Representing Groups with Imprecise Opinions

Hannu Nurmi*
Academy of Finland and
Department of Political Science
University of Turku
FI-20014 TURKU
Finland
e-mail:hnurmi@utu.fi

Janusz Kacprzyk
Systems Research Institute
Polish Academy of Sciences
Ul. Newelska 6
01-447 Warsaw
Poland
e-mail: kacprzyk@ibspan.waw.pl

Abstract

This paper deals with the composition of multi-member decision making bodies (committees) on the basis of fuzzy individual preference relations over candidates. The methods discussed are basically extensions of the principles of optimal representation as spelled out by Chamberlin and Courant. Since these principles are defined for non-fuzzy individual preference rankings, we focus primarily on the essential ambiguity in determining the degree of representation associated with each possible committee when the individual preferences are fuzzy. Once this ambiguity is solved, the composition of an optimal committee is relatively straightforward.

1 Introduction

The setting studied by social choice theory is typically one where we are given a set N of voters, set K of alternatives, a profile R of voter opinions over alternatives and our task is to design a method for electing a subset of K, i.e. a set of best alternatives or winner so that the voter opinions are reflected in the choice in a plausible (just, democratic, rational) way. Typically, it is assumed that the voter opinions can be represented as complete and transitive binary relations over K. The plausibility of the choice procedure is determined by the conditions one imposes on the method. The social choice theory is well known for the plethora of negative results stating the incompatibility of several intuitively desirable properties of choice methods. The best-known of these results is undoubtedly Arrow's (1961) theorem which states the incompatibility of universal domain, Pareto optimality, independence of irrelevant alternatives and non-dictatorship insofar as one is focusing on methods that result in complete and transitive relations over the alternatives. In other words, the result says that any rule that ends up with a ranking over alternatives must violate one or more of the just mentioned properties.

Arrow's result has given rise to a vast literature which we shall not go into here. For the purposes of the present paper it is important to point out, however, that Arrow's choice rule concept – a method that results in a (collective) ranking – was eventually replaced by another concept, viz. that of a social choice correspondence which associates with each preference profile and an alternative set a subset of alternatives. In other words, the standard literature settled for a somewhat less demanding output of the choice rule.

While the choice of a winning alternative or – should there be a tie between several alternatives – alternatives has received a lot of attention in the literature, the composition of multi-member bodies has been the focus of relatively few scholarly works. Yet, multi-member bodies exhibit new types of challenges for research that cannot be answered by simply iterating choices of a method that results in one winner. One of these challenges is the notion of representation.

2 Optimal representation

According to a widely held view a voting body is representative to the extent its composition is a "mirror image" of the composition of its constituency. This is, of course, a very crude way of describing the degree of representation. A somewhat more precise way is to require that the distribution of opinions in the representative body is the same as in the electorate. Chamberlin and Courant impose the following requirements on representation [1]. A committee member represents a voter to the extent that

1. the committee member "makes present" the

*This is a considerably abridged and modified version of [4] voter's opinions in the deliberations that take place within the committee,

2. the committee member is similarly responsive to various kinds of arguments presented in those deliberations as the voter, and
3. the committee member votes in the same way as the voter should the latter be present in the committee.

This is still too vague to provide a direct guideline for designing representative committees. To make the problem tractable we make resort to the standard assumptions of social choice theory mentioned above. More specifically, we assume that the voters are endowed with complete and transitive preference relations over candidates. In other words, rather than determining the degree of representation *a posteriori*, i.e. after the committee has been in office for some time, we assume that the voters know *a priori* which candidates best represent them.

We, thus, assume that the voters have complete and transitive binary preference relations over candidates and that we are to choose a maximally representative committee with k members on the basis of the preference information provided by the voters. Chamberlin and Courant suggest the following procedure. We first generate all k-member committees that can be formed out of the candidate set K. For each possible committee, we, then, compute the number of individuals whose most preferred candidate is present in the committee. Denote this number by n_1. It can obviously be any number between 0 and n, the total number of voters. Then count the number of voters whose first or second preference candidate is present in the committee and denote this by n_2. Continue in this manner until all ranks $1, \ldots, k$ have been considered. We denote the set of all k-member committees by C^k with elements $c_1, c_2, \ldots, c_s$. The value $C(c_i) = \sum_j^k n_j$, for each $i = 1, \ldots, s$ is the indicator of the representativeness of a committee: the higher the value, the better represented are the voters. Clearly, $k \times n$ is the maximum attainable value and is associated with a committee where each voter's first ranked candidate is present. Similarly, 0 is the minimum value of $C(c_i)$. This "worst possible" committee has the distinction that no voter ranks any committee member higher than $k + 1$th in his/her ranking.

It turns out that maximizing $C(c_i)$ over all possible k-member committees amounts to maximizing the sum of the Borda scores of committee members. The most representative one-member committee is the one consisting of the candidate with the maximum Borda score. A maximally representative k-member committee, on the other hand, is determined by a modified Borda count as follows. Let us define each voter's representative in a committee as the committee member getting the largest number of Borda points from that voter, i.e. the member ranked highest in the voter's ranking over candidates. Thus, each voter has a representative in each committee. Now, let $B(c_i)$ denote the sum over voters of the Borda points given to their representative in the committee c_i. The most representative committee is then defined as $c = arg\ max_i B(c_i)$, i.e. the committee where the sum of the Borda points given by each voter to his/her representative is maximal. This is, indeed, a modified Borda count since each voter gives only one score, viz. that of his/her representative.

3 Maximizing representation under fuzziness

Consider now the concept of representation in the context of fuzzy individual preference relations. Voter i's preference relation over candidates can be presented as:

$$\begin{matrix} - & r^i_{12} & \cdots & r^i_{1k} \\ r^i_{21} & - & \cdots & r^i_{2k} \\ \cdots & \cdots & \cdots & \cdots \\ r^i_{k1} & r^i_{k2} & \cdots & - \end{matrix}$$

Consider now voter i and a committee c_t consisting of k candidates as required. We are now primarily interested in finding the members of c_t that best represent i. Denote the set of these representatives by $B(i, c_t)$. Several plausible ways of finding the best representatives can be envisioned:

1. $B^i_{\text{sum}}(c_t) = \{j \in c_t | \sum_l r_{jl} \geq \sum_l r_{ql}, \forall q \in c_t\}$,
2. $B^i_{\text{min}}(c_t) = \{j \in c_t | \min_l r_{jl} \geq \min_l r_{ql}, \forall l \in K, \forall q \in c_t\}$,
3. $B^i_h(c_t) = \{j \in c_t | h(j) \geq h(q), \forall q \in c_t\}$ where $h(j) = p\ (\max_l r_{jl}) + (1-p)(\min_l r_{jl})$,
4. $B^i_{\text{cop}}(c_t) = \{j \in c_t | cop(j) \geq cop(q), \forall q \in c_t\}$ where $cop(j) = |\{l \in c_t | r_{jl} > r_{lj}, \forall l \in K\}|$

The first one determines the best representatives on the basis of the sums of the preference degrees obtained by candidates in all pairwise comparisons. This method is very much in the spirit of the Borda count. The second method looks at the minimum preference degree of each candidate when compared with all others and picks the candidate with the largest minimum. It is a variant of the min-max method in social choice theory. The third method is a version of Hurwicz's rule which maximizes the weighted sum of the smallest and largest preference degrees [2]. The fourth method is motivated by Copeland's rule in social choice theory. The Copeland winner is the candidate that defeats more candidates than any other candidate. In the setting of fuzzy preference relation $cop(j)$ is the number of candidates in c_s that are less preferred to j than j is preferred to them. In reciprocal preference matrices, $cop(j)$ is simply the number of entries larger than 0.5 on the j'th row.

Each of these methods singles out the best representatives of every voter in any given committee. Since each of the methods is based on a score, we can define a ranking of candidates in accordance with those scores. From the point of view of representation more important is, however, the ranking over committees ensuing from these methods. The most straightforward way to accomplish this is to define the score of committee c_t as follows:

$$S_t = \sum_{i \in N} \sum_{a \in c_t} \sum_{j \in K} r^i_{aj}.$$

Thus, the score of a committee is the sum of values given by voters to each of its members. The values, in turn, are the sums of preference degrees in all pairwise comparisons. This method is a variation of the Borda count. The most representative committee RC^B would then be:

$$RC^B = \{c_i \in C^k | S_i \geq S_j, \forall c_j \in C^k\}.$$

Although the Chamberlin-Courant approach is very close to the Borda count as well, the above method is not its most plausible fuzzy counterpart. Rather than summing the preference degrees over alternatives and voters, the Chamberlin-Courant approach sums the Borda scores of each voter's representative in any given committee. First we define

$$r^i_j = \sum_{q \in K} r^i_{jq}.$$

Then, for each committee c_t we define:

$$V_{it} = max_{j \in c_t} r^i_j.$$

This can be viewed as the value of the committee c_t to voter i as reflected by the value i assigns to his/her representative in c_t.

Now, the most representative committee in the sense of Chamberlin-Courant is:

$$RC^{CC}_{\text{sum}} = \{c_j \in C^k \mid \sum_i V_{ij} \geq \sum_i V_{iq},$$
$$\forall c_q \in C^k, i \in N, j \in K\}.$$

The RC^{CC}_{Sum} committee thus defined is based on the summation of preference degrees in individual preference matrices. In analogous manner one can define the most representative committee in the min-max sense. Let $\underline{r^i_j} = min_{q \in K} r_{jq}$. Now define, for each committee c_t and each voter i:

$$V'_{it} = max_{j \in c_t} \underline{r^i_j}.$$

Then the most representative committee in the min-max sense is:

$$RC'^{CC}_{\text{min}} = \{c_j \in C^k \mid \sum_i V'ij \geq \sum_i V'_{iq}, \forall c_q \in C^k\}.$$

The RC'^{CC}_{min} differs from the previous committee in using the min-max calculus to determine each voter's representative. In a way, RC'^{CC}_{min} mixes two kinds of maximands: the "utilitarian" and "Rawlsian". The former maximizes the average utility, while the latter maximizes the utility of the worst-off individual [5].

A purely Rawlsian committee can also be envisioned. This is obtained as follows:

$$RC^R = \{c_j \in C^k \mid min_i V'_{ij} \geq min_i V'_{iq}, \forall c_q \in C^k\}.$$

In similar vein, one can define Hurwicz and Copeland committees, RC^H and RC^{Co}, respectively. For a fixed value of $p^i \in [0,1]$, let $r^{iH}_j = p^i(\max_q r_{jq}) + (1 - p^i)(\min_q r_{jq})$ and $V^H_{it} = max_{j \in c_t} r^{iH}_j$. The set of most representative Hurwicz-type committees is, then:

$$RC^H = \{c_j \in C^k \mid \sum_i V^H_{ij} \geq V^H_{iq}, \forall c_q \in C^k\}.$$

Note that the value p^i is voter specific measure of his/her "optimism", i.e. the weight assigned to $\max_j r^i_{ij}$, i.e. the degree of preference assigned to each candidate in the comparison of its weakest competitor. Intuitively speaking the exclusive emphasis on strongest and weakest pairwise comparisons is somewhat questionable in voting contexts.

To define, the Copeland-type committee, let RC^{Co}, in turn, is based on the voters' value function $r^{iCo}_j = \mid \{q \in K \mid r_{jq} > r_{qj}\} \mid$ and the value function $V^{iCo}_{it} = max_{j \in c_t} r^{iCo}_j$. Now,

$$RC^{Co} = \{c_j \in C^k \mid \sum_i V_{ji}^{iCo} \geq \sum_i V_{qi}^{iCO},$$
$$\forall c_q \in C^k\}.$$

Of these four types of committees, the Rawlsian and Copeland types utilize the least amount of the voter preference information. The former looks at the minimal level preference of each candidate when compared with all others. The latter uses only the order information of preference degrees. Of course, if the aim is to economize on information usage, the very idea of resorting to fuzzy preference degrees loses much of its appeal.

4 Concluding remarks

The computational complexity issues notwithstanding the design of representation maximizing committees turns on the fundamental question: given a complete and transitive preference relation over candidates, how to determine the "winner"? This question has dominated much of the applied social choice literature, especially after Donald Saari presented a strong case for the Borda count which for a long time was considered inferior to Condorcet extension methods, e.g. Copeland's or Nanson's rule [6, 7, 3]. When the individual preferences are fuzzy we have several alternative ways of defining the winner, i.e. the best candidate in the voter's view. Some of these ways extend Condorcet's notion of winning into fuzzy environments, while others are more in accordance with Borda's views. Given the completeness of the individual preference relation, we can not only define the winner but also the ranking of the candidates. These rankings can then be used in defining the maximally representative committee.

References

[1] J. Chamberlin and P. Courant, Representative deliberations and representative decisions: Proportional representation and the Borda rule, The American Political Science Review 77 (1983) 718–733.

[2] J. Milnor, Games against nature, in: R. Thrall, C. Coombs and R. Davis (eds.) *Decision Processes*. Wiley, New York, 1954.

[3] H. Nurmi, Comparing Voting Systems, Dordrecht, D. Reidel 1987.

[4] H. Nurmi and J. Kacprzyk, Political representation: a perspective from fuzzy systems theory, New Mathematics and Natural Computation, forthcoming 2007.

[5] J. Rawls, A Theory of Justice, Oxford University Press, Oxford, 1971.

[6] D. Saari, Basic Geometry of Voting, Berlin-Heidelberg-New York, Springer Verlag, 1995.

[7] D. Saari, Which is better: The Condorcet or Borda winner?, Social Choice and Welfare 26 (2006) 107–129.