How to Take Into Account Monotonicity (and Other Properties) in Centroid Approach to Fuzzy and Intuitionistic Fuzzy Control

Olga Kosheleva and Vladik Kreinovich

University of Texas at El Paso
El Paso, Texas 79968, USA
{olgak,vladik}@utep.edu

1. Need for fuzzy control

- In many practical situations:
  - skilled human experts can make very effective control decision, very good medical recommendations, etc. and
  - a few very skilled experts can make almost perfect decisions and recommendations.
- It is therefore desirable to design an automatic system:
  - that would incorporate the knowledge and skills of such very skilled experts, and
  - that would help others make better decisions and recommendations.
- One of the difficulties in designing such systems is that in many cases, experts cannot explain their reasoning in precise terms.

2. Need for fuzzy control (cont-d)

- They provide rules that are formulated in terms in imprecise ("fuzzy") words from natural language like "small".
- To transform such knowledge into precise computer-understandable form, Lotfi Zadeh came up with techniques that he called *fuzzy*.

3. Fuzzy control: main ideas

- In the original version of fuzzy techniques, each imprecise property was described by assigning:
  - to each possible value x of the corresponding quantity,
  - a degree $\mu(x) \in [0, 1]$ to which the expert believes that this value satisfies the given property,
  - e.g., that x is small.
- Later on, several modifications were proposed.
- One of them is *intuitionistic* fuzzy logic in which:
  - in addition to the degree $\mu(x)$,
  - the expert is asked to provide a degree $\nu(x)$ to which, in the expert's opinion, the degree x *does not* satisfy the given property,
  - e.g., to what extent x is *not* small.

4. Fuzzy control: main ideas (cont-d)

- By using traditional fuzzy technique:
  - for each input x and for each possible value u,
  - we can estimate the degree $m(x, u)$ to which, based on the expert rules, u is a reasonable control value to apply.
- In the intuitionistic case, this value is supplemented by the value $n(x, u)$ to which u is *not* a reasonable control.
- If all we want is to provide a recommendation to a human expert, then this information is sufficient.
- However, in many practical cases, we want to automate this control (or decision making).
- In such cases, we need to transform such fuzzy recommendations into a value $\overline{u}(x)$ that we need to apply for the input x.
- Such transformation of fuzzy degrees into a single value is known as *defuzzification*.

5. Fuzzy control: main ideas (cont-d)

- The most widely used defuzzification – known as *centroid defuzzification* – has the following form:

$$\overline{u}(x) = \frac{\int m(x,u)\cdot u\,du}{\int m(x,u)\,du}.$$

- In the intuitionistic case, we first combine the value $m(x,u)$ and $n(x,u)$ into a single number – e.g., the number

$$c(x,y) = \frac{m(x,u) + (1 - n(x,u))}{2}.$$

- Then, we apply centroid defuzzification to this combined membership function:

$$\overline{u}(x) = \frac{\int c(x,u)\cdot u\,du}{\int c(x,u)\,du}.$$

6. Remaining question

- In many practical situations, in addition to expert rules, we have some additional information about the control function.
- For example, we may know that the control u should be an increasing function of x.
- For example, if we are describing braking, the larger the initial speed x, the more intense should be the braking.
- This monotonicity is usually reflected in the corresponding rules.
- However – due to the fuzzy character of these rules – it is reflected only approximately.
- As a result, in some cases, even when rules are monotonic, the fuzzy control is not everywhere monotonic.

7. Remaining question (cont-d)

- To avoid such situations, it is desirable to incorporate monotonicity – and other similar requirements – into fuzzy control.
- In this paper, we show how monotonicity – and other similar properties – can be incorporated into fuzzy and intuitionistic fuzzy control.

8. Idea

- Practitioners use centroid defuzzification because it is very empirically efficient.
- It is, in some reasonable sense, better than all other proposed defuzzification methods.
- This means that this method is optimal with respect to some reasonable criterion.
- So, a natural idea is:
  - to explicitly formulate this criterion, and then,
  - to solve the corresponding optimization problem under the monotonicity – (or any other) constraint.
- To come up with such a criterion, let us use general decision theory.

9. Decision theory: a brief reminder

- Decision theory describes how a rational person should make a decision.
- According to decision theory, a person's preferences can be described by a special function $U(x)$ called *utility*
  - out of all possible actions for which we get the result r_i with probability p_i,
  - the decision maker selects the action for which the following expression – known as *expected utility* – is the largest:

$$U = p_1 \cdot U(r_1) + p_2 \cdot U(r_2) + \ldots + p_n \cdot U(r_n).$$

- Utility is defined modulo adding a linear transformation.
- We can always choose a different starting point and a different measuring unit.

10. Decision theory: a brief reminder (cont-d)

- To make computations simpler, let us take, as a starting point, the ideal situation when we know the exact value of the best control.
- In this case, in all other situations, utility will be smaller than that – i.e., its values will be negative.

11. What do we need to do to apply this criterion

- This criterion is usually applied in situation when we know the probabilities and when we know the utility function.
- In our fuzzy control case, we do not have this information.
- So, to apply this approach to the fuzzy control situation, we need to estimate the probabilities and the utility.
- Let us do it.

12. How to estimate the probabilities

- To estimate the probabilities, let us take into account that one of the natural ways to estimate the fuzzy degree is to poll several experts.
- For example, to decide to what extend 30 C is hot, we can ask 10 experts.
- Suppose that 8 of them think that 30 degrees is hot.
- Then we assign, as the value $\mu(30)$ of the corresponding membership function, the degree $8/10 = 0.8$.
- From this viewpoint:
  - if for some input x, we have a control value u with a degree $m(x, u)$ to which u is reasonable,
  - we can interpret it as follows.
- Let us assume that we polled N experts.

13. How to estimate the probabilities (cont-d)

- This means that $\mu(x,u) = \dfrac{N(x,u)}{N}$, where $N(x,u)$ is the number of experts who believe that u is a reasonable control value for this x.
- Based on this formula, we can conclude that $N(x,u) = N \cdot m(x,u)$.
- So, if we have values $u_1, \ldots, u_n$ with degrees $m(x,u_1), \ldots, m(x,u_n)$, this means that:
  - we have $N \cdot m(x,u_1)$ experts who believe that u_1 is a reasonable control value,
  - we have $N \cdot m(x,u_2)$ experts who believe that u_2 is a reasonable control value,
  - $\ldots$, and
  - we have $N \cdot m(x,u_n)$ experts who believe that u_n is a reasonable control value.

14. How to estimate the probabilities (cont-d)

- So, overall, we have

$$\overline{N} \stackrel{\text{def}}{=} N\cdot m(x,u_1)+\ldots+N\cdot m(x,u_n) = N\cdot(m(x,u_1)+\ldots+m(x,u_m)) \text{ opinions.}$$

- We have no reasons to believe that one of these opinions is more valuable than others.
- So, it makes sense to assign to each of these opinions the same probability $p_i = \dfrac{1}{\overline{N}}$.
- Such an argument is often used in probability theory.
- It is known as *Laplace Indeterminacy Principle.*
- This principle makes perfect sense; e.g.:
  - if we have several suspects of a crime, and we have no reason to suspect one of them more than others,
  - then it makes sense to assign, to each of them, the same a priori probability.

15. How to estimate the probabilities (cont-d)

- With these probabilities, the expected utility takes the following form:

$$U(u) = \frac{N \cdot m(x, u_1)}{\overline{N}} \cdot U(x, u, u_1) + \ldots + \frac{N \cdot m(x, u_n)}{\overline{N}} \cdot U(x, u, u_n).$$

- Here, $U(x, u, u_i)$ is the utility of selecting the control value u when the best control value is u_i.

- All the terms in this sum have the same factor N and the same denominator $\overline{N}$.

- So, we can combine them together:

$$U(u) = \frac{N \cdot (m(x, u_1) \cdot U(x, u, u_1) + \ldots + m(x, u_n) \cdot U(x, u, u_n))}{N \cdot (m(x, u_1) + \ldots + m(x, u_m))}.$$

16. How to estimate the probabilities (cont-d)

- Dividing both numerator and denominator by N, we get a simplified expression

$$U(u) = \frac{m(x, u_1) \cdot U(x, u, u_1) + \ldots + m(x, u_n) \cdot U(x, u, u_n)}{m(x, u_1) + \ldots + m(x, u_m)}.$$

- To complete this description, we need to estimate the utility values.

17. How to estimate the utilities

- In most practical situations, we do not know how exactly utility depends on the control values.
- Such situations – when we do not know the exact form of dependence – is typical in many applications areas, e.g., in physics.
- In physics, the usual approach to such situations is to take into account that the dependence is usually smooth.
- Usually, smooth functions can be expanded in Taylor series over some small value.
- To get a good approximation to the function, we can just take the sum of the first few terms in the Taylor series.
- In the first approximation, we usually take:
  - the smallest number of terms
  - that are consistent with the qualitative understanding of the situations.

18. How to estimate the utilities (cont-d)

- Let us apply this approach here.
- We assume that the experts are reasonably good.
- This means that their estimates u_i are close to the actual best value u, i.e., that the difference $\Delta u_i \stackrel{\text{def}}{=} u_i - u$ is small.
- So, we can expand the expression $U(x, u, u_i) = U(x, u, u + \Delta u_i)$ in Taylor series in terms of Δu_i.
- We know that:
  - the utility $U(x, u, u + \Delta u_i)$ is equal to 0 when $u = u_i$, i.e., when $\Delta u_i = 0$, and
  - it is smaller than 0 for all other values Δu_i.
- This means that this function of Δu_i attains its maximum when $\Delta u_i = 0$.
- This fact implies that we cannot only keep terms which are linear in Δu_i – since linear functions do not attain their maxima.

19. How to estimate the utilities (cont-d)

- Thus, we need to consider quadratic terms.
- In general, such a quadratic expression takes the following form:

$$U(x,u,u+\Delta u_i)=a_0(x,u)+a_1(x,u)\cdot \Delta u_i+a_2(x,u)\cdot (\Delta u_i)^2.$$

- The fact that this value is equal to 0 when $\Delta u_i = 0$ implies that $a_0(x,u)=0$.
- The fact that the function attains its maximum for $\Delta u_i=0$ implies that the derivative of the expression is equal to 0 when $\Delta u_i=0$.
- This implies that $a_1(x,u)=0$; thus:

$$U(x,u,u+\Delta u_i)=a_2(x,u)\cdot (\Delta u_i)^2.$$

- That all utility values are non-negative means that $a_2(x,u)<0$.

20. How to estimate the utilities (cont-d)

- We can simplify this expression even further if we take into account that:
  - as we have mentioned earlier,
  - the utility usually smoothly depends on u.
- This means that the function $a_2(x,u)$ also smoothly depends on u.
- So, we can select some value u_0 close to all the expert estimates, so that $a(x,u)=a(x,u+\Delta u)$, where we denoted $\Delta u \stackrel{\text{def}}{=} u-u_0$.
- By applying the same Taylor approximation idea as before, we conclude that in the first approximation, we have

$$a_2(x,u)=a_2(x,u_0+\Delta u)=a_2(x,u_0)+a_3(x,u_0)\cdot\Delta u.$$

- Substituting this expression into the above formula, we get

$$U(x,u,u+\Delta u_i)=a_2(x,u_0)\cdot(\Delta u_i)^2+a_3(x,u_0)\cdot\Delta u\cdot(\Delta u_i)^2.$$

21. How to estimate the utilities (cont-d)

- The last term in this expression is already cubic in terms of small differences Δu and Δu_i.
- Since we are only considering quadratic approximation, we can safely ignore this terms and get:

$$U(x,u,u_i) = U(x,u,u+\Delta u_i) = a_2(x,u_0)\cdot(\Delta u_i)^2 = a_2(x,u_0)\cdot(u_i-u)^2.$$

22. Now, we get the final expression for utility

- Substituting this expression into the formula for $U(u)$, we conclude that

$$U(u) = \frac{m(x,u_1)\cdot a_2(x,u_0)\cdot (u_1 - u)^2 + \ldots + m(x,u_n)\cdot a_2(x,u_0)\cdot (u_n - u)}{m(x,u_1) + \ldots + m(x,u_m)}$$

- By definition of utility, for each x, we need to select the control value u for which the utility is the largest possible.
- If we multiply all the values of the objective function by a positive constant, this does not change which value is larger.
- E.g., who is the richest person does not change whether we count his/her amount in US dollars or in Euros.
- Thus, for the purpose of maximizing utility, we can:
  - multiply all the values of the utility function by the denominator of the expression, and
  - divide by the absolute value $|a_2(x, u_0)|$.

23. Now, we get the final expression for utility (cont-d)

- This way, we get the equivalent maximizing function

$$\overline{U}(u) \stackrel{\text{def}}{=} -(m(x,u_1)\cdot(u_1-u)^2+\ldots+m(x,u_n)\cdot(u_n-u)^2).$$

- Maximizing this expression is equivalent to minimizing its opposite

$$-\overline{U}(u) = m(x,u_1)\cdot(u_1-u)^2+\ldots+m(x,u_n)\cdot(u_n-u)^2.$$

- In the continuous case, the sum becomes the integral, so we want to minimize the expression

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

- From the computational viewpoint, the minimized expressions are a particular case of the general Least Squares problem.
- In this method, we minimize the sum of squares of some expressions.

24. If there are no additional constraints, this expression leads exactly to centroid defuzzification

- Indeed, if we differentiate the above expression with respect to u and equate the derivative to 0, we conclude that

$$2m(x, u_1) \cdot (u - u_1) + \ldots + 2m(x, u_n) \cdot (u - u_n) = 0.$$

- Let us divide both sides by 2, move all the terms not containing u to the right-hand side.
- Let us now group together all the terms proportional to u in the left-hand side.
- Then we get the following linear equation with one unknown:

$$(m(x, u_1) + \ldots + m(x, u_n)) \cdot u = m(x, u_1) \cdot u_1 + \ldots + m(x, u_n) \cdot u_n.$$

25. If there are no additional constraints, this expression leads exactly to centroid defuzzification (cont-d)

- Thus:

$$u = \frac{m(x, u_1) \cdot u_1 + \ldots + m(x, u_n) \cdot u_n}{m(x, u_1) + \ldots + m(x, u_n)}.$$

- In the continuous case, when sums turn into integrals, we get exactly the centroid defuzzification.

26. What to do if we have an additional constraint on the dependence $\overline{u}(x)$ – e.g., monotonicity

- So far, we only considered utility corresponding to a single input x.
- If we impose an additional restriction that relates controls corresponding to different inputs, then we need to consider the overall utility.
- This is the sum – or integral – of the expression $U(x)$ over all x.
  - similarly to what we argued about the possibility to ignore, in this approximation, the dependence of $a_2(x, u)$ on u,
  - we can argue that – at least for some range of inputs – we can also ignore the dependence on x, i.e., replace $a_2(x, u_0)$ with $a_2(x_0, u_0)$.
- In this case, maximizing the overall utility for all possible values of x is equivalent to minimizing the following integral

$$\int_x \int_u m(x, u) \cdot (u - \overline{u}(x))^2 \, du \, dx.$$

27. What to do if we have an additional constraint on the dependence $\overline{u}(x)$ – e.g., monotonicity (cont-d)

- Suppose that:
  - we know some additional constraint(s) on the function $\overline{u}(x)$,
  - e.g., that this dependence should be monotonic.
- Then we should find:
  - among all the functions $\underline{u}(x)$ that satisfy these constraints,
  - the one for which the above expression attains the smallest possible value.

28. How to actually find such control: case of monotonicity

- The minimized function is still the sum (integral) of squares, so it is still a particular case of the general Least Squares problem.
- For the case when the restriction is that the function $\overline{u}(x)$ is a (non-strictly) increasing function of x, the problem becomes as follows:
- To find an increasing solution to the Least Squares problem.
- In data processing, this problem is known as *isotonic regression*.
- There exist efficient algorithms for solving this problem.

29. What to do in the intuionistic fuzzy case

- In the intuitionistic fuzzy case, for each x and u:
  - in addition to $N \cdot m(x, u)$ experts who believe that u is a reasonable control value,
  - there are also $N \cdot n(x, u)$ experts who believe that u is *not* a reasonable control value.
- To take their opinions into account, we need to describe it in terms of utility.
- So, we need to describe the utility $V(x, u, u_i)$ corresponding to the case when:
  - the actual best control value is u
  - but the i-th expert says that u_i is not a good value.
- Similarly to the fuzzy case, we can select a starting point for the utility as the point for which $u_i = u$.

30. What to do in the intuionistic fuzzy case (cont-d)

- The difference is that now this is the worst case, when the expert is completely wrong.
- So, in this case, for all $u_i \neq u$, the utility is larger.
- Thus, the function $V(x, u, u_i)$ should now attain its *minimum* for $u_i = u$.
- Similarly to the fuzzy case, this means that:
  - we cannot restrict ourselves to linear terms in the Taylor expansion,
  - we need to take quadratic terms into account.
- If we restrict ourselves to quadratic terms, then, similarly to the fuzzy case, we get
$$V(x, u, u_i) = b_2(x, u_0) \cdot (u - u_i)^2.$$
- This time, the expression $b_2(x, u_0)$ is positive.

31. What to do in the intuionistic fuzzy case (cont-d)

- Let us now add up all terms corresponding to positive and negative opinions and perform the same simplifications as in the fuzzy case.
- We conclude that maximizing utility is equivalent to minimizing the expression

$$m(x,u_1)\cdot(u-u_1)^2 - r\cdot n(x,u_1)\cdot(u-u_1)^2+\ldots+$$
$$m(x,u_n)\cdot(u-u_n)^2 - r\cdot n(x,u_n)\cdot(u-u_n)^2.$$

- Here, $r \stackrel{\text{def}}{=} b(x,u_0)/|a(x,u_0)|$.
- In other words, we want to minimize the expression

$$c(x,u_1)\cdot(u-u_1)^2+\ldots+c(x,u_n)\cdot(u-u_n)^2.$$

- Here, $c(x,u_i)=m(x,u_i)-r\cdot n(x,u_i)$ is a combination of positive and negative membership functions.
- In the continuous case, thie means minimizing the expression

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

32. What to do in the intuionistic fuzzy case (cont-d)

- Similarly to the fuzzy case:
  - in the absence of additional constraints,
  - minimizing the above expression is equivalent to applying centroid defuzzification to the combined membership function.
- In the case of constraints, we need to find the function $\overline{u}(x)$ that:
  - among all the functions that satisfy the given constraints,
  - minimizes the expression

$$\int_x \int_u c(x,u) \cdot (u-\overline{u})^2 \, du \, dx.$$

- For the case of monotonicity, we can similarly use isotonic regression techniques to find the corresponding function.

33. Conclusion

- Let us summarize our results.
- In this paper, we discussed a problem of:
  - how to take into account additional a priori constraints on the control function $\overline{u}(x)$
  - in fuzzy and intuitionstic fuzzy control.
- To solve this problem, we used the general techniques of decision theory.
- By applying these techniques, we came up with the following conclusions.
- In the fuzzy case, we have a membership function $m(x, u)$ – that describes to what extent u is an appropriate control for the input x.

34. Conclusion (cont-d)

- To find the corresponding control, we need to find:
  - among all the functions $\overline{u}(x)$ that satisfy the given constraints,
  - the function that minimizes the following expression:

$$\int_x \int_u m(x,u) \cdot (u - \overline{u}(x))^2 \, du \, dx.$$

- In the intuitionistic fuzzy case, we also have the negative membership function $n(x,u)$.
- It describes to what extent u is *not* an appropriate control for the input x.
- Then, we need to minimize the following expression:

$$\int_x \int_u c(x,u) \cdot (u - \overline{u}(x))^2 \, du \, dx.$$

- Here $c(x,u) = m(x,u) - r \cdot n(x,u)$ is a combination of positive and negative membership functions.

35. Conclusion (cont-d)

- An important case is when the constraint means that the function $\overline{u}(x)$ should be an increasing function of x.
- Then, we can use efficient algorithms of isotonic regression to minimize the above expressions.

36. 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,