Finding Strategyproof Social Choice Functions via SAT Solving

Christian Geist

COST Meeting, Maastricht, The Netherlands
11 April 2014

Felix Brandt
Christian Geist
General idea and success cases

• Idea?
  ‣ Reduce to “small” instance (manually using induction)
  ‣ Solve base case on a computer (using universal problem solving approaches such as SAT)

• Successful?
  ‣ Tang/Lin, 2009: Famous impossibilities (Arrow, Sen, Muller-Satterthwaite, etc) for resolute social choice functions
  ‣ G./Endriss, 2011: Automated theorem search (through universal reduction step) for ranking sets of objects
  ‣ (Brandt/G./Seedig, 2014: Finding preference profiles for $k$-majority digraphs)
  ‣ (G., 2014: Finding preference profiles of given Condorcet dimension)

• Today:
  ‣ Method: more evolved technique to also treat strategyproofness for irresolute social choice functions
  ‣ Results: e.g., efficiency and strategyproofness are incompatible for a natural set extension
Results preview and related work

- Two notions of strategyproofness due to Kelly (1977) and Fishburn (1972) (see also Gärdenfors, 1979)
  - Impossibility: Pareto optimality is incompatible with Fishburn-SP
  - Possibility: There is a refinement of BP that is still Kelly-SP

- Closes gaps in the existing (axiomatic) literature on strategyproofness for irresolute social choice functions, e.g.,
  - Kelly (1977)
  - Barberá (1977)
  - Gärdenfors (1979)
  - Ching and Zhou (2002),
  - Brandt (2011)
  - Brandt and Brill (2011)
  - Sanver and Zwicker (2012)
Outline

• Preliminaries
  ▸ (Majoritarian) social choice functions
  ▸ Strategyproofness (incl. set extensions of preferences)

• Encoding into SAT
  ▸ Initial encoding
  ▸ Optimizations

• Main results
Preliminaries

• Finite sets of \( m \) alternatives, \( n \) voters
  ‣ Voters \( i \) with complete, anti-symmetric and transitive preference relations \( R_i \) over alternatives; strict part \( P_i \) (e.g., \( a P_i b P_i c \))
  ‣ Preference profiles \( R = (R_1, R_2, ..., R_n) \)

• A social choice function (SCF) is a function that maps preference profiles to non-empty subsets of alternatives
  ‣ An SCF \( f \) is resolute if \( |f(R)|=1 \) for all preference profiles \( R \)
  ‣ An SCF \( f \) is neutral if it treats all alternatives equally
  ‣ An SCF \( f \) is majoritarian if it is neutral and \( f(R) \) only depends on the pairwise majority comparisons of \( R \) (majority relation \( R_M \))

• Majoritarian SCFs are also known as tournament solutions
Tournament solution examples

- \( \text{TC}(R_M) \) selects the maximal elements of the transitive closure of \( R_M \)
- \( \text{UC}(R_M) \) consists of all alternatives that are not covered
  - \( x \) covers \( y \) if \( y R_M v \) implies \( x R_M v \) for all \( v \in V \)
- \( \text{BP}(R_M) \) defined based on game theory
  - Alternatives as actions; payoffs based on \( R_M \)
  - \( \text{BP}(R_M) \) consists of all alternatives with positive probability in some Nash equilibrium

\[
\begin{align*}
\text{TC}(R_M) &= \{a, b, c, d\} \\
\text{UC}(R_M) &= \{a, b, c\} \\
\text{BP}(R_M) &= \{a, b, c\}
\end{align*}
\]
A resolute SCF \( f \) is strategyproof if there is no \( R, R', i \in N \) such that \( R_j = R'_j \) for all \( j \neq i \) and \( f(R') \preceq_i f(R) \).

Theorem (Gibbard, 1973; Satterthwaite, 1975): Every strategyproof resolute SCF is either imposed or dictatorial.

- “resoluteness is a rather restrictive and unnatural assumption” (Gärdenfors; 1976 - a philosopher)
- “The Gibbard-Satterthwaite theorem [...] uses an assumption of singlevaluedness which is unreasonable” (Kelly; 1977 - an economist)
- “If there is a weakness to the Gibbard-Satterthwaite theorem, it is the assumption that winners are unique” (Taylor; 2005 - a mathematician)

Problem: Resolute SCFs are incompatible with anonymity and neutrality

Solution: Allow for sets of winners (irresolute SCFs)
- Natural next question: what preferences do voters have over sets of alternatives
Irresolute SCFs: Kelly’s extension

- How to deal with irresoluteness?
  - Assumption: A single alternative is eventually chosen, but the voters do not know anything about the tie-breaking mechanism.
  - Under this assumption, the preferences over choice sets are given by Kelly’s preference extension $R^K \subseteq A \times A$: $X R^K Y \iff \forall x \in X, y \in Y: (x R y)$
  - Example
    - $a P b P c$ implies that $\{a\} P^K \{a,b\} P^K \{b\} P^K \{b,c\}$
    - $\{a,c\}$ and $\{b\}$ are incomparable
    - $\{a,b\}$ and $\{a,b,c\}$ are incomparable(!)

- An SCF $f$ is $P^K$-strategyproof if there is no $R, R', i \in N$ such that $R_j = R'_j$ for all $j \neq i$ and $f(R') P^K_f(R)$
## What we know about Kelly-strategyproofness

<table>
<thead>
<tr>
<th>Kelly-strategyproof</th>
<th>manipulable</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pareto rule</td>
<td></td>
</tr>
<tr>
<td>Omninomination rule</td>
<td></td>
</tr>
<tr>
<td>Top cycle (TC), 1971</td>
<td></td>
</tr>
<tr>
<td>Uncovered set (UC), 1977</td>
<td></td>
</tr>
<tr>
<td>Minimal covering set (MC), 1988</td>
<td></td>
</tr>
<tr>
<td>Bipartisan set (BP), 1993</td>
<td></td>
</tr>
</tbody>
</table>

essentially everything else
Irresolute SCFs: Fishburn’s extension

- How to deal with irresoluteness?
  - Alternative assumption: There is an agent (with preferences unknown to the voters) who picks his most preferred alternative from the choice set, e.g., a chairman or one of the voters.
  - Under this assumption, the preferences over choice sets are given by Fishburn’s preference extension $R^F \subseteq F(U) \times F(U)$:
    $$X R^F Y \Leftrightarrow (\forall x \in X \setminus Y, y \in Y: x R y) \land (\forall x \in X, y \in Y \setminus X: y R x)$$
  - $X R^K Y \Rightarrow X R^F Y$ and hence $R^K \subseteq R^F$
  - Example
    - $a P b P c$ implies that $\{a, b\} P^F \{a, b, c\} P^F \{b, c\}$
    - $\{a, c\}$ and $\{b\}$ are still incomparable

- An SCF $f$ is $P^F$-strategyproof if there is no $R, R', i \in N$ such that $R_j = R'_j$ for all $j \neq i$ and $f(R') P^F_i f(R)$
What we know about Fishburn-strategyproofness

<table>
<thead>
<tr>
<th>Fishburn-strategyproof</th>
<th>manipulable</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pareto rule</td>
<td></td>
</tr>
<tr>
<td>Omninomination rule</td>
<td></td>
</tr>
<tr>
<td>Top cycle (TC), 1971</td>
<td></td>
</tr>
</tbody>
</table>

essentially everything else
Logically equivalent but simpler: tournament-strategyproofness

- A majoritarian SCF \( f \) is said to be \( P^E \)-tournament-strategyproof if there are no \( T, T' \) and \( P_\mu \supseteq (T - T') \) such that \( f(T') P^E_\mu f(T) \)

- Lemma. A majoritarian SCF is \( P^E \)-strategyproof iff it is \( P^E \)-tournament-strategyproof

- Enables more efficient check on a computer, but still large

<table>
<thead>
<tr>
<th>Alternatives</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Choice sets</td>
<td>15</td>
<td>31</td>
<td>63</td>
<td>127</td>
</tr>
<tr>
<td>Tournaments</td>
<td>64</td>
<td>1,024</td>
<td>32,768</td>
<td>(~ 2~)</td>
</tr>
<tr>
<td>Unlabeled tourn.</td>
<td>4</td>
<td>12</td>
<td>56</td>
<td>456</td>
</tr>
<tr>
<td>Majoritarian SCFs</td>
<td>50,625</td>
<td>(~ 10~)</td>
<td>(~ 10~)</td>
<td>(~ 10~)</td>
</tr>
</tbody>
</table>
Finding Strategyproof Social Choice Functions via SAT Solving

High-level system architecture

Input (setting, axioms) → CNF encoder → SAT solver → Results

LP solver → Tournament solver → CNF encoder

nauty → CNF encoder

Model decoder

SAT solver

Neutrality can be expressed as equality, and (iii) the orbit condition.

A direct encoding of this neutrality axiom, however, would lead to much more complexity in the generated clauses, which more than offset the potential advantage of reduced propositional variables. It, however, not only be tedious (quantification over all permutations), but also does our reformulation as canonical isomorphism encoding of both contextual as well as explicit axioms to CNF.

By design SAT solvers operate on propositional logic. Therefore, we want to check whether there exists a majoritarian SCF as a propositional formula and let a SAT solver decide whether this formula has a satisfying assignment. If the formula is unsatisfiable, we have shown that there is no SCF satisfying the respective axioms.

Apart from the explicit axioms, which we are going to describe in the next subsection, there are further axioms that need to be considered in order to fully model the context of majoritarian SCFs. For this purpose, an arbitrary function that maps tournaments to non-empty sets of its vertices will be called a choice function. Using our initial encoding we explain how this encoding can be optimized to increase the overall performance by orders of magnitude such that larger instances of up to 7 alternatives are solvable.

In formal terms this can be written as

\[ f(X) = \{X_1, X_2, \ldots, X_m \} \]

The second and third axioms together constitute strategyproofness by 

\[ \mathcal{P}(T, X) = \mathcal{P}(T, Y) \implies f(T, X) = f(T, Y) \]

Functionality is respected: for any preference profile \( T \) and any alternative \( X \) we have

\[ f(T, X) \neq \emptyset \]

A direct encoding of this neutrality axiom, however, would lead to much more complexity in the generated clauses, which more than offset the potential advantage of reduced propositional variables. It, however, not only be tedious (quantification over all permutations), but also does our reformulation as canonical isomorphism encoding of both contextual as well as explicit axioms to CNF.

By design SAT solvers operate on propositional logic. Therefore, we want to check whether there exists a majoritarian SCF as a propositional formula and let a SAT solver decide whether this formula has a satisfying assignment. If the formula is unsatisfiable, we have shown that there is no SCF satisfying the respective axioms.

Apart from the explicit axioms, which we are going to describe in the next subsection, there are further axioms that need to be considered in order to fully model the context of majoritarian SCFs. For this purpose, an arbitrary function that maps tournaments to non-empty sets of its vertices will be called a choice function. Using our initial encoding we explain how this encoding can be optimized to increase the overall performance by orders of magnitude such that larger instances of up to 7 alternatives are solvable.

In formal terms this can be written as

\[ f(X) = \{X_1, X_2, \ldots, X_m \} \]

The second and third axioms together constitute strategyproofness by 

\[ \mathcal{P}(T, X) = \mathcal{P}(T, Y) \implies f(T, X) = f(T, Y) \]

Functionality is respected: for any preference profile \( T \) and any alternative \( X \) we have

\[ f(T, X) \neq \emptyset \]

A direct encoding of this neutrality axiom, however, would lead to much more complexity in the generated clauses, which more than offset the potential advantage of reduced propositional variables. It, however, not only be tedious (quantification over all permutations), but also does our reformulation as canonical isomorphism encoding of both contextual as well as explicit axioms to CNF.

By design SAT solvers operate on propositional logic. Therefore, we want to check whether there exists a majoritarian SCF as a propositional formula and let a SAT solver decide whether this formula has a satisfying assignment. If the formula is unsatisfiable, we have shown that there is no SCF satisfying the respective axioms.

Apart from the explicit axioms, which we are going to describe in the next subsection, there are further axioms that need to be considered in order to fully model the context of majoritarian SCFs. For this purpose, an arbitrary function that maps tournaments to non-empty sets of its vertices will be called a choice function. Using our initial encoding we explain how this encoding can be optimized to increase the overall performance by orders of magnitude such that larger instances of up to 7 alternatives are solvable.

In formal terms this can be written as

\[ f(X) = \{X_1, X_2, \ldots, X_m \} \]

The second and third axioms together constitute strategyproofness by 

\[ \mathcal{P}(T, X) = \mathcal{P}(T, Y) \implies f(T, X) = f(T, Y) \]

Functionality is respected: for any preference profile \( T \) and any alternative \( X \) we have

\[ f(T, X) \neq \emptyset \]

A direct encoding of this neutrality axiom, however, would lead to much more complexity in the generated clauses, which more than offset the potential advantage of reduced propositional variables. It, however, not only be tedious (quantification over all permutations), but also does our reformulation as canonical isomorphism encoding of both contextual as well as explicit axioms to CNF.

By design SAT solvers operate on propositional logic. Therefore, we want to check whether there exists a majoritarian SCF as a propositional formula and let a SAT solver decide whether this formula has a satisfying assignment. If the formula is unsatisfiable, we have shown that there is no SCF satisfying the respective axioms.

Apart from the explicit axioms, which we are going to describe in the next subsection, there are further axioms that need to be considered in order to fully model the context of majoritarian SCFs. For this purpose, an arbitrary function that maps tournaments to non-empty sets of its vertices will be called a choice function. Using our initial encoding we explain how this encoding can be optimized to increase the overall performance by orders of magnitude such that larger instances of up to 7 alternatives are solvable.

In formal terms this can be written as

\[ f(X) = \{X_1, X_2, \ldots, X_m \} \]

The second and third axioms together constitute strategyproofness by 

\[ \mathcal{P}(T, X) = \mathcal{P}(T, Y) \implies f(T, X) = f(T, Y) \]

Functionality is respected: for any preference profile \( T \) and any alternative \( X \) we have

\[ f(T, X) \neq \emptyset \]

A direct encoding of this neutrality axiom, however, would lead to much more complexity in the generated clauses, which more than offset the potential advantage of reduced propositional variables. It, however, not only be tedious (quantification over all permutations), but also does our reformulation as canonical isomorphism encoding of both contextual as well as explicit axioms to CNF.

By design SAT solvers operate on propositional logic. Therefore, we want to check whether there exists a majoritarian SCF as a propositional formula and let a SAT solver decide whether this formula has a satisfying assignment. If the formula is unsatisfiable, we have shown that there is no SCF satisfying the respective axioms.

Apart from the explicit axioms, which we are going to describe in the next subsection, there are further axioms that need to be considered in order to fully model the context of majoritarian SCFs. For this purpose, an arbitrary function that maps tournaments to non-empty sets of its vertices will be called a choice function. Using our initial encoding we explain how this encoding can be optimized to increase the overall performance by orders of magnitude such that larger instances of up to 7 alternatives are solvable.

In formal terms this can be written as

\[ f(X) = \{X_1, X_2, \ldots, X_m \} \]

The second and third axioms together constitute strategyproofness by 

\[ \mathcal{P}(T, X) = \mathcal{P}(T, Y) \implies f(T, X) = f(T, Y) \]

Functionality is respected: for any preference profile \( T \) and any alternative \( X \) we have

\[ f(T, X) \neq \emptyset \]

A direct encoding of this neutrality axiom, however, would lead to much more complexity in the generated clauses, which more than offset the potential advantage of reduced propositional variables. It, however, not only be tedious (quantification over all permutations), but also does our reformulation as canonical isomorphism encoding of both contextual as well as explicit axioms to CNF.

By design SAT solvers operate on propositional logic. Therefore, we want to check whether there exists a majoritarian SCF as a propositional formula and let a SAT solver decide whether this formula has a satisfying assignment. If the formula is unsatisfiable, we have shown that there is no SCF satisfying the respective axioms.

Apart from the explicit axioms, which we are going to describe in the next subsection, there are further axioms that need to be considered in order to fully model the context of majoritarian SCFs. For this purpose, an arbitrary function that maps tournaments to non-empty sets of its vertices will be called a choice function. Using our initial encoding we explain how this encoding can be optimized to increase the overall performance by orders of magnitude such that larger instances of up to 7 alternatives are solvable.

In formal terms this can be written as

\[ f(X) = \{X_1, X_2, \ldots, X_m \} \]

The second and third axioms together constitute strategyproofness by 

\[ \mathcal{P}(T, X) = \mathcal{P}(T, Y) \implies f(T, X) = f(T, Y) \]

Functionality is respected: for any preference profile \( T \) and any alternative \( X \) we have

\[ f(T, X) \neq \emptyset \]
Basic encoding: goal and variables

• Goal: Encode full problem (of fixed size) into SAT
  ‣ Find propositional formula that is satisfiable iff base case is true

• Variable symbols $c_{T,X}$ to represent $f(T) = X$

• Explicit axioms
  ‣ (Tournament-)strategyproofness
  ‣ Pareto optimality
  ‣ ...

• Context axioms
  ‣ Functionality (of the choice function)
  ‣ Neutrality
Basic encoding: example axiom

- Apart from explicit axioms, 2 main contextual axioms
  - Functionality (of the choice function)
  - Neutrality

- Example: Functionality

\[(\forall T) ((\exists X) c_{T,X} \land \neg (c_{T,Y} \land c_{T,Z}))\]

\[\equiv \bigwedge_T \left( \left( \bigvee_X c_{T,X} \right) \land \bigwedge_{Y \neq Z} (\neg c_{T,Y} \lor \neg c_{T,Z}) \right)\]
Neutrality is not as innocent as it seems

- Formally: \( \pi(f(T)) = f(\pi(T)) \) for all tournaments \( T \) and permutations \( \pi \)
- Has implications across tournaments and even on single tournaments
  - *Across:* Isomorphic tournaments
    - *canonical tournaments* \( T_C \)
Neutrality is not as innocent as it seems

- Formally:
  \[ \pi(f(T)) = f(\pi(T)) \] for all tournaments \( T \) and permutations \( \pi \)

- Has implications across tournaments and even on single tournaments
  - Across: Isomorphic tournaments
    - canonical tournaments \( T_C \)
  - Within: Orbits
Neutrality is not as innocent as it seems

- Formally:
  \[ \pi(f(T)) = f(\pi(T)) \] for all tournaments \( T \) and permutations \( \pi \)

- Has implications across tournaments and even on single tournaments
  - Across: Isomorphic tournaments
  - Within: Orbits

- Lemma. Neutrality is equivalent to the conjunction of
  - Canonical isomorphism equality:
    \[ f(T) = \pi_T(f(T_C)) \]
  - Orbit condition:
    \[ O \subseteq f(T_C) \text{ or } O \cap f(T_C) = \emptyset \]

- Further optimizations of the encoding are possible
Main result: Pareto-optimility and Fishburn-strategyproofness are incompatible

- A SCF is **Pareto-optimal** if its choice sets never contain a Pareto dominated alternative

- **Theorem.** For $m \geq 5$ there is no majoritarian SCF $f$ that satisfies Fishburn-strategyproofness and Pareto-optimality
  - **Lemma.** Pareto-optimality $\iff$ refinement of UC
  - **Lemma.** Base case $m = 5$: automatic verification
    - Fishburn-strategyproofness
    - Refinement of UC
  - **Lemma.** $\exists$ strategyproof maj. SCF $f \subseteq UC$ for $m+1$ alternatives $\Rightarrow$
    $\exists$ strategyproof maj. SCF $f' \subseteq UC$ for $m$ alternatives
Positive result: Kelly-SP

- **Theorem.** There exists a refinement of BP which is still Kelly-strategyproof
  - BP is not the smallest majoritarian SCF satisfying Kelly-strategyproofness
  - The only strategyproof refinement on 5 alternatives
  - Not all desirable properties of BP carry over

- Defined like BP with the exception of:
Proof extraction is possible

Proof trace

Minimal UNSAT core

Human-readable proof
We successfully transferred SAT-based theorem proving to irresolute majoritarian social choice functions

- (Brief) introduction to
  - Irresolute SCFs
  - Majoritarian SCF (tournaments rather than preference profiles)
  - Kelly-/Fishburn-strategyproofness

- Encoding
  - Contextual and explicit axioms \(c(T, X)\)
  - Optimization techniques for improved performance

- Initial new results
  - Incompatibility of Pareto-optimality and Fishburn-strategyproofness
  - Kelly-strategyproof refinement of BP

- (Semi-automatic) proof extraction

- Universality and ease of adaptation most likely to enable further results