An Approach to Ordinal Decision Making

Ronald R. Yager

Machine Intelligence Institute, Iona College, New Rochelle, New York

ABSTRACT

We describe the problem of decision making under ignorance and discuss the need of introducing a decision attitude or disposition in solving this problem. We show the OWA operators provide a general framework for evaluating the alternatives when the payoffs are numeric values. We then turn to the problem of decision making when the payoffs are linguistic values drawn from an ordinal scale. A solution to this problem is then suggested using an ordinal version of the OWA aggregation method. We emphasize the role of the OWA weighting vector as a means for introducing the decision maker's attitude and investigate the representation of various prototypical decision attitudes. Two measures are provided for classifying the OWA weighting vectors used in the aggregations. The first of these measures is the degree of optimism implied by the vector. The second measures the degree of dispersion associated with the vector. In this work the interpretation of the decision making attitude, via the OWA weighting vector, as a kind of probability distribution is highlighted. In particular the i\textsuperscript{th} component in the vector is interpreted as the probability that the i\textsuperscript{th} best outcome will occur. This view helps unify the problems of decision making under ignorance and decision making under uncertainty and allows us to see the evaluation of an alternative, in both cases, as the formulation of an expected value. Furthermore this interpretation leads to a viewing of the measure of dispersion as a kind of entropy. In the ordinal environment this leads to the development of a new measure of entropy for ordinal probabilities.

KEYWORDS: decision making, OWA operators, ordinal probabilities, ordinal entropy

1. INTRODUCTION TO DECISION MAKING UNDER IGNORANCE

A classic problem in the decision making literature is the problem of decision making under ignorance. In this problem we have a decision
matrix of the type shown below:

\[
\begin{array}{c|ccc}
S_1 & S_j & S_n \\
\hline
A_1 & c_{11} & c_{1j} & c_{1n} \\
A_i & c_{ij} \\
A_q & c_{qi} \\
\end{array}
\]

In the above each \( A_i \) corresponds to a possible action (alternative). Each \( S_j \) corresponds to a value of a variable called the state of nature. The value \( c_{ij} \) corresponds to the payoff for selecting alternative \( A_i \) when the state of nature is \( S_j \). The problem faced by the decision maker is to select the best alternative. The choice of the best alternative of course depends upon the knowledge of the state of nature. In the environment of decision making under ignorance it is assumed that we have no knowledge about the state of nature. In this environment the decision maker replaces knowledge about the environment by assuming some particular decision making attitude. Among the decision attitudes discussed in the literature are the following:

1. **Pessimistic attitude.** Using this strategy, the decision maker selects for each alternative the worst possible outcome and then select the alternative that has the best worst outcome. This strategy is sometimes called the maximin strategy.

2. **Optimistic attitude.** Under this strategy, the decision maker selects for each alternative the best possible outcome and then selects the alternative that has the maximum best outcome.

3. **Hurwicz approach.** In this approach the decision maker selects some value \( \alpha \in [0, 1] \). Then for each alternative he takes the weighted average of the optimistic and pessimistic values:

\[
H = \alpha \text{Opt} + (1 - \alpha) \text{Pess}.
\]

He then chooses the alternative with the best \( H \).

4. **Normative approach.** In this approach for each alternative the decision maker takes the average of all the outcomes under that alternative and then selects the alternative with the best average.

The structure of the problem of decision making under ignorance can be expressed as follows:

1. For each alternative \( A_i \) calculate

\[
V_i = F(c_{i1}, c_{i2}, \ldots, c_{in}),
\]

where \( F \) is some aggregation function whose form depends upon the decision maker's assumed attitude.
2. Select the alternative $A^*$ such that its value $V^*$ is the maximum:

$$V^* = \max_i V_i.$$

The following table provides the function $F$ for the four strategies discussed above:

<table>
<thead>
<tr>
<th>Strategy</th>
<th>Aggregation function</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pessimistic</td>
<td>$F(c_{i1}, \ldots, c_{in}) = \min_j c_{ij}$</td>
</tr>
<tr>
<td>Optimistic</td>
<td>$F(c_{i1}, \ldots, c_{in}) = \max_j c_{ij}$</td>
</tr>
<tr>
<td>Hurwicz</td>
<td>$F(c_{i1}, \ldots, c_{in}) = \alpha \max_j c_{ij} + (1 - \alpha) \min_j c_{ij}$</td>
</tr>
<tr>
<td>Normative</td>
<td>$F(c_{i1}, \ldots, c_{in}) = \frac{1}{n} \sum_{j=1}^{n} c_{ij}$</td>
</tr>
</tbody>
</table>

In essence the function $F$ combines the different payoffs. It appears natural to begin to consider other functions that can be used to implement this aggregation. At the very least, such a function should be symmetric (commutative): the ordering of the arguments shouldn't matter. It also should be monotonic in the arguments: as the payoffs associated with an alternative in a row increase, the valuation of that alternative should not decrease. It should also be idempotent: if all the scores in one row are the same, then this score should be the value of that row.

The above structure implicitly assumes that the values in the payoff matrix are numeric values. In [1] Yager provided a general approach to the representation of the aggregation function in this numeric environment, which makes use of the OWA aggregation operators [2]. In that approach the decision attitude is captured by the form of the weighting vector used by the OWA aggregation operators. In this work we extend these ideas and this formulation to the situation in which the values of the payoff matrix are not numeric, but are linguistic values such as high, medium, and low. We note that Lamata and Moral [3] have also looked at this issue.

We first look at the general numeric approach suggested by Yager [1], and then turn to the problem in the linguistic ordinal environment. In investigating this problem, in addition to providing a structure for evaluating subjective linguistic decisions, a number of interesting and more generally useful formulations are developed. One of these results is a formulation for a weighted Max-Min aggregation when the objects are drawn from an ordinal scale. A second is the formulation of a measure of entropy in situations in which the probabilities are of an ordinal-linguistic nature.
2. A GENERAL APPROACH TO NUMERIC DECISION MAKING UNDER IGNORANCE

In [1] Yager introduced a whole family of operators to generalize the preceding aggregation. These operators are based upon the ordered weighted average (OWA) operators. As we shall subsequently see, the use of these operators provide a very interesting semantics to better understand the process of decision making under ignorance.

**Definition**  An ordered weighted averaging operator (OWA) of dimension \( n \) is a function

\[
F : \mathbb{R}^n \rightarrow \mathbb{R}
\]

that has associated with it a weighting vector

\[
W = \begin{bmatrix}
  w_1 \\
  w_2 \\
  \vdots \\
  w_n
\end{bmatrix}
\]

such that
1. \( w_i \in [0, 1] \),
2. \( \sum_i w_i = 1 \),

where for any set of values \( a_1, \ldots, a_n \)

\[
F(a_1, \ldots, a_n) = \sum_i w_i b_i,
\]

where \( b_i \) is the \( i \)th largest of the \( a_1, \ldots, a_n \).

A key feature of this aggregation process is the reordering of the elements, which essentially provides a nonlinear component to the aggregation.

Using these OWA aggregation functions, we can introduce a whole family of ways of evaluating alternatives under ignorance. It can easily be shown that these operators are monotone, symmetric, and idempotent.

Furthermore, in the following we show how the techniques described in the earlier part for handling decision under ignorance are special cases of this more general approach.
We see that for
\[ W = W^* = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \]
we get the purely optimistic decision, \( F(a_1, \ldots, a_n) = \text{Max}_i[a_i] \). For
\[ W = W_* = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} \]
we get the purely pessimistic model
\[ F(a_1, \ldots, a_n) = \text{Min}_i[a_i]. \]
For
\[ W_H = \begin{bmatrix} \alpha \\ 0 \\ \vdots \\ 0 \\ 1 - \alpha \end{bmatrix} \]
we get the Hurwicz model. For
\[ W_A = \begin{bmatrix} \frac{1}{n} \\ \vdots \\ \frac{1}{n} \end{bmatrix} \]
we get the normative case,
\[ F(a_1, \ldots, a_n) = \frac{1}{n} \sum a_i. \]

For different selections of \( W \) we obtain different models. In [1] Yager suggested using
\[ \alpha = \frac{1}{n-1} \sum_{i=1}^{n} w_i (n - i) \]
for measuring the degree of optimism associated with a particular selection of $W$. It can easily be shown that $\alpha$ always takes its value in the unit interval. We can easily see that

$$\alpha(W^*) = 1,$$
$$\alpha(W_+^*) = 0,$$
$$\alpha(W_R^*) = \alpha,$$
$$\alpha(W_A^*) = 0.5.$$ 

Thus, given some selection for $W$, we can measure the degree of optimism inherent in this procedure.

In particular, it is noted that the more of the weights are bunched near the top of $W$, the more optimistic the decision process, while if the weights are located near the bottom, the decision is pessimistic. The following theorem formalizes this observation.

**THEOREM** Assume $W$ and $W'$ are two OWA vectors of dimension $n$. Assume that for $q < r$ and $K$ positive

$$w'_q = w_q - K,$$
$$w_r = w_r + K,$$

while for all others

$$w_i = w'_i, \quad i \neq r, q.$$

Then

$$\alpha(W) > \alpha(W').$$

Proof

$$\alpha(W) = \frac{1}{n-1} \sum_{i=1}^{n} w_i(n - i),$$

$$\alpha(W') = \frac{1}{n-1} \sum_{i=1}^{n} w'_i(n - i)$$

$$= \frac{1}{n-1} \left[ \sum_{i \neq q, r} w_i(n - i) + w'_q(n - q) + w'_r(n - r) \right]$$

$$= \frac{1}{n-1} \left[ \sum_{i=1}^{n} w_i(n - i) + w_q(n - q) + w_r(n - r) - K(n - q) + K(n - r) \right],$$

$$\alpha(W') = \alpha(W) + \frac{1}{n} K(q - r).$$

Since $q < r$, we get our desired result $\alpha(W') \leq \alpha(W)$. \qed
Thus we see movement of weights down leads to a more pessimistic aggregation.

In [1] Yager suggested a semantics that can be associated with the use of the OWA operators in the environment of decision making under ignorance. Noticing that \( w_i \) have the properties of a probability distribution, \( w_i \in [0, 1] \) and \( \sum_i w_i = 1 \), it was suggested that we can view the weighting vector as a probability distribution. In this probability distribution \( w_i \) corresponds to the probability that the \( i \)th best thing will happen. Thus the pure optimist, with \( w_1 = 1 \), is essentially saying that the probability is one that the best thing will happen. The pure pessimist, with \( w_n = 1 \), says that the probability is one that the worst thing will happen. The neutral or normative decision maker is saying that all outcomes are equally likely.

In this view we see that the OWA aggregation

\[
F(a_1, \ldots, a_n) = \sum_i w_i b_i,
\]

where \( b_i \) is the value of the \( i \)th best outcome and \( w_i \) is interpreted as the probability of the \( i \)th best outcome, can be viewed as an expected value. Thus the OWA aggregation is a kind of expected value where constituents are positions of the outcome.

Used in this framework, we shall call the weighting vector \( W \) the \textit{dispositional probability distribution}.

Viewing the \( w_i \) as a kind of probability naturally allows us to introduce a measure of entropy associated with the vector \( W \),

\[
H(W) = -\sum_i w_i \ln w_i,
\]

to calculate the uncertainty associated with the dispositional probability distribution. We note that the pessimist and the optimist have zero entropy (they have assumed away all the uncertainty), while the neutral case has maximal entropy.

The introduction of the OWA aggregators as a means of evaluating the alternatives in the problem of decision making under ignorance provides us with a large array of potential ways of accomplishing the task. O'Hagan [4] suggested a methodology for easing the burden of selecting the appropriate weighting vector \( W \) in a given problem. He suggested that first we obtain from a decision maker a degree \( \alpha \) of optimism he wishes to use. Then we can select the weighting vector which has the maximal entropy for this degree of optimism. In particular, O'Hagan suggested solving the
following mathematical programming problem for the weights \( w_i \) associated with the aggregation:

\[
\text{Max} \quad -\sum_i w_i \ln w_i \\
\text{such that} \quad \sum_i w_i = 1, \\
\frac{1}{n-1} \sum_i w_i (n-1) = \alpha, \\
0 \leq w_i \leq 1.
\]

In [5] Filev and Yager investigate the analytic properties of the solution of this problem. In particular, they show that as the degree of optimism \( \alpha \) increases, the OWA aggregation value increases.

3. DECISION UNDER IGNORANCE WITH LINGUISTIC VALUES

In the preceding, in considering the problem of decision making under ignorance, we assumed the valuations for the outcomes associated with a given alternative (the \( c_{ij} \)) were drawn from a numeric scale. In many real world problems the information about the satisfaction associated with an outcome and a state of nature may be at best expressed in terms of some linguistic scale. For example, we may have values such as:

very high, 
high, 
medium, 
low, 
very low.

At best, one can associate with such a type of scoring a linear ordering amongst the valuations. Thus in the following we shall consider a scale

\[ L = \{L_1, L_2, \ldots, L_m\}, \]

such that \( L_i > L_j \) if \( i > j \), for use in evaluating the satisfaction of an alternative. Thus in the following we shall assume that the \( c_{ij} \in L \).

We note that the use of such a scale to provide the satisfaction information reduces the available operations to those of \( \text{Max} (\lor) \), \( \text{Min} (\land) \), and a type of negation [6].
Using the basic methodology of the previous section, we can suggest the following approach to decision making under ignorance with linguistic evaluations:

1. For each alternative $A_i$ calculate

   $$V_i = F(c_{i1}, c_{i2}, \ldots, c_{in}),$$

   where $F$ is some aggregation function whose form depends upon the decision maker's disposition.

2. Select the alternative $A^*$ such that its valuation $V^*$ is maximal:

   $$V^* = \max_i [V_i].$$

Because of the fact that the $c_{ij}$ are drawn from a linear scale, we are somewhat limited in the ways in which we can implement the function $F$. Two acceptable ways of implementing $F$ in this environment are the optimistic and pessimistic attitudes of the previous section:

- Optimistic: $F(c_{i1}, \ldots, c_{in}) = \max_j [c_{ij}].$
- Pessimistic: $F(c_{i1}, \ldots, c_{in}) = \min_j [c_{ij}].$

Using an ordinal form of the OWA operator introduced in [7], we can provide a much greater array of options for performing the aggregation $F$ used in the above methodology.

**Definition** A mapping $F : \mathbb{L}^n \rightarrow \mathbb{L}$

is called an ordinal OWA operator of dimension $n$ if it has an associated weighting vector $W = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix}$ such that

1. $w_j \in \mathbb{L}$,
2. $w_j \geq w_i$ if $j > i$,
3. $\max_j [w_j] = L_m$,

and where

$$F(a_1, \ldots, a_n) = \max_j [w_j \land b_j],$$

where $b_j$ is the $j$th largest of the $a_i$. 
We significantly note the second and third conditions on $W$, that is, the nondecreasing nature of the elements in $W$ and the fact that the last element in $W$ must attain the maximal element in $L$.

EXAMPLE Assume our scale is $L = \{L_1, L_2, L_3, L_4, L_5\}$, $m = 5$, and we are interested in aggregating $\langle L_3, L_2, L_4, L_1 \rangle$ using the following weighting vector:

$$W = \begin{bmatrix} L_1 \\ L_3 \\ L_5 \\ L_5 \end{bmatrix}.$$

Ordering the elements in $\langle L_3, L_2, L_4, L_1 \rangle$, we get

$$b_1 = L_4, \quad b_2 = L_3, \quad b_3 = L_2, \quad b_4 = L_1;$$

hence

$$\begin{align*}
F(L_3, L_2, L_4, L_1) &= \text{Max}[L_1 \land L_4, L_3 \land L_3, L_3 \land L_2, L_5 \land L_1], \\
F(L_3, L_2, L_4, L_1) &= \text{Max}[L_1, L_3, L_2, L_1] = L_1.
\end{align*}$$

The ordinal OWA operator can be shown to have the following properties [7]:

1. **Monotonicity.** $F(c_1, \ldots, c_n) \succeq F(d_1, \ldots, d_n)$ if $c_i \succeq d_i$ for all $i$.
2. **Idempotency.** If all $a_i = a$ for all $i$, then $F(a_1, \ldots, a_n) = a$.
3. **Commutativity.** The indexing of the arguments doesn't matter.

Using the concept of the ordinal OWA operator, we can provide a very general framework for suggesting various different solutions to the problem of aggregating outcomes under ignorance. In the discussion to follow we can consider that the weighting vector captures the attitude or disposition of the decision maker in this environment. We shall refer to $W$ when appropriate as the *dispositional vector*.

Using the structure of the ordinal OWA operator, we can easily implement the purely optimistic approach to decision making under risk in this ordinal environment. Consider the vector

$$W^* = \begin{bmatrix} L_m \\
\vdots \\
L_m \end{bmatrix}.$$
All weights in $W^*$ are equal to $L_m$, the maximal element in $L$. In this case, since $L_m \land c = c$ for any $c \in L$, we get

$$F^*(a_1, \ldots, a_n) = \max_i [L_m \land b_i] = \max_i [b_i] = b_1 = \max_i [a_i].$$

Thus, $W^*$ is the dispositional vector of the optimist.

The pure pessimist can also be modeled in the structure. Consider

$$W_* = \begin{bmatrix} L_1 \\ L_1 \\ \vdots \\ L_1 \\ L_m \end{bmatrix},$$

$$w_i = L_1 \quad \text{for} \quad i \neq n,$$

$$w_n = L_m.$$

Since for any $c \in L$ we have $L_1 \land c = L_1$ and $L_1 \lor c = c$, then in the case of $W_*$ we get

$$F_*(a_1, \ldots, a_n) = \max_i [L_i \land b_i] = L_1 \lor (L_m \land b_n)$$

$$= L_1 \lor b_n = b_n = \min_i [a_i],$$

and thus we get the pessimistic evaluation.

For simplicity, unless otherwise noted, in all of the following we shall assume that the arguments (the $a_i$) are indexed in such a manner that $a_i \geq a_j$ if $i < j$. If this indexing is followed, then

$$F(a_1, \ldots, a_n) = \max_i [w_i \land a_i].$$

We now show that the two vectors $W^*$ and $W_*$ provide the bounds on the ordinal OWA aggregation.

**Theorem** For any ordinal weighting $W$

$$F_*(a_1, \ldots, a_n) \leq F(a_1, \ldots, a_n) \leq F^*(a_1, \ldots, a_n).$$

**Proof** With $F(a_1, \ldots, a_n) = \max_j [w_j \land a_j]$, since $w_j \leq L_m$ for all $j$, we have

$$F(a_1, \ldots, a_n) \leq \max [L_m \land a_j] \leq F^*(a_1, \ldots, a_n).$$
Since \( w_j \geq L_1 \) for all \( j \), then

\[
F(a_1, \ldots, a_n) \geq \Max_{j \neq n} [L_1 \land a_j] \land L_m \land a_n \geq F_*(a_1, \ldots, a_n).
\]

Thus we see that the optimistic and pessimistic aggregations provide bounds on our aggregation process.

By appropriately selecting the weighting vector \( W \) we can provide for different formulations of aggregation, all lying between the optimistic and pessimistic cases. As we noted, each \( W \) can be seen to characterize some disposition or attitude on the part of the decision maker. Let us now look at some forms of \( W \) for which we can provide some semantics explaining the principle guiding its structure.

As we have noted, the optimist (Max) and pessimist (Min) are represented respectively by the vectors

\[
W_* = \begin{bmatrix}
L_m \\
\vdots \\
L_m
\end{bmatrix}, \quad W_* = \begin{bmatrix}
L_1 \\
\vdots \\
L_m
\end{bmatrix}.
\]

We now consider a vector denoted \( W^{[K]} \) and defined by

\[
W^{[K]}_i = \begin{cases}
L_1, & i = 1, \ldots, K - 1, \\
L_m, & i = K, \ldots, n.
\end{cases}
\]

In this situation \( F(a_1, \ldots, a_n) \), using the assumption that the \( a_i \)'s are indexed in decreasing order, has the following property:

\[
F(a_1, \ldots, a_n) = \Max_i [W^{[K]}_i \land a_i],
\]

\[
F(a_1, \ldots, a_n) = L_1 \land (a_1 \lor a_2 \lor \cdots \lor a_{K-1}) \lor (L_m) \land (a_K \lor \cdots \lor a_n),
\]

\[
F(a_1, \ldots, a_n) = (L_1 \land a_1) \lor (L_m \land a_K) = (L_1 \lor (L_m \land a_K) = L_m \land a_K,
\]

\[
F(a_1, \ldots, a_n) = a_K.
\]

Thus this vector selects the \( K \)th best value as the valuation. We notice that for \( K = 1 \) we get the optimist, \( W_* \), and for \( K = n \), we get \( W_* \).
Another and perhaps more interesting disposition weighting vector is the following:

$$W^{(L_K)} = \begin{bmatrix} L_K \\ L_K \\ \vdots \\ L_K \\ L_m \end{bmatrix}.$$ 

Here we have

$$w_i = L_K \quad \text{for} \quad i = 1, \ldots, n - 1,$$

$$w_n = L_m.$$

In this case

$$F(a_1, \ldots, a_n) = L_K \wedge (a_1 \lor a_2 \lor \cdots \lor a_{n-1}) \lor a_n,$$

$$F(a_1, \ldots, a_n) = (L_K \wedge a_1) \lor a_n,$$

$$F(a_1, \ldots, a_n) = (L_K \wedge \max_i [a_i]) \lor \min_i [a_i].$$

We shall call this the *Max-Min weighted average*. This formulation can be seen to be the analog to the Hurwicz type criteria used in the numeric case, $\alpha \max_i [a_i] + (1 - \alpha) \min_i [a_i]$.

For this vector we see that if $L_K = L_1$, then

$$F(a_1, \ldots, a_n) = \min_i [a_i]$$

and we have the pure pessimist. If $L_K = L_m$, then

$$F(a_1, \ldots, a_n) = \max_i [a] \lor \min_i [a] = \max_i [a_i],$$

which is the pure optimistic approach. In this environment we see that $L_K$ can be seen as some measure of optimism associated with the aggregation. It is obvious that as the optimism increases, $L_K$ increases, and the evaluation increases.

In the preceding we have essentially taken a weighted average of the two extremes, the maximum and minimum,

$$f(a_1, \ldots, a_n) = \left( L_K \wedge \max_i [a_i] \right) \lor \min_i [a_i]$$

with $L_K$ being our degree of optimism.
In this spirit we can consider a weighted average of any two symmetric aggregates. Assume the dimension of our aggregation is $n$. Let $q > p$ and $p + q = n + 1$. Consider the disposition

$$W_i = \begin{cases} L_1 & \text{for } i < p, \\ L_K & \text{for } p \leq i < q, \\ L_m & \text{for } i \geq q. \end{cases}$$

In this case

$$f(a_1, \ldots, a_n) = \max_i [w_i \land a_i],$$

$$f(a_1, \ldots, a_n) = \left( L_1 \land (a_1 \lor a_2 \lor \cdots \lor a_{p-1}) \right) \lor \left( L_K \land (a_p \lor \cdots \lor a_{q-1}) \right) \lor \left( L_m \land (a_q \lor \cdots \lor a_n) \right).$$

Since $a_i \geq a_j$ for $i < j$, then

$$f(a_1, \ldots, a_n) = (L_1 \land a_1) \lor (L_K \land a_p) \lor (L_m \land a_q),$$

$$f(a_1, \ldots, a_n) = L_1 \lor L_K \land a_p \lor a_q,$$

$$f(a_1, \ldots, a_n) = (L_K \land a_p) \lor a_q.$$

Thus, in this case we are taking a weighted average of the $p$ and $q$ best elements.

Before we consider the next class of aggregations, we induce a mapping called a unitor function [7].

**Definition** Assume $L$ is a linear scale with $m$ elements, $L = \{L_1, \ldots, L_m\}$. An ordinal unitor function is a mapping

$$H: [0,1] \rightarrow L$$

such that

$$H(r) = S_i, \quad \frac{i - 1}{m} \leq r < \frac{i}{m}, \quad i = 1, \ldots, m,$$

$$H(1) = S_m.$$

The unitor function can be seen to be an approximator to a linear function, $y = x$, in ordinary numeric cases. Figure 1 shows the unitor function for $m = 5$.

\begin{figure}[h]
\centering
\includegraphics[width=0.4\textwidth]{unitor_function.png}
\caption{An ordinal unitor function.}
\end{figure}
We now introduce, using the unitor function, a weighting $W_u$ which can be seen to be the analog of the normative aggregation, $w_i = 1/n$, in the numeric case. We define $W_u$ as a weighting vector of dimension $n$, with weights

$$w_j = H\left(\frac{j - 1}{n - 1}\right), \quad j = 1, \ldots, n.$$ 

Essentially this function divides the unit interval into $m$ pieces and maps onto the unit function $H$.

EXAMPLE Assume $m = 5$.

(a) $n = 3$:

$$w_1 = H(\frac{0}{3}) = L_1;$$

$$w_2 = H(\frac{1}{3}); \quad \text{thus } \frac{2}{3} \leq r < \frac{3}{3}; \quad \text{here } w_2 = L_3;$$

$$w_3 = H(1) = L_5.$$

(b) $n = 5$:

$$w_1 = H(\frac{0}{5}) = L_1;$$

$$w_2 = H(\frac{1}{5}); \quad \text{thus } \frac{1}{5} \leq r < \frac{2}{5}; \quad \text{hence } w_2 = L_2;$$

$$w_3 = H(\frac{2}{5}); \quad \text{thus } \frac{2}{5} \leq r < \frac{3}{5}; \quad \text{hence } w_3 = L_3;$$

$$w_4 = H(\frac{3}{5}); \quad \text{thus } \frac{3}{5} \leq r < \frac{4}{5}; \quad \text{hence } w_4 = L_4;$$

$$w_5 = H(1) = L_5.$$

Thus for $n = m$, we get $w_i = L_i$.

(c) $n = 8$:

$$w_1 = H(\frac{0}{8}) = L_1;$$

$$w_2 = H(\frac{1}{8}) = H(0.143); \quad \text{thus } 0 \leq r < \frac{1}{5}; \quad \text{hence } w_2 = L_1;$$

$$w_3 = H(\frac{2}{8}) = H(0.28); \quad \text{thus } \frac{1}{5} \leq r < \frac{2}{5}; \quad \text{hence } w_3 = L_2;$$

$$w_4 = H(\frac{3}{8}) = H(0.42); \quad \text{thus } \frac{2}{5} \leq r < \frac{3}{5}; \quad \text{hence } w_4 = L_3;$$

$$w_5 = H(\frac{4}{8}) = H(0.56); \quad \text{thus } \frac{3}{5} \leq r < \frac{4}{5}; \quad \text{hence } w_5 = L_3;$$

$$w_6 = H(\frac{5}{8}) = H(0.7); \quad \text{thus } \frac{4}{5} \leq r < \frac{5}{5}; \quad \text{hence } w_6 = L_4;$$

$$w_7 = H(\frac{6}{8}) = H(0.84); \quad \text{thus } \frac{5}{5} \leq r < \frac{6}{5}; \quad \text{hence } w_7 = L_5;$$

$$w_8 = H(\frac{7}{8}) = L_5.$
The following theorem shows the generality of the result obtained when \( n = m = 5 \).

**THEOREM**  For the weighting vector \( W_u \), when \( m = n \) we obtain \( w_i = L_i \).

**Proof**  Assume \( m = n = K \); then

\[
\frac{j - 1}{K - 1} - \frac{j - 1}{K} = \frac{j - 1}{(K - 1)K} [K - (K - 1)] = \frac{(j - 1)}{(K - 1)K} \geq 0;
\]

thus

\[
\frac{j - 1}{K - 1} \geq \frac{i - 1}{K}.
\]

Next consider

\[
\frac{j}{K} - \frac{j - 1}{K - 1} = \frac{1}{(K - 1)K} [j(K - 1) - (j - 1)K]
\]

\[
= \frac{1}{(K - 1)K} [j(K - 1) - K(j - 1)]
\]

\[
= \frac{1}{(K - 1)K} (jK - j - Kj + K)
\]

\[
= \frac{K - j}{(K - 1)K} \geq 0;
\]

hence

\[
\frac{i}{K} > \frac{j - 1}{K - 1}.
\]

---

4. MEASURES OF THE ORDINAL DISPOSITIONAL VECTORS

In [7] Yager introduced a measure of *orness* associated with the ordinal OWA vector. As we shall see, this measure can provide a measure, on the ordinal scale \( L \), of the degree of optimism associated with a vector \( W \).

**DEFINITION**  Assume \( F \) is an ordinal OWA operator of dimension \( n \) with dispositional vector

\[
W^T = [w_1, \ldots, w_n].
\]
The degree of optimism associated with this vector is defined as

$$\text{Optimism}(W) = F(h_1, \ldots, h_n),$$

where

$$h_j = H\left(\frac{n-j}{n-1}\right)$$

with $H$ being the unitor function associated using the scale $L$ given in the previous section. More specifically,

$$\text{Optimism}(W) = \cap_w \max_j \left[ w_j \land H\left(\frac{n-j}{n-1}\right) \right].$$

Let us look at the results we get when applying this measure to some of the dispositional vectors introduced in the earlier section.

Before proceeding we should note that $h_j$ is a nonincreasing function of $j$. In particular:

1. $h_1 = H\left(\frac{n-1}{n-1}\right) = H(1) = L_m$.
2. $h_n = H\left(\frac{n-n}{n-1}\right) = H(0) = L_1$.
3. Assume $n \geq j > j'$; then

$$\frac{n-j}{n-1} < \frac{n-j'}{n-1}$$

and hence

$$H\left(\frac{n-j}{n-1}\right) \leq H\left(\frac{n-j'}{n-1}\right);$$

thus

$$h_j \leq h_j'.$$

Consider first the case when $W = W^*$. In this case $W_j = L_m$ for all $j$. Hence

$$\alpha_{W^*} = \max_j [w_j \land h_j],$$

$$\alpha_{W^*} = \max_j [L_m \land h_j],$$

$$\alpha_{W^*} = L_m \land [h_1 \lor h_2 \lor \cdots \lor h_j].$$
Since $h_1 = L_m$, we get

$$\alpha_{w^*} = L_m \land L_m = L_m.$$ 

Thus in this case our degree of optimism is the maximal possible on our scale $L_m$.

Consider now $W = W_n$. In this case

$$w_i = L_1 \quad \text{for} \quad i < n,$$

$$w_n = L_m.$$ 

Using this, we get

$$\alpha_{w^*} = \max_{j} [w_j \land h_j],$$

$$\alpha_{w^*} = (L_1 \land (h_1 \lor h_2 \lor \cdots \lor h_{n-1})) \lor (L_m \land h_n),$$

$$\alpha_{w^*} = L_1 \lor (L_m \land L_1) = L_1.$$ 

Thus we see that the degree of optimism, as expected, is the minimal value possible $L_1$.

Let us now look at the Max-Min weighted average. Here we recall

$$f(a_1, \ldots, a_n) = (L_K \land a_1) \lor a_n.$$ 

In this aggregation

$$w_i = L_K \quad \text{for} \quad i < n,$$

$$w_n = L_m.$$ 

In this case

$$\alpha_{L_K} = \max_{j} [w_j \land h_j],$$

$$\alpha_{L_K} = (L_K \land (h_1 \lor h_2 \lor \cdots \lor h_{n-1})) \lor (L_m \land h_n).$$

Since $h_1 = L_m$ and $h_n = L_1$ and $h_j$ is nonincreasing, we get

$$\alpha_{L_K} = L_K \land (h_1 \lor L_1) \lor L_m = (L_K \land L_m) \lor (L_1 \land L_m),$$

$$\alpha_{L_K} = L_K \lor L_1,$$

$$\alpha_{L_K} = L_K.$$ 

Thus $L_K$ is the degree of optimism, as we conjectured earlier.
Consider now the case of $W^{[K]}$:

$$w_i = \begin{cases} L_1, & i < K, \\ L_m, & i \geq K. \end{cases}$$

In this case

$$\alpha_{w^{[K]}} = \max_j [w_i \land h_j],$$

$$\alpha_{w^{[K]}} = L_1 \land [h_1 \lor h_2 \lor \cdots \lor h_{K-1}] \lor L_m \land [h_K \lor h_{K+1} \lor \cdots \lor h_n],$$

$$\alpha_{w^{[K]}} = L_1 \lor L_m \cap h_K,$$

$$\alpha_{w^{[K]}} = h_K.$$

Thus

$$\alpha_{w^{[K]}} = h_K = H \left( \frac{n - K}{n - 1} \right).$$

Because of the nonincreasing nature of $h_K$, we see, as expected, that this decreases as $K$ increases. The actual formulation for $\alpha_{w^{[K]}}$ depends upon the relationship of $m$ and $n$. In particular, $h_K = L_i$, where

$$\frac{i - 1}{m} \leq \frac{n - K}{n - 1} < \frac{i}{m}.$$

Thus

$$i > \frac{(n - K)m}{n - 1}$$

and

$$i - 1 \leq \frac{(n - K)m}{n - 1},$$

$$i \leq \frac{(n - K)m}{n - 1} + \frac{n - 1}{n - 1}.$$

In the special case when $m = n - 1$, we get

$$i = n - K + 1;$$

hence $\alpha_{w^{[K]}} = L_{n - K + 1}$. 
We now consider the dispositional vector \( W_u \), where \( w_j = H((j - 1)/(n - 1)) \). We note that \( w_j \) is of course monotonically increasing. The measure of orness associated with this vector is

\[
\alpha = \text{Max} \left[ w_j \land h_j \right],
\]

\[
\alpha = \text{Max}_j \left[ H\left( \frac{j - 1}{n - 1} \right) \land H\left( \frac{n - j}{n - 1} \right) \right].
\]

For simplicity we first consider the case where \( n - 1 = m \). In this case

\[
H\left( \frac{j - 1}{n - 1} \right) = H\left( \frac{j - 1}{m} \right) = S_j \quad \text{and} \quad H\left( \frac{n - j}{n - 1} \right) = H\left( \frac{m + 1 - j}{m} \right);
\]

\[
\alpha = \text{Max}_j \left[ H\left( \frac{j - 1}{m} \right) \land H\left( \frac{m + 1 - j}{m} \right) \right]
\]

\[
= L_1 \land L_m \lor L_2 \land L_m \lor L_3 \land L_{m - 1}
\]

\[
= L_1 \land L_m \lor \text{Max}_{j=2,\ldots,n} \left[ L_j \land L_{m+2-j} \right].
\]

If \( m \) is even, it can be seen that

\[
\alpha = L_{m/2}.
\]

and if \( m \) is odd, then

\[
\alpha = L_{m+1/2}.
\]

The essential idea is that \( W_u \) tends to have a middle type optimism, just as in the neutral case.

In the following we shall suggest a measure of entropy that we can associate with the ordinal type weighting function. Assume our scale is \( L = \{L_1, \ldots, L_m\} \) and our weighting vector \( W \) is of dimension \( n \) with components \( w_i, i = 1, \ldots, n \). We first define the set

\[
Q = \{L_1\} \cup \{w_1, \ldots, w_n\}.
\]

Thus \( Q \) consists of all the scale values that appear in \( W \) plus the value \( L_1 \) if it doesn't appear in \( W \). Let \( p = \text{Card}(Q) - 1 \). We then define the entropy of \( W \), \( H(W) \), as

\[
H(W) = L_p.
\]
We first note that since \( w_n = L_m \), then for any \( Q \) we have \( \{L_1, L_m\} \subset Q \), and hence
\[
H(W) \geq L_1.
\]
In addition
\[
L \subset Q;
\]
hence
\[
H(W) \leq L_{m-1}.
\]
Thus we see that for any \( W \)
\[
L_1 \leq H(W) \leq L_{m-1}.
\]
The larger \( H(W) \), the more uncertainty.

**THEOREM** Assume \( W \) is a weighting vector of dimension \( n < m - 1 \). Then
\[
H(W) \leq L_n.
\]
Proof Assume \( n < m - 1 \); then \( \text{Card}(Q) \leq n + 1 \) and \( p \leq n + 1 - 1 \leq n \).

Thus, we see that if the dimension of \( W \) is less than \( m - 1 \), we have an upper bound on the entropy. Furthermore, this upper bound increases until \( n = m - 1 \).

Let us look at the measure of entropy we get for some of the prototypical weighting vectors we previously introduced.

Consider the pure optimist, \( W = W^* \); here \( w_i = L_m \) for all \( i \). In this case
\[
Q = \{L_1, L_m\},
\]
and thus \( p = 1 \), and hence \( H(W^*) = L_1 \), the smallest possible entropy. For, as we discussed, the optimist turns the situation into certainty by assuming the best thing will happen.

Consider the pure pessimist, \( W = W_* \); here
\[
w_i = L_1, \quad i \leq n,
\]
\[
w_n = L_m.
\]
In this case again \( Q = \{L_1, L_m\}, p = 1 \), and \( H(W_*) = L_1 \). Again, for this case we get the smallest entropy.
Consider now $W^{[K]}$. Here

$$w_i = \begin{cases} L_1, & i < K, \\ L_m, & i \geq K. \end{cases}$$

Again, we get $Q = \{L_1, L_m\}$, and thus $H(W^{[K]}) = L_1$. Thus, in the case when we are certain of the result, we have the value $a_K$, and the entropy is $L_1$, the smallest possible.

Consider now $W^{[L_K]}$. Here

$$w_i = L_K \quad \text{for} \quad i < n,$$
$$w_n = L_m.$$

In this case

$$Q = \{L_1, L_K, L_m\}$$

and we have $p = 2$ and $H(W^{[K]}) = L_2$. Thus this case has more uncertainty than $W^{[K]}$.

Consider now the unitor type weighting vector $W_u$, where $w_i = H((i - 1)/(m - 1))$. We first recall that this distribution was seen to be related to the uniform distribution in the numeric case. That being so, we would anticipate it would have the maximal entropy.

For the unitor type vector $W$, we see that if $n \geq m - 1$ then

$$Q = \{L_1, \ldots, L_m\} = L;$$

hence $p = m - 1$, and thus

$$H(W_u) = L_{m-1},$$

which is indeed the maximal entropy. If $n < m - 1$, we can show that $H(W_u) = L_n$, which is the largest possible entropy. Thus, we see that this measure of entropy is compatible with our expectations regarding the prototypical weighting vector.

In the following we shall try to provide some intuitive connection between the concept of entropy introduced here for these ordinal scales and the standard entropy measure. First we note that the dispositional vector $W$ can be viewed as a staircase function in the spirit of a cumulative probability distribution (see Figure 2).
The ordinal measure of entropy introduced here can be seen as related to the number of different individual steps we take in going from $L_1$ to $L_m$ via the $w_i$. The number of steps is in turn related to the size of the steps. The larger the steps, the fewer we need, and the smaller the cardinality of $Q$. In some sense, we can consider $W$ as inducing a probability distribution $\hat{P}_i$ as described in the following.

Assume

$$w_i = L_{f_i}.$$ 

where $f_i$ are integers in the interval 1 to $m$. We let

$$\hat{P}_i = \frac{f_i - f_{i-1}}{m - 1}, \quad i = 2, \ldots, m,$$

$$\hat{P}_1 = \frac{f_1 - 1}{m - 1}.$$

Consider now the classical entropy in this environment:

$$H(\hat{P}_i) = -\sum_i \hat{P}_i \ln \hat{P}_i.$$ 

Consider a case where we have $w_q^* = L_m$ and $w_{q-1}^* = L_1$. In this case $\hat{P}_q = 1$ and $\hat{P}_i = 0$ for all $i$. In this case the classical entropy is minimal. In our ordinal environment we have $Q = \{L_1, L_m\}$, and we also get minimal entropy. As the steps get smaller, the $\hat{P}_i$ become more dispersed and the classical entropy increases. However, if the steps get smaller, we have more different levels, and $Q$ gets bigger; thus the ordinal entropy increases. On the other hand, if the steps get bigger, the $\hat{P}_i$ become bigger for some $i$ and go to zero for other $i$. This results in a decrease in classical entropy. However, if the steps get bigger, we have less different values, and hence $Q$ becomes smaller—and this results in a decrease in the ordinal entropy.

In Yager's [2] original work on OWA operators, he introduced the entropy as a concept of dispersion, as a measure of the amount of the information in the arguments used in forming the OWA aggregation. In the following we see that the ordinal measure of entropy introduced is in this spirit. In particular, we see that as the number of arguments that appear in $F$ increases, the entropy increases.

Consider the aggregation

$$F(a_1, \ldots, a_n) = \max_i [w_i \land a_i],$$
where it is assumed that the $a_i$'s are indexed in decreasing order. We can provide a formulation for this that is more expressive for our purposes. Let us partition $W$ into classes

$$
V_1 = \{w_1, \ldots, w_{n_1}\},
$$
$$
V_2 = \{w_{n_1}, \ldots, w_{n_2}\},
$$
$$
\vdots
$$
$$
V_r = \{w_{n_{r-1}}, \ldots, w_n\}.
$$

such that $w_i$ and $w_j$ are in the same class if $w_i = w_j$.

We now place the $a_i$'s in classes $A_1, \ldots, A_r$ such that

$$
a_i \in A_j
$$

if $w_i \in V_j$. Let us denote

$$
V_i \land A_j = \operatorname{Max}_{w_j \in V_i, a_j \in A_i} [w_j \land a_j].
$$

We can express

$$
F(a_1, \ldots, a_n) = \operatorname{Max}_{i=1, \ldots, r} [V_i \land A_i].
$$

We first note that $r$ is equal to the number of different scale levels that appear in $W$. Consider $V_i \land A_j$. Since $w_j$ is the same for weights in $V_i$, then

$$
V_i \land A_j = V_i \land [\operatorname{Max}_{a_i \in A_i}].
$$

Because of the ordering of the $a_i$, we see that $\operatorname{Max}_{a_i \in A_i}$ is the first element associated with weight $w_i$. Thus, if we denote this as $d_i$, we set

$$
f(a_1, \ldots, a_n) = \operatorname{Max}_i [d_i \land V_i],
$$

where $V_i$ are the different scale values that appear in $W$, and $d_i$ is the largest aggregate appearing within that scale value.

Assume that $L_1 \neq w_1$, $L_1$ doesn't appear in the vector $W$. Then the number of terms appearing in $f(a_1, \ldots, a_n)$ is equal to $p$, and hence $L_p$ is a measure of the number of terms in the aggregation.

If $L_1 = w_1$, then

$$
f(a_1, \ldots, a_n) = \operatorname{Max}_{i=2} [d_i \land V_i].
$$

Furthermore, if the number of terms is equal to $p$, then $\operatorname{Card}(Q) = p + 1$, and again $H(P) = L_p$. Thus, we see that $H(P)$ is closely related to the
number of terms in the actual aggregation. We also see that $H(P)$ is closely related to the number of steps we take in going from $L_1$ to $L_m$ in $W$.

5. CONCLUSION

We have developed a structure for the evaluation of decisions under ignorance in situations in which the payoff matrix contains linguistic information. This structure is makes use of the ordinal version of the OWA aggregation operators. In this approach the decision maker’s dispositional attitude is represented by the weighting vector.

References


