How General Is Fuzzy Decision Making?

Olga Kosheleva[a] and **Vladik Kreinovich**[b]

[a]Teacher Education, Univ. of Texas at El Paso, El Paso, TX 79968, USA, `olgak@utep.edu`
[b]Computer Science, Univ. of Texas at El Paso, El Paso, TX 79968, USA, `vladik@utep.edu`

Abstract

In many practical situations, users describe their preferences in imprecise (fuzzy) terms. In such situations, fuzzy techniques are a natural way to describe these preferences in precise terms.

Of course, this description is only an approximation to the ideal decision making – that a person would perform if we took time to elicit his/her exact preferences. How accurate is this approximation? When can fuzzy decision making – potentially – describe the exact decision making, and when there is a limit to the accuracy of fuzzy approximations?

In this paper, we show that decision making can be precisely described in fuzzy terms if and only if different numerical characteristics describing the alternatives are independent – in the sense that if for two alternatives, all but one characteristics have the same value, then the preference between these two alternatives depends only on the differing characteristic and does not depend on the values of all other characteristics.

Keywords: Fuzzy decision making, Traditional decision making, Utility

1 Formulation of the Problem

Decision making in general: a brief reminder. In many practical situations, we make decisions, i.e., we select one of the alternatives. Let us first consider situations in which we have full information about the alternatives. For example, when we buy a house, we know its location, its price, its size, its age, etc.

The information about different alternatives is usually described by numerical values of different characteristics. For example, for a house, we may know its price, the distance to a nearby school, the square footage, etc.

Let us denote the number of such numerical characteristics by n, and the values of these characteristic by $x_1, \ldots, x_n$. In these terms, each alternative can be represented by the corresponding tuple of values

$$x = (x_1, \ldots, x_n).$$

Fuzzy decision making: a brief reminder. In some cases, people have a very clear description of how they want to make a decision. For example, a person with a big family may be interested mostly in the square footage. So, this person may have:

- a fixed amount of money a_1 that he/she is willing to pay for a house – so that the price x_1 does not exceed a_1,

- a fixed maximum distance a_2 from a school, so that the actual distance x_2 is smaller than or equal to a_2, and

- the desired square footage a_3 so that the actual square footage x_3 is larger than or equal to a_3.

Such a person is willing to buy any house $x = (x_1, x_2, x_3)$ for which $x_1 \leq a_1$, $x_2 \leq a_2$, and $x_3 \geq a_3$.

However, such decision makers are rare. Most of the time, when people make decisions, they do not formulate their decision criteria in precise terms. Instead, they formulate them by using imprecise (fuzzy) words from natural language.

For example, when you ask a person looking for a house what exactly he or she wants, this person will probably reply that he/she wants a house which is:

- located in a good neighborhood,

- reasonably large,

- not too expensive,

- not far away from the stores and entertainment district, etc.

All these terms – *good* neighborhood, *reasonably* large, etc. are imprecise.

A natural way to describe these criteria in precise terms is to use fuzzy techniques – techniques designed by Lotfi Zadeh specifically for translating words from natural language into precise, computer-understandable terms; see, e.g., [1, 4, 8, 12, 13, 15].

According to these techniques:

- First, for each characteristic *i*, we design a *membership function* μ_i that assigns, to each possible value x_i of this characteristic, a degree – on the scale from 0 to 1 – to which this value satisfies the decision maker.

- For each alternative $x = (x_1, \ldots, x_n)$, we use the membership functions μ_i to find the degrees $\mu_1(x_1)$, ..., $\mu_n(x_n)$ to which the value of each of *n* characteristics is satisfactory.

- Then, we use an appropriate "and"-operation (t-norm)
$$f_\& : [0,1] \times [0,1] \to [0,1]$$
to estimate the degree $\mu(x)$ to which the value of all *n* characteristics are satisfactory – i.e., to which the entire alternative *x* is satisfactory:
$$\mu(x) = f_\&(\mu_1(x_1), \ldots, \mu_n(x_n)). \qquad (1)$$

- After that, a reasonable idea is to select the alternative *x* for which this overall degree of satisfaction $\mu(x)$ is the largest possible.

This procedure is, of course, an approximation to the ideal exact decision making. Of course, every time we use imprecise words, what we get is an approximate description – in the case of decision making, it is an approximate description of our preferences.

There is a whole science of decision making that described how to elicit exact preferences and make exact decisions – we will recall its main ideas in the next section. Going ahead, according to decision theory, decisions of a rational decision maker are equivalent to maximizing the expected value of a special function *u* – called *utility function* – a function that describes the decision maker's preferences.

How accurately does fuzzy decision making approximate the exact one? A natural question is: how accurately does fuzzy decision making approximate the exact one? When can the actual decision making be approximated by the above-described fuzzy process with any given accuracy, and when it cannot be thus approximated?

This is what we study in this paper. As a result of this study, we provide an answer to this question.

2 Decision Theory: A Brief Reminder

What we do in this section. In this section, we provide a brief description of the traditional decision theory; for details, see, e.g., [2, 3, 6, 7, 9, 10, 14].

The main purpose of decision theory. The main objective of decision theory is to help people make reasonable decisions.

Of course, people have different tastes and different preferences. So, to be able to help a person (or a company) make a reasonable decision, we first need to describe this person's preferences.

How to elicit preferences: main idea. A natural way to elicit information about a person's preferences is:

- to provide this person with several real or hypothetical alternatives, and

- to ask which of these alternatives this person prefers.

For example, we can hypothetically propose to compare two houses on the same location with different prices and different values of square footage.

The idea of a complex alternative – i.e., an alternative with a probabilistic outcome. In addition to well-defined alternatives x, x', x'', etc., we can also ask a person to compare more complex situations, in which this person gets different alternatives with some probabilities.

This is not just a purely mathematical idea, this happens in real life all the time. Let us give a realistic example. Originally, a person had two alternatives:

- a somewhat worse alternative *x* (e.g., a house too far away), and

- a somewhat better alternative x'.

The buyer is finishing negotiations with the owner of the house x'. The buyer is sure that the owner will prefer to sell to this buyer – and not to this buyer's competitor (who is also willing to buy this house). However, the seller of the house x' is in a hurry, so the deal needs to be signed right away.

Suddenly, the buyer learns that a new, even better, house x'' has just appeared on the market. For this house, there are already many bidders, so it is expected to take some time (and some negotiations) for the owner of this new house to decide who to sell it to – and, of course, there is no guarantee that our buyer will get this house.

From the experience of a real estate agent, the buyer knows the probability p of winning the bid for the new house x''. So, the buyer has two choices:

- the buyer can ignore the new house and buy the house x';

- alternatively, the buyer can abandon negotiations about the house x' and bid for the new house; in this case, with probability p, the buyer will win the bid and get the new house x'', and with the remaining probability $1 - p$, the buyer will get the only remaining alternative x.

Such situations are quite realistic, so it makes sense to ask a decision maker:

- not only to compare original alternatives,

- but also to compare such "complex" alternatives – i.e., alternatives with a probabilistic outcome.

How to describe such complex alternatives in precise terms. Let us describe such situations in precise terms.

We have a finite list X of *actual* alternatives – e.g., actual houses that are or can be on the market. We can expand this list by adding finitely many *hypothetical* alternatives – e.g., ideal houses or houses which are real but which we know will not be on the market, such as the White House for those who are looking for a house in Washington, DC. Let us denote this expanded list of alternatives by $\mathscr{X}$.

A complex alternative means that we select some of the alternatives from the set $\mathscr{X}$ with different probabilities. In precise terms, a complex alternative is a probability measure on the set $\mathscr{X}$. In decision theory, such complex alternatives are known as *lotteries* – since in an actual lottery, we have different outcomes with different probabilities. The set of all such lotteries will be denoted by $\mathscr{L}$.

Each alternative x from the extended list $\mathscr{X}$ can be identified with a complex alternative in which:

- this alternative appears with probability 1, and

- all other alternatives have probability 0.

We assume that for every two lotteries L and L' from the set $\mathscr{L}$ (and thus, in particular, for every two alternatives from the set $\mathscr{X}$), the decision maker can decide:

- whether the lottery L is better; we will denote this preference by $L' < L$;

- or the lottery L' is better; we will denote this preference by $L < L'$;

- or the lotteries L and L' are of the same value to the decision maker; we will denote this situation by $L \sim L'$.

We assume that the decision maker is rational. This means, in particular, that:

- if L' is better than L ($L < L'$) and L'' is better than L' ($L' < L''$),

- then L'' should be better than L ($L < L''$).

In mathematical terms, the preference relation must be *transitive*.

Similarly, if $L < L'$ and $L' \sim L''$, then we must have $L < L''$, etc.

A numerical scale for preferences. Real-life alternative are rarely perfectly good, and rarely perfectly bad. As a result, in most practical situations, it is possible to add, to the set X of actual alternatives, the following two hypothetical alternatives:

- an alternative x_- which is worse than anything that we will actually encounter, i.e., in precise terms, worse than all alternatives from the set X; we will call this alternative *very bad*; and

- an alternative x_+ which is better than anything that we will actually encounter, i.e., in precise terms, better than all alternatives from the set X; we will call this alternative *very good*.

Thus, we get the set $\mathscr{X} = X \cup \{x_-, x_+\}$.

For example, in the situation of buying a house:

- the very good alternative x_+ may mean buying the White House, while

- the very bad alternative x_- means staying is the same crammed apartment as before.

(In principle, we can also add other hypothetical alternatives, i.e., get a set $\mathscr{X} \supset X \cup \{x_-, x_+\}$.)

For each real number p from the interval $[0, 1]$, we can consider a lottery in which:

- the probability of the very good alternative x_+ is equal to p,

- the probability of the very bad alternative x_- is equal to $1 - p$, and

- the probability of all other alternatives is 0.

We will denote this lottery by $L(p)$.

Now, we have a continuous family of lotteries characterized by the parameter p. Let us show how this family can be used to provide a numerical value to each alternative from the set X (or, more generally, to each lottery L in which we only alternatives from the set X).

Let $x \in X$ be an alternative. Then, we can compare it with the lotteries $L(p)$ corresponding to different values p. For each p, we have either $L(p) < x$, or $L(p) \sim x$, or $x \sim L(p)$.

For example, for a house described by characteristics $x = (x_1, \ldots, x_n)$ and for $p = 0.4$, we ask a person to compare the following two complex alternatives:

- getting this house x, and

- getting the White House with probability $p = 0.4$ and getting nothing with probability $1 - p = 0.6$.

In general:

- When $p = 0$, then the lottery $L(0)$ coincides with the very bad alternative x_-. Thus, because of our selection of the alternative x_-, we have $x_- < x$, i.e., $L(0) < x$.

- When $p = 1$, then the lottery $L(1)$ coincides with the very good alternative x_+. Thus, because of our selection of the alternative x_+, we have $x < x_+$, i.e., $x < L(1)$.

If we consider two such lotteries $L(p)$ and $L(q)$ with probabilities $p < q$, then clearly a user will prefer the lottery $L(q)$ in which:

- the probability of the very good alternative is higher, and

- the probability of the very bad alternative is lower:

$$L(p) < L(q).$$

Thus:

- if $L(q) < x$ (or $L(q) \sim x$) and $p < q$, then we must have $L(p) < x$; and

- if $x < L(p)$ (or $x \sim L(p)$) and $p < q$, then we must have $x < L(q)$.

One can prove that in this case, the value

$$u(x) \stackrel{\text{def}}{=} \sup\{p : L(p) < x\}$$

is equal to $\inf\{p : x < L(p)\}$, and that:

- for all $p < u(x)$, we have $L(p) < x$, and

- for all $p > u(x)$, we have $x < L(p)$.

This "threshold" probability value $u(x)$ is known as the *utility* of the alternative x.

A mapping that assigns, to some alternatives (actual, hypothetical, and/or complex), the value $u(x)$, is called the *utility function*.

Comment. Strictly speaking, we have defined utility only for alternatives from the set X. However, a similar definition can be stated for complex alternatives L formed by actual alternatives (from the set X), i.e., in mathematical terms, for probability distributions on the set X.

Practical indistinguishability. From the purely mathematical viewpoint, lotteries $L(p)$ and $L(q)$ corresponding to different values $p \neq q$ are different, even when these values are very close: e.g., when $p = 0.5$ and $q = 0.500001$.

However, in practice, we will probably not notice this difference – and if we *can* notice this difference, we will not notice the difference between $p = 0.5$ and $q = 0.5 + \varepsilon$ for some positive value $\varepsilon \ll 0.000001$. Indeed:

- Probabilities provided by the real estate agent are based on a finite sample of cases and are, thus, approximate.

- Even if we decide to use an actual fair coin to implement this lottery, this coin may be slightly flawed and thus, its probability may be somewhat different from 0.5.

In general, for sufficiently small $\varepsilon > 0$, we do not feel the difference between lotteries corresponding to probabilities p, $p - \varepsilon$, and $p + \varepsilon$. By definition of the utility $u(x)$, for each $\varepsilon > 0$, we have

$$L(u(x) - \varepsilon) < x < L(u(x) + \varepsilon).$$

Thus, we can say that from the practical viewpoint, the alternative x is actually equivalent to the lottery $L(u(x))$. We will denote this by $x \equiv L(u(x))$.

Which alternative should we select? As we have mentioned, each alternative x is equivalent to the lottery $L(u(x))$, where $u(x)$ is the utility of this alternative. Thus, comparing alternatives is equivalent to comparing the corresponding lotteries $L(u(x))$.

We have also mentioned that when we compare several lotteries $L(p)$, $L(q)$, ..., then the larger the probability of the very good alternative x_+, the better. Thus, we have to select the alternative x for which the utility value $u(x)$ is the largest.

Why expected utility. As we will show, one of the consequences of the above definition of utility is that in the case of uncertainty, we need to maximize the expected utility. Some folks – who are not very familiar with decision theory – mistakenly think that the maximization of expected utility is an additional (and not-well-justified) postulate, but it is not, it is a consequence of utility's definition.

Indeed, suppose that we have a complex alternative (lottery) L in which we get:

- an alternative $x^{(1)} \in X$ with probability p_1,

- an alternative $x^{(2)} \in X$ with probability p_2,

- ..., and

- an alternative $x^{(m)} \in X$ with probability p_m,

and all other alternatives with probability 0.

What is the utility $u(L)$ of this lottery? As we have mentioned, each alternative $x^{(i)}$ is equivalent to a lottery $L(u(x^{(i)}))$ in which we get:

- the very good alternative x_+ with probability equal to the utility $u(x^{(i)})$ of this alternative, and

- the very bad alternative x_- with the remaining probability $1 - u(x^{(i)})$.

So, if in the above lottery L, we replace each alternative $x^{(i)}$ with the equivalent lottery $L(u(x^{(i)}))$, we conclude that the lottery L is equivalent to the following 2-stage lottery:

- first, we select one of the alternatives $x^{(1)}$, ..., $x^{(m)}$, so that each alternative $x^{(i)}$ is selected with probability p_i, and

- then, depending on which alternative $x^{(i)}$ we selected, we select x_+ with probability $u(x^{(i)})$ and x_- with probability $1 - u(x^{(i)})$.

As a result of this 2-stage lottery, we get either x_+ or x_-. By considering all m possible ways to get x_+, in each of which we get x_+ with the probability $p_i \cdot u(x^{(i)})$, we conclude that the overall probability of selecting x_+ is equal to the sum of these values, i.e., to the sum

$$s = p_1 \cdot u(x^{(1)}) + \ldots + p_m \cdot u(x^{(m)}). \qquad (2)$$

Thus, the lottery L is equivalent to the lottery in which:

- we get x_+ with probability s,

- we get x_- with the remaining probability $1 - s$,

- and all other alternatives have probability 0.

By definition of utility, this means that the lottery L has utility $u(L) = s$.

And it so happens that the expression (2) – that describes this utility – is actually the expected value of the utility function $u(x)$. So, the principle of maximizing the expected utility indeed follows from the definition of utility.

3 So When Can Exact Decision Be Described in Fuzzy Terms?

The problem: reminder. Now that we recalled the traditional decision theory, let us go back to the original question.

In the general decision theory, when we select between alternative x characterized by values $x_1, \ldots, x_n$, the recommendation is to select the alternative x^{opt} for which the utility $u(x^{\text{opt}}) = u(x_1^{\text{opt}}, \ldots, x_n^{\text{opt}})$ is the largest possible.

In fuzzy decisions, we select the alternative for which the expression (1) attains its largest possible value. So, the question is: which preference relations (i.e., which utility functions) can be represented as maximizing the expression (1) for some membership functions and for some t-norm?

Let us reformulate fuzzy decision making in utility terms. To be able to compare the two approaches, let us perform some reformulations. Specifically, we reformulate the fuzzy decision making in terms which are closer to utilities.

To do this, we can take into account that:

- while, in principle, there exist many different "and"-operations (t-norms),

- it is known (see, e.g., [11]) that for every t-norm $f_\&$ and for every $\varepsilon > 0$, there exists a strictly increasing function $f : [0,1] \to [0,1]$ for which

for all a and b, we have $|f_\&(a,b) - g(a,b)| \leq \varepsilon$, where

$$g(a,b) \stackrel{\text{def}}{=} f^{-1}(f(a) \cdot f(b)), \qquad (3)$$

and f^{-1} denotes the inverse function.

From the practical viewpoint, sufficiently close degrees of certainty are practically indistinguishable. Indeed:

- If we are asked to mark our degree of confidence on a scale from 0 to 5, we can definitely meaningfully distinguish between the value 0.6 (corresponding to selecting 3 on a 0 to 5 scale) and the value 0.8 (corresponding to selecting 4 on this scale).

- However, when a person is asked to mark his/her degree of confidence on a scale from 0 to 100, hardly anyone can distinguish between, e.g., marks 80 and 81 – which correspond to degree 0.80 and 0.81.

From this viewpoint, for sufficiently small ε – e.g., for $\varepsilon \leq 0.01$ – there is no meaningful difference between the degrees $f_\&(a,b)$ and $g(a,b)$.

Thus, from the practical viewpoint, we can safely assume that the t-norm actually has the form (3).

For such a t-norm, the formula (1) turns into

$$g(\mu_1(x_1), \ldots, \mu_n(x_n)) =$$
$$f^{-1}(f(\mu_1(x_1)) \cdot \ldots \cdot f(\mu_n(x_n))). \qquad (4)$$

Since the function $f(x)$ is strictly increasing, maximizing the expression (4) is equivalent to maximizing the result of applying the function f to this value, i.e., to maximizing the expression

$$f(\mu_1(x_1)) \cdot \ldots \cdot f(\mu_n(x_n)). \qquad (5)$$

Since logarithm is also a strictly increasing function, maximizing the product (5) is equivalent to maximizing its logarithm. Since the logarithm of the product is equal to the sum of the logarithms, the problem of maximizing the logarithm of the expression (5) can be described as the problem of maximizing the sum

$$v_1(x_1) + \ldots + v_n(x_n), \qquad (6)$$

where we denoted $v_i(x_i) \stackrel{\text{def}}{=} \ln(f(\mu_i(x_i)))$.

Vice versa, if our decision problem can be described in the form (6), we can take, e.g., $f(x) = x$ (then $f_\&(a,b) = a \cdot b$), and $\mu_i(x_i) = C_i \cdot \exp(v_i(x_i))$, for some normalization constant C_i (such a constant is needed to

make sure that all the values of the resulting membership function do not exceed 1). One can easily see that for these "and"-operation and membership functions, maximizing the expression (1) is indeed equivalent to maximizing the formula (6).

Now the problem has been reformulated, so we can answer the original question. Now, the original problem – when decision making can be described in fuzzy terms – has been reformulated in precise terms:

When is a decision problem characterized by a utility function $u(x_1, \ldots, x_n)$ equivalent to maximizing the sum (6) – i.e., the sum of n expressions each of which depends only on one of the characteristics x_i describing the alternative?

Interestingly, this problem has already been solved in utility theory; see, e.g., [3, 5]. Namely, one can easily check that if our decision making is equivalent to maximizing the expression (6), this means that for us, the n characteristics are *independent* in the following sense:

- if we have two alternatives differing only by the values $x_i \neq x_i'$ of the i-th characteristic,

- then which of these two alternative is better depends only on the values x_i and x_i' and does not depend on the values of the other characteristics.

In precise terms:

- if
$$(x_1, \ldots, x_{i-1}, x_i, x_{i+1}, \ldots, x_n) <$$
$$(x_1, \ldots, x_{i-1}, x_i', x_{i+1}, \ldots, x_n)$$
for some values
$$x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n,$$

- then for any other values
$$x_1', \ldots, x_{i-1}', x_{i+1}', \ldots, x_n',$$
we will have a similar relation:
$$(x_1', \ldots, x_{i-1}', x_i, x_{i+1}', \ldots, x_n') <$$
$$(x_1', \ldots, x_{i-1}', x_i', x_{i+1}', \ldots, x_n').$$

This is true not only when we compare original alternatives, the same property holds if we consider complex alternatives (lotteries).

It has been proven that this independence property uniquely characterizes the possibility of representation (6); namely:

- if the above independence property holds,

- then the maximized utility can be represented as the sum (6) of the terms $v_i(x_i)$ each of which depends only on the corresponding characteristic.

So, we get the following answer to our questions.

4   Conclusion

We consider decision making problems, in which we compare alternatives x characterized by the values of several characteristics $x_1, \ldots, x_n$. We show that:

- such a decision problem can be represented in the equivalent form (1) corresponding to fuzzy decision making

- if and only if these characteristics are independent – in the formal sense described in the previous section.

Crudely speaking, independence means that:

- when for two alternatives, all characteristics but one have equal values,

- then our preference between these two alternatives depends only on the values of the differing characteristic and does not depend on the values of all other characteristics.

So:

- If this independence condition is satisfied, then fuzzy decision making can approximate the actual decision making as accurately as we want – in can even exactly represent the actual decision making.

- On the other hand, if the independence condition is not satisfied, then the decision making cannot be exactly represented in the fuzzy form – and thus, there is a limit on how accurately fuzzy decision making can approximate this decision making.

References

[1] R. Belohlavek, J. W. Dauben, G. J. Klir, Fuzzy Logic and Mathematics: A Historical Perspective, Oxford University Press, New York, 2017.

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

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

[4] G. Klir, B. Yuan, Fuzzy Sets and Fuzzy Logic, Prentice Hall, Upper Saddle River, New Jersey, 1995.

[5] O. Kosheleva, Symmetry-group justification of maximum entropy method and generalized maximum entropy methods in image processing, in: G. J. Erickson, J. T. Rychert, C. R. Smith (Eds.), Maximum Entropy and Bayesian Methods, Kluwer, Dordrecht, 1998, pp. 101–113.

[6] V. Kreinovich, Decision making under interval uncertainty (and beyond)", in: P. Guo, W. Pedrycz (Eds.), Human-centric decision-making models for social sciences. Springer Verlag, 2014, pp. 163–193

[7] R. D. Luce, R. Raiffa, Games and Decisions: Introduction and Critical Survey. Dover, New York, 1989.

[8] J. M. Mendel, Uncertain Rule-Based Fuzzy Systems: Introduction and New Directions, Springer, Cham, Switzerland, 2017.

[9] H. T. Nguyen, O. Kosheleva, V. Kreinovich, Decision making beyond Arrow's 'impossibility theorem', with the analysis of effects of collusion and mutual attraction, International Journal of Intelligent Systems, 2009, Vol. 24, No. 1, pp. 27–47.

[10] H. T. Nguyen, V. Kreinovich, B. Wu, G. Xiang, Computing Statistics under Interval and Fuzzy Uncertainty. Springer Verlag, Berlin, Heidelberg, 2012.

[11] H. T. Nguyen, V. Kreinovich, P. Wojciechowski, Strict Archimedean t-norms and t-conorms as universal approximators, International Journal of Approximate Reasoning, 1998, Vol. 18, No. 3–4, pp. 239–249.

[12] H. T. Nguyen, C. L. Walker, E. A. Walker, A First Course in Fuzzy Logic, Chapman and Hall/CRC, Boca Raton, Florida, 2019.

[13] V. Novák, I. Perfilieva, J. Močkoř, Mathematical Principles of Fuzzy Logic, Kluwer, Boston, Dordrecht, 1999.

[14] H. Raiffa, Decision Analysis. McGraw-Hill, Columbus, Ohio, 1997.

[15] L. A. Zadeh, Fuzzy sets, Information and Control, 1965, Vol. 8, pp. 338–353.