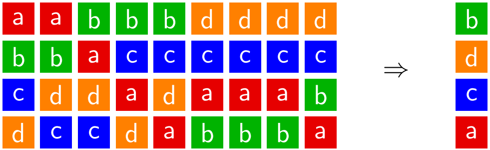
Rank Aggregation Using Scoring Rules

Niclas Boehmer[1], Robert Bredereck[2], Dominik Peters[3]

[1]TU Berlin [2]TU Clausthal [3]CNRS, LAMSADE, Université Paris Dauphine–PSL

AAAI 2023, 2023-02-09, Washington D.C.

Rank Aggregation



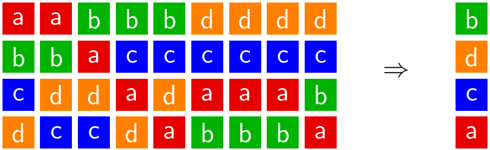
Rank Aggregation

- Fundamental task in social choice theory.
- Numerous applications besides preference aggregation:
  meta-search engines, sports tournaments, multi-criteria decisions, ...

Best-known method: Kemeny method

- good axiomatic properties
- maximum likelihood estimator if there is a ground truth
- but hard to compute when there are many candidates
- and also hard to verify if a ranking is optimal (thus hard to audit)

Rank Aggregation



Project Goal

Understand the simple alternative: scoring-based rank aggregation rules.

not so simple alternative: approximation algorithms based on derandomization etc.

Scores: for example plurality (number of times ranked first), Borda (average rank), veto (number of times not ranked last)

- How to aggregate rankings based on scoring rules?
- What are the properties (axiomatic, computational efficiency) of scoring-based rank aggregation rules?
- How well do these rules perform in practice?

Using Scoring Rules to Aggregate Rankings

Three ways of using scoring rules to aggregate rankings of m candidates:

Score

Rank the candidates in order of their score.

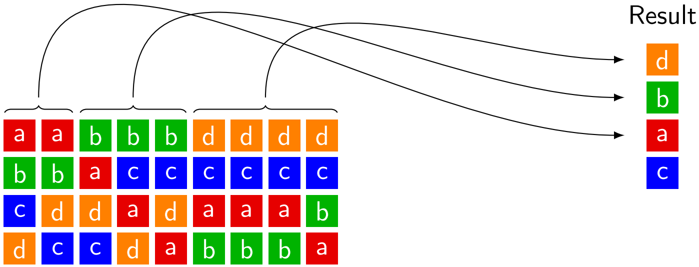
Sequential-Winner

- Take the candidate c with highest score; rank it at position $i := 1$.
- **Repeat:** Delete c from profile; re-calculate scores; compute new candidate c with the highest score & rank it at position $i := i + 1$.

Sequential-Loser

- Take the candidate c with lowest score; rank it at position $i := m$.
- **Repeat:** Delete c from profile; re-calculate scores; compute new candidate c with the lowest score & rank it at position $i := i - 1$.

Examples: Plurality Score



Examples: Sequential Plurality Winner

a	a	b	b	b	d	d	d	d
b	b	a	c	c	c	c	c	c
c	d	d	a	d	a	a	a	b
d	c	c	d	a	b	b	b	a

Result

d
c
a
b

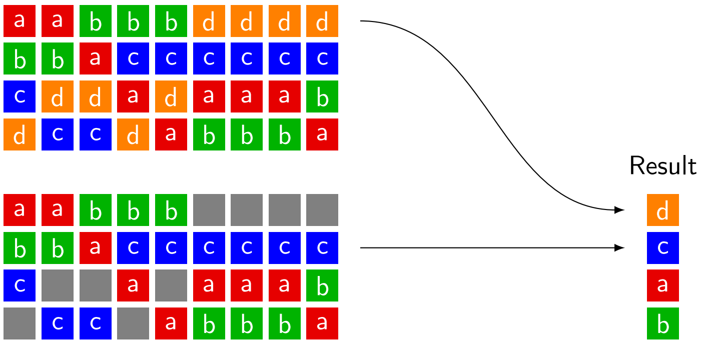
Examples: Sequential Plurality Winner

a	a	b	b	b	d	d	d	d
b	b	a	c	c	c	c	c	c
c	d	d	a	d	a	a	a	b
d	c	c	d	a	b	b	b	a

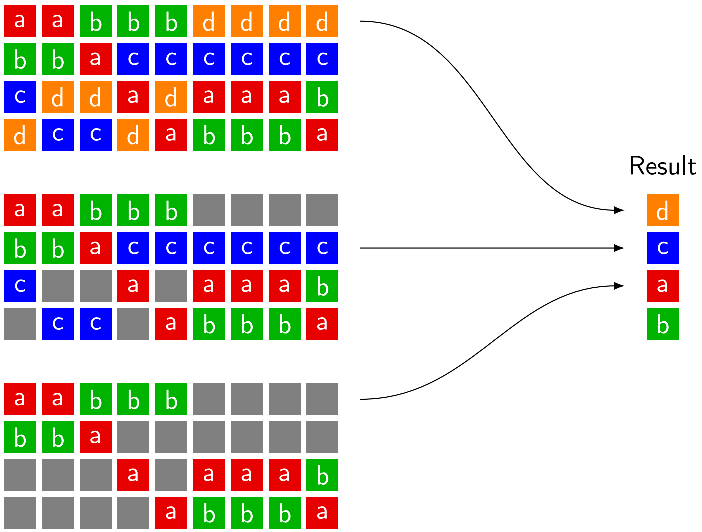
Result

d
c
a
b

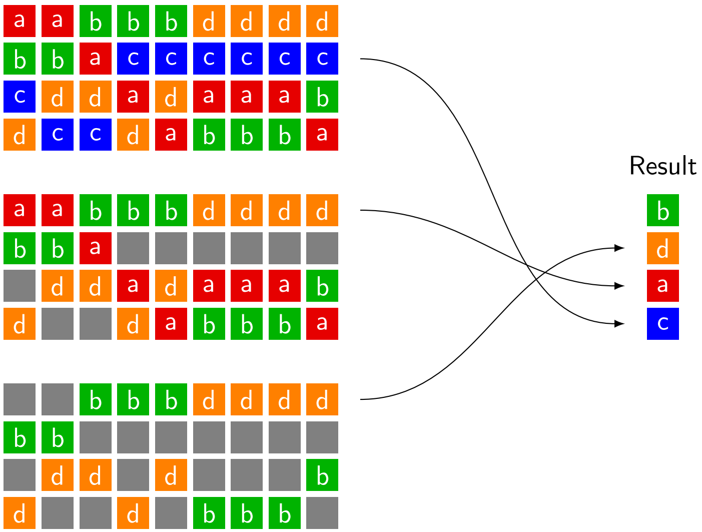
Examples: Sequential Plurality Winner



Examples: Sequential Plurality Winner



Examples: Sequential Plurality Loser



Axiomatic Properties

		Score			Seq.-Winner			Seq.-Loser		
	Kemeny	P	V	B	P	V	B	P	V	B
Independence at the top	✓				✓	✓	✓			
Independence at the bottom	✓							✓	✓	✓
Reinforcement	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Reinforcement at the top		✓	✓	✓	✓	✓	✓			
Reinforcement at the bottom		✓	✓	✓				✓	✓	✓
Condorcet winner at top	✓									✓
Copy majority	✓				✓				✓	
Independence of clones								✓		

K=Kemeny, P=Plurality, V=Veto, B=Borda

Much of this adapted from: Freeman, Brill, and Conitzer,
On the Axiomatic Characterization of Runoff Voting Rules, AAAI 2014.

Simulations I: How close to Kemeny?

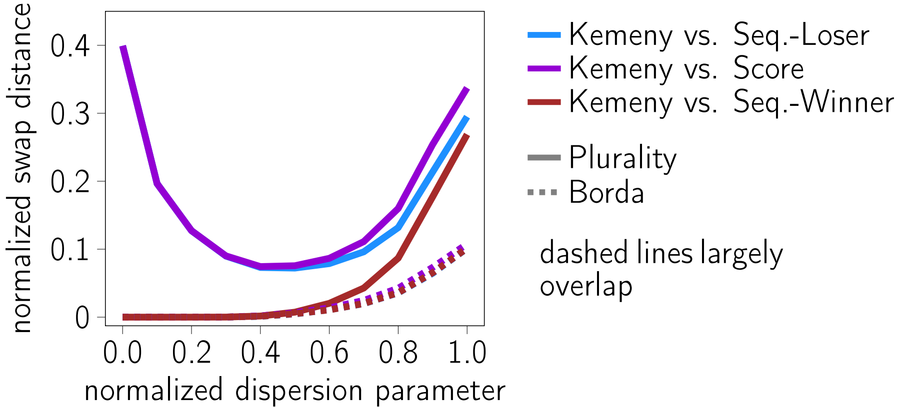


Figure: Pairwise average normalized swap distance on Mallows profiles with 10 candidates and 100 voters.

Seq.-Plurality-Winner produces rankings closest to Kemeny's ranking (among plurality-based rules).
All Borda-based rules are quite close.

Simulations II: How similar to each other?

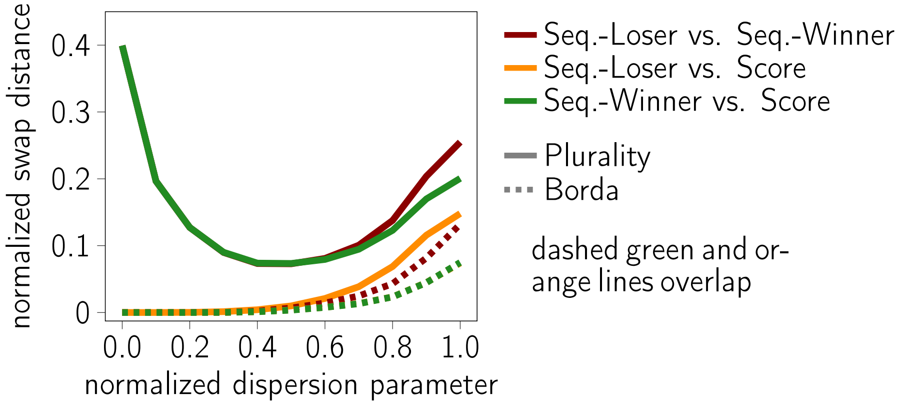
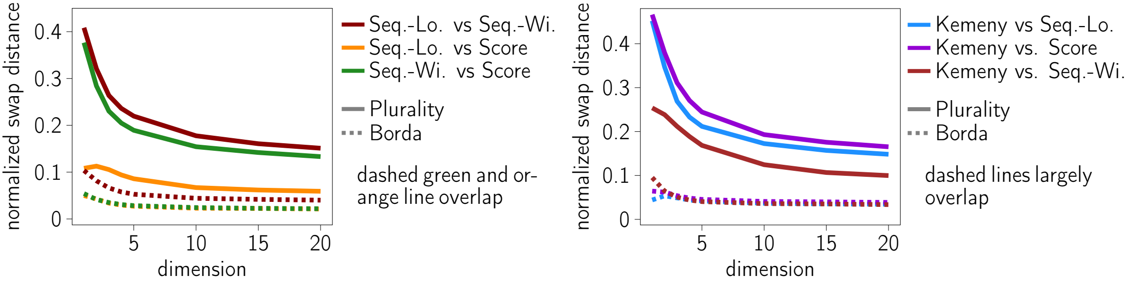


Figure: Pairwise average normalized swap distance on Mallows profiles with 10 candidates and 100 voters.

Seq.-Plurality-Loser and Plurality-Score are very similar, whereas Seq.-Plurality-Winner produces substantially different rankings.

Simulations III: Euclidean profiles



(a) Pairs of scoring-based methods

(b) Scoring-based vs. Kemeny ranking

Figure: Pairwise average normalized swap distance on Euclidean profiles with 10 candidates and 100 voters.

Simulations IV: Where are the disagreements?

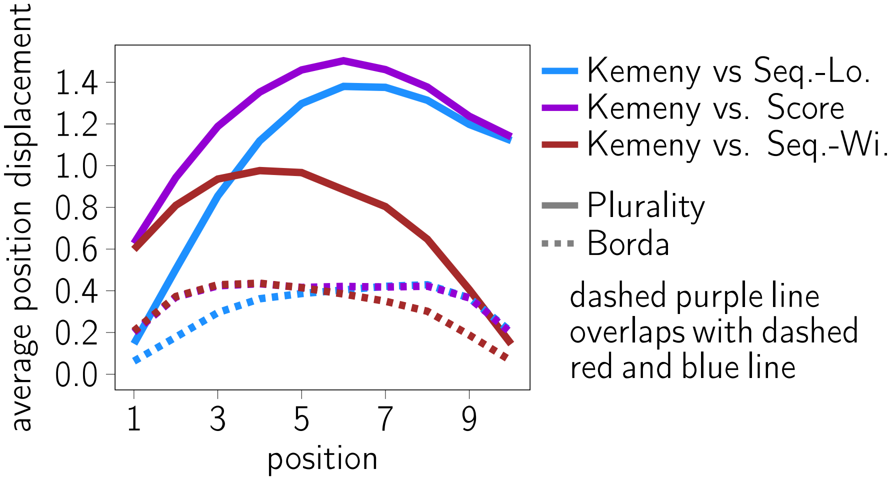
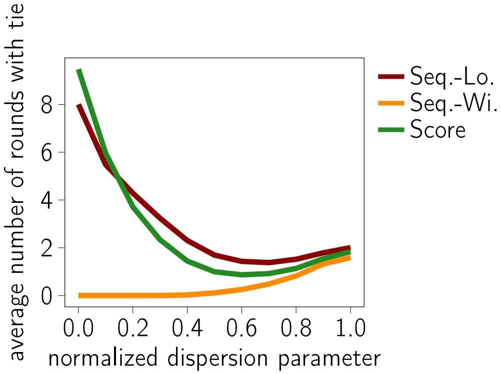
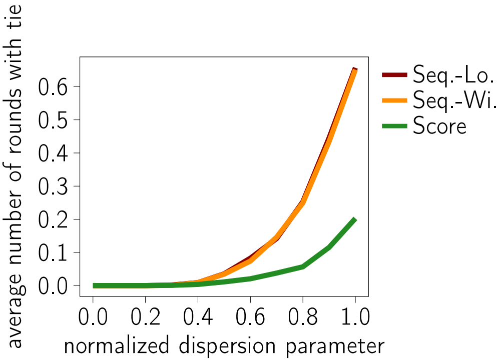


Figure: For pairs of rankings, average position displacement on each position for profiles generated using the Mallows model with $\mathrm{norm}\text{-}\phi = 0.8$ with 10 candidates and 100 voters.

Simulations V: Frequency of ties



(a) Plurality



(b) Borda

Figure: Average number of rounds in which a tie occurs for Seq.-Plurality-Winner, Seq.-Plurality-Loser, and Plurality-Score on profiles with 10 candidates and 100 voters sampled from the Mallows model.

Computational Complexity: Studied Problems

Position-k Determination for social preference function f	
Given:	A ranking profile P over candidate set C, a designated candidate $d \in C$, and an integer $k \in [|C|]$.
Question:	Is there a ranking $\succ$ selected by f on P where d is in position k, i.e., $\succ \in f(P)$ with $\mathrm{pos}(\succ, d) = k$?

Problem variants

- Top-k Determination: Can a given candidate be placed in top-k positions?
- Winner Determination: Special case for $k = 1$: Can a given candidate be placed in the top position?

Computational Complexity: Sequential-Loser

Sequential-*-Loser		n	m
Plurality (STV)	NP-c.	FPT	FPT
Veto (Coombs)	NP-c.	W[1]-h., XP	FPT
Borda (Baldwin)	NP-c.	NP-c. $(n = 8)$	FPT

Table: Hardness: Winner Determination; Algorithms: Position-k Determination.

- NP-hardness results stated or proven by Conitzer, Rognlie, and Xia (IJCAI '09) and Mattei, Narodytska, and Walsh (ECAI '14). Ours are a bit nicer and simpler IMHO.
- FPT: $m!$ dependence is obvious, can get 2^m by enumerating 'survival sets'. ETH lower bounds: cannot be solved in time $2^{o(m)} \cdot \text{poly}(n, m)$.
- Baldwin is hard for 8 voters: Use McGarvey's theorem very carefully. Same technique that shows Kemeny is hard for $n = 4/7$ voters.

Computational Complexity: Sequential Winner

Sequential-*-Winner		n	k	$n+k$	m
Plurality	NP-c.	W[1]-h., XP	W[1]-h., XP	FPT	FPT
Veto	NP-c.	FPT	W[2]-h., XP	FPT	FPT
Borda	NP-c.	NP-h. $(n = 8)$	W[1]-h., XP	?	FPT

Table: Hardness: Top-k Determination; Algorithms: Position-k Determination