CSC2556

Lecture 3

Approaches to Voting

Credit for several visuals: Ariel D. Procaccia
Announcement

• No class next week (1/30)

• Please use this time to work on the homework.
  ➢ I’ll post the full homework 1 by this weekend.

• You can also start thinking about the project idea!
Approaches to Voting

• What does an approach give us?
  ➢ A way to compare voting rules
  ➢ Hopefully a “uniquely optimal voting rule”

• Axiomatic Approach
• Distance Rationalizability
• Statistical Approach
• Utilitarian Approach
• ...
Axiomatic Approach

• Axiom: requirement that the voting rule should behave in a certain way

• Goal: define a set of reasonable axioms, and search for voting rules that satisfy them together
  - Ultimate hope: a unique voting rule satisfies the set of axioms simultaneously!
  - What often happens: no voting rule satisfies the axioms together 😞
Axiomatic Approach

• Weak axioms, satisfied by all popular voting rules

• **Unanimity:** If all voters have the same top choice, that alternative is the winner.

\[(\text{top}(>_{i}) = a \ \forall \ i \in N) \Rightarrow f(\succ) = a\]

➢ An even weaker version requires all rankings to be identical

• **Pareto optimality:** If all voters prefer \(a\) to \(b\), then \(b\) is not the winner.

\[(a >_{i} b \ \forall \ i \in N) \Rightarrow f(\succ) \neq b\]

• **Q:** What is the relation between these axioms?

➢ **Pareto optimality \(\Rightarrow\) Unanimity**
Axiomatic Approach

• **Anonymity:** Permuting votes does not change the winner (i.e., voter identities don’t matter).
  - E.g., these two profiles must have the same winner:
    - \{voter 1: \(a > b > c\), voter 2: \(b > c > a\)\}
    - \{voter 1: \(b > c > a\), voter 2: \(a > b > c\)\}

• **Neutrality:** Permuting alternative names just permutes the winner.
  - E.g., say \(a\) wins on \{voter 1: \(a > b > c\), voter 2: \(b > c > a\)\}
  - We permute all names: \(a \rightarrow b, b \rightarrow c,\) and \(c \rightarrow a\)
  - New profile: \{voter 1: \(b > c > a\), voter 2: \(c > a > b\)\}
  - Then, the new winner must be \(b\).
Axiomatic Approach

• Neutrality is tricky
  ➢ For deterministic rules, it is inconsistent with anonymity!
    o Imagine \{voter 1: a > b, voter 2: b > a\}
    o Without loss of generality, say \(a\) wins
    o Imagine a different profile: \{voter 1: b > a, voter 2: a > b\}
      • Neutrality: We just exchanged \(a \leftrightarrow b\), so winner is \(b\).
      • Anonymity: We just exchanged the votes, so winner stays \(a\).
  ➢ Typically, we only require neutrality for...
    o Randomized rules: E.g., a rule could satisfy both by choosing \(a\) and \(b\) as the winner with probability \(\frac{1}{2}\) each, on both profiles
    o Deterministic rules that return a set of tied winners: E.g., a rule could return \(\{a, b\}\) as tied winners on both profiles.
Axiomatic Approach

• Stronger but more subjective axioms

• **Majority consistency:** If a majority of voters have the same top choice, that alternative wins.
  \[ |\{i: \text{top}(>_i) = a\}| > \frac{n}{2} \Rightarrow f(\triangleright) = a \]

• **Condorcet consistency:** If \( a \) defeats every other alternative in a pairwise election, \( a \) wins.
  \[ |\{i: a >_i b\}| > \frac{n}{2}, \forall b \neq a \Rightarrow f(\triangleright) = a \]
Axiomatic Approach

• **Recall:** Condorcet consistency $\Rightarrow$ Majority consistency

• All positional scoring rules violate Condorcet consistency.

• Most positional scoring rules also violate majority consistency.
  - Plurality satisfies majority consistency.
Axiomatic Approach

• Consistency: If \( a \) is the winner on two profiles, it must be the winner on their union.

\[
f(\succeq_1) = a \land f(\succeq_2) = a \Rightarrow f(\succeq_1 + \succeq_2) = a
\]

➢ Example: \( \succeq_1 = \{a > b > c\}, \ \succeq_2 = \{a > c > b, b > c > a\} \)
➢ Then, \( \succeq_1 + \succeq_2 = \{a > b > c, a > c > b, b > c > a\} \)

• Theorem [Young ’75]:

➢ Subject to mild requirements, a voting rule is consistent if and only if it is a positional scoring rule!
Axiomatic Approach

• **Weak monotonicity:** If $a$ is the winner, and $a$ is “pushed up” in some votes, $a$ remains the winner.

  $f(\succ) = a \Rightarrow f(\succ') = a$, where
  
  $b >_i c \iff b >'_i c, \forall i \in N, b, c \in A\{a\}$ (Order of others preserved)
  $a >_i b \Rightarrow a >'_i b, \forall i \in N, b \in A\{a\}$ (a only improves)

• In contrast, **strong monotonicity** requires $f(\succ') = a$ even if $\succ'$ only satisfies the 2$^{nd}$ condition

  $f(\succ') = a$ even if $\succ'$ only satisfies the 2$^{nd}$ condition
  
  Too strong; only satisfied by dictatorial or non-onto rules

[GS Theorem]
Axiomatic Approach

• **Weak monotonicity:** If $a$ is the winner, and $a$ is “pushed up” in some votes, $a$ remains the winner.
  
  ➢ $f(\succ) = a \Rightarrow f(\succ') = a$, where
    - $b \succ_i c \iff b \succ'_i c$, $\forall i \in N$, $b, c \in A\setminus\{a\}$ (Order of others preserved)
    - $a \succ_i b \Rightarrow a \succ'_i b$, $\forall i \in N$, $b \in A\setminus\{a\}$ (a only improves)

• Weak monotonicity is satisfied by most voting rules
  
  ➢ Popular exceptions: STV, plurality with runoff
  ➢ But this helps STV be hard to manipulate
    - Theorem [Conitzer-Sandholm ‘06]: “Every weakly monotonic voting rule is easy to manipulate on average.”
### Axiomatic Approach

- STV violates weak monotonicity

<table>
<thead>
<tr>
<th>7 voters</th>
<th>5 voters</th>
<th>2 voters</th>
<th>6 voters</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>b</td>
<td>b</td>
<td>c</td>
</tr>
<tr>
<td>b</td>
<td>c</td>
<td>c</td>
<td>a</td>
</tr>
<tr>
<td>c</td>
<td>a</td>
<td>a</td>
<td>b</td>
</tr>
</tbody>
</table>

- First \( c \), then \( b \) eliminated
- Winner: \( a \)

<table>
<thead>
<tr>
<th>7 voters</th>
<th>5 voters</th>
<th>2 voters</th>
<th>6 voters</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>b</td>
<td>a</td>
<td>c</td>
</tr>
<tr>
<td>b</td>
<td>c</td>
<td>b</td>
<td>a</td>
</tr>
<tr>
<td>c</td>
<td>a</td>
<td>c</td>
<td>b</td>
</tr>
</tbody>
</table>

- First \( b \), then \( a \) eliminated
- Winner: \( c \)
Axiomatic Approach

• **Pareto optimality:** If $a \succ_i b$ for all voters $i$, then $f(\succ) \neq b$.

• **Relatively weak requirement**
  - Some rules that throw out alternatives early may violate this.
  - Example: voting trees
    - Alternatives move up by defeating opponent in pairwise election
    - $d$ may win even if all voters prefer $b$ to $d$ if $b$ loses to $e$ early, and $e$ loses to $c$
Axiomatic Approach

• Arrow’s Impossibility Theorem
  ➢ Applies to social welfare functions (profile → ranking)
  ➢ Independence of Irrelevant Alternatives (IIA): If the preferences of all voters between $a$ and $b$ are unchanged, the social preference between $a$ and $b$ should not change
  ➢ Pareto optimality: If all prefer $a$ to $b$, then the social preference should be $a > b$
  ➢ Theorem: IIA + Pareto optimality $\Rightarrow$ dictatorship.

• Interestingly, automated theorem provers can also prove Arrow’s and GS impossibilities!
Axiomatic Approach

• One can think of polynomial time computability as an axiom

- Two rules that attempt to make the pairwise comparison graph acyclic are NP-hard to compute:
  - Kemeny’s rule: invert edges with minimum total weight
  - Slater’s rule: invert minimum number of edges

- Both rules can be implemented by straightforward integer linear programs
  - For small instances (say, up to 20 alternatives), NP-hardness isn’t a practical concern.
Statistical Approach

• According to Condorcet [1785]:
  ➢ The purpose of voting is not merely to balance subjective opinions; it is a collective quest for the truth.
  ➢ Enlightened voters try to judge which alternative best serves society.

• Modern motivation due to human computation systems
  ➢ EteRNA: Select 8 RNA designs to synthesize so that the truly most stable design is likely one of them
Statistical Approach

• Traditionally well-explored for choosing a ranking

• For $m = 2$, the majority choice is most likely the true choice under any reasonable model.

• For $m \geq 3$: Condorcet suggested an approach, but the writing was too ambiguous to derive a well-defined voting rule.
Statistical Approach

• Young’s interpretation of Condorcet’s approach:
  ➢ Assume there is a ground truth ranking $\sigma^*$
  ➢ Each voter $i$ makes a noisy observation $\sigma_i$
  ➢ The observations are i.i.d. given the ground truth
    o $\Pr[\sigma|\sigma^*] \propto \phi^d(\sigma, \sigma^*)$
    o $d =$ Kendall-tau distance = #pairwise disagreements
    o Interesting tidbit: Normalization constant is independent of $\sigma^*$
      $$\Sigma_\sigma \phi^d(\sigma, \sigma^*) = 1 \cdot (1 + \phi) \cdot \ldots \cdot (1 + \varphi + \ldots + \varphi^{m-1})$$
  ➢ Which ranking is most likely to be the ground truth (maximum likelihood estimate – MLE)?
    o The ranking that Kemeny’s rule returns!
Statistical Approach

• The approach yields a uniquely optimal voting rule, but relies on a very specific distribution
  ➢ Other distributions will lead to different MLE rankings.
  ➢ Reasonable if sufficient data is available to estimate the distribution well
  ➢ Else, we may want robustness to a wide family of possible underlying distributions [Caragiannis et al. ’13, ’14]

• A connection to the axiomatic approach
  ➢ A voting rule can be MLE for some distribution only if it satisfies consistency. (Why?)
    o Maximin violates consistency, and therefore can never be MLE!
Implicit Utilitarian Approach

• **Utilities:** Voters have underlying numerical utilities
  - Utility of voter $i$ for alternative $a = u_i(a)$
    - Normalization: $\sum_a u_i(a) = 1$ for all voters $i$
  - Given utility vector $\vec{u}$, $sw(a, \vec{u}) = \sum_i u_i(a)$
  - **Goal:** choose $a^* \in \text{argmax}_a sw(a, \vec{u})$

• **Preferences:** Voters only report ranked preferences consistent with their utilities
  - $u_i(a) > u_i(b) \Rightarrow a >_i b$
  - Preference profile: $\succeq$
  - Cannot maximize welfare given only partial information
Implicit Utilitarian Approach

• **Modified goal:** Achieve the best worst-case approximation to social welfare

• **Distortion** of voting rule $f$

$$\max_{\vec{u}} \frac{\max_a \text{sw}(a, \vec{u})}{\text{sw}(f(\rightarrow), \vec{u})}$$

- Here, $\rightarrow$ are the preferences cast by voters when their utilities are $\vec{u}$
- If $f$ is randomized, we need $E[\text{sw}(f(\rightarrow), \vec{u})]$
Utilitarian Approach

• Pros:
  ➢ Uses minimal subjective assumptions
  ➢ Yields a uniquely optimal voting rule
    o One can define the distortion of $f$ on a given input $\succ$ by taking the worst case over all $\tilde{u}$ which would generate $\tilde{\succ}$
    o Optimal voting rule minimizes the distortion on every $\tilde{\succ}$ individually

• Cons:
  ➢ The optimal rule does not have an intuitive formula that humans can comprehend
  ➢ In some scenarios, the optimal rule is difficult to compute
Choosing One Alternative

• **Theorem** [Caragiannis et al. ’16]: Given ranked preferences, the optimal deterministic voting rule has $\Theta(m^2)$ distortion.

• **Proof:**
  - **Lower bound:** Construct a profile on which every deterministic voting rule has $\Omega(m^2)$ distortion.
  - **Upper bound:** Show *some* deterministic voting rule that has $O(m^2)$ distortion on every profile.
Choosing One Alternative

**Proof (lower bound):**

- Consider the profile on the right
- If the rule chooses $a_m$:
  - Infinite distortion. **WHY?**
- If the rule chooses $a_i$ for $i < m$:
  - Construct a bad utility profile $\tilde{u}$ as follows
    - Voters in column $i$ have utility $1/m$ for every alternative
    - All other voters have utility $1/2$ for their top two alternatives
  - $\text{sw}(a_i, \tilde{u}) = \frac{n}{m-1} \cdot \frac{1}{m}$, $\text{sw}(a_m, \tilde{u}) \geq \frac{n-n/(m-1)}{2}$
  - Distortion $= \Omega(m^2)$
Choosing One Alternative

• Proof (upper bound):
  ➢ Simply using plurality achieves $O(m^2)$ distortion.
    o WHY?

  ➢ Suppose plurality winner is $a$.
    o At least $n/m$ voters prefer $a$ the most, and thus have utility at least $1/m$ for $a$.

  ➢ $sw(a, \vec{u}) \geq n/m^2$
  ➢ $sw(a^*, \vec{u}) \leq n$ for every alternative $a^*$
  ➢ $O(m^2)$ distortion
Implicit Utilitarian Voting

• Plurality is as good as any other deterministic voting rule!

• Alternatively:
  ➢ If we must choose an alternative deterministically, ranked preferences provide no more useful information than top-place votes do, in the worst case.

• There’s more hope if we’re allowed to randomize.
Choosing One Alternative

• **Theorem** [Boutilier et al. ‘12]: Given ranked preferences, the optimal randomized voting rule has distortion $O(\sqrt{m} \cdot \log^* m)$, $\Omega(\sqrt{m})$.

• **Proof:**
  - **Lower bound:** Construct a profile on which every randomized voting rule $\Omega(\sqrt{m})$ distortion.
  - **Upper bound:** Show *some* randomized voting rule that has $O(\sqrt{m} \cdot \log^* m)$ distortion
    - We’ll do the much simpler $O(\sqrt{m \log m})$ distortion
Choosing One Alternative

- **Proof (lower bound):**
  - Consider a similar profile:
    - \( \sqrt{m} \) special alternatives
    - Voting rule must choose one of them (say \( a^* \)) w.p. at most \( 1/\sqrt{m} \)

- **Bad utility profile \( \vec{u} \):**
  - All voters ranking \( a^* \) first give utility 1 to \( a^* \)
  - All other voters give utility \( 1/m \) to each alternative
  - \( \frac{n}{\sqrt{m}} \leq sw(a^*, \vec{u}) \leq \frac{2n}{\sqrt{m}} \)
  - \( sw(a, \vec{u}) \leq n/m \) for every other \( a \).
  - Distortion lower bound: \( \sqrt{m}/3 \) (proof on the board!)
Choosing One Alternative

• Proof (upper bound):
  
  ➢ Given profile \( \succ \), define the harmonic score \( sc(a, \succ) \):
    
    o Each voter gives \( 1/k \) points to her \( k^{th} \) most preferred alternative
    o Take the sum of points across voters
    o \( sw(a, \vec{u}) \leq sc(a, \succ) \) (WHY?)
    o \( \sum_a sc(a, \succ) = n \cdot \sum_{k=1}^{m} 1/k = n \cdot H_m \leq n \cdot (\ln m + 1) \)

  ➢ Golden rule:
    
    o W.p. \( \frac{1}{2} \): Choose every \( a \) w.p. proportional to \( sc(a, \succ) \)
    o W.p. \( \frac{1}{2} \): Choose every \( a \) w.p. \( 1/m \) (uniformly at random)

  ➢ Distortion \( \leq 2 \sqrt{m \cdot (\ln m + 1)} \) (proof on the board!)
Optimal vs Near-Optimal Rules

• The distortion is often bad for large $m$
  - E.g., $\Theta(m^2)$ for deterministic rules.
  - But one can argue that the optimal alternative which minimizes distortion represents some meaningful aggregation of information.

• How difficult is it to find the optimal alternative?
  - Polynomial time computable for both deterministic (via a direct formula) and randomized (via a non-trivial LP) cases
Input Format

• What if we ask about underlying numerical utilities in a format other than ranking?

• Threshold approval votes
  ➢ Voting rule selects a threshold $\tau$, asks each voter $i$, for each alternative $a$, whether $u_i(a) \geq \tau$
  ➢ $O(\log m)$ distortion!

• Food for thought
  ➢ What is the tradeoff between the number of bits of information elicited and the distortion achieved?
  ➢ What is the best input format for a given number of bits?
Implicit Utilitarian Approach

• Extensions
  ➢ Selecting a subset of alternatives or a ranking
    o Lack of an obvious objective function
    o Has been studied for some natural objective functions
      [Caragiannis et al. ’16, ongoing work]
  ➢ Participatory budgeting [Benade et al. ’17]
  ➢ Graph problems
  ➢ Project idea: Replace numbers with rankings in any problem!

• Deployed