Quantifying the Classical Impossibility Theorems from Social Choice

Frank Feys

October 21, 2015

TU Delft Applied Logic Seminar

- 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, ...) $\Rightarrow$ 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)

$\underline{A \text{ vs. } B}$, $\underline{B \text{ vs. } C}$, and $\underline{C \text{ vs. } A}$

- Societal outcome: $A > B > C > A > B > C > \dots$

$\Rightarrow$ 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 $f : \{-1, 1\}^n \to \mathbb{R}$ are ubiquitous in theoretical computer science
- Each such f has Fourier expansion $f = \sum_{S \subseteq [n]} \widehat{f}(S)\chi_S$ where $\chi_S(x) := \prod_{i \in S} x_i$ (parity), and the $\widehat{f}(S)$'s are called Fourier coefficients
- For example,

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

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

- Note: Boolean functions $f : \{-1, 1\}^n \to \{-1, 1\}$ 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*
$\Rightarrow$ to "know" the Fourier expansion is to "know" those properties.

Example:

- In social choice, notion of "influence" $\mathrm{Inf}_i[f] \in [0,1]$ of voting rule $f : \{-1,1\}^n \to \{-1,1\}$
- Have a nice formula for this (under mild condition): $\mathrm{Inf}_i[f] = \widehat{f}(i)$
- Indeed:

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

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 *(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 ε-far from being "highly undesirable" (i.e., being a dictator or breaching Pareto condition), then its probability of having a cycle is *at least* δ (under ICA).

$$\Downarrow$$

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 ε-far from being "highly undesirable" (i.e., being a dictator or breaching Pareto condition), then its probability of having a cycle is *at least* δ (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 ε , but unfortunately

$$C \approx 2^{-10,000,000} = 0 \quad \text{for all practical purposes}$$

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

	voters						
	1	2	3	...			
a (+1) vs. b (−1)	$+1$	$+1$	-1	...	$=: x$	$\rightsquigarrow$	$f(x) \in \{-1, 1\}$
b (+1) vs. c (−1)	$+1$	-1	$+1$	...	$=: y$	$\rightsquigarrow$	$g(y) \in \{-1, 1\}$
c (+1) vs. a (−1)	-1	-1	$+1$	...	$=: z$	$\rightsquigarrow$	$h(z) \in \{-1, 1\}$

- Crucial! Note that the "forbidden" preferences (i.e., cycles) correspond to

$$\begin{pmatrix} +1 \\ +1 \\ +1 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} -1 \\ -1 \\ -1 \end{pmatrix}$$

- By ICA, columns are independent and uniformly distributed over the set

$$\left\{ \begin{pmatrix}+1\\+1\\-1\end{pmatrix}, \begin{pmatrix}+1\\-1\\-1\end{pmatrix}, \begin{pmatrix}+1\\-1\\+1\end{pmatrix}, \begin{pmatrix}-1\\+1\\-1\end{pmatrix}, \begin{pmatrix}-1\\+1\\+1\end{pmatrix}, \begin{pmatrix}-1\\-1\\+1\end{pmatrix} \right\}$$

- Crucial! Let NAE : $\{-1,1\}^n \to \{0,1\}$ be indicator of "Not-All-Equal"; its Fourier expansion is

$$\text{NAE}(t_1, t_2, t_3) = \frac{3}{4} - \frac{1}{4}t_1t_2 - \frac{1}{4}t_2t_3 - \frac{1}{4}t_1t_3$$

Kalai's idea: explicitly calculate

$$\Pr_{\text{ICA}}[\ \textbf{no cycles}\ ].$$

- Further analysis of Fourier expansion of NAE

$\Rightarrow$ Classical Arrow's Theorem PLUS *quantitative* version!

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 ≥ 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 voters, m candidates) is ε-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 ⇒ manipulation is easy *on average*
- At least, theoretically: degree of polynomial is MUCH too big:

  for Belgium ($n = 10^7$ and $m = 10$): Prob $\geq \frac{\varepsilon^{15}}{10^{684}}$
- **Conclusion:** *for real-life applications this result is "theoretically good", but practically insignificant*

PART 2:

Cognitive Biases in Small Meeting Sequential Voting

- 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 pervading 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*)

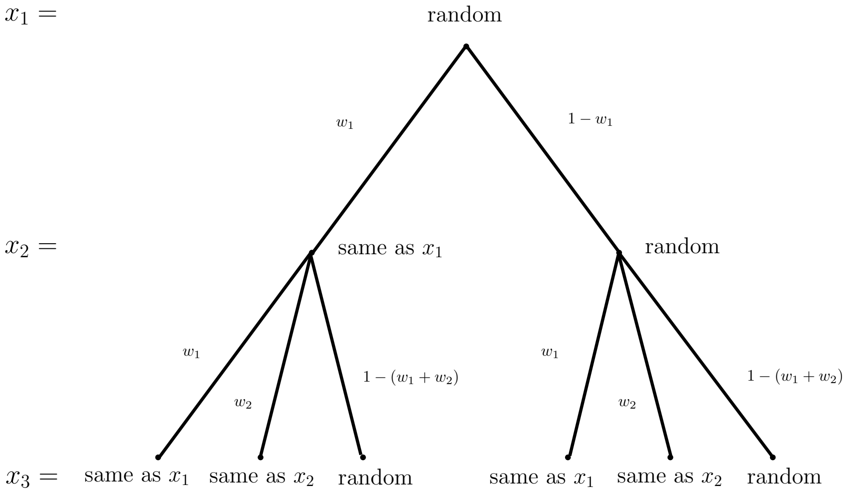
6 – Amos Tversky (dec., 1996) and Daniel Kahneman

- 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 negotation 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* SO**U**P, 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 ones's appearance/reputation ⇄ avoiding disagreement ⇄ status-quo bias ⇄ 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 $i \in \{2, 3, \ldots, n\}$,

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


$x_1 =$
random
w_1
$1 - w_1$
$x_2 =$
same as x_1
random
w_1
w_2
$1 - (w_1 + w_2)$
w_1
w_2
$1 - (w_1 + w_2)$
$x_3 =$
same as x_1
same as x_2
random
same as x_1
same as x_2
random

- Very simple model, Boolean-inspired $\Rightarrow$ 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

1. Fourier analysis on the Boolean cube is a useful technique also for social choice theory
2. Best illustration: strengthening of classical impossibility theorems (Arrow, Gibbard-Satterthwaite) into quantitative theorems
3. Social choice theory is linked with many other areas; there is a particularly interesting connection with cognitive sciences (biases)
4. We introduced a new, simple model to simulate the various *cognitive biases* present in small meeting decision-making
5. These topics are very important, also for real life