Quantifying the Classical Impossibility Theorems from Social Choice

Frank Feys

October 21, 2015

TU Delft Applied Logic Seminar
0 – Outline

- Introduction

- Part 1: Quantifying the Classical Impossibility Theorems

- Part 2: Cognitive Biases in Small Meeting Sequential Voting

- Conclusion
Example: going to the movies with the family

Suppose the 3 options are Star Wars (SW), The Pianist (TP), and Breakfast at Tiffany’s (BaT). The preferences might be

Dad: TP > SW > BaT
Mom: BaT > TP > SW
Son: SW > TP > BaT

Questions:

1. Which movie should they actually go see? How to decide?
2. How does the decision-making process depend on the social interactions taking place?

TODAY: To study such questions by (1) using the technique Fourier analysis on the Boolean cube, and (2) introducing biases
Social choice theory studies collective decision-making: how to aggregate individual “preferences” into a collective (societal) outcome?

“Preferences” might be votes, judgments, welfare, ... by any “agents” (persons, computers, ...) ⇒ very general framework

Some important questions:

- How to get a coherent societal outcome?
- Properties of different voting rules?
- Is there a “best” voting rule?
- Which rule is more “democratic” / “fair”?
- How about manipulability?
Breakthrough by Kenneth Arrow in 1951:

**ARROW'S IMPOSSIBILITY THEOREM** (informally stated):
There does not exist an “ideal” voting scheme when there are at least 3 candidates.

But what does “ideal” mean?

Arrow proposed use of axiomatic method, suggested desirable properties:

1. **Pareto condition:**
   *If everybody likes Star Wars more than The Pianist, then we definitely shouldn’t go watch The Pianist!*

2. **Independence of Irrelevant Alternatives (IIA)** = “no spoiler condition”

Note that the statement is qualitative: “there’s no ideal scheme”
Arrow’s Theorem is closely related to *Condorcet’s Paradox*: classical example

Voter 1: \( A > B > C \)
Voter 2: \( B > C > A \)
Voter 3: \( C > A > B \)

Condorcet’s idea (= IIA): do all pairwise competitions (with majority)

\[ A \text{ vs. } B, \quad B \text{ vs. } C, \quad \text{and} \quad C \text{ vs. } A \]

Societal outcome: \( A > B > C > A > B > C > \ldots \)  

\( \Rightarrow \) a CYCLE!

Alternative formulation of Arrow’s Theorem: occurrence of cycles is **inevitable**, for any reasonable rule
PART 1: Quantifying the Classical Impossibility Theorems

1. To teach you about the *quantified versions* of the classical impossibility theorems
2. To introduce you to the technique used to prove those theorems: *Fourier analysis on the Boolean cube*
3. To show why these things are *important*

PART 2: Cognitive Biases in Small Meeting Sequential Voting

1. To convince you that, alas, the *decision-making process* in small meetings is often *troubled by persistent biases*
2. To *introduce* you to some of those *biases*
3. To propose a very *simple model to simulate biases*
PART 1:

Quantifying the Classical Impossibility Theorems
Boolean functions $f : \{-1, 1\}^n \rightarrow \mathbb{R}$ are ubiquitous in theoretical computer science.

Each such $f$ has Fourier expansion $f = \sum_{S \subseteq [n]} \hat{f}(S) \chi_S$ where $\chi_S(x) := \prod_{i \in S} x_i$ (parity), and the $\hat{f}(S)$'s are called Fourier coefficients.

For example,

$$\text{Maj}_3(x_1, x_2, x_3) = \frac{1}{2}x_1 + \frac{1}{2}x_2 + \frac{1}{2}x_3 - \frac{1}{2}x_1x_2x_3$$

$\Rightarrow \hat{\text{Maj}}_3(\emptyset) = 0, \hat{\text{Maj}}_3(\{1\}) = \hat{\text{Maj}}_3(\{2\}) = \hat{\text{Maj}}_3(\{3\}) = \frac{1}{2}, \hat{\text{Maj}}_3(\{1, 2, 3\}) = -\frac{1}{2}$

Note: Boolean functions $f : \{-1, 1\}^n \rightarrow \{-1, 1\}$ can be thought of as voting rules.
ESSENCE OF BOOLEAN ANALYSIS:
All of the interesting combinatorial properties of Boolean functions are encoded in the Fourier coefficients
⇒ to “know” the Fourier expansion is to “know” those properties.

Example:
- In social choice, notion of “influence” $\text{Inf}_i[f] \in [0, 1]$ of voting rule $f : \{-1, 1\}^n \to \{-1, 1\}$
- Have a nice formula for this (under mild condition): $\text{Inf}_i[f] = \hat{f}(i)$
- Indeed:
  $$\text{Maj}_3(x_1, x_2, x_3) = \frac{1}{2}x_1 + \frac{1}{2}x_2 + \frac{1}{2}x_3 - \frac{1}{2}x_1x_2x_3$$
  so all voters have influence exactly $1/2$
• Number of applications of Boolean analysis in TCS last 20 years is huge (circuit theory, learning theory, cryptography, communication complexity, pseudorandomness, coding theory, ...)

• In last 15 years also social choice

**MAIN POINT FOR SOCIAL CHOICE:**

It allows us to get from *qualitative* to *quantitative* statements.

• This is very important: in real life we like to quantify things

• Conceptually: “changing the metric” (discrete $\Rightarrow$ Hamming)
QUANTITATIVE ARROW THEOREM \textit{(Kalai/Keller)}: For any \( \varepsilon > 0 \), there is a \( \delta = \delta(\varepsilon) \) such that, if a rule satisfies IIA, then: if the rule is at least \( \varepsilon \)-far from being “highly undesirable” (i.e., being a dictator or breaching Pareto condition), then its probability of having a cycle is at least \( \delta \) (under ICA).

LONG STORY SHORT:

“The more we want to avoid cycles, the more the election scheme will resemble a dictator function (under ICA).”
**Quantitative Arrow Theorem (Kalai/Keller):**

For any $\varepsilon > 0$, there is a $\delta = \delta(\varepsilon)$ such that, if a rule satisfies IIA, then:

if the rule is at least $\varepsilon$-far from being “highly undesirable” (i.e., being a dictator or breaching Pareto condition), then its probability of having a cycle is at least $\delta$ (under ICA).

- Very nice **theoretical** result: “continuity condition” also for Hamming metric

- **Bad news** though: it is negative result, and it **worsens** (aggravates) classical Arrow’s Theorem!

- Here $\delta(\varepsilon) = C \cdot \varepsilon^3$ (with $C$ natural constant)

- Fine poly-dependence on $\varepsilon$, but **unfortunately**

  $$ C \approx 2^{-10,000,000} \approx 0 \quad \text{for all practical purposes} $$
CLASSICAL ARROW’S THEOREM:
Let \( f, g, h : \{-1, 1\}^n \to \{-1, 1\} \) be unanimous (Pareto) and such that, when doing a 3-candidate election based on \((f, g, h)\), the outcome is never a cycle. Then \( f = g = h \), and they are a dictatorship function.

- \( f \) corresponds to \( a \ vs. \ b \)
- \( g \) corresponds to \( b \ vs. \ c \)
- \( h \) corresponds to \( c \ vs. \ a \)

<table>
<thead>
<tr>
<th>voters</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>...</th>
<th>( f(x) \in {-1, 1} )</th>
<th>( g(y) \in {-1, 1} )</th>
<th>( h(z) \in {-1, 1} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( a (+1) vs. b (-1) )</td>
<td>+1</td>
<td>+1</td>
<td>-1</td>
<td>...</td>
<td>( =: x ) ( \leadsto )</td>
<td>( f(x) \in {-1, 1} )</td>
<td>( g(y) \in {-1, 1} )</td>
</tr>
<tr>
<td>( b (+1) vs. c (-1) )</td>
<td>+1</td>
<td>-1</td>
<td>+1</td>
<td>...</td>
<td>( =: y ) ( \leadsto )</td>
<td>( g(y) \in {-1, 1} )</td>
<td>( h(z) \in {-1, 1} )</td>
</tr>
<tr>
<td>( c (+1) vs. a (-1) )</td>
<td>-1</td>
<td>-1</td>
<td>+1</td>
<td>...</td>
<td>( =: z ) ( \leadsto )</td>
<td>( h(z) \in {-1, 1} )</td>
<td>( h(z) \in {-1, 1} )</td>
</tr>
</tbody>
</table>

Crucial! Note that the “forbidden” preferences (i.e., cycles) correspond to
\[
\begin{pmatrix} +1 \\ +1 \\ +1 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} -1 \\ -1 \end{pmatrix}
\]
By ICA, columns are independent and uniformly distributed over the set
\[
\left\{ \left( \frac{+1}{+1} \right), \left( \frac{+1}{-1} \right), \left( \frac{-1}{+1} \right), \left( \frac{-1}{-1} \right) \right\}
\]

Crucial! Let \( \text{NAE} : \{-1, 1\}^n \rightarrow \{0, 1\} \) be indicator of “Not-All-Equal”; its Fourier expansion is
\[
\text{NAE}(t_1, t_2, t_3) = \frac{3}{4} - \frac{1}{4} t_1 t_2 - \frac{1}{4} t_2 t_3 - \frac{1}{4} t_1 t_3
\]

Kalai’s idea: explicitly calculate
\[
\Pr_{\text{ICA}}[\text{no cycles}].
\]

Further analysis of Fourier expansion of NAE
\[\Rightarrow\] Classical Arrow’s Theorem \PLUS quantitative version!

Quantifying the Classical Impossibility Theorems from Social Choice – Frank Feys
Important for theory as well as practice:

1. **Mathematical enrichment**
   - Continuity: impossibility theorems are resilient
   - New modus operandi: the method *itself* is quantitative

2. **Practice**
   - Introduces probabilistic perspective: gives us concrete probabilities
   - There is also a quantitative Gibbard-Satterthwaite Theorem
     ⇒ it has important consequences for computational complexity

*(on next slides…)*
A constructive manipulation means: misreporting your “true” preferences to gain a better outcome (w.r.t. your “true” preferences)

**GIBBARD-SATTERTHWAITE THEOREM** *(informally stated):* For a “reasonable” voting rule (with \( \geq 3 \) candidates), there are always people with an incentive to manipulate: “manipulation is unavoidable”.

- **Idea:** use computational hardness as a barrier against manipulation
  I.e., maybe voting rules exist for which manipulations are *hard to find*?
- **They do exist! Examples:** manipulation problem for
  - single transferable vote (STV)
  - variant of Copeland rule
  are *NP-complete*
- **Problem:** computational complexity is modeled as *worst-case*, but maybe manipulation is easy *on average*!? 
QUANTIFIED GIBBARD-SATTERTHWAITE THEOREM:
If social choice function \((n \text{ voters}, m \text{ candidates})\) is \(\varepsilon\)-far from family of non-manipulable functions NONMANIP, then probability of a profile being manipulable is bounded from below by a polynomial in \(\frac{1}{n}, \frac{1}{m}, \varepsilon\).

- Mossel and Rácz gave a proof, with

\[
p\left(\varepsilon, \frac{1}{n}, \frac{1}{m}\right) = \frac{\varepsilon^{15}}{10^{41}n^{68}m^{167}}
\]

- Quantified G-S Theorem \(\Rightarrow\) manipulation is easy on average

- At least, theoretically: degree of polynomial is \textbf{MUCH} too big:

  for Belgium \((n = 10^7 \text{ and } m = 10)\): \(\text{Prob} \geq \frac{\varepsilon^{15}}{10^{684}}\)

- **Conclusion:** for real-life applications this result is \textit{“theoretically good”}, but \textit{practically insignificant}
PART 2:
Cognitive Biases in Small Meeting Sequential Voting
6 – Problem Description

- Last chapter of thesis: introduction of new model to simulate the various cognitive biases that show up in small meetings.

- Kahneman, Tversky et al. have shown that decisions taken by committees are liable to pervasive biases.

- That’s very unfortunate!
  1. Bad decisions, waste of time and money
  2. But, it is even more sad because of Condorcet’s Jury Theorem / “Wisdom of the Crowd”:

    People as a group can, in principle, come to better decisions, but main problem is lack of independence (due to biases).

    Example: estimating the number of pennies in a glass jar.

- Kahneman got the Nobel Prize in Economics in 2002 (for developing behavioral economics).
Anchoring Effect: when making decisions, individuals tend to rely too much on the first piece of info put forward, the so-called “anchor” (E.g., “starting low” as negotiation skill when buying a car, judges in Germany)

Priming Effect: your actions and emotions can be *primed* by events of which you’re not even aware (E.g., EAT *primes* SOUP, when asked to fill in SO_P)

Halo Effect: people tend to like *everything* about a person/idea/argument whenever they like just *one* part of it (E.g., first impressions—they’re very important!)

Bandwagon Effect: social conformity, groupthink, herding (E.g., “likes” on Facebook, YouTube)

*Many, many more!* Regarding meetings, in particular:
- keeping up one’s appearance/reputation \(\Rightarrow\) avoiding disagreement \(\Rightarrow\)
- status-quo bias \(\Rightarrow\) bias towards “the obvious” (discarding private insights)
Setup:

- Small committee of \( n \) persons \( \{1, 2, \ldots, n\} \) has to decide on proposal: accept (1) or reject (0)?
- All votes are public, members vote **sequentially** (in order of index), and then the **majority rule** is applied.
- Each member has a **sway** \( w_i \in [0, 1] \), signifying that person’s “**weight**” (In English, “sway” is a synonym of “domination” “authority” “influence” “leadership”)
- Additional constraint: \( w_1 + \ldots + w_n \leq 1 \)
- Probabilistic model: if \( x_i \in \{0, 1\} \) is i’s vote,
  1. Individual 1 votes uniformly at random
  2. For any \( i \in \{2, 3, \ldots, n\} \),

\[
x_i \overset{\text{def}}{=} \begin{cases} 
  x_j & \text{with probability } w_j \quad (\forall j < i) \\
  \{0, 1\} \text{ uniformly at random} & \text{with probability } 1 - \sum_{j=1}^{i-1} w_j
\end{cases}
\]
$x_1 = \text{random}$

$x_2 = \text{same as } x_1$ \hspace{1cm} $1 - w_1$

\hspace{1cm} $w_1$

$w_1$

$w_2$

$1 - (w_1 + w_2)$

\hspace{1cm} $w_2$

$1 - (w_1 + w_2)$

$x_3 = \text{same as } x_1 \hspace{1cm} \text{same as } x_2 \hspace{1cm} \text{random}$

\hspace{1cm} $w_1$

$1 - (w_1 + w_2)$

$\text{same as } x_1 \hspace{1cm} \text{same as } x_2 \hspace{1cm} \text{random}$
Very simple model, Boolean-inspired ⇒ open to criticism

In thesis only the case $n = 3$ was analyzed (time constraints)

So, how to go about small meeting voting? Kahneman’s advice:

1. Before the meeting starts, all members secretly write down on paper a summary of their opinion

2. Who speaks first?
   - Either the first person to speak is picked uniformly at random (to avoid the same dominant personalities dominating the discussions)
   - Or people are required to speak in reverse order of “dominance” (sway)

3. Disagreement should be supported and even rewarded

Interesting future research: apply tools from logic to study such questions
Fourier analysis on the Boolean cube is a useful technique also for social choice theory.

Best illustration: strengthening of classical impossibility theorems (Arrow, Gibbard-Satterthwaite) into quantitative theorems.

Social choice theory is linked with many other areas; there is a particularly interesting connection with cognitive sciences (biases).

We introduced a new, simple model to simulate the various cognitive biases present in small meeting decision-making.

These topics are very important, also for real life.