2nd International Symposium on Imprecise Probabilities and Their Applications, Ithaca, New York, 2001

Characterizing Fuzzy Measures Used in Uncertainty Representation

Ronald R. Yager
Iona College
New Rochelle, NY 10801
yager@panix.com

Abstract

We introduce the fuzzy measure and its use in representing information about uncertain variables. The relationship between the fuzzy measure and the Dempster-Shafer belief structure is discussed. A method for generating the family of fuzzy measures associated with a D–S belief structure is described. We discuss the use of the Shapley index as a means for introducing an extension of concept of entropy to fuzzy measures. We introduce the cardinality index of a fuzzy measure and use it to define the attitudinal character of a fuzzy measure.

Keywords. fuzzy measures, cardinality based measures, entropy, attitude

1. Introduction

Fuzzy measures [1-3] can be used to represent information about an uncertain variable. The fuzzy measure conveys the users "opinion" of finding the value of the variable within a subset. In providing this information the user is reflecting at least two aspects of his view of the state of the world. The first reflects the capacity of each of the allowable values for being the actual value of the variable and hence involves information related to distinctions between the possible values. In a situation in which one outcome distinguishes itself as being the only achievable value we are in a state of complete certainty. On the other hand in the situation in which no distinction can be made regarding the attainability of any of the allowable outcomes, we are in a state of complete uncertainty. This aspect, based upon information emanating from distinctions, is closely related to our assessment of the situation on a scale with a dimension indicating a certainty-uncertainty component in our knowledge. In probability theory an often used quantification of this aspect of our knowledge about a variable is the Shannon entropy. In this work we use the Shapely entropy as a way of generalizing this idea to the fuzzy measure. When using a fuzzy measure the user is often reflecting an additional aspect of their view of the world different. This aspect becomes most nakedly apparent in the case of complete uncertainty. At one extreme is the case in which one uses a fuzzy measure which assigns the value of one to each subset except the null while at the other extreme is the case of a fuzzy measure which assigns a zero measure to each of the subsets except the universal subset which must be given the value one. The choice between these two is also a reflection about the users view of the state of the world. These two extreme fuzzy measures are both manifestations of complete uncertainty and hence the Shapley entropy can't differentiate between them. In this work we introduce a index, called the cardinality index, that allows us to distinguish between these two extreme fuzzy measures. We use this cardinality index to associate with each fuzzy measure a formal concept called its attitudinal character. These notions provide an additional dimension on which to classify fuzzy measures.

2. Fuzzy Measures and the Representation of Uncertainty

A fuzzy measure [2] on a finite space X is a mapping μ: $2^X \rightarrow [0,1]$ such that $\mu(\varnothing) = 0$, $\mu(X) = 1$ and if $A \subset B$ then $\mu(A) \leq \mu(B)$. Fuzzy measures can be used in the representation of knowledge about an uncertain variable. Let V be a variable taking its value in the set X. When using a fuzzy measure to represent knowledge about V $\mu(E)$ is interpreted as what we shall call the "certitude" that the value of V is contained in the subset E. A significant feature of the use of fuzzy measures is their ability to capture in a unified framework, many of the well established uncertainty calculi.. We shall briefly indicate some of these. If we know precisely the value of V, V = x, then $\mu_X(A) = 1$ for $x \in A$ and $\mu_X(A) = 0$ for $x \notin A$. Probabilistic uncertainty is represented by a fuzzy measure in which $\mu(A \cup B) = \mu(A) + \mu(B)$ for $A \cap B = \varnothing$. Possibilistic uncertainty [4] is characterized by a fuzzy measure in which $\mu(A) = \text{Max}[\mu(A), \mu(B)]$. A fuzzy measure is called a necessity measure [4] if it satisfies the property that $\mu(A \cap B) = \text{Min}[\mu(A), \mu(B)]$. Another class of fuzzy measures are those in which our certitude that the value of a variable lies in a subset is based upon the number of elements in the subset [5], we call these cardinality based measures. For these measure $\mu(E) = D_{Card(E)}$. Because of the monotonicity of fuzzy measures $D_j \geq D_i$ for $j > i$. Furthermore, $D_0 = 0$ and $D_n = 1$. One important special case of cardinality based measures, μ_N, occurs when $D_j = \frac{j}{n}$, in this case $\mu(E) = \frac{Card(E)}{n}$. It is noted that this measure is the same as the one used to represent lack of information in the case of probabilistic uncertainty. Another special case is μ^* where $D_0 = 0$ and $D_j = 1$ for all other j. It is noted that this measure is the same as the one used to represent lack of information in the case of possibilistic uncertainty. Another special case of cardinality based measure is μ_* having $D_n = 1$ and $D_j = 0$ for all other j, this measures is the same as the one used to represent lack of information in the case of the necessity measure

3. Fuzzy Measures and Dempster-Shafer Belief Structures

A fuzzy measure μ with respect to V provides a description of our knowledge about the variable. When using a fuzzy measure, although there exists some uncertainty with respect to the actual value of the variable, there exists no uncertainty with respect to our knowledge of the description of the uncertainty. An example of this kind of additional uncertainty would be a situation in which we only knew that $\mu(A)$ lies in some interval. The Dempster-Shafer belief structure [6] provides a framework for the representation of knowledge about the value of an variable which can be used when there exists some uncertainty regarding our knowledge of the underlying fuzzy measure [7]. If V is a variable taking its value in the space X a D-S belief structure is defined by a mapping m: $2^X \rightarrow [0, 1]$ such that $m(\varnothing) = 0$ and $\sum_{B \subseteq X} m(B) = 1$. We call the subsets B_j for which $m(B_j) > 0$ the focal elements, hence condition 2 becomes $\sum_{j=1}^{q} m(B_j) = 1$.

The use of a D-S belief structure to describe our knowledge about a variable provides only "partial information" about the underlying fuzzy measure. Given a D-S belief structure there are many possible fuzzy measures that can satisfy the constraints it imposes [7]. The two most well known fuzzy measures associated with a belief structure are the plausibility and belief measures. We can view a D-S belief structure as a constraint on the set of all fuzzy measures of X generating a set of possible fuzzy measures associated with the variable. In [7] we provided a methodology for generating the set of fuzzy measures that can be associated with a belief structure. In the following we describe this methodology.

Let m be a D-S belief structure with focal elements B_1, B_2, ..., B_q. For each focal element B_j let W_j be a vector of dimension $|B_j|$ whose components, $w_j(i)$, satisfy the conditions $w_j(i) \in [0, 1]$ and $\sum_{i=1}^{|B_j|} w_j(i) = 1$. We shall call these the allocation vectors. In [7] we show that a set measure defined by $\mu(E) = \sum_{j=1}^{q} (m(B_j) \sum_{i=1}^{|B_j \cap E|} w_j(i))$ is a fuzzy measure associated with the belief structure m. In the preceding if $|B_i \cap E| = 0$ then we take the sum as zero. Thus by selecting a collection $W = \langle W_1, W_2, ..., W_q \rangle$ of allocation vectors we can define a unique fuzzy measure associated with a belief structure. We point out some special cases . If all the W_j are such that $w_j(1) = 1$ then the resulting fuzzy measure is the plausibility measure. If all the W_j are selected such $w_j(|B_j|) = 1$, then this results in the belief measure. It can be easily shown that these two measures are the extremes, that is if μ is a fuzzy measure generated from a collection of allocation vectors W then for all $A \subset X$, $Pl(A) \geq \mu(A) \geq Bel(A)$.

While we can individually select each of associated allocation vectors it is often more interesting to select all the vectors in some consistent way as in the case of plausibility and belief.. One way of selecting the allocation vectors in a consistent manner is by use of a function f: $[0, 1] \rightarrow [0, 1]$ such that: $f(0) = 0$, $f(1) = 1$ and $f(x) \geq f(y)$ if $x \geq y$. Using this type of function we can globally define all the allocation vectors W_j as follows. For each B_j we define each W_j such that $w_j(i) = f(\frac{i}{|B_j|}) - f(\frac{i-1}{|B_j|})$. An interesting special case occurs when $F(x) = x$.. Here the W_j are such that $w_j(i) = \frac{1}{|B_j|}$, the weights in each allocation vector are uniformly distributed. In this case the resulting fuzzy measure is such that $\mu(E) = \sum_{j=1}^{q} m(B_j) \frac{|B_j \cap E|}{|B_j|}$. We note in this case that $\mu(\{x_k\}) = \sum_{j \text{ s.t. } x_k \in B_j} \frac{m(B_j)}{|B_j|}$ and that $\mu(E) = \sum_{k \text{ s.t. } x_k \in E} \mu(\{x_k\})$, thus the measure generated from this linear F is a probability measure.

We now consider the fuzzy measures resulting from some special cases of D-S belief structures. Consider the Bayesian belief structure, each focal element is a singleton, $B_j = \{x_j\}$. Let W_j be any collection of allocation vectors and μ be the fuzzy measure generated by these vectors, $\mu(E) = \sum_{j=1}^{q} (m(B_j) \sum_{i=1}^{|B_j \cap E|} w_j(i))$. Since each focal is a singleton, $|B_j| = 1$, thus each W_j consists of only one component with value one, $w_j(1) = 1$ for all j.. In this case we see that $\sum_{i=1}^{|B_j \cap E|} w_j(i) = 1$ if $B_j \cap E \neq \varnothing$, if $x_j \in E$ and $\sum_{i=1}^{|B_j \cap E|} w_j(i) = 0$ if $B_j \cap E = \varnothing$, if $x_j \notin E$. Thus here we always get $\mu(E) = \sum_{j \text{ s. t. } x_j \in E} m(B_j)$ which is a probability measure. Thus this Bayesian D-S structure is only compatible with this probability measure. The obvious reason for this is that the Bayesian D-S belief is a true probability distribution and therefore precisely specifies a unique fuzzy measure.

Consider the D-S belief structure in which $B_1 = X$ and $m(X) = 1$. This essentially corresponds to a case in which we have no information about the value of the variable other than that it lies in X. In this case we just need to choose one weighting vector W whose cardinality is $|X| = n$. Using this we get $\mu(E) = \sum_{j=1}^{|X \cap E|} w(j) = \sum_{j=1}^{|E|} w(j)$. Here $\mu(E)$ is the sum of the first $|E|$ elements in the vector W.

4. Entropy of A Fuzzy Measure

In probability theory an important tool is the Shannon entropy. With the aid of this tool we are able to express some quantification of the overall uncertainty associated with a probability distribution. Entropy can be seen as providing some indication about the distinction between the different outcomes with respect to there being the actual value. In [5] Yager suggested an extension of the Shannon entropy to the fuzzy measure, he called this the Shapley entropy. The definition makes use of the Shapley index associated with a fuzzy measure [8-10]. Let μ be a fuzzy measure on the space $X = \{x_1, ..., x_n\}$. For any $x_j \in X$ we define its Shapley index S_j as

$$S_j = \sum_{k=0}^{n-1} (\gamma_k \sum_{\substack{K \subset F_j \\ |K| = k}} (\mu(K \cup \{x_j\}) - \mu(K)))$$

where K is a subset of cardinality $|K|$, $F_j = X - \{x_j\}$ and $\gamma_k = \frac{(n-k-1)!\, k!}{n!}$.

This index can basically be seen as the average increase in "certitude" obtained by adding the element x_j to a set which doesn't contain it. Thus we see that the S_j are providing some information distinguishing between the different x_j. It can be shown that for any fuzzy measure μ, it is always the case that all $S_j \in [0, 1]$ and $\sum_{j=1}^{n} S_j$.

In [5] we suggested the use of these indices in the form of an extension of the concept of entropy to fuzzy measures. In particular he defined the Shapley entropy of a fuzzy measure as $H(\mu) = - \Sigma_j S_j \ln(S_j)$. It can be shown that in the case when μ is a probability measure this reduces to the Shannon entropy, we get $S_j = p_j$ for all j and hence $H(\mu) = - \Sigma_j p_j \ln(p_j)$.

It also can be shown [5] that in the case of any cardinality based measure that $S_j = \frac{1}{n}$ and $H(\mu) = \ln(n)$. This situation further supports the appropriateness of using this measure as a generalization of entropy. For as we have indicated the entropy is essentially providing some indication about the distinction between the elements in X regarding their appropriateness of being the value for V. In the case of a cardinality based measure we have no information distinguishing between the elements in X regarding the appropriateness as a solution for V. Thus the measures μ^*, μ_*, and μ_N as well as all other cardinality based measures have the same entropy, ln(n).

An interesting special case occurs with the D-S belief structure [11].

Theorem: Assume m is a D-S belief structure with q focal elements, B_j. If μ is any fuzzy measure obtained from m by use of a collection $W = [W_1, ..., W_q]$ of allocating vectors then the associated Shapley index is

$$S_i = \sum_{j=1}^{q} \frac{m(B_j)}{|B_j|} \cdot B_j(x_i).$$

Here we used $B_j(x_i)$ to indicate the characteristic function of B_j. The implication of this theorem is that all the fuzzy measures associated with a D-S structure have the same Shapley entropy, contain the same information regarding the distinction between the different elements.

5. Cardinality Index and Attitudinal Character of a Fuzzy Measure

Consider the two fuzzy measures μ^* where $\mu^*(\emptyset) = 0$ and $\mu^*(A) = 1$ for $A \neq \emptyset$; and μ_* where $\mu_*(X) = 1$ and $\mu_*(A) = 0$ for $A \neq X$. Both these are cardinality based measures and hence correspond to a situation in which we have no information distinguishing the elements in X. These two measures are clearly different in their nature. We see that μ^* is dealing with the complete lack of knowledge in a very optimistic way, it allocates complete certitude to finding the actual value of V in any non-null subset. The measure μ_* deals with the complete uncertainty in a very pessimistic way, it allocates no certitude finding the value of V to any set except X. These two fuzzy measures clearly display polar attitudes regarding the situation when faced with lack of information. Another fuzzy measure associated with complete lack of information is μ_N, where $\mu(E) = \frac{\text{Card (E)}}{n}$.

The preceding fuzzy measures have provided illustrations of different attitudes about the nature of uncertainty. In the following we shall introduce a characterization of a fuzzy measure which allows us to quantify these differing attitudes. We begin by defining a concept which we call the *cardinality index of a fuzzy measure*.

Definition: Let μ be a fuzzy measure on the set $X = \{x_1, x_2, ..., x_n\}$ we define the C_k, k = 0 to n - 1, as

$$C_k = \lambda_k \sum_{\substack{\text{all } K \\ |K| = k}} \left(\sum_{x \notin K} (\mu(K \cup \{x\}) - \mu(K))\right)$$

where $\lambda_k = \frac{(n - k - 1)!\, k!}{n!}$. We call C_k the k^{th} cardinality index of μ.

Essentially C_k measures the average gain in certitude in going from subsets of cardinality k to cardinality k + 1 It can be shown [11] that $C_k \in [0, 1]$ for all k = 0 to n - 1 and $\sum_{k=0}^{n-1} C_k = 1$.

We now evaluate the cardinality index for the cardinality based fuzzy measure. We recall here $\mu(E) = D_{|E|}$ where $0 = D_0 \leq D_1 \leq D_2 ..., \leq D_n = 1$. We can express this in terms of a collection of weights, w_i, i = 0 to n - 1 such that $w_i = D_{i+1} - D_i$. It's easy to see that $w_i \in [0, 1]$ and $\sum_{i=0}^{n-1} w_i = 1$. We see that w_i is the incremental gain in going from a set of cardinality i to one of cardinality i + 1. The cardinality index is $C_k = \lambda_k \sum_{\substack{\text{all } K \\ |K| = k}} \left(\sum_{x \notin K} (\mu(K \cup \{x\}) - \mu(K))\right)$. However, for this type of measure $\mu(K \cup \{x\}) - \mu(K) = w_k$ for all sets of cardinality k. From this it follows that $C_k = w_k$, the cardinality index for these measures is simply the differential weights. What should be strongly pointed out here is that as opposed to the Shapley index, which makes no distinction between any cardinality based measure, this index completely distinguishes between different cardinality based measures.

In [11] Yager used the cardinality indices to provide a characterization of a fuzzy measure which he called its attitudinal character. If μ is a fuzzy measure with cardinality indices C_k we define the *attitudinal* character of μ as $\mathbf{A\text{-}C}(\mu) = \frac{1}{n-1} \sum_{k=0}^{n-1} C_k (n - k - 1)$. In order to provide a semantics for this concept of attitudinal character we consider the three special cases of cardinality based measures, μ^*, μ_* and μ_N First we obtain the cardinality indices for these three measures. For μ^*, $C_0 = 1$ and $C_k = 0$ for all $k \neq 0$. For μ_*, $C_{n-1} = 1$ and $C_k = 0$ for $k \neq n-1$. For μ_N, where $\mu_N(E) = \frac{\text{Card E}}{n}$ we we see that $C_k = \frac{1}{n}$ for all k. We can use these values to obtain the attitudinal character of these measures: $\mathbf{A\text{–}C}(\mu^*) = 1$, $\mathbf{A\text{-}C}(\mu_N) = 0.5$ and $\mathbf{A\text{–}C}(\mu_*) = 0$. Since a cardinality based measure is a representation of a situation of complete uncertainty the distinction between the different measures is purely a reflection of an attitude in the sense that the μ^* is extremely optimistic, μ_* is pessimistic and μ_N is neutral. Based on this the attitudinal character can be seen as inducing a scale on the unit interval in which **A–C**(μ) = 1 indicates an optimistic type measure, **A–C**(μ) = 0 indicates a pessimistic type measure and **A–C**(μ) = 0.5 indicates a neutral type measure.

Consider the cardinality index in the case where μ is a probability measure with probability distribution $Prob(x_i) = p_i$. We can express $C_k = \lambda_k \sum_{i=1}^{n} \left(\sum_{\substack{K \text{ s.t.} \\ |K| = k \\ K \subseteq F_i}} (\mu(K \cup \{x_i\}) - \mu(K))\right)$ where $F_i = X - \{x_i\}$.

After some algebra we get $C_k = \frac{1}{n} \sum_{i=1}^{n} \frac{(n-k-1)!\,k!}{(n-1)!} \left(\sum_{\substack{K \text{ s.t.} \\ |K| = k \\ K \subseteq F_i}} (\mu(K \cup \{x_i\}) - \mu(K))\right)$

Realizing that $\frac{(n-1)}{(n-k-1)!\,k!}$ is the number of different subsets of cardinality k we can select from F_i and with $\mu(K \cup \{x_i\}) - \mu(K) = p_i$ for all K then we see that $C_k = \frac{1}{n} \sum_{i=1}^{n} p_i$. Since $\sum_{i=1}^{n} p_i = 1$ we obtain that $C_k = \frac{1}{n}$ for all k independent of the probability distribution. Two significant observations can be made. The first is to emphasize that all probability measures have the same cardinality indices, $C_k = \frac{1}{n}$ for all k.. Thus this characterization does not distinguish between probability distributions. The cardinality index can be seen as being orthogonal to the Shapley index; the Shapley index completely distinguishes between every probability measure, while being unable to make any distinction between cardinality based measure. All cardinality based measures have $S_i = \frac{1}{n}$ for all i. The cardinality index on the other hand completely distinguishes between cardinality based measures and makes no distinction between probability measures. Thus these two indices are characterizing orthogonal aspects of the fuzzy measure.

The second observation we would like to make is related to the fact that the cardinality indices of a probability measure are $C_k = \frac{1}{n}$ for all k. This corresponds to an ***A-C*** value of 0.5 which corresponds to a neutral value on are attitudinal scale. This situation is not unexpected in that the probability distribution offers no choice in allocation, it corresponds to D-S belief structure with no uncertainty with regard to the parameters.

We see that we now have two scales on which to characterize a fuzzy measure. The first is the uncertainty scale which is generated with the aid of the Shapley index via the Shapley entropy. The second is the attitude scale which is generated with the aid of cardinality index via the attitudinal characterization.

6. References

[1]. Grabisch, M., Murofushi, T. and Sugeno, M., Fuzzy Measures and Integrals, Physica-Verlag: Heidelberg, 1999.

[2]. Murofushi, T. and Sugeno, M., "Fuzzy measures and fuzzy integrals," in Fuzzy Measures and Integrals, edited by Grabisch, M., Murofushi, T. and Sugeno, M., Physica-Verlag: Heidelberg, 3-41, 1999.

[3]. Wang, Z. Y. and Klir, G. J., Fuzzy Measure Theory, Plenum Press: New York, 1992.

[4]. Dubois, D. and Prade, H., Possibility Theory : An Approach to Computerized Processing of Uncertainty, Plenum Press: New York, 1988.

[5]. Yager, R. R., "On the entropy of fuzzy measures," IEEE Transactions on Fuzzy Sets and Systems 8, 453-461, 2000.

[6]. Smets, P. and Kennes, R., "The transferable belief model," Artificial Intelligence 66, 191-234, 1994.

[7]. Yager, R. R., "A class of fuzzy measures generated from a Dempster-Shafer Belief Structure," International Journal of Intelligent Systems 14, 1239-1247., 1999.

[8]. Murofushi, T., "A technique for reading fuzzy measures (i): the Shapely value with respect to a fuzzy measure," Proceedings Second Fuzzy Workshop, Nagaoka, Japan (in Japanese), 39-48, 1992.

[9]. Shapley, L. S., "A value for n-person games," in Contributions to Game Theory, edited by Kuhn, H. W. and Tucker, A. W., Princeton University Press: Princeton, NJ, 307-317, 1953.

[10]. Grabisch, M., "Alternative representations of discrete fuzzy measures for decision making," Int. J. of Uncertainty, Fuzziness and Knowledge-Based Systems 5, 587-607, 1997.

[11]. Yager, R. R., "On the cardinality index and attitudinal character of fuzzy measures," Technical Report# MII–2101 Machine Intelligence Institute, Iona College, New Rochelle, NY, 2001.