Introduction to Social Choice

Lirong Xia

Rensselaer

Fall, 2016
Keep in mind

- Good science
  - What question does it answer?

- Good engineering
  - What problem does it solve?
Last class

How to model agents’ preferences?

Order theory

- linear orders
- weak orders
- partial orders

Utility theory

- preferences over lotteries
- risk attitudes: aversion, neutrality, seeking
Today

➢ Q: What problem does it solve?
➢ A: Aggregating agents’ preferences and make a joint decision by voting
Change the world: 2011 UK Referendum

- The second nationwide referendum in UK history
  - The first was in 1975
- Member of Parliament election:
  - Plurality rule ➔ Alternative vote rule
- 68% No vs. 32% Yes
- In 10/440 districts more voters said yes
  - 6 in London, Oxford, Cambridge, Edinburgh Central, and Glasgow Kelvin
- Why change?
- Why failed?
- Which voting rule is the best?
Social choice: Voting

Agents: \( n \) voters, \( N = \{1, \ldots, n\} \)

Alternatives: \( m \) candidates, \( A = \{a_1, \ldots, a_m\} \) or \( \{a, b, c, d, \ldots\} \)

Outcomes:
- winners (alternatives): \( O = A \). Social choice function
- rankings over alternatives: \( O = \text{Rankings}(A) \). Social welfare function

Preferences: \( R_j^* \) and \( R_j \) are full rankings over \( A \)

Voting rule: a function that maps each profile to an outcome
Popular voting rules

(a.k.a. what people have done in the past two centuries)
The Borda rule

\[ P = \{ \times 4, \times 3, \times 2, \times 2 \} \]

Borda \( P \) =

Borda scores:
\[ 2 \times 4 + 4 = 12 \]
\[ 2 \times 2 + 7 = 11 \]
\[ 2 \times 5 = 10 \]
Positional scoring rules

- Characterized by a score vector $s_1, \ldots, s_m$ in non-increasing order
- For each vote $R$, the alternative ranked in the $i$-th position gets $s_i$ points
- The alternative with the most total points is the winner
- Special cases
  - Borda: score vector $(m-1, m-2, \ldots, 0)$ [French academy of science 1784-1800, Slovenia, Naru]
  - $k$-approval: score vector $(1\ldots1, 0\ldots0)$
  - Plurality: score vector $(1, 0\ldots0)$ [UK, US]
  - Veto: score vector $(1\ldots1, 0)$
Example

\[ P = \{ \times 4, \times 3, \times 2, \times 2 \} \]

Borda

Plurality (1-approval)

Veto (2-approval)
Plurality with runoff

- The election has two rounds
  - First round, all alternatives except the two with the highest plurality scores drop out
  - Second round, the alternative preferred by more voters wins

- [used in France, Iran, North Carolina State]
Example: Plurality with runoff

\[ P = \{ \begin{array}{c}
\text{First round: drops out} \\
\text{Second round: defeats} \\
\text{Different from Plurality!}
\end{array} \]
Single transferable vote (STV)

- Also called instant run-off voting or alternative vote
- The election has $m-1$ rounds, in each round,
  - The alternative with the lowest plurality score drops out, and is removed from all votes
  - The last-remaining alternative is the winner
- [used in Australia and Ireland]

<table>
<thead>
<tr>
<th>Preference</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a &gt; b &gt; c &gt; d$</td>
<td>10</td>
</tr>
<tr>
<td>$d &gt; a &gt; b &gt; c$</td>
<td>7</td>
</tr>
<tr>
<td>$c &gt; d &gt; a &gt; b$</td>
<td>6</td>
</tr>
<tr>
<td>$b &gt; c &gt; d &gt; a$</td>
<td>3</td>
</tr>
</tbody>
</table>

\[ a \]
Other multi-round voting rules

➤ Baldwin’s rule
  • Borda+STV: in each round we eliminate one alternative with the lowest Borda score
  • break ties when necessary

➤ Nanson’s rule
  • Borda with multiple runoff: in each round we eliminate all alternatives whose Borda scores are below the average
  • [Marquette, Michigan, U. of Melbourne, U. of Adelaide]
The Copeland rule

- The **Copeland score** of an alternative is its total “pairwise wins”
  - the number of positive outgoing edges in the WMG
- The winner is the alternative with the highest Copeland score
- WMG-based
Example: Copeland

\[ P = \{ \begin{array}{ccc} \text{\textgreater} & \text{\textgreater} & \times 4 \\ \text{\textgreater} & \text{\textgreater} & \times 3 \\ \text{\textgreater} & \text{\textgreater} & \times 2 \\ \text{\textgreater} & \text{\textgreater} & \times 2 \end{array} \} \]

Copeland score:

\[ \begin{array}{ccc} \text{\textgreater} & : 2 \\ 1 & : 1 \\ \text{\textless} & : 0 \end{array} \]
The maximin rule

- A.k.a. Simpson or minimax
- The maximin score of an alternative $a$ is
  \[ MS_P(a) = \min_b (\# \{ a > b \text{ in } P \} - \# \{ b > a \text{ in } P \}) \]
  - the smallest pairwise defeats
- The winner is the alternative with the highest maximin score
- WMG-based
Example: maximin

\[ P = \{ (\text{\textbf{>}} \times 4, \text{\textbf{>}} \times 3), (\text{\textbf{>}} \times 2, \text{\textbf{>}} \times 2) \} \]

Maximin score:

\[ \text{\textbf{>}} : 1, \text{\textbf{>}} : -1, \text{\textbf{>}} : -1 \]
Ranked pairs

- Given the WMG
- Starting with an empty graph $G$, adding edges to $G$ in multiple rounds
  - In each round, choose the remaining edge with the highest weight
  - Add it to $G$ if this does not introduce cycles
  - Otherwise discard it
- The alternative at the top of $G$ is the winner
Example: ranked pairs

Q1: Is there always an alternative at the “top” of $G$?
Q2: Does it suffice to only consider positive edges?
The Schulze Rule

- In the WMG of a profile, the strength
  - of a path is the smallest weight on its edges
  - of a pair of alternatives \((a, b)\), denoted by \(S(a, b)\), is the largest strength of paths from \(a\) to \(b\)

- The Schulze winners are the alternatives \(a\) such that
  - for all alternatives \(a'\), \(S(a, a') \geq S(a', a)\)
  - \(S(a, b) = S(a, c) = S(a, d) = 6\rightarrow c\rightarrow b = 4\)
  - \(> 2 = S(b, a) = S(c, a) = S(d, a)\)
  - The (unique) winner is \(a\)
Ranked pairs and Schulze

- Ranked pairs [Tideman 1987] and Schulze [Schulze 1997]
  - Both satisfy anonymity, Condorcet consistency, monotonicity, immunity to clones, etc
  - Neither satisfy participation and consistency (these are not compatible with Condorcet consistency)

- Schulze rule has been used in elections at Wikimedia Foundation, the Pirate Party of Sweden and Germany, the Debian project, and the Gento Project
The Bucklin Rule

- An alternative $a$’s Bucklin score
  - smallest $k$ such that for the majority of agents, $a$ is ranked within top $k$

- Simplified Bucklin
  - Winners are the agents with the smallest Bucklin score
Kemeny’s rule

- Kendall tau distance
  - \( K(R,W) = \# \{ \text{different pairwise comparisons} \} \)

\[
K(\begin{array}{ccc}
  b & > & c \\
  c & > & a \\
  a & > & b \\
\end{array}, \begin{array}{ccc}
  a & > & b \\
  b & > & c \\
  c & > & a \\
\end{array}) = ?
\]

- Kemeny(\(D\)) = \( \arg\min_W K(D,W) = \arg\min_W \sum_{R \in D} K(R,W) \)

- For single winner, choose the top-ranked alternative in Kemeny(\(D\))

- [reveals the truth]
Weighted majority graph

Given a profile $P$, the weighted majority graph $\text{WMG}(P)$ is a weighted directed complete graph $(V,E,w)$ where

- $V = A$
- for every pair of alternatives $(a, b)$
  - $w(a \rightarrow b) = \#\{a > b \text{ in } P\} - \#\{b > a \text{ in } P\}$
  - $w(a \rightarrow b) = -w(b \rightarrow a)$
- $\text{WMG}$ (only showing positive edges) might be cyclic
  - Condorcet cycle: $\{a > b > c, b > c > a, c > a > b\}$
Example: WMG

\[ P = \{ \text{positive edges} \times 4, \text{positive edges} \times 2, \text{positive edges} \times 3 \} \]

\[ \text{WMG}(P) = \text{(positive edges)} \]
WMG-based voting rules

- A voting rule \( r \) is based on weighted majority graph, if for any profiles \( P_1, P_2 \),

\[
[ \text{WMG}(P_1) = \text{WMG}(P_2) ] \implies [ r(P_1) = r(P_2) ]
\]

- WMG-based rules can be redefined as a function that maps \{WMGs\} to \{outcomes\}

- **Example**: Borda is WMG-based
  - Proof: the Borda winner is the alternative with the highest sum over outgoing edges.
Voting with Prefpy

Implemented
- All positional scoring rules
- Bucklin, Copeland, maximin
- not well-tested for weak orders

Project ideas
- implementation of STV, ranked pairs, Kemeny
  - all are NP-hard to compute
- extends all rules to weak orders
Popular criteria for voting rules
(a.k.a. what people have done in the past 60 years)
How to evaluate and compare voting rules?

- No single numerical criteria
  - **Utilitarian**: the joint decision should maximize the total happiness of the agents
  - **Egalitarian**: the joint decision should maximize the worst agent’s happiness

- **Axioms**: properties that a “good” voting rules should satisfy
  - measures various aspects of preference aggregation
Fairness axioms

- **Anonymity**: names of the voters do not matter
  - Fairness for the voters

- **Non-dictatorship**: there is no dictator, whose top-ranked alternative is always the winner, no matter what the other votes are
  - Fairness for the voters

- **Neutrality**: names of the alternatives do not matter
  - Fairness for the alternatives
A truth-revealing axiom

- **Condorcet consistency**: Given a profile, if there exists a Condorcet winner, then it must win
  - The Condorcet winner beats all other alternatives in pairwise comparisons
  - The Condorcet winner only has positive outgoing edges in the WMG

- Why this is truth-revealing?
  - why Condorcet winner is the truth?
The Condorcet Jury theorem

[Condorcet 1785]

Given

- two alternatives \( \{a, b\} \). \( a \): liable, \( b \): not liable
- \( 0.5 < p < 1 \),

Suppose

- given the ground truth \((a\ or\ b)\), each voter’s preference is generated i.i.d., such that
  - w/p \( p \), the same as the ground truth
  - w/p \( 1-p \), different from the ground truth

Then, as \( n \to \infty \), the probability for the majority of agents’ preferences is the ground truth goes to 1

“lays, among other things, the foundations of the ideology of the democratic regime” (Paroush 1998)
Condorcet’s model

[Condorcet 1785]

• Given a “ground truth” ranking $W$ and $p > 1/2$, generate each pairwise comparison in $R$ independently as follows (suppose $c > d$ in $W$)

\[
\Pr( b > c > a \mid a > b > c ) = (1-p) p^2
\]

• Its MLE is Kemeny’s rule [Young JEP-95]
Truth revealing

Extended Condorcet Jury theorem

- **Given**
  - A ground truth ranking $W$
  - $0.5 < p < 1$

- **Suppose**
  - each agent’s preferences are generated i.i.d. according to Condorcet’s model

- Then, as $n \to \infty$, with probability that $\to 1$
  - the randomly generated profile has a Condorcet winner
  - The Condorcet winner is ranked at the top of $W$

- If $r$ satisfies Condorcet criterion, then as $n \to \infty$, $r$ will reveal the “correct” winner with probability that $\to 1$. 

Other axioms

- **Pareto optimality:** For any profile $D$, there is no alternative $c$ such that every voter prefers $c$ to $r(D)$

- **Consistency:** For any profiles $D_1$ and $D_2$, if $r(D_1) = r(D_2)$, then $r(D_1 \cup D_2) = r(D_1)$

- **Monotonicity:** For any profile $D_1$,
  - if we obtain $D_2$ by only raising the position of $r(D_1)$ in one vote,
  - then $r(D_1) = r(D_2)$
  - In other words, raising the position of the winner won’t hurt it
Which axiom is more important?

<table>
<thead>
<tr>
<th></th>
<th>Condorcet criterion</th>
<th>Consistency</th>
<th>Anonymity/neutrality, non-dictatorship, monotonicity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plurality</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>STV (alternative vote)</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
</tr>
</tbody>
</table>

- Some axioms are not compatible with others
- Which rule do you prefer?
An easy fact

- **Theorem.** For voting rules that selects a single winner, anonymity is not compatible with neutrality.
  - proof:

  - W.O.L.G.
  - Neutrality
  - Anonymity
Another easy fact [Fishburn APSR-74]

Theorem. No positional scoring rule satisfies Condorcet criterion:

- suppose $s_1 > s_2 > s_3$

3 Voters > >

2 Voters > >

1 Voter > >

1 Voter > >

is the Condorcet winner

CONTRADICTION

$3s_1 + 2s_2 + s_3$

$3s_1 + 3s_2 + 1s_3$
Arrow's impossibility theorem

- Recall: a social welfare function outputs a ranking over alternatives

- **Arrow's impossibility theorem.** No social welfare function satisfies the following four axioms
  - Non-dictatorship
  - Universal domain: agents can report any ranking
  - Unanimity: if \( a > b \) in all votes in \( D \), then \( a > b \) in \( r(D) \)
  - Independence of irrelevant alternatives (IIA): for two profiles \( D_1 = (R_1, \ldots, R_n) \) and \( D_2 = (R'_1, \ldots, R'_n) \) and any pair of alternatives \( a \) and \( b \)
    - if for all voter \( j \), the pairwise comparison between \( a \) and \( b \) in \( R_j \) is the same as that in \( R'_j \)
    - then the pairwise comparison between \( a \) and \( b \) are the same in \( r(D_1) \) as in \( r(D_2) \)
Other Not-So-Easy facts

- **Gibbard-Satterthwaite theorem**
  - Later in the “hard to manipulate” class

- **Axiomatic characterization**
  - Template: A voting rule satisfies axioms A1, A2, A2 if it is rule X
  - If you believe in A1 A2 A3 are the most desirable properties then X is optimal
  - (unrestricted domain+unanimity+IIA) dictatorships [Arrow]
  - (anonymity+neutrality+consistency+continuity) positional scoring rules [Young SIAMAM-75]
  - (neutrality+consistency+Condorocet consistency) Kemeny [Young&Levenglick SIAMAM-78]
Remembered all of these?

- Impressive! Now try a slightly larger tip of the iceberg at wiki
Change the world: 2011 UK Referendum

➢ The second nationwide referendum in UK history
  • The first was in 1975
➢ Member of Parliament election:
  Plurality rule ➔ Alternative vote rule
➢ 68% No vs. 32% Yes
➢ Why people want to change?
➢ Why it was not successful?
➢ Which voting rule is the best?
Wrap up

Voting rules

• positional scoring rules
• multi-round elimination rules
• WMG-based rules
• A Ground-truth revealing rule (Kemeny’s rule)

Criteria (axioms) for “good” rules

• Fairness axioms
• A ground-truth-revealing axiom (Condorcet consistency)
• Other axioms

Evaluation

• impossibility theorems
• Axiomatic characterization