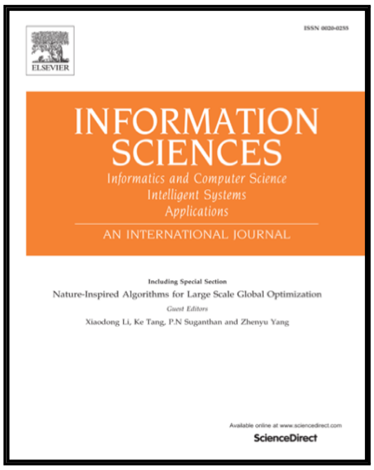
Accepted Manuscript



Categorization in Multi-Criteria Decision Making

Ronald R. Yager

PII: S0020-0255(17)30868-X
DOI: 10.1016/j.ins.2017.08.011
Reference: INS 13015

To appear in: *Information Sciences*

Received date: 29 March 2017
Revised date: 31 July 2017
Accepted date: 4 August 2017

Please cite this article as: Ronald R. Yager , Categorization in Multi-Criteria Decision Making, *Information Sciences* (2017), doi: 10.1016/j.ins.2017.08.011

Categorization in Multi-Criteria Decision Making

Ronald R. Yager

Machine Intelligence Institute Iona College New Rochelle, NY 10801

yager@panix.com

Technical Report #MII-3715R2

Abstract
We discuss the use of monotonic measures for the representation criteria importance information in multi-criteria decision-making. We show that the Choquet integral provides an appropriate method for the aggregation of the individual criteria satisfactions in the case where the relationship between criteria importance's is expressed using a measure. We describe the use of categories and the related idea of a categorization in formulating the structural relationship between multiple criteria. We show how we can model this categorization using a measure on the space of criteria, which in turn allows us to use the Choquet integral to evaluate an alternative's satisfaction to this type of multi-criteria decision problem. We look at a special categorization of the criteria that is closely to a prioritization of the criteria.

Keywords: Multi-Criteria, Set Measure, Aggregation, Categorization, Priority

1. Introduction

Multi-criteria appear in many modern technological tasks such as medical diagnosis, information retrieval, financial decision making and pattern recognition [1-5]. Collectively we shall refer to these as multi-criteria decision problems. Professor Janusz Kacprzyk has made important contributions this field [6-9]. In multi-criteria decision problems our interest is in selecting from some set of alternatives the one that best satisfies the criteria. Since it is generally difficult to rank alternatives based on their satisfaction's to multiple individual criteria a standard approach is to aggregate an alternative's satisfaction to the individual criteria to obtain a single scalar value corresponding to the alternative's overall satisfaction to the collection of criteria. These scalar values can then be used to rank the alternatives and enable a choice to be made. The aggregation of these multi-criteria satisfactions generally requires the use of some information regarding the importance of the individual criteria. The classic approach to this aggregation is to take a weighted average of an alternative's satisfaction to the individual criteria, the weights in this approach being the importance of the individual criteria. Implicit in this approach is an assumption that the individual criteria importance weights are additive. That is, for example, the importance of criteria weight of criteria one and two together is simply the addition of the two individual criteria importance weights. More generally, this assumes that the importance of any group of criteria taken together is simply the sum of the importance of the individual criteria. In many cases of decision making this simplifying assumption is not valid. For example, when selecting an employee the situation where the criteria of having a good education or considerable experience are interchangeable doesn't justify this assumption. More generally the situation in which the satisfaction of any one of a group of criteria is all that is needed does not satisfy the assumption of an additive relationship between individual criteria importance.

To model more complex relationships about the importance of subsets of criteria recent interest has focused on the use of a fuzzy measure [10-13]. In this approach, the additivity of the individual criteria importance's has been replaced by a monotonicity condition, if A and B are subsets of criteria such that A contains all the criteria in B then it is assumed that the importance of collection the A is at least as large as the collection B of criteria.

The use of this more general measure structure to represent our information about the importance of subsets of criteria complicates the process of aggregating the satisfactions of the individual criteria based on the importance information. The use of the simple weighted average of individual satisfactions does not always work. Here we show that the Choquet integral [14-18] provides an approach to the aggregation of the individual criteria satisfactions which generalizes the simple weighted average approach for additive weights to the case where the importance information is carried by a measure.

In some applications of multi-criteria decision making the criteria can be categorized, these categories can then used for expressing the information about criteria importance (see Zadeh [19]). Here, the collection of criteria in the same category shares a given amount of importance. This shared importance is distributed to the individual criteria in the category according to some rule, called the dispenser rule. Here, an alternative is evaluated as a weighted aggregation of the individual criteria satisfactions where the individual criteria weights is the sum of the amount of importance allocated to it by the categories to which it belongs. There can be many basis of this categorization. For example, criteria can be placed in the same category because of they are interchangeable. At the other extreme is one in which criteria can be placed in the same category because satisfying all of them is necessary. Here we shall look at the situation in which our information about the relationship between the criteria is expressed via a categorization. We show how we can use this category expressed importance information to obtain a measure-based representation of importance information. Once having this measure based representation we can use the Choquet integral to help in the evaluation of alternatives.

2. Aggregating Criteria Satisfactions using Measure Based Importance

Let $\mathcal{C} = \{C_1, \ldots, C_k, \ldots, C_q\}$ be a collection of criteria of interest in a decision problem. Here we shall use a measure μ to convey our information about the importance of subsets of criteria [10, 20]. In particular $\mu: 2^{\mathcal{C}} \to [0, 1]$ where

1) $\mu(\emptyset) = 0$

2) $\mu(\mathcal{C}) = 1$

3) $\mu(A) \geq \mu(B)$ if $B \subseteq A$

Thus here for any subset A of criteria $\mu(A) \in [0, 1]$ indicates the importance of this subset of criteria. We note that condition $\mu(\emptyset) = 0$ indicates that the importance of the null set is zero. The condition $\mu(\mathcal{C}) = 1$ tells us the importance of the whole set of criteria is one. Condition 3 says that if B is a smaller set of criteria then A, then B cannot have a larger importance.

The prototypical situation is the basic additive case where we have for each criteria C_k an importance α_k, $\mu(\{C_k\}) = \alpha_k$ and for any subset A of criteria $\mu_i(A) = \sum_{C_k \in A} \alpha_k$. We note for this situation since $\mu(\mathcal{C}) = 1$ we have that $\sum_{k=1}^{q} \mu(\{C_k\}) = \sum_{k=1}^{q} \alpha_k = 1$.

Using a measure μ to capture our importance provides an ability to model more sophisticated relationship between the criteria importance than the basic additive case. Here the use of a general measure μ can allow among other things the possibility that $\sum_{k=1}^{q} \mu(\{C_k\}) \neq 1$.

Thus if we have $\mu(\{C_k\}) = \alpha_k$ we can allow that $\sum_{k=1}^{q} \alpha_k \neq 1$.

Assume $X = \{x_i | i = 1 \text{ to } r\}$ are a set of alternatives and we are interested in choosing among those alternatives based upon their satisfactions to the criteria in C. Here $C_k(x_i) \in [0, 1]$ is the degree of satisfaction of criteria C_k by alternative x_i.

One way to select between these alternatives is to aggregate the alternative's individual criteria satisfactions guided by the importance information in μ, and then select the alternative with the largest aggregated value. Thus

$$D(x) = Agg_\mu(C_1(x), C_2(x), \ldots, C_q(x)).$$

For the situation where the importance relationship corresponds to the basic additive model then $D(x) = \sum_{k=1}^{q} \alpha_k\, C_k(x)$.

Our concern here is with the formulation of D in the general case of a measure based representation of the importance information. Let us look at some properties we desire of the formulation $D(x) = Agg_\mu(C_1(x), \ldots, C_q(x))$. In considering these properties we shall to some extent be guided by properties associated with the classic weighted average that comes from the case of additive importance's, $D(x) = \sum_{k=1}^{q} \alpha_k\, C_k(x)$.

A first feature we require of the general aggregator is that it is a mean like operator [21]. This requires this Agg_μ has the following three properties

1) Symmetry - It is indifferent to the index of the $C_k(x)$. More formally if $Q = \{1, \ldots, q\}$ and $\Pi : Q \rightarrow Q$ is a permutation operator then

$$Agg_\mu(C_1(x), \ldots, C_q(x)) = Agg_\mu(C_{\Pi(1)}(x), \ldots, C_{\Pi(q)}C(x))$$

2) Monotonicity with respect to the $C_k(x)$

If $C_k(x) \geq C_k(y)$ for all $k = 1, \ldots, q$ then

$$D(x) = Agg_\mu(C_1(x), \ldots, C_q(x)) \geq Agg_\mu(C_1(y), \ldots, C_q(y)) = D(y)$$

3) Boundedness

$$Min_k(C_k(x)) \leq Agg_\mu(C_1(x), \ldots, C_q(x)) \leq Max_k(C_k(x))$$

One implication of the boundedness is idempotency, if all $C_k(x) = a$ then $D(x) = a$.

Another feature we desire of the function Agg_μ is a kind of linearity. This is manifested in requiring the following two properties

4) Additivity: If $C_k(y) = C_k(x) + a$ for all k = 1 to q then $D(y) = D(x) + a$.

5) Positive homogeneity: If $C_k(y) = \lambda C_k(x)$ for all k and $\lambda \in [0, 1]$ then $D(y) = \lambda D(x)$

One fundamental feature associated with the classic weighted average based on additive weights is the following. If we move importance weight from a criteria with lesser satisfaction to one with greater satisfaction then D(x) increases. More formally if $C_1(x) > C_2(x)$ and $\tilde{\alpha}_1 = \alpha_1 + \Delta$ and $\tilde{\alpha}_2 = \alpha_2 - \Delta$ and $\tilde{\alpha}_k = \alpha_k$ for all other k = 2 to q then

$$\sum_{k=1}^{q} \tilde{\alpha}_k C_k(x) = \sum_{k=1}^{q} \alpha_k C_k(x) + \Delta C_1(x) - \Delta C_2(x) = \sum_{k=1}^{q} \alpha_k C_k(x) + \Delta (C_1(x) - C_2(x)) \geq \sum_{k=1}^{q} \alpha_k C_k(x)$$

In order to capture this feature in the more general case where our importance weights are expressed via a measure μ we require the following property of Agg_μ.

6) Dominance

Let ρ be an index function so that $\rho(j)$ is the index of j^{th} largest of the criteria satisfactions, here then $C_{\rho(j)}(x)$ is the j^{th} largest criteria satisfaction. Let H_j be the subset of criteria with the j largest satisfactions, $H_j = \{C_{\rho(1)}, \ldots, C_{\rho(j)}\}$. Assume μ_1 and μ_2 are two importance measures such that $\mu_2(H_j) \geq \mu_1(H_j)$ for all j = 1 to q, then by dominance we require our function Agg_μ to be such that

$$Agg_{\mu_2}(C_1(x), \ldots, C_q(x)) \geq Agg_{\mu_1}(C_1(x), \ldots, C_q(x))$$

The condition $\mu_2(H_j) \geq \mu_1(H_j)$ indicates that measure μ_2 has more importance associated with the j most satisfied criteria then μ_1.

Let us see that this property of dominance captures the situation in the case where our measure is the basic additive measure, $\mu_1(A) = \sum_{C_k \in A} \alpha_k$. Here with $H_j = \{C_{\rho(1)}, \ldots, C_{\rho(j)}\}$ and $\mu_1(H_j) = \sum_{i=1}^{j} \alpha_{\rho(i)}$. In particular $\mu_1(H_1) = \alpha_{\rho(1)}$ and $\mu_1(H_2) = \alpha_{\rho(1)} + \alpha_{\rho(2)}$ with $\mu_1(H_j) = \sum_{i=1}^{j} \alpha_{\rho}(i)$. Assume we now move some importance weight from $C_{\rho(2)}$ to $C_{\rho(1)}$ to form $\tilde{\mu}_2$, thus in the case of $\tilde{\mu}_2$ we have $\tilde{\alpha}_{\rho(1)} = \alpha_{\rho(1)} + \Delta$ and $\tilde{\alpha}_{\rho(2)} = \alpha_{\rho(2)} - \Delta$ and $\tilde{\alpha}_{\rho(j)} = \alpha_{\rho(j)}$ for all other j. In this case

$$\tilde{\mu}_2(H_1) = \tilde{\alpha}_{\rho(1)} = \alpha_{\rho(1)} + \Delta$$

$$\tilde{\mu}_2(H_2) = \tilde{\alpha}_{\rho(1)} + \tilde{\alpha}_{\rho(2)} = \alpha_{\rho(1)} + \Delta + \alpha_{\rho(2)} - \Delta = \alpha_{\rho(1)} + \alpha_{\rho(2)}$$

$$\tilde{\mu}_2(H_j) = \sum_{i=1}^{j} \alpha_{\rho(j)} = \sum_{i=1}^{j} \alpha_{\rho(j)} \text{ for j = 3 to q}$$

Thus here we have that $\tilde{\mu}_2(H_j) \geq \mu_1(H_j)$ for all j.

Thus we see that the condition $\tilde{\mu}_2(H_j) \geq \mu_1(H_j)$ for all j generalizes the idea of moving importance weight from less satisfied criteria to more satisfied criteria. Here then the requirement that $Agg_{\tilde{\mu}_2}(C_1(x), \ldots, C_q(x)) \geq Agg_{\mu_1}(C_1(x), \ldots, C_q(x))$ if $\tilde{\mu}_2(H_j) \geq \mu_1(H_1)$ for all j generalizes the property that if we move importance weight from less satisfied criteria to more satisfied criteria we should increase of overall satisfaction.

If μ_1 and μ_2 are two measures of importance we say that $\mu_2 \geq \mu_1$ if the $\mu_2(A) \geq \mu_1(A)$ for all A. We observe in the case $\mu_2 \geq \mu_1$ whatever elements constitute the H_j we have that $\mu_2(H_j) \geq \mu_1(H_j)$ for all H_j. Therefore we see in the situation where $\mu_2 \geq \mu_1$ we require that the Agg operator should satisfy $Agg_{\mu_2}(C_1(x), \ldots, C_q(x))) \geq Agg_{\mu_1}(C_1(x), \ldots, C_q(x))$ for any values of $C_k(x)$.

We now show that the Choquet integral expressed below can provide a formulation for $Agg_\mu(C_1(x), \ldots, C_q(x))$ that can satisfy all our requirements. Using the Choquet integral [21]

$$D(x) = Agg_\mu(C_1(x), \ldots, C_q(x)) = \sum_{j=1}^{q} (\mu(H_j) - \mu(H_{j-1}))C_{\rho(j)}(x))$$

In the above formula ρ is an index function such that ρ(j) is the index of the criteria with the j^{th} largest satisfaction and H_j is the subset of criteria with the j largest satisfactions.

We first observe that if we denote $\mu(H_j) - \mu(H_{j-1}) = w_j$ then each $w_j \geq 0$. In addition we have $\sum_{j=1}^{q} w_j = \sum_{j=1}^{q} (\mu(H_j) - \mu(H_{j-1})) = \mu(H_q) - \mu(H_0) = \mu(C) - \mu(\varnothing) = 1$. Thus the Choquet integral, $D(x) = \sum_{j=1}^{q} w_j C_{\rho(j)}(x)$, is a kind of weighted average of the criteria satisfactions.

It is well known that the Choquet integral satisfies conditions 1 – 5. [21, 22] Let us look at the sixth condition, dominance. To show that is condition is satisfied we shall rearrange the summation in the Choquet integral. Using some arithmetic manipulations we can show that

$$D(x) = \sum_{j=1}^{q} (\mu(H_j) - \mu(H_{j-1}))C_{\rho(j)}(x) = \sum_{j=1}^{q} \mu(H_j)(C_{\rho(j)}(x) - C_{\rho(j+1)}(x))$$

Consider the case where we have two measures of importance μ_1 and μ_2 such that $\mu_2 \geq \mu_1$, i.e. $\mu_2(A) \geq \mu_1(A)$ for all A. Consider

$$D_1(x) = Agg_{\mu_1}(C_1(x), \ldots, C_q(x)) = \sum_{j=1}^{q} \mu_1(H_j)\ (C_{\rho(j)}(x) - C_{\rho(j+1)}(x))$$

and

$$D_2(x) = Agg_{\mu_2}(C_1(x), \ldots, C_q(x)| = \sum_{j=1}^{q} \mu_2(H_j)\ (C_{\rho(j)}(x) - C_{\rho(j+1)}(x))$$

We see that $D_2(x) - D_1(x) = \sum_{j=1}^{q} (\mu_2(H_j) - \mu_1(H_1))\ (C_{\rho(j)}(x) - C_{\rho(j+1)}(x))$. Since $\mu_2(H_j) \geq \mu_1(H_j)$ for all j and $C_{\rho(j)}(x) \geq C_{\rho(j+1)}(x)$ for all j then $D_2(x) - D_1(x) \geq 0$ and hence $D_2(x) \geq D_1(x)$. Thus the sixth condition is satisfied.

3. Categorization of Criteria

Let $C = \{C_1, \ldots, C_q\}$ be a set of criteria that are relevant to a decision. A category F is any subset of criteria from C. A ***categorization*** F consists of a collection of categories, F_i for i = 1 to r. Thus a categorization consists of a collection of subsets of C. In a categorization there is no requirement for the constituent categories to be disjoint. In addition the categories in the categorization do not need to cover the whole set C, $\bigcup_{i-1}^{r} F_i \neq C$. Thus, a categorization does not require a partitioning of C.

Associated with each category F_i in the categorization F is a weigh $\lambda_i \in [0, 1]$ such that $\sum_{i=1}^{r} \lambda_i = 1$. Also associated with each category F_i in a categorization F is a ***dispenser*** rule which describes how the weight λ_i is dispensed depending on the criteria in the category satisfied. One extreme example of a dispenser rule is case where the whole weight λ_i is dispensed to the least satisfied criteria in the category, here all the criteria in the category can be seen as required to be satisfied. At the other extreme is the case in which the whole weight λ_i is dispensed to the most satisfied criteria in the category, here the criteria in a category can be seen as completely interchangeable. Intermediate to these extreme cases is one in which the whole weight λ_i is dispensed to the k^{th} most satisfied criteria in F_i.

Another special case of dispenser is a linear rule. Here rather when dispensing all the weight λ_i of a category F_i to one criterion we dispense it in a proportional manner, all the members of category get the same portion.

More generally we associate with a category F_i of cardinality $n_i = |F_i|$ a set of values $V_{ij}$for j = 0 to n_i so that

1) $V_{i0} = 0$

2) $V_{in_i} = 1$

3) $V_{ij+1} \geq V_{ij}$

Here V_{ij} is the proportion of the i^{th} category weight λ_i that is dispensed to the j most satisfied criteria in F_i.

We can see the examples of dispenser rules we previously described can be seen as special case of this formulation. The case where we require all the criteria in a category F_i to be satisfied is one in which $V_{in_i} = 1$ and $V_{ij} = 0$ for all other j. The case where we only need any element in F_i to be satisfied is one where $V_{i0} = 0$ and $V_{ij} = 1$ for $j \neq 0$. The case where we want at least K criteria can be modeled by one in which $V_{ij} = 0$ for $j < K$ and $V_{ij} = 1$ for $j \geq K$. Finally the case of linear dispenser rule can be modeled by one in which $V_{ij} = j/n_i$ for all j = 0 to n_i.

There are a number of alternate ways we can model the dispenser rule for category F_i. One useful ways is via a function monotonic function G_i: $[0, 1] \rightarrow [0, 1]$ having $G_i(0) = 0$ and $G_i(1) = 1$. In this case we obtain $V_{ij} = G_i(\frac{j}{n_i})$.

One advantage of using this function G_i is that we can easily compute the dispenser rule for different categories even if they have different number of elements. That is we can assign two categories of different cardinalities the same dispenser rule by using the same function G.

We now show how we can use a categorization $\mathcal{F}$ specifying a relationship between criteria from $\mathcal{C}$ to generate an importance measure μ on the space $\mathcal{C}$ of criteria. Having this measure μ we can use the Choquet integral to determine the overall satisfaction of an alternative x to the categorization $\mathcal{F}$ based upon its satisfaction to the individual criteria, the $C_i(x)$.

Consider now the set function μ on the space $\mathcal{C}$ of criteria so that for any subset A of criteria

$$\mu(A) = \sum_{i=1}^{r} \lambda_i V_{i|F_i \cap A|}$$

where $|F_i \cap A|$ is the number of elements in F_i that are in A. We now show that this has the basic properties of a measure

1. IF $A = \emptyset$ then $|F_i \cap A| = |F_i \cap \emptyset| = |\emptyset| = 0$. In this case $V_{io} = 0$ for all i and $\mu(\emptyset) = 0$

2) If $A = C$ then $|F_i \cap A) = |F_i \cap C| = |F_i| = n_i$. In case $V_{i|F_i \cap A|} = V_{i|F_i \cap C|} = V_{in_i} = 1$. Here then $\mu(C) = \sum_{i=1}^{r} \lambda_i = 1$,

3) If $A \subseteq B$ then $|F_i \cap A| \leq |F_i \cap B|$ and with $\mu(A) = \sum_{i=1}^{r} \lambda_i V_{i|F_i \cap A|}$ and $\mu(B) = \sum_{i=1}^{r} \lambda_i V_{i|F_i \cap B|}$ from the monotonicity of V_{ij} we see that $\mu(B) \geq \mu(A)$.

Let us now look at form of the decision function, D(x), generated by this F categorization based importance measure μ, $D(x) = \sum_{j=1}^{q} (\mu(H_j) - \mu(H_{j-1}))\, C_{\rho(j)}(x)$. Here with $\mu(H_j) = \sum_{i=1}^{r} \lambda_i V_{i|F_i \cap H_j|}$ and $\mu(H_{j-1}) = \sum_{i=1}^{r} \lambda_i V_{i|F_i \cap H_{j-1}|}$ we have

$$D(x) = \sum_{j=1}^{q} \left(\sum_{i=1}^{r} \lambda_i V_{i|F_i \cap H_j|} - \sum_{i=1}^{r} \lambda_i V_{i|F_i \cap H_{j-1}|}\right) C_{\rho(j)}(x)$$

$$D(x) = \sum_{j=1}^{q} \left(\sum_{i=1}^{r} \lambda_i (V_{i|F_i \cap H_j|} - V_{i|F_i \cap H_{j-1}|})\right) C_{\rho(j)}(x) \qquad \text{(I)}$$

$$D(x) = \sum_{i=1}^{r} \lambda_i \left(\sum_{j=1}^{q} (V_{i|F_i \cap H_j|} - V_{i|F_i \cap H_{j-1}|}\right) C_{\rho(j)}(x) \qquad \text{(II)}$$

We see that (I) allows us to view D(x) in terms of the ordered criteria satisfactions, the $C_{\rho(j)}(x)$. On the other hand (II) allows viewing D(x) in terms of categories.

In the following we shall focus on (II). Assume $C_{\rho(j)}$ is not in F_i. In this case we see that $F_i \cap H_j = F_i \cap H_{j-1}$ and hence $V_{i|F_i \cap H_j|} = V_{i|F_i \cap H_{j-1}|}$. In this case the term $C_{\rho(j)}(x)$ does not make any contribution to the inner sum.

Let ρ_i be a subpart of the function ρ just corresponding to the ordering of the elements in F_i. Thus $C_{\rho_{i(j)}}$ is the criteria in F_i with the j^{th} largest satisfaction. Using this we can express

$$\mathrm{D(x)} = \sum_{i=1}^{r} \lambda_i \left(\sum_{j=1}^{n_i} (V_{ij} - V_{ij-1}) C_{\rho_i(j)}(x)\right).$$

Let us denote $M_i(x) = \sum_{j=1}^{n_i} (V_{ij} - V_{ij-1}) C_{\rho_i(j)}(x)$. Using this we see the D(x) = $\sum_{i=1}^{r} \lambda_i M_i(x)$.

Let us look at the form of $M_i(x)$ for notable examples of category dispenser rules.

1) If V_{ij} is linear, $V_{ij} = \frac{j}{n_i}$ then $M_i(x) = \sum_{j=1}^{n_i} \frac{1}{n_i} C_{\rho_i(j)}(x)$. Here $M_i(x)$ is the average satisfaction of the criteria in F_i.

2) If V_{ij} is such that $V_{i0} = 0$ and $V_{ij} = 1$ for all $j \neq 0$ then $M_i(x) = \underset{C_k \in F_i}{Max} [C_k(x)]$

3) If V_{ij} is such that $V_{in_i} = 1$ and all other $V_{ij} = 0$ then $M_i(x) = \underset{C_k \in F_i}{Min} [C_k(x)]$

4. Nominal Importance of a Criterion in a Categorization

In [23] Shapley introduced the concept of the Shapley index associated with a measure. Assume μ is a measure on the space $\mathcal{C} = \{C_1, \ldots, C_q\}$ and the Shapley index S_i of the element C_i is defined as

$$S_i = \sum_{F \subseteq E_i} \frac{(q - Card(F) - 1)!\, Card(F)!}{q!} (\mu(F \cup \{C_i\}) - \mu(F))$$

where $E_i = \mathcal{C} - \{C_i\}$. Here then F is a subset of $\mathcal{C}$ not containing C_i. We see that Shapley index S_i is a kind of average gain in the measure μ by adding C_i to a subset of $\mathcal{C}$ not containing C_i.

The Shapley value associated with μ is the vector $(S_1, S_2, \ldots, S_q)$. It is known [24] that the following properties are satisfied by the S_i

1) $S_i \in [0, 1]$ for all i

2) $\sum_{i=1}^{q} S_i = 1$

In [24] Yager looked at the Shapley index for various types of measure. If μ is a basic additive measure with $\mu(\{C_i\}) = \alpha_i$ it can be shown that $S_i = \alpha_i$. [24]

In the framework of using the measure μ to capture the importance associated with a subset of $\mathcal{C}$ we see that the Shapley index S_i can be seen as some type of nominal or average importance associated with the criteria C_i.

In the case of the measure μ used to represent the categorization $\mathcal{F}$ it can be shown that

$S_k = \sum_{i=1}^{r} \frac{\lambda_i}{n_i} F_i(C_k)$ where $F_i(C_k) = 1$ if $C_k \in F_i$ and $F_i(C_k) = 0$ if $C_k \notin F_i$. It is interesting to observe here that the values of S_k are independent of the dispenser function V_{ij} associated the categories.

We see the formulation $D(x) = \sum_{i=1}^{r} \lambda_i (\sum_{j=1}^{n_i} (V_{ij} - V_{ij-1}) C_{\rho_i(j)}(x))$ depends on the dispenser function as manifested by the inclusion of V_{ij}. Consider the special cases where the V_{ij} is linear. Here $V_{ij} = \frac{j}{n_i}$ and $V_{ij} - V_{ij-1} = \frac{1}{n_i}$ and $D(x) = \sum_{i=1}^{r} \lambda_i (\sum_{j=1}^{n_i} \frac{1}{n_i} C_{\rho_i(j)}(x))$. After some algebraic manipulations we see that $D(x) = \sum_{k=1}^{q} S_k C_k(x)$. Thus we see that the Shapley values, the S_k, are the weights associated with criteria the C_k in the calculation of D(x) in the special case where all the categories have linear dispenser rule.

5. Prioritized Type Categorization

An interesting example of categorization is shown below. Assume for i = 1 to q we have the following categories in our categorization

$F_1 = \{C_1\}$

$F_2 = \{C_1, C_2\}$

…….

$F_i = \{(C_k| \text{ for } k = 1 \text{ to } i\}$

….

$F_q = \{C_1, \ldots, C_q\} = C$

Here the number of categories r = q, the number of criteria. Furthermore associated with each category F_i is a weight $\lambda_i \geq 0$ where $\sum_{i=1}^{q} \lambda_i = 1$. We note here that category F_i has cardinality $n_i = i$.

As we earlier showed the Shapley index is independent of the dispenser rule and hence $S_k = \sum_{i=1}^{q} \lambda_i \frac{1}{n_i} F_i(C_k)$ where $F_i(C_k) = 1$ if $C_k \in F_i$ and $F_i(C_k) = 0$ if $C_k \notin F$. Furthermore in this case we indicated that $n_i = i$, thus $S_k = \sum_{i=1}^{q} \lambda_i \frac{1}{i} F_i(C_k)$. For this categorization we see that for $i \geq k$ we have that $C_k \in F_i$. Thus here $F_i(C_k) = 1$ for $i \geq k$ and $F_i(C_k) = 0$ for $i < k$. Here then we get $S_k = \sum_{i=k}^{q} \frac{\lambda_i}{i}$. Thus the nominal importance associated with C_k in this categorization is $S_k = \sum_{i=k}^{q} \frac{\lambda_i}{i}$. Here then we have that $S_k \geq S_{k+1}$. Thus C_1 has the largest nominal importance and C_q the smallest.

Let us look at the calculation of D(x) using the measure μ induced by this categorization of the criteria in C for different formulation of the dispenser function. Since this is a special case of the categorization F studied earlier we can take advantage of many of the preceding results.

Here we shall look at D(x) for three prototypical forms of dispenser function V. Here we shall assume that all the F_i have the same form of dispenser function.

For the case when all the F_i have linear dispenser function then $V_{ij} = \frac{j}{n_i} = \frac{j}{i}$ and hence

$$D(x) = \sum_{k=1}^{q} \left(\sum_{i=1}^{k} \frac{\lambda_i}{i}\right) C_k(x).$$

The second prototypical case is one in which $V_{i0} = 0$ and $V_{ij} = 1$ for all j = 1 to n_i. Here F_i dispenses all of its weight if any of the criteria in F_i is satisfied. Here we see that

$$D^*(x) = \sum_{i=1}^{q} \lambda_i \underset{C_k \in F_i}{Max} [C_k(x)] = \sum_{i=1}^{q} \lambda_i \underset{k=1 \, to \, i}{Max} [C_k(x)]$$

The last case is where F_i dispenses all its weight if all the criteria in F_i are satisfied. In this

$$D_*(x) = \sum_{i=1}^{q} \lambda_i \underset{C_k \in F_i}{Min} [C_k(x)] = \sum_{i=1}^{q} \lambda_i \underset{k=1 \, to \, i}{Min} [C_k(x)]$$

We easily see that $D^*(x) \geq D_*(x)$ for any values of the $C_k(x)$.

We also observe that $\underset{k=1 \, to \, i}{Max} [C_k(x)] \geq C_1(x)$ for all i and $\underset{k=1 \, to \, i}{Min} [C_k(x)] \leq C_1(x)$ for all i.

Hence we see that $D^*(x) \geq \sum_{i=1}^{q} \lambda_i C_1(x) \geq C_1(x)$ and $D_*(x) \leq \sum_{i=1}^{q} \lambda_i C_1(x) \leq C_1(x)$

Let Π(j) be index function so that Π(j) is the index of j^{th} most satisfied criteria. Let us look at the formulation of $D^*(x)$ and $D_*(x)$ for different forms of the Π(j). First consider the case where Π(j) = j, here the ordering of the satisfactions is the same as the priority. We see that for this case $\underset{C_k \in F_i}{Max}[C_k(x)] = C_1(x)$ for all F_i. Here then we get that $D^*(x) = \sum_{i=1}^{q} \lambda_i C_1(x) = C_1(x) = \underset{k}{Max}[C_k(x)]$. On the other hand we see that $\underset{C_k \in F_i}{Min}[C_k(x)] = C_i(x)$ for all i and here $D_*(x) = \sum_{i=1}^{q} \lambda_i C_i(x)$.

Consider now the case where $\Pi(j)$ is inversely ordered to criteria, thus $\Pi(j) = q - 1 + j$. In this case we see that $\underset{C_k \in F_i}{Max}[C_k(x)] = C_i(x)$. Here we get $D^*(x) = \sum_{i=1}^{q} \lambda_i C_i(x)$. This is the same as $D_*(x)$ for the case where $\Pi(j) = j$. For this inverse ordering $\underset{C_k \in F_i}{Min}[C_k(x)] = C_1(x)$ for all i. Here then we get $D_*(x) = \sum_{i=1}^{q} \lambda_i C_1(x) = C_1(x) = \underset{k}{Min}[C_k(x)]$.

We shall say that C_j has priority over $C_{j+1,}$ a prioritization exists between C_j and C_{j+1}, if even the largest increase in the satisfaction to C_{j+1} generally cannot compensate for even the smallest loss in satisfaction to C_j. We now show that the categorization that is manifested as D_* provides the basic properties of a prioritization between C_j and C_{j+1}.

Assume that $C_j(x) = C_{j+1}(x) = a$ and all other $C_i(x) = 1$. We see here that from $D_*(x) = \sum_{i=1}^{q} \lambda_i \underset{k=1 to i}{Min}[C_k(x)]$ with these values we get $D_*(x) = \sum_{i=1}^{j-1} \lambda_i + a\sum_{i=j}^{q} \lambda_i$. Assume now for $\Delta >> \varepsilon$ we increase C_{j+1} to $C_{j+1}(x) = a + \Delta$ and we decrease C_j so that $C_j(x) = a - \varepsilon$ and leave all other $C_i(x) = 1$. Here we see that $D_*(x) = \sum_{i=1}^{j-1} \lambda_i + (a - \varepsilon)\sum_{i=j}^{q} \lambda_i < D_*(x)$.

Now assume we have $C_j(x) = a$ and $C_{j+1}(x) = b$ where $b > a$. Here again we get $D_1(x) = \sum_{i=1}^{j-1} \lambda_i + a\sum_{i=j}^{q} \lambda_i$. Now assume we again we increase C_{j+1} so that $C_{j+1} = b + \Delta$ and decrease C_j so that $C_j(x) = a - \varepsilon$. Here again we see that $D_*(x) = \sum_{i=1}^{j-1} \lambda_i + (a - \varepsilon)\sum_{i=j}^{q} \lambda_i < D_1(x)$.

Finally consider the situation where $a > b$. Here we get $D_1(x) = \sum_{i=1}^{j-1} \lambda_i + a\lambda_j + b\sum_{i=j+1}^{q} \lambda_i$. Now assume we increase C_{j+1} so that $C_{j+1} = b + \Delta$ and decrease C_j so that $C_j = a - \varepsilon$. Here we get $D_*(x) = \sum_{i=1}^{j} \lambda_i + (a - \varepsilon)\lambda_j + \text{Min}[(b + \Delta), (a - \varepsilon)] \sum_{i=j+1}^{q} \lambda_i$ Only if Δ is so large such that (b + Δ) > (a - ∈) then $D_*(x) = \sum_{i=1}^{j} \lambda_i + (a - \varepsilon)\lambda_j + (a - \varepsilon) \sum_{i=j+1}^{q} \lambda_i > D_*(x)$. Only in this case we can obtain some comparison however we are bounded by (a - ε).

6. Conclusion

We discussed the use of monotonic measures for the representation criteria importance information in multi-criteria decision-making. We showed that the Choquet integral provides an appropriate method for the aggregation of the individual criteria satisfactions in this case where the relationship between criteria importance is expressed using a measure. We described the use of categories and the related idea of a categorization in expressing the structural relationship between multiple criteria. We showed how we could model this categorization using a measure on the space of criteria, which in turn allowed us to use the Choquet integral to evaluate an alternative's satisfaction to this type of multi-criteria decision problem. We looked at a special categorization of the criteria that is closely to a prioritization of the criteria.

7. References

[1] Zopounidis, C. and Doumpos, M., "Multiple Criteria Decision Making: Applications in Management and Engineering," Springer: Heidelberg, 2017.
[2] Greco, S., Ehrgott, M., and Figueira, J. R., Multiple Criteria Decision Analysis: State of the Art Surveys. Heidelberg: Springer, 2016.
[3] Yager, R. R., "Pythagorean membership grades in multi-criteria decision making," IEEE Transaction on Fuzzy Systems 22, pp. 958-965, 2014.
[4] Koksalan, M., Wallenius, J., and Zionts, S., Multiple Criteria Decision Making: From Early History to the 21st Century. Singapore: World Scientific Publishing, 2011.
[5] Pedrycz, W., Ekel, P., and Parreiras, R., Fuzzy Multicriteria Decision-Making: Models, Methods and Applications. New York: John Wiley, 2010.
[6] Kacprzyk, J., Multistage Decision Making under Fuzziness. Cologne: Verlag TUV Rheinland, 1983.
[7] Kacprzyk, J. and Fedrizzi, M., Multiperson Decision Making Using Fuzzy Sets And Possibility Theory. Hingham, MA: Kluwer, 1990.
[8] Wilbik, A. and Kacprzyk, J., "A multi-criteria evaluation of linguistic summaries of time sries via a measure of informativeness," In: Rutkowski L., Scherer R., Tadeusiewicz R., Zadeh L.A., Zurada J.M. (eds) Artificial Intelligence and Soft Computing. ICAISC 2010. Lecture Notes in Computer Science, vol 6113. Springer, Berlin, Heidelberg, 2010.
[9] Ventre, A. G. S., Maturo, A., Hoskova-Mayerova, S., and Kacprzyk, J., "Multicriteria and Multiagent Decision Making with Applications to Economics and Social Sciences," Springer,

Heidelberg, 2013.

[10] Yager, R. R. and Alajlan, N., "On the measure based formulation of multi-criteria decision functions," Information Sciences 370-371, pp. 256-269, 2016.

[11] Grabisch, M. and Labreuche, C., "Fuzzy measures and integrals in MCDA," In Multiple Criteria Decision Analysis : State of the Art Surveys, S. Greco, M. Ehrgott, J.R. Figueira (Eds.), Springer: Berlin, 2016.

[12] Yager, R. R., "Multi-criteria decision making using fuzzy measures," Cybernetics and Systems 46, pp. 150-171, 2015.

[13] Marichal, J. L., "On Choquet and Sugeno integrals as aggregation functions," in *Fuzzy Measures and Integrals*, Grabisch, M., Murofushi, T., and Sugeno, M., Eds., ed Heidelberg: Springer, 2000, pp. 247-272.

[14] Yager, R. R., "Evaluating Choquet integrals whose arguments are probability distributions," IEEE Transactions on Fuzzy Systems 24, pp. 957-965, 2016.

[15] Wang, Z., Yang, R., and Leung, K.-S., Nonlinear Integrals and Their Applications in Data Mining. Singapore: World Scientific, 2010.

[16] Klement, E. P., Mesiar, R., and Pap, E., "A universal integral as common frame for Choquet and Sugeno," IEEE Transaction of Fuzzy Systems 18, pp. 178-187, 2010.

[17] Wang, Z. and Klir, G. J., Generalized Measure Theory. New York: Springer, 2009.

[18] Choquet, G., "Theory of Capacities," Annales de l'Institut Fourier 5, pp. 131-295, 1953.

[19] Zadeh, L. A., "A very simple formula for aggregation and multicriteria 0ptimization," International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems 24, pp. 961-962, 2016.

[20] Torra, V., Narukawa, Y., and Sugeno, M., Non-Additive Measures: Theory and Applications. Heidelberg: Springer, 2013.

[21] Beliakov, G., Pradera, A., and Calvo, T., Aggregation Functions: A Guide for Practitioners. Heidelberg: Springer, 2007.

[22] Klement, E. P. and Mesiar, R., "On the axiomatization of some classes of discrete universal integrals," Knowledge-Based Systems 28, pp. 13-18, 2012.

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

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