Popular Ranking

Anke van Zuylen, [a] Frans Schalekamp, [b] David P. Williamson [c]

[a] *Max-Planck-Institut für Informatik, Saarbrücken, Germany,*
anke@mpi-inf.mpg.de

[b] *Independent*

[c] *School of ORIE, Cornell University, Ithaca, NY, USA.*

Key words: Rank aggregation, Kemeny Rank Aggregation, Popular Ranking

1 Introduction

How do you aggregate the preferences of multiple agents in a fair manner? This is a question that has occupied researchers for several centuries. Suppose we have k voters who each give their preferences on n candidates. How should the candidates be ranked to best represent the input? Marquis de Condorcet [5] showed that there may not exist a "winner": a candidate who beats all other candidates in a pairwise majority vote. Borda [4] and Condorcet [5] (and many others after them) proposed different ways of aggregating the preferences of the voters, and argued over which method is the right one. Only in the middle of the 20th century, Arrow [2] showed that there is no right method: there exists no aggregation method that simulataneously satisfies three natural criteria (non-dictatorship, independence of irrelevant alternatives and Pareto efficiency).

This negative result notwithstanding, we still want to find aggregate rankings based on voters' inputs. In this paper, we consider the case when, rather than selecting a winner, we would like to find a permutation of the candidates that represents the voters' inputs. Each voter's input is assumed to be a permutation of the candidates, where a candidate is ranked above another candidate, if the voter prefers the former to the latter candidate. The goal is to find a permutation that minimizes the sum of the distances to the voters' permutations, where in principle any distance(-like) function on permutations can be used, e.g. Kendall distance or Footrule distance.

Young & Levenglick [9] show that the Kendall distance is the unique distance function such that the permutation(s) that minimize it have three desirable properties of being neutral, consistent and Condorcet. The latter property means that, if there exists a permutation such that the order of every pair of elements is the order preferred by a majority, then this permutation has minimum distance to the voters' permutations. This distance was already proposed by Kemeny [6] for other reasons ([6] defines axioms on the *distance function*, and finds that the Kendall distance adheres to the axioms), and the problem of finding an optimal ranking with respect to this criterion is now known as Kemeny Rank Aggregation.

In this paper, we suggest a new way of thinking about this problem. Suppose instead of minimizing the total distance from the voters' inputs, we want to find a permutation that makes a majority of the voters "happy"? Of course, a voter is happy when we follow her opinion exactly, and we cannot do this simultaneously for a majority of the voters, unless a majority of the voters is in total agreement. Therefore, our goal is to find a permutation such that there exists no other permutation that a majority of the voters prefer, in the sense that their distance to the alternative permutation is smaller. We call such a permutation a *popular ranking*.

Unfortunately, we show that such a permutation is unlikely to exist: it only exists if Condorcet's paradox does not occur. Even worse than this, we show that if Condorcet's paradox does not occur, then it may still be the case that no popular ranking exists. The only positive news in this context is, perhaps paradoxically, an NP-hardness result: we show that if Condorcet's paradox does not occur, then we can efficiently compute a permutation, which may or may not be popular, but for which the voters will have to solve an NP-hard problem to compute a permutation that a majority of them prefer.

Related Work: Our work is inspired by Abraham *et al.* [1] where the notion of *popular matchings* is introduced. Popular ranking is also related to the problem of designing a voting mechanism in which the voters do not have an incentive to lie about their preferences. However, rather than considering deviations of a single voter, a popular solution is robust against deviations of a majority of the voters. We show that, if the input does not contain Condorcet's paradox, then there is a solution that may or may not be popular, but for which it is computationally hard for a majority of the voters to manipulate the output to their advantage. This result has a similar flavor as a result by Bartholdi *et al.* [3], who demonstrate a voting rule for deciding the "winner" of an election, for which it is computationally hard for a single voter to manipulate the output.

2 Popular Ranking

We are given a set of alternatives $[n]$ (where the notation $[n]$ means $\{1, 2, \ldots, n\}$) and a set of voters $[k]$, where each voter ℓ has a complete ordering of the alternatives. We will denote these complete orderings of a voter ℓ as a list of the alternatives, where an alternative earlier in the list is preferred to elements that succeed it, and use the notation $\pi_\ell^{-1} : [n] \to [n]$ (the use of "${}^{-1}$" will become clear shortly), where $\pi_\ell^{-1}(i)$ is the alternative at position i in the ordering of voter ℓ. Note that we can interpret π_ℓ^{-1} as a permutation. Further, the inverse of π_ℓ^{-1}, which we will denote by π_ℓ, is well defined and can be interpretated as the position of the alternatives in the list of voter ℓ. We will use $\mathrm{list}(\pi_\ell)$ to denote the ordered sequence $(\pi_\ell^{-1}(1), \pi_\ell^{-1}(2), \ldots, \pi^{-1}(n))$.

The Kendall distance between two permutations π, σ, denoted by $K(\pi, \sigma)$, is defined as the number of pairwise disagreements of π and σ, i.e. $K(\pi, \sigma) = \#\{i, j :$

$\pi(i) < \pi(j)$ and $\sigma(i) > \sigma(j)\} + \#\{i, j : \pi(i) > \pi(j)$ and $\sigma(i) < \sigma(j)\}$.

Definition 2.1 *We say a permutation* π *is* popular, *if* $\not\exists \pi'$ *such that* $K(\pi_\ell, \pi') < K(\pi_\ell, \pi)$ *for a strict majority of the voters* $\ell \in [k]$.

We define the *majority graph* $G = (V, A)$ for an instance as the directed graph which has a vertex for every $i \in [n]$ and an arc (i, j) if a majority of the voters $\ell \in [k]$ has $\pi_\ell(i) < \pi_\ell(j)$. Condorcet observed that such a graph may have a cycle; this is known as "Concorcet's paradox".

Lemma 2.1 *No popular ranking exists if the majority graph has a directed cycle.*

Proof (sketch) (sketch) If we order the elements from left to right according to a ranking π, then there must be some arc (i, j) in the graph that is a *back arc*, i.e. for which $\pi(j) < \pi(i)$. Let π' be the permutation we obtain by swapping i and j, i.e. $\pi'(i) = \pi(j), \pi'(j) = \pi(i)$ and $\pi'(t) = \pi(t)$ for all $t \neq i, j$. Then one can show that a strict majority of the voters prefer π' to π, namely the voters ℓ who have $\pi_\ell(i) < \pi_\ell(j)$. ∎

If the majority graph is acyclic, then a popular ranking could exist. We consider the case when the majority graph is a tournament, i.e. for every i, j exactly one of the arcs (i, j) and (j, i) is in G. Note that the majority graph is always a tournament if the number of voters is odd. By Lemma 2.1, the only permutation that *could* be popular is the permutation we obtain by topologically sorting the majority tournament. However, it is not the case that this ranking is always a popular ranking, as we show in the full version of this paper [8].Even though the topologically sort of the majority tournament is not necessarily a popular ranking, it turns out that it is a "good" permutation in the sense that it is NP-hard to find a ranking that a majority of the voters prefer. The proof of the following theorem is given in the full version [8].

Theorem 2.2 *Given an input to the popular rank aggregation problem with an acyclic majority graph, it is NP-hard to find a ranking* ρ *that a majority of the voters* S *prefers to a topological sort of the majority graph, even if* S *is given.*

3 Directions

We have seen that a popular ranking does not always exist, even if the majority graph has no cycles. Perhaps popularity is asking for too much and we should relax our objective. It is an interesting question whether there exists a suitable relaxation of the notion of popularity, so that one can get positive results. One way of relaxing the notion is looking for rankings with *least-unpopularity-factor* (McCutchen [7] introduced this notion for matchings). The bad news is that it can be shown that the unpopularity factor of the permutation π we obtain by topologically sorting the majority tournament may be unbounded. It is an open question however whether there exists a permutation with bounded unpopularity (and if so, what this uniform bound is) and whether such a permutation can be found in polynomial time.

Acknowledgements
The first author thanks Chien-Chung Huang for helpful discussions.

References

[1] D. J. Abraham, R. W. Irving, T. Kavitha, and K. Mehlhorn. Popular matchings. *SIAM J. Comput.*, 37(4):1030–1045, 2007.

[2] K. J. Arrow. *Social choice and individual values.* Yale University Press, 1951.

[3] J. Bartholdi, C. Tovey, and M. Trick. The computational difficulty of manipulating an election. *Social Choice and Welfare*, 6:227–241, 1989.

[4] J. Borda. *Memoire sur les elections au scrutin.* Histoire de l'Academie Royal des Sciences, 1781.

[5] M. Condorcet. *Sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix.* L'Imprimerie Royale, Paris, 1785.

[6] J. Kemeny. Mathematics without numbers. *Daedalus*, 88:575–591, 1959.

[7] R. McCutchen. The least-unpopularity-factor and least-unpopularity-margin criteria for matching problems with one-sided preferences. In *LATIN'08*, pages 593–604, Búzios, Brazil, 2008.

[8] A. van Zuylen, F. Schalekamp, and D. P. Williamson. Popular ranking. Full version available at `http://www.mpi-inf.mpg.de/~anke/`.

[9] H. P. Young and A. Levenglick. A consistent extension of Condorcet's election principle. *SIAM Journal on Applied Mathematics*, 35(2):285–300, 1978.