Computational Social Choice: Spring 2019

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

Plan for Today

The second scenario of social choice we review in this course is that of *voting*, and *preference aggregation* more generally.

Today we start by defining the *model* and many different *voting rules*.

We then discuss examples for two types of *characterisation results:*

- *normative characterisation:* using the axiomatic method to single out voting rules that satisfy certain intuitively appealing properties
- *epistemic characterisation:* using probabilistic modelling to single out voting rules most likely to recover the "truly" best alternative

For full details see Zwicker (2016) and Elkind and Slinko (2016).

W.S. Zwicker. Introduction to the Theory of Voting. In F. Brandt et al. (eds.), *Handbook of Computational Social Choice*. CUP, 2016.

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

Three Voting Rules

Suppose n *voters* choose from a set of m *alternatives* by stating their preferences in the form of *linear orders* over the alternatives.

Here are three *voting rules* (there are many more):

- *Plurality:* elect the alternative ranked first most often (i.e., each voter assigns 1 point to an alternative of her choice, and the alternative receiving the most points wins)
- *Plurality with runoff :* run a plurality election and retain the two front-runners; then run a majority contest between them
- *Borda:* each voter gives $m-1$ points to the alternative she ranks first, $m-2$ to the alternative she ranks second, etc.; and the alternative with the most points wins

Example: Choosing a Beverage for Lunch

Consider this election, with nine *voters* having to choose from three *alternatives* (namely what beverage to order for a common lunch):

2 Germans:	Beer	$\succ$	Wine	$\succ$	Milk
3 Frenchmen:	Wine	$\succ$	Beer	$\succ$	Milk
4 Dutchmen:	Milk	$\succ$	Beer	$\succ$	Wine

Which beverage *wins* the election for

- the plurality rule?
- plurality with runoff?
- the Borda rule?

The Model

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

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

Each member i of a finite set $N = \{1, \ldots, n\}$ of *voters* supplies us with a ballot $\succ_i$, giving rise to a *profile* $\boldsymbol{\succ} = (\succ_1, \ldots, \succ_n) \in \mathcal{L}(A)^n$ (sometimes denoted $\boldsymbol{R} = (R_1, \ldots, R_n) \in \mathcal{L}(A)^n$).

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

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

If $|F(\boldsymbol{\succ})| = 1$ for all profiles $\boldsymbol{\succ}$, then F is called *resolute*.

Most natural voting rules are *irresolute* and have to be paired with a *tie-breaking rule* to always select a unique election winner.

<u>Examples:</u> random tie-breaking, lexicographic tie-breaking

The Condorcet Principle

An alternative that beats every other alternative in pairwise majority contests is called a *Condorcet winner*. Sometimes there is no CW:

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

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

Voter 3: $c \succ a \succ b$

This is the famous *Condorcet Paradox* (the majority relation is cyclic).

The *Condorcet Principle* says that, *if it exists*, only the CW should win. Voting rules that satisfy this principle are called *Condorcet extensions*.

The *Borda* rule is a voting rule that violates the Condorcet Principle:

3 voters: $c \succ b \succ a$

2 voters: $b \succ a \succ c$

M.J.A.N. de Caritat (Marquis de Condorcet). *Essai sur l'application de l'analyse à la probabilté des décisions rendues a la pluralité des voix*. Paris, 1785.

Positional Scoring Rules

We can generalise the idea underlying the Borda rule as follows:

A *positional scoring rule* (PSR) is defined by a so-called *scoring vector* $\boldsymbol{s} = (s_1, \ldots, s_m) \in \mathbb{R}^m$ with $s_1 \geqslant s_2 \geqslant \cdots \geqslant s_m$ and $s_1 > s_m$.

Each voter submits a ranking of the m alternatives. Each alternative receives s_i points for every voter putting it at the ith position.

The alternative(s) with the highest score (sum of points) win(s).

Examples:

- *Borda rule* = PSR with scoring vector $(m-1, m-2, \ldots, 0)$
- *Plurality rule* = PSR with scoring vector $(1, 0, \ldots, 0)$
- *Antiplurality* (or *veto*) *rule* = PSR with scoring vector $(1, \ldots, 1, 0)$
- For any $k < m$, *k-approval* = PSR with $(\underbrace{1, \ldots, 1}_{k}, 0, \ldots, 0)$

Positional Scoring Rules and the Condorcet Principle

Consider this example with three alternatives and seven voters:

3 voters: $a \succ b \succ c$
2 voters: $b \succ c \succ a$
1 voter: $b \succ a \succ c$
1 voter: $c \succ a \succ b$

So a is the *Condorcet winner:* a beats both b and c (with 4 out of 7).
But any *positional scoring rule* makes b win (because $s_1 \geqslant s_2 \geqslant s_3$):

a: $3 \cdot s_1 + 2 \cdot s_2 + 2 \cdot s_3$
b: $3 \cdot s_1 + 3 \cdot s_2 + 1 \cdot s_3$
c: $1 \cdot s_1 + 2 \cdot s_2 + 4 \cdot s_3$

Thus, *no positional scoring rule* for three (or more) alternatives can possibly satisfy the *Condorcet Principle*.

Copeland Rule and Majority Graph

Under the *Copeland rule* an alternative gets $+1$ point for every pairwise majority contest won and -1 point for every such contest lost.

Exercise: *Show that the Copeland rule is a Condorcet extension.*

Copeland winners can be computed from the *majority graph* (with an edge from x to y whenever x beats y in a pairwise majority contest).

Exercise: *How can you characterise the Condorcet winner (if it exists) in graph-theoretical terms in a given majority graph?*

A.H. Copeland. *A "Reasonable" Social Welfare Function*. Seminar on Mathematics in Social Sciences, University of Michigan, 1951.

F. Brandt, M. Brill, and P. Harrenstein. Tournament Solutions. In F. Brandt et al. (eds.), *Handbook of Computational Social Choice*. CUP, 2016.

Kemeny Rule and Weighted Majority Graph

Under the *Kemeny rule* an alternative wins if it is maximal in a ranking that minimises the sum of pairwise disagreements with the ballots:

(1) For every possible ranking $\succ$, count the number of triples (i, x, y) s.t. $\succ$ disagrees with voter i on the ranking of alternatives x and y.

(2) Find all rankings $\succ$ that have a minimal score in the above sense.

(3) Elect any alternative that is maximal in such a "closest" ranking.

Exercise: *Show that the Kemeny rule is a Condorcet extension.*

The Kemeny rule needs more information than just the majority graph. But it can be computed from the *weighted majority graph*.

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

F. Fischer, O. Hudry, and R. Niedermeier. Weighted Tournament Solutions. In F. Brandt et al. (eds.), *Handbook of Computational Social Choice*. CUP, 2016.

More Voting Rules

Here are a few more voting rules (still not all there is!):

- *STV (single transferable vote):* Keep eliminating plurality loser.
- *Slater:* Find ranking that minimises number of edges in majority graph we'd have to switch. Elect top alternative in that ranking.
- *Ranked-Pairs:* Build a full ranking by "locking in" ordered pairs in order of majority strength (but avoid cycles). Elect top alternative.
- *Young:* Elect alternative x that minimises the number of voters we need to remove before x becomes the Condorcet winner.
- *Dodgson:* Elect alternative x that minimises the number of swaps of adjacent alternatives in the profile we need to perform before x becomes the Condorcet winner. (Note difference to Kemeny!)

Trivia: Dodgson is also known as Lewis Carroll (*Alice in Wonderland*).

Computational Complexity of Winner Determination

We can also classify voting rules according to the computational complexity of the *winner determination problem*. Omitting details:

Proposition 1 *For any positional scoring rule, the election winners can be computed in polynomial time.*

Theorem 2 (Brill and Fischer, 2012) *Deciding whether a given alternative is a winner under the ranked-pairs rule is NP-complete.*

Theorem 3 (Hemaspaandra et al., 1997) *Deciding whether a given alternative is a Dodgson winner is complete for parallel access to NP.*

M. Brill and F. Fischer. The Price of Neutrality for the Ranked Pairs Method. AAAI-2012.

E. Hemaspaandra, L.A. Hemaspaandra, and J. Rothe. Exact Analysis of Dodgson Elections. *Journal of the ACM*, 44(6):806–825, 1997.

Nonstandard Ballots

We defined voting rules over profiles of strict linear orders (even if some rules, e.g., plurality, don't use all information). Other options:

- *Approval voting:* You can approve of any subset of the alternatives. The alternative with the most approvals wins.
- *Even-and-equal cumulative voting:* You vote as for AV, but 1 point gets split evenly amongst the alternatives you approve.
- *Range voting:* You vote by dividing 100 points amongst the alternatives as you see fit (as long every share is an integer).
- *Majority judgment:* You award a grade to each of the alternatives ("excellent", "good", etc.). Highest median grade wins.

The most important of these is approval voting.

Remark: *k-approval* and *approval voting* are very different rules!

Normative Characterisation: The Axiomatic Method

So many different voting rules! *How do you choose?*

One approach is to use the *axiomatic method* to identify voting rules of *normative* appeal. We will see one example for a classical result.

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 $\pi : A \to A$ (with π 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

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 $\boldsymbol{R}$.

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 $\boldsymbol{R'}$ s.t. $N^{\boldsymbol{R}}_{x^\star \succ y} \subseteq N^{\boldsymbol{R}'}_{x^\star \succ y}$ and $N^{\boldsymbol{R}}_{y \succ z} = N^{\boldsymbol{R'}}_{y \succ z}$ for all $y, z \in A \setminus \{x^\star\}$.

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 4 (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) $\rightsquigarrow$ no ties.

There are two possible ballots: $a \succ b$ and $b \succ a$.
Anonymity $\rightsquigarrow$ only *number of ballots* of each type matters.

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

- Whenever $|N^{\boldsymbol{R}}_{a \succ b}| = |N^{\boldsymbol{R}}_{b \succ a}| + 1$, then only a wins.
  By *PR*, a wins whenever $|N^{\boldsymbol{R}}_{a \succ b}| > |N^{\boldsymbol{R}}_{b \succ a}|$. By *neutrality*, b wins otherwise. But this is just what the simple majority rule does. ✓
- There exist a profile $\boldsymbol{R}$ with $|N^{\boldsymbol{R}}_{a \succ b}| = |N^{\boldsymbol{R}}_{b \succ a}| + 1$, yet b wins.
  Suppose one a-voter switches to b, yielding $\boldsymbol{R'}$. By *PR*, now only b wins. But now $|N^{\boldsymbol{R'}}_{b \succ a}| = |N^{\boldsymbol{R'}}_{a \succ b}| + 1$, which is symmetric to the earlier situation, so by *neutrality* a should win. Contradiction. ✓

Epistemic Characterisation: 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: Ballots as Noisy Signals

Consider the following scenario:

- two alternatives: a and b
- either $a \succ b$ or $b \succ a$ (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 $a \succ b$.
What can we infer, given this observation (let's call it E)?

- Probability for E to happen *in case $a \succ b$* is correct:

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

- Probability for E to happen *in case $b \succ a$* is correct:

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

<u>Thus:</u> $P(E \mid a \succ b) \,/\, P(E \mid b \succ a) \;=\; 0.75^4/0.25^4 \;=\; 81.$
From Bayes and assuming uniform priors $[P(a \succ b) = P(b \succ a)]$:
Given E, a being better is 81 times as likely as b 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 5 (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. 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 6 (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. UAI-2005.

Summary

We have introduced a large number of voting rules:

- *Positional scoring rules:* Borda, plurality, antiplurality, k-approval
- *Condorcet extensions:* Copeland, Slater, Kemeny, ...
- And many more: plurality-with-runoff, STV, approval, ...

They differ in their *intuitive appeal*, the amount of *information* they require (e.g., majority graph), and their *computational complexity*.

We have then seen two approaches to characterising a voting rule:

- as the only rule satisfying certain *axioms* (normative desiderata)
- as optimally *tracking the truth* (under some probabilistic model)

What next? More applications of the axiomatic method.