SHADOWS OF FUZZY SETS

by

L. A. Zadeh

Memorandum No. ERL-M336

March 1966

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720

Published in Problems in Transmission of Information, March 1966. Vol. 2, pp. 37-44. (In Russian)

Shadows of Fuzzy Sets

L. A. Zadeh *

In information theory as well as in many other fields of science it is customary to treat uncertainty and imprecision through the concepts and methods of probability theory. The almost exclusive reliance on probability theory for this purpose obscures the fact that there are many situations in which the source of imprecision is not a random variable but a class or classes which do not possess sharply defined boundaries. For example, the "class" of real numbers which are much greater than 10 is clearly not a precisely defined set of objects. The same is true of the "class" of good strategies in a game, the "class" of handwritten characters representing the letter A, the "class" of intelligent men, the "class" of systems which are approximately equivalent to a specified system, etc. In fact, on closer examination it appears that most of the classes of objects encountered in the real world are of this fuzzy, not sharply defined, type. In such classes, an object need not necessarily either belong or not belong to a class; there may be intermediate grades of membership. Thus, to describe the degree of belonging to such a class requires the use of a multivalued logic with a possibly continuous infinity of truth values.

In a recent paper,[1] a conceptual framework for dealing with classes in which there may be grades of membership intermediate between full membership and non-membership was outlined. A central concept in this framework is that of a fuzzy set, that is, a "class" with a possibly continuous infinity of grades of membership in it. More specifically, let X be a

---

Department of Electrical Engineering and Computer Sciences, University of California, Berkeley, California. This work was supported in part by the National Science Foundation under Grant GP-2413.

collection of objects (points) with a generic object (point) denoted by x. Thus, $X = \{x\}$. Then, a <u>fuzzy set</u> A in X is characterized by a membership function, $\mu_A(x)$ (or simply μ_A), which assigns to each x a number in the interval [0,1] which represents the grade of membership of x in A. Thus, the nearer the value of $\mu_A(x)$ to unity, the higher the grade of membership of x in A; and conversely, the smaller the value of $\mu_A(x)$, the lower the grade of membership of x in A.

Consider, for example, the fuzzy set in E^1: $A = \{x | x \gg 10\}$. In this case, A may be characterized - subjectively, of course - by a membership function μ_A whose typical values may be: $\mu_A(10) = 0$; $\mu_A(50) = 0.3$; $\mu_A(100) = 0.9$; $\mu_A(200) = 1$, etc.

The concept of a fuzzy set provides a natural way of formulating the problem of abstraction - a problem which plays a central role in pattern classification, heuristic programming and many other fields.[2] Specifically, assume that we are given a finite number of samples from a fuzzy set A, e.g., N pairs of the form $(x_1, \mu_A(x_1)), \ldots, (x_n, \mu_A(x_n))$, where x_i, $i = 1, \ldots, n$, is a point in X and $\mu_A(x_i)$ is its grade of membership in A. Then, an abstraction on this set of samples consists in estimating the membership function μ_A of A from the sample values $(x_1,\mu_A(x_1)), \ldots, (x_n,\mu_A(x_n))$. In the form stated, this is not, of course, a mathematically well-posed problem, since there is no provision for assessing the goodness of the estimate of μ_A. To make abstraction mathematically meaningful, it is necessary to have some a priori information about the class of functions to which μ_A belongs and have some way of comparing μ_A with its estimate. What is disconcerting about the problem of abstraction is that the human mind can perform abstraction very effectively even when the problems involved are not mathematically well defined. It is this lack of understanding of abstraction processes and our consequent inability to instruct machines on how to perform them that is at the center of many unresolved problems in heuristic programming, pattern classification and related problem areas.

It should be noted that, from the point of view of fuzzy sets, when one says, for example, "Eugene is a tall man," one gives a pair (Eugene, μ_A) where A is the fuzzy set of tall men and μ_A is its membership function. Usually, $\mu_A(x)$ would be known only for a finite set of sample men and thus, in general, one would have to perform an abstraction to estimate μ_A.

As a first step toward the development of systematic techniques for performing abstraction on finite samples from fuzzy sets, it is necessary to construct a mathematical framework for manipulating fuzzy sets and studying their properties. In this note, we shall concern ourselves with one particular aspect of such sets, namely, the notion of a <u>shadow</u> of a fuzzy set and certain properties related to the dual notions of convexity and concavity. Although the theory of fuzzy sets appears to have considerable relevance to problems in pattern classification, optimization under fuzzy constraints and transmission of information, we shall not touch upon these and other applications in the present discussion.

In order to make our discussion self-contained, we summarize below some of the basic definitions relating to fuzzy sets.

1. Two fuzzy sets A and B are <u>equal</u>, written as A = B, if and only if $\mu_A(x) = \mu_B(x)$ for all x. In the sequel, this relation will be written as $\mu_A = \mu_B$, with the understanding that the suppression of x indicates that the equality holds for all x. The same convention will be used in all cases where an equality or inequality holds for all x.

2. The <u>complement</u> of a fuzzy set A is a fuzzy set A' whose membership

function is expressed by

(1) $$\mu_{A'} = 1 - \mu_A$$

3. A fuzzy set A is contained in a fuzzy set B, written as $A \subset B$, if and only if

(2) $$\mu_A \leq \mu_B$$

4. The union of two fuzzy sets A and B is denoted by $A \cup B$ and is defined as the smallest fuzzy set containing both A and B. The membership function of $A \cup B$ is given by

(3) $$\mu_{A \cup B}(x) = \mathrm{Max}\ [\mu_A(x), \mu_B(x)]$$

5. Similarly, the intersection of two fuzzy sets A and B is denoted by $A \cap B$ and is defined as the largest fuzzy set contained in both A and B. The membership function of $A \cap B$ is expressed by

(4) $$\mu_{A \cap B}(x) = \mathrm{Min}\ [\mu_A(x), \mu_B(x)]$$

6. The union and intersection of A and B are special cases of the convex combination of A, B and a third fuzzy set Λ. Specifically, the membership function of the convex combination - denoted by $(A, B: \Lambda)$ - of A, B and Λ is expressed by

(5) $$\mu_{(A,B;\ \Lambda)}(x) = \mu_\Lambda(x)\ \mu_A(x) + (1 - \mu_\Lambda(x))\ \mu_B(x).$$

In what follows, we assume for concreteness that X is a real Euclidean n-space E^n. In such a space, a fuzzy set A is convex if and only if the sets $\Gamma_\alpha = \{x | \mu_A(x) \geq \alpha\}$ are convex for all $\alpha > 0$. Equivalently, A is convex if and only if the inequality

(6) $$\mu_A(\lambda x_1 + (1 - \lambda) x_2) \geq \mathrm{Min}\ [\mu_A(x_1), \mu_A(x_2)]$$

holds for all x_1, x_2 in E^n and all λ in the interval [0,1].

A fuzzy set A is concave if A is the complement of a convex set. This implies that for a concave set the inequality (6) is replaced by the dual inequality

$$(7) \quad \mu_A (\lambda x_1, + (1 - \lambda) x_2) \leq \text{Max} [\mu_A (x_1), \mu_A (x_2)]$$

It is easy to show [1] that convexity is preserved under intersections. Dually, concavity is preserved under unions.

The <u>convex hull</u> of A is denoted by convA and is defined as the smallest convex fuzzy set containing A. Similarly, the <u>concave core</u> of A is denoted by concA and is defined as the largest concave fuzzy set contained in A.

We are now ready to define the notion of a shadow of a fuzzy set and examine some of its basic properties.

Let p_0 and H be, respectively, a point and a hyperplane in E^n. Then, a <u>point shadow</u> of A on H is a fuzzy set S(A) in H whose membership function $\mu_{S(A)}(x)$ is defined as follows: Let L be a line passing through p_0, with L intersecting H at a point h. Then,

$$(8) \qquad \mu_{S(A)}(h) = \sup_{x \in L} \mu_A(x) \quad , h \in H$$

$$\mu_{S(A)}(x) = 0 \quad , x \notin H$$

Note that we use the suggestive term "point-shadow" to describe this fuzzy set because it bears resemblance to the shadow thrown by a cloud A on a plane H, with p_0 acting as a point source of light.

Dual to the notion of a point-shadow is that of a <u>complementary point-shadow</u>, C(A), which is defined as the complement of $S(A')$ on H, where A' is the complement of A. More explicitly,

$$(9) \qquad \mu_{C(A)}(h) = \text{Inf}\ \mu_{(A)}(x) \quad , h \in H$$

$$\mu_{C(A)}(x) = 0 \quad , x \notin H$$

The transformation S which takes A into S(A) will be referred to as <u>point-projection</u> of A on H <u>with respect to</u> p_0. In the special case where

p_0 is a point at infinity and the lines L are orthogonal to H, we shall refer to S(A) and S as <u>orthogonal shadow</u> and <u>orthogonal projection</u>, respectively. For example if H is the coordinate plane $H = \{x | x_1 = 0\}$, $x = (x_1, \ldots, x_n)$, then the orthogonal shadow of A on H is characterized by the membership function

$$(10) \qquad \mu_{S(A)}(x_2, \ldots, x_n) = \sup_{x_1} \mu_A(x_1, \ldots, x_n) \quad , \; x \varepsilon H$$
$$= 0 \quad , \quad x \notin H$$

In the sequel, we shall frequently use the terms shadow and projection without the adjective "point" or "orthogonal", relying on the context to indicate the specific meaning in which these terms should be understood.

We proceed to establish several basic properties of shadows and complementary shadows of fuzzy sets. Most of these properties are immediate consequences of the defining relations (8) and (9).

<u>Homogeneity</u>. Let kA denote a fuzzy set whose membership function is given by

$$(11) \qquad \mu_{kA}(x) = k\mu_A(x)$$

where k is a constant, $0 \leq k \leq 1$. Then clearly

$$(12) \qquad S(kA) = kS(A)$$

<u>Monotonicity</u>. This property is expressed by the relation

$$(13) \qquad A \subset B \Rightarrow S(A) \subset S(B)$$

and is an immediate consequence of

$$(14) \qquad \forall x \, [\mu_A(x) \leq \mu_B(x)] \Rightarrow \sup_{x \varepsilon L} \mu_A(x) \leq \sup_{x \varepsilon H} \mu_B(x)$$

<u>Distributivity</u>. For any fuzzy sets A and B, we have

$$(15) \qquad S(A \cup B) = S(A) \cup S(B)$$

which implies that S is distributive with respect to U. This follows at once from the identity

$$(16) \quad \underset{x\varepsilon L}{\mathrm{Sup}}\ \mathrm{Max}\ [\mu_A(x),\ \mu_B(x)] = \mathrm{Max}\ [\underset{x\varepsilon L}{\mathrm{Sup}}\ \mu_A(x),\ \underset{x\varepsilon L}{\mathrm{Sup}}\ \mu_B(x)]$$

In connection with (15), it is natural to raise the question: Is S distributive with respect to $\cap$, that is, is it true that

$$(17) \qquad S(A \cap B) = S(A) \cap S(B).$$

In this case, the corresponding relation in terms of membership functions reads

$$(18) \quad \underset{x\varepsilon L}{\mathrm{Sup}}\ \mathrm{Min}\ [\mu_A(x),\ \mu_B(x)] = \mathrm{Min}\ [\underset{x\varepsilon L}{\mathrm{Sup}}\ \mu_A(x),\ \underset{x\varepsilon L}{\mathrm{Sup}}\ \mu_B(x)]$$

This equality does not hold for arbitrary μ_A and μ_B. However, it can be made valid by suitably restricting $\mu_A(x)$ and $\mu_B(x)$, as in the minimax theorem.[3] For arbitrary μ_A and μ_B, one can assert that

$$(19) \qquad S(A \cap B) \subset S(A) \cap S(B)$$

since (see (43) et seq.)

$$(20) \quad \underset{x\varepsilon L}{\mathrm{Sup}}\ \mathrm{Min}\ [\mu_A(x),\ \mu_B(x)] < \mathrm{Min}\ [\underset{x\varepsilon L}{\mathrm{Sup}}\ \mu_A(x),\ \underset{x\varepsilon L}{\mathrm{Sup}}\ \mu_B(x)]$$

Note that by combining (11) and (15), we have for any constants k_1 and k_2 in [0,1],

$$(21) \qquad S(k_1A \cup k_2B) = k_1S(A) \cup k_2S(B)$$

This identity indicates that S is a linear transformation, with the restriction that $k_1, k_2 \in [0,1]$. Note also that S is idempotent, i.e., $S^2(A) = S(S(A)) = S(A)$.

Invariance of convexity and concavity under projections. Let A be a convex fuzzy set in E^n and let S(A) be an orthogonal shadow of A on a hyperplane H. Then S(A) is a convex fuzzy set in H. Dually, if A is concave, then so is C(A) (complementary shadow of A).

Proof. It will suffice to prove the assertion relating to convexity. To this end, let h_1 and h_2 be two arbitrary points in H and let h be any point in H defined by

$$h = \lambda\, h_1 + (1 - \lambda)h_2 \tag{22}$$

with $\lambda \in [0,1]$. Let L_1, and L_2 and L be lines orthogonal to H and passing through h_1, h_2 and h respectively.

By the definition of Sup, for every $\varepsilon > 0$ there will be at least two points x_1 and x_2 in L_1 and L_2 such that

$$\sup_{x\in L_1} \mu_A(x) - \mu_A(x_1) \leq \varepsilon \tag{23}$$

$$\sup_{x\in L_2} \mu_A(x) - \mu_A(x_2) \leq \varepsilon \tag{24}$$

Now, since A is a convex fuzzy set, we have

$$\mu_A(\lambda x_1 + (1 - \lambda)\, x_2) \geq \text{Min}\,[\mu_A(x_1), \mu_A(x_2)] \tag{25}$$

which in view of (23) and (24) implies

$$\mu_A(\lambda x_1 + (1 - \lambda)x_2 \geq \text{Min}\,[\sup_{x\in L_1} \mu_A(x), \sup_{x\in L_2} \mu_A(x)] - \varepsilon \tag{26}$$

Noting that

$$\mu_{S(A)}(h) = \sup_{x\in L} \mu_A(x) \geq \mu_A(\lambda x_1 + (1 - \lambda)x_2) \tag{27}$$

and

$$\mu_{S(A)}(h_1) = \sup_{x\in L_1} \mu_A(x) \tag{28}$$

$$\mu_{S(A)}(h_2) = \sup_{x\in L_2} \mu_A(x) \tag{29}$$

we can infer from (26)

$$\mu_{S(A)}(h) \geq \text{Min}\,[\mu_{S(A)}(h_1), \mu_{S(A)}(h_2)] - \varepsilon \tag{30}$$

for all $\varepsilon > 0$, and hence

(31) $$\mu_{S(A)}(h) \geq \text{Min}\,[\mu_{S(A)}(h_1), \mu_{S(A)}(h_2)],$$

which demonstrates that S(A) is a convex fuzzy set in H. Q.E.D.

Bounds in terms of shadows and complementary shadows. The shadows and complementary shadows of A on a set of hyperplanes provide an obvious means of placing upper and lower bounds on A. Such bounds are useful when A has to be estimated from the knowledge of its shadows, as is frequently the case in problems involving optimization under fuzzy constraints.

Specifically, let A be a fuzzy set in E^n with membership function $\mu_A(x_1 \ldots, x_n)$ and let H_i denote the i^{th} coordinate hyperplane $H_i = \{x | x_i = 0\}$, $i = 1, \ldots, n$. Let S_i and C_i denote, respectively, the shadow and complementary shadow of A on H_i, with the membership functions of S_i and C_i given by

(32) $$\begin{aligned}\mu_{S_i}(x) &= \sup_{x_i} \mu_A(x_1, \ldots, x_n) \quad &, x \in H_i \\ &= 0 \quad &, x \notin H_i\end{aligned}$$

(33) $$\begin{aligned}\mu_{C_i}(x) &= \inf_{x_i} \mu_A(x_1, \ldots, x_n) \quad &, x \in H_i \\ &= 0 \quad &, x \notin H_i\end{aligned}$$

Consider now cylindrical fuzzy sets $\bar{S}_i$ and $\bar{C}_i$ generated by S_i and C_i via the membership functions

(34) $$\mu_{\bar{S}_i}(x) = \sup_{x_i} \mu_A(x_1, \ldots, x_n)$$

(35) $$\mu_{\bar{C}_i}(x) = \inf_{x_i} \mu_A(x_1, \ldots, x_n)$$

In terms of these cylindrical fuzzy sets, A can be bounded from above and below by the intersection of the $\bar{S}_i$ and the union of the $\bar{C}_i$, $i = 1, \ldots, n$.

Thus

(36) $$\bigcup_{i=1}^{n} \bar{C}_i \subset A \subset \bigcap_{i-1}^{n} \bar{S}_i$$

This relation is an immediate consequence of the inequalities

(37) $$\inf_{\substack{x_i \\ i-1,\ldots,n.}} \mu_A(x_1, \ldots ,x_n) \leq \mu_A(x_1, \ldots ,x_n) \leq \sup_{x_i} \mu_A(x_1, \ldots ,x_n).$$

When A is a convex or concave fuzzy set and the H_i constitute the set of all hyperplanes in E^n, the inequalities in (36) can be replaced by equalities. More concretely, we can assert that: If A and B are convex sets and S(A) = S(B) for all p_0 (and a fixed H), then A = B. (Dually, the same conclusion holds for concave sets and complementary shadows.)

Proof. It will be sufficient to show that if A ≠ B, then there exists a p_0 such that S(A) ≠ S(B).

Assuming that A ≠ B, let x_0 be a point at which $\mu_A(x_0) \neq \mu_B(x_0)$, e.g., for concreteness, $\mu_A(x_0) = \alpha > \mu_B(x_0) = \beta$. Since B is a convex set, the set $\Gamma_\beta = \{x | \mu_B(x) > \beta\}$ is a convex set and hence there exists a hyperplane F supporting Γ_β and passing through x_0. In relation to F, we have $\mu_B(x) \leq \beta$ for all x on F and on the side of F not containing Γ_β.

Now let p_0 be an arbitrarily chosen point on F, and let L be a line passing through p_0 and x_0. At the intersection, h, of this line with H (which may be at infinity), we have

$$\mu_{S(B)}(h) \leq \beta$$

but on the other hand $\mu_{S(A)}(h) \geq \alpha$ since $\mu_A(x_0) = \alpha$. Consequently, $\mu_{S(A)}(h) \neq \mu_{S(B)}(h)$. Q.E.D.

In the case of orthogonal shadows, the statement of the property in question becomes: If A and B are convex sets and S(A) = S(B) for all H, then A = B. More generally, if A and B are not necessarily convex, then the conclusion A = B would be replaced by the weaker equality conv A = conv B, where conv A and conv B denote the convex hulls of A and B, respectively.

<u>Degree of separability</u>. In [1], the classical separation theorem for convex sets was generalized to convex fuzzy sets in the following manner. Let A and B be two convex fuzzy sets in E^n and let M be the maximal grade in the intersection of A and B, i.e.,

$$M = \operatorname*{Sup}_{x} \operatorname{Min} [\mu_A(x), \mu_B(x)] \tag{38}$$

Then, (a) there exists a hyperplane H such that $\mu_A(x) \leq M$ for all x on one side of H and $\mu_B(x) \leq M$ for all x on the other side of H; and (b) there does not exist a number $M' < M$ for which this is true. For this reason, the number D = 1-M is called the degree of separability of A and B.

In connection with applications to pattern classification, it is of interest to inquire if the degree of separability can be increased by projecting A and B on a hyperplane. The answer to this question can readily be shown to be in the negative.

For simplicity, let n = 2 and consider the shadows of convex fuzzy sets A and B on the hyperplane $\{x | x_2 = 0\}$. By (10), the membership functions of these convex shadows on the hyperplane $\{x | x_2 = 0\}$ are given by

$$\mu_{S(A)}(x_1) = \operatorname*{Sup}_{x_2} \mu_A(x_1, x_2) \tag{39}$$

$$\mu_{S(B)}(x_1) = \operatorname*{Sup}_{x_2} \mu_B(x_1, x_2) \tag{40}$$

Now the degree of separability for A and B can be expressed as

$$D = 1 - \operatorname*{Sup}_{x_1} \operatorname*{Sup}_{x_2} \operatorname{Min} [\mu_A(x_1, x_2), \mu_B(x_1, x_2)] \tag{41}$$

whereas the degree of separability for S(A) and S(B) is given by

$$D_s = 1 - \operatorname*{Sup}_{x_1} \operatorname{Min} [\operatorname*{Sup}_{x_2} \mu_A(x_1, x_2) \;,\; \operatorname*{Sup}_{x_2} \mu_B(x_1, x_2)] \tag{42}$$

Thus, to show that $D \leq p$ it suffices to show that, for all x_1,

(43) $$\operatorname{Sup}_{x_2} \operatorname{Min} [\mu_A(x_1, x_2), \mu_B(x_1, x_2)] \leq \operatorname{Min} [\operatorname{Sup}_{x_2} \mu_A(x_1, x_2), \operatorname{Sup}_{x_2} \mu_B(x_1, x_2]$$

This inequality follows at once by noting that the inequalities

(44) $$\operatorname{Sup}_{x_2} \mu_A(x_1, x_2) \geq \mu_A(x_1, x_2) \qquad \text{for all } x_1$$

and

(45) $$\operatorname{Sup}_{x_2} \mu_B(x_1, x_2) \geq \mu_B(x_1, x_2) \qquad \text{for all } x_1$$

imply that for all x_1

(46) $$\operatorname{Min} [\operatorname{Sup}_{x_2} \mu_A(x_1, x_2), \operatorname{Sup}_{x_2} \mu_B(x_1, x_2)] \geq \operatorname{Min} [\mu_A(x_1, x_2), \mu_B(x_1, x_2)]$$

and hence imply (43).

Concluding remarks. Essentially, the notion of the shadow of a fuzzy set plays the same role in the theory of fuzzy sets as the notion of a marginal distribution plays in the theory of probability. In this note, we touched only upon some of the more elementary properties of shadows of fuzzy sets and did not concern ourselves with applications. Although work on these is still in its preliminary stages, it appears that the concept of a shadow along with some of the other notions sketched in this note may have useful applications in several areas, particularly in pattern classification and optimization under fuzzy constraints.

References

1. L. A. Zadeh, "Fuzzy Sets," Information and Control, Vol. 8, pp. 338-353, June, 1965.

2. R. Bellman, R. Kalaba and L. A. Zadeh, "Abstraction and Pattern Classification," Journal of Math. Anal. and Appl., Vol. 13, pp. 1-7, January, 1966.

3. S. Karlin, Mathematical Methods and Theory in Games, Programming and Economics, Addison-Wesley Publishing Co., Inc., Reading, Mass., p. 28 et seq.