Semi-Heuristic Target-Based Fuzzy Decision Procedures: Towards a New Interval Justification

Christian Servin
Program of Computational Science
University of Texas at El Paso
500 W. University
El Paso, TX 79968, USA
christians@miners.utep.edu

Van-Nam Huynh and Yoshiteru Nakamori
Japan Advanced Institute of
Science and Technology (JAIST)
1-1 Asahidai, Nomi, Ishikawa 923-1292, Japan
{huynh, nakamori}@jaist.ac.jp

***Abstract*—To more adequately describe human decision making, V.-N. Huynh, Y. Nakamori, and others proposed a special semi-heuristic target-based fuzzy decision procedure. A usual justification for this procedure is based on the selection of the simplest possible membership functions and "and"- and "or"-operations; if we use more complex membership functions and "and"- and "or"-operations, we get different results. Interestingly, in practical applications, the procedure based on the simplest choices most adequately describes human preferences. It is therefore desirable to come up with a justification that explains this empirical fact. Such a justification is proposed in this paper.**

I. Formulation of the Problem

Traditional approach to decision making. Traditional decision making techniques (see, e.g., [1], [2], [9], [11], [14]) deal with the problems in which the quality of each possible alternative is characterized by the values of several quantities. For example, when we buy a car, we are interested in its cost, its energy efficiency, its power, size, etc. Specifically, for each of these quantities, we usually have some desirable range of values.

Sometimes, there is only one alternative that satisfies all these requirements. In other real-life situations, there are several different alternatives all of which satisfy all these requirements. In such cases, the traditional decision making approach usually assumes that there is an objective function that describes the user's preferences; the corresponding techniques then enable us to select an alternative with the largest possible value of this objective function.

Traditional approach to decision making: limitations. The traditional approach to decision making assumes:

- that the user knows exactly what he or she wants — i.e., knows the objective function – and
- that the user also knows exactly what he or she will get as a result of each possible decision.

In practice, the user is often uncertain:

- the user is often uncertain about his or her own preferences, and
- the user is often uncertain about possible consequences of different decisions.

It is therefore desirable to take this uncertainty into account when we describe decision making.

Fuzzy target approach. To describe actual decision making, the authors of [3], [4], [5], [6], [7], [15], [16] proposed an alternative approach. In this approach, to properly take uncertainty into account, for each numerical characteristic of a possible decision, we form two fuzzy sets:

- first, we form a fuzzy set $\mu_i(x)$ describing the users' ideal value;
- then, we form the fuzzy set $\mu_a(x)$ describing the users' impression of the actual value.

For example, a person wants a well done steak, and the steak comes out as medium well done. In this case, we form a fuzzy set $\mu_i(x)$ corresponding to "well done", and we form a fuzzy set $\mu_a(x)$ corresponding to "medium well done".

How can we estimate to what extent the actual result is satisfactory? All we know is the membership functions corresponding to ideal (desired) and actual. If these sets were crisp, then we could say that it is possible that the proposed solution is satisfactory if some of the possibly actual values is also desired. In the fuzzy case, when we only have degrees describing to what exact each value is possible and to what extent each value is desired, we can only talk about the degree to which the proposed solution can be desired. To find this degree, we can use the fact that a possible decision is satisfactory if:

- either the actual value is x_1, and this value is desired,
- or the actual value is x_2, and this value is desired,
- $\ldots$,

where x_1, x_2, $\ldots$, go over all possible values of the desired quantity.

We know the membership functions $\mu_i(x)$ and $\mu_a(x)$. This means that for each value x_k, we know the the degree $\mu_a(x_k)$ with which this value is actual, and the degree $\mu_a i(x_k)$ to which this value is desired. If we use $\min(a, b)$ to describe "and" (the simplest possible choice of an "and"-operation [10], [13]), then we can estimate the degree to which the value x_k is both actual *and* desired as

$$\min(\mu_a(x_k), \mu_i(x_k)).$$

If we now use $\max(a,b)$ to describe "or" (the simplest possible choice of an "or"-operation [10], [13]), then we can estimate the degree d to which the two fuzzy sets match as

$$d = \max_x \min(\mu_a(x), \mu_i(x)).$$

How can we elicit the corresponding membership functions? In principle, membership functions can have different shapes. It is known, however, that in many applications (e.g., in intelligent control), the actual shape of a membership function does not affect the result. Thus, it is reasonable to use the simplest possible membership functions – symmetric triangular ones [10], [13].

To describe a symmetric triangular function, it is sufficient to know the support of this function, i.e., the interval $[\underline{x}, \overline{x}]$ on which this function is different from 0. This interval can also be described as $[\widetilde{x} - \Delta_x, \widetilde{x} + \Delta_x]$, where:

- $\widetilde{x} = \dfrac{\underline{x} + \overline{x}}{2}$ is the interval's midpoint, and
- $\Delta_x = \dfrac{\overline{x} - \underline{x}}{2}$ is the interval's half-width.

The corresponding membership function:

- linearly increases from 0 to 1 on the first half-interval $[\widetilde{x} - \Delta_x, \widetilde{x}]$, and then
- linearly decreases from 1 to 0 on the second half-interval $[\widetilde{x}, \widetilde{x} + \Delta_x]$.

As we have just mentioned, once we know the interval, we can uniquely determine the corresponding membership function. So, to elicit the membership function from the user, it is sufficient to elicit the corresponding interval. How can we elicit the interval from the user? To elicit this interval, we can simply ask the users which values are possible, and then take the smallest of these possible values as $\underline{x}$ and the largest of these possible values as $\overline{x}$.

So, to get the membership function $\mu_i(x)$ describing the desired situation, we can ask the user for all the values $a_1, \ldots, a_n$ which, in their opinion, satisfy the requirement, and then take the smallest of these values as $\underline{a}$ and the largest of these values as $\overline{a}$.

Similarly, to get the membership function $\mu_a(x)$ describing the result of a proposed decision, we can ask the user for all the values $b_1, \ldots, b_m$ which, in their opinion, satisfy the corresponding property (like "medium well done"), and then take the smallest of these values as $\underline{b}$ and the largest of these values as $\overline{b}$.

Fuzzy target approach: successes. The above approach works well, e.g., in predicting how the customers buying handcrafted souvenirs select among "almost-desirable" souvenirs when their "ideal" souvenir is not available.

Fuzzy target approach: the remaining problem. it is somewhat heuristic: it is based on selecting the simplest possible membership function and the simplest possible "and"- and "or"-operations. If we use more complex membership functions and "and"- and "or"-operations, we will get different results.

The fact that the existing approach works well in practice indicates that there should be a justification for this method which goes beyond simplicity. Such a justification would give us more confidence in using these techniques.

What we do in this paper. In this paper, we provide a justification for the above semi-heuristic target-based fuzzy decision procedure.

II. Solution to the Problem

What we know: reminder. As we have mentioned earlier, while the procedure that we want to justify uses fuzzy techniques, all we know is two intervals:

- an interval $[\underline{a}, \overline{a}] = [\widetilde{a} - \Delta_a, \widetilde{a} + \Delta_a]$ describing the set of all desired values, and
- an interval $[\underline{b}, \overline{b}] = [\widetilde{b} - \Delta_b, \widetilde{b} + \Delta_b]$ describing the set of all the values which are possible for a given decision.

The formula that we need to justify. Let us describe an explicit expression for the formula that we need to justify – the formula describing the degree d to which the proposed decision leads to the desired result.

The above procedure is symmetric with respect to changing a and b. So, if necessary, we can swap a and b. Thus, without losing generality, we can assume that $\widetilde{a} \le \widetilde{b}$.

One can prove that the maximum d of the function $\min(\mu_a(x), \mu_i(x))$ is attained when the values $\mu_a(x)$ and $\mu_i(x)$ are equal, i.e., at a point x_m for which $\mu_a(x_m) = \mu_i(x_m)$. Indeed, if, e.g., $\mu_a(x_m) > \mu_i(x_m)$, then $\min(\mu_a(x_m), \mu_i(x_m)) = \mu_i(x_m)$. In this case, we cannot have $\mu_i(x_m) = 1$, so we must have $\mu_i(x_m) < 1$. In this case, by modifying x_m a little bit, we can increase the value $\mu_i(x)$ and thus, achieve a larger value of the $\min(\mu_a(x), \mu_i(x))$ – which contradicts to our assumption that the function attains maximum at x_m. Similarly, the maximum cannot be attained when $\mu_a(x_m) < \mu_i(x_m)$, so it has to be attained when $\mu_a(x_m) = \mu_i(x_m)$.

In this case, the desired maximum d is equal to $d = \mu_a(x_m) = \mu_i(x_m)$.

Since $\widetilde{a} \le \widetilde{b}$, the equality $\mu_a(x_m) = \mu_i(x_m)$ is attained when:

- the value on the decreasing part of the a-membership function $\mu_a(x)$ (that linearly goes from 1 at $\widetilde{a}$ to 0 at $\widetilde{a} + \Delta_a$) coincides with
- the value on the increasing part of the b-membership function $\mu_b(x)$ (that linearly goes from 0 at $\widetilde{b} - \Delta_b$ to 1 for $\widetilde{b}$).

By applying the general formula

$$y = y_1 + \frac{y_2 - y_1}{x_2 - x_2} \cdot (x - x_1)$$

for a straight line that takes value y_1 at $x = x_1$ and value y_2 at $x = x_2$, we conclude that

$$\mu_a(x) = 1 - \frac{x - \widetilde{a}}{\Delta_a} \text{ and } \mu_b(x) = \frac{x - (\widetilde{b} - \Delta_b)}{\Delta_b}.$$

Thus, the condition $\mu_a(x_m) = \mu_i(x_m)$ takes the form

$$1 - \frac{x_m - \widetilde{a}}{\Delta_a} = \frac{x_m - (\widetilde{b} - \Delta_b)}{\Delta_b}.$$

By opening the parentheses, we get

$$1 - \frac{x_m - \widetilde{a}}{\Delta_a} = \frac{x_m - \widetilde{b} + \Delta_b}{\Delta_b} = 1 - \frac{\widetilde{b} - x_m}{\Delta_b},$$

or, equivalently,

$$\frac{x_m - \widetilde{a}}{\Delta_a} = \frac{\widetilde{b} - x_m}{\Delta_b}.$$

Multiplying both side by Δ_a and Δ_b, we get

$$(x_m - \widetilde{a}) \cdot \Delta_b = (\widetilde{b} - x_m) \cdot \Delta_a.$$

Moving all the terms containing the unknown x_m into the left-hand side and all the other terms into the right-hand side, we get

$$x_m \cdot (\Delta_a + \Delta_b) = \widetilde{a} \cdot \Delta_b + \widetilde{b} \cdot \Delta_a,$$

hence

$$x_m = \frac{\widetilde{a} \cdot \Delta_b + \widetilde{b} \cdot \Delta_a}{\Delta_a + \Delta_b}.$$

Substituting this value into the formula for $\mu_a(x)$, we conclude that the desired maximum value d is equal to

$$d = \mu_a(x_m) = 1 - \frac{x_m - \widetilde{a}}{\Delta_a}.$$

Here,

$$x_m - \widetilde{a} = \frac{\widetilde{a} \cdot \Delta_b + \widetilde{b} \cdot \Delta_a}{\Delta_a + \Delta_b} - \widetilde{a}.$$

By moving both terms in the right-hand side to the common denominator, we conclude that

$$x_m - \widetilde{a} = \frac{\widetilde{a} \cdot \Delta_b + \widetilde{b} \cdot \Delta_a - \widetilde{a} \cdot \Delta_a - \widetilde{a} \cdot \Delta_b}{\Delta_a + \Delta_b} = \frac{\widetilde{b} \cdot \Delta_a - \widetilde{a} \cdot \Delta_a}{\Delta_a + \Delta_b}.$$

Thus,

$$d = 1 - \frac{\widetilde{b} - \widetilde{a}}{\Delta_a + \Delta_b}. \tag{1}$$

This is the formula that we need to justify.

Our main idea. If we knew the exact values of a and b, then we would be able to conclude that either $a = b$, or $a < b$, or $b < a$.

In reality, we know the values a and b with uncertainty. So, even if the actual values a and b are the same, we may get approximate values which are different; it is reasonable to assume that if the actual values are the same, then we have the same probability $\mathrm{Prob}(a > b)$ and $\mathrm{Prob}(b > a)$ of observing $a > b$ and $a < b$ are the same, i.e., that both these probabilities are equal to 1/2. If the probabilities that $a > b$ and that $a < b$ differ, this is an indication that the actual value a and b are different.

Thus, it is reasonable to use the difference

$$|\mathrm{Prob}(a > b) - \mathrm{Prob}(b > a)|$$

as the degree to which a and b may be different.

How we can estimate the probabilities $\mathrm{Prob}(a > b)$ and $\mathrm{Prob}(b > a)$. If we know the exact values of a and b, then we can check the inequality $a > b$ by computing the difference $a - b$ and comparing this difference to 0.

In real life, as we have mentioned, we only know a and b with interval uncertainty, i.e., we only know that

$$a \in [\widetilde{a} - \Delta_a, \widetilde{a} + \Delta_a] \text{ and } b \in [\widetilde{b} - \Delta_b, \widetilde{b} + \Delta_b].$$

In this case, we do not know the exact value of the difference $a - b$, we only know the range of possible values of this difference.

Such a range can be computed by using interval arithmetic; see, e.g., [12]. Namely:

- the smallest possible value of the difference $a - b$ is attained when a attains its smallest possible value $\widetilde{a} - \Delta_a$ and b attains its largest possible value $\widetilde{b} + \Delta_b$; the resulting difference is equal to
$$\widetilde{a} - \Delta_a - (\widetilde{b} + \Delta_b) = (\widetilde{a} - \widetilde{b}) - (\Delta_a + \Delta_b);$$
- the largest possible value of the difference $a - b$ is attained when a attains its largest possible value $\widetilde{a} + \Delta_a$ and b attains its smallest possible value $\widetilde{b} - \Delta_b$; the resulting difference is equal to
$$\widetilde{a} + \Delta_a - (\widetilde{b} - \Delta_b) = (\widetilde{a} - \widetilde{b}) + (\Delta_a + \Delta_b).$$

Thus, the only thing we know about the difference $a - b$ is that this difference belongs to the interval

$$[(\widetilde{a} - \widetilde{b}) - (\Delta_a + \Delta_b), (\widetilde{a} - \widetilde{b}) + (\Delta_a + \Delta_b)].$$

We do not have any reason to assume that some values from this interval are more probable and some are less probable. It is therefore reasonable to assume that all the values from this interval are equally probable, i.e., that the corresponding probability distribution is *uniform* on this interval.

Comment. This argument is widely used in data processing; it is called *Laplace principle of indifference* or *Laplace principle of insufficient reason*. Its most adequate mathematical description is the Maximum Entropy approach (see, e.g, [8]) – widely used in statistics – according to which, if several different probability distributions $\rho(x)$ are consistent with our knowledge, we should select the *least informative* one, i.e., the one for which the entropy $S = -\int \rho(x) \cdot \ln(\rho(x))\, dx$ is the largest possible. In our case, all we know about the corresponding probability density function $\rho(x)$ is that it is located on a given interval $[\underline{c}, \overline{c}]$, i.e., that $\rho(x) = 0$ outside this interval. Thus, in all integrations containing $\rho(x)$, we can skip the parts where this function is 0 and only consider values from the interval $[\underline{c}, \overline{c}]$. So, to find the appropriate distribution, we must maximize the entropy

$$S = -\int_{\underline{c}}^{\overline{c}} \rho(x) \cdot \ln(\rho(x))\, dx \to \max$$

under the constraints that $\rho(x) \geq 0$ and

$$\int_{\underline{c}}^{\overline{c}} \rho(x)\,dx = 1.$$

By using Lagrange multiplier method, we can reduce this constraint optimization to unconstrained optimization problem

$$-\int_{\underline{c}}^{\overline{c}} \rho(x) \cdot \ln(\rho(x))\,dx + \lambda \cdot \left(\int_{\underline{c}}^{\overline{c}} \rho(x)\,dx - 1\right) \to \max.$$

Differentiating the objective function with respect to $\rho(x)$, we conclude that

$$-\rho(x) - 1 + \lambda = 0,$$

i.e., that $\ln(\rho(x)) = \lambda - 1$ and thus, $\rho(x) = \exp(\lambda - 1)$. This value is the same for all $x \in [\underline{c}, \overline{c}]$, so we indeed get a uniform distribution.

How we can estimate the probabilities $\mathrm{Prob}(a > b)$ **and** $\mathrm{Prob}(b > a)$ **(cont-d).** In our approach:

- as an estimate for the probability $\mathrm{Prob}(a > b)$, we take the probability $\mathrm{Prob}(a - b > 0)$ that the difference $a - b$ is positive, and
- as an estimate for the probability $\mathrm{Prob}(a < b)$, we take the probability $\mathrm{Prob}(a - b < 0)$ that the difference $a - b$ is negative.

Now that we have assumed that the probability distribution on the set of all the values $a - b$ is uniformly distributed on the interval

$$[(\widetilde{a} - \widetilde{b}) - (\Delta_a + \Delta_b), (\widetilde{a} - \widetilde{b}) + (\Delta_a + \Delta_b)],$$

we can find the numerical values for both probabilities. Namely,

- values $a - b > 0$ form a subinterval

$$(0, (\widetilde{a} - \widetilde{b}) + (\Delta_a + \Delta_b)];$$

- values $a - b < 0$ form a subinterval

$$[(\widetilde{a} - \widetilde{b}) - (\Delta_a + \Delta_b), 0).$$

In a uniform distribution, the probability to be in a subinterval is proportional to the width of this subinterval. The coefficient of proportionality can be found from the condition that the overall probability to be in the entire interval is equal to 1. Thus, when we have a uniform distribution on an arbitrary interval $[\underline{c}, \overline{c}]$, the probability p to be in a subinterval $[\underline{d}, \overline{d}] \subseteq [\underline{c}, \overline{c}]$ is equal to the ratio of the widths of these intervals:

$$p = \frac{\overline{d} - \underline{d}}{\overline{c} - \underline{c}}.$$

In our case, the width of the big interval is equal to

$$((\widetilde{a} - \widetilde{b}) + (\Delta_a + \Delta_b)) - ((\widetilde{a} - \widetilde{b}) - (\Delta_a + \Delta_b)) = 2 \cdot (\Delta_a + \Delta_b),$$

and thus, the probabilities $\mathrm{Prob}(a > b)$ and $\mathrm{Prob}(a < b)$ are equal to

$$\mathrm{Prob}(a > b) = \frac{(\widetilde{a} - \widetilde{b}) + (\Delta_a + \Delta_b)}{2 \cdot (\Delta_a + \Delta_b)};$$

$$\mathrm{Prob}(a < b) = \frac{(\Delta_a + \Delta_b) - (\widetilde{a} - \widetilde{b})}{2 \cdot (\Delta_a + \Delta_b)}.$$

So, the desired difference $\mathrm{Prob}(a > b) - \mathrm{Prob}(b > a)$ takes the form

$$\mathrm{Prob}(a > b) - \mathrm{Prob}(b > a) =$$

$$\frac{(\widetilde{a} - \widetilde{b}) + (\Delta_a + \Delta_b)}{2 \cdot (\Delta_a + \Delta_b)} - \frac{(\Delta_a + \Delta_b) - (\widetilde{a} - \widetilde{b})}{2 \cdot (\Delta_a + \Delta_b)} =$$

$$\frac{2 \cdot (\widetilde{a} - \widetilde{b})}{2 \cdot (\Delta_a + \Delta_b)} = \frac{\widetilde{a} - \widetilde{b}}{\Delta_a + \Delta_b}.$$

Since $\widetilde{a} \geq \overline{b}$, we can conclude that

$$|\mathrm{Prob}(a > b) - \mathrm{Prob}(b > a)| = \frac{\widetilde{a} - \widetilde{b}}{\Delta_a + \Delta_b}. \qquad (2)$$

Conclusion. By comparing:

- the above formula (1) for the degree d to which the alternative a fits the fuzzy target b
- with the formula (2) for the probability

$$|\mathrm{Prob}(a > b) - \mathrm{Prob}(b > a)|$$

with which the alternative a and the fuzzy target b are different,

we can see that

$$d + |\mathrm{Prob}(a > b) - \mathrm{Prob}(b > a)| = 1.$$

Thus, the degree d can be described, in reasonable probabilistic terms, as

$$d = 1 - |\mathrm{Prob}(a > b) - \mathrm{Prob}(b > a)|.$$

We have therefore produced a new justification for the above complex formula for d, the justification that does not use any simplifying assumptions and which is, therefore, applicable in the general case.

Acknowledgments

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

References

[1] P. C. Fishburn, *Utility Theory for Decision Making*, John Wiley & Sons Inc., New York, 1969.

[2] P. C. Fishburn, *Nonlinear Preference and Utility Theory*, The John Hopkins Press, Baltimore, Maryland, 1988.

[3] A. Hiramatsu, V.-N. Huynh, and Y. Nakamori, "A behavioral decision model based on fuzzy targets in decision making using weather information", *Journal of Advanced Computational Intelligence and Intelligent Informatics*, 2008, Vol. 12, No. 5, pp. 435–442.

[4] V.-N. Huynh and Y. Nakamori, "Behavioral decision analysis using fuzzy targets", *Proceedings of the Fifth International Conference of the Thailand Econometric Society*, Chiang Mai, Thailand, January 12–13, 2012.

[5] V.-N. Huynh, Y. Nakamori, and J. Lawry, "A probability-based approach to comparison of fuzzy numbers and applications to target oriented decision making", *IEEE Transactions on Fuzzy Systems*, 2008, Vol. 16, No. 2, pp. 371–387.

[6] V.-N. Huynh, Y. Nakamori, M. Ryoke, and T. B. Ho, "Decision making under uncertainty with fuzzy targets", *Fuzzy Optimization and Decision Making*, 2007, Vol. 6, No. 3, pp. 255–278.

[7] V.-N. Huynh, H. B. Yan, and Y. Nakamori, "A target-based decision making approach to consumer-oriented evaluation model for Japanese traditional crafts", *IEEE Transactions on Engineering Management*, 2010, Vol. 57, No. 4, pp. 575–588.
[8] E. T. Jaynes, *Probability Theory: The Logic of Science*, Cambridge University Press, Cambridge, Massachusetts, 2003.
[9] R. L. Keeney and H. Raiffa, *Decisions with Multiple Objectives*, John Wiley and Sons, New York, 1976.
[10] G. J. Klir and B. Yuan, *Fuzzy Sets and Fuzzy Logic*, Prentice Hall, Upper Saddle River, New Jersey, 1995.
[11] R. D. Luce and R. Raiffa, *Games and Decisions: Introduction and Critical Survey*, Dover, New York, 1989.
[12] R. E. Moore, R. B. Kearfott, and M. J. Cloud, *Introduction to Interval Analysis*, SIAM Press, Philadelphia, Pennsylvania, 2009.
[13] H. T. Nguyen and E. A. Walker, *First Course In Fuzzy Logic*, CRC Press, Boca Raton, Florida, 2006.
[14] H. Raiffa, *Decision Analysis*, Addison-Wesley, Reading, Massachusetts, 1970.
[15] H.-B. Yan, V.-N. Huynh, and Y. Nakamori, "Target-oriented decision analysis with different target preferences", in V. Torra, Y. Narukawa, and M. Inuiguchi (Eds.), *Proceedings of the The 6th International Conference on Modeling Decisions for Artificial Intelligence MDAI'2009*, Awaji Island, Japan, November 30 – December 2, 2009, Springer Lecture Notes in Computer Science, Vol. 5861, pp. 182–193.
[16] H.-B. Yan, V.-N. Huynh, and Y. Nakamori, "A group nonadditive multi-attribute consumer-oriented Kansei evaluation model with an application to traditional crafts", *Annals of Operations Research*, in press.