MODELING DEMOCRACY NOTES

JALEN CHRYSOS

1. About the Class

- Professor is Moon Duchin, who we should address as Moon.
- This is an “interdisciplinary PhD class.”
- There will be Problem Sets. There will be Proofs.
- There is an option to give a presentation.[1] Past projects have turned into publications.
- No laptops (oops).

[1] “For those of you looking to do anything later in life, this is good practice.”

2. Voting Systems

A voting system takes as an input the ranked preferences of a population on some finite number of options and outputs a collective choice (either of a subset of those options or a ranking of them). Here are some examples[2]:

First, we have voting rules which only take into account the pairwise comparison graph (PWCG), a directed graph on the set of candidates with weights of edges corresponding to the margins of match-ups between any pair of candidates in isolation.

<u>PWCG-Based Rules</u>:

- *PWC*: Whoever wins most head-to-head match-ups wins.
- *Sequential*: A tournament structure is fixed ahead of time. In each match-up, the winner is determined by the ranked-choice ballots. The winner of the tournament is selected.
- *Smith*: The winner is selected (using some other procedure) from within the Smith Set; that is, the smallest subset of candidates S such that no candidate outside S beats any candidate in S.
- *Beatpath*:
  - A *beatpath* is a directed path in the PWCG.
  - The *strength* of a beatpath is the least margin of any link in the path.
  - A candidate A *dominates* B if the strongest beatpath from A to B is stronger than the strongest beatpath from B to A.
  - In the Beatpath voting rule, the winner is selected (somehow) from among the candidates which are not dominated by any other candidate.
- *Ranked Pairs*: Edges in the PWCG are sequentially "activated" in order of greatest margin first, skipping any that would produce a directed cycle. The resulting graph will have a Condorcet candidate within that graph, and that candidate is selected.

Next, we have methods which repeatedly eliminate candidates using a runoff system:

<u>Runoff-Based Rules:</u>

- *Top-Two*: The top two in terms of first-choice votes proceed to a second runoff election. This can be determined by the relative ranking of the top two in the profile. Or alternatively, in the real world there can be a second election later, allowing time for more campaigning.
- *IRV (Instant-Runoff Voting)*: The candidates with least first-choice votes are repeatedly eliminated and their votes redistributed until one remains (only one election).
- *STV (Single Transferable Vote)*: In a situation where K winners must be selected, select a threshold T (often $V/(K+1)$ where V is the number of voters) and run the following algorithm until K candidates are elected:
  (1) If any candidate has more than T first-choice votes:
    - The one with the most votes, A, is elected.
    - The voters who elected A "spend" the portion of their remaining ballot corresponding to the portion that was necessary to reach T, so that the weight of their ballot for the remainder of the decision procedure is multiplied by $(M-T)/M$ where M is the (weighted) total of first-choice ballots for A.
    - A is removed from the pool of candidates and all ballots are consolidated.
    - Repeat step 1.
  (2) Otherwise, if no candidate has T first-choice votes, remove the candidate with the least first-choice votes. Then try step 1 again.
- *Coombs*: Like IRV except that the candidate with most last-choice votes is repeatedly eliminated.

[2]Note that many of these only identify sets of winners, within which tiebreaks must be conducted.

Next we have voting rules that are metric-based. A metric on ballots can be defined as follows: let the distance between two rankings be the least number of elementary moves required to get from one ballot to the other, where an elementary move consists of either swapping two adjacent-ranking candidates or (if ballots are permitted to be incomplete) adding/removing a candidate in last place. Then the following methods are possible:

Metric-Based Rules:

- *Dodgson*: From any profile, there is a least distance to an alternate profile where candidate A is Condorcet. Dodgson voting elects the candidate for which this distance is minimized.
- *Kemeny*: For every ordering of the candidates, we compute the distance to each ballot in the profile and sum these. The ordering in which this distance is minimized is selected, and a winner or multiple winners can be derived from this order.

Some rules privilege information about precise ranking levels rather than relative match-ups:

Rank-Specific Rules:

- *Plurality*: Whoever gets the most first-choice votes wins.
- *Borda*: Candidates are awarded a_1 points for each first place ranking, a_2 points for each second place ranking, etc. Typically $a_i = N - i$, where N is the number of candidates.
- *Condo-Borda*: If there is a Condorcet winner (i.e. option which wins all direct match-ups), it wins. Otherwise, proceed by Borda.
- *Secondality*: The candidate with the second-most first-choice votes wins (I think this is mostly a joke).
- *Plurality Veto*: Each candidate is awarded one point for each first-choice vote they receive. The voters are ordered in some way (which does effect the outcome) and each voter's ballot removes one point from their least-favored candidate remaining. When a candidate has lost all points, they are eliminated.

And finally we have the simplest of all voting rules:

Dictatorship:

- *Dictatorship*: One particular voter decides the result.

3. Criteria for Fairness

Voting systems can be judged based on various criteria for "fairness." Below is a list of these.

First, there are criteria based on guaranteeing sufficiently common ballot features are reflected in the outcome:

Consensus-Based Criteria:

- *Pareto Efficient*: If all ballots rank A first, then A must win.
- *Unanimity-fair*: If all ballots rank A above B, then the outcome must have A above B.
- *Majority-fair*: If the majority of ballots rank A first, then A must win.
- *Condorcet-fair*: If A is Condorcet then A must win.
- *Smith-fair*: Winners must come from the Smith set.

Next there are robustness criteria which restrict the effect that given changes to the ballots can have on the outcome:

Profile-Robustness Criteria:

- *Monotonic*: If A is winning and one ballot is altered by swapping A upwards, A still wins.
- *Strictly Monotonic*: If A is winning and one ballot is altered by any swap that doesn't swap A downwards, A still wins.
- *Strategy-Proof*: If altering a single ballot changes the winner from A to B, then the original ballot had $A > B$ and the altered ballot had $B > A$[3].
- *Voter Anonymity*: The voting rule does not depend on the order of ballots.
- *Non-Dictatorship*: The outcome does not depend on only one voter.

Then there are criteria restricting the effect that given changes to the candidate pool can have on the outcome:

Candidate-Robustness Criteria:

- *No Spoilers*: If A is winning, removing B cannot cause A to lose, nor can adding in a candidate C, unless C is the sole winner. In these scenarios, B and C are called spoilers.
- *No Weak Spoilers*: There can be no spoilers that are weak, i.e. not in the Smith set.
- *Candidate Anonymity*: The voting rule does not depend on the order of candidates.

And finally, here is one important criterion which requires robustness in both senses:

General Robustness Criteria:

- *Independence of Irrelevant Alternatives*: The relative outcome of candidates A and B cannot be affected by swapping any two candidates other than A and B on a given ballot *or* by adding or subtracting other candidates. In other words, whether $A < B$ in the outcome depends only on how each voter ranked A relative to B[4].

3.1. Relationships Between Fairness Criteria.

To understand these criteria better, one should see how they relate, which are stronger and weaker, and so on. To begin with, here is a diagram

[3] If this is not true, then voters may be incentivized to report their preferences falsely. For example, if ballot 1 and ballot 2 both have $A > B$, then a voter whose true preferences are ballot 2 should switch to ballot 1.

[4] If one has Voter Anonymity and Candidate Anonymity (two very reasonable expectations), IIA would imply that whether $A \succ B$ depends only on the *number* of ballots preferring $A > B$ vs $B > A$, and by candidate anonymity this decision procedure must be the same for all pairs of candidates, so in effect the voting rule must be determined

of the implications:

$$\begin{tikzcd}
& & \mathsf{SF} \arrow[d, Rightarrow] & & \\
& & \mathsf{CF} \arrow[d, Rightarrow] & \mathsf{SP} \arrow[d, Rightarrow] & \mathsf{IIA} \arrow[d, Rightarrow] \\
\mathsf{UF} \arrow[dr, Rightarrow] & & \mathsf{MF} \arrow[dl, Rightarrow] & \mathsf{SM} \arrow[d, Rightarrow] & \mathsf{NS} \arrow[d, Rightarrow] \\
& \mathsf{PE} & & \mathsf{M} & \mathsf{NWS}
\end{tikzcd}$$

Now, in many places one can

4. Impossibility Results

Having established some desirable fairness properties, we're forced to bite the bullet: many of them can never coexist in one voting rule.

Incompatible Fairness Criteria:

- *Monotonic* and *No Spoilers*:

  Proof. Consider the following simple profile with a Condorcet cycle:

$\times 1$	$\times 1$	$\times 1$
A	B	C
B	C	A
C	A	B

  In principle, it should be difficult to name a winner here due to the symmetry. Indeed, by Monotonicity, A would lose if B were removed, B would lose if C were removed, and C would lose if A were removed. Thus if any of A, B, C were winning, they could be made to lose by removing a candidate, making the removed candidate a spoiler. So there are either spoilers or the system declares no winner. □

- (**Arrow**) *Unanimity Fair*, *Independence of Irrelevant Alternatives* and *Non-Dictatorship*:

  Proof. Consider the following sequence of profiles:

  $$P_1 = \begin{array}{|c|c|}\hline \times 1 & \times 1 \\ \hline A & B \\ B & A \\ C & C \\ \hline \end{array} \qquad P_2 = \begin{array}{|c|c|}\hline \times 1 & \times 1 \\ \hline A & B \\ B & C \\ C & A \\ \hline \end{array} \qquad P_3 = \begin{array}{|c|c|}\hline \times 1 & \times 1 \\ \hline A & C \\ B & B \\ C & A \\ \hline \end{array}$$

  In profile P_1, a unanimity-fair voting rule must rank C last. Suppose WLOG that it ranks A first. Then in P_2, by IIA, the voting rule must still return $A > B > C$ because the ranking of B relative to C hasn't changed. And moving from P_2 to P_3, the relative rankings of (A, B) and (A, C) haven't changed, so the voting rule must still give □

- (**Muller-Satterthwaite**) *Pareto Efficient*, *Strongly Monotonic* and *Non-Dictatorship*:
- (**Gibbard-Satterthwaite**) *Pareto Efficient*, *Strategy-Proof* and *Non-Dictatorship*:

If one considers Monotonicity, Unanimity-fairness, Pareto Efficiency, and Non-Dictatorship to be essential requirements of a fair voting system (and they do seem like reasonable requests), then these impossibility theorems can be interpreted as saying that the properties *No Spoilers*, *Independence of Irrelevant Alternatives*, *Strongly Monotonic*, and *Strategy-Proof* can never be achieved in any fair ranking-based voting system.