A Crash Course on Computational Social Choice and Fair Division

Vincent Conitzer, Duke University
ACM Conference on Economics and Computation (EC), 2018

comsoc mailing list: https://lists.duke.edu/sympa/subscribe/comsoc
Lirong Xia  
(Ph.D. 2011, now at RPI)

Markus Brill  
(postdoc 2013-2015, now at TU Berlin)

Rupert Freeman  
(Ph.D. 2018, joining MSR NYC for postdoc)
Voting

$n$ voters… … each produce a ranking of $m$ alternatives…

\[ b > a > c \]

\[ a > c > b \]

… which a social preference function (SPF) maps to one or more aggregate rankings.

\[ a > b > c \]

… or, a social choice function (SCF) just produces one or more winners.

\[ a \]
Plurality

\[
\begin{array}{ccc}
1 & 0 & 0 \\
2 & 1 & 0 \\
\end{array}
\]

\[
b > a > c \\
a > c > b \\
a > b > c
\]
Borda

\[
\begin{align*}
2 & \quad 1 & \quad 0 \\
\end{align*}
\]

\[
\begin{align*}
b & > a & > c \\
\end{align*}
\]

\[
\begin{align*}
a & > c & > b \\
\end{align*}
\]

\[
\begin{align*}
a & > b & > c \\
\end{align*}
\]
Instant runoff voting / single transferable vote (STV)

\[ b > a > c \]

\[ a > b > b \]

\[ a > b > c \]
Kemeny

- Natural interpretation as maximum likelihood estimate of the “correct” ranking [Young 1988, 1995]

- 2 disagreements ↔ 3*3 - 2 = 7 agreements (maximum)
Kemeny on pairwise election graphs

- Final ranking = acyclic tournament graph
  - Edge (a, b) means a ranked above b
  - Acyclic = no cycles, tournament = edge between every pair
- Kemeny ranking seeks to minimize the total weight of the inverted edges

pairwise election graph

Kemeny ranking

NP-hard even with 4 voters [Dwork et al. 2001]
Integer programs scale reasonably [C., Davenport, Kalagnanam 2006]
Ranking Ph.D. applicants
(briefly described in C. [2010])

• Input: Rankings of subsets of the (non-eliminated) applicants

• Output: (one) Kemeny ranking of the (non-eliminated) applicants
Choosing a rule

• How do we choose a rule from all of these rules?
• How do we know that there does not exist another, “perfect” rule?
• Let us look at some criteria that we would like our voting rule to satisfy
Condorcet criterion

- A candidate is the Condorcet winner if it wins all of its pairwise elections
- Does not always exist…
- … but the Condorcet criterion says that if it does exist, it should win

- Many rules do not satisfy this
- E.g., for plurality:
  - b > a > c > d
  - c > a > b > d
  - d > a > b > c
- a is the Condorcet winner, but it does not win under plurality
An SPF f is said to be consistent if the following holds:
- Suppose $V_1$ and $V_2$ are two voting profiles (multisets) such that f produces the same ranking on both
- Then f should produce the same ranking on their union.

Which of our rules satisfy this?
Consistency (SCF sense)

- An SCF $f$ is said to be consistent if the following holds:
  - Suppose $V_1$ and $V_2$ are two voting profiles (multisets) such that $f$ produces the same winner on both.
  - Then $f$ should produce the same winner on their union.

- Which of our rules satisfy this?

- Consistency properties are closely related to interpretability as MLE of the truth [C., Rognlie, Xia 2009]
Some axiomatizations

• **Theorem [Young 1975]**. An SCF is symmetric, consistent, and continuous if and only if it is a positional scoring rule.

• **Theorem [Young and Levenglick 1978]**. An SPF is neutral, consistent, and Condorcet if and only if it is the Kemeny SPF.

• **Theorem [Freeman, Brill, C. 2014]**. An SPF satisfies independence of bottom alternatives, consistency at the bottom, independence of clones (& some minor conditions) if and only if it is the STV SPF.
Manipulability

• Sometimes, a voter is better off revealing her preferences insincerely, AKA manipulating

• E.g., plurality
  – Suppose a voter prefers $a > b > c$
  – Also suppose she knows that the other votes are
    • 2 times $b > c > a$
    • 2 times $c > a > b$
  – Voting truthfully will lead to a tie between $b$ and $c$
  – She would be better off voting, e.g., $b > a > c$, guaranteeing $b$ wins
Gibbard-Satterthwaite impossibility theorem

• Suppose there are at least 3 alternatives
• There exists no rule that is simultaneously:
  – non-imposing/onto (for every alternative, there are some votes that would make that alternative win),
  – nondictatorial (there does not exist a voter such that the rule simply always selects that voter’s first-ranked alternative as the winner), and
  – nonmanipulable/strategy-proof
Single-peaked preferences

• Suppose candidates are ordered on a line
• Every voter prefers candidates that are closer to her most preferred candidate
• Let every voter report only her most preferred candidate ("peak")
• Choose the median voter’s peak as the winner
  – This will also be the Condorcet winner
• Nonmanipulable!

Impossibility results do not necessarily hold when the space of preferences is restricted
Moulin’s characterization

- Slight generalization: add **phantom** voters, then choose the median of real+phantom voters
- **Theorem** [Moulin 1980]. Under single-peaked preferences, an SCF is strategy-proof, Pareto efficient, and anonymous if and only if it is such a generalized median rule.

```
  v_5  v_4
  p_1  v_2
  a_1  a_2
      p_2
  a_3  v_1
      p_4
      v_3
  a_4  a_5
```
Computational hardness as a barrier to manipulation

- A (successful) manipulation is a way of misreporting one’s preferences that leads to a better result for oneself
- Gibbard-Satterthwaite only tells us that for some instances, successful manipulations exist
- It does not say that these manipulations are always easy to find
- Do voting rules exist for which manipulations are computationally hard to find?
A formal computational problem

• The simplest version of the manipulation problem:
• **CONSTRUCTIVE-MANIPULATION:**
  – We are given a voting rule \( r \), the (unweighted) votes of the other voters, and an alternative \( p \).
  – We are asked if we can cast our (single) vote to make \( p \) win.
• E.g., for the Borda rule:
  – Voter 1 votes \( A > B > C \)
  – Voter 2 votes \( B > A > C \)
  – Voter 3 votes \( C > A > B \)
• Borda scores are now: \( A: 4, B: 3, C: 2 \)
• Can we make \( B \) win?
• Answer: YES. Vote \( B > C > A \) (Borda scores: \( A: 4, B: 5, C: 3 \))
Early research

• Theorem. CONSTRUCTIVE-MANIPULATION is NP-complete for the second-order Copeland rule. [Bartholdi, Tovey, Trick 1989]
  – Second order Copeland = alternative’s score is sum of Copeland scores of alternatives it defeats

• Theorem. CONSTRUCTIVE-MANIPULATION is NP-complete for the STV rule. [Bartholdi, Orlin 1991]

• Most other rules are easy to manipulate (in P)
Ranked pairs rule [Tideman 1987]

- Order pairwise elections by decreasing strength of victory
- Successively “lock in” results of pairwise elections unless it causes a cycle

```
Final ranking: c > a > b > d
```

- **Theorem.** CONSTRUCTIVE-MANIPULATION is NP-complete for the ranked pairs rule [Xia et al. IJCAI 2009]
Many manipulation problems…

Table from: C. & Walsh, Barriers to Manipulation, Chapter 6 in *Handbook of Computational Social Choice*

<table>
<thead>
<tr>
<th># alternatives</th>
<th># manipulators</th>
<th>unweighted votes, constructive manipulation</th>
<th>weighted votes, constructive</th>
<th>weighted votes, destructive</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>plurality</td>
<td>1</td>
<td>P</td>
<td>P</td>
<td>P</td>
</tr>
<tr>
<td>plurality with runoff</td>
<td>≥2</td>
<td>P</td>
<td>P</td>
<td>NP-c</td>
</tr>
<tr>
<td>veto</td>
<td>1</td>
<td>P</td>
<td>NP-c</td>
<td>NP-c</td>
</tr>
<tr>
<td>cup</td>
<td>1</td>
<td>P</td>
<td>NP-c</td>
<td>NP-c</td>
</tr>
<tr>
<td>Copeland</td>
<td>1</td>
<td>P</td>
<td>NP-c</td>
<td>NP-c</td>
</tr>
<tr>
<td>Borda</td>
<td>1</td>
<td>NP-c</td>
<td>P</td>
<td>NP-c</td>
</tr>
<tr>
<td>Nanson</td>
<td>1</td>
<td>NP-c</td>
<td>P</td>
<td>NP-c</td>
</tr>
<tr>
<td>Baldwin</td>
<td>1</td>
<td>NP-c</td>
<td>P</td>
<td>NP-c</td>
</tr>
<tr>
<td>Black</td>
<td>1</td>
<td>NP-c</td>
<td>P</td>
<td>NP-c</td>
</tr>
<tr>
<td>STV</td>
<td>1</td>
<td>NP-c</td>
<td>P</td>
<td>NP-c</td>
</tr>
<tr>
<td>maximin</td>
<td>1</td>
<td>NP-c</td>
<td>P</td>
<td>NP-c</td>
</tr>
<tr>
<td>Bucklin</td>
<td>1</td>
<td>P</td>
<td>NP-c</td>
<td>NP-c</td>
</tr>
<tr>
<td>fallback</td>
<td>1</td>
<td>P</td>
<td>P</td>
<td>NP-c</td>
</tr>
<tr>
<td>ranked pairs</td>
<td>1</td>
<td>NP-c</td>
<td>P</td>
<td>P</td>
</tr>
<tr>
<td>Schulze</td>
<td>1</td>
<td>P</td>
<td>P</td>
<td>P</td>
</tr>
</tbody>
</table>
STV manipulation algorithm

[C., Sandholm, Lang JACM 2007]

- rescue d
  - c eliminated
    - no choice for manipulator
      - b eliminated
        - no choice for manipulator
          - d eliminated
            - rescue a
              - don’t rescue a
            - don’t rescue a
              - …
    - d eliminated
      - rescue a
        - …
      - don’t rescue a
        - …
  - don’t rescue d
    - …

Runs in $O(((1+\sqrt{5})/2)^m)$ time (worst case)
Runtime on random votes [Walsh 2011]
Fine – how about another rule?

- **Heuristic algorithms and/or experimental (simulation) evaluation**

- **Quantitative versions of Gibbard-Satterthwaite** showing that under certain conditions, for some voter, even a random manipulation on a random instance has significant probability of succeeding
  [Friedgut, Kalai, Nisan 2008; Xia & C. 2008; Dobzinski & Procaccia 2008; Isaksson, Kindler, Mossel 2010; Mossel & Racz 2013]

  "for a social choice function $f$ on $k \geq 3$ alternatives and $n$ voters, which is $\epsilon$-far from the family of nonmanipulable functions, a uniformly chosen voter profile is manipulable with probability at least inverse polynomial in $n$, $k$, and $\epsilon^{-1}$."
Just a bit about fair allocation of resources

• Suppose we have $m$ items and $n$ agents
• Agent $i$ values item $j$ at $v_{ij}$ (additive valuations)
• Who should receive what? (no payments!)
• One solution: $\max \sum_i \sum_j v_{ij} x_{ij}$
• Downsides?
• Better: $\max$ Nash welfare, $\max \prod_i (\sum_j v_{ij} x_{ij})$
• Does it matter if items are divisible?
Eisenberg-Gale convex program

- Max $\Sigma_i \log u_i$
- subject to:
  - for all $i$, $u_i = \Sigma_j v_{ij} x_{ij}$
  - for all $j$, $\Sigma_i x_{ij} \leq 1$
  - for all $i$ and $j$, $x_{ij} \geq 0$

- Finding the optimal integer solution (indivisible items) is NP-hard [Ramezani and Endriss 2010], can be approximated efficiently in a sense [Cole and Gkatzelis 2015]
Competitive equilibrium from equal incomes (CEEI)

Nash welfare: $4 \times 40 \times 2 = 320$

Note: $(4-10\varepsilon) \times 40 \times (2+5\varepsilon) = 320 - 2000\varepsilon^2$
Nice properties of the max Nash welfare solution

- With divisible items, it constitutes a competitive equilibrium from equal incomes!
  - Follows from KKT conditions on convex program
  - Instant corollaries: envy-free, proportional

- With indivisible items:
  - envy-free up to one good [Caragiannis et al. 2016]
  - proportional up to one good (can be generalized to public decisions) [C., Freeman, Shah 2017]