Topics in Computational Social Choice 2026

Théo Delemazure

Institute for Language, Logic and Computation

University of Amsterdam

Experimental Analyses

January 9th 2026

The social choice pipeline

Preference profile → **Social choice function** → **Collective decision**

The social choice pipeline

Plurality	Borda	IRV	Approval	…

Which voting rule to chose?

Axiomatic analysis: which normative properties are satisfied by the different voting rules?

Complexity analysis: how hard is it to compute the result? To manipulate?

Experimental analysis: how do the different voting rules compare in practice?

Typical experimental questions

- How much time does it take to compute the result of a social choice function? (related to **computational complexity**)
- How **similar** are different rules?
- How often does a rule satisfies/fails some property? (related to **axiomatic analysis**)
- Which rule return the **“best”** winner? (when “best” have a meaning)

Experiments on the **preference profiles**

- How **similar** are two different *profiles*?
- How often does a *profile* satisfies/fails some property? (related to **axiomatic analysis**)
- Can we **learn** some underlying structure of the electorate from the profile?

Part 1
Implementing voting rules

Computational complexity of voting rules

Polynomial: Plurality, Veto, Borda, any positional scoring rule, Copeland, IRV (*tie-breaker*), Approval, MAV, S-PAV, S-CCAV, …

NP-hard: Kemeny, Young, IRV (*parallel-universe*), PAV, CCAV, …

Computational complexity of voting rules

Kemeny rule:

$$F(P) = \text{argmin}_{\succ^* \in L(A)} \sum_{i \in V} d_{KT}(\succ_i, \succ^*)$$

Kendall-tau distance:

$$d_{KT}(\succ_1, \succ_2) = |\{ x, y \in C : x \succ_1 y \text{ and } y \succ_2 x\}|$$

Going around computational hardness

Algorithmic techniques to make the computation faster. **Examples:**

- You can sometimes **reduce the search space** by removing options that are clearly not optimal.
- You can sometimes **use heuristics to “guess” the optimal solution** and checking it.
- For NP-hard problem, you can encode the rule (or the problem) in an **ILP solver or a SAT solver** to solve it more efficiently.

Sometimes the computation becomes easier if **we fix some parameter(s)** (number of candidates, size of the committee, size of approval ballots, allowed preferences): field of **parametrized complexity.**

You could also design **approximation algorithms** running in polynomial time.

Other computational complexity problems

Many problems in social choice have been studied **under the computational complexity angle** and shown to be hard to solve:

- Strategic voting, control and bribery.
- Identifying some structure in preferences (e.g., (near)-singlepeakedness).
- Possible winners when preferences are incomplete.
- ...

The algorithmic aspects of **how to compute things efficiently** are often underrepresented in COMSOC research.

Tools for voting rules

Mainly **Python libraries**, but you can also find implementation for other languages online.

- `pref_voting` for preferential voting (Pacuit and Holliday)
- `whalrus` for preferential voting (Durand)
- `abcvoting` for approval-based committee voting (Lackner)
- `pabutools` for participatory budgeting (Rey et al)

Part 2

Generating preferences

Reminder: Model and notations

Fix a finite set $A = \{a, b, c, \ldots\}$ of **alternatives** with $|A| = m \geq 2$.

A ranking is a linear order over the alternatives $\succ \in L(A)$.

Each **voter** of the finite set $N = \{1, \ldots, n\}$ supplies a preference ranking (their ballot) $\succ_i$, giving rise to a **preference profile** $P = (\succ_1, \ldots, \succ_n) \in (L(A))^n$.

A **voting rule** for A and N selects one or (in case of ties) more winners for every such profile:

$$F : (L(A))^n \to 2^A \setminus \{\emptyset\}$$

The Impartial Culture model

In the **Impartial Culture (IC) model**, every possible ranking as the same probability to be sampled i.i.d. for each voter:

$$P(\succ_i = \succ) = \frac{1}{m!}$$

For all ranking $\succ \in L(A)$ and all voter $i \in V$.

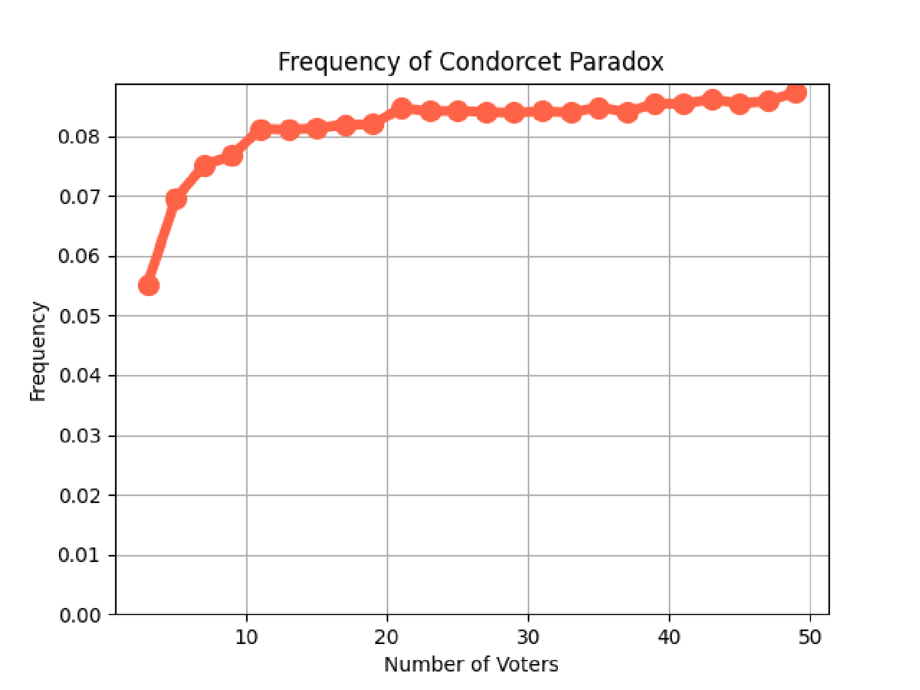
Remark: IC is very simplistic and unrealistic so it should not be the only model used, but it is a frequently used model, so it serves as a baseline.

Example of use case

First used in [Guibault, 1952] to compute **the probability of a Condorcet paradox** (but without computer simulation).

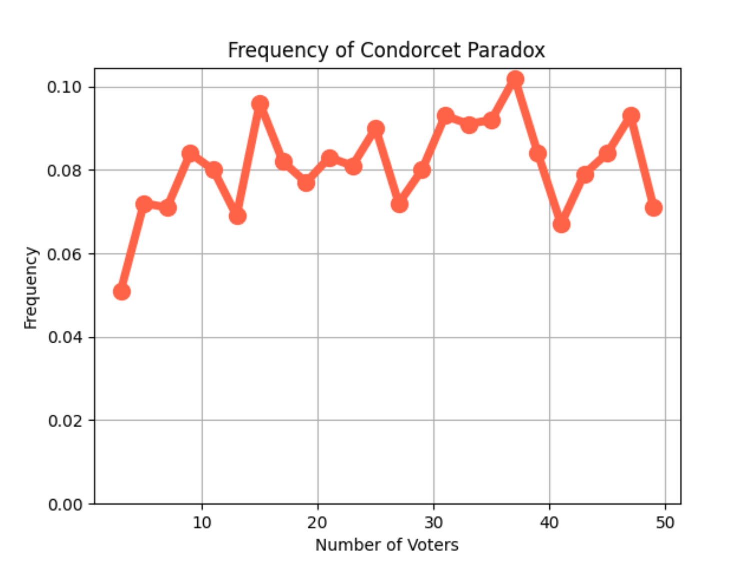
Initialize $c = 0$
Repeat t times:
1. Sample a preference profile P using IC.
2. If there is a Condorcet paradox in P, $c \mathrel{+}= 1$

Return c/t



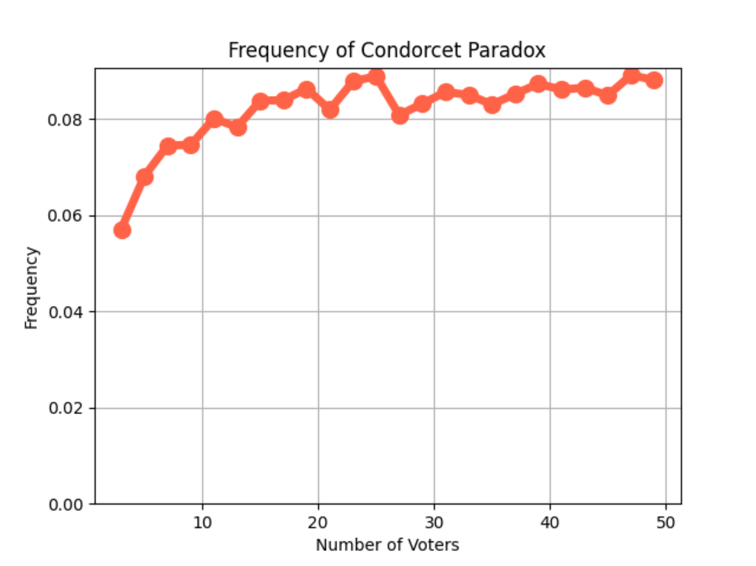
Example of use case

Good practice: sample as many profiles as possible.



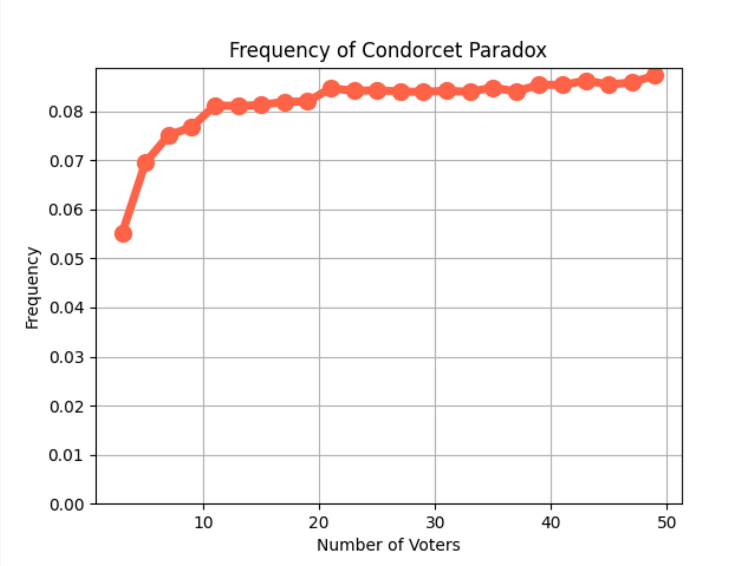
$t = 1\,000$

2s



$t = 10\,000$

18s



$t = 100\,000$

3min

Example of use case

We can similarly compute **similarities** between rules, and how often rules satisfies some properties.

Question: Can we use this culture to say which rule selects the “best” candidate?

Variants of IC

Impartial Anonymous Culture (IAC): Every vote distribution is equally likely to occur.

Impartial Anonymous and Neutral Culture (IANC): Every vote distribution is equally likely to occur.

Remark: in general, we only use IC as they give very similar results.

Exercise: why are there different than IC?

Variants of IC

With $n = 2$ and $m = 2$:

		Voter 1: $a \succ b$	Voter 1: $b \succ a$
Voter 2	$a \succ b$	$a \succ b$ $a \succ b$	$a \succ b$ $b \succ a$
Voter 2	$b \succ a$	$b \succ a$ $a \succ b$	$b \succ a$ $b \succ a$

IC

$a \succ b$ / $a \succ b$ $\quad$ $a \succ b$ / $b \succ a$ $\quad$ $b \succ a$ / $b \succ a$

IAC

$a \succ b$ / $a \succ b$ $\quad$ $a \succ b$ / $b \succ a$

IANC

Mallows' model

In a **Mallows' model**, all rankings are noisy approximations of a ground truth ranking. More formally, there exists a **central ranking** $\succ^*$ such that it is more likely to sample rankings closer to $\succ^*$. The distance between rankings is computed with the **Kendall-tau distance**.

Then, we sample rankings based on the central ranking $\succ^*$ and a **dispersion parameter** $\phi \in [0,1]$:

$$P(\succ_i = \succ \mid \succ^*, \phi) = \frac{\phi^{d_{KT}(\succ, \succ^*)}}{K}$$

with K a normalization constant.

Exercise: what happens when $\phi = 0$? And when $\phi = 1$?

Mallows' model [Mallows, 1957]

Remark: The Kemeny rule is the Maximum Likelihood Estimator (MLE) for Mallows' model central ranking.

Question: Do you expect Condorcet paradox to be more or less frequent for Mallows' model than for IC?

Question: Can you use this model to say which rule selects the "best" alternative?

Mixture of Mallows

In a **mixture of k Mallows**, there are k central rankings $(\succ_1^*, \ldots, \succ_k^*)$ and probabilities $(p_1, \ldots, p_k)$ with $\sum p_j = 1$. For each voter, we select one Mallows according to the probabilities $\left(p_j\right)_j$ and we draw a random ranking according to the Mallows model with central ranking $\succ_j^*$ and dispersion ϕ.

This enables to have **more diversity** in the preferences.

Remark: you can get an "antagonistic" culture by taking 2 Mallows with reversed central rankings.

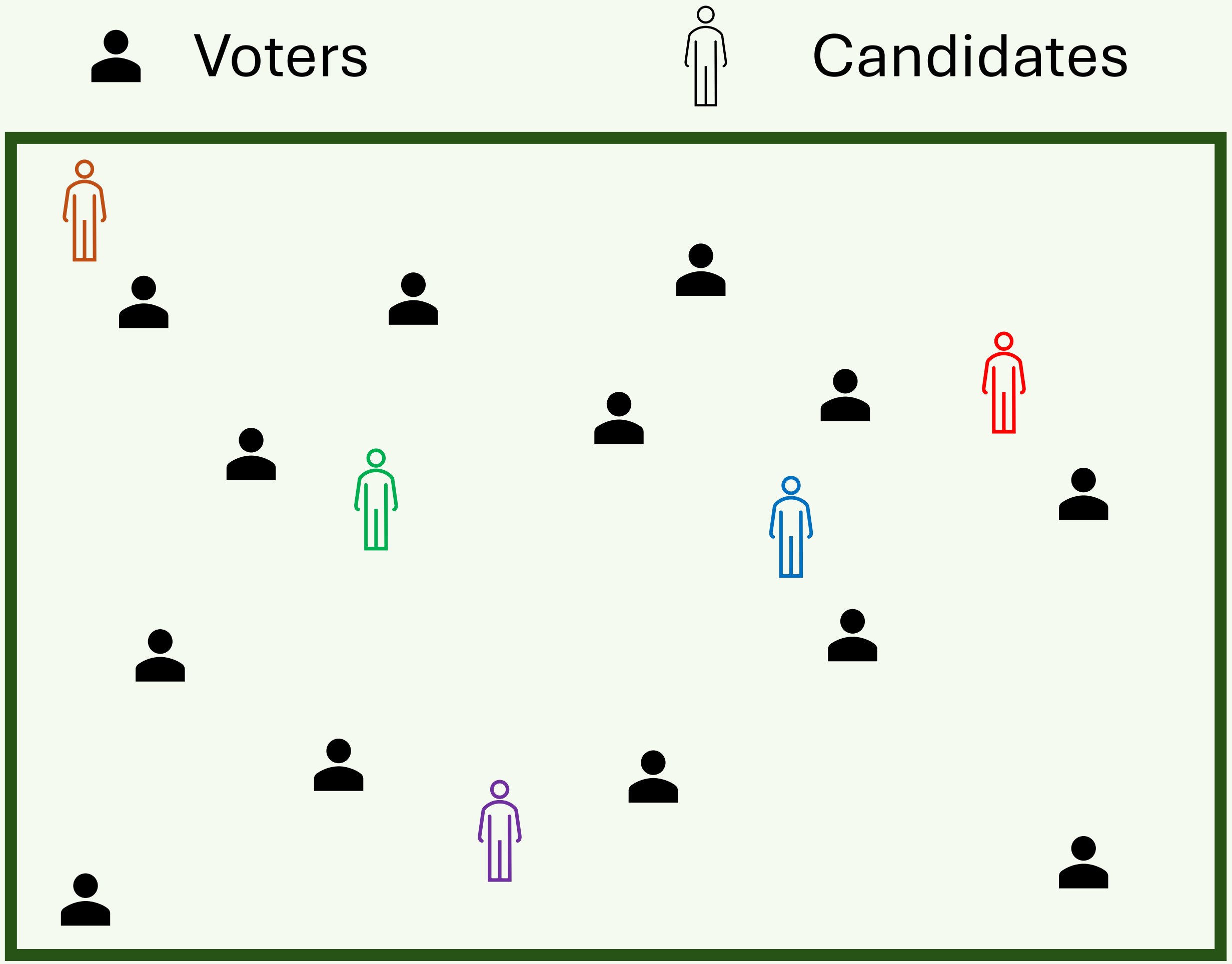
The Urn model [Eggenberger and Pólya, 1923]

We sample rankings based on a **contagion parameter** $\alpha \in [0,1]$ using the following algorithm:

1. Start with an urn containing **one copy** of each possible ranking.
2. Repeat n times:
    1. Draw a random ranking from the urn and add it to the profile.
    2. Put back the ranking in the urn, together with $\alpha \cdot m!$ additional copies.

Question: what happens when $\alpha = 0$? When $\alpha = 1$?

Euclidean preferences [Enelow and Hinich, 1984]

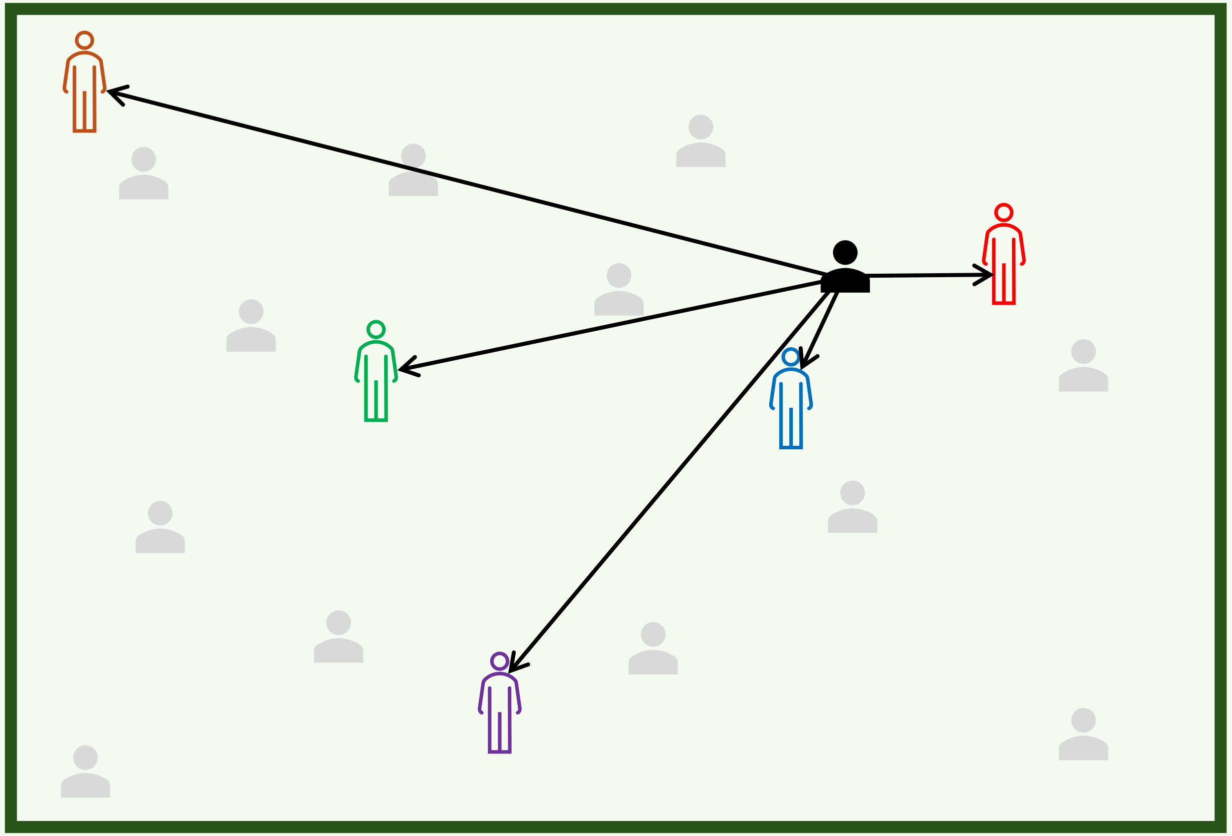


Positions of voters and candidates are sampled randomly in a ***d*-dimensional space.**

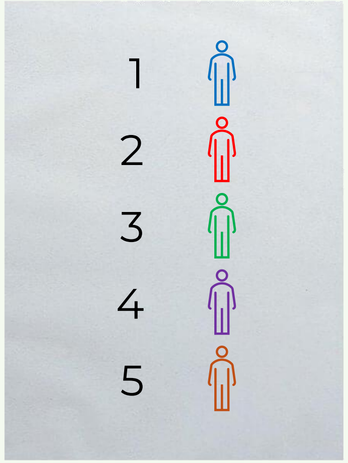
We can use different **distributions**: uniform, Gaussian, multi-pole...

Euclidean preferences

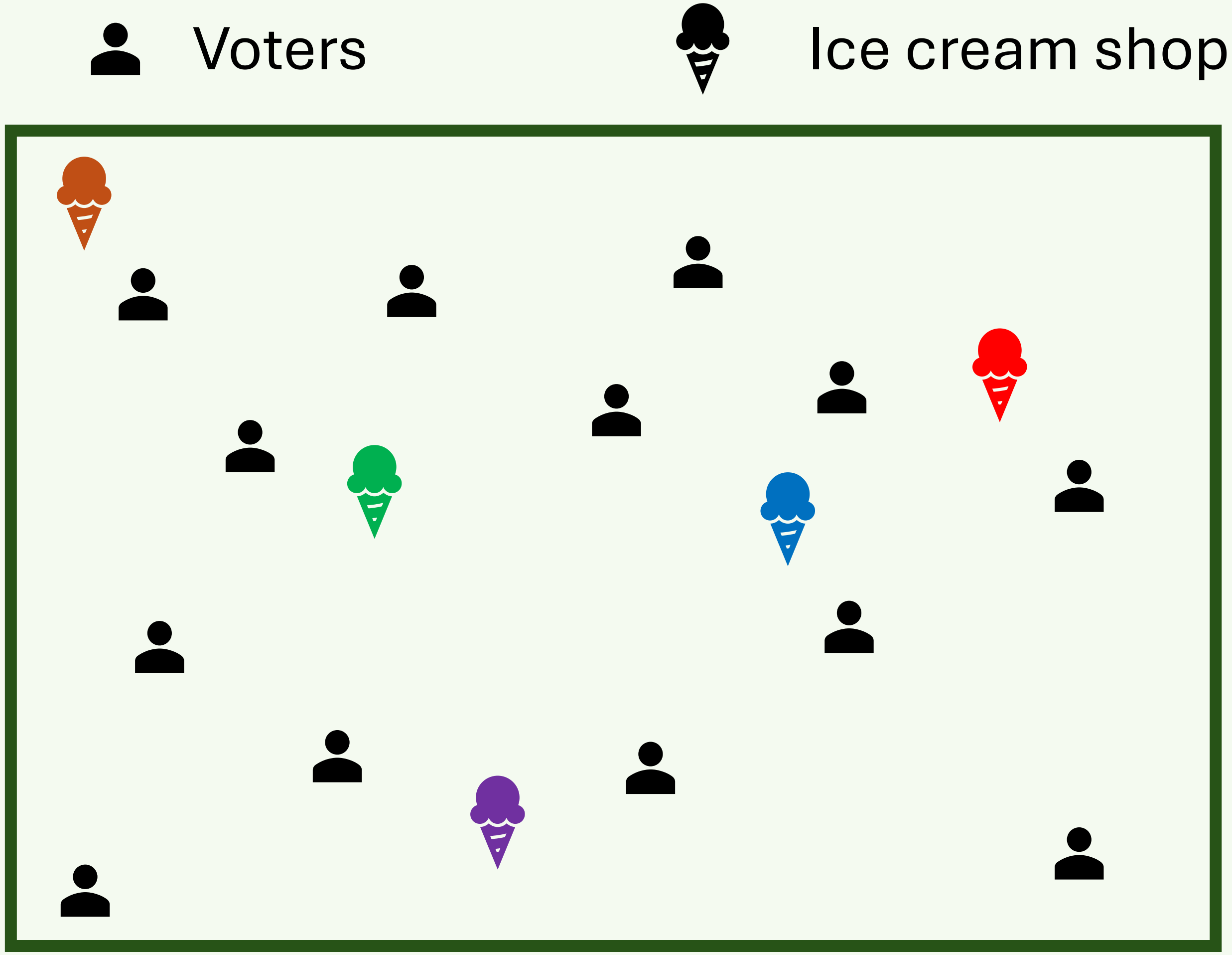
Voters Candidates



We can derive a ranking by saying that voters prefer candidates that are **closer** to them:



Euclidean preferences



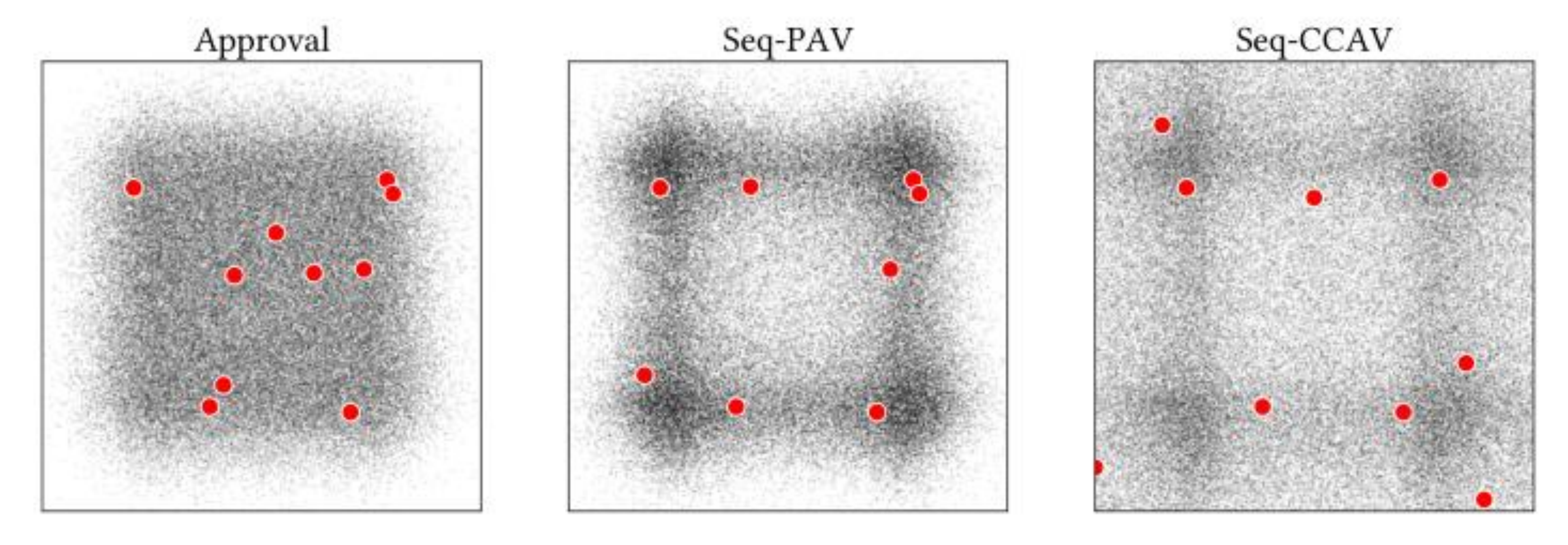
Remark: this is linked to the **facility location** problem.

Exercise: how can we derive **approval preferences** from these positions?

Example of use case: committee voting

Positions of voters and candidates are sampled uniformly at random in $[0,1]^2$.

Question: for ABC rules, what are the positions of the winners?

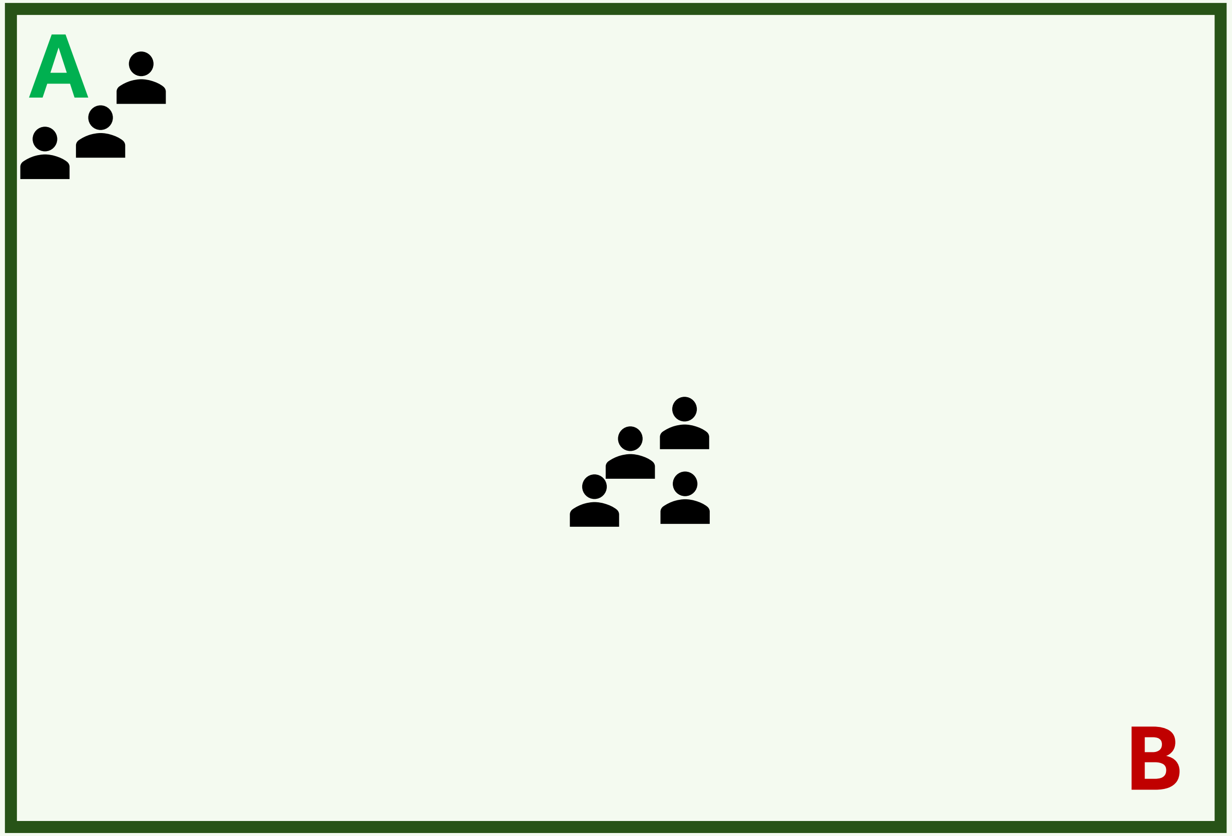


Finding the “best” alternative

The **“best”** alternative is...

...the one that minimizes the sum of the distance to each voter: **Utilitarian approach.**
...the one that minimizes the distance to the furthest voter: **Egalitarian approach.**

Finding the **“best” alternative**



$3 \times A \succ B \qquad 4 \times B \succ A$

Majority rule selects **B**, but **A** is better according to both objective.

Field of **metric distortion**: given a voting rule, how much worse than the optimal do we get in the worst case/average case?

Utilities as preferences

Similarly, you could first draw **utilities $u_i(x)$** for each voter and each alternative and derive preferences from these utilities, such that $u_i(a) > u_i(b)$ implies $a \succ_i b$.

As for Euclidean preferences, this also allows for utilitarian/egalitarian evaluations of the rules (using for instance **distortion**).

Question: if every utility is drawn randomly between [0,1] and we derive rankings from these utilities (assuming no ties), what do we get?

The map of elections [Szufa et al, 2020]

- They sampled preference profiles using **every common probabilistic model.**
- They looked **how similar** are each pair of profile (using some distance function) and they use it to build **a 2D embedding** of these profiles.

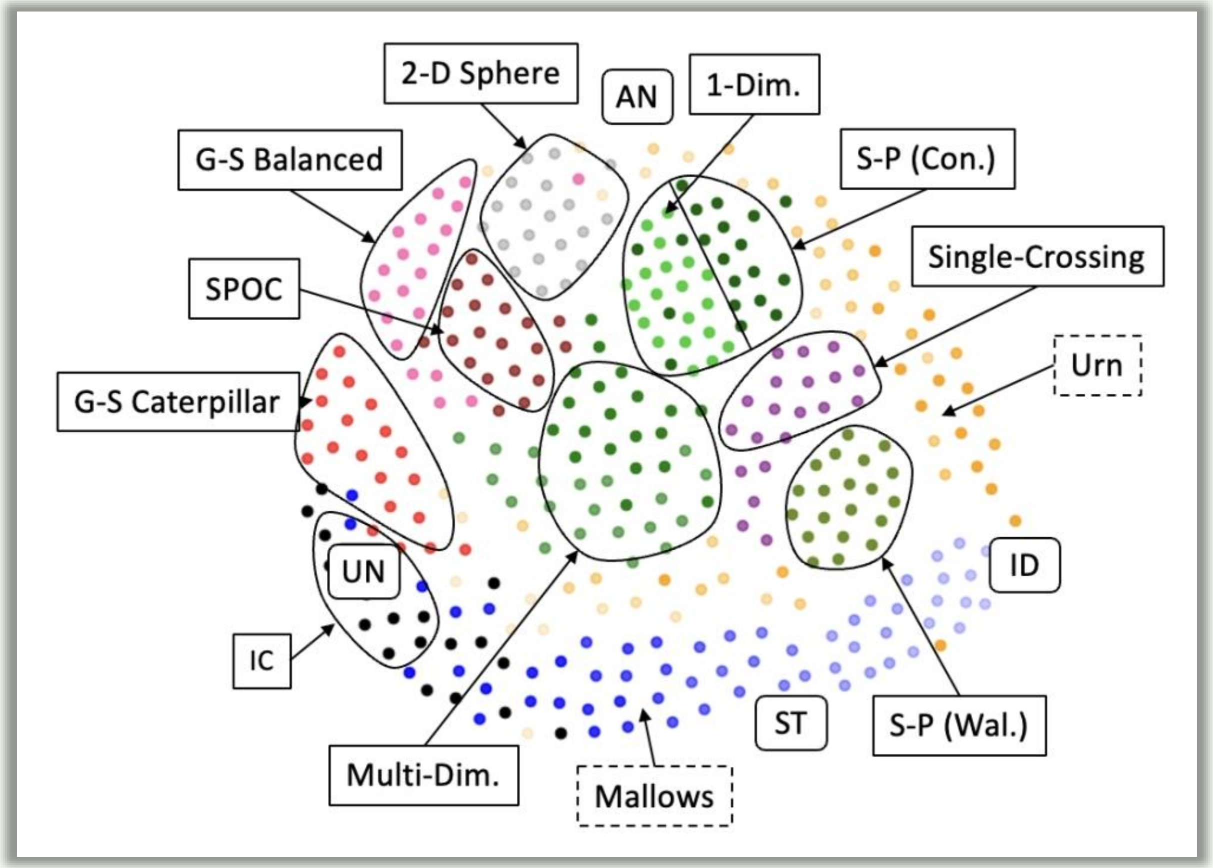


Figure. Map of elections with the isomorphic swap distance. Picture from *Boehmer et al. (2022b)*

The map of elections [Szufa et al, 2020]

This map can be used as a **baseline**, and to see where interesting things are happening in the space of elections.

It is also good to test the **robustness** of a result.

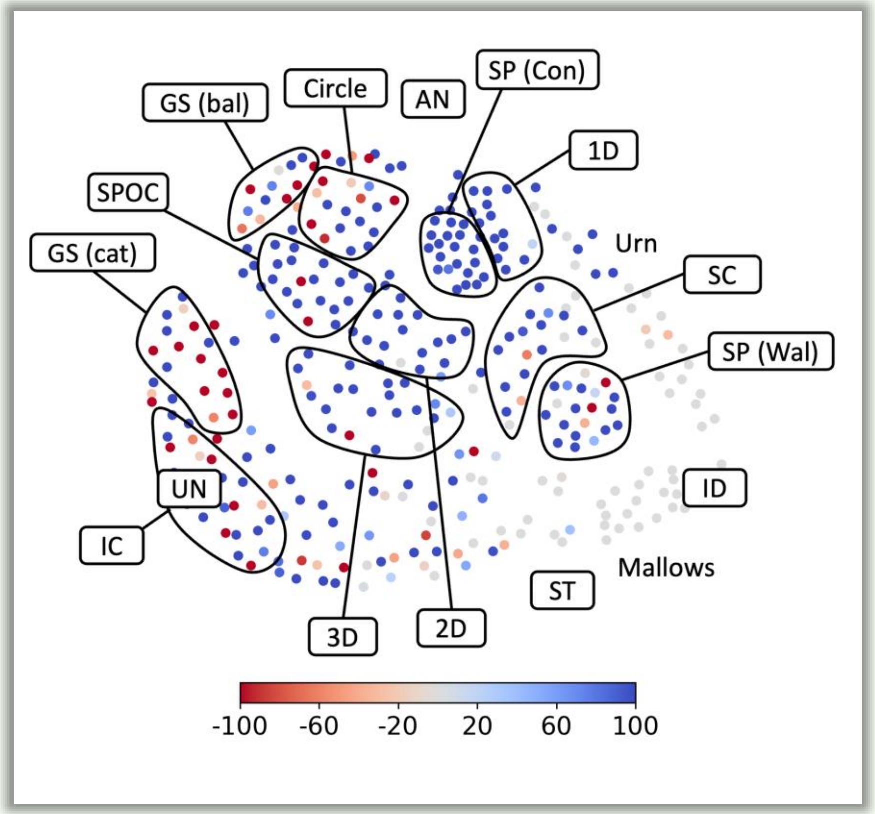


Figure. Map of elections colored based on which of two tested rules return winners with higher Borda scores.

Get **creative**

You may want to design more interesting (or complicated) probabilistic models if you want to illustrate something very specific with your experiment.

Example: if you want to test the robustness of facility location rules, you can design a model in which voter have noisy approximation of their distance to the alternatives.

Good practices

- Run the experiments **for different values of the parameters** (in particular, the number of voters and alternatives, but also parameters of the model).

- For each set of values of parameters, run the experiment **on a lot of profiles** and take the average.

- Run the experiments on profiles sampled using **various models** (can be done easily with map of election).

- Give enough details in your paper (and provide the code on some repository) for **reproducibility**.

Tools to generate data

Python libraries:

- `prefsampling` for preference samplers (Rey and Szufa)
- `mapel` for the map of election (Szufa and Kaczmarczyk)

You can also take a look at the following **paper**: *Guide to Numerical Experiments on Elections in Computational Social Choice* [Boehmer et al, 2024]

Part 3
Collecting real preferences

Libraries of preferences

Preflib.org (Mattei, Walsh, Rey) for various datasets

Pabulib.org (Faliszewski et al.) for participatory budgeting data

Voting Experiment Library (Delemazure)

Finding preferences

There are many datasets **freely available online** that could be converted into preferences.

Ratings: movies (Movielens, Netflix), restaurants (Tripadvisor), books, ...
Competitions: Spotify charts, Tour de France, Formula 1, Eurovision, ... [Boehmer, 2023]; sport competitions
Parliaments: vote on bills (approval?)

Note that almost any **binary data** can fit the approval model (e.g., clicks of users on a website, likes on tweets)

Finding preferences: be creative!

Example: on Tiermaker.com, people can order alternatives in Tierlists, and these are easily downloadable. This can be converted to weak orders or approval preferences.

Remark: even if scraping is sometimes used to collect data, there are good argument to say that it is unethical and is not always appreciated in conferences.

Collecting preferences

You can also run **your own experiment**. Different main kinds of experiments:

- Lab experiments
- In situ experiments
- Online experiments
- Polling institutes

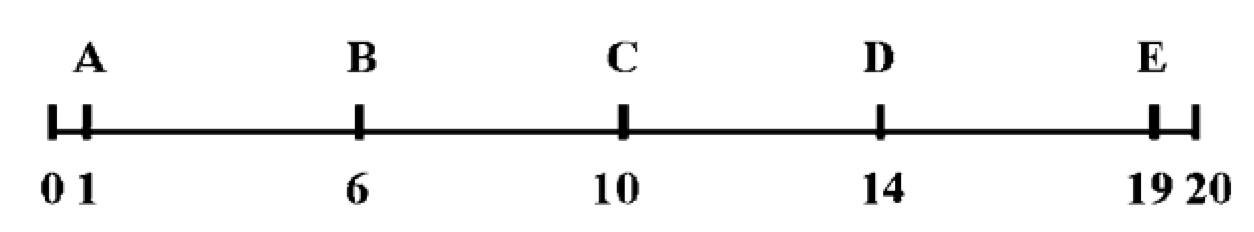
Lab experiment

Researcher want to test one or several **hypothesis**, and generally prepare a very specific voting scenario, with fake alternatives on which participants are told their preferences.

Participants are **paid** and get a better reward if their favorite alternative is elected.

Typically used for experiments **on strategic behavior of voters**.

Example: alternatives are A, B, C, D, E. You are at position 3. Your reward is $20 - d$ where d is your distance to the winner.



In situ experiment

Researchers directly find participants **on the field** (e.g. at the voting station) and ask them to vote with alternative voting methods (generally involving a preferential or approval ballots).

This has been done during **every French presidential** election since 2002.

It needs a lot of **organization** (in particular to get the required authorization from the city to conduct the experiment).

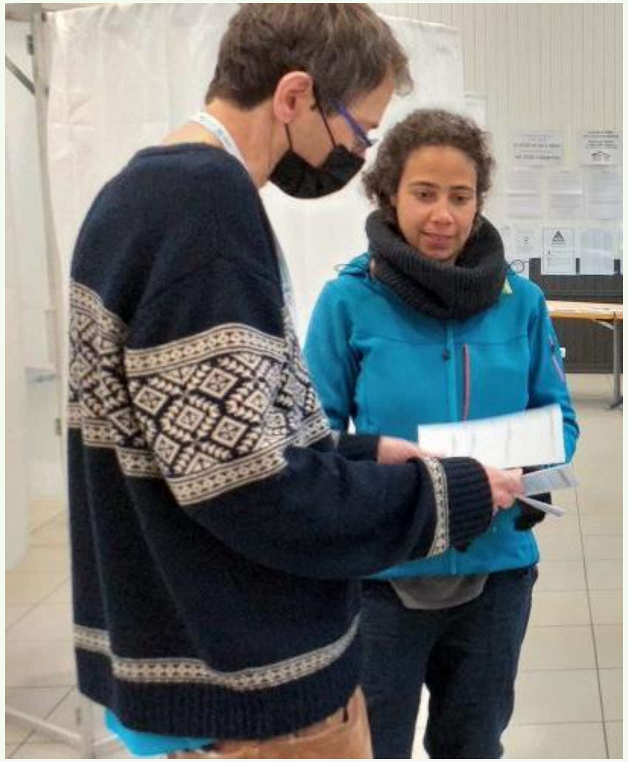


Fig. A researcher explaining how to vote with approval voting to a participant of the experiment.

Online experiment

Build a website (or use an existing one) to ask participants their preferences using a specific ballot format.

We generally use **real alternatives**, for instance ice cream flavors of political candidates.

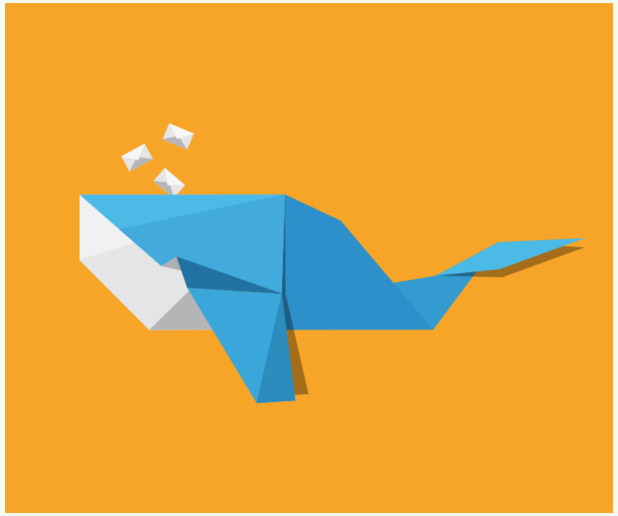
Easier to organize, you just need to share the link to enough people.

Remark: this gives a very unrepresentative sample of participants, so this should not be treated the same as a poll.

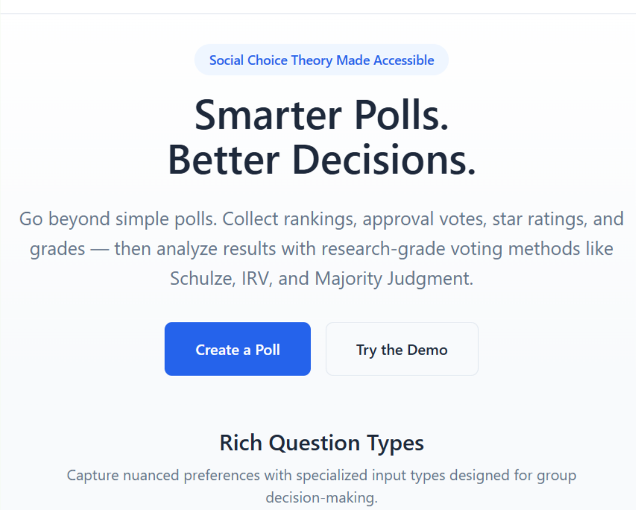
Examples (in French): vote.imag.fr (40k participants in 2017), vote.lamsade.fr (3k participants in 2024)

Tools for alternative voting rules

If you want to run a vote with voting rules based on rankings, approval ballots or some other ballot format, you can use **one of the following platforms**:



[Whale](#)
(Sylvain Bouveret)



Pref.tools/vote/
(Dominik Peters)

Statistical tests

To assess the statistical significance of your results, it can be useful to run a **statistical test**, such as a chi-squared test or a regression analysis.

Not very frequent in COMSOC because we are generally studying how voting rules/algorithms behave. Much more use by **behavioral economists and political scientists** to talk about voters' voting behaviors.