Which Random-Set Representation of a Fuzzy Set Is the Simplest?

Vladik Kreinovich, Olga Kosheleva, and Hung T. Nguyen

Abstract One of the ways to elicit membership degrees is by polling. For example, we ask a group of people how many believe that 30 C is hot. If 8 out of ten say that it is hot, we assign the degree 8/10 to the statement "30 C is hot". In precise mathematical terms, polling can be described via so-called random sets. It is known that every fuzzy set can be obtained this way, i.e., that every fuzzy set can be represented by an appropriate random set. Moreover, it is known that for many fuzzy sets, there are several different random-set representations. From the computational viewpoint, it is desirable to use the random sets which are the simplest, i.e., which contains the smallest possible number of elements. So, the natural questions are: what is the simplest random-set representation of a given fuzzy set? and is such simplest representation unique or are there several different random-set representations with the same number of elements? In this paper, we answer both questions: we show that for almost all fuzzy sense (in some reasonable sense), there are several different simplest random-set representations, and that the known α-cut representation (where probabilities are assigned to α-cuts of the fuzzy set) is one of them.

Vladik Kreinovich
Department of Computer Science, University of Texas at El Paso
500 W. University, El Paso, Texas 79968, USA, e-mail: vladik@utep.edu

Olga Kosheleva
Department of Teacher Education, University of Texas at El Paso
500 W. University, El Paso, Texas 79968, USA, e-mail: olgak@utep.edu

Hung T. Nguyen
Department of Mathematical Sciences, New Mexico State University
Las Cruces, New Mexico 88003, USA, and Faculty of Economics
Chiang Mai University, Thailand, e-mail: hunguyen@nmsu.edu

1 Formulation of the Problem

Our objective is to look for the simplest random-set representation of a fuzzy set. In this section, we explain what exactly is the problem and why this problem is important.

To understand why this problem is important, we need to understand why do we need random-set representations of fuzzy sets in the first place. And to answer that question, we need to recall how and why fuzzy techniques appeared. So let us start with the origin of the fuzzy techniques.

Why fuzzy techniques? In the 1960s, Lotfi Zadeh, who was then one of the leading researchers in optimal control – and an author of the most popular textbook on this subject – was thinking about the puzzle that has been bothering him – and others in the research community – for quite some time. The puzzle was that in many practical situations like chemical engineering, human controllers were much more effective in controlling the plant than the most sophisticated automatic controllers, controllers that used state-of-the-art optimization techniques.

From the purely mathematical viewpoint, this may sound like a paradox: how can some control be better than the optimal control? However, from the practical viewpoint, there is no paradox: what we call optimal controller are optimal with respect not necessarily of the *actual* plant, but with respect to the available mathematical *model* of this plant. In this sense, the fact that control which is *optimal* with respect to these models was *not optimal* with respect to the real plant means that there is a discrepancy between these models and the actual plant, that these models need improvement.

Similarly, the fact that human controllers can control better than the model-based automatic controllers means that human controllers have some additional knowledge about the plant, knowledge that is not yet incorporated in the models. So, a natural idea is to elicit this knowledge from the experts and add this knowledge to the model.

Most human experts were absolutely willing to share their expertise, but it turns out that it is not easy to add this knowledge to the model. The main obstacle was that when experts described how they control, they used imprecise ("fuzzy") words from natural language like "small" or "very small". We humans understand these words, but computers do not. To capture this knowledge and to be able to add it to the model, Zadeh invented a special techniques that he called *fuzzy* [1, 3, 5, 9, 10, 12].

Main idea behind fuzzy techniques: a brief reminder. In the ideal case, experts may say something like "If the concentration x of a certain gas becomes larger than 0.1, then perform some action A" (e.g., switch on an additional filter). This is a clear and precise statement. It means that:

- if x is smaller than or equal to 0.1, do not perform the action A, while
- if $x > 0.1$, then perform the action A.

Unfortunately, the actual expert rules are often not as precise. (Those that are precise have probably already been incorporated into the model.) The actual rules sounds more like this: "If the concentration x of a certain has becomes sufficiently large, then perform some action A." The very fact that the expert cannot formulate a more precise rule means that the expert him/herself is not always 100% sure how to act:

- when the concentration x is very small, e.g., 0.01, the expert is probably 100% sure that this value is not sufficiently large, and thus, the action A should not be peformed;
- when x is large, e.g., $x = 0.3$, then the expert is 100% sure that this value is sufficiently large and thus, the action A should be performed;
- however, in the intermediate cases, e.g., when $x = 0.09$ or $x = 0.11$, the expert is not sure whether this value is sufficiently large; when asked, he/she may reply that it somewhat large (or something like that).

How can we describe this imprecise information in a computer-understandable form?

For a precise statement like $x > 0.1$, for each x, this statement is either true or false. For an imprecise statement "x is sufficiently large", this statement is sometimes neither absolutely true nor absolutely false – according to the expert, it is true to some extent. In a computer, "true" is usually represented as 1, and "false" as 0. It is therefore reasonable to described intermediate degrees of confidence – such as "somewhat true" – by numbers between 0 and 1. These numbers are called *truth values*, or, alternatively, *membership degrees* (since they generalize 1 or 0 describing whether x is a member of the set $\{x : x > 0.1\}$ of all the values which are greater than 0.1).

So, to describe the expert-used natural-language property like "x is sufficiently large" in precise terms, we need to assign, to each real number x, the degree $\mu(x) \in [0, 1]$ to which, according to the expert, x satisfies this property. The function $\mu(x)$ that assigns these degrees is called a *membership function*, or, alternatively, a *fuzzy set*.

Comment. Usually, we consider fuzzy sets for which $\mu(x) = 1$ for some x.

But how do we elicit these degrees? To utilize the above idea, we need to elicit, from the expert, the values $\mu(x)$ corresponding to different real numbers x. A natural way to do it is to explicitly ask the expert to mark, for each x, his/her degree of confidence in the given statement by a number $\mu(x)$ from the interval $[0, 1]$, so that:

- $\mu(x) = 0$ means that the expert is absolutely sure that the given property is not satisfied for this x,
- $\mu(x) = 1$ means that the expert is absolutely sure that the given property is satisfied for this x, and
- values $\mu(x)$ between 0 and 1 correspond to intermediate degrees of belief.

The expert may not be accustomed for such evaluations, but we are all accustomed to evaluations on a scale:

- students evaluate their instructors on a scale from 0 to 4,
- we all evaluate the quality of a service on a scale from 0 to 10, etc.

So, if the expert cannot easily mark his/her degrees of a 0-to-1 scale, we can simply ask him/her to mark this degree on a 0-to-10 scale and divide the results by 10.

This "Likert-scale" approach is indeed the main way of eliciting membership functions. It works for many experts.

For some experts, this usual approach does not work. Many people have no trouble marking their degree of confidence on a scale – but not all people. There is a significant portion of people who are not comfortable marking on a scale (one of the authors VK is one of these people, as many other people with math background).

The existence of such people is not a problem for student evaluations or for a hotel interested in the customer's opinion about their new breakfast menu: such unusual people are in a minority, and there are many others who can provide the desired degrees.

However, for describing expert estimate, we do not want to dismiss any experts – since there usually only a few top experts, and we do not want to miss the opinion of each of them. What can we do?

A solution is simple: polling. The way we describe the problem makes it sound like an unusual problem requiring some creative ideas. However, this a problem that we encounter all the time. For example, shall a Democrat or a Republican represent our district in the state legislature? If you ask people, many of them will mark it on a scale: e.g., a Libertarian may support Republicans 60% on economic issues, but support Democrats 40% on human issues of gay rights etc. In practice, no one asks for these marking: people just vote for one candidate or another. If out of n voters, n_c voted for the candidate c, then the portion n_c/n of the electorate who voted for this candidate can serve as a measure of this candidate's support.

Similarly, for each x, we can simply ask experts whether they believe that this x is sufficiently large. If out of n experts, $n(x)$ said that x is sufficiently large, then we can take $\mu(x) = n(x)/n$.

Random set is, in effect, just a fancy mathematical term for polling. How can we describe polling in precise terms?

If we select an expert randomly, so that each expert has the same probability $1/n$ of being selected, then $n(x)/n$ is exactly the probability that a randomly selected expert will say that the given value x is sufficiently large.

Let us describe it in precise terms. Each expert:

- for some x, says that this x is sufficiently large, while
- for some other values x, this expert will say that x is not sufficiently large.

So, the opinions of an expert i can be describe by the set S_i of all the values x about which this expert things that they are sufficiently large. In these terms, to fully describe the expert opinions, we need to describe the corresponding sets $S_1, \ldots, S_n$.

Selecting an expert with probability $1/n$ means selecting one of these sets S_i with equal probability $1/n$. This is what is called a *random set* (see, e.g., [6, 8]):

- a *random number* is when we have different numbers with different probabilities;
- similarly, a *random set* if when we have different sets with different probabilities.

In these terms, we have a random set $\mathcal{S}$ in which each of the sets $S_1, \ldots, S_n$ appear with the same probability $P(\{S_i\}) = 1/n$. In these terms, the value $\mu(x) = n(x)/n$

can be described as the probability that x is contained in the randomly selected set, i.e.,

$$\mu(x) = P(x \in \mathcal{S}). \tag{1}$$

This is what is known as a *random-set representation of a fuzzy set*.

First comment: we may have different probabilities. In political voting, usually, everyone has the same vote. This corresponds to the case when each set whose probability is not 0 has the exact probability $1/n$. However, in describing expert opinions, we may want to assign more weights $w_i > 1$ to opinions of top experts and smaller weights $w_i < 1$ to the opinions of less experienced experts. Then, to get the degree $\mu(x)$, instead of simply counting number of experts who believe that x has the given property, we can count them with their weights, and take the ratio

$$\mu(x) = \frac{\sum\limits_{i: x \in S_i} w_i}{\sum\limits_{j=1}^{n} w_j}.$$

In random set terms, this is equivalent to still using the formula (1), but with different probabilities of different sets:

$$P(\{S_i\}) = \frac{w_i}{\sum\limits_{j=1}^{n} w_j}.$$

Second comment: we can only do it for finitely many values x. In practice, we can only ask finitely many questions to an expert, so we can only elicit the values $\mu(x)$ for a finite set of real numbers $X = \{x_1, \ldots, x_m\}$.

Towards the first natural question. Let us briefly recall what we have just discussed. The usual way of eliciting a membership function is to ask the expert to assign, to each value x – or, to be precise, to each value $x \in \{x_1, \ldots, x_m\}$ – the corresponding degree $\mu(x) \in [0, 1]$. This way, we can, in principle, get all possible functions $\mu : X \mapsto [0, 1]$ from the finite set $X \subset \mathbb{R}$ of the real line to the interval $[0, 1]$.

Then, we described an alternative elicitation mechanism in which we have a random set – in which we have sets $S_1, \ldots, S_n \subseteq X$ with probabilities $P_i \stackrel{\text{def}}{=} P(\{S_i\})$ that add up to 1, and we take $\mu(x) = P(x \in \mathcal{S})$, i.e.,

$$\mu(x) = \sum_{i: x \in S_i} P(\{S_i\}).$$

The first natural question is as follows.

Natural question 1. *Can any membership function be obtained this way?*

The answer to this natural question. The answer (positive) to this natural question was, in effect, provided in [7]. Actually, for each membership function $\mu(x)$, two different random sets can be provided that lead to $\mu(x)$:

- One possibility is to assume that each element $x \in X$ belongs to the set S with probability $\mu(x)$ and does not belong to S with the remaining probability $1-\mu(x)$, and that different elements $x \in X$ are independent. In this case, for each set $S \subseteq X$, its probability $P(\{S\})$ is equal to

$$P(\{S\}) = \left(\prod_{i:i\in S} \mu(x)\right) \cdot \left(\prod_{x:x\notin S} (1-\mu(x))\right). \quad (2)$$

- The second possibility is to sort the values $\mu_1 \stackrel{\text{def}}{=} \mu(x_1), \ldots, \mu_m \stackrel{\text{def}}{=} \mu(x_m)$ in the increasing order:

$$\mu_{(1)} \le \ldots \le \mu_{(m)},$$

  and consider the random set $\mathcal{S}$ in which possible sets are α-cuts

$$S_j \stackrel{\text{def}}{=} \{x : \mu(x) \ge \mu_{(j)}\},$$

  and we have $P(\{S_j\}) = \mu_{(j)} - \mu_{(j-1)}$ for $j > 1$ and $P(\{S_1\}) = \mu_{(1)}$ for $i = 1$. We also need to assign probability $P(\{\emptyset\}) = 1 - \mu_{(m)}$ to the empty set $\emptyset$.

Comment. As we have mentioned, usually, we consider fuzzy sets for which $\mu(x) = 1$ for some x. In this case, $\mu_{(m)} = 1$, so we do not need to assign a non-zero probability to the empty set.

From the computational viewpoint, what is the difference between there two representations? From the purely mathematical viewpoint, both representations are possible. However, as we have mentioned, the main goal of fuzzy techniques is applications – and in applications, we need to perform computations with fuzzy degrees. From this viewpoint, there is a big difference between these two representations:

- In the first representation, in the general case when all values μ_i are different from 0 and 1, all 2^m sets $S \subseteq \{1, \ldots, m\}$ have non-zero probability. (If we have $\mu(x_0) = 1$ for some x_0, then we have 2^{m-1} sets with non-zero probability – since sets not containing x_0 will have 0 probability.)
- In the second representation, we only need $m + 1$ sets with non-zero probability. (In the case when $\mu(x_0) = 1$ for some x_0, we only need m such sets.)

The difference is that dealing with 2^m objects is not computationally feasible: already for reasonable m – e.g., for $m = 300$ – we get more computational steps than the lifetime of the Universe; see, e.g., [2, 4, 11]. In contrast, dealing with $m+1$ or m sets is quite feasible.

In general, the fewer objects we have, the more feasible are the corresponding computations. Thus, we naturally arrive at the following two questions.

Natural question 2. *Is the number of non-zero-probability sets corresponding to the α-cut representation the smallest possible, or can we further decrease the number of such sets?*

Natural question 3. *If the number of non-zero-probability sets cannot be further decreased, is the α-cut representation the only one with this smallest number of non-zero-probability sets?*

What we do in this paper. In this paper, we provide answers to both these questions.

2 Definitions and the Main Results

Discussion. Let us start with the second natural question. Of course, there are cases when fewer than m non-zero-probability sets are sufficient: for example, a fuzzy set in which $\mu(x_1) = 1$ and $\mu(x_i) = 0$ can described by a random set $\mathcal{S}$ that is equal to $\{x_1\}$ with probability 1.

What we will show is that *for almost all* fuzzy sets, we cannot have fewer than m non-zero-probability sets. First, we need to explain what we mean by "almost all".

Definitions. *Let a set $X = \{x_1, \ldots, x_m\} \subset \mathbb{R}$ be given.*

- *By a* fuzzy set*, we understand a function $\mu : X \mapsto [0, 1]$ for which $\mu(x_{j_0}) = 1$ for some j_0.*
- *By a* random set*, we mean a probability measure $\mathcal{S}$ on the set 2^X of all subsets of the set X, i.e., a class of sets $\{S_1, \ldots, S_n\}$ with probabilities $P_i = P(\{S_i\}) > 0$ that add up to 1.*
- *We say that a random set* represents *a fuzzy set $\mu(x)$ if for every x, we have*

$$\mu(x) = P(x \in \mathcal{S}) = \sum_{i: x \in S_i} P_i.$$

- *To each fuzzy set $\mu(x)$ with $\mu(x_{j_0}) = 1$, we can put into correspondence a point*

$$(\mu(x_1), \ldots, \mu(x_{j_0-1}), \mu(x_{j_0+1}), \ldots, \mu(x_n)) \in [0, 1]^{m-1}$$

  in the $(m-1)$-dimensional unit cube $[0, 1]^{m-1}$. Thus, the class of all fuzzy sets is mapped to the union U of m unit cubes corresponding to $j_0 = 1, \ldots, m$.
- *On the union U, we can define the usual Lebesgue measure.*
- *We say that a property if true* for almost all fuzzy sets *if fuzzy sets that do not have this property form a set of measure 0 in the union U.*

Proposition. *The following two properties hold for almost all fuzzy sets $\mu(x)$:*

- *any random set representing $\mu(x)$ has at least m non-zero-probability sets S_j;*
- *when $m \geq 3$, there exists at least two different random sets representing $\mu(x)$.*

Discussion. Thus, we have answers to both Natural Questions 2 and 3:

- The answer to Natural Question 2 is that, in almost all cases, the α-cut representation of a fuzzy set is indeed the shortest.
- The answer to Natural Question 3 is that, in almost all cases, the α-cut representation of a fuzzy set is not the only shortest one.

3 Proof

0°. For convenience, for each fuzzy set $\mu(x)$, let us denote by j_0, the index for which $\mu(x_{j_0}) = 1$, and let us denote $\mu_j \stackrel{\text{def}}{=} \mu(x_j)$.

1°. Let us first prove the first statement.

In the proof of this statement, we will use linear algebra over the set Q of all rational numbers. It is very similar to the usual linear algebra, the only difference is that we are only allowing rational coefficients in linear combinations. In other words:

- we say that a number v is a *Q-linear combination* of the numbers $v_1, \ldots, v_k$ if $v = q_1 \cdot v_1 + \ldots + q_k \cdot v_k$ for some rational numbers $q_1, \ldots, q_k$;
- we say that numbers $v_1, \ldots, v_k$ are *Q-linearly dependent* if some linear combination v of these values is equal to 0 while not all q_i are equal to 0;
- we say that numbers $v_1, \ldots, v_k$ are *Q-linearly independent* if they are not Q-linearly dependent;
- by a *Q-linear space* generated by the numbers $v_1, \ldots, v_k$, we mean the set of all possible Q-linear combinations of these numbers;
- by the *Q-dimension* of a Q-linear space we mean the smallest amount of numbers that generate this space.

1.1°. No, let us consider the class C of all fuzzy sets for which $m - 1$ values $\mu_1, \ldots, \mu_{j_0-1}, \mu_{j_0+1}, \ldots, \mu_n$ are Q-linearly independent, i.e., for which an equality

$$q_1 \cdot \mu_1 + \ldots + q_{j_0-1} \cdot \mu_{j_0-1} + q_{j_0-1} \cdot \mu_{j_0-1} + \ldots + q_m \cdot \mu_0 = 0, \tag{3}$$

where $q_j \in Q$, implies that $q_j = 0$ for all j.

In linear algebra terms, this means that the Q-dimension d of the Q-linear space generated by these values μ_j is equal to $m - 1$.

1.2°. Let us prove that the complement to the class C has measure 0.

Indeed, the class C can obtained if we subtract, from the class of all fuzzy sets, the union of the following sets S_q:

- for each tuple $q = (q_1, \ldots, q_{j_0-1}, q_{j_0+1}, \ldots, q_m)$ of rational numbers $q_j \in Q$ not all of which are equal to 0,
- by S_q, we denote the class of all fuzzy sets that satisfy the corresponding property (3).

Each set S_q forms a hyperplane in the cube and is, thus, of Lebesgue measure 0.

There are countably many rational numbers. Thus, there are countable many tuples q – so, countably many sets S_q. The union of countably many sets of measure 0 also has measure 0. Thus, the complement to the set C indeed has measure 0.

1.3°. Let us now prove that for any fuzzy set $\mu(x)$ from the class C, every random set representing $\mu(x)$ has to have at least m different elements.

Let $\mathcal{S}$ be a random set with m' non-zero-probability elements $S_1, \ldots, S_{m'}$ that represents $\mu(x)$. Let $P_1, \ldots, P_{m'}$ be the probabilities corresponding to $\mathcal{S}$. The probabilities P_i satisfy the condition $P_1 + \ldots + P_{m'} = 1$. Hence, $P_{m'} = 1 - P_1 - \ldots - P_{m'-1}$ is a Q-linear combination of $m' - 1$ values $P_1, \ldots, P_{m'-1}$. Thus, $m' - 1$ numbers $P_1, \ldots, P_{m'-1}$ generate the Q-linear space L generated by all these probabilities.

Since the random set $\mathcal{S}$ represent the fuzzy set $\mu(x)$, each value $\mu(x_i)$ can be represented as a sum of the probabilities P_i. Thus, all the values μ_j can be represented as Q-linear combinations of $m' - 1$ probabilities. Hence, the Q-dimension d of the set generated by all the values μ_j is smaller than or equal to $m' - 1$: $d \le m' - 1$.

We have shown that for all $\mu(x) \in C$, we have $d = m - 1$. Thus, we have $m - 1 \le m' - 1$, i.e., indeed, $m' \ge m$. The proposition is proven.

2°. Let us now prove that for almost all fuzzy sets $\mu(x)$, there exist a random set with m non-zero-probability elements which is different from the α-cut one and that also represents the fuzzy set $\mu(x)$. We will prove that this is true for all fuzzy sets from the class C.

By definition of the class C, for each fuzzy set from this class, all the values μ_j – with the exception of the value $\mu_{j_0} = 1$ – are Q-linearly independent. This implies, in particular, that these values are all different. Thus, when we sort them in increasing order, all inequalities are strict:

$$\mu_{(1)} < \ldots < \mu_{(m-1)} < \mu_{(m)} = 1.$$

Without losing generality, we can assume that this was the original enumeration of the values x_j, i.e., that

$$\mu_1 < \ldots < \mu_{m-1} < \mu_m = 1.$$

In the α-cut representation of this fuzzy sets, we have the following m sets:

- the set $S_m = \{x_m\}$ with probability $P_m = 1 - \mu_{m-1}$;
- the set $S_{m-1} = \{x_{m-1}, x_m\}$ with probability $P_{m-1} = \mu_{m-1} - \mu_{m-2}$,
- the set $S_{m-2} = \{x_{m-2}, x_{m-1}, x_m\}$ with probability $P_{m-1} = \mu_{m-2} - \mu_{m-3}$,
- ...
- the set $S_i = \{x_j, x_{j+1}, \ldots, x_m\}$ with probability $P_i = \mu_j - \mu_{j-1}$,
- ...
- the set $S_2 = \{x_2, x_3, \ldots, x_m\}$ with probability $P_2 = \mu_2 - \mu_1$, and
- the set $S_1 = \{x_1, x_2, \ldots, x_m\}$ with probability $P_1 = \mu_1$.

We can form a different random-set representation of the same fuzzy set if we replace the sets S_m, S_{m-1}, and S_{m-2} with different three sets S'_m, S'_{m-1}, and S'_{m-2} while keeping all other sets and their probabilities unchanged. Depending on the relation between $1 - \mu_{m-1}$ and $\mu_{m-2} - \mu_{m-3}$, we have the following two options of selecting the sets S'_i.

2.1°. When $1 - \mu_{m-1} \ge \mu_{m-2} - \mu_{m-3}$, we take:

- the set $S'_m = \{x_m\}$ with probability $P'_m = (1 - \mu_{m-1}) - (\mu_{m-2} - \mu_{m-3})$
- the set $S'_{m-1} = \{x_{m-1}, x_m\}$ with probability $P'_{m-1} = \mu_{m-1} - \mu_{m-3}$,
- the set $S'_{m-2} = \{x_{m-2}, x_m\}$ with probability $P'_{m-2} = \mu_{m-2} - \mu_{m-3}$, and

- we take $S'_i = S_i$ with probability $P'_i = P_i$ for all $i < m-2$.

Let us show that this new random set $\mathcal{S}'$ represents the same fuzzy set as the original random set $\mathcal{S}$, i.e., that for each j, we have

$$\sum_{i:x_j \in S'_i} P'_i = \sum_{i:x_j \in S_i} P_i = \mu_j.$$

2.1.1°. For $j < m-2$, the only sets that contain x_j are sets S'_i for $i \le j$. For these sets, we have $S'_i = S_i$ and $P'_i = P_i$, so indeed

$$\sum_{i:x_j \in S'_i} P'_i = \sum_{i:x_j \in S_i} P_i.$$

So, to prove that the new random set represents the same fuzzy set, it is sufficient to consider three remaining elements: x_m, x_{m-1}, and x_{m-2}.

2.1.2°. For x_m, we have:

$$\sum_{i:x_m \in S'_i} P'_i =$$

$$(1-\mu_{m-1}) - (\mu_{m-2} - \mu_{m-3}) + (\mu_{m-1} - \mu_{m-3}) + (\mu_{m-1} - \mu_{m-3}) + \sum_{i=1}^{m-3} P'_i. \quad (4)$$

Here, the last sum in the right-hand side is equal to

$$\sum_{i=1}^{m-3} P'_i = \sum_{i=1}^{m-3} P_i = \mu_1 + (\mu_2 - \mu_1) + \ldots + (\mu_{m-3} - \mu_{m-4}).$$

One can check that if we open the parentheses, then all the terms cancel each other except for μ_{m-3}, so

$$\sum_{i=1}^{m-3} P'_i = \mu_{m-3}. \quad (5)$$

If we substitute this expression into the formula (4) and open parentheses, the formula (4) has the form

$$\sum_{i:x_m \in S'_i} P'_i = 1 - \mu_{m-1} - \mu_{m-2} + \mu_{m-3} + \mu_{m-1} - \mu_{m-3} + \mu_{m-1} - \mu_{m-3} + \mu_{m-3}.$$

One can see that all the terms cancel each other except for the term 1, so we have

$$\sum_{i:x_m \in S_i} P'_i = 1 = \mu_m.$$

This is exactly what we wanted.

2.1.3°. For x_{m-1}, we have:

$$\sum_{i:x_{m-1}\in S'_i} P'_i = (\mu_{m-1} - \mu_{m-3}) + \sum_{i=1}^{m-3} P'_i.$$

Due to the formula (5), we have

$$\sum_{i:x_{m-1}\in S'_i} P'_i = (\mu_{m-1} - \mu_{m-3}) + \mu_{m-3} = \mu_{m-1},$$

also exactly what we wanted.

2.1.4°. For x_{m-2}, we have:

$$\sum_{i:x_{m-2}\in S'_i} P'_i = (\mu_{m-2} - \mu_{m-3}) + \sum_{i=1}^{m-3} P'_i.$$

Due to the formula (5), we have

$$\sum_{i:x_{m-2}\in S'_i} P'_i = (\mu_{m-2} - \mu_{m-3}) + \mu_{m-3} = \mu_{m-2},$$

also exactly what we wanted.

2.2°. When $1 - \mu_{m-1} < \mu_{m-2} - \mu_{m-3}$, we take:

- the set $S'_m = \{x_{m-1}, x_{m-2}, x_m\}$ with probability $P'_m = (\mu_{m-2} - \mu_{m-3}) - (1 - \mu_{m-1})$,
- the set $S'_{m-1} = \{x_{m-1}, x_m\}$ with probability $P'_{m-1} = 1 - \mu_{m-2}$,
- the set $S'_{m-2} = \{x_{m-2}, x_m\}$ with probability $P'_{m-2} = 1 - \mu_{m-1}$, and
- we take $S'_i = S_i$ with probability $P'_i = P_i$ for all $i < m - 2$.

Let us show that this new random set $\mathcal{S}'$ represents the same fuzzy set as the original random set $\mathcal{S}$, i.e., that for each j, we have

$$\sum_{i:x_j\in S'_i} P'_i = \sum_{i:x_j\in S_i} P_i = \mu_j.$$

2.2.1°. For $j < m - 2$, the only sets that contain x_j are sets S'_i for $i \le j$. For these sets, we have $S'_i = S_i$ and $P'_i = P_i$, so indeed

$$\sum_{i:x_j\in S'_i} P'_i = \sum_{i:x_j\in S_i} P_i.$$

So, to prove that the new random set represents the same fuzzy set, it is sufficient to consider three remaining elements: x_m, x_{m-1}, and x_{m-2}.

2.2.2°. For x_m, we have:

$$\sum_{i:x_m\in S'_i} P'_i = (\mu_{m-2}-\mu_{m-3}) - (1-\mu_{m-1}) + (1-\mu_{m-2}) + (1-\mu_{m-1}) + \sum_{i=1}^{m-3} P'_i.$$

Substituting the expression (5) for the last sum in the right-hand side and opening the parentheses, we get

$$\sum_{i:x_m\in S'_i} P'_i = \mu_{m-2}-\mu_{m-3} - 1 + \mu_{m-1} + 1 - \mu_{m-2} + 1 - \mu_{m-1} + \mu_{m-3}.$$

One can see that all the terms cancel each other, except one 1, so we get

$$\sum_{i:x_m\in S_i} P'_i = 1 = \mu_m,$$

exactly what we wanted.

2.2.3°. For x_{m-1}, we have:

$$\sum_{i:x_{m-1}\in S'_i} P'_i = (\mu_{m-2}-\mu_{m-3}) - (1-\mu_{m-1}) + (1-\mu_{m-2}) + \sum_{i=1}^{m-3} P'_i.$$

Substituting the expression (5) for the last sum in the right-hand side and opening the parentheses, we get

$$\sum_{i:x_{m-1}\in S'_i} P'_i = \mu_{m-2}-\mu_{m-3} - 1 + \mu_{m-1} + 1 - \mu_{m-2} + \mu_{m-3}.$$

One can see that all the terms cancel each other, except μ_{m-1}, so we get

$$\sum_{i:x_{m-1}\in S_i} P'_i = \mu_{m-1},$$

exactly what we wanted.

2.2.4°. Finally, for x_{m-2}, we have:

$$\sum_{i:x_{m-2}\in S'_i} P'_i = (\mu_{m-2}-\mu_{m-3}) - (1-\mu_{m-1}) + (1-\mu_{m-1}) + \sum_{i=1}^{m-3} P'_i.$$

Substituting the expression (5) for the last sum in the right-hand side and opening the parentheses, we get

$$\sum_{i:x_{m-2}\in S'_i} P'_i = \mu_{m-2}-\mu_{m-3} - 1 + \mu_{m-1} + 1 - \mu_{m-1} + \mu_{m-3}.$$

One can see that all the terms cancel each other, except μ_{m-2}, so we get

$$\sum_{i:x_{m-2}\in S_i} P'_i = \mu_{m-2},$$

exactly what we wanted.

The proposition is thus proven.

Acknowlegments

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), EAR-2225395 (Center for Collective Impact in Earthquake Science C-CIES), and by the AT&T Fellowship in Information Technology.

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, and by a grant from the Hungarian National Research, Development and Innovation Office (NRDI).

References

1. R. Belohlavek, J. W. Dauben, and G. J. Klir, *Fuzzy Logic and Mathematics: A Historical Perspective*, Oxford University Press, New York, 2017.
2. Th. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein, *Introduction to Algorithms*, MIT Press, Cambridge, Massachusetts, 2022.
3. G. Klir and B. Yuan, *Fuzzy Sets and Fuzzy Logic*, Prentice Hall, Upper Saddle River, New Jersey, 1995.
4. V. Kreinovich, A. Lakeyev, J. Rohn, and P. Kahl, *Computational Complexity and Feasibility of Data Processing and Interval Computations*, Kluwer, Dordrecht, 1998.
5. J. M. Mendel, *Uncertain Rule-Based Fuzzy Systems: Introduction and New Directions*, Springer, Cham, Switzerland, 2017.
6. I. Molchanov, *Theory of Random Sets*, Springer, Cham, Switzerland, 2017.
7. H. T. Nguyen, “On random sets and belief functions", *Journal of Mathematical Analysis and Applications*, 1978, Vol. 65, pp. 531–542.
8. H. T. Nguyen, *An Introduction to Random Sets*, Chapman & Hall/CRC, Boca Raton, Florida, 2006.
9. H. T. Nguyen, C. L. Walker, and E. A. Walker, *A First Course in Fuzzy Logic*, Chapman and Hall/CRC, Boca Raton, Florida, 2019.
10. V. Novák, I. Perfilieva, and J. Močkoř, *Mathematical Principles of Fuzzy Logic*, Kluwer, Boston, Dordrecht, 1999.
11. C. Papadimitriou, *Computational Complexity*, Addison-Wesley, Reading, Massachusetts, 1994.
12. L. A. Zadeh, “Fuzzy sets”, *Information and Control*, 1965, Vol. 8, pp. 338–353.