# Computational Social Choice: Spring 2017

Ulle Endriss
Institute for Logic, Language and Computation
University of Amsterdam

# Plan for Today

The broad aim for today is to show how we can *characterise* voting rules in terms of their properties. We review three approaches:

- *Axiomatic method:* to characterise a (family of) voting rule(s) as the only one satisfying certain axioms
- *Maximum likelihood estimation:* to characterise a voting rule as computing the most likely “correct” winner, given $n$ distorted copies of an objectively “correct” ranking (the ballots)
- *Distance-based rationalisation:* to characterise a voting rule in terms of a notion of *consensus* (profiles where outcomes are clear) and a notion of *distance* (from such a consensus profile)

Under the first approach we think of voting a a *compromise-seeking* activity (so we need to be fair, etc.). Under the second approach we think of voting as a *truth-finding* activity (e.g., amongst experts).

# Reminder: Formal Framework

Need to choose from a finite set $X = \{x_1, \ldots, x_m\}$ of *alternatives*.

Let $\mathcal{L}(X)$ denote the set of all strict linear orders on $X$. We use elements of $\mathcal{L}(X)$ to model (true) *preferences* and (declared) *ballots*.

Each member of a finite set $N = \{1, \ldots, n\}$ of *voters* supplies us with a ballot, giving rise to a *profile* $\boldsymbol{R} = (R_1, \ldots, R_n) \in \mathcal{L}(X)^n$.

A *voting rule* (or *social choice function*) for $N$ and $X$ selects one or more winners for every such profile:

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

If $|F(\boldsymbol{R})| = 1$ for all profiles $\boldsymbol{R} \in \mathcal{L}(X)^n$, then $F$ is called *resolute*.

If $F$ is resolute, we usually write $F(\boldsymbol{R}) = x^\star$ instead of $F(\boldsymbol{R}) = \{x^\star\}$.

Notation: Write $N^{\boldsymbol{R}}_{x \succ y} = \{i \in N \mid (x, y) \in R_i\}$ for the set of voters who rank alternative $x$ above alternative $y$ in profile $\boldsymbol{R}$.

## Axioms: Anonymity and Neutrality

Two basic fairness requirements for a voting rule $F$:

- $F$ is *anonymous* if $F(R_1, \ldots, R_n) = F(R_{\pi(1)}, \ldots, R_{\pi(n)})$ for any profile $(R_1, \ldots, R_n)$ and any permutation $\pi : N \to N$.
- $F$ is *neutral* if $F(\pi(\boldsymbol{R})) = \pi(F(\boldsymbol{R}))$ for any profile $\boldsymbol{R}$ and any permutation $\pi : X \to X$ (with $\pi$ extended to profiles and sets of alternatives in the natural manner).

Thus: A is symmetry w.r.t. voters. N is symmetry w.r.t. alternatives.

## Axiom: Positive Responsiveness

A (not necessarily resolute) voting rule satisfies *positive responsiveness* if, whenever some voter raises a (possibly tied) winner $x^\star$ in her ballot, then $x^\star$ will become the *unique* winner. Formally:

- $F$ is *positively responsive* if $x^\star \in F(\boldsymbol{R})$ implies $\{x^\star\} = F(\boldsymbol{R'})$ for any alternative $x^\star$ and any two *distinct* profiles $\boldsymbol{R}$ and $\boldsymbol{R'}$ with $N^{\boldsymbol{R}}_{x^\star \succ y} \subseteq N^{\boldsymbol{R}'}_{x^\star \succ y}$ and $N^{\boldsymbol{R}}_{y \succ z} = N^{\boldsymbol{R'}}_{y \succ z}$ for all $y, z \in X \setminus \{x^\star\}$.

Thus, this is a monotonicity requirement (we'll see others later on).

## May's Theorem

When there are only *two alternatives*, then all the voting rules we have seen coincide with the *simple majority rule*. Good news:

**Theorem 1 (May, 1952)** *A voting rule for two alternatives satisfies anonymity, neutrality, and positive responsiveness if and only if that rule is the simple majority rule.*

This provides a good justification for using this rule (arguing in favour of "majority" directly is harder than arguing for anonymity etc.).

K.O. May. A Set of Independent Necessary and Sufficient Conditions for Simple Majority Decisions. *Econometrica*, 20(4):680–684, 1952.

# Proof Sketch

Clearly, the simple majority rule satisfies all three properties. ✓

Now for the other direction:

Assume the number of voters is *odd* (other case: similar) $\leadsto$ no ties.

There are two possible ballots: $x \succ y$ and $y \succ x$.
Anonymity $\leadsto$ only *number of ballots* of each type matters.

Consider all possible profiles $\boldsymbol{R}$. Distinguish two cases:

- Whenever $|N^{\boldsymbol{R}}_{x \succ y}| = |N^{\boldsymbol{R}}_{y \succ x}| + 1$, then only $x$ wins.
  By *PR*, $x$ wins whenever $|N^{\boldsymbol{R}}_{x \succ y}| > |N^{\boldsymbol{R}}_{y \succ x}|$. By *neutrality*, $y$ wins otherwise. But this is just what the simple majority rule does. ✓

- There exist a profile $\boldsymbol{R}$ with $|N^{\boldsymbol{R}}_{x \succ y}| = |N^{\boldsymbol{R}}_{y \succ x}| + 1$, yet $y$ wins.
  Suppose one $x$-voter switches to $y$, yielding $\boldsymbol{R'}$. By *PR*, now only $y$ wins. But now $|N^{\boldsymbol{R'}}_{y \succ x}| = |N^{\boldsymbol{R'}}_{x \succ y}| + 1$, which is symmetric to the earlier situation, so by *neutrality* $x$ should win. Contradiction. ✓

# Young's Theorem

Another seminal result (which we won't discuss in detail here) is known as *Young's Theorem*. It provides a characterisation of the *PSR*'s. The core axiom is *reinforcement* (a.k.a. *consistency*):

- $F$ satisfies *reinforcement* if, whenever we split the electorate into two groups and some alternative were to win for both groups, then it will also win for the full electorate. More precisely:

$$F(\boldsymbol{R}) \cap F(\boldsymbol{R'}) \neq \emptyset \;\;\Rightarrow\;\; F(\boldsymbol{R} \oplus \boldsymbol{R'}) = F(\boldsymbol{R}) \cap F(\boldsymbol{R'})$$

Young showed that $F$ is a (*generalised*) *positional scoring rule* iff it satisfies *anonymity*, *neutrality*, *reinforcement*, and a technical condition known as *continuity*.

Here, "generalised" means that the scoring vector need not be decreasing.

H.P. Young. Social Choice Scoring Functions. *SIAM Journal on Applied Mathematics*, 28(4):824–838, 1975.

## Voting as Truth-Tracking

An alternative interpretation of "voting":

- There exists an objectively "correct" ranking of the alternatives.
- The voters want to identify the correct ranking (or winner), but cannot tell with certainty which ranking is correct. Their ballots reflect what they believe to be true.
- We want to estimate the most likely ranking (or winner), given the ballots we observe. Can we use a voting rule to do this?

## Example

Consider the following scenario:

- two alternatives: $x$ and $y$
- either $x \succ y$ or $y \succ x$ (we don't know which and have no priors)
- 20 voters/experts with probability 75% each of getting it right

Now suppose we observe that 12/20 voters say $x \succ y$.
What can we infer, given this observation (let's call it $E$)?

- Probability for $E$ to happen *in case $x \succ y$* is correct:

$$P(E \mid x \succ y) \quad = \quad \binom{20}{12} \cdot 0.75^{12} \cdot 0.25^{8}$$

- Probability for $E$ to happen *in case $y \succ x$* is correct:

$$P(E \mid y \succ x) \quad = \quad \binom{20}{8} \cdot 0.75^{8} \cdot 0.25^{12}$$

<u>Thus:</u> $P(E \mid x \succ y)/P(E \mid y \succ x) \ = \ 0.75^4/0.25^4 \ = \ 81.$
From Bayes and assuming uniform priors $[P(x \succ y) = P(y \succ x)]$:
Given $E$, $x$ being better is 81 times as likely as $y$ being better.

# The Condorcet Jury Theorem

For the case of two alternatives, the simple majority rule is the best choice also under the truth-tracking perspective:

**Theorem 2 (Condorcet, 1785)** *Suppose a jury of $n$ voters need to select the better of two alternatives and each voter independently makes the correct decision with the same probability $p > \frac{1}{2}$. Then the probability that the simple majority rule returns the correct decision increases monotonically in $n$ and approaches $1$ as $n$ goes to infinity.*

Proof sketch: By the law of large numbers, the number of voters making the correct choice approaches $p \cdot n > \frac{1}{2} \cdot n$. ✓

For modern expositions see Nitzan (2010) and Young (1995).

Writings of the Marquis de Condorcet. In I. McLean and A. Urken (eds.), *Classics of Social Choice*, University of Michigan Press, 1995.

S. Nitzan. *Collective Preference and Choice*. Cambridge University Press, 2010.

H.P. Young. Optimal Voting Rules. *J. Economic Perspectives*, 9(1):51–64, 1995.

## Characterising Voting Rules via Noise Models

For $n$ alternatives, Young (1995) shows that, if the probability of a voter to rank a given pair correctly is $p > \frac{1}{2}$, then the voting rule selecting the most likely winner coincides with the *Kemeny rule*.

Conitzer and Sandholm (2005) ask a general question:

- For a given voting rule $F$, can we design a "*noise model*" such that $F$ is a *maximum likelihood estimator* for the winner?

H.P. Young. Optimal Voting Rules. *J. Economic Perspectives*, 9(1):51–64, 1995.

V. Conitzer and T. Sandholm. Common Voting Rules as Maximum Likelihood Estimators. Proc. UAI-2005.

E. Elkind and A. Slinko. Rationalizations of Voting Rules. In F. Brandt et al. (eds.), *Handbook of Computational Social Choice*. CUP, 2016.

# The Borda Rule as a Maximum Likelihood Estimator

It *is* possible for the Borda rule:

**Proposition 3 (Conitzer and Sandholm, 2005)** *If each voter independently ranks the true winner at position $k$ with probability $\frac{2^{m-k}}{2^m-1}$, then the maximum likelihood estimator is the Borda rule.*

Proof: Let $r_i(x)$ be the position at which voter $i$ ranks alternative $x$.

Probability to observe the actual ballot profile if $x$ is the true winner:

$$\frac{\prod_{i \in N} 2^{m-r_i(x)}}{(2^m-1)^n} = \frac{2^{\sum_{i \in N} m-r_i(x)}}{(2^m-1)^n} = \frac{2^{\mathrm{BordaScore}(x)}}{(2^m-1)^n}$$

Hence, $x$ has maximal likelihood of being the true winner iff $x$ has a maximal Borda score. ✓

V. Conitzer and T. Sandholm. Common Voting Rules as Maximum Likelihood Estimators. Proc. UAI-2005.

# Characterisation via Consensus and Distance

Recall: Rules such as The *Dodgson* and *Young* compute the "closest" profile with a Condorcet winner and then elect that Condorcet winner.

This suggests a general method for defining a voting rule:

- Fix a class of *consensus profiles:* profiles in which there is a clear (set of) winner(s). (And specify *who* wins.)
- Fix a metric to measure the *distance* between two profiles.
- This induces a *voting rule:* for a given profile, find the closest consensus profile(s) and elect the corresponding winner(s).

T. Meskanen and H. Nurmi. Closeness Counts in Social Choice. In M. Braham and F. Steffen (eds.), *Power, Freedom, and Voting*, Springer-Verlag, 2008.

E. Elkind and A. Slinko. Rationalizations of Voting Rules. In F. Brandt et al. (eds.), *Handbook of Computational Social Choice*. CUP, 2016.

## Notions of Consensus

Four natural definitions for what constitutes a consensus profile:

- *Condorcet Winner:* there exists a Condorcet winner $x$ ($\rightsquigarrow$ $x$ wins)
- *Majority Winner:* there exists an alternative $x$ that is ranked first by an absolute majority of the voters ($\rightsquigarrow$ $x$ wins)
- *Unanimous Winner:* there exists an alternative $x$ that is ranked first by all voters ($\rightsquigarrow$ $x$ wins)
- *Unanimous Ranking:* all voters report exactly the same ranking ($\rightsquigarrow$ the top alternative in that unanimous ranking wins)

(Other definitions are possible.)

## Ways of Measuring Distance

Two natural definitions of distance between profiles $\boldsymbol{R}$ and $\boldsymbol{R'}$:

- *Swap distance:* minimal number of pairs of adjacent alternatives that need to get swapped to get from $\boldsymbol{R}$ to $\boldsymbol{R'}$.

  Equivalently: distance between two ballots = number of pairs of alternatives with distinct relative ranking (so-called *Kendall tau distance*); sum over voters to get distance between two profiles.

$$\frac{1}{2} \cdot \sum_{i \in N} \#\{(x,y) \in X^2 \mid \mathbb{1}_{i \in N^{\boldsymbol{R}}_{x \succ y}} \neq \mathbb{1}_{i \in N^{\boldsymbol{R'}}_{x \succ y}}\}$$

- *Discrete distance:* distance between two ballots is $0$ if they are the same and $1$ otherwise; sum over voters to get profile distance.

$$\#\{i \in N \mid R_i \neq R'_i\}$$

(Other definitions are possible.)

## Examples

Two voting rules for which the standard definition is already formulated in terms of consensus and distance:

- Dodgson Rule = Condorcet Winner + Swap Distance
- Kemeny Rule = Unanimous Ranking + Swap Distance

How about other rules? Borda? Plurality?

Writings of C.L. Dodgson. In I. McLean and A. Urken (eds.), *Classics of Social Choice*, University of Michigan Press, 1995.

J. Kemeny. Mathematics without Numbers. *Daedalus*, 88:571–591, 1959.

# Characterisation of the Borda Rule

Recall: the Borda rule is the PSR with vector $(m{-}1, m{-}2, \ldots, 0)$.

**Proposition 4 (Farkas and Nitzan, 1979)** *Borda is characterised by the unanimous winner consensus criterion and the swap distance.*

Proof sketch: The swap distance between a given ballot that ranks $x$ at position $k$ and the closest ballot that ranks $x$ at the top is $k{-}1$. Thus, if voter $i$ ranks $x$ at position $k$, she gives $-(k{-}1)$ points to $x$. This corresponds to the PSR with vector $(0, -1, -2, \ldots, -(m{-}1))$, which is equivalent to the Borda rule. ✓

Remark: So Dodgson, Kemeny, and Borda are all *rationalisable* via the same notion of distance!

D. Farkas and S. Nitzan. The Borda Rule and Pareto Stability: A Comment. *Econometrica*, 47(5):1305–1306, 1979.

## Characterisation of the Plurality Rule

Recall: the plurality rule is the PSR with scoring vector $(1, 0, \ldots, 0)$.

**Proposition 5 (Nitzan, 1981)** *Plurality is characterised by the unanimous winner consensus criterion and the discrete distance.*

Proof: Immediate. ✓

Remark 1: to be precise, Nitzan used a slightly different distance

Remark 2: also works with Majority Winner + discrete distance, but doesn't work with Condorcet Winner or Unanimous Ranking

S. Nitzan. Some Measures of Closeness to Unanimity and their Implications. *Theory and Decision*, 13(2):129–138, 1981.

## Summary

We have seen three approaches to characterising a voting rule:

- as the only rule satisfying certain *axioms*;
- as returning *the most likely "true" winner*, given the noisy signals the voters have received about the "true" ranking; and
- as computing the closest *consensus* profile (w.r.t. some *distance*) with a clear winner and returning that winner.

All three approaches are (potentially) useful

- to better understand particular voting rules;
- to explain why there are so many "natural" voting rules; and
- to help prove general results about families of voting rules.

**What next?** More applications of the axiomatic method.