How to Gauge the Accuracy of Fuzzy-Control Recommendations: A Simple Idea

Patricia Melin[1], Oscar Castillo[1],
Andrzej Pownuk[2], Olga Kosheleva[2], and Vladik Kreinovich[2]

[1]Department of Computer Science, Tijuana Institute of Technology,
Tijuana, Baja California, Mexico,
ocastillo@tectijuana.mx, pmelin@tectijuana.mx
[2]University of Texas at El Paso, El Paso, TX 79968, USA
ampownuk@utep.edu, olgak@utep.edu, vladik@utep.edu

1. Need to Gauge Accuracy of Fuzzy Recommendations

- Fuzzy logic has been successfully applied to many different application areas, e.g., in control.
- A natural question is: with what accuracy do we need to implement this recommendation?
- In many applications, this is an important question:
  - it is often much easier to implement the control value approximately,
  - but maybe a more accurate actuator is needed?
- To answer this question, we must be able to gauge the accuracy of the corresponding recommendations.

2. Such Gauging Is Possible for Probabilistic Uncertainty

- Probabilistic uncertainty means that instead of the exact value x, we only know a probability distribution.
- This distribution can be described, e.g., by the probability density $\rho(x)$.
- If we need to select a single value x, a natural idea is to select, e.g., the mean value $\overline{\mathrm{x}} = \int x \cdot \rho(x)\,dx$.
- A natural measure of accuracy is the mean square deviation from the mean, known as the standard deviation:

$$\sigma \stackrel{\text{def}}{=} \sqrt{\int (x - \overline{\mathrm{x}})^2\,dx}.$$

- We need a similar formula for the fuzzy case.

3. How We Elicit Fuzzy Degrees: A Brief Reminder

- For each possible value x of the corresponding quantity, we ask the expert to mark:
  - on a scale from 0 to 1,
  - his/her degree of confidence that x satisfies the given property.
- For example, we ask the expert to specify the degree to which the value x is small.
- In some cases, this is all we need.
- However, in many other cases, we get a *non-normalized* membership function, for which $\max\limits_x \mu(x) < 1$.
- Most fuzzy techniques assume that the membership function is normalized.

4. How We Elicit Fuzzy Degrees (cont-d)

- So, we sometimes need to perform an additional step to get an easy-to-process membership function.
- Namely, we *normalize* the original values $\mu(x)$ by dividing them by the largest of the values $\mu(y)$:

$$\mu'(x) \stackrel{\text{def}}{=} \frac{\mu(x)}{\max\limits_y \mu(y)}.$$

- Sometimes, the experts have some subjective probabilities $\rho(x)$ assigned to different values x.
- In this case, when asked to indicate their degree of certainty, they list $\mu(x) = \rho(x)$.
- After normalizing this $\mu(x)$, we get the membership function $\mu(x) = \dfrac{\rho(x)}{\max\limits_y \rho(y)}$.

5. Let Us Use This Idea to Gauge the Accuracy of Fuzzy Recommendations

- We assign, to each probability density function $\rho(x)$, a membership function $\mu(x) = \dfrac{\rho(x)}{\max\limits_y \rho(y)}$.
- Vice versa, if we know that $\mu(x)$ was obtained by normalizing some $\rho(x)$, we can uniquely reconstruct $\rho(x)$:

$$\rho(x) = \frac{\mu(x)}{\int \mu(y)\,dy}.$$

- Our idea is then to use the probabilistic formulas corresponding to this artificial distribution.
- At first glance, this does not make sense.
- The probabilistic measure of accuracy is based on the assumption that we use the mean.
- But don't we use something else in fuzzy?

6. Let Us Use This Idea to Gauge the Accuracy of Fuzzy Recommendations (cont-d)

- Don't we use something else in fuzzy?
- Actually, not really.
- The mean of the distribution $\rho(x) = \dfrac{\mu(x)}{\int \mu(y)\,dy}$ is

$$\bar{\mathrm{x}} = \int x \cdot \rho(x)\,dx = \frac{\int x \cdot \mu(x)\,dx}{\int \mu(x)\,dx}.$$

- This is the centroid defuzzification – one of the main ways to transform $\mu(x)$ into a control recommendation.
- Since the above idea makes sense, let us use it to gauge the accuracy of the fuzzy control recommendation.

7. Resulting Recommendation

- For a given membership function $\mu(x)$, we usually generate the result $\overline{\mathrm{x}}$ of its centroid defuzzification.
- We should also generate, as a measure of the accuracy of this recommendation, the following value σ:

$$\sigma^2 = \int (x - \overline{\mathrm{x}})^2 \cdot \rho(x)\,dx = \frac{\int (x - \overline{\mathrm{x}})^2 \cdot \mu(x)\,dx}{\int \mu(x)\,dx} =$$

$$\frac{\int x^2 \cdot \mu(x)\,dx}{\int \mu(x)\,dx} - \left(\frac{\int x \cdot \mu(x)\,dx}{\int \mu(x)\,dx}\right)^2.$$

8. But What Should We Do in the Interval-Valued Fuzzy Case?

- Often, experts cannot tell us the exact values $\mu(x)$.
- Instead, for each x, they tell us the interval $[\underline{\mu}(x), \overline{\mu}(x)]$ of possible value of degree of confidence $\mu(x)$.
- For different functions $\mu(x) \in [\underline{\mu}(x), \overline{\mu}(x)]$, we get different values σ^2.
- It is desirable to find the range of possible values σ^2 when $\mu(x) \in [\underline{\mu}(x), \overline{\mu}(x)]$.

9. Resulting Algorithm

- For all possible pairs $\underline{x} < \overline{x}$, we compute $\sigma^2(\mu^-)$ and $\sigma^2(\mu^+)$, where:
  - $\mu^+(x) = \overline{\mu}(x)$ when $x < \underline{x}$ or $x > \overline{x}$, and $\mu^+(x) = \underline{\mu}(x)$ when $\underline{x} < x < \overline{x}$;
  - $\mu^-(x) = \underline{\mu}(x)$ when $x < \underline{x}$ or $x > \overline{x}$, and $\mu^-(x) = \overline{\mu}(x)$ when $\underline{x} < x < \overline{x}$.
- As the upper bound for σ^2, we take the maximum of the values $\sigma^2(\mu^+)$ corresponding to different pairs $\underline{x} < \overline{x}$.
- As the lower bound for σ^2, we take the minimum of the values $\sigma^2(\mu^-)$ corresponding to different pairs $\underline{x} < \overline{x}$.

10. Proof

- According to calculus, when $f(z)$ attains max on $[\underline{z}, \overline{z}]$ at $z_0 \in [\underline{z}, \overline{z}]$, then we have one of the three cases:
  - we can have $z_0 \in (\underline{z}, \overline{z})$, in which case $\dfrac{df}{dz}(z_0) = 0$;
  - we can have $z_0 = \underline{z}$, in which case $\dfrac{df}{dz}(z_0) \le 0$, or
  - we can have $z_0 = \overline{z}$, in which case $\dfrac{df}{dz}(z_0) \ge 0$.
- Similarly, when $f(z)$ attains min on $[\underline{z}, \overline{z}]$ at $z_0 \in [\underline{z}, \overline{z}]$, then we have one of the three cases:
  - we can have $z_0 \in (\underline{z}, \overline{z})$, in which case $\dfrac{df}{dz}(z_0) = 0$;
  - we can have $z_0 = \underline{z}$, in this case, in which case $\dfrac{df}{dz}(z_0) \ge 0$, or
  - we can have $z_0 = \overline{z}$, in which case $\dfrac{df}{dz}(z_0) \le 0$.

11. Proof (cont-d)

- Let us apply this to the dependence of σ on $\mu(a)$.
- Here, since $\int \mu(x)\,dx \approx \sum \mu(x_i)\cdot \Delta x_i$, we get:

$$\frac{\partial(\int \mu(x)\,dx)}{\partial(\mu(a))} = \Delta x, \quad \frac{\partial(\int x\cdot \mu(x)\,dx)}{\partial(\mu(a))} = a\cdot \Delta x \text{ and}$$

$$\frac{\partial(\int x^2\cdot \mu(x)\,dx)}{\partial(\mu(a))} = a^2\cdot \Delta x.$$

- By using the usual rules for differentiating the ratio, for the composition, and for the square, we get:

$$\frac{\partial(\sigma^2)}{\partial(\mu(a))} = \Delta x\cdot S(a), \text{ where } S(a) \stackrel{\text{def}}{=} \frac{a^2}{\int \mu(x)\,dx} -$$

$$\frac{\int x^2\cdot \mu(x)\,dx}{\left(\int \mu(x)\,dx\right)^2} - 2\cdot \overline{\mathrm{x}}\cdot \left(\frac{a}{\int \mu(x)\,dx} - \frac{\int x\cdot \mu(x)\,dx}{\left(\int \mu(x)\,dx\right)^2}\right).$$

12. Proof (cont-d)

- We are only interested in the sign of the derivative $\dfrac{\partial(\sigma^2)}{\partial(\mu(a))}$, and this is exactly the sign of $S(a)$.
- Similarly, the sign of $S(a)$ is the same as the sign of $s(a) \stackrel{\text{def}}{=} S(a) \cdot \int \mu(y)\,dy$, for which:

$$s(a) = a^2 - ((\overline{\mathrm{x}})^2 + \sigma^2) - 2 \cdot \overline{\mathrm{x}} \cdot (a - \overline{\mathrm{x}}).$$

- If we know the roots $\underline{x} < \overline{x}$ of this quadratic expression, we can conclude that this quadratic expression $s(a)$ is:
  - positive when $a < \underline{x}$ and
  - negative when $a > \overline{x}$.
- Here, the value $a = \overline{\mathrm{x}}$ is between $\underline{x}$ and $\overline{x}$, since for this value a, we have $s(\overline{\mathrm{x}}) = -\sigma^2 < 0$.

13. Proof (final)

- Thus, due to calculus, when $a < \underline{x}$ or $a > \overline{x}$:
  - to find $\overline{\sigma^2}$, we must take $\mu(a) = \overline{\mu}(a)$ and
  - to find $\underline{\sigma^2}$, we must take $\mu(a) = \underline{\mu}(a)$.
- When $\underline{x} < a < \overline{a}$, then, vice versa:
  - we need to take $\mu(a) = \underline{\mu}(a)$ to find $\overline{\sigma^2}$ and
  - we must take $\mu(a) = \overline{\mu}(a)$ to find $\underline{\sigma^2}$.
- This mathematical conclusion makes perfect sense:
  - to get the largest σ^2, we concentrate the distribution as much as possible on values far from $\bar{\mathrm{x}}$;
  - to get the smallest σ^2, we concentrate it as much as possible on values close to the mean $\bar{\mathrm{x}}$.
- Thus, we arrive at the above algorithm.

14. Acknowledgments

This work was supported in part by NSF grant HRD-1242122.