Institute for Logic, Language and Computation

Diversity, Agreement, and Polarization in Elections

Dennis Lindberg, Matteo Mercuri

15/01/2026

The Problem

- Agreement has been disproportionally studied over diversity and polarisation.
- $A(E)$ has been characterised axiomatically.
- $1 - A(E)$ as a measure of disagreement (not diversity nor polarization).
- Polarisation creates a loss of diveristy.
- **Main goal**: *design election indices that distinguish these notions*.

Structure of the Paper

- Defining diversity, agreement, and polarization (for ordinal elections).
- Finding diversity and polarization indices.
- The indices are based on the k-Kemeny problem.
- Computation of k-Kemeny distance.
- To evaluate the indices, *maps of elections* are used.

Defining elections

- Election $E = (C, V)$.
- $p_E(a, b)$: the fraction of voters in E that prefer a over b.
- Three characteristic elections:
  - **Identity (ID):** all votes are identical: perfect agreement.
  - **Antagonism (AN):** exactly half of the voters have one preference order; other half has the reversed one: perfect polarization.
  - **Uniformity (UN):** contains the same number of copies of every possible preference order: perfect diversity.

Kemeny Rankings and Swap Distance

- $swap(u, v)$: swap distance, the minimal number of swaps of consecutive candidates required to transform u into v.
- **Kemeny ranking** of $E = (C, V)$: a linear order over C that minimizes the sum of its swap distances to the votes from V.
- $dswap(E, F)$: the isomorphic swap distance between two elections $E = (C, V)$ and $F = (D, U)$, s. t. $|C| = |D|$, $V = (v_1, ..., v_n)$, and $U = (u_1, ..., u_n)$.

Defining Maps of Elections

- **Map of elections**: a collection of elections represented on a 2D plane as points.
- The Euclidean distances between the points reflect the similarity between the elections.
- Maps created using isomorphic swap distance (the candidate sets considered are small).
- Compute the distance between each two elections, then run the multidimensional scaling algorithm (MDS) to find an embedding of points on a plane that reflects the computed distances.

Defining diversity, agreement, and polarization

- **Election index**: a function that given an election outputs a real number.
- The **agreement index** of an election $E = (C, V)$:

$$A(E) = (\sum_{\{a,b\} \subseteq C} |p_E(a, b) - pE(b, a)|)/(\binom{|C|}{2})$$

  - The index takes values between 0 and 1.
  - 0 means perfect disagreement; 1 means perfect agreement.
  - We have A(ID) = 1 and A(UN) = A(AN) = 0.

Diversity and Polarization Indices

- The diversity and polarization indices (main contribution).
- Defined on top of a generalization of the Kemeny ranking problem:

  $k-$Kemeny rankings of election $E = (C, V)$ are the elements of a set $\Lambda = \lambda_1, ..., \lambda_k$ of $k-$linear orders over C that minimize $\sum\{v \in V\} min_{i \in [k]} swap(v, \lambda_i)$.
- The $k-$Kemeny distance, $k_k(E)$, is equal to this minimum.
- Finding $k-$Kemeny rankings: finding an optimal split of votes into k groups; then minimizing the sum of each group's distance to its Kemeny ranking.
- $1-$Kemeny distance: the distance of the voters from the (standard) Kemeny ranking.

The diversity index

- **Desiderata**: diversity index high for UN, small for AN and ID.
- For ID, $1-$Kemeny distance is equal to 0.
- For both UN and AN, $1-$Kemeny distance is equal to the maximal possible value: $|V| \cdot \binom{|C|}{2}/2$.
- For $k \geq 2$, the $k-$Kemeny distance of AN is 0; in UN non-negligible positive distances arise.
- Then, the **diversity index** is a normalized sum of all $k-$Kemeny distances:

$$D(E) = (\sum_{k \in [|V|]} k_k(E)/k)/(|V| \cdot \binom{|C|}{2})$$

- Finding a more robust normalization is desirable: diversity of UN grows slightly faster than linearly with the growing number of candidates.

The polarization index

- Strategy: look at AN and the drop from the maximal possible value of the 1−Kemeny distance to zero for the 2−Kemeny (characteristic of polarised elections).
- Divide by $|V| \cdot \binom{|C|}{2}/2$ for normalisation; the index takes values between 0, for the lowest polarization, and 1.
- The **polarization index** of an election $E = (C, V)$:

$$P(E) = 2(k_1(E) - k_2(E))/(|V|) \cdot \binom{|C|}{2})$$

- P(AN) polarization is one, while for ID it is zero; with a growing number of candidates the polarization of UN approaches zero.

k-Kemeny is Hard to Compute

Decision variant of k-`Kemeny` is **NP-Complete**.

The problem is equivalent to the k-`Median` facility location problem:

- voters = clients
- k central rankings = k facilities
- swap distance = distance metric

IDEA: Leverage efficient approximation algorithms for k-`Median`.

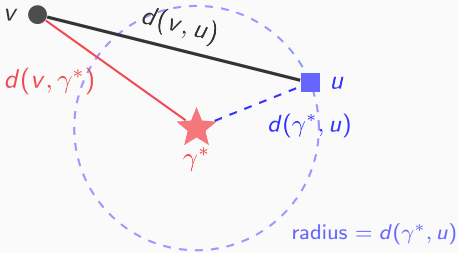
PROBLEM: Scales with the number of facilities, i.e., the $m!$ possible preference rankings...

SOLUTION: k-`Kemeny Among Votes`; restrict search space to preference rankings that appear in the input.

k-Kemeny Among Votes

Theorem. *An α-approximation for k-Kemeny Among Votes is a 2α-approximation for k-Kemeny.*

Proof.



k-Kemeny Among Votes is NP-Complete

Max K-cover (NP-hard)

- Universe $X = \{x_1, \ldots, x_N\}$, and subsets $\mathcal{S} = \{S_1, \ldots, S_M\}$
- Goal: Pick $K \leq M$ subsets to maximize coverage

Reduction to k-Kemeny Among Votes

- Set Voters: Many copies of v_j per set S_j $\quad + $ preferences
- Element Voters: One e_i per element x_i $\quad + $ preferences

$$\text{swap}(e_i, v_j) = \begin{cases} 3 & \text{if } x_i \in S_j \\ 3 + C & \text{otherwise} \end{cases}$$

Equivalences

Pick k central rankings	$\iff$	Pick K sets
Element close to center	$\iff$	Element covered
Minimize total distance	$\iff$	Maximize coverage

k-Kemeny Among Votes Approximations

Algorithm	Approximation Ratio
Greedy	?
Local search	$6 + \frac{4}{p}$
Combined heuristic	$6 + \frac{4}{p}$

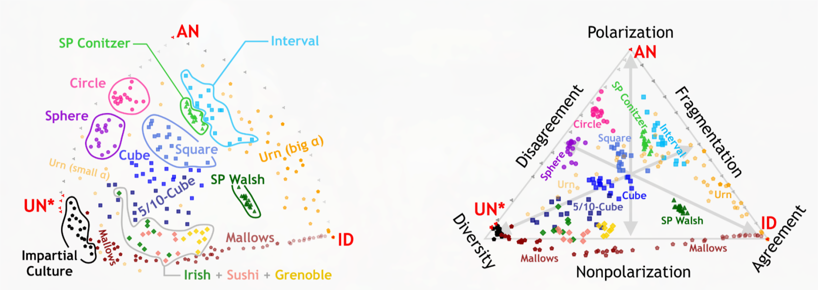


Figure 1: Map of elections (left), and plot where coordinates are agreement and diversity (right).

Understanding the Map of Elections

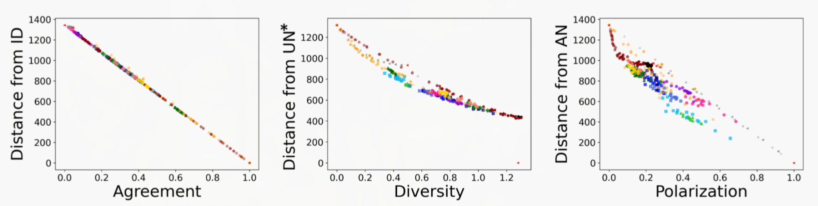


Figure 2: Correlation coefficients below −0.9.

Summary

- $1 - A_E$ captures neither diversity nor polarization, but disagreement.
- (something on compass elections?)
- (something on diversity/polarization indices?) ... k-Kemeny unites all three indices.
- Maps of elections can be understood in terms of Agreement, Diversity, and Polarization.