How General Is Fuzzy Decision Making?

Olga Kosheleva and Vladik Kreinovich

University of Texas at El Paso 500 W. Universty El Paso, Texas 79968, USA olgak@utep.edu, vladik@utep.edu



1. 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 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.



2. Decision making in general (cont-d)

• 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

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



3. 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 .



4. Fuzzy decision making (cont-d)

- Such a person is willing to buy any house $x = (x_1, x_2, x_3)$ for which $x_1 \le a_1, x_2 \le a_2$, and $x_3 \ge a_3$.
- 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.
- A person looking for a house will probably say 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.



5. Fuzzy decision making (cont-d)

- 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*.
- These techniques were designed by Lotfi Zadeh specifically for translating:
 - words from natural language
 - into precise, computer-understandable terms.
- 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.



6. Fuzzy decision making (cont-d)

- For each alternative $x = (x_1, \ldots, x_n)$:
 - we use the membership functions μ_i
 - to find the degrees $\mu_1(x_1), \ldots, \mu_n(x_n)$ 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:

$$\mu(x) = f_{\&}(\mu_1(x_1), \dots, \mu_n(x_n)).$$

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



7. 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 describes:
 - how to elicit exact preferences and
 - how to make exact decisions.
- A natural question is: how accurately does fuzzy decision making approximate the exact one?
- When can the actual decision making be approximated by a fuzzy process with any given accuracy?
- When it cannot be thus approximated?



8. Decision theory: a brief reminder

- 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.



9. 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.



10. The idea of a complex alternative – i.e., an alternative with a probabilistic outcome

- 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 for the owner of this new house to decide whom to sell,
 - 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 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 the second 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.



14. How to describe such complex alternatives 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.
- Example:
 - ideal houses or
 - houses which are real but which we know will not be on the market, such as the White House.
- ullet Let us denote this expanded list of alternatives by \mathcal{X} .
- A complex alternative means that we select some of the alternatives from the set \mathcal{X} with different probabilities.
- In precise terms, a complex alternative is a probability measure on the set \mathcal{X} .



15. Describing complex alternatives (cont-d)

- In decision theory, such complex alternatives are known as *lotteries*.
- The set of all such lotteries will be denoted by \mathcal{L} .
- Each alternative x from the extended list \mathcal{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', 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.



18. 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., worse than all $x \in X$.
- We will call this alternative very bad.
- An alternative x_+ which is better than anything that we will actually encounter, i.e., better than all $x \in X$.
- We will call this alternative very good.
- Thus, we get the set $\mathcal{X} = X \cup \{x_-, x_+\}$.



- 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 $\mathcal{X} \supset X \cup \{x_-, x_+\}$.)
- 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.
- More generally, we can provide a value to each lottery L in which we only alternatives from 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 $x \sim L(p)$.
- Example: a house described by characteristics $x = (x_1, \ldots, x_n), p = 0.4.$



- Then, 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.
- 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).



- Let us 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 $u(x) \stackrel{\text{def}}{=} \sup\{p : L(p) < x\}$ 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.
- By a *utility function*, we mean a mapping that assigns:
 - to some alternatives (actual, hypothetical, and/or complex),
 - the value u(x).



- 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 X),
 - i.e., in mathematical terms, for probability distributions on the set X.



25. 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.
- \bullet Some folks may notice this difference.
- But they will not notice the difference between p = 0.5 and $q = 0.5 + \varepsilon$ for some positive value $\varepsilon \ll 0.000001$.
- There are two reasons for this.



26. Practical indistinguishability (cont-d)

- Probabilities provided by the real estate agent:
 - 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.
- 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 $p + \varepsilon$.



27. Practical indistinguishability (cont-d)

- By definition of the utility u(x), for each $\varepsilon > 0$, we have $L(u(x) \varepsilon) < x < L(u(x) + \varepsilon)$.
- Thus, from the practical viewpoint, the alternative x is actually equivalent to the lottery L(u(x)).
- We will denote this by $x \equiv L(u(x))$.



28. 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 $L(p), L(q), \ldots$,
 - 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.



29. 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.



30. Why expected utility (cont-d)

- Indeed, suppose that we have a complex alternative (lottery) L in which we get:
 - an alternative $x^{(1)} \in X$ with probability $p_1, \ldots,$
 - 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 $1 u(x^{(i)})$.
- In the above lottery L, we can replace each alternative $x^{(i)}$ with the equivalent lottery $L(u(x^{(i)}))$.



31. Why expected utility (cont-d)

- We conclude that the lottery L is equivalent to the following 2-stage lottery:
 - first, we select one of the alternatives $x^{(1)}, \ldots, x^{(m)}$, so that each $x^{(i)}$ is selected with prob. p_i .
 - then, depending on which alternative $x^{(i)}$ we selected, we select x_+ with prob. $u(x^{(i)})$ or x_- .
- As a result of this 2-stage lottery, we get either x_+ or x_- .
- There are m possible ways to get x_+ , in each of which we get x_+ with the probability $p_i \cdot u(x^{(i)})$.
- So, the overall probability of selecting x_+ is equal to the sum of these values

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



32. Why expected utility (cont-d)

- Thus, the lottery L is equivalent to the lottery L(s) in which:
 - we get x_+ with probability s,
 - we get x_{-} with the 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.
- It so happens that the expression for s 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.



33. So when can exact decision be described in fuzzy terms?

- 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 ,
 - we select the alternative x^{opt} for which the utility $u(x^{\text{opt}}) = u(x_1^{\text{opt}}, \dots, x_n^{\text{opt}})$ is the largest.
- In fuzzy decisions, we select the alternative for which the expression $\mu(x)$ attains its largest possible value:

$$\mu(x) = f_{\&}(\mu_1(x_1), \dots, \mu_n(x_n)).$$

• Which preference relations can be represented as maximizing the expression $\mu(x)$ for some μ_i and $f_{\&}$?



34. 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.
- In principle, there exist many different "and"-operations (t-norms).
- It is known that for every t-norm $f_{\&}$, 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 $g(a,b) \stackrel{\text{def}}{=} f^{-1}(f(a) \cdot f(b))$.



- From the practical viewpoint, sufficiently close degrees of certainty are practically indistinguishable.
- Let us assume that 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).
- Suppose now that a person 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.



- So, for small $\varepsilon > 0$ e.g., for $\varepsilon \le 0.01$ there is no 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

$$g(a,b)=f^{-1}(f(a)\cdot f(b)).$$

- For such a t-norm, the formula for $\mu(x)$ turns into $g(\mu_1(x_1), \dots, \mu_n(x_n)) = f^{-1}(f(\mu_1(x_1)) \cdot \dots \cdot f(\mu_n(x_n))).$
- The function f(x) is strictly increasing.
- So, maximizing this expression is equivalent to maximizing the result of applying f to this value:

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

• Logarithm is also a strictly increasing function.

Decision making in . . . Fuzzy decision . . . Decision Theory: A... A numerical scale for . . . Practical . . . Why expected utility So When Can Exact . . . Home Page Title Page **>>** Page 37 of 47 Go Back Full Screen Close Quit

- So, maximizing this product is equivalent to maximizing its logarithm.
- The logarithm of the product is equal to the sum of the logarithms:

$$v_1(x_1) + \ldots + v_n(x_n)$$
, where $v_i(x_i) \stackrel{\text{def}}{=} \ln(f(\mu_i(x_i)))$.

- Vice versa:
 - if our decision problem can be described in this form,
 - 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 C_i ;
 - $-C_i$ 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 $\mu(x)$ is indeed equivalent to maximizing the sum.



39. Now the problem has been reformulated, so we can answer the original question

- The original problem was: when decision making can be described in fuzzy terms?
- Now we reformulate it in precise terms.
- When is:
 - a decision problem characterized by a utility function $u(x_1, \ldots, x_n)$
 - equivalent to maximizing the sum

$$v_1(x_1) + \ldots + v_n(x_n)?$$

• Interestingly, this problem has already been solved in utility theory.



40. We can answer the original question (cont-d)

- Namely, one can easily check that:
 - if our decision making is equivalent to maximizing the sum,
 - this means that for us, the n characteristics are independent in the following sense.
- Suppose that 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.



41. We can answer the original question (cont-d)

- In precise terms:
 - if for some values $x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n$:

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

- 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).

Decision making in . . . Fuzzy decision . . . Decision Theory: A... A numerical scale for . . . Practical . . . Why expected utility So When Can Exact . . . Home Page Title Page **>>** Page 42 of 47 Go Back Full Screen Close Quit

42. We can answer the original question (cont-d)

- It has been proven that:
 - this independence property
 - uniquely characterizes the possibility of representation as the sum.
- So, we get the following answer to our questions.



43. 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 fuzzy form
 - if and only if these characteristics are independent
 - in the above formal sense.



44. Conclusion (cont-d)

- Independence means that when for two alternatives, all characteristics but one have equal values, then:
 - our preference depends only on the values of the differing characteristic and
 - does not depend on the values of all other characteristics.
- If this independence condition is satisfied, then fuzzy decision making:
 - can approximate the actual decision making as accurately as we want
 - it can even exactly represent the actual decision making.



45. Conclusion (cont-d)

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



46. Acknowledgment

This work was supported in part by the 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).

It was also supported by the program of the development of the Scientific-Educational Mathematical Center of Volga Federal District No. 075-02-2020-1478.

