Knowledge-Based Systems 92 (2016) 1–8

Contents lists available at ScienceDirect

Knowledge-Based Systems

journal homepage: www.elsevier.com/locate/knosys

Uncertainty modeling using fuzzy measures

Ronald R. Yager*

Machine Intelligence Institute, Iona College, New Rochelle, NY 10801, United States

ARTICLE INFO

Article history:
Received 6 June 2015
Revised 12 August 2015
Accepted 3 October 2015
Available online 20 October 2015

Keywords:
Uncertainty
Decision-making
Information-fusion
Fuzzy measures

ABSTRACT

We consider the representation of information about the value of an uncertain variable using a monotonic set measure, fuzzy measure. We introduce a number of notable measures that are useful for the representation of this kind of information. We look at the formulation of the concept of entropy in this framework. The Choquet integral is introduced as a tool to help obtain expected value like formulations in the case of measure type uncertainty. It is shown how this can be used in decision making with alternatives having uncertain payoffs. A formulation for the variance associated with measure based uncertain information is provided. The issue of the fusion of multiple pieces of measure-represented information is investigated.

1. Introduction

The need to model uncertain information arises in many tasks in computational intelligence. Further compounding the situation, is the fact that in many current technological applications uncertain information can appear, not only in the classic probabilistic format, but can come in many different modalities. Here we investigate the use of a monotonic set measure, fuzzy measure [1,2], for the representation of uncertain information about the value of a variable. The use of a fuzzy measure to represent uncertain information provides a unified framework for the representation of many different modes of uncertain information. This is particularly useful in the task of multi-source information fusion and decision making under uncertainty. Given our large body of experience in working in the probabilistic environment it appears only prudent to try to take advantage of this as we move to the more general measure representation of uncertain information. With this in mind our focus here is on formulating some of the important concepts and methodologies used in the probabilistic environment in the framework of a measure based representation of uncertain information.

After first describing the basics of the use of a measure to model uncertain information we show how the concept of entropy [3,4] can be expressed in the case of a measure. More significantly we illustrate how the Choquet integral [5,6] can be used to move an idea similar to expected value to the measure representation environment. We show this has immediate application to the problem of decision making under measure based-uncertainty. More generally we show that the Choquet can be used to obtain an expected-like value for any function of a measure type uncertain variable. This immediately allows us to formulate the concept of variance in this measure environment. In probability theory an important idea is the probability of an event, where an event A is crisp or fuzzy subset of the domain of the uncertain variable. Here, since the idea of measure is broader than that of probability, we use the less specific terminology of "anticipation" and speak of the anticipation of an event in an analogous manner to speaking of the probability of an event. We provide a formulation for the anticipation of an event that reduces to the probability of an event in the case where the measure is a probabilistic one. We describe a method for aggregating fuzzy measures which results in fuzzy measures. This is shown to provide a basis for an approach to a measure based technology for the fusion of multi-source uncertain information.

2. Fuzzy measures for uncertainty modeling

A fuzzy measure μ on the space $X = \{x_1, \ldots, x_n\}$ is a set mapping $2^X \to [0, 1]$ having the properties that $\mu(\emptyset) = 0$, $\mu(X) = 1$ and $\mu(A) \geq \mu(B)$ if $B \subseteq A$ [1,7]. Here we shall simply use the term measure for these objects. We observe that a fuzzy measure essentially associates with subsets of the space X a value in the unit interval. We note that a measure having the property that $\mu(A) \in \{0, 1\}$ for all A is called a binary measure.

Assume V is an uncertain variable taking its value in the space X. We can represent our knowledge about the value of V using a measure μ. Here for any subset A of X, $\mu(A)$ indicates our anticipation of finding the value of V in A [7]. We see the appropriateness of the use of this measure since the properties of a measure are very natural for this type of representation. The property $\mu(\emptyset) = 0$ reflects the fact that we have zero anticipation of finding the value of V in the null set. The property $\mu(X) = 1$ reflects the fact that we have complete

* Corresponding author.
E-mail address: yager@panix.com

http://dx.doi.org/10.1016/j.knosys.2015.10.001
0950-7051/

anticipation of finding the value of V in the set X. The last condition reflects the fact that as the size of a set increases our anticipation of finding the value of V in it cannot decrease. In passing we note that our use of the term anticipation is generic for as we shall see for some measures we use a more specific term then anticipation, such as probability.

An important related concept is the dual of a measure. If μ is a measure on X then its dual $\hat{\mu}$ is also a measure on X defined so that $\hat{\mu}(A) = 1 - \mu(\bar{A})$. If $\mu(A)$ is interpreted as the anticipation of finding the value V in A then the dual is representing the anticipation of not finding the value of V in not A. We note that μ and $\hat{\mu}$ come in unique pairs since $\hat{\hat{\mu}} = \mu$. In the special case where $\mu(\bar{A}) = 1 - \mu(A)$ we see that $\hat{\mu}(A) = 1 - (1 - \mu(A)) = \mu(A)$. Thus in this situation the anticipation of finding the value of V in A is the same as not finding the value of V in not A.

Two important measures in the perspective of hard-soft information fusion are probability and possibility measures [8]. In addition to their importance in the representation of hard and soft information one important feature shared by these measures is that they are decomposable in the sense that for these measures the measure of any set A can be expressed in terms of the measures of the individual elements comprising the set A. This has the effect of greatly reducing the problem of obtaining the value of the measure for different subsets. We just need a small number of parameters.

A probability measure on the space X, often called an additive measure, is defined in terms of a collection of parameters p_i for $i = 1$ to n such that $p_i \in [0, 1]$ with $\sum_{i=1}^{n} p_i = 1$ and $\mu(\{x_i\}) = p_i$. Furthermore this measure is additive in the sense that $\mu(A) = \sum_{x_i \in A} p_i$. For a probability measure we use the more specific term probability instead of anticipation. For this measure p_i is denoted as the probability that V is x_i and $\mu(A)$ is referred to as the probability that V lies in A or more succinctly as the probability of A. We emphasize here that a probability measure only requires the n parameters p_i to completely define it.

Two important cases of probability measure are those corresponding to complete certainty and complete uncertainty. Complete certainty is the case where for some x_k, $p_k = 1$ and $p_j = 0$ for all $j \neq k$. This represents the situation where we know that $V = x_k$. It is easy to see for this type of measure $\mu(A) = 1$ for any A such that $x_k \in A$ and $\mu(A) = 0$ for any A such that $x_k \notin A$. Thus complete certainty is a binary measure. The second special case of probability measure, complete uncertainty, is one in which $p_j = 1/n$ for all x_j. Here we see that $\mu(A) = \frac{\text{Card}(A)}{n} = \frac{|A|}{n}$.

All probability measures have the very important property that $\mu(A) = 1 - \mu(\bar{A})$. We see this since $\mu(A) = \sum_{x_j \in A} p_j$ and $\mu(\bar{A}) = \sum_{x_j \notin A} p_j$ and since $A \cup \bar{A} = X$ we have $\mu(A) + \mu(\bar{A}) = 1$. From this it follows that for a probability measure $\hat{\mu}(A) = \mu(A)$, they are self-dual; the dual of A is equal to $\mu(A)$. This is a very important property and as we shall subsequently see it is especially useful in decision-making and question answering.

A possibility measure [9–11] on the space X is also defined in terms of a collection of n parameters, π_i for $i = 1$ to n such that $\pi_i \in [0, 1]$ and $\text{Max}_i[\pi_i] = 1$. For this measure $\mu(A) = \text{Max}_{x_i \in A}[\pi_i]$. Here $\mu(\{x_i\}) = \pi_i$ and it is called the possibility of x_i. For this measure π_i indicates the possibility that V is x_i. Here $\mu(A)$ indicates the possibility that V lies in A. It is simply referred to as the possibility of A. It is easy to show that for the measure $\mu(A \cup B) = \text{Max}[\mu(A), \mu(B)]$ [1].

Two important examples of possibility measures are those modeling complete certainty and complete uncertainty. For complete certainty we have for some x_k, that $\pi_k = 1$ and $\pi_j = 0$ for all $j \neq k$. This is the case where we know that $V = x_k$. It is easy to see for this type of measure $\mu(A) = 1$ for any A such that $x_k \in A$ and $\mu(A) = 0$ for any A such that $x_k \notin A$. We see that this is the same as the probability measure in the case of complete certainty. For the second special case, complete uncertainty, we have $\pi_j = 1$ for all j. From this it follows that for all $A \neq \emptyset$ we have $\mu(A) = 1$. We see this representation of complete uncertainty is different then for the case of probability measures.

We note that while the union of possibilities is compositional, $\mu(A \cup B) = \text{Max}[\mu(A), \mu(B)]$, the intersection is not compositional, $\mu(A \cap B) \leq \text{Min}[\mu(A), \mu(B)]$.

We note that the dual of a possibility measure, denoted $\hat{\mu}$, is such that $\hat{\mu}(A) = 1 - \mu(\bar{A})$, and is called a necessity measure [12]. It is easy to show that if μ is a possibility measure then $\hat{\mu}(A) \leq \mu(A)$ for all A and that $\hat{\mu}(A \cap B) = \text{Min}[\hat{\mu}(A), \hat{\mu}(B)]$. Furthermore since $\hat{\mu}(A) + \mu(\bar{A}) = 1$ and for possibility measures we have $\hat{\mu}(A) \leq \mu(A)$ we see $\mu(A) + \mu(\bar{A}) \geq 1$ for possibility measures.

If μ is a possibility measure with $\mu(\{x_j\}) = \pi_j$ and the subset $F_j = X - \{x_j\}$ then

$$\hat{\mu}(Fj) = 1 - \mu(\bar{F}_j) = 1 - m(\{x_j\}) = 1 - \pi_j.$$

Let $E = X - F$ then $E = \bigcap_{x_j \notin E} F_j$. Then $\hat{\mu}(E)\ \text{Min}[(1 - \pi_j)]_{j,\ x_j \notin E}$

Another important class of measures for the representation of information about an uncertain variable is the cardinality-based measure. Here the measure of a set just depends on the number of elements in the set, independent of which elements are in the set. Thus our anticipation of finding the value of V in the set A, $\mu(A)$, just depends on the number of elements in A. More formally, we define a cardinality-based measure using a set of parameters, $\alpha_j \in [0, 1]$ for $j = 0$ to n, such that $\alpha_0 = 0$, $\alpha_n = 1$ and $\alpha_{j+1} \geq \alpha_j$. For these measures $\mu(A) = \alpha_{|A|}$ where $|A|$ is the cardinality of A. We emphasize for this type of measure for all x_i, we have $\mu(\{x_i\}) = \alpha_1$. Thus all elements in X have the same anticipation being the value of V. We emphasize here that these measures also only require a small set of parameters for their complete modeling, the α_j.

Two important measures in this class are μ^* and μ_*. For μ^* we define $\alpha_j = 1$ for $j \neq 0$ and for μ_* we define $\alpha_j = 0$ for $j \neq 1$. Thus $\mu^*(A) = 1$ for $A \neq \emptyset$ and $\mu^*(\emptyset) = 0$ and $\mu_*(A) = 0$ for $A \neq X$ and $\mu^*(X) = 1$. We note that both of these measures are representing situations in which we have no information about the value of V other then it lies in X, however μ^* is representing it in a very optimistic way while μ_* is a more pessimistic representation. Thus in the case of cardinality-based measures the parameters are capturing some aspects of an optimism-pessimism scale.

Another special case of the cardinality-based measure is a ‘tipping’ measure. Here $\alpha_j = 0$ for $j < k$ and $\alpha_j = 1$ for $j \geq k$. Thus here $\mu(A) = 1$ if A has at least k elements and otherwise $\mu(A) = 0$. We note that while these last three examples are also binary measures, generally cardinality-based measures are not necessarily binary measures.

3. Entropy of a measure

A fundamental idea in probability theory is the concept of entropy [13]. The entropy of a probability distribution provides a measure of the overall uncertainty associated with a probability distribution. If P is a probability distribution on the space $X = \{x_1, \ldots, x_n\}$ where p_i is the probability that $V = x_i$ the Shannon entropy [4] of P is $H(P) = -\sum_{i=1}^{n} p_i \ln(p_i)$. It is well known that H(P) takes its maximal value, $\ln(n)$, for the case where $p_i = 1/n$ for all x_i. It takes its minimal value $H(P) = 0$ when P is such that there exists one x_k such that $p_k = 1$ and all other $p_j = 0$. In this case we know exactly the value of V. We see the larger the entropy the more the uncertainty, thus entropy is a measure of uncertainty.

In [14] Yager extended the idea of entropy to situations where our knowledge about the value of the variable V is expressed using a measure μ on the domain X of the variable. Marichal and Roubens [15] as well as Dukhovny [16] have also suggested extensions of the Shannon entropy in the framework of measures. We note Honda [17] provides a comprehensive discussion of the idea of entropy for measures.

Yager’s definition of entropy of a measure is based on the use of the Shapley index [18]. If μ is a measure on the space X then for any