Fuzzy Sets Can Be
Interpreted as Limits of Crisp Sets,
and This Can Help to Fuzzify Crisp Notions

Olga Kosheleva[1], Vladik Kreinovich[1], and Thavatchai Ngamsantivong[2]

[1]University of Texas at El Paso, El Paso, TX 79968, USA,
`olgak@utep.edu`, `vladik@utep.edu`
[2]Computer and Information Science, Faculty of Applied Sciences,
King Mongkut's University of Technology North Bangkok,
Bangkok 10800 Thailand, `tvc@kmutnb.ac.th`

Abstract. Fuzzy sets have been originally introduced as *generalizations* of crisp sets, and this is how they are usually considered. From the mathematical viewpoint, the problem with this approach is that most notions allow many different generalizations, so every time we try to generalize some notions to fuzzy sets, we have numerous alternatives. In this paper, we show that fuzzy sets can be alternatively viewed as *limits* of crisp sets. As a result, for some notions, we can come up with a unique generalization – as the limit of the results of applying this notion to the corresponding crisp sets.

1   Formulation of the Problem: Too Many Different Fuzzifications

Crisp sets: brief reminder. Many properties are well-defined and objective ("crisp"). For example, a real number x is either positive or not positive, it is either smaller than 10 or not, etc. Each such crisp property can be described by a (crisp) *set* – namely, by the set S of all the objects that satisfy this property.

For each such set S and for each object x, either the object x belongs to the set S ($x \in S$) or the object x does not belong to the set S ($x \notin S$). A set S can therefore be equivalently described by its *characteristic function* $\mu_S(x)$ that assigns, to each object x, the truth value of the statement $x \in S$. In other words:

- If $x \in S$, i.e., if the object x satisfies the property defining the set, then $\mu_S(x) = 1$.
- On the other hand, if $x \notin S$, i.e., if the object x does not satisfy the desired property, then we take $\mu_S(x) = 0$.

Need for fuzzy sets. Humans routinely deal with properties which are not fully well-defined and not fully objective, such as "small", "young", etc. A large portion of our knowledge, of our experience, is described in terms of such properties. To deal with such imprecise ("fuzzy") properties, L. Zadeh introduced the notion of a *fuzzy set*; see, e.g., [1–3]. The main idea behind fuzzy sets is that for

fuzzy notions S, we no longer have a clear division into objects which absolutely satisfy this notion and objects which absolutely do not satisfy this notion.

For a crisp property like "positive", if we continuously increase a number x from negative values to positive ones, we first have numbers which are absolutely not positive, and then, at $x = 0$, we abruptly switch to numbers which are absolutely positive. In contrast, for a property like "small", as we increase values x from small to not small, we do not abruptly switch from small numbers to non-small ones, the transition is continuous. When the value x is very small, this value will be classified by everyone as absolutely small. Similarly, a very large value will be classified by everyone as absolutely not small. However, for intermediate values, we may differ whether this value is small or nor, and even a single person can hesitate.

To capture this phenomenon of a "smooth" transition between true and false, Zadeh decided to use values between 0 ("false") and 1 ("true") to describe the intermediate states of our beliefs. As a result, a *fuzzy set S* can be mathematically defined as a function $\mu_S(x)$ that assigns, to each possible object x into a number $\mu_S(x)$ from the interval $[0, 1]$.

Fuzzy sets have been very successful. Zadeh's idea of capturing the fuzziness of human reasoning has led to numerous successful applications, in control, in clustering, etc.; see, e.g., [1, 2].

Important problem: how to fuzzify? It is not always easy to apply fuzzy techniques. One of the reasons for this is that there are many alternative fuzzy techniques, and it is not clear which of these techniques we should use.

The reason for this variety of techniques is that, as defined above, the concept of a fuzzy set is a *generalization* of the concept of a crisp set. A crisp set can be defined as function that assigns, to each object x, a value from the 2-element set $\{0, 1\}$, while a fuzzy set is defined as a function that assigns to each x an element of the more general set $[0, 1]$. To extend a concept from crisp sets to fuzzy sets means that we need to extend an operation defined for two truth value 0 and 1 to all possible intermediate values.

This happens, e.g., when we define a complement to a set. In the classical case, this is easy: if an element x belongs to the set S, this element does not belong to the complement $-S$, and vice versa. In other words, if $\chi_S(x) = 1$, then $\chi_{-S}(x) = 0$, and if $\chi_S(x) = 0$, then $\chi_{-S}(x) = 1$. So, a complement is described by an operation that maps 0 to 1 and 1 to 0. We would like to generalize this operation to general fuzzy sets, i.e., to general numbers from the interval $[0, 1]$. It is clear that there are many different ways to extend an operation.

That is why in fuzzy applications, there are many generalizations of negation, many generalization of the union (different t-conorms) and of the intersection (different t-norms), different definitions of probability of a fuzzy event, etc. Which of these generalization should we choose? This is often not clear.

What we do in this paper. In this paper, we show that fuzzy sets can be also naturally interpreted as *limits* of crisp sets. Mathematically, this equivalent reformulation in terms of limits is equivalent to the original definition.

Computationally, the reformulation in terms of limits may even be much worse that the original definition, since we are replacing a reasonably simple notion of a function (from the set X of all objects to the interval $[0, 1]$), a notion routinely studied in high school, with a much more complex concept of limit of sets, a concept that is only studied by professional mathematicians.

However, from the viewpoint of generalizations, the limit definition has a clear advantage: once a fuzzy set S is represented as a limit of a sequence of crisp sets S_1, S_2, ..., then we can define, e.g., the probability of this fuzzy set as a limit of the probabilities of the corresponding crisp sets.

We will show that in some cases, this idea indeed enables us to select one definition among many. We will also show that this idea is not a panacea: sometimes, when the sequence S_n tends to S, the corresponding values do not tend to a limit; in such situations, we still have to choose an appropriate fuzzification.

We hope that, in addition to the above-described pragmatic use of this idea, it will also lead to an even wider acceptance of fuzzy set techniques: one may be reluctant to use generalizations, but it is a natural idea to use limits: this is why we routinely use real numbers which are, in effect, nothing else but limits of directly observable rational numbers; that is why we routinely use infinities which are nothing else but limits of real numbers, etc.

2 Polling Interpretation of Fuzzy Properties Naturally Leads to Fuzzy Sets as Limits of Crisp Sets

Polling interpretation of fuzzy properties: reminder. One of the standard ways to elicit the membership degrees $\mu_S(x)$ is by polling. We ask several (N) experts whether they would agree that an object x satisfies the corresponding property (e.g., whether the given number is small). If M out of N folks say that the given value x is small, we take the ratio $\dfrac{M}{N}$ as the desired degree $\mu_S(x)$.

Polling: logistic challenges. When a variable x only takes finitely many possible values, we can simply repeat the above procedure for all these values, and get all the desired degrees $\mu_S(x)$, i.e., get the full description of the corresponding fuzzy set. In reality, there are infinitely many possible values of each quantity x. The first challenge us that we can only ask folks about finitely many different values x.

The second challenge is that it is difficult for many people to think in abstract terms. When I see a person, I can tell whether, in my opinion the person is short or not; when I experience a certain temperature, I can tell whether this temperature corresponds to "warm" or not. However, if for me, a fuzzy threshold between short and non-short lies around 170 cm, and none of my friends are of this threshold height, it will be difficult for me to decide whether someone of

height exactly 170 cm is short or not, without actually observing such a person. Because of this, not only we are limited to finitely many possible values x, but we are also limited to values x corresponding to actual objects shown to the polled experts.

This leads to the third challenge: that we may be polling, e.g., medical doctors or geoscientists located at different parts of the word. It would be too expensive to fly the same patients to all the medical doctors, or to fly all geoscientists to the same earth formation. Realistically, each expert deals with his/her own values x, and our goal is to combine this data.

What is the direct result of polling. As a result of polling, we get a finite collection C of values x, some of which are marked, by an expert, as having the property S while others are marked as not having the property S. Let $C^+ \subseteq C$ denote the collection of all the values that experts marked as satisfying the desired property.

We are dealing with real-life objects. For two real-life objects, the probability that they have the exact same value of some quantity x is 0. We can therefore safely assume that all these values are different. For example, if we pick two rocks, it is highly improbable that they will have the exact same weight; their weights may be close, but they cannot be exactly equal.

How polling techniques process this information. Based on the given set of samples marked by experts, how can we estimate the degrees $\mu_S(x)$ corresponding to different values x? For each value x, it is highly improbable that one of the experts actually dealt with this very value; we can only hope to find close values for which an expert has expressed his or her opinion. So, we take a neighborhood $(x - \varepsilon, x + \varepsilon)$ of the desired value x and, in this neighborhood, count the proportion of points which were marked by an expert as satisfying the property S. In other words, as an estimate for $\mu_S(x)$, we take the ratio

$$\mu_S(x) \approx \frac{\#(C^+ \cap (x - \varepsilon, x + \varepsilon))}{\#(C \cap (x - \varepsilon, x + \varepsilon))}, \tag{1}$$

where $\#(s)$ denote the number of elements in a set s.

To get a more accurate estimate, we need to elicit more opinions from the experts, which would enable us to get more points in the set C. This means that we are not just dealing with a single set C of such values, we get an increasing sequence of finite sets C_1, C_2, ..., C_n, ... corresponding to increasing number of points ($\#(C_n) \to +\infty$), so that

$$\mu_S(x) \approx \frac{\#(C_n^+ \cap (x - \varepsilon, x + \varepsilon))}{\#(C_n \cap (x - \varepsilon, x + \varepsilon))} \tag{2}$$

for all n.

The more points, the more accurate the estimate; thus, the most accurate estimate corresponds to the limit $n \to +\infty$:

$$\mu_S(x) \approx \lim_{n \to +\infty} \frac{\#(C_n^+ \cap (x - \varepsilon, x + \varepsilon))}{\#(C_n \cap (x - \varepsilon, x + \varepsilon))}. \tag{3}$$

This limit, however, is not yet an exact value of $\mu_S(x)$, since this limit represents not just a single value x but the whole interval $(x - \varepsilon, x + \varepsilon)$. To get the exact value of $\mu_S(x)$, we therefore need to perform one more transition to a limit: by taking $\varepsilon \to 0$. After that, we should be able to get the exact value of the membership degree:

$$\mu_S(x) = \lim_{\varepsilon \to 0} \lim_{n \to +\infty} \frac{\#(C_n^+ \cap (x - \varepsilon, x + \varepsilon))}{\#(C_n \cap (x - \varepsilon, x + \varepsilon))}. \tag{4}$$

Let us show how this procedure can lead to a limit interpretation of fuzzy sets.

From finite lists C_n to crisp sets S_n: crisp case. Let us start our analysis of the situation with the case when the desired property is crisp. Suppose that we have a finite set C_n of values for each of which experts decided whether this value satisfies the desired property or not. For each given value x, how can we then decide whether the given value x satisfied the desired property?

A natural idea, as we have mentioned, is to check whether some value which is close to x have been classified by experts. The closer this already-classified value to x, the more confident we are that this element and the desired value x both satisfy or both do not satisfy the property S. Thus, a reasonable idea is to look for the element from C_n which is the closest to x:

- if this closest element satisfies the property S, then we conclude that the given value x also satisfies the property S;
- is this closest element does not satisfy the property S, then we conclude that the given value x does not satisfy S either.

Of course, there will be few threshold cases when the value x is exactly in between two values, one classifies as satisfying S and another classified as not satisfying S, but these values are rare, so we can arbitrarily classify them to S or to a complement to S.

Thus, for each n, we divide the set X of all possible values of x into two sets:

- the set of all the values x which are, based on the set C_n, classified as satisfying the property S; we will denote this set by S_n; and
- the set of all the values x which are, based on the set C_n, classified as not satisfying the property S; this second set is simply a complement $X - S_n$.

As we elicit more and more opinions from experts, we get sets C_n which have more and more points; moreover, we get more and more points within each interval. So in the limit, when we increase n, the corresponding sets S_n and $X - S_n$ becomes closer and closer to the actual sets S and $X - S$, in the sense that:

- if the value x actually satisfies the property S, i.e., if $x \in S$, then most probably, starting with some sufficiently large n, it will be recognized by this procedure as having this property;

– similarly, if the value x actually does not satisfy the property S, i.e., if $x \notin S$, then most probably, starting with some sufficiently large n, it will be recognized by this procedure as not having this property.

In other words, if we form the values $\chi_{S_n}(x)$, then for sufficiently large n, these values will coincide with $\chi(S)$, i.e., we will have $\chi_S(x) = \lim\limits_{n \to +\infty} \chi_{S_n}(x)$.

From finite lists C_n to crisp sets S_n: general (fuzzy) case. What happens in the general (fuzzy) case? In this case, based on each set of observations C_n, we can also subdivide the entire set X into two crisp subsets:

– the set S_n of all the values x for which the closest point from C_n is classified as having the property S, and
– the set $X - S_n$ of all the values x for which the closest point from C_n is classified as not having the property S.

So far, the description is similar to the corresponding description of the crisp case. The difference is what happens in the intermediate values x for which the experts differ. For such intermediate values, if we start with a randomly selected collection of values around x, out of which experts classify a proportion $\mu_S(x)$ as satisfying the property S, then, as one can easily check, the proportion of points assigned to the set S_n will also be approximately the same. In other words, we will have

$$\mu_S(x) \approx \frac{\text{len}(S_n \cap (x - \varepsilon, x + \varepsilon))}{\text{len}(x - \varepsilon, x + \varepsilon)}, \tag{5}$$

where $\text{len}(s)$ denotes the total length of the set S:

– for an interval, it is exactly its length;
– for a union of several disjoint intervals, it is the sum of their lengths.

Similarly to the above formulas (3)–(4), to get an accurate value $\mu_S(x)$, we need to take more and more points n and narrower and narrower interval $(x - \varepsilon, x + \varepsilon)$. Then, we get

$$\mu_S(x) = \lim_{\varepsilon \to 0} \lim_{n \to +\infty} \frac{\text{len}(S_n \cap (x - \varepsilon, x + \varepsilon))}{\text{len}(x - \varepsilon, x + \varepsilon)}. \tag{6}$$

Resulting idea: fuzzy set as a limit of crisp sets. Similarly to describing a crisp set S as a limit of the corresponding crisp sets S_n, we can thus formally describe a fuzzy set as a limit of crisp sets if the formula (6) is satisfied.

We can now define operations on fuzzy sets as limits of operations on the corresponding sets S_n – when such a limit exists. Let us describe this idea in precise terms.

3 Fuzzy Sets as Limits of Crisp Sets: Definitions and Results

Definition 1. *We say that a fuzzy set S with a membership function $\mu_S(x)$ is a limit of a sequence of (crisp) sets S_n, and denote it as $S_n \to S$, if for every $x \in X$, the formula (6) holds.*

Comment. A similar definition can be formulated for fuzzy subsets of a plane, a 3-D space, or, more generally, a multi-D space; in this case:

– to describe a neighborhood of a point $x = (x_1, \ldots, x_d)$, it is reasonable to use, e.g., boxes

$$(x_1 - \varepsilon_1, x_1 + \varepsilon_1) \times \ldots \times (x_d - \varepsilon_d, x_d + \varepsilon_d)$$

instead of intervals;
– instead of a length of a set, we need to use a more general Lebesgue measure; e.g., area for sets in a plane, volume for 3-D sets. etc.

Let us first show that the above intuitive idea indeed works, i.e., that generic fuzzy sets can indeed be represented as limits of crisp sets.

Proposition 1. *Every fuzzy set with a continuous membership function can be represented as a limit of crisp sets.*

Proof. One way to describe the corresponding set S_n is to divide the real axis into intervals $\left[\dfrac{k}{n}, \dfrac{k+1}{n} \right)$ corresponding to different integers k, and to divide each such interval of length $\dfrac{1}{n}$ into two parts:

– a part $\left[\dfrac{k}{n}, \dfrac{k}{n} + \mu\left(\dfrac{k}{n}\right) \cdot \dfrac{1}{n} \right)$, a portion $\mu\left(\dfrac{k}{n}\right)$, is assigned to the set S_n, while
– the remaining part $\left(\dfrac{k}{n} + \mu\left(\dfrac{k}{n}\right) \cdot \dfrac{1}{n}, \dfrac{k+1}{n} \right)$ is assigned to the complement of the set S_n.

One can easily check that for the resulting sequence of sets

$$S_n = \bigcup_k \left[\frac{k}{n}, \frac{k}{n} + \mu\left(\frac{k}{n}\right) \cdot \frac{1}{n} \right),$$

the equation (6) holds for every x. The statement is proven.

Let us show that this enables us to uniquely describe probability of a fuzzy set.

Definition 2. *Let $\rho(x)$ be a continuous probability density, let $P(s) \overset{\text{def}}{=} \int_s \rho(x)\, dx$ be the resulting probability measure, and let S be a fuzzy set. We say that a real number $P(S)$ is a probability of a fuzzy set S if for every sequence of crisp set S_n with $S_n \to S$, we have $P(S_n) \to P(S)$.*

Proposition 2. *For every fuzzy set S with a continuous membership function $\mu(x)$, its probability is well-defined and equal to $P(S) = \int \mu(x) \cdot \rho(x)\, dx$.*

Comment. This result provides one more justification for the original Zadeh's definition of the probability of a fuzzy set [4].

Another case when the limit idea enables us to select a unique generalization is the case of a complement.

Definition 3. *We say that a fuzzy set S' is a complement to a fuzzy set S if for every sequence S_n of crisp sets for which $S_n \to S$, we have $-S_n \to S'$ for the sequence of their complements $-S_n$.*

Proposition 3. *For every fuzzy set S with a continuous membership function $\mu(x)$, its complement S' is well-defined and its membership function is equal to $\mu_{S'}(x) = 1 - \mu(S)$.*

The limit idea is not a panacea. While the above idea works well for defining probability, it is not a panacea. Let us show, for example, that this idea does not lead to a unique definition of a union or intersection of two fuzzy sets.

Indeed, ideally, we should be able to define the intersection of two fuzzy sets S and S' in a similar manner:

- we say that a fuzzy set is a union $S \cup S'$ of fuzzy sets S and S' if for every two sequences of crisp sets $S_n \to S$ and $S_n' \to S'$ imply $S_n \cup S_n' \to S \cup S'$;
- similarly, we say that a fuzzy set is an intersection $S \cap S'$ of fuzzy sets S and S' if for every two sequences of crisp sets $S_n \to S$ and $S_n' \to S'$ imply $S_n \cap S_n' \to S \cap S'$.

Alas, it turns out that for different sequences S_n and S_n', we get different limits. Indeed, let us consider, for example, the two identical fuzzy sets $S = S'$ both corresponding to complete ignorance $\mu_S(x) = \mu_{S'}(x) = 0.5$. For this membership function, the construction from the proof of Proposition 1 leads to the sets

$$S_n = \bigcup_k \left[\frac{k}{n}, \frac{k}{n} + \frac{1}{2} \cdot \frac{1}{n} \right). \tag{7}$$

If we use these sets as sequences S_n and S_n' corresponding to both fuzzy sets S and S', then we get $S_n \cup S_n' = S_n \cap S_n' = S_n$, and thus, for the limit fuzzy sets, we get $\mu(x) = 0.5$ for all x.

Alternatively, we can still use S_n as a sequence of crisp sets approximating the set S while using a different sequence

$$S_n' = \bigcup_k \left[\frac{k}{n} + \frac{1}{2} \cdot \frac{1}{n}, \frac{k+1}{n} \right) \tag{8}$$

to approximate the set S'. In this case:

- the union $S_n \cup S_n'$ is the whole real line, so in the limit (6), we get $\mu(x) = 1$ for all x;

– on the other hand, the intersection $S_n \cap S'_n$ consists of midpoints of all intervals $\left[\dfrac{k}{n}, \dfrac{k+1}{n} \right)$, so here, in the limit (6), we have $\mu(x) = 0$.

Comment. In mathematical terms, the property that $S_n \to S$ implies $P(S_n) \to P(S)$ is known as continuity. In these terms, we can say that:

– probability is a continuous function of sets (in the sense of convergence $S_n \to S$);
– complement is a continuous operation, while
– union and intersection are not continuous operations.

For such discontinuous operations, instead of a single limit value, we have an *interval* of possible limit values. So maybe we can extend the limit idea to interval-valued fuzzy sets?

Acknowledgments

References

1. G. J. Klir and B. Yuan, *Fuzzy Sets and Fuzzy Logic*, Prentice Hall, Upper Saddle River, New Jersey, 1995.
2. H. T. Nguyen and E. A. Walker, *First Course In Fuzzy Logic*, CRC Press, Boca Raton, Florida, 2006.
3. L. A. Zadeh, "Fuzzy sets", *Information and Control*, 1965, Vol. 8, pp. 338–353.
4. L. A. Zadeh, "Probabilisty measures of fuzzy events", *Journal of Mathematical Analysis and Applications*, 1968, Vol. 23, No. 2, pp. 421–427.