Computational Social Choice: Autumn 2010

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

Plan for Today

We have already seen that voters will sometimes have an incentive not to truthfully reveal their preferences when they vote.

Today we shall discuss two important theorems that show that this kind of *strategic manipulation* is impossible to avoid:

- the *Gibbard-Satterthwaite Theorem* [1973/1975]
- the *Duggan-Schwartz Theorem* [2000]

The latter generalises the former by considering irresolute voting procedures, where voters have to strategise wrt. *sets* of winners.

This raises a more general question, and we will spend the final part of the lecture on an introduction to a subfield of SCT concerned with the problem of *extending preferences* over objects to sets of objects.

Example

Let's have another look at our favourite example:

49%:	Bush $\succ$ Gore $\succ$ Nader
20%:	Gore $\succ$ Nader $\succ$ Bush
20%:	Gore $\succ$ Bush $\succ$ Nader
11%:	Nader $\succ$ Gore $\succ$ Bush

Under the plurality rule, the Nader supporters could *manipulate:* pretend they like Gore best and improve the result.

Ballots and Preferences

We have to distinguish a voter's ballot and her true preferences:

- We continue to stipulate that a *ballot* should be a linear order over the alternatives in $\mathcal{X}$, i.e., ballots are elements of $\mathcal{L}(\mathcal{X})$.

  [More generally, we could introduce a *ballot language* $\mathcal{B}(\mathcal{X})$.]

- We shall also assume that each voter has a *preference order* over the alternatives in $\mathcal{X}$. We shall assume that these are linear orders on $\mathcal{X}$ as well, i.e., also preferences are elements of $\mathcal{L}(\mathcal{X})$.

  [Or we could work with a *class of preference structures* $\mathcal{P}(\mathcal{X})$.]

Truthfulness, Manipulation, Strategy-Proofness

For now, we will only deal with *resolute* voting procedures.

Let $\succ_i$ be the true preference of voter i and let b_i be the ballot of i. Some important terminology:

- A voter i is said to vote *truthfully* if her ballot b_i coincides with her actual preference order $\succ_i$.
- A voter i is said to *manipulate* (successfully) if she does not vote truthfully and thereby improves the outcome (wrt. $\succ_i$).
- A *resolute* voting procedure F is called *immune to manipulation* (or *strategy-proof*) if there exist no profile $\underline{b} = (b_1, \ldots, b_n)$ and no voter i such that $F(\underline{b}) \succ_i F(b_1, \ldots, b_{i-1}, \succ_i, b_{i+1}, \ldots, b_n)$ — with $\succ_i$ lifted from alternatives to singletons in the natural manner.

  In other words: under a strategy-proof voting procedure no voter will ever have an incentive to misrepresent her preferences.

The Gibbard-Satterthwaite Theorem

Recall: a resolute voting procedure F is *surjective* if for any alternative x there exists a ballot profile $\underline{b}$ such that $F(\underline{b}) = \{x\}$

Theorem 1 (Gibbard-Satterthwaite) *Any resolute voting procedure for $\geqslant 3$ alternatives that is surjective and strategy-proof is dictatorial.*

Remarks:

- a *surprising* result + not applicable in case of *two* alternatives
- The opposite direction is clear: *dictatorial* $\Rightarrow$ *strategy-proof*
- *Random* procedures don't count (but might be "strategy-proof").

A. Gibbard. Manipulation of Voting Schemes: A General Result. *Econometrica*, 41(4):587–601, 1973.

M.A. Satterthwaite. Strategy-proofness and Arrow's Conditions. *Journal of Economic Theory*, 10:187–217, 1975.

Proof

We shall prove the G-S Theorem to be a corollary of the Muller-Satterthwaite Theorem (even if, historically, G-S came first).

Recall the *Muller-Satterthwaite Theorem:*

- Any *resolute* voting procedure for $\geqslant 3$ alternatives that is *surjective* and *strongly monotonic* is *dictatorial*.

We shall prove a lemma showing that strategy-proofness implies strong monotonicity (and we'll be done). ✓

For short proofs of G-S, see also Barberà (1983) and Benoît (2000).

S. Barberà. Strategy-Proofness and Pivotal Voters: A Direct Proof the Gibbard-Satterthwaite Theorem. *International Economic Review*, 24(2):413–417, 1983.

J.-P. Benoît. The Gibbard-Satterthwaite Theorem: A Simple Proof. *Economic Letters*, 69(3):319–322, 2000.

Strategy-Proofness implies Strong Monotonicity

Lemma 1 *Any resolute voting procedure that is strategy-proof must also be strongly monotonic.*

- SP: no incentive to vote untruthfully
- SM: $F(\underline{b}) = \{x\} \Rightarrow F(\underline{b}') = \{x\}$ if $\forall y : \underline{b}(x \succ y) \subseteq \underline{b}'(x \succ y)$

Proof: We'll prove the contrapositive. So assume F is *not* SM. So there exist alternatives x, x' with $x \neq x'$ and profiles $\underline{b}, \underline{b}'$ such that:

- $\underline{b}(x \succ y) \subseteq \underline{b}'(x \succ y)$ for all alternatives y, including x' $(\star)$
- $F(\underline{b}) = \{x\}$ and $F(\underline{b}') = \{x'\}$

Moving from $\underline{b}$ to $\underline{b}'$, there must be a *first* voter affecting the winner. So w.l.o.g., may assume $\underline{b}$ and $\underline{b}'$ differ only wrt. voter i. Two cases:

- $i \in \underline{b}'(x \succ x')$: if i's true preferences are as in $\underline{b}'$, she can benefit from voting instead as in $\underline{b}$ $\Rightarrow$ ↯ [SP]
- $i \notin \underline{b}'(x \succ x') \Rightarrow^{(\star)} i \notin \underline{b}(x \succ x') \Rightarrow i \in \underline{b}(x' \succ x)$: if i's true pref's are as in $\underline{b}$, she can benefit from voting as in $\underline{b}'$ $\Rightarrow$ ↯ [SP]

Shortcomings of Resolute Voting Procedures

The Gibbard-Satterthwaite Theorem (like Muller-Satterthwaite, but unlike Arrow) only applies to *resolute* voting procedures.

But the restriction to resolute procedures is problematic:

- No "natural" voting procedure is resolute (w/o tie-breaking rule).
- We can get very basic impossibilities for resolute procedures:

  Fact: *No resolute* voting procedure for 2 *voters* and 2 *alternatives* can be both *anonymous* and *neutral*.

  Proof: Consider the case where the voters' rankings differ . . . ✓

We therefore should really be analysing *irresolute* voting procedures . . .

Manipulability wrt. Psychological Assumptions

To analyse manipulability when we might get a set of winners, we need to make assumptions on how voters rank *sets of alternatives*, e.g.:

- A voter is an *optimist* if she prefers X over Y whenever she prefers her favourite $x \in X$ over her favourite $y \in Y$.
- A voter is an *pessimist* if she prefers X over Y whenever she prefers her least preferred $x \in X$ over her least preferred $y \in Y$.

Now we can speak about manipulability by certain types of voters:

- F is called *immune to manipulation by optimistic voters* if no optimistic voter can ever benefit from voting untruthfully.
- F is called *immune to manipulation by pessimistic voters* if no pessimistic voter can ever benefit from voting untruthfully.

Other Axioms

Let F be a voting procedure.

- Recall: a dictator can impose a unique winner. A variation:
  - A voter is a *weak dictator* (or a *nominator*) for F if her top-ranked alternative is always *one of* the winners under F.
  - F is called *weakly dictatorial* if it has a weak dictator; otherwise F is called *strongly nondictatorial*.
- Recall: F is *nonimposed* if for any alternative x there exists a ballot profile $\underline{b}$ such that $F(\underline{b}) = \{x\}$.

The Duggan-Schwartz Theorem

There are several extensions of the Gibbard-Satterthwaite Theorem for irresolute voting procedure. The Duggan-Schwartz Theorem is usually regarded as the strongest of these results.

Our statement of the theorem follows Taylor (2002):

Theorem 2 (Duggan and Schwartz, 2000) *Any voting procedure for $\geqslant 3$ alternatives that is nonimposed and immune to manipulation by both optimistic and pessimistic voters is weakly dictatorial.*

Proof: Omitted.

Note that the Gibbard-Satterthwaite Theorem is a direct corollary.

J. Duggan and T. Schwartz. Strategic Manipulation w/o Resoluteness or Shared Beliefs: Gibbard-Satterthwaite Generalized. *Soc. Choice Welf.*, 17(1):85–93, 2000.

A.D. Taylor. The Manipulability of Voting Systems. *The American Mathematical Monthly*, 109(4)321–337, 2002.

Digression: Ranking Sets of Objects

Recall: to analyse strategic behaviour for irresolute voting procedures we had to make assumptions on how voters rank sets of winners.

That's an interesting problem (and a subfield of SCT) in its own right:

- Given: preference order $\succeq$ declared over *elements* of set $\mathcal{X}$.
- Question: what can we say about the corresponding preference order $\hat{\succeq}$ over *nonempty subsets* of $\mathcal{X}$?
  - Answer: in general, nothing
  - But what are *reasonable principles* for extending preferences?

Besides manipulation in voting theory, other applications include decision making under complete uncertainty.

S. Barberà, W. Bossert, and P.K. Pattanaik. Ranking sets of objects. In *Handbook of Utility Theory*, volume 2. Kluwer Academic Publishers, 2004.

Formal Framework

Three components:

- finite set $\mathcal{X}$ of alternatives (options, candidates, ...)
- total order $\succeq$ on $\mathcal{X}$ (preferences over alternatives)
- weak order $\hat{\succeq}$ on $2^{\mathcal{X}} \setminus \{\emptyset\}$ (preferences over nonempty sets)

preorder:	reflexive, transitive
weak order:	reflexive, transitive, complete
total order:	reflexive, transitive, complete, antisymmetric
linear order:	strict part of total order

Some questions you may ask:

- If we only know $\mathcal{X}$ and $\succeq$, what properties of $\hat{\succeq}$ should we *reasonably* be able to infer?
- What are interesting axioms to impose on structures $\langle \mathcal{X}, \succeq, \hat{\succeq} \rangle$?

Examples

(1) You know $a \succ b \succ c$.
Can you infer $\{a\} \mathrel{\hat{\succ}} \{b, c\}$?

(2) You know $a \succ b \succ c$.
Can you infer anything regarding $\{b\}$ and $\{a, c\}$?

(3) You know $a \succ b \succ c \succ d$.
Can you infer $\{a, b, d\} \mathrel{\hat{\succeq}} \{a, c, d\}$?

Interpretations

Note that there are different possible interpretations to such sets:

(A) You will get one of the elements, but cannot control which.
(B) You can choose one of the elements.
(C) You will get the full set.

Kelly Principle

The extension axiom:

(EXT) $\{a\} \hat{\succ} \{b\}$ if $a \succ b$

Two further axioms:

(MAX) $\{\max(A)\} \hat{\succeq} A$ $[\max(A) =$ best element in A wrt. $\succ]$

(MIN) $A \hat{\succeq} \{\min(A)\}$ $[\min(A) =$ worst element in A wrt. $\succ]$

The *Kelly Principle* = (EXT) + (MAX) + (MIN). That is:

- $A \hat{\succ} B$ if all elements in A are strictly better than all those in B
- $A \hat{\succeq} B$ if all elements in A are at least as good as all those in B

J.S. Kelly. Strategy-Proofness and Social Choice Functions without Single-Valuedness. *Econometrica*, 45(2):439–446, 1977.

Gärdenfors Principle

Two axioms:

(GF1) $A \cup \{b\} \mathrel{\hat{\succ}} A$ if $b \succ a$ for all $a \in A$

(GF2) $A \mathrel{\hat{\succ}} A \cup \{b\}$ if $a \succ b$ for all $a \in A$

The *Gärdenfors Principle* $=$ (GF1) $+$ (GF2):

If I can get from A to B by means of a (nonempty) sequence of steps, each involving deleting the best element or adding a new worst element, then A is strictly better than B.

The Gärdenfors Principle *entails* the Kelly Principle, but not *vice versa*.

P. Gärdenfors. Manipulation of Social Choice Functions. *Journal of Economic Theory*. 1392):217–228, 1976.

Independence

The *independence* axiom:

(IND) $A \cup \{c\} \mathrel{\hat{\succeq}} B \cup \{c\}$ if $A \mathrel{\hat{\succ}} B$ and $c \not\in A \cup B$

That is: if you (strictly) prefer A over B, then that preference should not get inverted when we add a new object c to both sets.

The Kannai-Peleg Theorem

The 1984 paper by Yakar Kannai and Bezalel Peleg is considered *the* seminal contribution to the axiomatic study of *ranking sets of objects*.

Theorem 3 (Kannai and Peleg, 1984) *If $|\mathcal{X}| \geqslant 6$, then no weak order $\hat{\succeq}$ satisfies both the Gärdenfors Principle and independence.*

Probably the first paper treating the problem of preference extension as a *problem in its own right*, from an *axiomatic* perspective.

- For Kelly and Gärdenfors (and others), the problem has been more of a side issue (when studying manipulation in voting).
- Work on the problem of ranking sets of objects itself published before 1984 is descriptive rather than axiomatic.

Y. Kannai and B. Peleg. A Note on the Extension of an Order on a Set to the Power Set. *Journal of Economic Theory*, 32(1):172–175, 1984.

Lemma

Recall the axioms:

$$\left.\begin{array}{ll}\text{(GF1)}\ A \cup \{b\} \mathrel{\hat{\succ}} A & \text{if } b \succ a \text{ for all } a \in A \\ \text{(GF2)}\ A \mathrel{\hat{\succ}} A \cup \{b\} & \text{if } a \succ b \text{ for all } a \in A\end{array}\right\}\ \text{Gärdenfors Principle}$$

$$\text{(IND)}\ A \cup \{c\} \mathrel{\hat{\succeq}} B \cup \{c\} \ \text{ if } A \mathrel{\hat{\succ}} B \text{ and } c \notin A \cup B$$

Lemma 2 *Gärdenfors* + (IND) *entails* $A \mathrel{\hat{\sim}} \{\max(A), \min(A)\}$.

Proof:

- If $|A| \leqslant 2$, then $A = \{\max(A), \min(A)\}$. ✓
- If $|A| > 2$:
  - $\{\max(A)\} \mathrel{\hat{\succ}} A \setminus \{\min(A)\}$ by repeated application of (GF2), and thus $\{\max(A), \min(A)\} \mathrel{\hat{\succeq}} A$ by (IND).
  - $A \setminus \{\max(A)\} \mathrel{\hat{\succ}} \{\min(A)\}$ by repeated application of (GF1), and thus $A \mathrel{\hat{\succeq}} \{\max A, \min(A)\}$ by (IND).

Hence, $A \mathrel{\hat{\sim}} \{\max(A), \min(A)\}$. ✓

Proof of the Kannai-Peleg Theorem

Theorem: If $|\mathcal{X}| \geqslant 6$, then no weak order $\hat{\succeq}$ satisfies both the Gärdenfors Principle and independence.

Proof: Suppose $a_6 \succ a_5 \succ a_4 \succ a_3 \succ a_2 \succ a_1$.

Claim: $\{a_2, a_5\} \hat{\succeq} \{a_4\}$ $(*)$

Proof of claim: if not, then $\{a_4\} \hat{\succ} \{a_2, a_5\}$, as $\hat{\succeq}$ is complete

$\Rightarrow \{a_1, a_4\} \hat{\succeq} \{a_1, a_2, a_5\}$ by (IND)

$\Rightarrow \{a_1, a_2, a_3, a_4\} \hat{\succeq} \{a_1, a_2, a_3, a_4, a_5\}$ by Lemma $\Rightarrow \lightning$ [GP]

Hence: $\{a_2, a_5\} \hat{\succ} \{a_3\}$ from $(*)$ and $\{a_4\} \hat{\succ} \{a_3\}$ [GP]

$\Rightarrow \{a_2, a_5, a_6\} \hat{\succeq} \{a_3, a_6\}$ by (IND)

$\Rightarrow \{a_2, a_3, a_4, a_5, a_6\} \hat{\succeq} \{a_3, a_4, a_5, a_6\}$ by Lemma $\Rightarrow \lightning$ [GP]

Done. ✓

Remark: Note that there *are* *preorders* satisfying all axioms, e.g.:

$A \hat{\succeq} B :\Leftrightarrow \max(A) \succeq \max(B)$ and $\min(A) \succeq \min(B)$

Automated Theorem Search

Recall our discussion of approaches for automated verification of theorems in social choice theory (lecture on impossibility theorems).

For *ranking sets of objects*, we can go one step further and even *discover new theorems* automatically (Geist, 2010):

- Introduce a logic for expressing axioms (many-sorted FOL).
- Identify syntactic conditions on axioms under which any impossibility verified for $|\mathcal{X}| = k$ generalises to all larger domains.
- For a fixed domain, axioms can be expressed in propositional logic.
- Impossibility for a fixed domain can be checked by a SAT solver.
- Can search over all combinations of axioms from a given set and thereby discover all impossibilities (found 84 impossibility theorems for 20 axioms, that apply to any domain $\mathcal{X}$ with $|\mathcal{X}| \geqslant 8$).

C. Geist. *Automated Search for Impossibility Theorems in Choice Theory: Ranking Sets of Objects*. Master of Logic thesis, ILLC, University of Amsterdam, 2010.

Summary

In the first part of the lecture, we have seen that *strategic manipulation* is a major problem in voting:

- *Gibbard-Satterthwaite:* only dictatorships are strategy-proof amongst the surjective and resolute voting procedures.
- *Duggan-Schwartz:* similar result for irresolute procedures

Strategic manipulation is a central issue that connects social choice theory with game theory and mechanism design.

In the second part of the lecture, we have looked closer into the problem of *ranking sets of objects*, which was raised by our analysis of strategic considerations for irresolute voting procedures:

- *Kannai-Peleg:* impossible to satisfy both *dominance* (Gärdenfors) and *independence* principles (for large domains)

What next?

While the Gibbard-Satterthwaite and the Duggan-Schwartz Theorem show that it is *impossible* to always prevent strategic manipulation, next week we'll discuss three approaches to limit the damage:

- *Domain restrictions:* Can we do better if we assume that certain (combinations of) preferences will never occur?
- *Changes in the framework:* To what extent do the impossibilities depend on (possibly dispensable) details of our framework?
- *Complexity barriers:* Even if manipulation is *possible* in principle, maybe it is sometimes (or can be made) computationally hard?