International Journal of Approximate Reasoning 50 (2009) 1279–1286

Contents lists available at ScienceDirect

International Journal of Approximate Reasoning

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

On generalized Bonferroni mean operators for multi-criteria aggregation

Ronald R. Yager

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

ARTICLE INFO

Article history:
Received 9 February 2009
Received in revised form 20 May 2009
Accepted 8 June 2009
Available online 14 June 2009

Keywords:
Aggregation operators
OWA operators
Multi-criteria modeling
Decision-making

ABSTRACT

We introduce the idea of multi-criteria aggregation functions and describe a number of properties desired in such functions. We emphasize the importance of having an aggregation function capture the expressed interrelationship between the criteria. A number of standard aggregation functions are introduced. We next introduce the Bonferroni mean operator. We provide an interpretation of this operator as involving a product of each argument with the average of the other arguments, a combined averaging and "anding" operator. This allows us to suggest generalizations of this operator by replacing the simple averaging by other mean type operators as well as associating differing importances with the arguments.

1. Introduction

Problems involving multi-criteria are pervasive in many areas of modern technology. Not only do they appear in decision-making but they also arise in such diverse areas as pattern recognition, information retrieval, case based reasoning and database querying among others. A central problem in multi-criteria problems is the aggregation of the satisfactions to the individual criteria to obtain a measure of satisfaction to the overall collection of criteria. This aggregation process must be guided by the interrelationship of the individual criteria, the criteria organization. As many different types of criteria relationships exist in the real world there is a need for many types of formal aggregation operations to enable the modeling of these numerous types of relationships. In response to this need a formal mathematical discipline called aggregation theory is emerging [1–4]. Here we contribute to this theory by looking at the Bonferroni mean operator [5,6] and suggesting some generalizations that enhance its modeling capability. We provide an interpretation of this operator as involving a product of each argument with the average of the other arguments, a combined averaging and "anding" operator. This allows us to suggest generalizations of this operator by replacing the simple averaging by other mean type operators such as the OWA operator and Choquet integral as well as associating differing importances with the arguments. We that various extensions of the Bonferroni mean can model different degrees of hard and soft partial conjunctions [7].

2. Multi-criteria aggregation functions

In multi-criteria decision-making have a collection $A_1,\ldots,A_n$ of criteria and a set $X = \{x_1,\ldots,x_m\}$ of alternatives. For each alternative x_i we have a value $A_j(x_i) \in [0, 1]$ indicating the degree to which alternative x_i satisfies criteria A_j. Our objective is to develop some procedure to select from these alternatives the one that best satisfies the collection of criteria. One property often required of such a procedure is what Arrow [8] called indifference to irrelevant alternatives. Essentially this

E-mail address: yager@panix.com

0888-613X/$ - see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.ijar.2009.06.004

property assures that the decision procedure is such that we can't affect the outcome by introducing alternatives whose sole purpose is to disturb the process. Formally this property requires that our procedure is such that if the application of procedure to $X = \{x_1, \ldots, x_m\}$ selects x^*, Procedure$(x_1, \ldots, x_m) \rightarrow x^*$, then application of Procedure to $\{x_1, \ldots, x_m, x_{m+1}\}$ must result in either x^* or x_{m+1}.

One way to guarantee this property of indifference to irrelevant alternatives is to obtain for each alternative x_j a valuation of $D(x_j)$ using a function $D(x_j) = F(A_1(x_j), \ldots, A_n(x_j))$ and then select the alternative with largest value of D. A function such as F is called a pointwise valuation function. The important feature here is that $D(x_j)$ just depends on the satisfaction of the criteria by x_j, it does not depend on the satisfactions by any of the other alternatives.

In addition to the above other properties are desired in the valuation procedure. One of these is monotonicity, if x_j and x_k are two alternatives such that $A_i(x_j) \geqslant A_i(x_k)$ for all A_i then we require $D(x_j) \geqslant D(x_k)$. Another property is what we shall call grounding, if $A_i(x_j) = 0$ for all i then $D(x_j) = 0$. If in addition $D(x_j) = 1$ if all $A_i\,(x_j) = 1$, a condition we shall refer to as being standard, then F is what is called an aggregation function [1,9]. Letting $I = [0,1]$ then formally an aggregation function is a mapping Agg: $I^n \rightarrow I$ having the properties: $\text{Agg}(0, \ldots, 0) = 0$, $\text{Agg}(1, \ldots, 1) = 1$ and $\text{Agg}(a_1, \ldots, a_n) \geqslant \text{Agg}(b_1, \ldots, b_n)$ if $a_i \geqslant b_i$ for all i. We shall use the terms aggregation functions and aggregation operators synonymously.

Since the function F should be consistently chosen for all alternatives the pointwise nature of F allows us to simply focus on just one typical alternative, x, in discussing F. In the following we shall generally use a_j to indicate $A_j(x)$.

The actual choice of the aggregation function should be a reflection of our knowledge of the relational organization of the criteria. In the following we shall discuss some notable aggregation functions and indicate the type of criteria relationships they can model.

One formulation is $D(x) = T(a_1, \ldots, a_n)$ where T is a t-norm operator [10]. These aggregation functions are used to model situations when *all* the criteria are required to be satisfied by a solution. Notable among this class of functions are the following: $D(x) = \text{Min}(a_1, \ldots, a_n)$, $D(x) = \prod_{j=1}^{n} a_j$ and $D(x) = \text{Max}\left(0, \sum_{j=1}^{n} a_j - (n-1)\right)$.

Another class of functions is $D(x) = S(a_1, \ldots, a_n)$ where S is a t-conorm operator [10]. These are used to model situations where the satisfaction to *any* of the criteria is sufficient. Notable among this class of functions are the following: $D(x) = \text{Max}(a_j), D(x) = 1 - \prod_{j=1}^{n}(1 - a_j)$ and $D(x) = \text{Min}\left(1, \sum_{j=1}^{n} a_j\right)$.

A general class of functions that can be used to formulate the aggregation function F is the OWA operator [11]. Assume $w_j \in [0,1]$ are a collection of parameters that sum to one. Letting $\pi(j)$ be the index of the jth largest of the a_i the OWA aggregation is calculated as

$$D(x) = F(a_1, \ldots, a_n) = \sum_{j=1}^{n} w_j a_{\pi(j)}$$

The w_j are referred to as the OWA weights and collectively they can viewed as a vector W whose jth component is w_j. The OWA operators are mean type aggregation functions [11].

By assigning different values to the OWA weights we can obtain a wide class of formulations for the aggregation function F. If $w_1 = 1$ and $w_j = 0$ for $j \neq 1$ then $D(x) = \text{Max}_i(a_i)$ and if $w_n = 1$ and $w_j = 0$ for $j \neq n$ then $D(x) = \text{Min}_i(a_i)$. If $w_j = 1/n$ for all j we get the usual average, $D(x) = \frac{1}{n}\sum_{i=1}^{n} a_i$. We can associate with an OWA operator a measure called its attitudinal character [11] defined as $A\text{–}C(W) = \sum_{j=1}^{n} w_j \frac{n-j}{n-1}$. It can be shown that $A\text{–}C(W) \in [0,1]$. We note for the case of Max, $A\text{–}C(W) = 1$, for the case of Min, $A\text{–}C(W) = 0$ and for the average, $A\text{–}C(W) = 0.5$. Another measure associated with the OWA operator is the measure of dispersion [11,12], which is defined $\text{Disp}(W) = -\sum_{j=1}^{n} w_j \ln(w_j)$.

In [13] Yager suggested a useful approach to obtain the OWA operator. Consider the class of functions $f{:}[0,1] \rightarrow [0,1]$ such that $f(0) = 0$, $f(1) = 1$ and $f(x) \geqslant f(y)$ if $x \geqslant y$. We refer to these as BUM functions. Using these functions we can generate valid weights for an OWA operator, $w_j = f\left(\frac{j}{n}\right) - f\left(\frac{j-1}{n}\right)$. An important example is the case where $f(x) = x$, here we get $w_j = 1/n$. In [13] Yager related these BUM functions to Zadeh's concept [14] of linguistic quantifiers. This enabled the formulation of OWA operators based on linguistically expressed specifications.

The aggregation of criteria using the BUM function can be easily extended to the case where each of the criteria has an importance weight, $u_i \in [0,1]$. If we let $u_{\pi(j)}$ indicate the importance weight of the criteria with the jth largest value for a_i then we generate the OWA weights as $w_j = f\left(\frac{T_j}{T}\right) - f\left(\frac{T_{j-1}}{T}\right)$ where $T_j = \sum_{k=1}^{j} u_{\pi(k)}$ and $T = \sum_{i=1}^{n} u_i$. In this special case where $f(x) = x$ we obtain that $w_j = u_j$ and hence we get the usual weighted average.

3. Bonferroni mean operators

The wide variety of possible relationships between the criteria in multi-criteria problems motivates great interest in seeking aggregation functions that can be used to model these various possibilities. Here we investigate the capabilities of a class of aggregation operators called Bonferroni means. The Bonferroni mean was originally introduced in [5] and discussed more recently in [1,6].

Let $(a_1, \ldots, a_n)$ be a collection of values so that $a_i \in [0,1]$. Assume p and $q \geqslant 0$, then the general Bonferroni mean of these values is defined as