea on the example of approximating numbers from the interval $[0, 1]$.

12. General idea: the more options, the better the approximation (cont-d)

- Let us gauge the quality of an approximation by the worst-case absolute value w of the difference between:
  - the actual value and
  - its approximation.
- If we only have one option to approximate, then the best option is to select $x_1 = 0.5$, then $w = 0.5$.
- If we are allowed 2 options, then the best way is to have $x_1 = 0.25$ and $x_2 = 0.75$, in which case $w = 0.25$.
- In general, if we are allowed n options, then the best idea is to have

$$x_1 = \frac{1}{2n}, x_2 = \frac{3}{2n}, \ldots, x_i = \frac{2i-1}{2n}, \ldots, x_n = \frac{2n-1}{2n}.$$

- In this case $w = \dfrac{1}{2n}$.

13. Consequence: the more parameters, the better the approximation (cont-d)

- In practice, each parameter is represented in a computer with some accuracy ε.
- Let us denote the width of the range of possible values of this parameter by W.
- This means that we can have W/ε possible distinguishable values of this parameter.
- Thus, if we have p parameters, we have $(W/\varepsilon)^p$ possible options.
- Hence clearly, the more parameters we have, the more options we have and thus, the more accurate will be the approximation.

14. How is all this related to type-1 and interval-valued fuzzy techniques

- In principle, each degree $m(x)$ can be any number from the interval $[0,1]$.
- There are infinitely many real numbers on the interval $[0,1]$.
- However, in a computer, we can only represent the degree $m(x)$ with some accuracy.
- As a result, we have only finitely many possible values of the degree:

$$m_1 < m_2 < \ldots < m_{M-1} < m_M \text{ for some } M.$$

- Similarly, there are infinitely many actual values x of the corresponding quantity.
- However, in a computer, we can only represent finitely many values

$$x_1 < x_2 < \ldots < x_{N-1} < x_N \text{ for some } N.$$

15. How is all this related to type-1 and interval-valued fuzzy techniques (cont-d)

- Thus, a general membership function can be represented by a finite number of degrees $m(x_1), \ldots, m(x_N)$.
- Each of these degrees takes values from the finite set $\mathcal{M} \stackrel{\text{def}}{=} \{m_1, \ldots, m_M\}$: $m(x_i) \in \mathcal{M}$.
- Usual membership functions:
  - first increase (in general, non-strictly) from 0 to 1,
  - then maybe take the value 1 for some time,
  - then decrease from 1 to 0.
- So, for some k for which the value $m(x_i)$ is the largest, we have:

$$m(x_1) \leq m(x_2) \leq \ldots \leq m(x_{k-1}) \leq m(x_k) \geq m(x_{k+1}) \geq \ldots \geq m(x_N).$$

- So, we have N values $m(x_i)$.

16. How is all this related to type-1 and interval-valued fuzzy techniques (cont-d)

- These values must satisfy the above monotonicity constraint for some k.
- What happens in the interval-valued case?
- In this case, for each x, we have an interval $[\underline{m}(x), \overline{m}(x)]$ of possible values of degree $m(x)$, i.e., in effect, two values $\underline{m}(x) \le \overline{m}(x)$; so:
  - when we consider the case of the same number of parameters as with the type-1 representation,
  - we need to only consider $N' = N/2$ possible values of x.
- Let us denote these values by $x'_1 < x'_2 < \ldots < x'_{N'}$.
- So we get the values $\underline{m}(x'_i)$ and $\overline{m}(x'_i)$.

17. How is all this related to type-1 and interval-valued fuzzy techniques (cont-d)

- The values of each of the lower and upper membership functions $\underline{m}(x)$ are $\overline{m}(x)$ must satisfy conditions similar to the type-1 case:

$$\underline{m}(x'_1) \leq \underline{m}(x'_2) \leq \ldots \leq \underline{m}(x'_{k'-1}) \leq \underline{m}(x'_{k'}) \geq \underline{m}(x'_{k'+1}) \geq \ldots \geq \underline{m}(x'_{N'});$$

$$\overline{m}(x'_1) \leq \overline{m}(x'_2) \leq \ldots \leq \overline{m}(x'_{k'-1}) \leq \overline{m}(x'_{k'}) \geq \overline{m}(x'_{k'+1}) \geq \ldots \geq \overline{m}(x'_{N'}).$$

- We also need constraints $\underline{m}(x'_i) \leq \overline{m}(x'_i)$ for all i.

18. Resulting theoretical explanation

- We plan to show that for the same N and M:
  - there are more tuples that satisfy the type-2 constraints
  - than tuples that satisfy the type-1 constraints.
- This will clearly imply that interval-valued fuzzy techniques can lead to a more accurate approximation than type-1.
- And this is exactly the empirical fact that needs to be explained; indeed:
  - for every sequence of values $m(x_i)$ that satisfies the type-1 monotonicity constraints,
  - we can form the interval-valued approximation in which:
    - for $2i \le k$, we have $\underline{m}(x'_i) = m(x_{2i-1})$ and $\overline{m}(x'_i) = m(x_{2i})$, and
    - for $2i > k$, we have $\underline{m}(x'_i) = m(x_{2i})$ and $\overline{m}(x'_i) = m(x_{2i-1})$.
- One can check that in this case, type-2 conditions are automatically satisfied.

19. Resulting theoretical explanation (cont-d)

- This means that interval-valued approximation has at least as many approximating options as the type-1 approximation.
- However, there are interval-valued approximations that cannot be obtained this way.
- Indeed, in each scheme obtained this way, when $2i \le k$, we have $\overline{m}(x'_i) \le \underline{m}(x'_{i+1})$.
- However, this is not always true for interval-valued fuzzy sets: e.g., we can have $\underline{m}(x'_{i+1}) < 1$ but $\overline{m}(x'_i) = 1$.
- Thus, with the same number of parameters:
  - interval-valued approximation scheme indeed contains more approximating options than the type-1 approximating scheme, and
  - thus, potentially leads to higher approximation accuracy.

20. Resulting theoretical explanation (cont-d)

- We provided a detailed explanation only for interval-valued techniques.
- A similar argument works for general type-2 techniques: in this case, we have even fewer constraints.

21. A word of caution

- We have explained why:
  - interval-valued and type-2 fuzzy methods often provide more accurate description of the data
  - than more traditional type-1 fuzzy techniques that use the same number of parameters.
- Accuracy is important.
- However, higher accuracy does not mean that we should use interval-valued and type-2 methods in all such applications.
- In general, interval-valued and type-2 algorithms require more computational steps that type-1 methods.
- So they require more time, more energy, etc. than type-1 techniques.

22. A word of caution (cont-d)

- As a result, interval-valued and type-2 methods may not be applicable in situations when our resources are limited.
- This happens, e.g.:
  - when we are solving a real-time control problem and/or
  - when we use a computational device with limited energy and computation budget.

23. 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),
  - HRD-1834620 and HRD-2034030 (CAHSI Includes),
  - EAR-2225395 (Center for Collective Impact in Earthquake Science C-CIES),
- by the AT&T Fellowship in Information Technology, and
- by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) Focus Program SPP 100+ 2388, Grant Nr. 501624329,