AN EXTENSION OF THE CONDORCET CRITERION

AND

KEMENY ORDERS

by

Michel Truchon

CRÉFA and Département d'économique
Pavillon J.A. de Sève
Université Laval
Québec, Canada, G1K 7P4
tél.: 418-656-5102
fax: 418-656-7798
email: mtru@ecn.ulaval.ca

cahier 9813

cahier 98-15
du Centre de Recherche en Économie et Finance Appliquées
CRÉFA

October 1998

I am most grateful to John Galbraith and Bryan Campbell for allowing me to work with a set of data that they had collected on 24 olympic competitions. I would also like to thank Philippe De Donder, Mohamed Sidi Drissi-Bakhkhat and Pascal St-Amour for their comments. However, I remain responsible for any errors or weaknesses. Finally, I acknowledge the financial support of the Social Sciences and Humanities Research Council of Canada.

Abstract

The usual Condorcet Criterion says that if an alternative is ranked ahead of all other alternatives by an absolute majority of voters, it should be declared the winner. The following partial extension of this criterion to other ranks is proposed: If an alternative is consistently ranked ahead of another alternative by an absolute majority of voters, it should be ahead in the final ranking. The term "consistently" refers to the absence of cycles in the majority relation involving these two alternatives. If there are cycles, this criterion gives partial orders that can be completed with the Kemeny rule. An algorithm to construct Kemeny orders is presented. It is based on a result saying that a complete Kemeny order over all alternatives can be obtained by splicing together Kemeny orders on the subsets of an admissible partition of the alternatives underlying the Extended Condorcet Criterion.

Key words: aggregation, Condorcet Criterion, Kemeny orders, algorithm.

Journal of Economic Literature classification: D710, D720.

Résumé

Le critère usuel de Condorcet exige que, si une alternative est classée avant toutes les autres par une majorité de votants, elle devrait être déclarée vainqueur. Une extension partielle de ce critère aux autres rangs est proposée: Si une alternative est classée avant une autre de manière cohérente par une majorité de votants, elle devrait l’être dans le classement final. La cohérence réfère à l’absence de cycle dans la relation majoritaire impliquant ces deux alternatives. En cas de cycles, ce critère donne des ordres partiels, qui peuvent être complétés avec la règle de Kemeny. Un algorithme pour la construction des ordres de Kemeny est présenté. Il s’appuie sur un résultat affirmant qu’un ordre de Kemeny peut être obtenu en juxtaposant des ordres de Kemeny sur les sous-ensembles d’une partition des alternatives sous-jacente au critère de Condorcet généralisé.

1. Introduction

The usual Condorcet Criterion found in the literature says that if an alternative or a candidate is ranked ahead of all other alternatives by an absolute majority of voters, it should be declared the winner. Such an alternative may fail to exist because of a cycle in the majority relation. When it exists, it is called the Condorcet winner.

Condorcet was preoccupied not only with the winner but also with the whole ranking of all alternatives. This may pose a problem, again because of the possibility of a cycle in the majority relation. Yet, a partial extension of the Condorcet Criterion to other ranks can be done as follows: If an alternative is consistently ranked ahead of another alternative by an absolute majority of voters, it should be ahead in the final ranking. The term "consistently" refers to the absence of cycles involving these two alternatives.

A first objective of this paper is to propose a formalization of this idea, called the Extended Condorcet Criterion (XCC). In essence, it says that if the set of alternatives can be partitioned in such a way that all members of a subset of this partition defeat all alternatives belonging to subsets with a higher index, then the former should obtain a better rank than the latter. A partition satisfying the above property is said admissible for (XCC). If there is a cycle in the weak majority relation and, in particular, if there is a subset of alternatives that all tie, then this cycle or subset should belong to a same subset of any admissible partition.

This work was motivated by the comparison of aggregation procedures in figure skating. Ties could not be neglected despite the complication that they usually bring. Indeed, two skaters may tie even if the number of judges is odd because a judge may give the same rank to more than one skater. Ties may create cycles in the weak majority relation and increase the length of existing ones. Many ties and many cycles of the weak majority relation have actually been found in the data of Olympic Games analyzed in Truchon (1998).

In the presence of cycles, (XCC) gives a partial order of the alternatives. The question is then: how to complete this order to obtain a complete ranking? To answer this question while remaining in the spirit of (XCC), let us turn to another question addressed by Condorcet (1785): which final ranking is most likely to be the true ranking of the alternatives? Neglecting the possibility of ties, Condorcet showed that if the pairwise ranking of the alternatives under the majority rule does not involve any cycle, then it yields a complete order that has maximum likelihood of being the true order, under the assumption that every voter chooses the best of two alternatives with a probability larger than one half and that this judgment is independent between pairs and voters.

Condorcet gave indications on how to break cycles that might occur. However, his prescription is not completely clear. Young (1988) shows that a correct application of Condorcet's maximum likelihood approach leads to an order that has the maximum pairwise support from the voters. Such an order is often called a Kemeny order because it involves the minimum number of pairwise inversions with the individual rankings. Kemeny (1959) proposes this number as a distance between an order and a profile of individual rankings. Such a ranking is also a median ranking for those composing a profile. In this sense, it represents a best compromise between the possibly conflicting views of the voters.

It is shown in this paper that a Kemeny order satisfies (XCC). This is a good justification for (XCC) and its name. Conversely, the Kemeny-Young approach is a natural complement to (XCC). Indeed, a complete Kemeny order over all alternatives can be obtained by splicing together Kemeny orders on the subsets of an admissible partition for (XCC). This is a very useful result since the Kemeny-Young approach may become prohibitive when the number of alternatives is large. Note that in the finest possible admissible partition for (XCC), many of the subsets may be singletons, which eases the task considerably.

A second objective of this paper is to present an algorithm, based on the above result, to construct complete Kemeny orders. The key to this algorithm is the construction of an admissible partition. Actually, if some alternatives tie within a subset of such a partition, then this subset can possibly be further partitioned, thus enhancing the performance of the algorithm.

The paper is organized as follows. Section 2 introduces the notation including the formal description of (XCC). For the sake of completeness, Section 3 first presents the Kemeny-Young approach with sufficient details. It then develops the results on which the algorithm is based. This algorithm is presented in Section 4. In Section 5, it is explained how we can deal with the multiplicity of Kemeny orders. A brief conclusion makes up Section 6.

2. Notation

Let X be the set of alternatives, with cardinality $|X| = m$, and N be the set of voters, with $|N| = n$, an odd number. Each voter j is assumed to have a *weak order* or *ranking* r^j of the alternatives in X. The element r_s^j of this vector is the rank of alternative s. A ranking with no tie for a rank is an *order* on X. An order can be represented alternatively as a sequence $s_1 s_2 ...$, where s_1 and s_2 are respectively the alternatives with ranks 1 and 2, etc.

Let $\Re$ be the set of all possible rank vectors r. A *profile of rankings* is an $m \times n$ matrix $R = (r^1, ..., r^n) \in \Re^n$. A *ranking rule* is a mapping $FR : \Re^n \to \Re$. $FR(R)$ is the final ranking resulting from profile R. In the language of the theory of social choice, FR is also a *social welfare function*.

Next, let us define $\nu_{st}(R) = |\{j \in N : r_s^j < r_t^j\}|$ and the complete binary relation M on X by $s\,M\,t \Leftrightarrow \nu_{st} \geq \nu_{ts}$. We write $\overline{M}$ for the asymmetric component of M, i.e. the relation defined by $s\,\overline{M}\,t \Leftrightarrow \nu_{st} > \nu_{ts}$ and T for the symmetric component of M, i.e. the relation defined by $s\,T\,t \Leftrightarrow \nu_{st} = \nu_{ts}$. They are respectively the *majority relation*, the *strict majority relation* and the *ex aequo relation* on X. We can read $s\,\overline{M}\,t$ as s defeats t and $s\,T\,t$ as s ties with t.

A cycle of M is a subset $\{x_1, ..., x_k\} \subset X$ such that $x_i\,M\,x_{i+1}$, $i = 1, ..., k-1$, and $x_k\,M\,x_1$. One defines similarly cycles of $\overline{M}$ and T. Equivalently, a cycle of T is a subset $S \subset X$ such that $s\,T\,t \;\forall\, s,t \in S$. Cycles of $\overline{M}$ and T are obviously cycles of M.

The usual Condorcet Criterion reads as follows:

Condorcet Criterion (CC)

$$\forall\, s \in X, \forall\, t \in X, t \neq s : s\,\overline{M}\,t \Rightarrow FR_s(R) = 1 \text{ and } FR_t(R) > 1$$

An s satisfying (CC) may not exist because of a cycle in the majority relation M. When there exists an $s \in X$ satisfying (CC), this s is called the *Condorcet winner*. Note that there may exist a Condorcet winner even if there are cycles over some subsets of alternatives.

To extend (CC) to other ranks, let $\wp_0(X)$ be the class of partitions $\boldsymbol{X} = \{X_1, ..., X_p\}$ of X, satisfying:

$$\forall\ X_\alpha, X_\beta \in \boldsymbol{X} \text{ with } \alpha < \beta,\ \forall\ s \in X_\alpha,\ \forall\ t \in X_\beta : s\overline{M}t$$

Notice that, if there is a cycle of M over some subset of alternatives, then these alternatives must belong to a same subset X_α of any partition in $\wp_0(X)$. In particular, this must be the case for two alternatives s and t such that $s\,T\,t$. In the finest partition of this class, the sets X_α are cycles of maximal length of M or singletons. X_1 is also called the *top cycle* of M or the *Condorcet set*, a solution concept introduced by Good (1971) and Schwartz (1972) for the strict majority relation. X_2 is the top cycle on $X \setminus X_1$, etc.

Extended Condorcet Criterion (XCC)

For any partition $\boldsymbol{X} \in \wp_0(X)$, the following must hold:

$$\forall\ X_\alpha, X_\beta \in \boldsymbol{X} \text{ with } \alpha < \beta,\ \forall\ s \in X_\alpha,\ \forall\ t \in X_\beta : FR_s(R) < FR_t(R)$$

If $M = \overline{M}$ and if $\overline{M}$ contains no cycles, then all X_α of the finest partition of $\wp_0(X)$ are singletons and a final ranking $FR(R)$ satisfying (XCC) is a complete order. When M contains a cycle, (XCC) does not say how to rank alternatives within an X_α of the finest partition in $\wp_0(X)$. In particular, it does not imply that $FR_s = FR_t$ if $s\,T\,t$. In other words, (XCC) yields only a partial order in these circumstances. A complete final ranking, consistent with (XCC), can be obtained with the maximum likelihood or Kemeny approach.

3. Kemeny orders

Assuming that the rankings r^j in some profile R are independent evaluations of the true ranking of the alternatives, an interesting question is: which final ranking $FR(R)$ is most likely to be the true ranking of the alternatives. This is precisely the question addressed by Condorcet (1785). His objective was to justify the majority principle. Condorcet's approach is one of the first applications of statistical hypothesis testing and maximum likelihood estimation. He assumes that every voter chooses the best of two alternatives with a probability p satisfying $1/2 < p \leq 1$, and that this judgment is independent between pairs and voters. If the binary relation M is an order on X, then it is the solution to his problem, i.e. the most probable order on X. If M contains a cycle, Condorcet's prescription is to eliminate some of the propositions ($s\,M\,t$ is a proposition), starting with the one with the weakest majority and so on until the cycle disappears. This works fine if $m = 3$ but may give ambiguous results or a partial order for $m \geq 4$.

Young (1988) develops a correct application of Condorcet's maximum likelihood approach. In the case $m = 3$, it goes as follows. If the true order on the set $\{a, b, c\}$ is *abc*, then, neglecting a multiplicative constant and the argument R in $\nu_{st}(R)$, the conditional probability of observing a profile of votes R is given by:

$$p^{\nu_{ab}}(1-p)^{\nu_{ba}}p^{\nu_{ac}}(1-p)^{\nu_{ca}}p^{\nu_{bc}}(1-p)^{\nu_{cb}} = p^{\nu_{ab}+\nu_{ac}+\nu_{bc}}(1-p)^{\nu_{ba}+\nu_{ca}+\nu_{cb}}$$

The probability of observing the same profile, conditional on the true order being *acb*, is given by:

$$p^{\nu_{ac}}(1-p)^{\nu_{ca}}p^{\nu_{ab}}(1-p)^{\nu_{ba}}p^{\nu_{cb}}(1-p)^{\nu_{bc}} = p^{\nu_{ac}+\nu_{ab}+\nu_{cb}}(1-p)^{\nu_{ca}+\nu_{ba}+\nu_{bc}}$$

Hence if $p > ½$, *abc* is more probable than *acb* as on order if and only if $\nu_{ab}+\nu_{ac}+\nu_{bc} > \nu_{ac}+\nu_{ab}+\nu_{cb}$. More generally, let:

$$K(r,\ R) = \sum_{s\in X}\sum_{\substack{t\in X\\ r_s<r_t}} \nu_{st}(R).$$

An order r^* on X is a solution of Condorcet's problem or a *maximum likelihood order* if it is a solution of $\max_{r\in\Re} K(r, R)$. The value of $K(r, R)$ may be seen as the total number of pairwise supports for r in profile R, i.e. the total number of voters who rank pairs of alternatives as in r. A maximum likelihood order is thus one that has the maximum total support from the voters as expressed in R.

This problem may be given a different expression using a notion of distance for orders proposed by Kemeny (1959). It is presented here in a slightly modified form to accommodate the fact that weak orders may be found in profiles of rankings. Given an order r, a weak order $r^j \in \Re$, and two alternatives $s, t \in X$, define:

$$\delta_{st}(r,\ r^j) = \begin{cases} 1 & \text{if } r_s < r_t \text{ and } r_t^j \le r_s^j \\ 0 & \text{otherwise} \end{cases}$$

$$\text{and } \Delta(r,\ r^j) = \textstyle\sum_{s\in X}\sum_{t\in X}\delta_{st}(r,\ r^j)$$

The value of $\delta_{st}(r, r^j)$ indicates whether there is a disagreement in the relative ranking of s and t between r and r^j. $\Delta(r, r^j)$ is the total number of such disagreements between r and r^j. The function Δ is a distance on the set $\Re$, with the restriction that its first argument must be an order. One can then define a "distance" d between an order r and a profile R by: $d(r, R) = \sum_{j=1}^{n}\Delta(r, r^j)$. In plain words, $d(r, R)$ is the total number of disagreements between an order r and all the rankings in profile R.

A *Kemeny order* for a profile R is an order r^K solving $\min_{r \in \Re} d(r, R)$, i.e. an order that is closest to the given profile according to the "distance" d or an order that has the minimum number of disagreements with the profile. A Kemeny order is also a median order for the rankings in the profile. As such, it represents the best compromise between the different opinions of the voters.

The following lemmas and corollaries give some of the properties of this order. The first one asserts that a Kemeny order is a maximum likelihood order. From this lemma, we can reassert that a maximum likelihood order is one that has the maximum number of agreements with the profile.

Lemma 1: A Kemeny order for a profile R is an order solving $\max_{r \in \Re} K(r, R)$.

Proof. $$d(r,\ R) = \sum_{s \in X} \sum_{t \in X} \sum_{j=1}^{n} \delta_{st}(r,\ r^j) = \sum_{s \in X} \sum_{\substack{t \in X \\ r_s < r_t}} (n - \nu_{st}(R)) = \frac{m(m-1)n}{2} - \sum_{s \in X} \sum_{\substack{t \in X \\ r_s < r_t}} \nu_{st}(R),$$ hence the result. □

Lemma 2: Suppose $r^K = (1, 2, ..., m)$ is a Kemeny order for a given profile R. Then $\nu_{s,\,s+1}(R) \geq \nu_{s+1,\,s}(R)$, $s = 1, ..., m-1$, or, equivalently, $1\,\mathcal{M}\,2\,\mathcal{M} \ldots \mathcal{M}\,m$.

Proof. For any $s = 1, ..., m-1$, consider the order $r = (1, 2, ..., s-1, s+1, s, s+2, ..., m)$. By the proof of Lemma 1, $d(r^K, R) - d(r, R) = \nu_{s,\,s+1} - \nu_{s+1,\,s}$, which cannot be negative if r is a Kemeny order. □

Corollary 3: Given a Kemeny order for a given profile R, if there exists a Condorcet winner under this profile, it must be the alternative ranked first in the Kemeny order.

Proof. Let $r^K = (1, 2, ..., m)$ be a Kemeny order. From Lemma 2, $\nu_{s-1,\,s}(R) \geq \nu_{s,\,s-1}(R)$, $s = 2, ..., m$. Thus none of the alternatives $s = 2, ..., m$ may be a Condorcet winner, leaving 1 as the Condorcet winner. □

Corollary 4: A Kemeny order r^K satisfies (XCC).

Proof. Suppose r^K violates (XCC), i.e. there exists a partition $\boldsymbol{X} \in \wp_0(X)$, $X_\alpha, X_\beta \in \boldsymbol{X}$ with $\alpha < \beta$, $s \in X_\alpha$ and $t \in X_\beta$ such that $r_t^K \leq r_s^K$. By definition of $\wp_0(X)$, we must have $s\,\overline{\mathcal{M}}\,t$. Thus, by Lemma 2, there must exist other alternatives, say $a, ..., k$, between t and s in the Kemeny order. Using Lemma 2 again, we must have $t\,\mathcal{M}\,a\,\mathcal{M} \ldots \mathcal{M}\,k\,\mathcal{M}\,s$. Since we also have $s\,\overline{\mathcal{M}}\,t$, there is a cycle over the set $\{t, a, ..., b, s\}$. Using the definition of $\wp_0(X)$ again, $\{t,a, ..., b,s\}$ should belong to the same set of the partition $\boldsymbol{X}$. We thus have a contradiction since, at the outset, s and t belonged to different X_α and X_β. □

The next result provides an easy way to find complete Kemeny orders. In essence, it says that the latter can be constructed by the concatenation of Kemeny orders on each of the sets of a partition $\boldsymbol{X} \in \wp_0(X)$. Recall that an order can take the form x or r, where x_i is the alternative whose rank is i while r_s is the rank of alternative s.

Theorem 5: Take any partition $\boldsymbol{X} = \{X_1, ..., X_p\} \in \wp_0(X)$ and a vector $x^* = (x_1, ..., x_p)$ where x_α is a Kemeny order on X_α under profile R restricted to X_α, $\alpha = 1, ..., p$. Then x^* or equivalently the corresponding r^* is a Kemeny order on X.

Proof. Suppose that there exists an order r on X such that $K(r, R) > K(r^*, R)$. Then r cannot be different from r^* in respect only to alternatives who belong to the same X_α since this would violate the assumption that x_α is a Kemeny order on X_α. Thus there exist $X_\alpha, X_\beta \in \boldsymbol{X}$ with $\alpha < \beta$, $s \in X_\alpha$ and $t \in X_\beta$ such that $r_t \leq r_s$ instead of $r_s < r_t$ as in r^*. By Corollary 4, r cannot be a Kemeny order since it violates (XCC). If there were orders r such that $K(r, R) > K(r^*, R)$, there would be a Kemeny order among them. Thus there is no such order and r^* is a Kemeny order. □

If $M \neq \overline{M}$, we can go one step further in partitioning X. Given a partition $\boldsymbol{X} \in \wp_0(X)$, if there exists an $X_\alpha \in \boldsymbol{X}$ and a subset $S \subset X_\alpha$ such that $sMt\ \forall\ s \in S, \forall\ t \in X_\alpha$, then X_α can be further partitioned into $\{S, X_\alpha \backslash S\}$ to give another (finer) partition of X. Note that if $|S| > 1$, then $sTt\ \forall\ s,t \in S$. This refinement can possibly be repeated on $X_\alpha \backslash S$ and so on. Call this *enlarged class of partitions* $\wp(X)$.

Similarly, given a partition $\boldsymbol{X} \in \wp(X)$, if there exists an $X_\alpha \in \boldsymbol{X}$ and a subset $S \subset X_\alpha$ such that $tMs\ \ \forall\ s \in S, \forall\ t \in X_\alpha$, then X_α can be further partitioned into $\{X_\alpha \backslash S, S\}$ to give another (finer) partition of X and so on. Let these finer partitions also belong to $\wp(X)$. The class $\wp(X)$ will be actually larger than $\wp_0(X)$ only if $M \neq \overline{M}$. The cycles of $\overline{M}$ and T still belong to a same subset of any partition of $\wp(X)$. However, some cycles of M may have been broken in the refining process that leads from $\wp_0(X)$ to $\wp(X)$. The following theorem justifies the above enlargement of the class of partitions.

Theorem 6: Given a partition $\boldsymbol{X} = \{X_1, ..., X_p\} \in \wp(X)$ and some $X_\alpha \in \boldsymbol{X}$, if there exists a subset $S \subset X_\alpha$ such that $sMt\ \forall\ s \in S, \forall\ t \in X_\alpha$, then there exists a Kemeny order on X_α in which the elements of S occupy the first $|S|$ ranks in any order we wish. Similarly, if there exists a subset $S \subset X_\alpha$ such that $tMs\ \ \forall\ s \in S, \forall\ t \in X_\alpha$, then there exists a Kemeny order on X_α in which the elements of S occupy the last $|S|$ ranks in any order we wish. Thus, in either case, a Kemeny order on X_α can be obtained by splicing any order on S with a Kemeny order on $X_\alpha \backslash S$.

Proof. Consider the case where some X_α can be partitioned into $\{S, X_\alpha\backslash S\}$ and let r_α be any Kemeny order on X_α. Take any $s \in S$ and suppose that it does not occupy the first rank in r_α. Let t be the alternative just before s in r_α. Combining the assumption on S and Lemma 2, we must have $\nu_{st} = \nu_{ts}$. Hence s can be moved up one rank without decreasing $K(r_\alpha, R_\alpha)$. This gives us another Kemeny order. This argument may be repeated until s reaches the first position. If we apply this argument to all elements in S in the reverse order in which we want them, we will end up with a Kemeny order in which the elements of S will occupy the first $|S|$ ranks in the chosen order. The other case (the element of S in the last positions) is handled in a similar way. The last affirmation follows at once. □

Combining Theorems 5 and 6 yields the following corollary.

Corollary 7: Take any partition $\boldsymbol{X} = \{X_1, ..., X_p\} \in \wp(X)$ and a vector $x^* = (x_1, ..., x_p)$ where x_α is a Kemeny order on X_α under profile R restricted to X_α, $\alpha = 1,..., p$. Then x^* or equivalently the corresponding r^* is a Kemeny order on X.

4. An algorithm for the construction of Kemeny orders.

First, a partition $\boldsymbol{X} = \{X_1, ..., X_p\} \in \wp(X)$ must be constructed. This can be done with the following two stage procedure. In the first stage, the subsets are constructed from the beginning, i.e. in the order X_1, X_2, etc. In the second, they are constructed from the end, i.e. in the order X_p, X_{p-1}, etc. The algorithm attempts to construct a partition as fine as possible. However, there may remain a large residual subset when the procedure stops.

There is a non-negative and integer parameter γ in this procedure to be set by the user. It controls the fineness of the partition, including the size of the residual subset. The choice of a value for γ is a matter of compromise. With $\gamma = 0$, the partition belongs to the subclass $\wp_0(X) \subset \wp(X)$. The greater the value of γ the finer the partition. On the other hand, some Kemeny orders may escape us with too high a value of γ. This will be illustrated after the presentation of the algorithm.

First stage: Suppose $X_1, X_2, ..., X_{i-1}$ have been constructed.

Step i : Let $X^i = X \backslash (X_1 \cup X_2 \cup ... \cup X_{i-1})$.

If $X^i = \varnothing$, stop the procedure. $\{X_1, X_2, ..., X_{i-1}\}$ is the desired partition.

If $X^i \neq \varnothing$, let $S = \{s \in X^i : s\,M\,t \ \forall\ t \in X^i\}$;

If $0 < |S| \leq \gamma$ or if $[|S| > \gamma$ and $s\,\overline{M}\,t \ \forall\ s \in S, \forall\ t \in X^i \backslash S]$,

set $X_i = S$ and go to step $i + 1$;

Otherwise, go to the second stage.

Second stage: Suppose $X_1, X_2, ..., X_i$ have been constructed in the first stage and $X_p, X_{p-1}, ...X_{k+1}$ in this second stage.

Step k : Let $X^k = X \backslash (X_1 \cup X_2 \cup ... \cup X_i \cup X_p \cup X_{p-1} \cup ... \cup X_{k+1})$ and $S = \{s \in X^k : t\,M\,s \ \forall\ t \in X^k\}$

If $0 < |S| \leq \gamma$ or if $[|S| > \gamma$ and $t\,\overline{M}\,s \ \forall\ s \in S, \forall\ t \in X^k \backslash S]$,

set $X_k = S$ and go to step $k - 1$;

Otherwise, set $X_k = X^k$ and stop the procedure.

$\{X_1, X_2, ..., X_i, X_k, ..., X_{p-1}, X_p\}$ is the desired partition.

Except for the residual X_k, the subsets of the partition constructed by the algorithm are either singletons or cycles of T. If there are no cycles of M other than those of T, then only the first stage of the algorithm is used and there is no residual subset. Moreover, with $\gamma = 0$, we get the finest possible partition of $\wp_0(X)$ but not necessarily of $\wp(X)$.

Still with $\gamma = 0$, if there is a unique cycle in M other than those of T, then the algorithm gives again the finest possible partition of $\wp_0(X)$ and the residual subset X_k is made up of this cycle. If there is more than one cycle in M other than those of T, then the residual subset contains all these cycles and all alternatives that are between those of the different cycles according to the majority relation.

Once a partition is obtained, a Kemeny order on each subset must be found and these orders are spliced together to yield a complete Kemeny order on X. For singletons, this is trivial. For cycles of T, any order will do. For this reason, it will be argued below that the members of these subsets could be declared ex aequo. Finally, the Kemeny orders on the residual subset X_k can be found by simple enumeration. When performing the latter, most orders are quickly eliminated because they do not satisfy Lemma 2. Yet, the size of the residual subset may pose some difficulty, hence the need to keep it as small as possible.

If the residual subset is made up of cycles of $\bar{M}$ (no ties), the algorithm presented here is unable to reduce its size whatever the value of γ, even if there are many disjoint cycles (with other alternatives between them) in this subset. However, raising the value of γ may help in diminishing the size of the residual subset if some of the cycles that it contains are cycles of M but not of $\bar{M}$, i.e. contain ties. Note that allowing for ties may increase the frequency of cycles as well as their length. As a compensation, we obtain this possibility of reducing the size of the residual subset even if it is made up of a single cycle in M. This is illustrated with real examples taken from the data of Olympic Games analyzed in Truchon (1998).

The first example comes from the men free program of the 1992 olympic games. The residual subset X_k, with γ = 0, contains 11 skaters, namely skaters ranked 4 to 14 according to their raw marks. The reason is that there is a cycle of M on the subset {4, 5, 6, 7, 8} and another one on the subset {12, 13, 14}, which the algorithm does not identify. However, with γ = 1 the residual subset is split into {{4}, {5, 6, 7, 8}, {9}, {10}, {11}, {12}, {13}, {14}}. This is possible, thanks to Theorem 6, since $4Mt \ \ \forall\ t \in \{5, ..., 14\}$ and $tM14 \ \ \forall\ t \in \{4, ..., 13\}$. Once 14 is removed from the original residual subset, 13, 12, ..., 9 can be removed in turn to form singletons in a new partition. The new residual subset {5, 6, 7, 8} cannot be broken further since there is a cycle of $\bar{M}$ on this subset.

Another example comes from the women short program of the 1988 olympic games. In the results of this competition, there is a cycle of M over 9 skaters, namely skaters ranked 12 to 20 according to their raw marks. With γ = 0, these nine skaters make up the residual subset X_k. Applying the Kemeny rule over a set of this size is already costly. With γ = 1, the residual subset is split into {{12}, {13}, {14}, {15, ..., 20}}. This is possible because $12Mt$ for all other t in this subset. Once 12 is removed from this subset, 13 and 14 can be removed in turn. The new residual subset {15, ..., 20} cannot be broken further because there is a cycle of $\bar{M}$ on this subset.

Why not set γ as high as possible, say $|X|$? The men short program of the 1976 olympic games provides a good illustration. The data yield $3T4T6$, $5T7T8$, $3\overline{M}t\ \forall\ t \neq 3, 4$; $4\overline{M}t\ \forall\ t \neq 3, 4, 6$; $5\overline{M}6$, $5\overline{M}8$, $7\overline{M}6$, and $8\overline{M}6$. Hence, there is a cycle of M on the subset {3, 4, 5, 6, 7, 8} namely $4T3\overline{M}5T7T8T\overline{M}6$ and $6T4$. With $\gamma = 1$, skater 6 is removed from this subset. With $\gamma > 1$, this subset is broken further into {{3, 4}, {5, 7}, {8}, {6}}. If all that is wanted is a Kemeny order over {3, 4, 5, 6, 7, 8}, then one should indeed set $\gamma = |X|$ to ease up the search of such an order. However, Kemeny orders are not necessarily unique and we may be interested in knowing all of them. In the case at hand, consider the subset {5, 7, 8}. There are 3 Kemeny orders on this subset namely: 5 7 8, 7 5 8, and 5 8 7. With $\gamma > 1$, this last order would never show up. Thus, if we are interested in identifying all Kemeny orders, we should limit ourselves to $\gamma = 0$ or 1. This is the choice that has been done for the computations presented in Truchon (1998). However, there could be circumstances calling for a higher value of γ.

Notice that, even with $\gamma = 1$, some Kemeny orders could formally escape us but the latter are easy to catch up. Consider the set $\{a, b, c, d\}$ and suppose that it can be split into $\{\{a\}, \{b, c, d\}\}$ because aTb, $a\overline{M}c$ and $a\overline{M}d$. If *bcd* turns out to be a Kemeny order on $\{b, c, d\}$, then clearly *bacd* and *abcd* are both Kemeny orders on $\{a, b, c, d\}$, which is easy to identify. This case of multiple Kemeny orders is possible only if there are at least four alternatives in the subset to be partitioned. Indeed, suppose that the 3-element set $\{a, b, c\}$ is split into $\{\{a\}, \{b, c\}\}$ with $\gamma = 1$ because we have say aTb and $a\overline{M}c$. Then we should also have $c\overline{M}b$. With $b\overline{M}c$, the splitting would rather be $\{\{a, b\}, \{c\}\}$ and with bTc, no splitting would occur. But, with aTb, $a\overline{M}c$, and $c\overline{M}b$, there is a unique Kemeny order: *acb*.

From the above examples, we can appreciate the compromise to be made in setting the value of γ. A good strategy is to start with a low value of γ, say $\gamma = 0$ or 1, and then increase this value if the residual subset is too large. Increasing this value may be necessary since finding Kemeny order by enumeration becomes prohibitive as the number of alternatives increases. A procedure written by the author in Mathematica gives the Kemeny orders on a set of 7 alternatives in less than 3 seconds on a Pentium 200. Computing time goes to 26 seconds with eight alternatives and to 13 minutes with nine alternatives. This dramatic increase in time is partly due to a lack of memory. With more alternatives, finding a Kemeny order directly is almost out of reach. Recall that the number of possible orders on a set of m objects is $m!$.

A more efficient program could be written or a more powerful computer be used but this would only push the problem to a higher number of alternatives. Another avenue would be to use branch and bound techniques as in Barthelemy, Guénoche, and Hudry (1989) but these techniques also have their limits. Hence, whatever the technique, program or computer that are used, working with partitions as fine as possible and splicing the Kemeny orders on the subsets of this partition is almost inescapable.

We could also do better if we could distinguish all cycles rather than capture them in a single residual subset. However, this is not an easy task. The algorithm proposed here is far more simple and probably more efficient. In the men free program of 1992 reported above, it does a lot better than simply identifying the two cycles. With $\gamma = 1$, it breaks one of them into singletons and reduces the other one to a shorter cycle of $\overline{M}$. In the women short program of 1988, there is only one long cycle of M, which is reduced to a much shorter cycle of $\overline{M}$ by the algorithm.

Actually, this algorithm has proved very efficient, with $\gamma = 1$, in computing all Kemeny orders in the 24 olympic games, yielding all the results in matters of seconds. It should be sufficient for most practical applications involving the construction of Kemeny orders.

5. Dealing with multiple Kemeny orders.

A Kemeny order is not necessarily unique. The following rule can be applied to handle the occurrence of multiple Kemeny orders. Given a set $\{r^1, ..., r^k\}$ of Kemeny orders, consider the weak order r^m defined by:

$$\forall\ s,t \in\ X : r_s^m \leq r_t^m \Leftrightarrow \textstyle\sum_{q=1}^{k} r_s^q \leq \sum_{q=1}^{k} r_t^q$$

This weak order is a ranking according to the mean ranks of alternatives over all Kemeny orders. It will be called the *mean Kemeny ranking* if it weakly agrees with at least one order in $\{r^1, ..., r^k\}$, i.e. if there exists an order $r^q \in\ \{r^1, ..., r^k\}$ such that:

$$\forall\ s,t \in\ X : r_s^q < r_t^q \Rightarrow r_s^m \leq r_t^m$$

If $r_s^m < r_t^m$, this means that there are more Kemeny orders in which s is ranked ahead of t than Kemeny orders in which s is placed after t. Thus, if a Kemeny order is chosen at random, the probability that s be ranked ahead of t is higher than the probability that it be ranked after t. In r^m, alternatives are thus ranked according to these probabilities. In particular, two alternatives obtain the same rank if $r_s^m = r_t^m$. Thus, choosing r^m over other Kemeny orders makes sense if r^m weakly agrees with one Kemeny order.

However, it would be inconsistent with the Kemeny-Young approach to choose r^m if it is not a Kemeny order, since it is then less probable than any Kemeny order. In this case, a Kemeny order could be chosen at random or according to some other criterion.

With this approach, we look for Kemeny orders but we may end up with a weak order as a final choice. An alternative approach would consist in working with the set of weak orders instead of orders at the outset but this would be costly. For example, there are 75 weak orders on a set of four alternatives compared to 24 orders. The above approach is thus more practical. In Truchon (1998), the mean Kemeny ranking was actually chosen when it existed. Otherwise, a Kemeny order as close as possible to the official olympic ranking was retained. The only instances where the mean Kemeny ranking failed to exist were ones with only two Kemeny orders.

One case in which r^m gives the same rank to two alternatives s and t is when $\nu_{st} = \nu_{ts}$ and when in addition s and t are adjacent in any Kemeny order. Indeed, in this case, for any Kemeny order in which s is ahead of t, there is another one in which the only difference is that the positions of s and t are interchanged. In particular, all alternatives of a cycle of T in a partition of $\wp(X)$ obtain the same rank under r^m. Consistent with this remark, all elements of each subset of the partition constructed by the algorithm, other than the residual subset X_k, should be declared ex aequo.

There are other instances in which some alternatives could be declared ex aequo. Truchon (1998) reports an example where three Kemeny orders are obtained on the set {A, B, C, D, E, F} namely: EABFCD, EABDFC, and EABCDF. This means that CDF, DFC and FCD have the same likelihood. Not surprisingly, there were a cycle C $\bar{M}$ D $\bar{M}$ F $\bar{M}$ C over the subset {C, D, F}. There is good ground here to declare these three alternatives ex aequo since they have the same likelihood of being in any of the last three positions. This is what happens under r^m.

6. Conclusion

An extension of the usual Condorcet Criterion to other ranks has been proposed. This Extended Condorcet Criterion gives partial orders when there are cycles of the weak majority relation including ties between alternatives. These partial orders may be completed by reverting to the Kemeny rule. An algorithm to construct Kemeny orders has also been presented. The latter has proved to be very efficient on the data of 24 olympic competitions reported in Truchon (1998). There were as many as 24 skaters in these competitions. A total of 15 cycles of the weak majority relation, some involving 9 skaters, have been found in these data. It was thus important to have an efficient procedure to find the Kemeny orders.

References

Barthelemy, J.P, A. Guénoche, and O. Hudry, "Median Linear Orders: Heuristics and a Branch and Bound Algorithm," *European Journal of Operational Research*, 42 (1989), 313-325.

Condorcet, Marquis de, "Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix," Paris: Imprimerie royale, 1785. Also reproduced in Condorcet, *Sur les élections et autres textes*, edited by O. de Bernon, Fayard, 1986.

Good, I.J., "A Note on Condorcet Sets," *Public Choice*, 10 (1971), 97-101.

Kemeny, J. "Mathematics without Numbers," *Daedalus*, 88 (1959), 571-591.

Schwartz, T., "Rationality and the Myth of the Maximum," *Noûs*, 6 (1972), 97-117.

Truchon M., "Figure Skating and the Theory of Social Choice," Cahier de recherche 9814, Département d'économique, Université Laval, October 1998.

Young, H.P., "Condorcet's Theory of Voting," *American Political Science Review*, 82 (1988), 1231-1244.