Towards less manipulable voting systems

François Durand, Fabien Mathieu, Ludovic Noirie

Inria / Alcatel-Lucent Bell Labs France
Displexity workshop

April 5, 2013

Introduction: voting systems

Context:

- Origins in **politics**.
- Applications in any situation of **collective choice**.

Questions:

- Is there a natural way to select a **reasonable winner**?
- Can we **trust** the electors?
- If not, is it possible to design a voting system that is **resistant to manipulation**?

Terminology warning: "manipulation"
$=$ internal manipulation by electors themselves
$=$ tactical voting
$\neq$ "bribery" (somebody pays electors to change their votes), etc.

Plan

Previously on voting systems. . .

- In quest for a reasonable winner
- Presentation of manipulability
- Reducing manipulability, step 1: condorcification

Reducing manipulability, step 2: slicing

- Introduction
- Idea of the theorem in a particular case
- Generalization of the theorem

Conclusion and future work

Plan

A simplified framework

n **electors**.
m **candidates** named A, B, C...

Each elector i has a **binary relation** r_i over the candidates, that represents her **preferences**.

- Example of i's preferences: $\text{A} \sim \text{B} \succ \text{D} \succ \text{C}$.

Voting system $f : (r_1, \ldots, r_n) \to v \in \{\text{A}, \text{B}, \text{C} \ldots\}$.

2 candidates: May's theorem

Plurality (= "uninominal à un tour"):

- Each elector votes for one candidate.
- The candidate with most votes gets elected.

May's theorem (1952): plurality is the only anonymous, neutral and positively responsive voting system for 2 candidates.

3 candidates or more: Condorcet winner

Independence of irrelevant alternatives (IIA): if we remove one of the losing candidates, the winner should remain the same.

E.g. if C wins, then she must win any electoral duel versus A or B. If it is the case, we say that she's a **Condorcet winner**.

Example:

	Electors		
	40	35	25
Preferences	A	B	C
	C	C	A
	B	A	B

"Majority matrix":

	A	B	C	Victories
A	[redacted]			0
B		[redacted]		0
C			[redacted]	0

3 candidates or more: Condorcet winner

Independence of irrelevant alternatives (IIA): if we remove one of the losing candidates, the winner should remain the same.

E.g. if C wins, then she must win any electoral duel versus A or B. If it is the case, we say that she's a **Condorcet winner**.

Example:

	Electors		
	40	35	25
Preferences	A	B	C
	C	C	A
	B	A	B

"Majority matrix":

	A	B	C	Victories
A		**65**		1
B	35			0
C				0

3 candidates or more: Condorcet winner

Independence of irrelevant alternatives (IIA): if we remove one of the losing candidates, the winner should remain the same.

E.g. if C wins, then she must win any electoral duel versus A or B. If it is the case, we say that she's a **Condorcet winner**.

Example:

	Electors		
	40	35	25
Preferences	A	B	C
	C	C	A
	B	A	B

“Majority matrix”:

	A	B	C	Victories
A		**65**	40	1
B	35			0
C	**60**			1

3 candidates or more: Condorcet winner

Independence of irrelevant alternatives (IIA): if we remove one of the losing candidates, the winner should remain the same.

E.g. if C wins, then she must win any electoral duel versus A or B. If it is the case, we say that she's a **Condorcet winner**.

Example:

	Electors		
	40	35	25
Preferences	A	B	C
	C	C	A
	B	A	B

"Majority matrix":

	A	B	C	Victories
A		**65**	40	1
B	35		35	0
C	**60**	**65**		2

3 candidates or more: Condorcet winner

Independence of irrelevant alternatives (IIA): if we remove one of the losing candidates, the winner should remain the same.

E.g. if C wins, then she must win any electoral duel versus A or B. If it is the case, we say that she's a **Condorcet winner**.

Example:

	Electors		
	40	35	25
Preferences	A	B	C
	C	C	A
	B	A	B

"Majority matrix":

	A	B	C	Victories
A		**65**	40	1
B	35		35	0
C	**60**	**65**		2

If we want to extend Plurality for $m \geq 3$ and respect IIA, then C must be elected. We then say that our voting system respects **Condorcet criterion**.

3 candidates or more: Condorcet's paradox

Example:

	Electors		
	40	35	25
Preferences	A	B	C
	B	C	A
	C	A	B

"Majority matrix":

	A	B	C	Victories
A		**65**	40	1
B	35		**75**	1
C	**60**	25		1

Condorcet's paradox (1785): A defeats B, B defeats C and C defeats A.

It's not possible to extend Plurality for $m \geq 3$ while respecting independence of irrelevant alternatives (IIA).

Arrow's theorem

We would like a voting system with the following properties.

- **Non-dictatorship:** there is not one elector who always decides alone.
- **Unanimity:** whenever all electors prefers A to B, candidate B cannot get elected.
- **Independence of irrelevant alternatives (IIA):** if we remove one of the losing candidates, the winner should remain the same.

Arrow's theorem

We would like a voting system with the following properties.

- **Non-dictatorship:** there is not one elector who always decides alone.
- **Unanimity:** whenever all electors prefers A to B, candidate B cannot get elected.
- **Independence of irrelevant alternatives (IIA):** if we remove one of the losing candidates, the winner should remain the same.

Arrow's theorem (1951): for $m \geq 3$ candidates, such a voting system does not exist.

$\Rightarrow$ For $m \geq 3$ candidates, there is no "natural", canonical way to aggregate binary relations of preferences from several electors in order to choose a winning candidate.

Extending the framework of preferences

Old framework: binary relations of preferences only.

Example of extended framework:

- Each elector has a **utility vector** about the candidates, e.g. $(10, 10, 0, 2)$.
- This utility vector induces a binary relation of preferences over the candidates, e.g. A $\sim$ B $\succ$ D $\succ$ C.

Extending the framework of preferences

Old framework: binary relations of preferences only.

Example of extended framework:

- Each elector has a **utility vector** about the candidates, e.g. $(10, 10, 0, 2)$.
- This utility vector induces a binary relation of preferences over the candidates, e.g. $\mathrm{A} \sim \mathrm{B} \succ \mathrm{D} \succ \mathrm{C}$.

General extended framework:

- Each elector i has a **state** $\omega_i \in \Omega_i$,
- This state contains enough information so that we can extract her **binary relation of preference** $r_i = R_i(\omega_i)$.

Voting system $f : (\omega_1, \ldots, \omega_n) \to v \in \{\mathrm{A}, \mathrm{B}, \mathrm{C} \ldots\}$.

Escaping Arrow's theorem

Example: range voting.

- Each elector gives her utility vector, that is, a note for each candidate.
- The candidate with highest average (or median) wins.

This voting system is **non-dictatorial**, **unanimous** and **independent of irrelevant alternatives**... and infinitely many other voting systems are too!

So... have we won? Have we found a voting system that is fully satisfying?

Plan

Manipulability: an example

Voting system: range voting.

		Sincere			
		A	B		
Electors	80%	7	5		
	20%	2	5		
Average		6	5		
Winner		A			

Manipulability: an example

Voting system: range voting.

		Sincere		Tactical	
		A	B	A	B
Electors	80%	7	5	7	5
	20%	2	5	0	10
Average		6	5	5,6	6
Winner		A		B	

We say that this situation is **manipulable** for this voting system:

- A subset of electors, by casting a tactical ballot, may change the result to a candidate they prefer.
- I.e., sincere voting is not a **strong Nash equilibrium**.

Gibbard's theorem

Gibbard's theorem (1973): for any non-dictatorial voting system with at least 3 eligible candidates, there exists a situation that is manipulable by one elector.

I.e.: this situation is not even a **weak Nash equilibrium**.

Manipulability rate

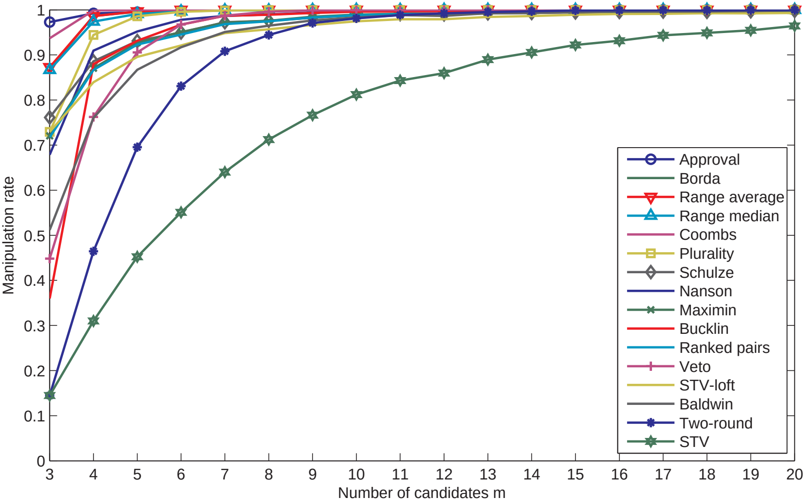
We draw a situation $\omega = (\omega_1, \ldots, \omega_n)$ (states of all electors) according to a **probability measure** P.

Manipulability rate: what is the probability that this situation ω is manipulable for voting system f?

$$\rho_P(f) = P(\omega \text{ is } f\text{-manipulable}).$$

Manipulability is quite frequent

P = "Uniform spherical culture", $n = 33$ electors



Plan

Condorcification

Durand, Mathieu and Noirie (2012)

The ***condorcification*** of f is a new voting system f^c:

- Whenever there exists a Condorcet winner, designate her;
- Otherwise, use f.

If f has reasonable properties, then f^c is **at most as manipulable** as f:

- Any situation ω manipulable for f^c is manipulable for f;
- In particular, $\rho_P(f^c) \leq \rho_P(f)$ for any probability measure P.

If f meets a simple condition, then f^c is **strictly less manipulable** than f. It is the case for all classical voting systems that do not meet Condorcet criterion already (except veto).

Plan

Goal

Some voting systems may not depend on binary relations of preference only.

- Example: range voting (a note for each candidate).

Intuition:

- Binary relations of preference = **necessary information** (define Condorcet winner and coalitions).
- More information than that = **more opportunities for lies**.

Our goal: restrict to binary relations of preferences while reducing manipulability.

- Restrict the scope of research for a voting system with the lowest manipulability.

Reminder of the framework

Electoral space:

- n: number of electors
- m: number of candidates
- ω_i: elector i's state (utilities, etc.)
- $r_i = R_i(\omega_i)$: elector i's binary relation of preference over the candidates

Voting system:

$$f : (\omega_1, \ldots, \omega_n) \to v \in \{\mathrm{A}, \mathrm{B} \ldots\}.$$

Manipulability of voting system f: a more formal definition

Situation $\omega = (\omega_1, \ldots, \omega_n)$ is ***f-manipulable to situation*** ψ iff:

$$\begin{cases} f(\psi) \neq f(\omega), \\ \forall \text{ elector } i, \psi_i \neq \omega_i \Rightarrow f(\psi) \succ_{R_i(\omega_i)} f(\omega). \end{cases}$$

Situation ω is ***f-manipulable*** iff there exists such a ψ.

Manipulability indicator of f:

$$M_f(\omega) = \left| \begin{array}{l} 1 \text{ if } \omega \text{ is } f\text{-manipulable}, \\ 0 \text{ otherwise.} \end{array} \right.$$

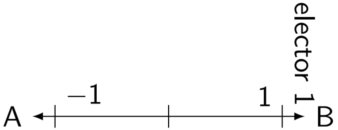
With P a probability distribution used to draw the situation ω, the ***manipulability rate*** of f is:

$$\rho_P(f) = \int M_f(\omega) P(d\omega).$$

Plan

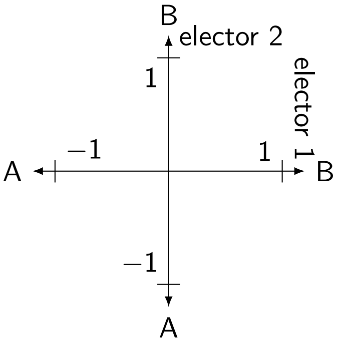
Example

2 electors, 2 candidates named A ($-$) and B ($+$).
$\omega_i = (r_i, y_i)$, where $y_i \in]0, 1]$ is an intensity of preference.



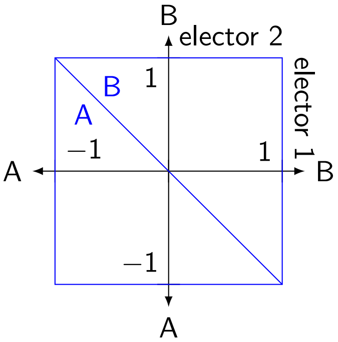
Example

2 electors, 2 candidates named A ($-$) and B ($+$).
$\omega_i = (r_i, y_i)$, where $y_i \in]0, 1]$ is an intensity of preference.



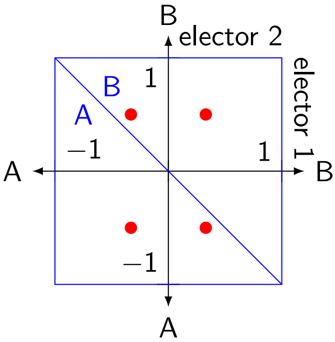
Example

2 electors, 2 candidates named A ($-$) and B ($+$).
$\omega_i = (r_i, y_i)$, where $y_i \in]0, 1]$ is an intensity of preference.
Voting system f: elect sign of the sum.



Example

2 electors, 2 candidates named A ($-$) and B ($+$).
$\omega_i = (r_i, y_i)$, where $y_i \in]0, 1]$ is an intensity of preference.
Voting system f: elect sign of the sum.

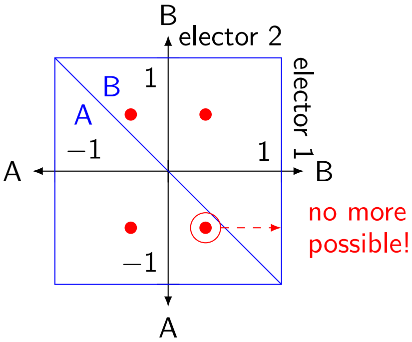


E.g. do as if $y_1 = \frac{1}{3}$ and $y_2 = \frac{1}{2}$, whatever electors say.

- ***Sliced*** voting system f_y.
- Depends on r_1 and r_2 only.

Example

2 electors, 2 candidates named A $(-)$ and B $(+)$.
$\omega_i = (r_i, y_i)$, where $y_i \in]0,1]$ is an intensity of preference.
Voting system f: elect sign of the sum.



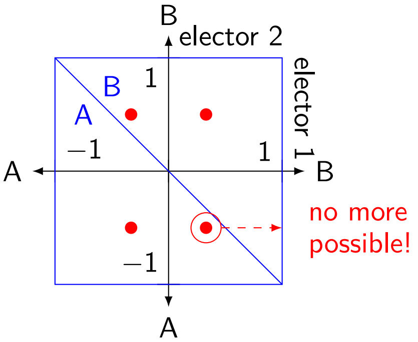
E.g. do as if $y_1 = \frac{1}{3}$ and $y_2 = \frac{1}{2}$, whatever electors say.

- ***Sliced*** voting system f_y.
- Depends on r_1 and r_2 only.

Red situations are less manipulable than before!

Example

2 electors, 2 candidates named A ($-$) and B ($+$).
$\omega_i = (r_i, y_i)$, where $y_i \in]0, 1]$ is an intensity of preference.
Voting system f: elect sign of the sum.



E.g. do as if $y_1 = \frac{1}{3}$ and $y_2 = \frac{1}{2}$, whatever electors say.

- ***Sliced*** voting system f_y.
- Depends on r_1 and r_2 only.

Red situations are less manipulable than before!

If r and y are independent, then f's manipulability rate is the average of its restrictions to such red figures.

A particular case for the slicing theorem

Elector i's state: $\omega_i = (r_i, y_i)$

- r_i: elector i's binary relation of preference over the candidates
- y_i: **additional information** about elector i's preferences

Probability laws:

- $\omega = (\omega_1, \ldots, \omega_n) \in \Omega \sim P$
- $R(\omega) = (r_1, \ldots, r_n) \in \mathcal{R} \sim \mu$ (law induced by P)
- $Y(\omega) = (y_1, \ldots, y_n) \in \mathcal{Y} \sim \nu$ (law induced by P)

Assumption: random variables R and Y are **independent**.

- $(\Omega, P) = (\mathcal{R}, \mu) \times (\mathcal{Y}, \nu)$.

Does **not** mean that electors are independent!

Slicing theorem

Durand, Mathieu and Noirie (2013)

Lemma

For situations of type (r, y), slice f_y is less manipulable than f:

$$M_{f_y}(r, y) \leq M_f(r, y).$$

Sketch of proof: the same candidate is elected, but opportunities for tactical voting are limited (red situations in previous example).

Theorem

There exists y such that $\rho_P(f_y) \leq \rho_P(f)$.

Remark

If f respects Condorcet criterion, then any slice f_y does.

- Final voting system f_y is interesting (not dictatorial).

Proof of the theorem

For any y, we have:

$$\begin{aligned}\rho_P(f_y) &= \int \mu(dr) \int \nu(dy') M_{f_y}(r, y') && \text{(Fubbini)} \\ &= \int \mu(dr) M_{f_y}(r, y). && (f_y \text{ depends on } r \text{ only})\end{aligned}$$

Manipulability of f:

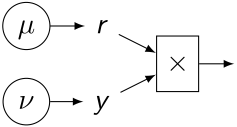
$$\begin{aligned}\rho_P(f) &= \int \nu(dy) \int \mu(dr) M_f(r, y) && \text{(Fubbini)} \\ &\geq \int \nu(dy) \underbrace{\int \mu(dr) M_{f_y}(r, y)}_{=\rho_P(f_y)} && \text{(lemma)}\end{aligned}$$

So, there exists y such that $\rho_P(f_y) \leq \rho_P(f)$.

Plan

Decomposed electoral space

Particular case seen before

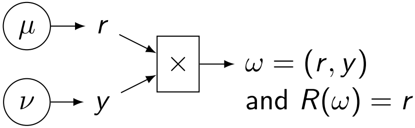


$\omega = (r, y)$ and $R(\omega) = r$

$$(\Omega, P) = (\mathcal{R}, \mu) \times (\mathcal{Y}, \nu)$$

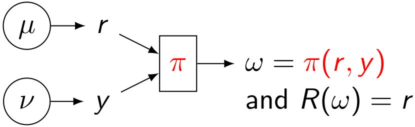
Decomposed electoral space

Particular case seen before



$$(\Omega, P) = (\mathcal{R}, \mu) \times (\mathcal{Y}, \nu)$$

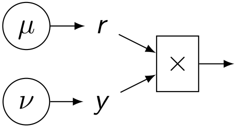
General case



$$(\Omega, P) = \pi\big((\mathcal{R}, \mu) \times (\mathcal{Y}, \nu)\big)$$

Decomposed electoral space

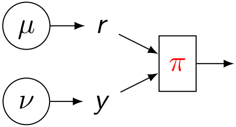
Particular case seen before



$\omega = (r, y)$
and $R(\omega) = r$

$$(\Omega, P) = (\mathcal{R}, \mu) \times (\mathcal{Y}, \nu)$$

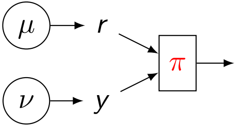
General case



$\omega = \pi(r, y)$
and $R(\omega) = r$

$$(\Omega, P) = \pi\big((\mathcal{R}, \mu) \times (\mathcal{Y}, \nu)\big)$$

Simpler (but equivalent) general case



$\omega = y(r)$
and $R(\omega) = r$

Each y_i is a function $r_i \to \omega_i$.
π is the evaluation.

Example of decomposed electoral space

$n = 2$ electors, $m = 2$ candidates.

μ draws equiprobably:

1. Each elector has binary relation $\mathrm{A} \succ \mathrm{B}$,
2. Each elector has binary relation $\mathrm{B} \succ \mathrm{A}$.

ν always draws functions y_1 and y_2 that each does:

$$\begin{array}{rcl} \mathrm{A} \succ \mathrm{B} & \rightarrow & \mathcal{A} = (\mathrm{A} \succ \mathrm{B}, \text{orange}) \\ \mathrm{B} \succ \mathrm{A} & \rightarrow & \mathcal{B} = (\mathrm{B} \succ \mathrm{A}, \text{blue}). \end{array}$$

Then $P = \pi(\mu \times \nu)$ draws equiprobably:

1. Each elector is in state $\mathcal{A} = (\mathrm{A} \succ \mathrm{B}, \text{orange})$,
2. Each elector is in state $\mathcal{B} = (\mathrm{B} \succ \mathrm{A}, \text{blue})$.

Example of R-decomposable electoral space

$n = 2$ electors, $m = 2$ candidates.

P draws equiprobably:

1. Each elector is in state $\mathcal{A} = (\text{A} \succ \text{B}, \text{orange})$,
2. Each elector is in state $\mathcal{B} = (\text{B} \succ \text{A}, \text{blue})$.

Then:

- μ (the law of R) is like before,
- we can exhibit ν like before.

We say that space (Ω, P) is R-***decomposable***.

Definition of R-decomposable electoral space

What is given:

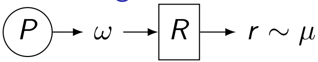
- $\omega = (\omega_1, \dots, \omega_n) \in \Omega \sim P$
- $R(\omega) = (r_1, \dots, r_n) \in \mathcal{R} \sim \mu$ (law induced by P)

We want a probability measure ν that draws n functions y_i such that:

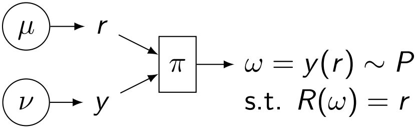
$$\begin{cases} P = \pi(\mu \times \nu), \text{where } \pi \text{ is the evaluation,} \\ \left(\forall r \in R(\Omega), R(y(r)) = r\right), \quad \nu\text{-almost surely for } y. \end{cases}$$

One elector: lemma of complementary random variable

What is given



What we want



Lemma

- (Ω, P) a probability space.
- R a random variable with values in a measurable space $\mathcal{R}$.

Assumption: $\mathcal{R}$ is **finite** (and endowed with discrete σ-algebra).
Then (Ω, P) **is** R**-decomposable**.

Sketch of proof

Independently for each possible value r, we choose $y(r)$ according to P_r (conditional probability knowing r).

Several independent electors

Proposition

If electors are **independent**, then the electoral space is **R-decomposable**.

Sketch of proof

Apply lemma of complementary random variable to each elector.

Consequence

Slicing works: for any voting system f, there exists a sliced voting system f_y that is at most as manipulable as f.

Other criteria of decomposability

Independence is not a necessary condition.
Cf. example with orange and blue electors.

- We know other sufficient conditions.
- We also know a necessary condition.
- But we know no simple equivalent condition for the R-decomposability of an electoral space.

Plan

Existence of an optimal voting system

P: probability measure used to draw population preferences (with reasonable properties = decomposability).

When looking for a voting system whose **manipulation rate is minimal** (among systems with reasonable properties), we can restrict the research to the class of those that:

- **depend on binary relations of preference only**,
- **respect Condorcet criterion**.

For each pair (n, m), there is a finite number of such voting systems.

- **An optimal voting system exists.**

Future work: know more about optimal voting systems

For each pair (n, m), number of voting systems that depend on binary relations of preference only and respect Condorcet criterion:

$$\sim m^{\left(\left(2^{(m^2)}\right)^n\right)}.$$

For the moment, we can give an optimal voting system explicitly for **very small** values: $n \leq 5$ electors and $m \leq 4$ candidates...

Conclusion

- For voting systems that rely on binary relations of preferences only, there is no canonical way to choose the winner (Arrow's theorem).
- For $m \geq 3$ candidates, all non-dictatorial voting systems are subject to manipulation (Gibbard's theorem).
- We can limit manipulability by **condorcification** and **slicing**.
- There exists an **optimal voting system** that depends on binary relations of preferences only and respects Condorcet criterion.

Thanks for your attention! Questions?

Why would manipulability be a problem?

If electors do not vote sincerely, the collective decision relies on **false information**.

If electors do vote sincerely, they may be frustrated and find the system nonsensical, since a non-sincere ballot, **misrepresenting** their preferences, would have **defended these preferences better**.

Restriction to sincere strategies (*reduction*)

Two-round system:

- Round 1: Fabien votes for A.
- Round 2, A versus B: Fabien votes for B.

Obviously, Fabien is not sincere!

Solution:

- do only one round,
- ask preferences directly,
- determine the corresponding sincere strategy automatically.

= ***reduced*** **voting system**

People may still lie about their preferences. But they can no more use strategies that are obviously insincere!

Manipulability: an example

Voting system: plurality.

	Electors		
	40	35	25
Preferences	A	B	C
	B	C	A
	C	A	B
Sincere ballot	**A**	B	C

Manipulability: an example

Voting system: plurality.

	Electors		
	40	35	25
Preferences	A	B	C
	B	C	A
	C	A	B
Sincere ballot	**A**	B	C
Manipulation	A	C	**C**

We say that this situation is **manipulable** for this voting system:

- A subset of electors, by casting a tactical ballot, may change the result to a candidate that they prefer.
- I.e., sincere voting is not a **strong Nash equilibrium**.