Computational Social Choice

Edith Elkind University of Oxford

Rank Aggregation

- Rank aggregation (also known as voting) is the problem of aggregating several ordered lists of alternatives
- Input:
 - a set of alternatives (candidates) $C = \{c_1, ..., c_m\}$
 - a set of voters V = {1, ..., n}
 - for each voter, a total order (ranking) over C
- Output:
 - a winner
 - a set of winners
 - a total ranking of the alternatives

Rank Aggregation: Examples

- What movie should the Simpson family watch?
 - - : Frozen > Paddington > Minions





: Paddington > Minions > Frozen







Rank Aggregation: Examples

- Which PhD applicant should the algorithmic game theory group at Oxford accept?
 - Paul: X > Y > Z
 - Elias: Y > X > Z
 - Edith: Z > Y > X



Rank Aggregation: Examples

- United Kingdom (specific precinct)
 - candidates: Conservatives (C), Labour (L), Liberal Democrats (LD)
 - 60 000 voters
 - $-25\ 000\ voters\ prefer\ C\ to\ LD\ to\ L:\ C > LD > L$
 - 20 000 voters: L > LD > C
 - 11 000 voters: LD > L > C
 - 4 000 voters: LD > C > L

Example: Competition for a Fellowship

- Candidates: 50 students
- Voters: 15 panel members
 - each panel member has a ranking of the candidates
 (or perhaps top 10 candidates)
- Goal: select 10 students who will get a fellowship
- Asking each panel member to vote for her favorite candidate is not appropriate:

– at most 7 students can get 2 or more votes

Example: Ranking of the Universities

- A panel of experts is supposed to rank UK universities
 - Expert 1: Cambridge > Oxford > UCL > Imperial
 - Expert 2: Oxford > Cambridge > Imperial > UCL
 - Expert 3: UCL > Cambridge > Oxford > Imperial
 - Expert 4: Oxford > Imperial > Cambridge > UCL
 - Expert 5: Imperial > Cambridge > UCL > Oxford
- Goal: produce a total ranking of 4 universities

Example: Ranking of the Universities

- A panel is supposed to rank UK universities
- Rankings are based on 5 different criteria:
 - reputation ranking
 - grant income
 - student satisfaction
 - number of research papers published
 - average salary after graduation
- Rankings:
 - criterion 1: Cambridge > Oxford > Imperial > UCL
 - criterion 2: Oxford > Cambridge > UCL > Imperial
 - criterion 3: UCL > Cambridge > Oxford > Imperial
 - criterion 4: Oxford > Imperial > Cambridge > UCL
 - criterion 5: Imperial > Cambridge > UCL > Oxford
- Should all criteria have the same weight?

Part 1: the zoo of voting rules

Single-Winner Rules: Plurality

- <u>Plurality</u>:
 - each voter names his favorite candidate
 - candidates with the largest number of votes win
 - if two or more candidates get the highest score,
 the winner is chosen using some tie-breaking rule
- For 2 candidates,
 Plurality selects the majority winner

Political Voting

- United Kingdom (specific precinct)
 - candidates: Conservatives (C), Labour (L), Liberal Democrats (LD)
 - 60 000 voters
 - $-25\ 000\ voters\ prefer\ C\ to\ LD\ to\ L:\ C > LD > L$
 - 20 000 voters: L > LD > C
 - 11 000 voters: LD > L > C
 - 4 000 voters: LD > C > L
 - Plurality outcome: C wins with 25000 votes

Single-Winner Rules: Plurality

- Plurality is obviously the best voting rule if there are only 2 candidates
- However, for 3 candidates it may behave in an undesirable way
 - the majority of voters may prefer some other alternative to the current winner
 - voters have an incentive to vote non-truthfully

Plurality: Example Revisited

- United Kingdom elections:
 - 25 000 voters: C > LD > L
 - 20 000 voters: L > LD > C
 - 11 000 voters: LD > L > C
 - -4000 voters: LD > C > L
- Outcome under Plurality:
 - C wins with 25000 votes
- Undesirable properties:
 - 31 000 voters prefer L to C, 35 000 voters prefer LD to C
 - the voters with ranking LD > L > C would be better off voting L

Two-Round Elections

- 1. All voters vote for their favorite candidate
- 2. All but the two highest-scoring candidates are eliminated
- 3. The voters are asked to vote again over the remaining candidates

This rule is known as Plurality with Runoff; used in France for presidential elections

Plurality With Runoff: Example

- United Kingdom elections:
 - 25 000 voters: C > LD > L
 - 20 000 voters: L > LD > C
 - 11 000 voters: LD > L > C
 - -4000 voters: LD > C > L
- 1st round: C: 25 000, L: 20 000, LD: 15 000

- noone has more than 30 000 votes, so LD is eliminated

• 2nd round: C: 25 000+4 000, L: 20 000 + 11 000

L gets the majority of votes, so it wins

Multi-Round Elections

- 1. All voters vote for their favorite candidate
- 2. If some candidate gets more than 50% of the votes, he is declared the winner
- 3. Otherwise, the candidate with the smallest number of votes is eliminated
- 4. The voters are asked to vote again over the remaining candidates
- 5. The process repeats until some candidate gets a majority of votes

Single Transferable Vote

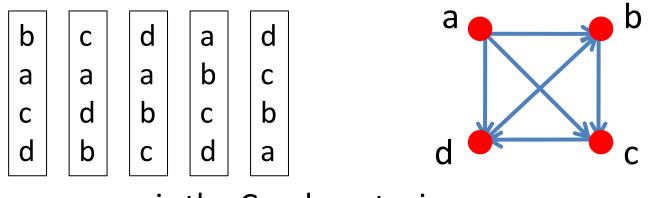
- Multi-round elections often produce a more appealing outcome than Plurality
- However, they are hard to implement:
 voters have to come to voting booths many times
- <u>Single Transferable Vote</u>: an implementation of multi-round elections in a single round of voting
 - each voter submits a total ranking of candidates
 - the election authority simulates multi-winner elections based on the information in the ballots (assuming that all voters always vote for their most preferred available candidate)
- UK had a referendum of switching to STV on May 5th, 2011 - but the decision was "no"

How Good are Plurality With Runoff and STV?

- United Kingdom elections:
 - 25 000 voters: C > LD > L
 - 20 000 voters: L > LD > C
 - 11 000 voters: LD > L > C
 - 4 000 voters: LD > C > L
- Plurality chooses C, STV chooses L
- Yet, 40 000 voters prefer LD to L and 35 000 voters prefer LD to C
- Under both Plurality and STV, more than 50% of voters would have preferred a different candidate
- Under STV, the voters who rank C first would be better off voting for LD

Condorcet Winners

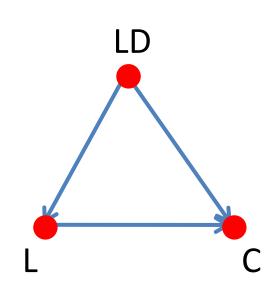
- Suppose that each of the n voters has a ranking of all m candidates
- <u>Definition</u>: a candidate c wins a pairwise election against a candidate d if more than half of the voters rank c above d
- A candidate is said to be a Condorcet winner if he wins in all pairwise elections he participates in



a is the Condorcet winner

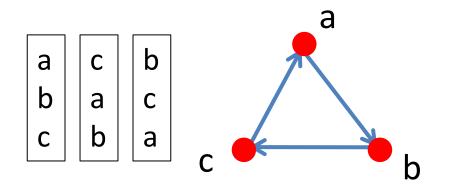
Condorcet Consistency

- A voting rule is said to be Condorcet-consistent if it selects the Condorcet winner whenever it exists
- United Kingdom elections:
 - 25 000 voters: C > LD > L
 - 20 000 voters: L > LD > C
 - 11 000 voters: LD > L > C
 - -4000 voters: LD > C > L
- Plurality chooses C, STV chooses L
- LD is the Condorcet winner
 - even though it has the smallest number of voters who rank it first
- Hence, neither Plurality nor STV are Condorcet-consistent



Do Elections Always Have Condorcet Winners?

- 2 voters rank a above b
- 2 voters rank b above c
- 2 voters rank c above a
- No Condorcet winner!
- <u>Definition</u>: G is a pairwise majority graph for an election E with a candidate set C if its vertex set is C and there is an edge from a to b iff majority of voters prefer a to b
- <u>Theorem</u>: any directed graph with no 2-cycles can arise as a pairwise majority graph



Condorcet-Consistent Rules: Copeland

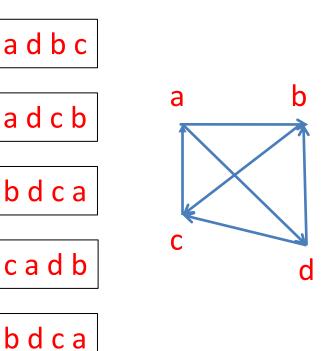
- A Condorcet-consistent rule must elect a Condorcet winner when one exists
 - how can we extend this principle if there is no Condorcet winner?
- <u>Copeland rule</u>: each candidate gets
 - 1 point for each pairwise election he wins
 - 0.5 points for each pairwise election he ties
 - the candidate with the largest number of points wins
- In an m-candidate election, if a Condorcet winner exists, he gets m-1 point, all other candidates get at most m-2 points

Condorcet-Consistent Rules: Maximim

- <u>Maximin rule</u>: the score of each candidate is the number of votes he gets in his worst pairwise election
 - the candidate with the highest score wins
- In an n-voter election, if a Condorcet winner exists
 - his Maximin score is greater than n/2,
 - everyone else's Maximin score is less than n/2

Condorcet-Consistent Rules: Dodgson

- - NP-hard to compute
- Dodgson winner(s): the candidate(s) with the smallest Dodgson score



Scoring Rules

- Condorcet-consistent rules are hard to explain to voters

 implementation is non-trivial
- Alternative: scoring rules
- A scoring rule for an election with m candidates is given by a vector (s₁, ..., s_m), s₁ ≥ ... ≥ s_m
 - each candidate gets s_i points from each voter who ranks him i-th
 - candidate with the maximum number of points wins
- Plurality is a scoring rule with score vector (1, 0, ..., 0)
- Borda: (m-1, m-2, ..., 2, 1, 0)
- k-approval: (1, ..., 1, 0, ..., 0)

equivalent to allowing voters to vote for k candidates

Competition for a Fellowship Revisited

- Candidates: 50 students
- Voters: 15 panel members
 - each panel member has a ranking of the candidates (or perhaps top 10 candidates)
- Goal: select 10 students who will get a fellowship
- 10-approval (aka Bloc):
 - each voter is asked to vote for top 10 candidates
- Truncated Borda:
 - each voter is asked to identify top 10 candidates, and order them
 - each student gets 11 i points from each voter who ranks him in position i
- In either case, students with top 10 total scores win

Scoring Rules

- United Kingdom elections:
 - 25 000 voters: C > LD > L
 - 20 000 voters: L > LD > C
 - 11 000 voters: LD > L > C
 - -4000 voters: LD > C > L
- Plurality: C wins with 25 000 points
- Borda:
 - C gets 2 x 25 000 + 1 x 4 000 = 54 000 points
 - L gets 2 x 20 000 + 1 x 11 000 = 51 000 points
 - LD gets 1 x 45 000 + 2 x 15 000 = 75 000 points
- 2-approval:
 - C: 29 000, L: 31 000, LD: 60 000

Scoring Rules: Pro and Contra

- Scoring rules are easy to understand and implement
- They take into account preferences other than just the voter's top choice
- However, no scoring rule is Condorcet-consistent
 - Borda:

- a is the Condorcet winner, yet a gets 8 points, while b gets 10
- Borda rule is very easy to manipulate:
 - 3 voters: a > b > c > d > e
 1 voter: b > a > c > d > e a: 15, b: 13

 - if the last voter, who prefers b, votes b > c > d > e > a, a loses 3 points, so b wins

Bucklin's rule

- How do we choose k for k-approval?
- One possible answer: adaptively
- Let k* be the smallest value of k such that there is a candidate ranked in top k positions by more than n/2 voters
- Bucklin rule: output all k*-approval winners
- Alternative interpretation:
 - for k =1,, m do
 - ask each voter to name their top k candidates
 - stop when some candidate is named by a majority
 - report all such candidates

Schulze's Rule

- Consider the weighted majority graph
 - the weight of the edge AB is the number of voters who prefer A to B
 - only keep edges whose weight is $\geq n/2$
- Strength of a path from A to B: min weight along that path
- p[A, B]: strength of the strongest path from A to B
- A is a winner if $p[A, B] \ge p[B, A]$ for all B

- always exists

Ties?

- All rules defined so far may produce multiple winners
- In a sense, this is unavoidable
 - suppose input election contains a single copy of each of the m! permutations of candidates
- Tie-breaking:
 - lexicographic (based on a candidate order)
 - randomized
 - uniform over top-scoring candidates
 - pick a random voter, ask her to break the tie

Rankings: Social Welfare Functions

- Score-based rules can be used to produce rankings: order candidates by score
 - not just scoring functions, but also Copeland, Maximin, etc.

- Kemeny rule:
 - for two votes u,v, let $d(u,v)=\# \{(A,B): A >_u B, B >_v A\}$
 - find a ranking that minimizes
 the total distance to votes

Ranking of the Universities: Borda

- A panel of experts is supposed to rank UK universities
 - Expert 1: Cambridge > Oxford > UCL > Imperial
 - Expert 2: Oxford > Cambridge > Imperial > UCL
 - Expert 3: UCL > Cambridge > Oxford > Imperial
 - Expert 4: Oxford > Imperial > Cambridge > UCL
 - Expert 5: IMperial > Cambridge > UCL > Oxford
- Goal: produce a total ranking of 4 universities
- Borda rule:
 - each university gets 4-i points from each expert who ranks it in position i
 - Cambridge: 10, Oxford: 9, UCL: 5, Imperial: 6

Ranking of the Universities: Kemeny

- A panel of experts is supposed to rank UK universities
 - Expert 1: Cambridge > Oxford > UCL > Imperial
 - Expert 2: Oxford > Cambridge > Imperial > UCL
 - Expert 3: UCL > Cambridge > Oxford > Imperial
 - Expert 4: Oxford > Imperial > Cambridge > UCL
 - Expert 5: Imperial > Cambridge > UCL > Oxford
- Goal: produce a total ranking of 4 universities
- Kemeny rule:
 - need to score each of the 24 possible rankings
 - e.g., Oxford > Cambridge > UCL > Imperial scores 5+5+3+1

Complexity of Winner Determination

- Can we efficiently compute the outcome of a voting rule?
 - poly-time algorithms: scoring rules, Copeland, Maximin, Schulze
 - NP-hard: Dodgson, Kemeny
 - it's complicated: STV
 - we can run STV breaking ties in some way and find some winner
 - it is NP-hard to decide whether a given candidate is a winner for some way of breaking ties

Part 2: Justifying Voting Rules

Desirable Properties of Voting Rules

- Anonymity: all voters are treated in the same way
 +: all
- Neutrality: all candidates are treated in the same way
 - +: all (ties?)
- Condorcet consistency
 - +: Copeland, Maximin, Dodgson, Schulze
 - -: Plurality, Plurality with Runoff, STV, Borda

Criteria for Voting Rules: Single-Winner Elections

- Consistency: consider two elections with disjoint sets of voters over the same set of candidates. If c wins in both elections, he should also win when we merge these two elections
 - +: scoring rules
 - -: (nearly) everything ese
- Pareto efficiency: if all voters rank a above b, b should not win +: all

Criteria for Voting Rules: Single-Winner Elections

- Monotonicity: if c wins, and some voter moves
 c higher in her ranking, without changing the
 order of other candidates, then c still wins
 - +: Plurality, Copeland, Maximin, Borda, Schulze
 - -: Plurality with Runoff, STV
 - Example (STV):

A moves to the top in the first 2 votes

Criteria for Voting Rules: Rankings

- Pareto efficiency: if all voters rank a above b, in the final ranking a should appear above b
- Monotonicity: if some voter moves c up in their ranking, in the overall ranking c goes up
- Independence of irrelevant alternatives (IIA): if a is ranked above b in the current election, and we permute the candidates in each vote without changing the relative order of a and b, then a should be ranked above b in the resulting election

Dictatorship

- There is a very simple rule that produces a ranking of alternatives and satisfies all of our criteria: dictatorship
- This rule simply copies the ranking of some fixed voter
- Satisfies monotonicity, Pareto-optimality, IIA
- Truthful voting is a dominant strategy
- Is usually not an acceptable voting rule for obvious reasons

Arrow's Theorem [1951]

- Suppose there are at least 3 candidates
- Then any voting rule that produces a ranking of all candidates and is simultaneously:
 - Pareto efficient and
 - independent of irrelevant alternatives
 - is a dictatorship

"There is no perfect voting rule"

Gibbard-Satterthwaite Theorem

- Suppose there are at least 3 candidates. Then for any voting rule that is not a dictatorship there exists a list of voters' preferences such that some voter v has an incentive to vote non-truthfully
 - v can change his vote so that the winner is a candidate that v ranks higher than the original winner
- No voting rule is resistant to manipulative behavior!

Voting as Preference Aggregation

• What movie should the Simpson family watch?



: Frozen > Paddington > Minions





: Paddington > Minions > Frozen





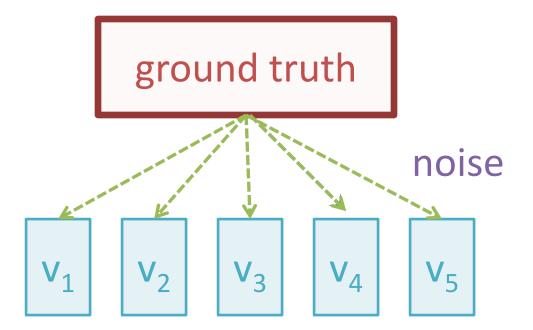


Voting as a Way to Uncover Truth

- Which cleaning company should we hire?
 - Adam: A > B > C
 - Ben: C > B > A
 - Charlie: B > C > A
- Which PhD applicant should we accept?
 - Paul: X > Y > Z
 - Elias: Y > X > Z
 - Edith: Z > Y > X
- Medieval church elections
- Crowdsourcing



Voting as Maximum Likelihood Estimation



Which true state of the world is most likely to generate the observed votes?

History

- Marquis de Condorcet (1785), Essai sur l'application de l'analyse a la probabilité des décisions rendues a la pluralité des voix
- H. Peyton Young (1988), Condorcet's theory of voting, Am. Pol. Sci. Review
- Elkind, Shah (2014), Choosing the most probable without eliminating the irrational: voting on intransitive domains, UAI'14

Condorcet-Young-Mallows Model

- m alternatives, n voters: V = (v₁, ..., v_n)
- Ground truth = ranking of the alternatives
- Votes = rankings of the alternatives
- Noise:
 - $\text{ fix } \frac{1}{2}$
 - ground truth: u
 - each vote is an outcome of the following process:
 - pick a fresh pair of alternatives a, b; assume a >_u b
 - rank them as a > b w.p. p and as b > a w.p. 1-p
 - if this produces a cycle, restart

Most Likely Ranking [Young'88]

- Kemeny distance: d(u, v) = |{(a, b): a >_u b, b >_v a}|
- $\phi = p/(1-p)$
- $Pr[v] \sim p^{m(m-1)/2 d(u, v)} (1-p)^{d(u, v)}$
- $Pr[V] = Pr[v_1] \times ... \times Pr[v_n] \sim \phi^{-\Sigma_i d (u, v_i)}$
- $\Pr[V] \sim \phi^{-d(u, V)}$
- Most likely ranking: one that minimizes the total distance to votes
 - Kemeny's rule

Rankings vs. Winners

- Finding the most likely ranking: Kemeny's rule
- Finding the most likely winner?
- s_R(a): cumulative likelihood of rankings where a is ranked first
- $s_R(a) = \sum u: top(u) = a \phi^{-d}(u, V)$
- Which a maximizes s_R(a)?

Most Likely Winner [Y'88, PRS'12]

- $s_R(a) = \sum_{u: top(u)=a} \phi^{-d(u, V)}$
- s_R(a): sum of (m-1)! non-positive powers of
- $p \rightarrow 1$, $\phi = p/(1-p) \rightarrow \infty$ (low noise):

 the set of most likely winners is a subset of Kemeny winners

• $p \rightarrow 1/2$, $\phi = p/(1-p) \rightarrow 1$ (high noise):

 the set of most likely winners is a subset of Borda winners

Part 3: Domain Restrictions

Difficulties

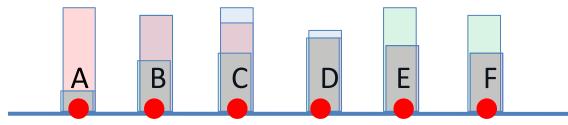
• <u>Problem</u>:

with no assumption on preference structure

- majority cycles may occur
- all voting rules are manipulable
- computing outcomes of some voting rules is NP-hard
- Solution: restrict the preference domain

Single-Peaked Preferences

- <u>Definition</u>: a vote v is single-peaked (SP) wrt an ordering < of candidates (axis) if it holds that:
 - if top(v) < D < E, v prefers D to E</p>
 - if A < B < top(v), v prefers B to A</p>
- Example:
 - voter 1: C > B > D > E > F > A
 - voter 2: A > B > C > D > E > F
 - voter 3: E > F > D > C > B > A



Example: Political Voting

- United Kingdom (specific precinct)
 - candidates: Conservatives (C), Labour (L), Liberal Democrats (LD)
 - 60 000 voters
 - $-25\ 000\ voters\ prefer\ C\ to\ LD\ to\ L:\ C > LD > L$
 - 20 000 voters: L > LD > C
 - 11 000 voters: LD > L > C
 - 4 000 voters: LD > C > L



Example: Temperature

• Perfect water temperature?

+16 +20 +23 +25 +27 +30

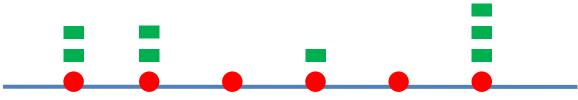


SP Preferences: Transitivity

- <u>Theorem</u>: in single-peaked elections with an odd number of voters the majority relation is transitive
 - if more than n/2 voters prefer a to b and more than n/2 voters prefer b to c then more than n/2 voters prefer a to c
- <u>Lemma</u>: each single-peaked election with an odd number of voters has a Condorcet winner (CW))
- Proof of the theorem (assuming the lemma):
 - by the lemma, there is a CW, say a
 - delete a from all votes; the profile remains SP
 - use induction

SP Preferences: Condorcet Winners

- <u>Lemma</u>: in single-peaked elections with an odd number of voters there exists a Condorcet winner (CW))
 - ask each voter v to vote for one candidate
 - let C(v) denote the vote of voter v
 - order voters by C(v), breaking ties arbitrarily
 - if we have n = 2k+1 voters, $top(v_{k+1})$ is a CW
 - even n: if we have n = 2k voters, all candidates between $top(v_k)$ and $top(v_{k+1})$ are weak CWs

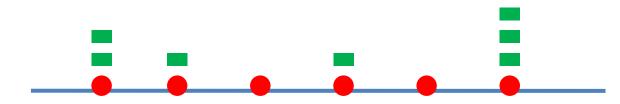


Transitivity: Consequences

- <u>Theorem</u>: in a <u>single-peaked</u> election with an odd number of voters the winning ranking under the Kemeny rule can be computed in <u>polynomial time</u>
 - <u>Lemma</u>: if the majority relation is transitive, the
 Kemeny ranking coincides with the majority relation.

SP Preferences: Circumventing Gibbard-Satterthwaite

- Suppose we have n = 2k+1 voters
- Median voter rule:
 - consider an election that is single-peaked wrt <</p>
 - ask each voter v to vote for one candidate
 - let C(v) denote the vote of voter v
 - order voters by C(v), breaking ties arbitrarily
 - output $C^* = C(v_{k+1})$



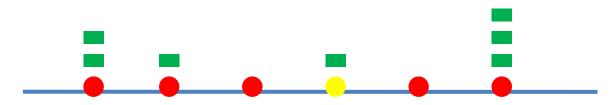
SP Preferences: Median Is Truthful

- <u>Theorem</u>: under the median voter rule, it is a dominant strategy to vote for one's top choice
- Consider a voter v_i in our order
 - -i = k+1: v_i gets his most preferred outcome
 - -i < k+1 (i > k+1 is symmetric):
 - if v_i votes C, C ≤ C*, v_{k+1} remains the median voter, so the outcome does not change



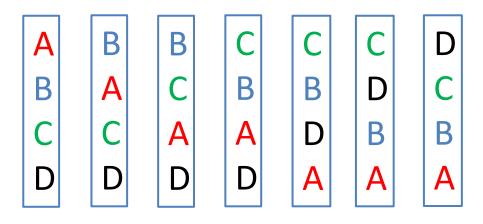
SP Preferences: Median is Truthful

- <u>Theorem</u>: under the median voter rule, it is a dominant strategy to vote for one's top choice
- Consider a voter v_i in our order
 - -i = k+1: v_i gets his most preferred outcome
 - i < k+1 (i > k+1 is symmetric):
 - if v_i votes C, C ≤ C*, v_{k+1} remains the median voter, so the outcome does not change
 - if v_i votes C, C* < C, either v_i (with his new vote) or v_{k+2} becomes the median voter, so the outcome gets worse for v_i



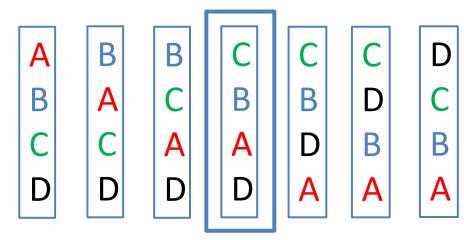
Single-Crossing Preferences

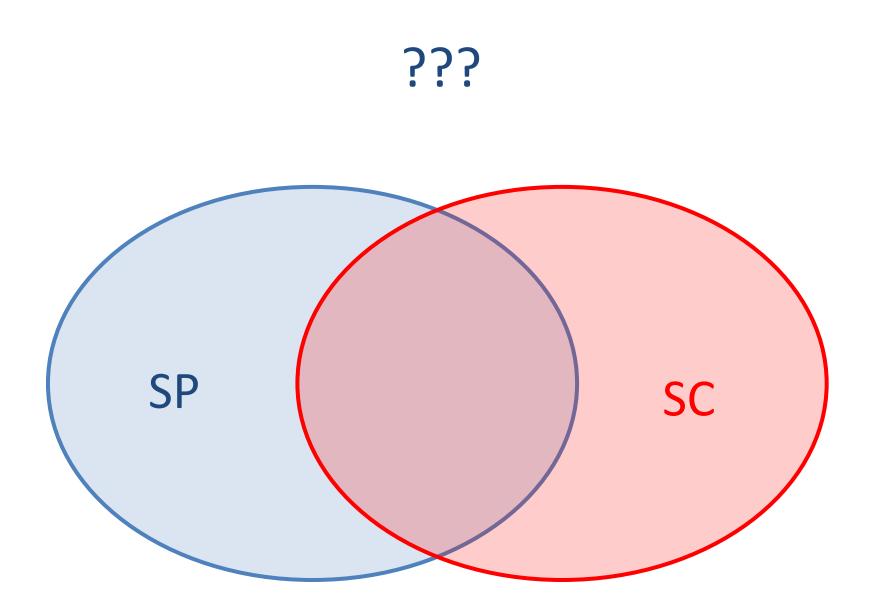
<u>Definition</u>: a profile is single-crossing (SC) wrt an ordering of voters $(v_1, ..., v_n)$ if for each pair of candidates A, B there exists an $i \in \{0, ..., n\}$ such that voters $v_1, ..., v_i$ prefer A to B, and voters $v_{i+1}, ..., v_n$ prefer B to A



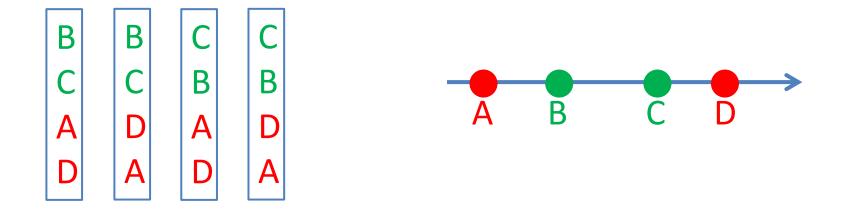
SC Preferences: Majority is Transitive

- <u>Claim</u>: in single-crossing elections, the majority relation is (weakly) transitive
 - we will prove the claim for n=2k+1 voters
 - consider the ranking of voter v_{k+1}
 - if v_{k+1} prefers B to A, so do $\geq k$ other voters
- <u>Claim</u>: the SC order of voters is essentially unique



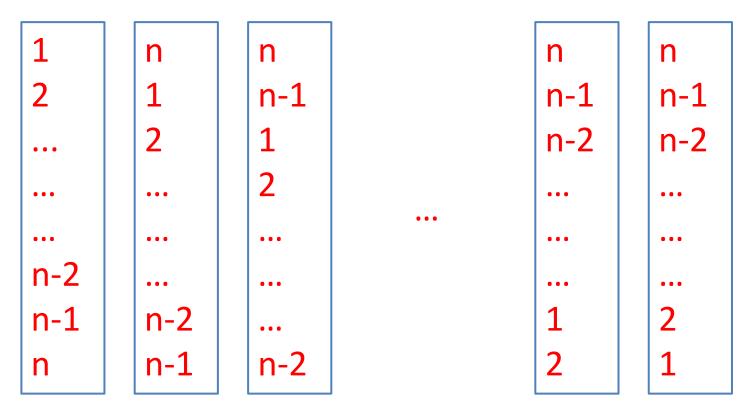


Single-Peaked Profile That Is Not Single-Crossing

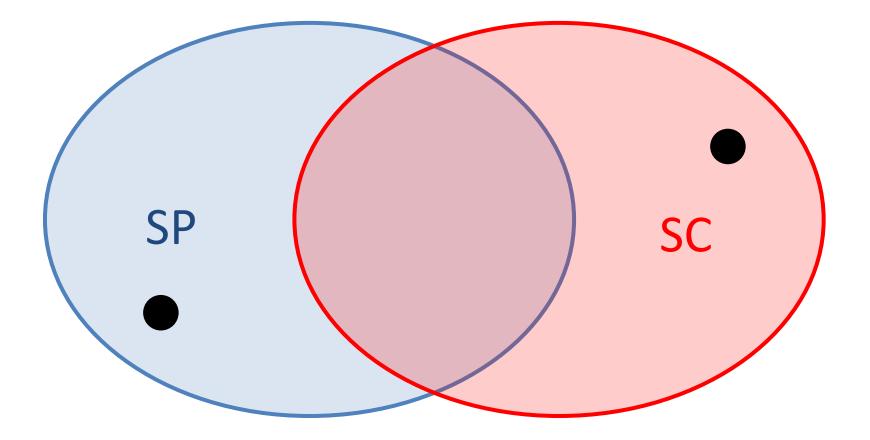


v₁ and v₂ have to be adjacent (because of B, C)
v₃ and v₄ have to be adjacent (because of B, C)
v₁ and v₃ have to be adjacent (because of A, D)
v₂ and v₄ have to be adjacent (because of A, D) a contradiction

Single-Crossing Profile That Is Not Single-Peaked

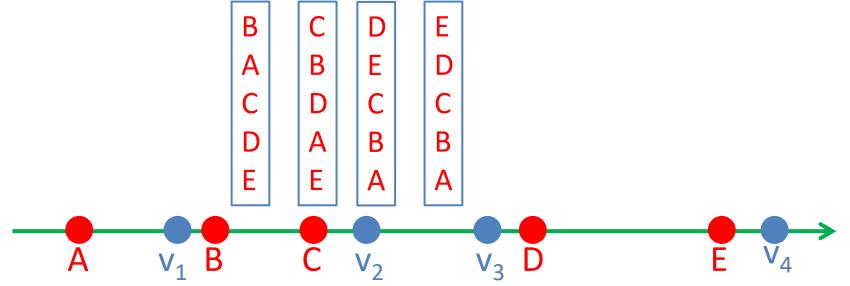


Each candidate is ranked last exactly once



1D-Euclidean Preferences

- Both voters and candidates are points in \mathbb{R}
- v prefers A to B if |v A| < |v B|
- <u>Observation</u>: 1D-Euclidean preferences are
 - single-peaked (wrt ordering of candidates on the line)
 - single-crossing (wrt ordering of voters on the line)



$1\text{-Euc} = SP \cap SC?$

- <u>Observation</u>: There exists a preference profile that is SP and SC, but not 1-Euclidean
 - v_1 : B C D E A F v_2 : D E C B A F v_3 : D E F C B A

- SC wrt $v_1 < v_2 < v_3$, SP wrt A < B < C < D < E < F
- Not 1-Euclidean:

 $- (x(A) + x(E))/2 < x(v_1) < (x(B) + x(C))/2$ - (x(C) + x(D))/2 < x(v_2) < (x(A) + x(F))/2

 $-(x(B) + x(F))/2 < x(v_3) < (x(D) + x(E))/2$

