AD-A044 408 TEXAS UNIV AT AUSTIN CENTER FOR CYBERNETIC STUDIES F/G 12/1

A COMMITTEE RANKING PROBLEM USING ORDINAL UTILITIES.(U)

JUL 77 R D ARMSTRONG, W D COOK N00014-75-C-0569

UNCLASSIFIED CCS-298 NL

1 OF 1

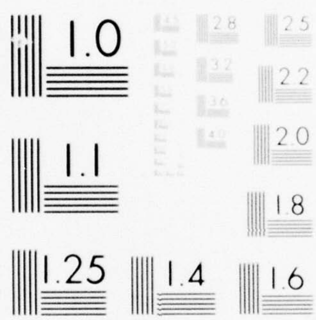
AD A044408

END

DATE FILMED

10-77

DDC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

12

See 1473

CENTER FOR CYBERNETIC STUDIES

The University of Texas
Austin, Texas 78712

Research Report CCS 298

A COMMITTEE RANKING PROBLEM USING ORDINAL UTILITIES

by

R. D. Armstrong
W. D. Cook*

July 1977

*Faculty of Administrative Studies, York University, Toronto, Ontario, Canada.

This research was partly supported by NRC Grant #A8966, and NSF Grant MCS-00100, and by Project NR047-021, ONR Contract N00014-75-C-0569 with the Center for Cybernetic Studies, The University of Texas. Reproduction in whole or in part is permitted for any purpose of the United States Government.

CENTER FOR CYBERNETIC STUDIES

A. Charnes, Director
Business-Economics Building, 203E
The University of Texas
Austin, Texas 78712
(512) 471-1821

ABSTRACT

This paper investigates aspects of obtaining a concensus ranking from ordinal rankings supplied by committee members. A procedure for calculating the distance between rankings which are represented by vectors is given along with an algorithm for deriving a min-max concensus ranking. In addition, an approach to the committee ranking problem which makes allowance for weights, partial rankings and multiple attributes is discussed.

1. *Introduction*

In [5] Kemeny and Snell define the median of a set of m rankings $\{A^{\ell}\}_{\ell=1}^{m}$ of n objects to be that ranking B for which

$$\sum_{\ell=1}^{m} d(A^{\ell},B) = \sum_{\ell}\sum_{i,j} |a_{ij}^{\ell}-b_{ij}| \tag{1.1}$$

is minimized. (a_{ij}^{ℓ}) is the matrix representation of the ranking A^{ℓ}; i.e., $a_{ij}^{\ell} = +1$ if i is preferred to j, -1 if j is preferred to i and 0 if i and j are tied in preference. The determination of B is important in cases where the A^{ℓ} are rankings provided by members of a committee, in that it represents one form of compromise or concensus. Cook and Saipe [4] describe a branch and bound algorithm for solving this problem. Some computational results relating to this algorithm are given in [3], [8] and [9].

In this article we investigate some further aspects of the concensus problem in committee ranking. In section 3, we consider the ℓ^{∞} norm rather than the ℓ^{1} norm of (1.1). In this case we minimize the maximum of the deviations rather than minimizing the total deviation. Since computer storage is a problem if the above matrix representation of rankings is preserved, we give an alternate representation. A formula for computing distances under this revised representation is derived in section 2. This will then be used to compute bounds in the algorithm which follows. An illustrative example is given along with a brief discussion of the computational aspects.

Section 4 presents a more general approach to the committee rahking problem. We investigate partial rankings in connection with the concept of a median and discuss its derivation. In addition, weighted rankings are examined. We conclude with a brief discussion of weight assignment.

2. *Computational Procedure for Calculating d(A,B)*

Kemeny and Snell [5] define the distance between two rankings A^k and A^ℓ to be

$$d(A^\ell,A^k) = \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} |a_{ij}^\ell - a_{ij}^k|. \tag{2.1}$$

Clearly, the computations involved in calculating this distance is straightforward when A^ℓ and A^k are available explicitly. However, if there are several committee members and a ranking matrix is saved for each member, a large amount of computer storage would be consumed. Because the ranking matrices are skew symmetric, the amount of necessary storage can be substantially reduced by saving only the elements above the diagonal. Additionally, because all the elements of the ranking matrices are 0, -1 or $+1$, the data could be packed in the computer. The latter technique would cause the computer code to be machine dependent, however, while the former would still require

$n(n-1)/2$ storage locations for each committee member. These drawbacks can be overcome by saving the ranking matrices on peripheral storage devices, but if it is necessary to access the rankings frequently this approach can be expensive and time consuming. In this section we present an alternative method for saving committee member rankings and calculating distance.

Two vector representations of the ℓ^{th} committee member's ranking of n objects are given by the following:

$$q^{\ell} = (q_1^{\ell}, q_2^{\ell}, \ldots, q_n^{\ell})^T$$

where $q_j^{\ell} = i$ implies that committee member ℓ ranks object j in the i^{th} position. Ties are indicated by assigning the (same) average ranking to tied objects. For example, a ranking

$$\begin{pmatrix} 5, \quad 6, \quad 9 \\ 7 \\ 4 \\ 3, \quad 2, \\ 1, \quad 8 \end{pmatrix} \qquad (2.2)$$

((2.2) is the notation used by Kemeny and Snell and is unsuitable for digital computers) is given by

$$q^T = (8.5,\ 6.5,\ 6.5,\ 5,\ 2,\ 2,\ 4,\ 8.5,\ 2).$$

The second vector representation is:

$$r^{\ell} = (r^{\ell}(1), r^{\ell}(2), \ldots, r^{\ell}(n))^T$$

where $r^{\ell}(i) = j$ implies that committee member ℓ ranks object j in the i^{th} position. $r(i) = -j$ indicates that the ℓ^{th} committee member is indifferent between object j and object $r^{\ell}(i+1)$. For the previous ranking

$$r^T = (-5, -6, 9, 7, 4, -3, 2, -1, 8)$$

These two vector representations are, of course, redundant as all the necessary information concerning a ranking can be obtained from either one. However, even if both representations are saved the storage requirements are considerably less than for the matrix representation.

It will now be shown how $d(A^{\ell}, A^{k})$ can be calculated efficiently using q and r. Through the remainder of the section the superscripts on q and r will be omitted. It will be assumed that *q always has a superscript ℓ and that r always has a superscript k.* Also, since r is a column vector, when object j is below object i it is meant that when looking down r i is encountered before j. Hence, in the previous example, 9 is above 7 but below 6.

Before we state and prove a theorem which will be instrumental in developing the procedure to calculate $d(A^{k}, A^{\ell})$, the following definitions are needed.

$T^{\rho}_{r(j)}$	The number of objects tied with $r(j)$ and below $r(j)$ in the k^{th} committee member's ranking, and also tied with $r(j)$ in the ℓ^{th} committee member's ranking.

$T^{+}_{r(j)}$	The number of objects tied with $r(j)$ and below $r(j)$ in the k^{th} committee member's ranking, and $r(j)$ preferred to them in the ℓ^{th} committee member's ranking.
$T^{-}_{r(j)}$	The number of objects tied with $r(j)$ and below $r(j)$ in the k^{th} committee member's ranking, and preferred to $r(j)$ in the ℓ^{th} committee member's ranking.
$T^{*}_{r(j)}$	The number of objects above $r(j)$ in the k^{th} committee member's ranking and tied with $r(j)$ in the ℓ^{th} committee member's ranking.
$H_{r(j)}$	The number of objects above $r(j)$ in the k^{th} committee member's ranking and $r(j)$ preferred to them in the ℓ^{th} committee member's ranking:

□ **Theorem 2.1**: The distance between A^k and A^ℓ as defined by (2.1) is given by:

$$\sum_{j=1}^{n} (2H_{r(j)} + T^{*}_{r(j)} + T^{+}_{r(j)} - T^{-}_{r(j)} - T^{0}_{r(j)}) \tag{2.3}$$

Proof: It will be proven by induction that

$$\sum_{j=1}^{M} (2H_{r(j)} + T^{*}_{r(j)} + T^{+}_{r(j)} - T^{-}_{r(j)} - T^{0}_{r(j)} + 2q_{r(j)} - 2j)$$

$$= \frac{1}{2} \sum_{i,j} |a^{k}_{ij} - a^{\ell}_{ij}| \tag{2.4}$$

where $\sum_{i,j}$ is a summation over the rows and columns $r(1), r(2), \ldots, r(M)$, and $M \leq n$. Then since

$\sum_{j=1}^{n} q_{r(j)} = \sum_{j=1}^{n} j$, the proof of the theorem follows:

Assume that (2.4) is true for $M = m-1$ (the proof for $M = 1$ is very similar to the one about to be given for

$M = m$ and is omitted). Then to prove (2.3) is true for $M = m$ it must be shown that the inclusion of column and and row $r(m)$ increases the summation of the right hand side by

$$2H_{r(m)} + T^{*}_{r(m)} + T^{+}_{r(m)} - T^{-}_{r(m)} - T^{0}_{r(m)} + 2(q_{r(m)} - m).$$

Note that all the rows and columns corresponding to objects above $r(m)$ have already been included in the summation with $M = m-1$.

Let β be the number of objects tied with $r(m)$ in the ℓ^{th} committee member's ranking. There are $q_{r(m)} - \frac{\beta}{2} - 1$ objects preferred to $r(m)$ in the ℓ^{th} committee member's ranking. Of these, the rows and columns corresponding to $(m-1) - H_{r(m)} - T^{*}_{r(m)}$ of them have already been included in the summation with $M = m - 1$. By including row and column $r(m)$ in summation (2.4), non-zero quantities that were not present before will be added in the following cases:

CASE	TOTAL CONTRIBUTION
1. The case where $r(m)$ is preferred to certain objects in committee member k's ranking and the reverse is true in committee member ℓ's ranking.	$2[(q_{r(m)} - \frac{\beta}{2} - 1) - ((m-1)-H_{r(m)}-T^{*}_{r(m)})-T^{-}_{r(m)}]$

CASE	TOTAL CONTRIBUTION
2. The case where objects are tied with $r(m)$ *and* below $r(m)$ in committee member k's ranking but not tied with $r(m)$ in committee member ℓ's ranking.	$T^{+}_{r(m)} + T^{-}_{r(m)}$
3. The case where $r(m)$ is preferred to certain objects in committee member k's ranking and is tied with $r(m)$ in committee member ℓ's ranking.	$\beta - T^{0}_{r(m)} - T^{*}_{r(m)}$

Therefore, the summation on the right hand side of (2.4) increases by

$$2H_{r(m)} + T^{*}_{r(m)} + T^{+}_{r(m)} - T^{-}_{r(m)} - T^{0}_{r(m)} + 2(q_{r(m)}-m)$$

and this proves the validity of (2.4). The proof of the theorem now follows by letting $M = n$ and observing that

$$\sum_{j=1}^{n} \dot{q}_{r(j)} = \sum_{j=1}^{n} j.$$

□

Although the method of calculating distance given in the theorem appears much more cumbersome than the direct application of (2.1), the structure of (2.3) makes the implementation in a computer code quite efficient. It is not felt that a detailed discussion of such a code is appropriate here, but a copy of a FORTRAN version of a code designed to calculate distance via (2.3) can be obtained upon request from the authors. An advantage of (2.3) is the previously mentioned reduction in storage requirements. Also, the rankings will generally be supplied in a form such as q or r and (2.3) will not require the construction of the ranking matrix.

We conclude this section with an example calculating the distance between two rankings by using (2.3).

Let

$$r^k = (-5, -6, 9, 7, 4, -3, 2, -1, 8)$$

and

$$q^\ell = (3, 1.5, 1.5, 8.5, 5.5, 4, 7, 8.5, 5.5)$$

The first ranking is the previous one and the second ranking is

$$\begin{pmatrix} 3, & 2 \\ & 1 \\ & 6 \\ 5, & 9 \\ & 7 \\ 8, & 4 \end{pmatrix}$$

$j = 1,\ r(1) = -5,\ q_5 = 5.5,\ T_5^0 = 1,\ T_5^+ = 0,\ T_5^- = 1,\ T_5^* = 0$

$H_5 = 0$

$j = 2,\ r(2) = -6,\ q_6 = 4,\ T_6^0 = 0,\ T_6^+ = 1,\ T_6^- = 0,\ T_6^* = 0$

$H_6 = 1$

$j = 3,\ r(3) = 9,\ q_9 = 5.5,\ T_9^0 = 0,\ T_9^+ = 0,\ T_9^- = 0,\ T_9^* = 1$

$H_9 = 0$

$$j = 4,\ r(4) = 7,\ q_7 = 7,\ T_7^0 = T_7^* = T_7^- = H_7 = 0$$

$$j = 5,\ r(5) = 4,\ q_4 = 8.5,\ T_4^0 = T_4^+ = T_4^- = T_4^* = H_4 = 0$$

$$j = 6,\ r(6) = -3,\ q_3 = 1.5,\ T_5^0 = 1,\ T_5^+ = T_5^- = T_5^* = 0,\ H_6 = 5$$

$$j = 7,\ r(7) = 2,\ q_2 = 1.5,\ T_2^0 = T_2^+ = T_2^- = 0,\ T_2^* = 1,\ H_2 = 5$$

$$j = 8,\ r(8) = 1,\ q_1 = 3,\ T_3^0 = 0,\ T_3^+ = 1,\ T_3^- = T_3^* = 0,\ H_3 = 5$$

$$j = 9,\ r(9) = 8,\ q_8 = 8.5,\ T_8^0 = T_8^+ = T_8^- = 0,\ T_8^* = 1,\ H_8 = 0.$$

By performing the summation of (2.3), the distance between the two rankings is found to be *34*.

3. *The L^∞ Norm Censensus Ranking*

This section will develop a branch and bound [7] algorithm to obtain an L^∞ norm or min-max consensus ranking for m rankings of n objects. This type of consensus ranking minimizes the maximum distance between itself and any ranking supplied by a committee member. The **algorithm to be described here will consider only** *complete* **rankings (see [7]) as candidates for the consensus ranking. The modifications to allow ties in the consensus ranking are obvious but will generally increase the size of the solution tree exponentially. We denote the min-max consensus ranking by the matrix** B **and let** $d^* = max\ (A^\ell, B)$. **Let**

$$d^{k\ell} = d(A^k, A^\ell),\ k = 1,\ 2,\ \ldots,\ m-1$$
$$\ell = i+1,\ \ldots,\ m$$

The algorithm begins by calculating the distance between every pair of rankings supplied by the committee members. A lower bound on d^* is then given by

$$LB = \langle d^{*k_o \ell_o}/2 \rangle \qquad (3.1)$$

where

$$d^{*k_o \ell_o} = \max_{k,\ell} d^{k\ell}$$

and $\langle X \rangle$ is the smallest integer greater than or equal to X.

A candidate consensus ranking is obtained by summing q^{ℓ_o} and q^{k_o} (the two rankings yielding $d^{*k_o \ell_o}$), and inspecting the magnitude of the components in the resulting vector. The object with the smallest entry will be most preferred in the candidate consensus ranking; the object with the next smallest entry will be ranked second and so on. Since we are only allowing complete ranking, ties are not allowed, and they will be broken in an arbitrary manner. The candidate ranking obtained by this process is denoted by Γ. A similar procedure for obtaining a consensus ranking is given by Kendall [6]. He sums all m q's before ordering and shows that the resulting ranking is a least squares consensus ranking when "distance" is defined differently (see [6], p.114). The Γ that we have obtained should be a good estimate of B because the min-max ranking represents a compromise between the most widely divergent committee members.

The next step in the algorithm is to calculate the distance between Γ and A^ℓ, $\ell = 1, 2, \ldots, m$. We define

$$UB = \max d(\Gamma, A^k), \quad k = 1, 2, \ldots, m. \tag{3.2}$$

If $UB = LB$, then the algorithm terminates with $B = \Gamma$. Otherwise, a branch and bound procedure is undertaken to obtain B.

The dichotomy performed at any stage of the tree search will be of the form

$$i > j \text{ or } j > i,$$

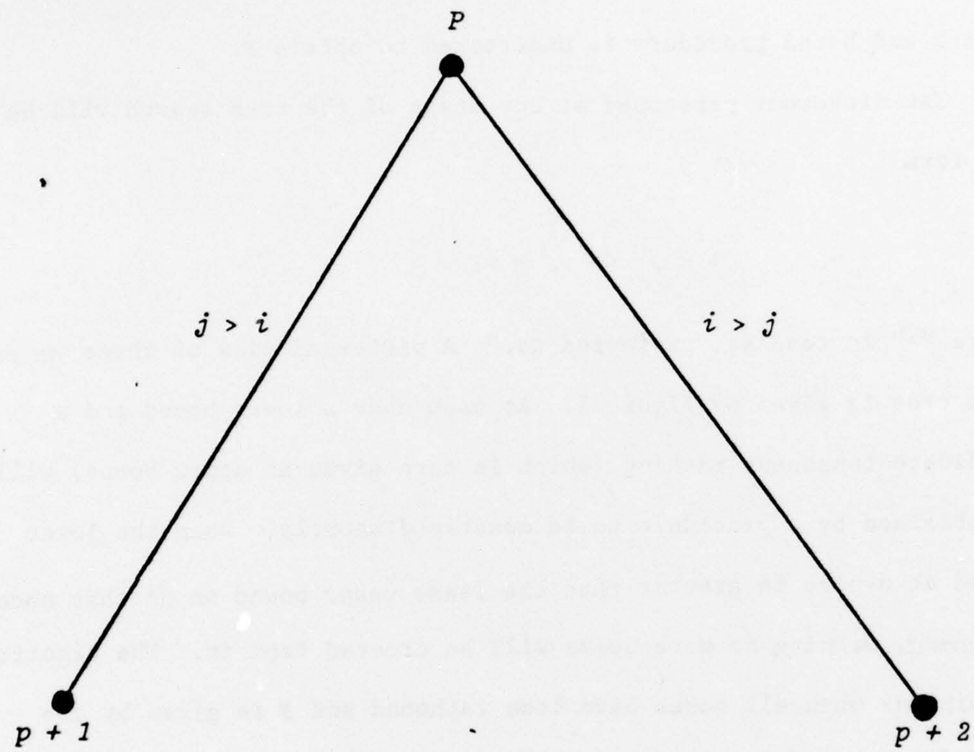
where ">" is read as "preferred to." A pictorial view of three *nodes* in a tree is given by figure 1. At each node a lower bound and a candidate consensus ranking (which in turn gives an upper bound) will be obtained by a procedure to be described shortly. When the lower bound at a node is greater than the least upper bound on d^* that node is *fathomed*, meaning no more nodes will be created from it. The algorithm terminates when all nodes have been fathomed and B is given by the candidate consensus ranking yielding the least upper bound.

Let P be the total number of *live nodes* (i.e., nodes that are not fathomed and have no branches emanating from them) at any stage in the branch and bound process. Let LB_p and Γ_p, $p = 1, 2, \ldots, P$ be, respectively, the lower bound and candidate consensus ranking at node p. From Γ_p the upper bound at node p, UB_p, is calculated from

$$UB_p = \max_k d(\Gamma_p, A^k)$$

and the least upper bound on d^* is

$$UB = \min_p \{UB_p\}.$$



<u>Figure 1</u>. A picture of three nodes from a possible branch and bound tree. At node $p + 1$ object j must be preferred to object i in any candidate consensus ranking and at node $p + 2$ the reverse must be true.

It will now be shown how Γ_p is obtained. Let C be a complete ranking which will be manipulated through pairwise interchanges to obtain Γ_p. Before the interchanges are attempted $C = \Gamma_s$ where s the *immediate predecessor* of node p or, for $p = 1$, $C = \Gamma$. If C satisfies the constraints placed on the ranking at that particular node we place $\Gamma_p = C$ and calculate UB_p. Otherwise, the interchanges necessary to make C feasible are performed. Notice that because of the transitive relationship between objects the constraint just imposed on the ranking may, in conjucture with preceeding constraints, create other constraints.

Once C is feasible *pairwise adjacent* (we assume C is represented in the form given by r) interchanges are attempted to improve C while maintaining feasibility. To determine when an improvement is possible we first calculate and record

$$d(C,A^k), \quad k = 1, 2, \ldots, n.$$

Let $d(C,A^{k_o}) = \max_k d(C,A^k)$. To improve the min-max distance between C and the A^k changes in the ranking given by C must make it closer to the k_oth committee member's ranking. Every pairwise adjacent change such as this will decrease $d(C,A^{k_o})$ by two or one. To see if an overall improvement has been made, all $d(C,A^k)$ within two of $d(C,A^{k_o})$ are checked to see what an effect such a change will make on these distances. If none of them would increase to $d(C,A^{k_o})$, the interchange is made, all distances are updated and a new A^{k_o} determined if necessary. The process is repeated until no improvement is possible and then Γ_p is placed equal to C. Notice that UB_p has already been calculated while determining Γ_p.

The calculation of LB_p involves determining penalties for the restrictions placed on Γ_p and adding them to the $d^{k\ell}$'s. The initial lower bound of (3.1) was obtained from the following axiom given by Kemeny and Snell [5, p.9]

$$d(A^k,B) + d(A^\ell,B) \geq d(A^k,A^\ell). \qquad (3.3)$$

Intuitively, (3.3) says that the best B can do is to split the difference between A^ℓ and A^k. As additional restrictions are placed on the structure of B a positive quantity may be added to original lower bound. To obtain this penalty we define $I_p^{k,\ell}$ to be the weighted total number of conflicts common to A^k and A^ℓ which are forced by the restrictions at node p. A weight of one is given when a conflict involves a tie and weight of two otherwise. A lower bound LB_p is obtained from the following theorem.

□ Theorem 3.1: A lower bound on the distance between any ranking which satisfies the restrictions at node p and the A^ℓ furthest from that ranking is given by

$$LB_p = \max_{k,\ell} I_p^{k,\ell} + \langle d^{k\ell}/2 \rangle. \qquad (3.4)$$

Proof: The distance $d^{k\ell}$ is the weighted sum of the conflicts between A^k and A^ℓ where a weight of one is given to a conflict involving a tie and a weight of two otherwise. The best we can hope to do with these conflicts in the min-max sense is to split them. Any additional conflicts enforced by the restrictions which are common to *both* A^k and A^ℓ must be added because all original conflicts are still present. Thus, by taking the maximum of

$I_p^{k,\ell} + \langle d^{k\ell}/2\rangle$ over all k and ℓ LB_p is obtained.

□ This completes the proof.

We will summarize with a step by step statement of the algorithm which will describe possible criteria for constructing the branch and bound tree.

STEP 1: Calculate $d^{k\ell}$ for $1 \le k < \ell \le n$ and determine $d^{*k_o\ell_o}$ equal to the minimum $d^{k\ell}$. Sum q^{ℓ_o} and q^{k_o} and derive the ranking Γ in the manner previously described. Calculate LB from (3.1) and UB from (3.2). If $LB = UB$, go to step 7. Otherwise, set $C = \Gamma$, $p = 1$ and go to step 2.

STEP 2: Attempt to improve C through pairwise adjacent interchanges, updating C and UB_p when improvements are made. After it is determined that no improvement is possible from pairwise adjacent interchanges, set $\Gamma_p = C$. If $UB_p < UB$, set $UB = UB_p$ and save Γ_p as the best complete ranking found so far. Go to step 3.

STEP 3: Search over all live nodes to find the one with the least lower bound. If no live nodes exist, go to step 7. Otherwise, denote the node with the least lower bound by node s. Go to step 4.

STEP 4: Choose the two objects furthest from one another in Γ_s and whose relationship is not already constrained. Let these two objects be given by Y and Z with Y preferred to Z in Γ_s. Create two new nodes from s by restricting $Y > Z$ at node $p + 1$ and $Z > Y$ at node $p + 2$. Check over the previous restrictions to see if, in conjunction with either $Y > Z$ or $Z > Y$, any additional restrictions are automatically imposed from the transitive relationship of the ranking. If any restrictions are forced, add them to the list of restrictions at the appropriate node. Go to step 5.

STEP 5: Calculate $I_{p+1}^{k,\ell}$ for $1 \le k < \ell \le n$ and determine LB_{p+1} from (3.4). If $LB_{p+1} > UB$, fathom node $p + 1$. Otherwise, set $\Gamma_p = \Gamma_s$ and go to step 6.

STEP 6: Calculate $I_{p+2}^{k,\ell}$ for $1 \le k < \ell \le n$ and determine LB_{p+2} from (3.3). If $LB_{p+2} > UB$, fathom node $p + 2$. Otherwise, set $C = \Gamma_s$ and force feasibility by interchanging the rankings of objects. Calculate $d(C,A^k)$, $k = 1, 2, \ldots, n$. Set $p = p + 2$ and go to step 2.

STEP 7: An optimal min-max consensus ranking has been obtained. The ranking yielding UB is the optimal ranking B and $d^* = UB$.

Example: To demonstrate the algorithm we will solve a sample problem with three committee members and five objects. The ranking are as follows:

$$q^1 = (1, 2, 3, 4, 5),\quad q^2 = (1.5, 3, 5, 1.5, 4),\quad q^3 = (4, 2, 2, 2, 5)$$

$$r^1 = (1, 2, 3, 4, 5),\ r^2 = (-1, 4, 2, 5, 3),\ r^3 = (-2, -3, 4, 1, 5)$$

By calculating $d^{k\ell}$ from (2.3),

$$d_{12} = 7,\quad d_{13} = 9 \text{ and } d_{23} = 10.$$

Committee members 2 and 3 are furthest apart, so q^2 and q^3 are summed to obtain $(5.5, 5, 7, 3.5, 9)$. The ranking Γ is then given by $(4,2,1,3,5)$, where this representation is defined as r.

$$LB = 5 \text{ and } UB = 8$$

$$d(\Gamma,A^1) = 8,\ d(\Gamma,A^2) = 5$$

and $d(\Gamma, A^3) = 5$.

C is placed equal to Γ, and by attempting pairwise adjacent interchanges it is seen that C can be improved by interchanging objects 4 and 2. We then have

$$UB = UB_1 = 7 \text{ and } \Gamma_1 = (2, 4, 1, 3, 5).$$

The first dichotomy is $2 > 5$ or $5 > 2$. Corresponding to the constraint $5 > 2$, $I^{2,3} = 2$ because $5 > 2$ is in conflict with the rankings of committee member's 2 and 3. Thus, the lower bound on any ranking with $5 > 2$ is $5 + 2 = 7$. But $UB = 7$, hence, rankings with $5 > 2$ need not be considered. Although it was not stated explicitly in the algorithm, it is clear that a branch with the constraint $5 > 2$ is not required and the constraint $2 > 5$ can be imposed at node 1. Similarly, the algorithm will impose the restrictions $2 > 3$ and $4 > 5$ at node 1.

The next dichotomy to be considered is $4 > 3$ or $3 > 4$. The latter restrictions, along with those already present, require $3 > 5$ and $2 > 4$. Nodes 2 and 3 are now created. The constraints at node 2 are $2 > 5$, $2 > 3$, $4 > 5$ and $4 > 3$. The constraints at node 3 are $2 > 5$, $2 > 3$, $4 > 5$, $3 > 4$, $3 > 5$ and $2 > 4$.

$$I_2^{1,2} = 0, \quad I_2^{1,3} = 1, \quad I_2^{2,3} = 0$$

and $$UB_2 = 1 + 5 = 6.$$

$$I_3^{1,2} = 0, \quad I_3^{1,3} = 0, \quad I_3^{2,3} = 2$$

and $UB_3 = 7$. Thus, node 3 is fathomed.

The algorithm next considers node 2 for further partitioning. The dichotomy $5 > 1$ or $1 > 5$ is performed, and it is seen that the restriction $5 > 1$ implies $4 > 1$ and $2 > 1$. Hence, $5 > 1$ yields $I^{1,2} = 5$ and $1 > 5$ is imposed at node 2.

Nodes 4 and 5 are now created by the dichotomy $2 > 1$ or $1 > 2$. At node 4 we have the constraints $1 > 5$ and $2 > 1$ in addition to those previously listed at node 2. At node 5, the additional constraints $1 > 5$, $1 > 2$ and $1 > 5$ are present.

$$\Gamma_4 = (2, 4, 1, 3, 5), \quad LB_4 = 6$$

and

$$UB_4 = 7$$

At node 5

$$LB_5 = 6, \quad C = (1, 4, 2, 3, 5)$$

Interchanging the order of 1 and 4, $\Gamma_5 = (4,1,2,3,5)$ and $UB_5 = 7$. Still at node 5, it is seen that $4 > 2$ or $2 > 4$ provide upper bounds greater than or equal to UB and node 5 is fathomed.

Returning to the only live node (node 4), the dichotomy $4 > 1$ or $1 > 4$ is applied and two more nodes created. At node 6 we have $UB_6 = LB_6 = 7$ and $_6 = (2,4,1,3,5)$. Node 6 is fathomed.

At node 7, $1 > 4$ implies $1 > 5$ and $1 > 3$. $UB_7 = 7$ and $LB_7 = 6$, with $\Gamma_7 = (2,1,4,3,5)$. By attempting further dichotomies it is shown that the restrictions $3 > 5$ and $2 > 4$ are forced. The latter restriction yields $I^{2,3} = 2$ and $LB_7 = 7$. There are no longer any live nodes on the tree and the algorithm terminates.

Alternate optimal ranking are given by (*2,1,4,3,5*), (*4,1,2,3,5*) and (*2,4,1,3,5*) with $d^* = 7$.

As is the case with every branch and bound algorithm there are many ways that it can be implemented. Refinements in the algorithm presented here, along with a thorough computational study will be the subject of future research by the authors.

4. *General Ranking Structures*

In many decision making situations involving committee ranking and the associated problem of concensus derivation, a more general frameword than that posed by Kemeny and Snell must be adopted. We cite a few examples.

Military officers' promotion boards are assembled for the purpose of deciding which members of a group of candidates (officers) should be elevated to the next highest rank. Each senior officer on the committee must assess each candidate in relation to a number of attributes, and must base the assessment on information contained in performance evaluation reports as well as any additional information the senior officer in question might possess. Any given board member may be capable of ranking only a proper *subset* of the candidates - not necessarily being able to rank all of them in an order of preference. One committee member may provide more information (or be of a higher rank) than another member, and, thus, should be weighted differently in terms of his or her contribution.

The assessment of proposed *R* and *D* projects by grants committees is an area that promises to be one in which the concept of priority ranking and concensus will assume a major role. Since the committee members come from a variety of disciplines we again are confronted with the possibility of having to deal with partial rankings and weights.

Federal spending in defence construction has become an issue of increasing importance in recent times. See Brightwell et. al. [2]. The problem is a multiphase one in which funds must be dispensed among various major military bases, then across numerous sections on each base, and finally across a set of proposed projects within each section. A general ranking framework allowing for such things as partial ranks and weights is required to be able to deal with problems of this nature.

To see how one can deal with partial ranking and weighting we investigate the original model due to Kemeny and Snell [5], as well as the model of the previous sections and see how the consensus concept is approached.

Weighted Rankings

First let us examine the effect of associating weights to committee members. With reference to the ℓ^1 norm, Kemeny and Snell's definition becomes

$$M(\hat{B}) \equiv \sum_{\ell=1}^{m} w^{\ell} d(A^{\ell},\hat{B}) = \underset{B}{\text{minimum}} \sum_{\ell=1}^{m} \sum_{i=1}^{n} \sum_{j=1}^{n} w^{\ell} |a_{ij}^{\ell} - b_{ij}|, \qquad (4.1)$$

with w^{ℓ} representing the relative influence of member ℓ and $\hat{B}$ the median ranking. In [4] we define i to be an "eligible immediate predecessor" of j if in any ranking B in which i immediately preceeds j, no improvement in $\sum_{\ell=1}^{m} d(A^{\ell},B)$ is obtained by interchanging i and j or by tying them. An "eligible immediate successor" and "tie" are defined analogously. For the general case (4.1) analogous definitions are obvious.

Let m_{ij} denote the total member weight in favor of not tying objects i and j, and r_{ij} the total weight ranking i and j in the non-natural (reverse) order. That is

$$m_{ij} = \sum_{\ell=1}^{m} w^{\ell} |a_{ij}^{\ell}| \tag{4.2a}$$

and

$$r_{ij} = \sum_{=1}^{m} w^{\ell} \delta_{ij} \text{ where } \delta_{ij}^{\ell} = \begin{cases} 1 \text{ if } a_{ij}^{\ell} = -1 \\ 0 \text{ otherwise.} \end{cases} \tag{4.2b}$$

Theorem 4.1: For any two objects i and j $(i < j)$ i is an eligible immediate predecessor of j if $r_{ij} \le m_{ij} - (\sum_{\ell=1}^{m} w^{\ell})/2$,

an eligible immediate successor of j if $r_{ij} \ge (\sum_{\ell=1}^{m} w^{\ell})/2$,

and is eligible to tie with j if $m_{ij} - (\sum_{\ell=1}^{m} w^{\ell})/2 < r_{ij} < (\sum_{\ell=1}^{m} w^{\ell})/2$.

The proof of this theorem follows in a straightforward manner from that given in [4] for the special case (1.1), and is therefore omitted.

Defining m_{ij} and r_{ij} as per (4.2a) and (4.2b) respectively, and utilizing theorem 4.1, we can apply the algorithm of [4] in the usual manner to derive the median $\hat{B}$ as defined in (4.1).

Turning to the ℓ^{∞} norm problem we note that the solution procedure is based on the distance between *pairs of committee members* rather than on the distance between each committee member and the unknown median, as in the case of the ℓ^{1} norm. Reconsidering (3.1), for example, the lower bound is defined thus since in the case of equally weighted members the optimal position relative to any pair is the point midway between that pair. With a general set of weights $w^1, w^2, \ldots, w^m$ the optimal position is at distances $w^\ell d^{k\ell}/(w^k+w^\ell)$ and $w^k d^{k\ell}/(w^k+w^\ell)$ from members k and ℓ respectively. (3.1) then becomes

$$LB = \langle \max_{k,\ell} \{w^k w^\ell d^{k\ell}/(w^k+w^\ell)\}\rangle .$$

(3.2) becomes

$$UB = \max_{\ell} w^\ell d(\Gamma, A^\ell).$$

(3.4) would need to be redefined accordingly.

Partial Ranking: Single Attribute

The definition as proposed by Kemeny and Snell covers only special cases to the extent that it is assumed that each member supplies an "all inclusive" (a.i.) ranking of the objects, that is, a preference ordering of all n objects. In many instances an individual committee member may be capable of supplying a preference ranking corresponding to only a subset of the objects. Thus, let us assume that each member ℓ supplies a ranking of all candidates or objects in a subset $S^{\ell} \subseteq \{1, 2, \ldots, n\}$. We consider first the case in which individuals are ranked on the basis of a single attribute. In this instance if $i > j > k$ then automatically $i > k$ by this member, and we get no conflicts. This, we shall see, is not the case in the multiple attribute problem.

It is noted that in case $S^{\ell_0} \cap S^{\ell} = \phi$ for all $\ell \neq \ell_0$, S^{ℓ_0} can be ranked as a set of objects on its own. As far as the existing committee is concerned, then, no object in S^{ℓ_0} can be directly compared to any object in $\tilde{S}^{\ell_0}$, the compliment of S^{ℓ_0}. In the R and D example, S^{ℓ_0} might represent all projects in the humanities category with $\tilde{S}^{\ell_0}$ denoting all other projects. If no comparison can be drawn between those elements inside S^{ℓ_0} and those outside, a ranking for each of S^{ℓ_0} and its compliment would need to be supplied, and then the R and D budget split up between these two categories by means of some other method.

We assume, therefore, that for each member ℓ_1 there exists a member ℓ_2 such that

$$S^{\ell_1} \cap S^{\ell_2} \neq \phi. \tag{4.3}$$

In this case each object will be comparable with every other object. The ranking A^{ℓ} supplied by member ℓ includes the elements of S^{ℓ} only. With reference to the median, let us define a matrix representation of A^{ℓ} by (a_{ij}^{ℓ}) where

$$a_{ij}^{\ell} = \begin{cases} +1 \text{ if } i \text{ is preferred to } j \\ -1 \text{ if } j \text{ is preferred to } i \\ 0 \text{ if either } i \text{ and } j \text{ are tied or are not compared.} \end{cases} \qquad (4.4)$$

In order to determine a median ranking $\hat{B}$, which will be that ranking satisfying (4.1), we need to derive conditions under which an object i is an eligible predecessor, successor or tie of another object j. We adopt the convention that the final decision as to the preference ordering of any pair (i,j) must be a function only of the orderings provided by those members who compared i and j. A member who does not commit himself to comparing i and j cannot influence their ordering in $\hat{B}$. Let S_{ij} denote those committee members ℓ who compared i and j when their rankings A^{ℓ} were constructed. r_{ij} and m_{ij} are defined as in (4.2a) and (4.2b). We state without proof the following:

□ Theorem 4.2: For any two objects i and j $(i < j)$, i is an eligible immediate predecessor of j if $r_{ij} \le m_{ij} - (\sum_{\ell \in S_{ij}} w^{\ell})/2$

an eligible immediate successor of j if $r_{ij} \ge (\sum_{\ell \in S_{ij}} w^{\ell})/2$

and is eligible to tie with j if $m_{ij} - (\sum_{\ell \in S_{ij}} w^{\ell})/2$

$< r_{ij} < (\sum_{\ell \in S_{ij}} w^{\ell})/2.$

Theorem 4.2 used in connection with the branch and bound method of [4] will lead to an a.i. median ranking $\hat{B}$ as a result of assumption (2.3).

In reference to determining the min max (ℓ^{∞}) ranking, distances $d^{k\ell}$ are computed as before allowing of course for the modification given in (4.4). Lower and upper bounds are computed as indicated in the previous subsection. The sets S_{ij} play no part in developing the min max ranking.

Multiple Attributes

In many instances candidates are assessed on the basis of a set of R attributes. For example, in the construction projects case attention must be paid to safety and security implications, impact on economy of resources, long run effect on operations, and duration over which the benefit of the project will be realized.

For each member ℓ let $S^{\ell r}$ denote that subset of $\{1, 2, \ldots, m\}$ of objects on which member ℓ can make a decision regarding a preference ranking according to attribute r. Let $(a_{ij}^{\ell r})$ represent the matrix associated with the ranking of $S^{\ell r}$. For one reason or another one attribute may be deemed more important or less important than some other attribute. In the assessment of defence construction project proposals, projects classified as pertaining mostly to security and safety, for example, are considered to be of higher priority than those categorized as general maintenance projects. The assignment of attribute weights is generally a subjective task. In addition, these weights can vary from year to year particularly in the construction projects example. Specifically let v^{r_0} be a weight expressing the relative importance of attribute r_0 to the other attributes. In determining the median, for example, we then weight each matrix $(a_{ij}^{\ell r_0})$ by the factor $w^{\ell} v^{r_0} / \sum_{r=1}^{R} v^{r}$. The median $\hat{B}$ must be defined by

$$M(\hat{B}) = \underset{B}{\text{minimum}} \sum_{\ell} \sum_{r} \sum_{i} \sum_{j} (w^{\ell} v^{r} / \sum_{r} v^{r}) |a_{ij}^{\ell r} - b_{ij}|. \qquad (2.5)$$

Weights as Functions of Information Added

The problem of assigning weights to committee members is an extremely unsatisfying issue to deal with. In almost any problem involving group decision making there is always the question as to whether one member's contribution to the deliberations is of more value than that of another member. In the priority ranking problem described above, each member utilizes information available to him to rank objects (candidates, projects, etc.) on the basis of various attributes possessed by those objects. The value of the information can be judged in various ways. Some of these are: 1) the number of years of experience the member has had on similar committees; 2) the rank or professional status of the member; 3) the personal experience of the member in the trade involved. (A mathematician, in assessing a research proposal, could call upon his own experience as a researcher in the technical area relating to the proposal); 4) number of years a member has been personally associated or involved with the project or candidate in question. (The qualifications of a candidate being considered for promotion can be better assessed by an officer who personally supervised that candidate for some length of time than by an officer who has had no personal connection with that candidate.)

It is clear from the preceding paragraph that numerous factors must be brought into focus in assigning weights to the decisions of various committee members. Factors other than those mentioned in (1), (2), (3) and (4) would be those associated with the special problem being considered at the time. Every ranking problem is characterized by special considerations and a need for a vast amount of subjective judgement. Consequently, it appears that the only general statement which can be made concerning the determination of weights, is that weights assigned to committee members should be a function of the information contributed by those members when delegated to rank a set of objects. There are, of course, as many definitions of information as there are problems requiring the input of the said information by decision makers. Generally, the information contributed by member ℓ on object i relative to attribute r can be defined as

$$I_i^{\ell r} = \begin{cases} f(x_1, x_2, \ldots, x_p) & \text{for each } \ell \text{ for which } i \in S_{\ell r} \\ 0 \text{ if } i \notin S_{\ell r}. \end{cases} \tag{4.6}$$

The parameters $x_1, x_2, x_3, \ldots, x_p$ might represent quantitative measures of (1) to (4) in the preceding paragraph, in addition to any other factors specific to the problem in question. f may be simply an average of the values represented by $x_1, x_2, \ldots x_p$. f can, however, be a more formal measure of overall contribution of the member. This decision would need to be tailored to the particular problem under investigation.

Given definition (4.6), the weight assigned to member ℓ_0 for his inputs regarding the r^{th} attribute of object i can be defined as

$$w_i^{\ell_0 r} = I_i^{\ell_0 r} / \sum_{\ell=1}^{m} I_i^{\ell r} . \tag{4.7}$$

The definition of the median as given in (2.5) illustrates that it is the pairwise comparison of objects which must be weighted, and not individual objects. A reasonable weighting factor on the pair (i,j) (that is on the term $|a_{ij}^{\ell_0 r} - b_{ij}|$) is, therefore,

$$w_{ij}^{\ell_{0}r} = \begin{cases} (w_i^{\ell_{0}r} + w_j^{\ell_{0}r})/2 \text{ if both } w_i^{\ell_{0}r} > 0 \text{ and } w_j^{\ell_{0}r} > 0 \\ 0 \text{ otherwise.} \end{cases} \quad (4.8)$$

Thus, in determining a median preference ranking of the objects in $\{1, 2, ..., n\}$ the relative ordering of each pair (i,j) is a function of the decisions of those members who compare i and j, and only those members. The weight assigned to a member who ranked i and j on the basis of some attribute is the average of the ratios of his contributed information to the total information supplied by all members who ranked i and j relative to that attribute. (4.5) then becomes

$$M(\hat{B}) = \underset{B}{minimum} \sum_{\ell} \sum_{r} \sum_{i} \sum_{j} (w_{ij}^{\ell r} v^r / \sum_{r} v^r) |a_{ij}^{\ell r} - b_{ij}|. \quad (4.9)$$

The purpose of this subsection has been to outline a methodology for determining weights to be assigned to committee members' decisions. The concept of information is an intuitively natural one to explore in this regard. The difficulty arises from trying to quantify what we mean by information. In addition, there is the related problem of removing bias from committee decisions. In a recent paper by Warner [10] a methodology is given for extracting information from advocates through the preparation of pro and con briefs on a topic of controversy. The method is not immediately applicable to the preference ranking problem, although it does give some promise of providing a systematic unbiased means of assigning weights. The investigation of such a method in connection with the ranking problem may be the topic of later research.

REFERENCES

1. S.A. Brightwell and W.D. Cook, "Unifying Group Decisions Relating to Cardinal Preference Ranking; with an Application to the Operation of a Military Promotion Board,: submitted to *Cahiers du Centre d'Etudes de Recherche Operationnelle*.

2. S.A. Brightwell, W.D. Cook and S.L. Mehndiratta, "Priority Assignment to Command Construction Projects,: *INFOR*, Vol. 13, #3 (1975).

3. W.D. Cook and H. Farr, "A Computer Program for Determining the Median of a Set of Rankings," *Defence Research Board Staff Note* (1975).

4. W.D. Cook and A.L. Saipe, "Committee Approach to Priority Planning: The Median Ranking Method,: forthcoming in *Cahiers du Centre d'Etudes de Recherche Operationnelle* (1976).

5. J.E. Kemeny and L.J. Snell, "Preference Ranking: An Axiomatic Approach," in *Mathematical Models in the Social Sciences*, Ginn Pub. Co., N.Y. (1962).

6. M.G. Kendall, *Rank Correlation Methods*, 3rd. edition, Hafner Pub. Co., N.Y. (1962).

7. E.R. Lawler and D.E. Wood, "Branch and Bound Methods: A Survey," *Operations Research*, Vol. 14 (1966).

8. E.L. Leese, "A Program to Determine the Median of a Set of Rankings," *DMS* (Directorate of Mathematics and Statistics) Staff Note # 4/74.

9. A. Norman, "A Program to Determine the Median of a Set of Rankings Using a Branch and Bound Technique," *DMS* Staff Note # 9/74.

10. S. Warner, "Advocate Scoring for Unbiased Information," forthcoming in *JASA*.

Unclassified
Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author) Center for Cybernetic Studies The University of Texas	2a. REPORT SECURITY CLASSIFICATION Unclassified 2b. GROUP

3. REPORT TITLE

A Committee Ranking Problem Using Ordinal Utilities.

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)

5. AUTHOR(S) (First name, middle initial, last name)

Ronald D. Armstrong
Wade D. Cook

12 34p.

6. REPORT DATE July 1977	7a. TOTAL NO. OF PAGES 30	7b. NO. OF REFS 10
8a. CONTRACT OR GRANT NO. N00014-75-C-0569, VNRC-A8966 b. PROJECT NO. NR047-021 c. d.	9a. ORIGINATOR'S REPORT NUMBER(S) Center for Cybernetic Studies Research Report, CCS-298 9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)	

10. DISTRIBUTION STATEMENT

This document has been approved for public release and sale; its distribution is unlimited.

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY Office of Naval Research (Code 434) Washington, D.C.

13. ABSTRACT

This paper investigates aspects of obtaining a concensus ranking from ordinal rankings supplied by committee members. A procedure for calculating the distance between rankings which are represented by vectors is given along with an algorithm for deriving a min-max concensus ranking. In addition, an approach to the committee ranking problem which makes allowance for weights, partial rankings and multiple attributes is discussed.

DD FORM 1 NOV 65 1473 (PAGE 1)
S/N 0101-807-6811

406197

Unclassified
Security Classification

A-31408

Unclassified
Security Classification

14 KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Committee Ranking						
Ordinal Utilities						
Branch and Bound						
Median Ranking						

DD FORM 1 NOV 65 1473 (BACK)
S/N 0102-014-6800

Unclassified
Security Classification

A-31409