COV886: Special Module in Algorithms **Spring 2022**

Lecture 1: Voting Rules

Lecturer: Rohit Vaish Scribes: Akash Agrawal, Aniket Mishra

1 Computational Social Choice: An Introduction

In this section, we introduce the field of Computational Social Choice, and outline the key problems of interest that we will cover in the course.

1.1 What is Social Choice?

One of the most common problems that human society has faced has been making decisions when in a group. Whether it be deciding which restaurant your friend circle goes to, which college should be allotted to which student or how to divide the workload of a group project fairly - you have all faced these problems in your day-to-day life. Solutions to these kinds of problems have been pondered upon for a long time as well, and still continues to be an active area of research.

Formalization of these thoughts led to the birth of what we now call Social Choice Theory. It is an area of economics that concerns itself with collective decision making, through the use of mathematical models and rules/guidelines, to create systems where groups can take decisions based on individual preferences, that lead to an outcome that has some desired "properties".

To take an example, a classical Social Choice problem is known as Fair Cake Cutting - Is it possible to divide a cake (or some other resource) between multiple interested parties, such that everyone believes it to be a "fair" division? Here, "fairness" is one such desirable property, although a vague one that requires further definition.

It may be apparent that many such problems may have solutions that are computationally hard to find. This led to the rise of the field of Computational Social Choice. Instead of just being concerned with whether we can have a system that provides a decision-making solution with some desired properties, it also concerns itself with the influence that computation would have on the selection of these properties and the system as well. Returning to the above example, Computational Social Choice would be interested in the questions - "How quickly can we find a fair division, given a certain definition of fairness?" or "Which definition of fairness would be both a useful property as well as computationally easy to find for the Cake Cutting problem?". We will consider this example in detail later on in the course.

1.2 Scope of this course

First, we define what a social choice procedure is -

Definition 1 (Social Choice Procedure). A **Social Choice Procedure** is a function that maps the individual preferences of each agent in a population to a "societal outcome" for a certain Social Choice Problem.

This course will cover various Social Choice Procedures and their design, from the three core areas in classical social choice - Voting, Matching and Fair Division. The primary focus will be on establishing the theoretical basis of the problems from Social Choice Theory and then exploring the computational aspects of those problems.

The three core areas are explained in brief below -

1. **Voting**: Given the individual preferences of the voters over various candidates in an election, design a voting rule (social choice function) that returns the winner of the election, or alternatively, the unified preferences of the voting group over the candidates.
2. **Matching**: Given two groups of agents, each having individual preferences over the agents in the other group, for each agent in the group, determine a matching process to match sets of agents together.
3. **Fair Division**: Given a set of resources or costs to be divided among a group and the preferences of each member of the group, determine an allocation mechanism to allocate various resources/costs to each agent in the group.

2 Voting

2.1 The Voting Problem

The setting is as follows: there's a list of candidates C ($|C| = m$) and a preference profile P over C. A preference profile for a set of n voters V is a list of n preference lists, $P = (P_1, P_2, ..., P_n)$ s.t.

1. Each list P_i is a strict preference ordering, i.e., between two candidates x and y, a voter will always have a strict preference ($x \prec y$ or $x \succ y$) but never an indifference ($x \sim y$) or weak preference ($x \preceq y$ or $x \succeq y$).
2. Each P_i is complete, i.e., each voter has a preference ordering among all m candidates.
3. Each P_i is transitive, i.e., no voter has a cyclic preference ordering. For example, A voter i cannot have such a preference ordering over candidates x, y, z : $x \prec y$, $y \prec z$, $z \prec x$.

The problem, now, is to create a voting rule that somehow aggregate all of these voters' preferences to obtain a winner among the candidates, i.e., the candidate most preferred by the group. We denote this voting rule by a function $f(\dot{})$, that maps a preference profile P to a winner $f(P) \in C$. Such a voting rule is called a **Social Choice Function**.

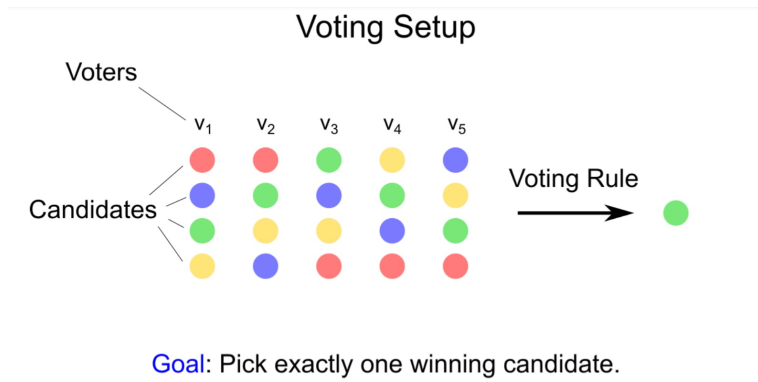

Figure 1: A pictorial representation of the Voting Problem

Note: Sometimes, a voting rule is defined to map a preference profile P not to a single winner, but to a complete, weak ranking order over the candidates. Such a voting rule is called a **Social Welfare Function**. We do not discuss them in this lecture and all discussion hereon is in the context of a Social Choice Function, f. However, the reader may find transforming various properties and voting rules discussed to their Social Welfare counterparts a useful and fun exercise.

2.2 Why are we interested in studying voting?

Voting is a decision making problem that groups of all sizes and importance face on a daily basis - whether it is your friend group deciding on which restaurant to go to for dinner or the world's largest democracy deciding on who will be a part of its government for the next term. It may seem a simple problem on the onset, but determining what the group prefers overall based on what each individual in the group prefers often becomes a complicated and confusing issue, full of paradoxes and counter-intuitive results. Hence, voting becomes an interesting problem to study.

Further, it is important to note what we are interested in when studying voting, i.e., the desirable properties that we might want a voting rule to have. Some such basic properties are noted below informally and we will look into them and others more formally in the next lecture.

Definition 2 (Anonymity)**.** A voting rule is considered **Anonymous** if each pair of voters is interchangeable, i.e., if the preferences of any two voters were to be swapped, the winner given by the voting rule would not be altered. Mathematically, if P' is a preference profile obtained by swapping the preference lists of a pair of voters in P, then $f(P') = f(P)$.

Definition 3 (Non-Dictatorial)**.** A voting rule is considered to be **Non-Dictatorial** if there is no dictator, i.e., the outcome of the voting rule does not coincide with the preferences of one single voter for every possible preference profile of the entire group. In simpler terms, there should not be any one voter who single-handedly determines the result of the voting. Formally, voter i is a Dictator if $f(P)$ coincides with i's top preference for all preference profiles P. This is a weaker assumption than Anonymity.

Definition 4 (Neutrality)**.** A voting rule is considered **Neutral** if each pair of candidates is interchangeable, i.e., if on swapping the position of candidates x and y in the preference list of each voter, the result of the election given by the voting rule is obtained by swapping the positions x and y in the original result.

Definition 5 (Ontoness)**.** A voting rule is considered to be **Onto** if there is no candidate that is unelectable, i.e., there is no candidate x for whom there does not exist any P s.t. $f(P) = x$. In simpler terms, every candidate should have a chance at winning the election. This is also sometimes called the Non-Imposition property.

Definition 6 (Pareto Optimality)**.** A voting rule is considered to be **Pareto Optimal** if it never elects a Pareto dominated candidate. We say that candidate x is Pareto-dominated by another candidate y in a preference profile if y is more preferred to x for every voter. In mathematical terms, for a Pareto optimal voting rule f and a preference profile P s.t $\exists x, y \in C$ with x Pareto dominated by y, $f(P) \neq x$. This seems to be a very reasonable property to have as y is a candidate that is more preferred by every voter in this preference profile.

Definition 7 (Strategyproof)**.** A voting rule is considered to be **Strategyproof** if it is not susceptible to single-voter manipulations. A single voter manipulation occurs when a single voter misreports its actual preferences in the preference profile and this results in a more preferred winner for that voter. Mathematically, a strategyproof voting rule requires that $\forall P$, if P' is some preference profile obtained from P by altering only ith voters preferences, P_i and $f(P) = x \in C, f(P') = y \in C$, then $x \succeq y$ in P_i. This is also sometimes called the Truthfulness property as no voter can improve by misreporting their preference and hence, has no incentive to lie.

Definition 8 (Monotonicity)**.** A voting rule is considered to be **Monotone** if increasing a candidate's support never makes it worse off, i.e., if for some P, $f(P) = x$ and we create a new preference profile P' from P s.t. x in P'_i is at a higher or equal preference level than in $P_i, \forall i$ and the candidates below x in P_i still remain below x in P'_i, then $f(P') = x$. In simpler terms, the voting rule is monotone, if it never punishes a candidate for having a better support among the voters.

Definition 9 (Participation Criterion)**.** A voting rule is considered to promote **Absenteeism** if there are voters who could benefit (obtain a more preferred result) from not participating in the election, i.e., not cast their votes. Mathematically, if for some P, $f(P) = x$ then on creating a new preference profile P' by discarding some "absent" voters V_A, i.e., $P' = P \setminus \{P_i : i \in V_A\}$ s.t. $f(P') = y$, then $y \succ x$ for voters in V_A.
Ideally, we would like to have voting rules that do not promote absenteeism so as to have increased turn-out in elections.

2.3 Voting with 2 candidates

Before we consider the voting problem in general, let us look at a simpler case - where there are only two candidates from which the winner is to be chosen. While there are many possible voting rules, a voting rule called **Majority** is the only sensible one. The Majority rule is as follows: of the two candidates, the winner is the one who is preferred by more number of voters.

Exercise: Reason whether the Majority rule follows the properties outlined above.

Note: You might be wondering what happens in the case of a tie. Tie-breaking can be done in various ways - lexicographic, a deciding vote, etc. However, this will not be the focus of this lecture. Choose any preferred tie-breaker you want for this rule as well as the rules that follow. We follow the lexicographic tie breaker in this lecture unless indicated otherwise.

3 Voting with Multiple Candidates

Now, we are prepared to look at voting rules that can decide a winner between any number of candidates. While there are numerous possibilities for choosing a voting rule, we explore six popular ones classified into three broad classes.

Exercise: For each of the rules outlined below, try to reason whether they follow the five properties in section **2.2** or not, as well as whether they reduce to the Majority rule when only two candidates are considered.

3.1 Scoring Based Rules

The first class of rules we consider are the most straight forward ones. In these rules, to aggregate the individual preferences into a group decision, we assign each candidate a score based on the preference list of each voter, and we choose the winner based on the aggregated scores of the candidates over the entire preference profile. Mathematically, we can define a scoring based voting rule $f_{SC}()$ that using a scoring function $SC()$ as follows:

$$f_{SC}(P) = \arg\max_{x} \sum_{i \in V} SC(x, P_i)$$

where, $x \in C$, P_i is the preference list of the ith voter.

3.1.1 Plurality

The Plurality voting rule is one of the simplest rules out there. It simply declares the candidate who is the top preference of the most number of people as the winner. Defining it in terms of a scoring based rule:

$$f_{PLU}(P) = \arg\max_x \sum_{i \in V} PLU(x, P_i) \quad x \in C$$

where $PLU(x, P_i) = 1$ if x is the top-most candidate in P_i and 0 otherwise.
Hence, the aggregated plurality score of the candidate is the number of voters which rank it as their first choice; and the candidate with the highest plurality score wins (see figure 2).
However, this simplicity comes with its fair share of problems: it doesn't take into account the

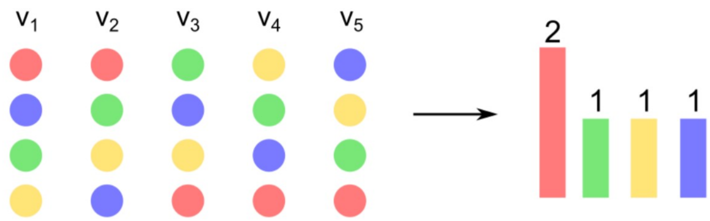

Figure 2: A Plurality Based Election. The red candidate, having the highest plurality score, wins.

other preferences of voters (save for their first choice), it is manipulable (for elections with more than 3 candidates), and it isn't even Condorcet consistent (a desirable property which we'll read about in section 3.3). Even in the shown example, you can note that though Red is the winner, a majority (three out of five) of the voters would have preferred either of Blue, Green or Yellow over Red.

3.1.2 Borda Count

First described by Jean-Charles de Borda as an alternative to plurality that will take into account the entire preference order of the voters, the Borda Count voting rule is commonly used in elections for various academic bodies as well as popularity contests such as Eurovision. Each candidate is given a Borda score by each voter, and the candidate with the highest aggregated Borda score wins (see figure 3). The Borda score of a candidate in an m-candidate election is defined as follows:

$$BORDA(x, P_i) = m - k \quad \text{if } x \text{ is the } k^{th} \text{ ranked candidate in } P_i$$

$$f_{BORDA}(P) = \arg\max_x \sum_{i \in V} BORDA(x, P_i) \quad x \in C$$

While Borda Count improves upon Plurality as a more inclusive rule, that takes into account other preferences of the voters, it still has its share of problems. It is still susceptible to manipulations, as indicated in by the example in figure 4.

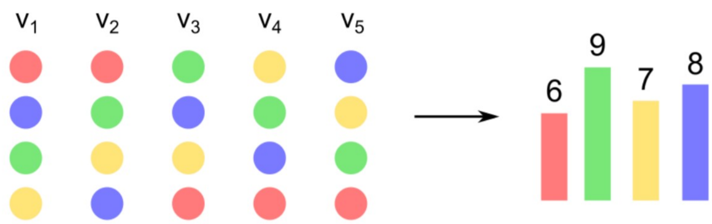

Figure 3: A Borda-Rule Based Election. Green, having the highest borda score, wins.

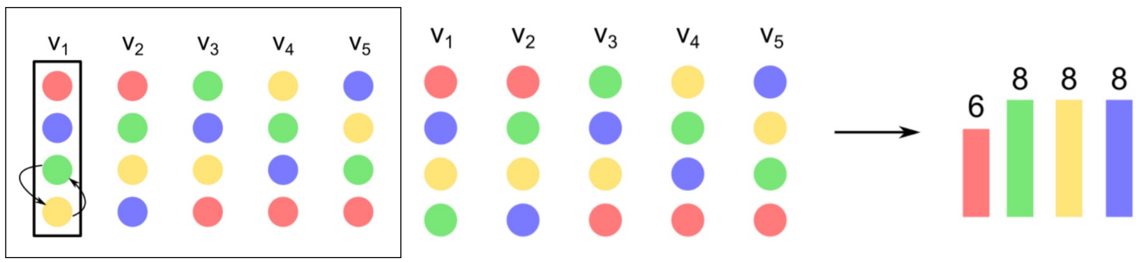

Figure 4: (a) Consider the following Manipulation by voter V_1 in the above example. (b) In the now modified election, Blue is declared the winner (through a lexicographic tie break), and hence V_1 is happier as Blue was a better candidate for him than Green. This shows that Borda Count is not strategyproof.

3.2 Rules based on Runoffs

Another way to define a voting rule is to have multiple rounds of elections, each resulting in a smaller candidate pool until a single winner remains. Such a voting rule is said to have Runoffs. Now, we will see some examples of Runoff voting rules.

3.2.1 Plurality with Runoff

Based on the Plurality rule, there are 2 rounds. In the first round, we select the 2 candidates with the highest plurality score. In the second round, we only have a smaller election between these two candidates using the Majority rule, and the winner is declared the Plurality with Runoff winner.

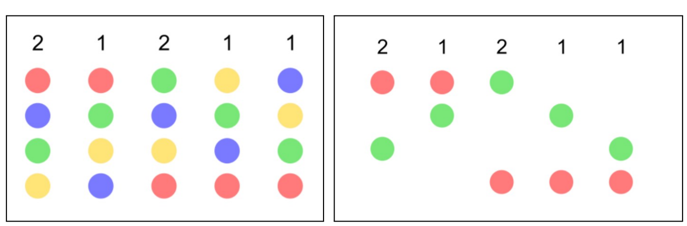

Figure 5: (a) Round 1: Red (Plurality = 3) and Green (Plurality = 2) move to the next round (numbers over a preference list indicate the no. of voters with that preference). (b) Round 2: Green is declared the winner as it is the majority preference (4 out of 7).

A system similar to this one is followed in the French presidential election. There, voters are asked to vote for 1 candidate in the first round. If no candidate receives more than 50% votes in the first round, the 2 candidates ranked highest (i.e. have highest plurality score) go head-to-head in a second round, where voters vote again. If no voter changes their preference from the first round to the next, the result of the election is the same as plurality with runoff.

3.2.2 Single Transferable Vote - Instant Runoff (STV)

This is another widely used round-wise voting system based on Plurality. In each round, the candidate who gets the lowest plurality score is eliminated, and the voters whose top priority for that round was that candidate will have their votes cast to their next most preferred candidates in the next round. In an m-candidate election, STV will take $m-1$ rounds to decide the winner (see figure 6).

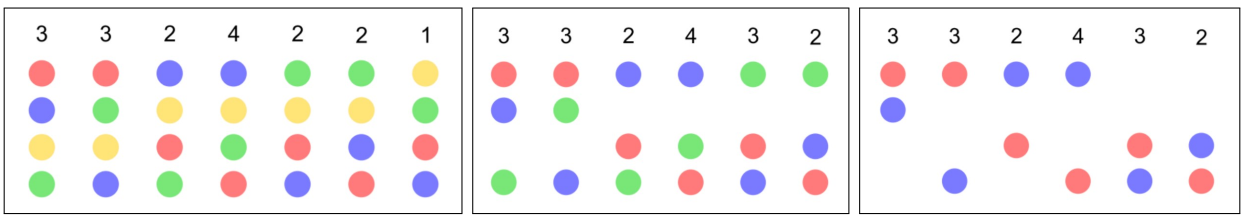

Figure 6: (a) Round 1: Yellow (Plurality = 1) is eliminated. (b) Round 2: Green (Plurality = 5) is eliminated. (c) Red is declared the winner as it is the majority preference (9 out of 17).

One issue with STV is that it is non-monotonic i.e. there is a possibility a candidate becomes worse-off even when its support among the voter only improves (see figure 7).

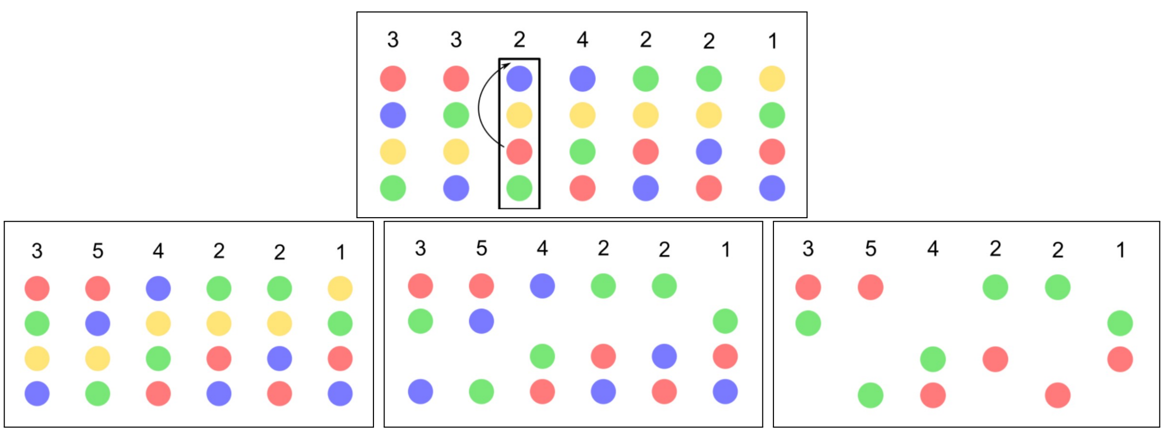

Figure 7: (a) Consider the following change from the above example, 2 more voters maker Red their most preferred choice, and leave others in the same order as before. Now, let's see what happens. (b) Round 1: Yellow (Plurality = 1) is eliminated. (c) Round 2: Blue (Plurality = 4) is eliminated instead of Green here. (d) Green is declared the winner as it is the majority preference (9 out of 17), making Red worse off than before even though his support had increased.

3.3 Condorcet Consistency and Condorcet based rules

Another very interesting property that we may want a voting rule to have is called Condercet Consistency, which was first proposed by Nicolas de Condorcet, a contemporary of Borda. It is defined as follows :-

Definition 10. A **Condorcet winner** is a candidate which defeats every other candidate in a pairwise head-to-head election (i.e. if we disregard all the other candidates in the preference profile and only look at the relative rankings of the two candidates for each voter) (see figure 8). A voting rule is said to be **Condorcet consistent** if it chooses a Condorcet winner whenever there is one.

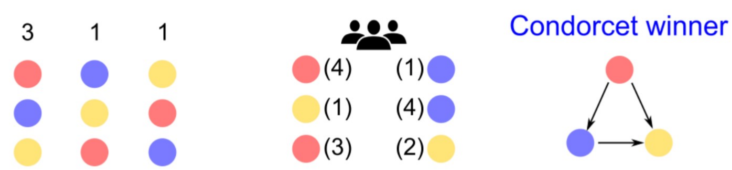

Figure 8: A 5-voter election that has Red as a Condorcet winner.

However, it isn't necessary that a Condorcet winner will exist for each preference profile. The **Condorcet paradox** is that collective preferences can be cyclic even when individual voters' preferences are acyclic. In other words, each voters' preferences being transitive doesn't necessarily lead to the collective's preferences being transitive. (see figure 9)

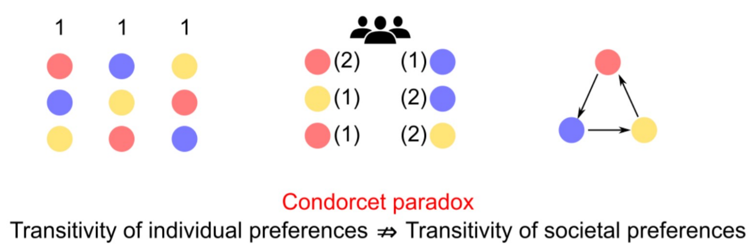

Figure 9: A very simple 3-voter example that has no condorcet winner, demonstrating the Condorcet Paradox.

3.3.1 Copeland

Copeland's method is a Condorcet consistent voting rule. Here, each candidate is given a score based on how they fare against each other candidate in a pairwise election, i.e. for an m-candidate election, we have $\binom{m}{2}$ pairwise 2-candidate elections. For every win, a candidate is given 1 point; for every tie, 0.5 points; and for every loss, 0 points (see figure 10).

Figure 10: A election based on the Copeland Rule. Blue is declared the winner (note that it is also the Condorcet winner)

It is easy to see that Copeland's method is Condorcet consistent – if there is a Condorcet winner, they will defeat every other candidate; and hence will have a score of $m - 1$. Any other candidate will have at least one loss (to the Condorcet winner), and hence will have $\leqslant m - 2$ points. Hence, the Condorcet winner will also be declared the Copeland winner.

However, it may be the case that a voter might benefit (or at least not lose) by not voting at all. Hence, Copeland's method can potentially promote absenteeism (see figure 11).

Figure 11: An example of how Copeland may promote Absenteeism. If the three voters shown faded decide not to vote, Copeland would instead have declared Red as the winner, which is a more preferred outcome for the absentees. Hence, they have an incentive not to vote. (Also note that in this example, there is no Condorcet winner)

3.3.2 Schulze

Now, we come to Schulze's voting rule. This rule is slightly complicated – so bear with us. In this rule, the winner is the candidate who "chain-beats" all the other candidates. The chain-beats relation is defined as follows:

- Make an $m \times m$ matrix M where cell (x, y) contains (number of voters who prefer x to y) - (number of voters who prefer y to x). Each cell (x, x) is has value 0. Note that M is skew-symmetric.
- Make a graph $G = (C, E)$ where there is a directed edge from x to y $(x, y \in C)$ with weight $M_{x,y}$. This graph will also be called the Schulze graph of the preference profile in this text.
- The strength of any path in this graph is defined to be equal to the weight of the least-weight (or weakest) edge in it. Let $S(x, y)$ be the strength of the "strongest" path from x to y.

- A candidate a chain-beats a candidate b if $S(a,b) > S(b,a)$.

We'll use "$a >> b$" to denote that a chain-beats b. Now, a candidate x is a potential Schulze winner if $S(x,y) \geqslant S(y,x)\ \forall y \in C \backslash \{x\}$. In case there are multiple such candidates, then we will use the tie-breaking order to determine which one is the winner out of them.

Let us look at an example:-

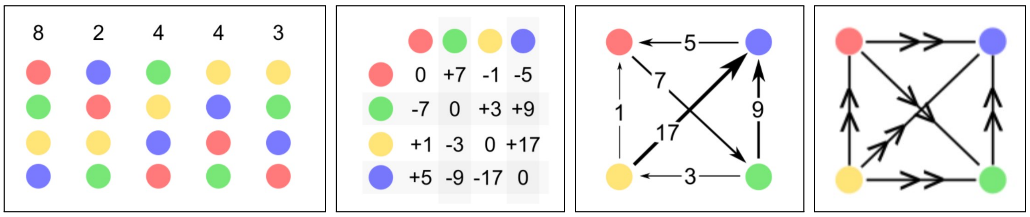

Figure 12: (a) A sample 4-candidate 21-voter election for the Schulze Voting rule. (b) The matrix M calculated for the example. (c) The Schulze Graph G generated using M. (d) The Chain-Beats relation graph obtained from G, which indicates Yellow to be the winner as it chain-beats all other candidates.

Exercise: Assure yourself that the Chain-Beats relation graph has been correctly obtained for the above example. Here are the strength values, for checking your work -

$S(r,b) = 7, S(r,y) = 3, S(r,g) = 7, S(b,y) = 3, S(b,g) = 5, S(y,g) = 5,$

$S(b,r) = 5, S(y,r) = 5, S(g,r) = 5, S(y,b) = 17, S(g,b) = 9, S(g,y) = 3$

Now, we will prove that the chain-beats relation is transitive. Before that, we'll also prove an obvious, yet powerful fact about the Schulze graph.

Lemma 1. *For any $z \in C$, $S(x,y) \geqslant min(S(x,z), S(z,y))$.*

Proof. The strength of the strongest path from x to y is at least as large as the strength of any other path from x to y. Now, $x \to z \to y$ is an $x \to y$ path in the Schulze graph, whose strength is dictated by its min-weight edge. The weight of this edge is either $S(x,z)$ (if it lies on the $x \to z$ part of the path, since $S(x,z) =$ the weight of the weakest edge in the strongest $x \to z$ path) or $S(z,y)$ (if it lies on the $z \to y$ part); hence, the weight of this edge is $\min(S(x,z), S(y,z))$. □

Lemma 2. *(Transitivity of chain-beats) If $x >> y$ and $y >> z$, then $x >> z$.*

Proof. Suppose, for contradiction that x does not chain-beat z. Then $S(z,x) \geqslant S(x,z)$. Now, we know that

$$S(x,y) > S(y,x) \geqslant \min(S(y,z), S(z,x))$$

Let us consider two cases depending on whether $S(y,z) \leqslant S(z,x)$ or $S(y,z) > S(z,x)$:

Case 1: $S(y,z) \leqslant S(z,x)$

So, $S(x, y) > S(y, z)$.

$$S(z, y) \geqslant \min(S(z, x), S(x, y))$$

$$S(z, x) \geqslant S(y, z),\ S(x, y) > S(y, z)$$

This means that

$$S(z, y) \geqslant S(y, z)$$

which contradicts $y >> z$.

Case 2: $S(y, z) > S(z, x)$

So, $S(x, y) > S(z, x)$.

$$S(x, z) \geqslant \min(S(x, y), S(y, z))$$

$$S(x, y) > S(z, x),\ S(y, z) > S(z, x)$$

This means that

$$S(x, z) > S(z, x)$$

which contradicts $S(z, x) \geqslant S(x, z)$.

Therefore, x chain-beats z. □

Now, we'll prove that Schulze's rule will always give us a winner (unlike the Condorcet rule).

Theorem 1. *For each preference profile, Schulze's rule determines a unique winner.*

Proof. For any pair x, y, one of the following are true: $x >> y$, $y >> x$, or $S(x, y) = S(y, x)$. Now, let $W(x) = \{y \in C : x >> y \text{ or } S(x, y) = S(y, x)\}$ and $L(x) = \{y \in C : y >> x\}$. We need to prove that $\exists x \in C$ such that $W(x) = C \backslash \{x\}$ (there is a candidate x for which $S(x, y) \geqslant S(y, x) \forall y \in C \backslash \{x\}$ i.e. a potential Schulze winner).

We'll prove this by contradiction. Assume that there's no such x. Then, let $a \in C$ such that $|W(a)| \geqslant |W(y)|\ \forall y \in C$. Then, $L(a) \neq \phi$ (otherwise $W(a) = C \backslash \{x\}$). This is because $\forall x \in C$, $x \cup W(x) \cup L(x) = C$.

Let $b \in L(a)$. Then, $|W(b)| \leqslant |W(a)| => |L(b)| \geqslant |L(a)|$. Let $c \in L(b) \backslash L(a)$; then, $a >> c$ (since $c \in W(a)$) and $c >> b$; therefore, $a >> b$, which is a contradiction.

Therefore, a is such that $\forall y \in C \backslash \{a\}$, $S(a, y) >= S(y, a)$. Hence, a is a potential Schulze winner. Now, we have proved that there must be at least one potential Schulze winner. The Schulze winner is the top ranked candidate in the tie-breaking order out of all potential Schulze winners.

□

We can also see that Schulze is Condorcet consistent. If there is a Condorcet winner a, then each path from any $b \in C$ to a in the Schulze graph will have its last edge of the form (c, a), where $c \in C$. Now, the weight of this edge is $< \frac{m}{2}$, since c loses to a in a pairwise election;

hence, the strength of any path will have to be $< \frac{m}{2}$. Now, there exists at least one path from a to b that has strength $> \frac{m}{2}$, which is the path containing only one edge (a, b). Hence, $S(a, b) \geqslant$ strength of this single edge path $> S(b, a)$. So a is the Schulze winner.

4 Conclusion and Next Topic

To conclude our discussion on various types of voting rules, we would like to address the following example in figure 13.

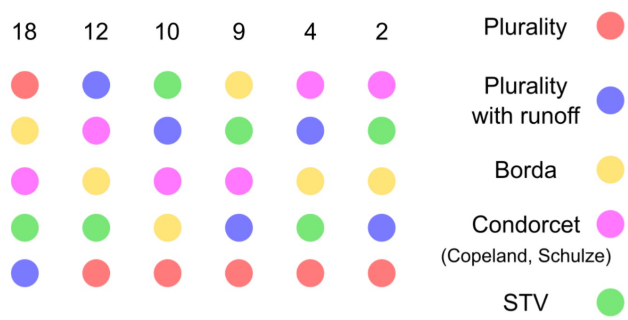

Figure 13: An example that demonstrates how the voting rules that we looked at can all arrive at different winners on the same preference profile. This is why voting is such a paradoxical problem.

Exercise: Show that the winners declared by the voting rules in the above example are correct.

For convenience, a table which shows which properties are held by the various rules discussed in the lecture is attached on the next page.

	Majority	Plurality	Borda Count	Plurality with Runoff
Anonymity	✓	✓	✓	✓
Non-dictatorial	✓	✓	✓	✓
Neutrality	✓	✓	✓	✓
Ontoness	✓	✓	✓	✓
Condorcet consistency	✓	✗	✗	✗
Pareto Optimality	✓	✓	✓	✓
Strategyproof	✓	✗	✗	✗
Monotonicity	✓	✓	✓	✗
Anti-Absenteeism	✓	✓	✓	✗
Computation	$\mathcal{O}(n)$	$\mathcal{O}(n)$	$\mathcal{O}(n)$	$\mathcal{O}(n)$

Table 1: Voting rules and their properties

	STV	Copeland	Schulze	Dictatorship
Anonymity	✓	✓	✓	✗
Non-dictatorial	✓	✓	✓	✗
Neutrality	✓	✓	✓	✓
Ontoness	✓	✓	✓	✓
Condorcet consistency	✗	✓	✓	✗
Pareto Optimality	✓	✓	✓	✓
Strategyproof	✗	✗	✗	✓
Monotonicity	✗	✓	✓	✓
Anti-Absenteeism	✗	✗	✗	✗
computation	$\mathcal{O}(n^2)$	$\mathcal{O}(n^2)$	$\mathcal{O}(n^2)$	$\mathcal{O}(1)$

Table 2: Voting rules and their properties (continued)

This completes the first lecture. In the next lecture, we will see how truthful voting rules are impossible.