Condorcet’s Principle and the Preference Reversal Paradox

Dominik Peters
Department of Computer Science
University of Oxford
Under many voting rules, it is sometimes better to report the *exact opposite* of your true preference ordering.

<table>
<thead>
<tr>
<th>3</th>
<th>3</th>
<th>4</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>b</td>
<td>c</td>
<td>d</td>
<td>d</td>
</tr>
<tr>
<td>b</td>
<td>d</td>
<td>a</td>
<td>c</td>
<td>c</td>
</tr>
<tr>
<td>d</td>
<td>c</td>
<td>b</td>
<td>a</td>
<td>b</td>
</tr>
<tr>
<td>c</td>
<td>a</td>
<td>d</td>
<td>b</td>
<td>a</td>
</tr>
</tbody>
</table>

*d* is Kemeny winner
The Paradox

- Under many voting rules, it is sometimes better to report the *exact opposite* of your true preference ordering.

<table>
<thead>
<tr>
<th></th>
<th>3</th>
<th>3</th>
<th>4</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>b</td>
<td>c</td>
<td>d</td>
<td>d</td>
<td></td>
</tr>
<tr>
<td>b</td>
<td>d</td>
<td>a</td>
<td>c</td>
<td>c</td>
<td></td>
</tr>
<tr>
<td>d</td>
<td>c</td>
<td>b</td>
<td>a</td>
<td>b</td>
<td></td>
</tr>
<tr>
<td>c</td>
<td>a</td>
<td>d</td>
<td>b</td>
<td>a</td>
<td></td>
</tr>
</tbody>
</table>

$d$ is Kemeny winner

<table>
<thead>
<tr>
<th></th>
<th>3</th>
<th>3</th>
<th>3</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>b</td>
<td>c</td>
<td>d</td>
<td>d</td>
<td>d</td>
<td>d</td>
</tr>
<tr>
<td>b</td>
<td>d</td>
<td>a</td>
<td>b</td>
<td>c</td>
<td>c</td>
<td>c</td>
</tr>
<tr>
<td>d</td>
<td>c</td>
<td>b</td>
<td>a</td>
<td>a</td>
<td>b</td>
<td>d</td>
</tr>
<tr>
<td>c</td>
<td>a</td>
<td>d</td>
<td>c</td>
<td>b</td>
<td>a</td>
<td></td>
</tr>
</tbody>
</table>
The Paradox

- Under many voting rules, it is sometimes better to report the *exact opposite* of your true preference ordering.

<table>
<thead>
<tr>
<th></th>
<th>3</th>
<th>3</th>
<th>4</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>b</td>
<td>c</td>
<td>d</td>
<td>d</td>
<td></td>
</tr>
<tr>
<td>b</td>
<td>d</td>
<td>a</td>
<td>c</td>
<td>c</td>
<td></td>
</tr>
<tr>
<td>d</td>
<td>c</td>
<td>b</td>
<td>a</td>
<td>b</td>
<td></td>
</tr>
<tr>
<td>c</td>
<td>a</td>
<td>d</td>
<td>b</td>
<td>a</td>
<td></td>
</tr>
</tbody>
</table>

*d is Kemeny winner*

<table>
<thead>
<tr>
<th></th>
<th>3</th>
<th>3</th>
<th>3</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>b</td>
<td>c</td>
<td>d</td>
<td>d</td>
<td>d</td>
<td>d</td>
</tr>
<tr>
<td>b</td>
<td>d</td>
<td>a</td>
<td>b</td>
<td>c</td>
<td>c</td>
<td></td>
</tr>
<tr>
<td>d</td>
<td>c</td>
<td>b</td>
<td>a</td>
<td>a</td>
<td>b</td>
<td></td>
</tr>
<tr>
<td>c</td>
<td>a</td>
<td>d</td>
<td>c</td>
<td>b</td>
<td>a</td>
<td></td>
</tr>
</tbody>
</table>

*b is Kemeny winner*
The Paradox

- Under many voting rules, it is sometimes better to report the exact opposite of your true preference ordering.

<table>
<thead>
<tr>
<th>3</th>
<th>3</th>
<th>4</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>b</td>
<td>c</td>
<td>d</td>
<td>d</td>
</tr>
<tr>
<td>b</td>
<td>d</td>
<td>a</td>
<td>c</td>
<td>c</td>
</tr>
<tr>
<td>d</td>
<td>c</td>
<td>b</td>
<td>a</td>
<td>b</td>
</tr>
<tr>
<td>c</td>
<td>a</td>
<td>d</td>
<td>b</td>
<td>a</td>
</tr>
</tbody>
</table>

*d is Kemeny winner*

<table>
<thead>
<tr>
<th>3</th>
<th>3</th>
<th>3</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>b</td>
<td>c</td>
<td>d</td>
<td>d</td>
<td>d</td>
</tr>
<tr>
<td>b</td>
<td>d</td>
<td>a</td>
<td>b</td>
<td>c</td>
<td>c</td>
</tr>
<tr>
<td>d</td>
<td>c</td>
<td>b</td>
<td>a</td>
<td>a</td>
<td>b</td>
</tr>
<tr>
<td>c</td>
<td>a</td>
<td>d</td>
<td>c</td>
<td>b</td>
<td>a</td>
</tr>
</tbody>
</table>

*b is Kemeny winner*
Half-way monotonicity

- A voting rule satisfies **half-way monotonicity** if it avoids the paradox.
- => always weakly better to report truth rather than reverse.
- Introduced by Sanver and Zwicker (IJGT 2009).

**Good rules:**
- Borda
- Plurality
- Scoring rules

**Bad rules:**
- STV/Baldwin/Coombs
- Kemeny/maximin/Copeland/Schulze/Ranked Pairs
- Condorcet extensions i.e. rules that select the Condorcet winner if it exists
Sanver and Zwicker (2009): No Condorcet extension satisfies half-way monotonicity if there are $m \geq 4$ alternatives and sufficiently many voters.

Proof needs $n = 702 + \Omega(m!)$ voters, and $n$ even.

Unfortunately, their proof contains an arithmetical mistake, non-trivial to fix.

Is the paradox a problem in small electorates?

We show the result holds for $n = 15$. 
Proof.

Preference Reversal Paradox

Fig. 1. Proof diagram of the proof of Theorem 5.3.

Fig. 2. Proof diagram of the proof of Theorem 5.4.
Comparison to Participation

- Voting rule satisfies **participation** if it is always weakly better to participate than to abstain (→ avoid **no-show paradox**).
- participation → half-way monotonicity
- Moulin (1988): No Condorcet extension satisfies participation. → our result is stronger
### Number of Voters Required

<table>
<thead>
<tr>
<th>n</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
<th>16</th>
<th>17</th>
<th>18</th>
<th>19</th>
<th>20</th>
<th>21</th>
<th>22</th>
<th>23</th>
<th>24</th>
<th>25</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>participation</td>
<td>✔️</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Brandt, Geist, P. (AAMAS 2016): For $m = 4$, participation is compatible with Condorcet consistency for $n = 11$, but incompatible for $n \geq 12$. 
### Number of Voters Required

<table>
<thead>
<tr>
<th>$n$</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
<th>16</th>
<th>17</th>
<th>18</th>
<th>19</th>
<th>20</th>
<th>21</th>
<th>22</th>
<th>23</th>
<th>24</th>
<th>25</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>participation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>half-way monotonicity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Brandt, Geist, P. (AAMAS 2016): For $m = 4$, participation is compatible with Condorcet consistency for $n = 11$, but incompatible for $n \geq 12$. 
Number of Voters Required

<table>
<thead>
<tr>
<th>n</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
<th>16</th>
<th>17</th>
<th>18</th>
<th>19</th>
<th>20</th>
<th>21</th>
<th>22</th>
<th>23</th>
<th>24</th>
<th>25</th>
</tr>
</thead>
<tbody>
<tr>
<td>participation</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>half-way monotonicity</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
</tbody>
</table>

Brandt, Geist, P. (AAMAS 2016): For \( m = 4 \), participation is compatible with Condorcet consistency for \( n = 11 \), but incompatible for \( n \geq 12 \).
### Number of Voters Required

<table>
<thead>
<tr>
<th>Participation</th>
<th>Half-way Monotonicity</th>
</tr>
</thead>
<tbody>
<tr>
<td>n = 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25</td>
<td></td>
</tr>
</tbody>
</table>

Brandt, Geist, P. (AAMAS 2016): For \( m = 4 \), participation is compatible with Condorcet consistency for \( n = 11 \), but incompatible for \( n \geq 12 \).
Method: Impossibility Theorems via SAT Solving

- For fixed $n$ and $m$, write a formula of propositional logic whose models correspond to good voting rules.
- Give to state-of-the-art SAT solver (e.g., lingeling, glucose)
- SAT $\Rightarrow$ good rule given by a look-up table
- UNSAT $\Rightarrow$ impossibility result
- Induction steps on $n$ and $m$

Encodings

- **functionality of** $f$, i.e., that $f(P) = a$ for exactly one alternative, which means that there is at least one and at most one such alternative:

  $$\varphi_{\text{functionality}} \equiv \bigwedge_{P \in A!^N} \left( \bigvee_{a \in A} x_{P,a} \right) \land \bigwedge_{a \neq b \in A} (\neg x_{P,a} \lor \neg x_{P,b})$$

- **Condorcet-consistency**: for $a \in A$, let $C_a \subseteq A!^N$ be the set of profiles in which $a$ is the Condorcet winner.

  $$\varphi_{\text{Condorcet}} \equiv \bigwedge_{a \in A} \left( \bigwedge_{P \in C_a} x_{P,a} \right)$$

- **half-way monotonicity**: if $a, b \in A$ are such that $a >_i b$ for voter $i$ in profile $P \in A!^N$, and $f(P) = b$, then $f(P_{-i}, >_{i}^{\text{rev}}) \neq a$.

  $$\varphi_{\text{half-way monotonicity}} \equiv \bigwedge_{i \in N} \bigwedge_{P \in A!^N} \left( \bigwedge_{a, b \in A} (\neg \nu_{P,b} \lor \neg \nu(P_{-i}, >_{i}^{\text{rev}}, a)) \right) .$$
Satisfiability → Lookup-Table

| A#1(1,1,1,1,1) | A#2(2,2,0,2,0) | A#2(2,2,2,2,2) | B#2(2,0,0,2,0) |
| A#1(1,1,1,1,1) | A#2(2,2,0,2,0) | A#2(2,2,0,2,0) | B#2(0,0,2,2,0) |
| A#1(1,1,1,1,1) | B#2(2,0,0,2,0) | B#2(0,0,2,2,0) | C#2(0,2,0,2,0) |
| A#1(1,1,1,1,1) | A#2(2,0,0,2,0) | A#2(2,0,0,2,0) | C#2(0,2,0,2,0) |
| A#1(1,1,1,1,1) | C#2(0,0,2,2,0) | C#2(0,2,0,2,0) | A#1(1,1,1,1,1) |
| B#1(1,1,1,1,1) | A#2(2,2,2,2,2) | B#2(2,2,0,2,0) | B#2(2,2,2,2,2) |
| B#1(1,1,1,1,1) | A#2(2,2,0,2,0) | B#2(2,2,0,2,0) | B#2(2,2,0,2,0) |
| B#1(1,1,1,1,1) | A#2(2,2,2,2,2) | B#2(2,2,0,2,0) | B#2(2,2,2,2,2) |
| B#1(1,1,1,1,1) | A#2(2,2,0,2,0) | B#2(2,2,2,2,2) | B#2(2,2,0,2,0) |
| B#1(1,1,1,1,1) | A#2(2,2,2,2,2) | B#2(2,2,2,2,2) | B#2(2,2,0,2,0) |
| B#1(1,1,1,1,1) | A#2(2,2,2,2,2) | B#2(2,2,2,2,2) | B#2(2,2,0,2,0) |
| A#1(1,1,1,1,1) | A#2(2,2,2,2,2) | B#2(2,2,2,2,2) | B#2(2,2,2,2,2) |
| A#1(1,1,1,1,1) | A#2(2,2,2,2,2) | B#2(2,2,2,2,2) | B#2(2,2,2,2,2) |
Unsatisfiability – Proof?

c reading input file majority-resolute.cnf
c no embedded options
c found 'p cnf 648 12048' header
c read 648 variables, 12048 clauses, 24144 literals in 0.00 seconds
s UNSATISFIABLE

c 0.000 38% simplifying
c 0.000 0% search

c 0.000 100% all

c 0 decisions, 0.0 decisions/sec
c 0 conflicts, 0.0 conflicts/sec
c 241 propagations, 1.2 megaprops/sec
c 0.0 seconds, 0.1 MB
Unsatisfiability – Proof?

- Use Minimal Unsatisfiable Sets (MUS)
- Subselection of the clauses, already unsatisfiable
- Removing any clause yields satisfiability
- \( \Rightarrow \) every clause in MUS is “necessary proof step”
- Fast tools to find an MUS (e.g. MUSer2, MARCO)
- If small, can obtain a human-readable proof
Proof.

Fig. 1. Proof diagram of the proof of Theorem 5.3.

Fig. 2. Proof diagram of the proof of Theorem 5.4.
Other Examples of the Method

- No Pareto-optimal tournament solution is strategyproof (Brandt and Geist, JAIR 2016) nor satisfies participation (Brandl et al, IJCAI 2015)

- No probabilistic voting rule is efficient and strategyproof (Brandl, Brandt, Eberl, Geist, JACM 2017)

- Impossibilities for set extensions (Geist and Endriss, JAIR 2011).

- Proofs of Arrow’s and Gibbard–Satterthwaite Theorem (Tang and Lin, AIJ 2009)

- No strategyproof and proportional multiwinner rules (P., working paper)

- many other possible applications..
Results in the Paper

- Incompatibility for $n = 15$ and $24$
- Induction steps for $m + 1$ and $n + 2$
- Set-valued voting rules (optimists and pessimists)
- Strong preference reversal paradox (which most Condorcet extensions except maximin exhibit)
- A version of the Gibbard–Satterthwaite Theorem
Every anonymous, unanimous voting rule is either needlessly or egregiously manipulable. \((n \geq 15\) odd, \(m \geq 4\))

Idea for this type of result appears in Zwicker (2016, Handbook of COMSOC)

Proof:

- If the voting rule is a Condorcet extension, then it is manipulable by preference reversal.
- If the voting rule is not a Condorcet extension, then by the Campbell–Kelly Theorem (2003, Econ Theory), the rule is manipulable between two profiles with a Condorcet winner, since the Condorcet rule is the only (non-trivial) strategyproof rule on Condorcet profiles.
Future Work

- Find other versions of G–S with restricted manipulations
  - e.g., known that G–S holds even if manipulators only perform a single swap (Caragiannis et al, EC 2012; Sato, JET 2013)
- Analogue of G–S result for Duggan–Schwartz
- Find concise description of positive examples
- What happens for $m \geq 5$?
Condorcet’s Principle and the Preference Reversal Paradox

Dominik Peters
Department of Computer Science
University of Oxford
dominik-peters.de