Shadows of Fuzzy Sets – A Natural Way to Describe 2-D and Multi-D Fuzzy Uncertainty in Linguistic Terms

Hung T. Nguyen[1], Berlin Wu[2], and Vladik Kreinovich[3]

[1]Department of Mathematical Sciences
New Mexico State University
Las Cruces, NM 88003, USA
email hunguyen@nmsu.edu

[2]Department of Mathematical Sciences
National Chengchi University
Taipei, Taiwan
email berlin@math.nccu.edu.tw

[3]Department of Computer Science
University of Texas at El Paso
El Paso, TX 79968, USA
email vladik@cs.utep.edu

Abstract

Fuzzy information processing systems start with expert knowledge which is usually formulated in terms of words from natural language. This knowledge is then usually reformulated in computer-friendly terms of membership functions, and the system transform these input membership functions into the membership functions which describe the result of fuzzy data processing. It is then desirable to translate this fuzzy information back from the computer-friendly membership functions language to the human-friendly natural language. In a 1-D case, when we are interested in a single quantity y, it is usually easy to describe the resulting membership function by a word from natural language, because most words do describe 1-D case and there are, therefore, so many of them that the corresponding membership functions form a dense set in the class of all possible membership functions. The problem becomes more complicated in 2-D and multi-D cases, when we are interested in several quantities $y_1, \ldots, y_m$, because there are fewer words which describe the relation between several quantities. To describe such fuzzy information in terms of a natural language, L. Zadeh proposed, in 1966, to use words to describe fuzzy information about different *combinations* $y = f(y_1, \ldots, y_m)$ of the desired variables. This idea is similar to the use of marginal distributions in probability theory. The corresponding terms are called *shadows* of the original fuzzy set. The main question is: do we lose any information in this translation? Zadeh has shown that under certain conditions, the original fuzzy set can be uniquely reconstructed from its shadows. In this paper, we prove that for appropriately chosen shadows, the reconstruction is *always* unique. Thus, the translation from the original membership function into the linguistic terms which describe different combinations y is lossless.

1 Membership Functions As a Computer-Friendly Translation of Natural Language Terms

Humans often describe their knowledge by terms from natural language like "young", "large", etc. If we want a computer to be able to use this knowledge, we must reformulate it in terms which are understandable to a computer. One of the main objectives of fuzzy methodology is to provide such a translation. Fuzzy logic describes each natural language term t defined on a set X by the corresponding *membership function*

$\mu_t(x) : X \to [0,1]$, a function which describes, for each element x of the domain X, to what extent this element x satisfies the property t.

Fuzzy methodology provides us with the tools (t-norms, t-conorms, fuzzy inference rules, etc.) which are able to process these functions. A typical application of these tools is to the following situation:

- We are interested in the values of some quantities $y_1, \ldots, y_m$ about which we have no direct knowledge (e.g., we may be interested to know how the economy will grow in the next few years).
- What we do know is the relation between these quantities y_i and some other quantities $x_1, \ldots, x_n$ about which we have some (fuzzy) knowledge. For example, for an economy, we may know how it was growing in the past, we may know some specific parameters characterizing its common state, etc. The rules connecting x_i and y_j are also typically described not in precise mathematical form, but rather by words from natural language.

Fuzzy methodology enables us to transform a fuzzy knowledge about x_i and the fuzzy rules which connect x_i and y_j into a fuzzy knowledge about y_j, i.e., into the membership function $\mu(\vec{y})$ on the set of all possible values of $\vec{y} = (y_1, \ldots, y_m)$ (see, e.g., [1, 2]). In short, we get the desired information about y_j, but we get it in terms of membership functions.

2 It Is Desirable to Translate The Result of Fuzzy Data Processing Back Into the Natural Language

A membership function is not something which is natural for a human to understand and to use, it was invented as a way of representing human fuzzy knowledge in a language which is understandable for a computer. From this viewpoint, the fact that the result of using traditional fuzzy methodology is a membership function means that this result is not presented in a very user-friendly form; for the user's convenience, we must translate the result of computer's information processing from the computer-native language of membership functions into the human-friendly natural language.

3 Such a Translation Is Relatively Easy in 1-D Cases

This translation is relatively easy in a 1-D case, when we are interested only in the value of a single quantity y_1. This easiness comes from the fact that most words from natural language characterize a single quantity ("young", "small", etc.), so there are plenty of different membership functions corresponding to different words of this type. Since there are many such membership functions stemming from natural language terms, these functions form a dense net in the set of all possible membership functions. Therefore, often, for a membership function $\mu(y)$ produced by the fuzzy system, we are able to find a natural language term t for which the corresponding membership function $\mu_t(y)$ is close enough to $\mu(y)$. It is then natural to return this term t as the result of fuzzy information processing.

4 2-D and Multi-D Cases: The Idea of Shadows of a Fuzzy Set

For 2-D and multi-D problems, when we are interested in the values of several quantities $y_1, \ldots, y_m$, $m \geq 2$, the situation is radically different. There are much fewer terms from natural language which describe the relation between several quantities than in 1-D case: e.g., we can say that y_1 is much larger than y_2, etc. So, in such cases, when the fuzzy system produces a membership function $\mu(\vec{y}) = \mu(y_1, \ldots, y_m)$, we are often unable to find a natural-language term which describes this membership function. How can we, in this situation, describe the membership function in (user-friendly) linguistic terms?

To describe such fuzzy information in terms of natural language, L. Zadeh proposed, in [4], to use words to describe fuzzy information about different *combinations* $y = f(y_1, \ldots, y_m)$ of the desired variables. The corresponding 1-D fuzzy set is called a *shadow* of the original multi-D fuzzy set. (This idea is similar to the use of marginal distributions in probability theory.) For example, we can use *linear* combinations

$$f(y_1, \ldots, y_m) = a_0 + a_1 \cdot y_1 + \ldots + a_m \cdot y_m; \tag{1}$$

or, we can use more general *quadratic* combinations

$$f(y_1,\dots,y_m)=a_0+\sum_{i=1}^m a_i\cdot y_i+\sum_{i=1}^m\sum_{j=1}^m a_{ij}\cdot y_i\cdot y_j. \tag{2}$$

Shadows corresponding to linear combination functions $f(\vec{y})$ will be called *linear* shadows, and shadows corresponding to quadratic combination functions $f(\vec{y})$ will be called *quadratic* shadows. For each such combination, in order to describe the information about this combination in terms of natural language, we do the following:

- first, we find the membership function $\mu_f(y)$ which corresponds to this combination, and then
- we find a term from natural language which is the best in describing the 1-D membership function $\mu_f(y)$.

5 The Main Problem of Shadow Theory and What We Are Planning To Do

In short, Zadeh's idea is to describe the original multi-D membership function by several 1-D membership functions corresponding to different combinations, and then describe each of this functions by a term from natural language. As a result, we get a natural-language description of the original fuzzy information, e.g., about (y_1,y_2), as a collection of statements of the type: y_1 is large; y_2 is small; y_1+y_2 is medium; etc.

The important question is: do we lose any information in this translation?

A similar question appears in *tomography*, where we reconstruct the image from sections.

Zadeh has shown that under certain conditions, the original fuzzy set can be uniquely reconstructed from its shadows. In this paper, we prove that for appropriately chosen shadows, the reconstruction is *always* unique. Thus, the translation from the original membership function into the linguistic terms which describe different combinations y is lossless.

6 The Idea of a Shadow Reformulated in Terms of Sets

In order to formulate and prove our result, we will first recall and use some basic definitions. The membership function $\mu_f(y)$ corresponding to the shadow can be described by the extension principle:

$$\mu_f(y)=\max_{\vec{y}:\ f(\vec{y})=y}\mu(\vec{y}). \tag{3}$$

We would like to somewhat simplify this formula. Namely, one can easily see that for any value y, the resulting value $\mu_f(y)$ does not use all the information about the combination function $f(y_1,\dots,y_m)$; it only uses the level set $\{\vec{y}:\ f(\vec{y})=y\}$ corresponding to this combination function. Thus, if any other function $g(\vec{y})$ has the same level set for this y, we will get $\mu_f(y)=\mu_g(y)$. To describe this fact, we can reformulate the formula (3) as a two-step procedure:

- first, for each combination function $f(\vec{y})$ and value y, we form a set

$$S=\{\vec{y}\,|\,f(\vec{y})=y\}; \tag{4}$$

- second, for every such set S, we compute the value

$$\mu(S)=\max_{\vec{y}\in S}\mu(\vec{y}). \tag{5}$$

The formula (5) has a natural interpretation in terms of fuzziness and possibility theory: namely, if we interpret $\mu(\vec{y})$ as the degree of possibility that the "actual" (unknown) value $\vec{y}_{\rm act}$ coincides with the given $\vec{y}$, then $\mu(S)$ describes the degree of possibility that the actual value $\vec{y}_{\rm act}$ is in S. Indeed, $\vec{y}_{\rm act}$ is in S if and only if it coincides with one of the vectors $\vec{y}\in S$. Thus, the degree to which it is possible that $\vec{y}_{\rm act}$ is in S can be computed as a degree of possibility of the following statement:

$$\exists\vec{y}_{\vec{y}\in S}(\vec{y}=\vec{y}_{\rm act}).$$

We assumed that the degree of possibility of each statement $\vec{y} = \vec{y}_{\rm act}$ is equal to $\mu(\vec{y})$; the existential quantifier is, in essence, an infinite "or" operation, so we can use the simple fuzzy "or" operation max to describe it. Therefore, we get the formula (5).

What did we gain by this reformulation?

- In the original Zadeh's formulation, we fix a class of functions $f(\vec{y})$, and translate the original membership function $\mu(\vec{y})$ into several membership functions $\mu_f(y)$ corresponding to different functions from this class. In these terms, the main question is as follows: if we know all these functions $\mu_f(y)$, can we reconstruct the original function $\mu(\vec{y})$ uniquely?

- In the new formulation, instead of fixing a class of functions, we fix a class $\mathcal{S}$ of *sets*. For each set $S \in \mathcal{S}$, we have the degree of possibility $\mu(S)$ of this set which is determined by the formula (5). In these terms, the above question takes the following form: if we know $\mu(S)$ for all $S \in \mathcal{S}$, can we uniquely reconstruct the original fuzzy set (membership function) $\mu(\vec{y})$?

7 Reconstructing a Fuzzy Set From Its Shadows: Heuristic Idea

In the previous text, we gave an intuitive explanation for the formula (5) which describe the degree $\mu(S)$ in terms of the values $\mu(\vec{y})$. Similar arguments can describe $\mu(\vec{y})$ in terms of $\mu(S)$. Indeed, suppose that we know, for each set $S \in \mathcal{S}$, the degree to which it is possible that the actual vector $\vec{y}_{\rm act}$ is in this set S. A vector $\vec{y}$ is possible if all sets which contain $\vec{y}$ are possible. Thus, the degree to which it is possible that $\vec{y}_{\rm act}$ is equal to $\vec{y}_{\rm act}$ can be computed as a degree of possibility of the following statement:

$$\forall S_{\vec{y} \in S} (\vec{y}_{\rm act} \in S).$$

We assumed that the degree of possibility of each statement $\vec{y}_{\rm act} \in S$ is equal to $\mu(S)$; the universal quantifier is, in essence, an infinite "and" operation, so we can use the simple fuzzy "and" operation min to describe it. Therefore, we get the following heuristic formula:

$$\mu(\vec{y}) = \min_{S : \vec{y} \in S} \mu(S). \tag{6}$$

The question is: when is this heuristic formula correct? We will describe our results in the following two sections.

8 First Result: For a Convex Fuzzy Set, Linear Shadows Reconstruct It Uniquely

In 1-D case, a fuzzy set is called *convex* if the corresponding membership function is continuous, and for every $\alpha > 0$, the corresponding α-cut is bounded and convex. We can use a similar definition in multi-D case:

Definition 1. *A fuzzy set $\mu : R^m \to [0,1]$ is called convex if the membership function μ is continuous, and for every $\alpha > 0$, its α-cut $\{\vec{y} \,|\, \mu(\vec{y}) \ge \alpha\}$ is bounded and convex.*

Theorem 1. *A convex fuzzy set can be uniquely reconstructed from its linear shadows.*

Comments.

- For reader's convenience, all the proofs are placed in a special (last) section.

- As we will see from the proof, not only is reconstruction possible, but this reconstruction can be done by using the formula (6).

9 Second Result: For a General Fuzzy Set, Quadratic Shadows Reconstruct It Uniquely

Theorem 2. *There exists a fuzzy set which cannot be uniquely reconstructed from its linear shadows.*

Since linear functions are not enough, the natural next step is to use quadratic functions. This is already sufficient:

Theorem 3. *A fuzzy set can be uniquely reconstructed from its quadratic shadows.*

Comment. Similarly to Theorem 1, not only is reconstruction possible, but this reconstruction can be done by using the formula (6).

10 Proofs

10.1 Proof of Theorem 1

Let us show that this reconstruction can be done by using the formula (6). If $\vec{y} \in S$, then, due to formula (5), we have $\mu(\vec{y}) \le \mu(S)$. Thus,

$$\mu(\vec{y}) \le \min_{S:\vec{y}\in S} \mu(S). \tag{7}$$

So, to complete our proof, it is sufficient to show that we cannot have

$$\mu(\vec{y}) < \min_{S:\vec{y}\in S} \mu(S). \tag{8}$$

We will prove this impossibility by reduction to a contradiction. Assume that (8) is true for some $\vec{y}^{(0)}$. Let us denote $\mu(\vec{y}^{(0)})$ by α, and

$$\min_{S:\vec{y}^{(0)}\in S} \mu(S)$$

by $\beta > \alpha$. Let us compute the value $\gamma = (\alpha+\beta)/2$. Then, $\alpha < \gamma < \beta$. Since μ is a membership function of a convex fuzzy set, its γ-cut F_γ is closed and convex. For the point $\vec{y}$, the value of the membership function is $\alpha < \gamma$ and therefore, this point is outside the closed convex γ-cut F_γ. Therefore, by the known properties of convex sets (see, e.g., [3]), there exists a hyperplane S_0 which contains $\vec{y}^{(0)}$ and which is completely outside F_γ. Since S_0 is outside $F_\gamma = \{\vec{y}\,|\,\mu(\vec{y}) \ge \gamma\}$, for all points $\vec{y} \in S$, we have $\mu(\vec{y}) < \gamma$, and therefore,

$$\mu(S_0) = \max_{\vec{y}\in S_0} \mu(\vec{y}) \le \gamma. \tag{9}$$

Since $\vec{y} \in S_0$, we can conclude that

$$\beta = \min_{S:\vec{y}^{(0)}\in S} \mu(S) \le \mu(S_0) \le \gamma,$$

which contradicts to the fact that $\beta > \gamma$. This contradiction shows that the inequality (8) is impossible and therefore, that (6) is indeed true. The theorem is proven.

10.2 Proof of Theorem 2

For simplicity, let us consider 2-D case, in which $\vec{y} = (y_1, y_2)$. Let μ_1 be a membership function which is a characteristic function of a closed unit disk, i.e., $\mu_1(\vec{y}) = 1$ if $(\vec{y})^2 \le 1$, and $\mu_1(\vec{y}) = 0$ otherwise. Let μ_2 be a membership function which is a characteristic function of a unit circle, i.e., $\mu_1(\vec{y}) = 1$ if $(\vec{y})^2 = 1$, and $\mu_1(\vec{y}) = 0$ otherwise. In a 2-D case, to describe linear shadows, it is sufficient to describe the values $\mu(S)$ for all straight lines. We will show that for each straight line S, $\mu_1(S) = \mu_2(S)$. Indeed, by definition, $\mu_1(S)$ is equal to 1 if S contains points from a disk (i.e., intersects with a disk), and to 0 else. Similarly, $\mu_2(S)$ is equal to 1 if S intersects with a circle, and 0 else.

- If a straight line intersects with a circle, then, of course, it intersects with a disk.
- On the other hand, if a straight line (which is infinite) intersects with a disk, it cannot stay within this bounded disk and it has to go outside somewhere. So, if a straight line intersects with a unit disk, it must also intersect with a unit circle.

In other words, a straight line S intersects with a disk if and only if it intersects with a unit circle. Thus, for every straight line S, $\mu_1(S) = \mu_2(S)$. However, $\mu_1 \ne \mu_2$. The theorem is proven.

10.3 Proof of Theorem 3

The proof of Theorem 2 is very straightforward. Namely, let us show fix a vector $\vec{y}^{(0)}$ and show how we can reconstruct the value $\mu(\vec{y}^{(0)})$ from the values $\mu(S)$ known for different quadratic sets S, i.e., for sets

$$S = \{\vec{y} \mid f(\vec{y}) = y\}$$

corresponding to quadratic functions $f(\vec{y})$. Indeed, as a particular example of such a quadratic set, we can take a set $S_0 \in \mathcal{S}$ corresponding to the distance function $f(y_1, \ldots, y_m) = (y_1 - y_1^{(0)})^2 + \ldots + (y_m - y_m^{(0)})^2$ and to $y = 0$. This set S_0 consists of a single point $\vec{y}^{(0)}$ and therefore, for this set S_0, the formula (5) leads to $\mu(S_0) = \mu(\vec{y}^{(0)})$. So, for each $\vec{y}^{(0)}$, the value $\mu(\vec{y}^{(0)})$ can indeed be uniquely reconstructed from the values $\mu(S)$.

Let us now show that this reconstruction can be done by using the formula (6). We have already proven, in the proof of Theorem 1, that the inequality (7) always holds. On the other hand, we have shown that for some S_0 for which $\vec{x} \in S_0$, we have $\mu(\vec{y}) = \mu(S_0)$; thus,

$$\mu(\vec{y}) = \mu(S_0) \geq \min_{S:\vec{y}\in S} \mu(S). \tag{10}$$

Combining the two inequalities (7) and (10), we get the desired formula (6). The theorem is proven.

Acknowledgments

This work was supported in part by NASA under cooperative agreement NCC5-209, by NSF grants No. DUE-9750858 and CDA-9522207, by United Space Alliance, grant No. NAS 9-20000 (PWO C0C67713A6), by the Future Aerospace Science and Technology Program (FAST) Center for Structural Integrity of Aerospace Systems, effort sponsored by the Air Force Office of Scientific Research, Air Force Materiel Command, USAF, under grant number F49620-95-1-0518, and by the National Security Agency under Grant No. MDA904-98-1-0561 and MDA904-98-1-0564.

The authors are thankful to Lotfi Zadeh who attracted their attention to his paper on shadows on fuzzy sets, and to Abe Kandel for fruitful discussions.

References

[1] G. Klir and B. Yuan, *Fuzzy sets and fuzzy logic: theory and applications,* Prentice Hall, Upper Saddle River, NJ, 1995.

[2] H. T. Nguyen and E. A. Walker, *A first course in fuzzy logic,* CRC Press, Boca Raton, Florida, 1999.

[3] R. T. Rockafeller and T. R. Rockafeller, *Convex analysis*, Princeton Univ. Press, Princeton, 1970.

[4] L. A. Zadeh, "Shadows of fuzzy sets", *Problems in Transferring of Information*, 1966, No. 2, pp. 37–44 (in Russian); Engl. translation in: G. J. Klir and B. Yuan (eds.), *Fuzzy Sets, Fuzzy Logic, and Fuzzy Systems. Selected Papers by Lotfi A. Zadeh*, World Scientific, Singapore, 1996, pp. 51–59.