Convergence In Iterative Voting

Panagiotis Grontas

Theoretical Computer Science II,

24.07.2014

Introduction

* Multi Agent Systems
  * Select course of action
  * How: Aggregate the preferences of individual agents by voting
  * Social choice theory
* Problems:
  * All voting rules **can** be manipulated (Gibbard - Satterthwaite)
  * Strategic Behavior: Game theory to the rescue
  * But …
  * Distorted equilibria
    * The result cannot be changed
    * But, there might be individual unhappy voters
    * **Example:** Everybody votes for the same candidate
  * Cannot force agents honest preferences

Introduction

- A solution:
  - Manipulate the manipulation!
  - How: Use it to **converge** to a stable state, where no agent **wants** to further manipulate the game
  - Stable state: Nash Equilibrium
- Iterative voting (**Meir, 2010**)
  - All agents vote and view the result
  - Individual Preferences are not revealed
  - Unhappy agents can change their votes
  - Repeat until everybody is happy

Results Summary

- Under Restrictions
  - Plurality Voting Converges
  - Antiplurality Converges
- Other voting rules do not converge, regardless of restrictions
  - Borda
  - K-approval
- Restrictions:
  - Initial Vote (truthful or not)
  - Voter weights
  - Type of agent action (best/better reply)
  - Tie breaking rules

The model

* V: set of n voters/players
  * $|V| = n$
  * k strategic voters
  * $n - k$ truthful
* C: set of m candidates
  * $|C| = m$
  * Score at time t: $\boldsymbol{s_t}$(c)
* Truthful voters forgotten after voting
  * Initial score: $\hat{\boldsymbol{s}}_{\boldsymbol{t}}$(c)
* Strategic Voters
  * Myopic Greedy Moves
  * One Change At A Time
* Winners at time t
  * $\boldsymbol{o_t}$
* Potential Next Winners
  * A voter can make them win in the next move
  * $\boldsymbol{W_t}$

The model – (2)

* Voting rule $f: A^n \to 2^{C^*}$
  * A: all preference orders for C (real/announced)
  * Scoring Rules: $(a_1, a_2, \cdots, a_{m-1}, 0)$
  * $a_i \geq a_{i+1}$
  * Examples:
    * **Plurality:** $(1,0, \cdots, 0,0)$
    * **Veto**: $(1,1, \cdots, 1,0)$ or
      * $(0,0, \cdots, 0, -1)$
    * **Borda:**$(m-1, m-2, \cdots, 1,0)$
    * **K-Approval:** $(1,1, \cdots, 0,0)$

* **Maximin**
  * $\forall\, x, y \in C$
  * $\mathrm{N}(x, y) = |\{ i: x \succ_i y\}|$
  * $S(x) = \min_y \{N(x, y)\}$
  * Winner: $max_x S$

The model – (3)

* Tie breaking:
  * t: $2^{C^*} \rightarrow$ C
  * Alphabetical Tie – Breaking
    * $\boldsymbol{O_t}$: Singleton
  * Randomized Tie Breaking
    * Sets of possible winners
  * Linear Order Tie Breaking
    * If $\forall\ x, y \ \in D \ \subseteq C: t(D) = x$
    * and $x, y \ \in D'$
    * Then $t(D') \neq y$

Non convergence - example

* $|V| = 9,\ k=3$
* $C = \{a, b, c, d\}$
* Initial score (2,2,2,0)
* Plurality with Alphabetical Tie Breaking
* Preference Profile 1: $d \succ_1 a \succ_1 b \succ_1 c$
* Preference Profile 2: $c \succ_2 b \succ_2 a \succ_2 d$
* Preference Profile 3: $d \succ_3 a \succ_3 b \succ_3 c$
* Initial Move (truthful): $(d, c, d)\{c\}$ (2,2,3,2)

$$(\boldsymbol{b}, c, d)\{b\} \xrightarrow[(2,3,3,1)]{} (b, c, \boldsymbol{a})\{a\} \xrightarrow[(3,3,3,0)]{} (b, \boldsymbol{b}, a)\{b\} \xrightarrow[(3,4,2,0)]{}$$

$$(\boldsymbol{c}, b, a)\{a\} \xrightarrow[(3,3,3,0)]{} (c, \boldsymbol{c}, a)\{c\} \xrightarrow[(3,2,4,0)]{} (\boldsymbol{b}, c, a)\{a\}$$

Plurality Convergence

Theorem

Plurality with **alphabetical tie breaking** converges to a Nash Equilibrium from **any starting state** in at most ***m+(m-1)n*** steps, if all players respond with **restricted best replies**

Improvement Steps

At time t a player responds with $a \rightarrow b$

- **Type 1:** $\boldsymbol{a \neq o_{t-1}}$ **and** $\boldsymbol{b = o_t}$
- **Type 2:** $\boldsymbol{a = o_{t-1}}$ **and** $\boldsymbol{b \neq o_t}$
- **Type 3:** $\boldsymbol{a = o_{t-1}}$ **and** $\boldsymbol{b = o_t}$

Note: A type 4 move ($\boldsymbol{a \neq o_{t-1}}$ **and** $\boldsymbol{b \neq o_t}$) has no impact on the winner set

Restricted Best Reply: Type 1 and Type 3 improvement steps,

Improvement Steps Practice

* $|V| = 3, k = 2$
* $C = \{a, b, c\}$
* Initial score (1,0,0)
* Alphabetical Tie Breaking
* Preference Profile 1: $a \succ_1 \; b \succ_1 \; c$
* Preference Profile 2: $c \succ_2 \; b \succ_2 \; a$
* $(b, c)\{a\} \xrightarrow[\substack{c \to b \\ Type\ 1}]{} (b, \boldsymbol{b})\{b\} \xrightarrow[\substack{b \to \mathbf{C} \\ Type\ 2}]{} (\boldsymbol{c}, b)\{a\}$
* $\xrightarrow[\substack{b \to \mathbf{C} \\ Type\ 1}]{} (c, \boldsymbol{c})\{c\} \xrightarrow[\substack{c \to a \\ Type\ 3}]{} (\boldsymbol{a}, c)\{a\}$

An observation

- Simultaneous responses might not converge
  - $C = \{a, b, c\}$
  - $|V| = 4$, $k=2$
  - Initial score (0,0,2)
  - Preference Profile 1: $a \succ_1 b \succ_1 c$
  - Preference Profile 2: $b \succ_2 a \succ_2 c$
  - $(\boldsymbol{a}, \boldsymbol{b})\{c\} \rightarrow (\boldsymbol{b}, \boldsymbol{a})\{c\} \rightarrow (\boldsymbol{a}, \boldsymbol{b})\{c\}$

Plurality Convergence Proof (1)

Lemma 1

The set of **potential** next winners **never increases** over time

$$\forall\, t < t' : W_{t'} \subseteq W_t$$

Proof

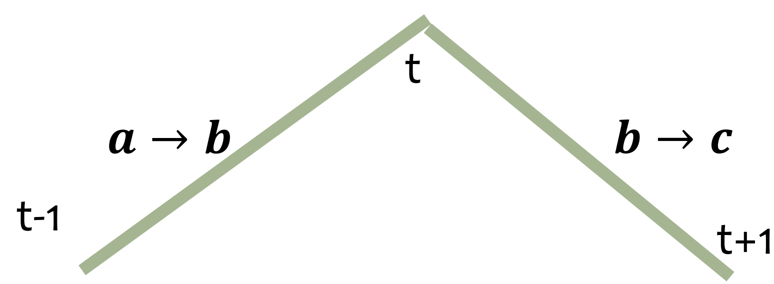
All these take place at time t+1

- $\forall\, t$: $c \in W_t \Rightarrow c \in W_{t-1}$
- Assume $c \in W_t$:
  - An RBR step $a \rightarrow c$ yields $O_{t+1} = c$
- Step increases c's score:
  - $s_{t+1}(c) = s_t(c) + 1$
- By winner definition: $\forall y \in C$
  - $s_{t+1}(c) \geq s_{t+1}(y) \implies s_t(c) + 1 \geq s_{t+1}(y)$

Plurality Convergence Proof (2)

At time t, an improvement step $a \rightarrow b$ occurs

* C maintains score at time t
* $s_t(c) = s_{t-1}(c)$
* $b = o_t$ (from RBR def)
* **c can beat b at time t**
  * $s_t(c) + 1 \geq s_{t+1}(b)$
  * $s_{t-1}(c) + 1 \geq s_{t+1}(b)$ $s_{t+1}(b) = s_t(b) - 1 = s_{t-1}(b)$ - b fell back to the same score



Plurality Convergence Proof (3)

* **c can beat a at time t**
  * **$a \rightarrow b$ is a type 3 move**
    * $a = o_{t-1}$
    * Moves are: $a \rightarrow b$ *followed by* $b \rightarrow c$
    * Replace with $a \rightarrow c$
    * Scores at time t+1 = scores at time t
    * As a result c can win at time t

Plurality Convergence Proof (4)

- **$a \rightarrow b$ is a type 1 move**
  - $a \neq o_{t-1}$ and $b = o_t$
  - This means: $\exists a', a' = o_{t-1}, a' \neq a$
  - $s_{t-1}(a') \geq s_{t-1}(y)\ \forall y \in C$
  - $s_{t-1}(a') = s_t(a') = s_{t+1}(a')$
  - $s_{t-1}(c) + 1 = s_t(c) + 1 \geq s_t(a') = s_{t-1}(a') \geq s_{t-1}(a)$
- **c can beat all the rest at time t**
  - **The rest are unaffected** $s_{t-1}(c) + 1 \geq s_{t-1}(y)$
- **As a result** $c \in W_{t-1}$

Plurality Convergence Proof (5)

Lemma 2

- There are at most *m-1* type 3 moves *for each* voter
- There are at most *m* type 1 moves

Proof

- A step of type 3, $a \rightarrow b$, is an improvement step
- So for each voter: b $\succ_i$ a
- There are *m-1* such possible improvements

***

- For all type 1 moves $a \rightarrow b, a \notin W_t$
- If $a \in W_t$ then $b \rightarrow a$ would make a win at time $t + 1$
- But $a \rightarrow b \rightarrow a$ results in a not winning (type 1 def)
- Conclusion: Type 1 moves only decrease the number of winners
- At most m such moves

Total: $\boldsymbol{n(m-1)+m}$

Anti-Plurality Convergence

Theorem

Anti-Plurality with **alphabetical tie breaking** converges to a Nash Equilibrium from any starting state in at most $\boldsymbol{mn}$ steps, if all players respond with **restricted best replies**

Improvement Steps

At time t a player responds with $-a \rightarrow -b$

- **Type 1:** $\boldsymbol{a \neq o_t}$ **and** $\boldsymbol{b = o_{t-1}}$
- **Type 2:** $\boldsymbol{a = o_t}$ **and** $\boldsymbol{b \neq o_{t-1}}$
- **Type 3:** $\boldsymbol{a = o_t}$ **and** $\boldsymbol{b = o_{t-1}}$

Note: A type 4 move ($\boldsymbol{a \neq o_{t-1}}$ **and** $\boldsymbol{b \neq o_t}$) has no impact on the winner set

Restricted Best Reply: Type 1 and Type 3 improvement steps (Veto the previous winner)

Proof

Lemma 1

A type 2 step $-a \rightarrow -c$ can be replaced with a type 3 step $-a \rightarrow -b$ without changing the winner but increasing the margin of victory of a

Proof

- A step of type 2, $-a \rightarrow -c$ results in:
  - a being a winner
  - Increasing $s_t(a)$
  - Decreasing $s_t(c)$
- By type 2 definition c is not a previous winner
- If we replace with type 3 we get:
  - a being a winner (winner set is the same)
  - Increasing $s_t(a)$
  - Decreasing $s_t(b)$ where b the previous winner

Proof (2)

Lemma 2

Over time the **set of potential next winners** never **decreases**

$$\forall\, t < t' : W_t \subseteq W_{t'}$$

Proof

- $\forall\, t$: $c \in W_{t-1} \Rightarrow c \in W_t$
- At time t improvement step $-a \rightarrow -b$ occurs
- Prove that $s_t(c) + 1 \geq s_t(y)\ \forall y \in C$
- Three cases to consider: **a, b, rest**
- **Rest**
  - $s_{t-1}(c) + 1 \geq s_{t-1}(y)$ since $c \in W_{t-1}$
  - But nobody changed
  - $s_t(c) + 1 \geq s_t(y)$
  - $c \in W_t$

Proof (3)

- **For b = $\boldsymbol{o_{t-1}}$**
  - $s_{t-1}(c) + 1 \succeq s_{t-1}(b)$
  - b was vetoed $\rightarrow$ score decreased $s_{t-1}(b) > s_t(b)$
  - c was unaffected
  - $s_t(c) + 1 \succeq s_t(b)$
  - $c \in W_t$
- **For $\boldsymbol{a}$**
  - **If** $-a \rightarrow -b$ **Type 3 step:**
    - $-c \rightarrow -b$ at time t is equivalent in score with
    - $-a \rightarrow -b$ at time t
    - $-c \rightarrow -a$ at time $t + 1$
  - **As a result nothing changes for c, b**
  - $c \in W_t$

Proof (4)

- **If $-a \to -b$ Type 1 step:**
  - $O_t = b'$, $\mathrm{b}' \neq a, b$ (somebody else was the winner)
  - This means: $s_t(b') \succeq s_t(a)$
  - $s_{t-1}(b')$=$s_t(b')$
  - $s_t(c) + 1 = s_{t-1}(c) + 1 \succeq s_{t-1}(b') = s_t(b') \succeq s_t(a)$
  - $c \in W_t$

Proof (5)

Lemma 3

- Each voter has at most 1 type 1 moves
- Each voter has at most $m - 1$ type 3 moves

Proof

- A step of type 3, $-a \rightarrow -b$, is an improvement step
- So for each voter: $\alpha \succ_i \ b$
- There are *m-1* such possible improvements

***

- Consider $-a \rightarrow -b$ step of type 1 for a voter v
- It will be the first improvement step
- If not a was vetoed in the past
- This means that a was a winner sometime
- And before that a potential next winner …
- Since W_t does not ever decrease, a is still
- Then $-a \rightarrow -b$ makes him a winner. Not a type 1 move

Randomised Tie Breaking

- Sets of current winners
- How to define best response?
- Approaches:
  - Cardinal Utilities
    - Consistent with preference ordering
    - Convergence from truthful state
    - Not guaranteed from arbitrary state
  - Stochastic Dominance
    - $W \succeq_v W'$ if the probability of selecting the k preferred candidates with winner set W is no less than with W'
    - Convergence

Borda and Best Replies

- May not converge even with alphabetical tie breaking
- **Counterexample**
- $|V| = 2, k = 2$
- $C = \{a, b, c\}$
- Preference Profile 1: $a \succ_1 \; b \succ_1 \; c$
- Preference Profile 2: $b \succ_2 \; c \succ_2 \; a$
- Moves:
- $(abc, bca)\{b\} \rightarrow (\boldsymbol{acb}, bca)\{a\} \rightarrow (acb, \boldsymbol{cba})\{c\} \rightarrow (\boldsymbol{abc}, cba)\{a\} \rightarrow (abc, \boldsymbol{bca})\{b\} \rightarrow \ldots$
- Order of players matters
- May not converge even without tie breaking

2-approval and Best Replies

- May not converge even with alphabetical tie breaking
- **Counterexample**
- $|V| = 2, k = 2$
- $C = \{a, b, c, d\}$
- Preference Profile 1: $a \succ_1 \; c \succ_1 \; d \succ_1 \; b$
- Preference Profile 2: $d \succ_1 \; b \succ_1 \; c \succ_1 \; a$
- <u>Moves:</u>
- $(ac, db)\{a\} \rightarrow (ac, \boldsymbol{dc})\{c\} \rightarrow (\boldsymbol{ab}, dc)\{a\} \rightarrow (ab, \boldsymbol{db})\{b\} \rightarrow$ $(\boldsymbol{ac}, db)\{a\} \rightarrow \ldots$

Maximin – (1)

- **May not converge**
- **Counterexample**
- $|V| = 2, k = 2$
- $C = \{a, b, c, d\}$
- Preference Profile 1:
  - $c \succ_1 \ d \succ_1 \ b \succ_1 \ a$
- Preference Profile 2:
  - $b \succ_1 \ d \succ_1 \ c \succ_1 \ a$

- Tie breaking rule:
  - if
    - b = c = d → b
    - b = c → c
    - a = b = c → b
    - a = b = c =d → a
    - c = d → c
    - b = d → b
  - Else
    - a

Maximin – (2)

* Truthful start
  * V1: $c \succ_1 d \succ_1 b \succ_1 a$
  * V2: $b \succ_1 d \succ_1 c \succ_1 a$
  * **S (a,b,c,d) = (0,1,1,1) {b}**
  * V1: $\boldsymbol{c \succ_1 b \succ_1 d \succ_1 a}$
  * V2: $b \succ_1 d \succ_1 c \succ_1 a$
  * **S (a,b,c,d) = (0,1,1,0) {c}**
  * V1: $c \succ_1 b \succ_1 d \succ_1 a$
  * V2: $\boldsymbol{a \succ_1 b \succ_1 d \succ_1 c}$
  * **S (a,b,c,d) = (1,1,1,0) {b}**
* V1: $\boldsymbol{c \succ_1 d \succ_1 b \succ_1 a}$
* V2: $a \succ_1 b \succ_1 d \succ_1 c$
* **S (a,b,c,d) = (1,1,1,1) {c}**
* V1: $c \succ_1 d \succ_1 b \succ_1 a$
* V2: $b \succ_1 d \succ_1 c \succ_1 a$
* **S (a,b,c,d) = (0,1,1,1) {b}**
* **…**

Open Problems

* Reply definition for convergence
* General convergence conditions applicable to different social choice functions

References

1. R. Meir, M. Polukarov, J. S. Rosenschein, and N. R. Jennings. **Convergence to equilibria of plurality voting.** In AAAI, pages 823–828, 2010
2. R. Reyhani and M. Wilson. **Best reply dynamics for scoring rules.** In ECAI, volume 2, pages 672–677, 2012.
3. O. Lev and J. S. Rosenschein. **Convergence of iterative voting**. In AAMAS, volume 2, pages 611–618, 2012
4. Obraztsova, S., Markakis, E., Polukarov, M., Rabinovich, Z., Jennings, N. R. (2014) **On the Convergence of Iterative Voting: How Restrictive should Restricted dynamics be?** In: *5th International Workshop on Computational Social Choice (ComSoc)*