Social Choice (6/2/17 to 20/2/17)

7 Introduction

Social choice is concerned with aggregating preferences or votes and then coming up with a rule to decide one or more winners. In some cases there may be no winner.

Understanding social choice is important because it is fundamental to any democratic system which depends on elcitation of preferences from the concerned people. Examples are elections and referendums.

Definition 7 (Election). *An* `election` *is a pair* $\mathcal{E} = (C, V)$ *where* C *is a finite set of candidates or alternatives and* V *is a multi-set*[1] *of votes in the form of a ranking (generally linear order but can be a partial order in some cases) over* C*. Votes and some times elections are also called preference profiles.*

Definition 8 (Voting system (VS)). *A* `voting system` *is a rule that given an election* $\mathcal{E}$ *decides the winner(s) of* $\mathcal{E}$*. It is possible that a voting system produces no winners.*

There can be three types of VS each represented by a function.

`Social choice correspondence` *is a function* $F : \{\mathcal{E} = (C, V) | V \text{ a preference profile}\} \rightarrow \mathcal{P}(C)$*, where* $\mathcal{P}(C)$ *is the power set of* C*. So for any election* $\mathcal{E}$*,* $F(\mathcal{E}) \subseteq C$*,* $F(\mathcal{E})$*, possibly empty, are called the winners of the election.*

`Social choice function` *is a function* $f : \{\mathcal{E} = (C, V) | V \text{ a preference profile}\} \rightarrow C$*. A social choice function picks a single winner.*

`Social welfare function` *is a function* $f_w : \{\mathcal{E} = (C, V) | V \text{ a preference profile}\} \rightarrow \rho(C)$ *where* ρ *is a ranking function for* C *and produces a ranked ordering of the elements of* C*.*

The vote is a preference (usually a linear order) over C. To express that A is preferred to B we write $A > B$. If there are four candidates A, B, C,D there can be a total of 4! or 24 possible preferences. The preference relation is total, transitive, assymetric (that is $A > B \implies B \not> A$). We will often omit the relation symbol while writing the preference. For example, $B > C > A$ will be written as $B\ C\ A$. The first/leftmost alternative is the most preferred and the last/rightmost is the least preferred. In rare cases the relation $>$ may not be a total order if a voter does not have a preference between 2 or more candidates. For example $A > B, C > D$ - no preference between B and C.

8 Voting systems

A large number of voting systems are in use and have also been studied theoretically. In this section we will describe some well known and widely used voting systems. In what follows we always assume we are given an election $\mathcal{E} = (C, V)$. We also assume $|C| = m$ and $|V| = n$ - that is there are m candidates and n votes or equivalently voters. Each voter casts one vote which is a linear order on C.

8.1 Scoring protocol based systems

A scoring protcol system is based on a *scoring vector* $\alpha = (\alpha_1, \ldots, \alpha_m)$ where $\alpha_i \geq \alpha_{i+1}$. Assuming a ballot or vote has the candidates arranged as: $(x_1, \ldots, x_m)$ where each $x_i \in C$ and $x_i \neq x_j, i \neq j$. The points scored by candidate $c_i \in C$ is calculated as follows: Assume j is the position of candidate c_i on a vote $v_k \in V$ (that is candidate c_i is the j^{th} preference in vote v_k) then Points$(c_i) = \sum_{k=1}^{n} \alpha_j$. The winner(s) are the candidates with maximum points. If needed there can be a tie breaking rule to choose one winner from a set of winners. Different scoring vectors lead to different voting systems. The vector $(1, 0, \ldots, 0)$ gives the **plurality** based system where only the first preference matters. This gives rise to the first past the post system as in India. **Borda's system** uses the scoring vector $(m-1, m-2, \ldots, 0)$.

[1] V is a multi-set since multiple voters can have the same preference ranking.

The veto system uses $(1, 1, 1, \ldots, 1, 0)$. The vector $(\underbrace{1, 1, \ldots, 1}_{k}, 0, 0, \ldots, 0)$ is used by the **k-approval** system which generalizes the plurality and veto systems. If $k = 1$ we have plurality and $k = m - 1$ gives veto.

Examples: Consider election $\mathcal{E} = (\{A, B, C\}, V)$ with the votes in V shown in table 3 below.

Votes (V)	Voting system and points								
	Plurality			Veto			Borda		
	A	B	C	A	B	C	A	B	C
A B C	1	0	0	1	1	0	2	1	0
B C A	0	1	0	0	1	1	0	2	1
A B C	1	0	0	1	1	0	2	1	0
Winner	②	1	0	2	③	1	④	④	1

Table 3: Example of scoring based systems (winners circled).

Notice how the different voting systems produce very different winners.

8.2 Pairwise comparison based systems

In these systems the $\binom{m}{2}$ pairwise contests are used to decide winner(s).

8.2.1 Condorcet voting system

A **Condorcet (weak Condorcet) winner** is the candidate who wins the 1-1 contest against all other $(m - 1)$ candidates by more than half (at least half) of the votes cast. If n is odd we can only have a Condorcet winner while if n is even we can have weak Condorcet winner(s). A Condorcet winner when one exists is unique. A weak Condorcet winner need not be unique. For example, in a 3 candidates election with n votes with n even we can have $A|B$ tied which means both A and B win this 1-1 contest; $A|C$ won by A; $B|C$ won by B giving us two weak Condorcet winners, A and B. Similarly, we can define the **Condorcet (weak Condorcet) loser** as the candidate who loses by more than half (at least half) of the votes cast in the 1-1 contests with all other $(m - 1)$ candidates.

Example: Consider the election $\mathcal{E} = (\{A, B, C, D\}, V)$ with the votes given in table 4.

	1-1 results					
Votes (V)	$A\|B$	$A\|C$	$A\|D$	$B\|C$	$B\|D$	$C\|D$
B D C A	B	C	D	B	B	D
D B C A	B	C	D	B	D	D
C A B D	A	C	A	C	B	C
A C D B	A	A	A	C	D	C
A C B D	A	A	A	C	B	C
Winner	A	Ⓒ	A	Ⓒ	B	Ⓒ

Table 4: A Condorcet voting system, C is the Condorcet winner.

An unfortunate aspect of the Condorcet voting system is the possiblity that no winner exists. For example consider the election in table 5 with $C = \{A, B, C\}$:

Condorcet winner determination can also be done by representing the 1-1 results as a digraph called the **majority graph** G_m. The nodes in G_m are the candidates and there is a directed edge from A to B if A wins the 1-1 contest

	1-1 results		
Votes (V)	$A\|B$	$A\|C$	$B\|C$
A B C	A	A	B
C A B	A	C	C
B C A	A	C	B
Winners	A	C	B

Table 5: An election where there is no Condorcet winner.

$A|B$. If both are tied then the edge is bi-directional. A Condorcet winner will have $(m-1)$ outward directed edges with no incoming edge. Similarly, a Condorcet loser will have $(m-1)$ incoming edges with no outward edge. A weak Condorcet winner will also have $(m-1)$ outward edges but will have at least one incoming edge. Similarly, the Condorcet loser will have $(m-1)$ incoming edges but will also have at least one outgoing edge. Figures 6 and 7 show the digraph representation for the elections in tables 4 and 5. Notice the directed cycle in figure 7 which is called the **Condorcet or top cycle**. A top cycle that involves the candidates with the maximum 1-1 wins in the majority graph implies there is no Condorcet winner. This is also called the Condorcet paradox.

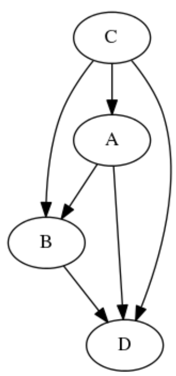

Figure 6: Majority graph for election in table 4. C is a Condorcet winner and D the Condercet loser.

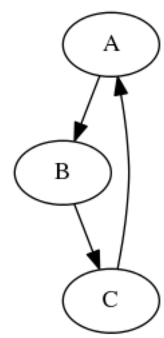

Figure 7: Top cycle in the election in table 5. No winner.

Definition 9 (Condorcet criterion)**.** *A VS satisfies the* `Condorcet (weak Condorcet) criterion` *if it chooses as winner the Condorcet (weak Condorcet) winner if one exists. Also, called Condorcet consistent.*

Note that the definition allows the VS to select a winner when there is no Condorcet winner.

8.2.2 Copeland systems

The Copeland system gives points for 1-1 contests as follows:
Winner-1; Loser-0; Tie-$\alpha \in [0,1]$, α a rational number. So, they are often called Copeland$^\alpha$ systems. Llull gave $\alpha = 1$, Copeland originally used $\alpha = \frac{1}{2}$. Many systems that give different numbers of points for winning and a tie can be normalized to a Copeland$^\alpha$ system. For example, many leagues give: Winner-3; loser-0; tie-1 gives a system with $\alpha = \frac{1}{3}$ after normalization.

Theorem 4. *If there is a Condorcet winner in an election with m candidates then the Condorcet winner is the only one to get a maximum Copeland$^\alpha$ score of $(m-1)$ and is the Copeland$^\alpha$ winner for all α.*

Example: Consider the election $\mathcal{E} = (\{A, B, C, D\}, V)$ with the votes as in table 6. The election in table 6 has a A as the weak Condorcet winner and B as the weak Condorcet loser. For weak Condorcet winners the value of α in the Copeland system is important. Consider the the two tables 7 and 8.

	1-1 results					
Votes (V)	$A\|B$	$A\|C$	$A\|D$	$B\|C$	$B\|D$	$C\|D$
A D C B	A	A	A	C	D	D
C D B A	B	C	D	C	D	C
C D B A	B	C	D	C	D	C
B D A C	B	A	D	B	B	D
A C D B	A	A	A	C	D	C
A C B D	A	A	A	C	B	C
Winner	A,B	A	A,D	C	D	C
Votes	3:3	4:2	3:3	5:1	4:2	4:2

Table 6: Election with weak Condorcet winner A, weak Condorcet loser B.

Candidate	Copeland score
A	$1 + 2\alpha$
B	α
C	2
D	$1 + \alpha$

Table 7: Copeland α scores for table 6.

α	Winner
0	C
$0 < \alpha < \frac{1}{2}$	C
$\frac{1}{2}$	A, C
$\alpha > \frac{1}{2}$	A

Table 8: How α value changes winner in Copeland.

8.2.3 Dodgson voting system

If the election has a Condorcet winner then she is also the Dodgson winner. Otherwise, we calculate a Dodgson score (dScore) for each candidate. This score is the minimum number of nearest neighbour swaps in votes that will make the candidate a Condorcet winner. The Dodgson winner is the candidate with the minimum dScore. For the election in table 6 we have the following dScores: dScore(A)=2; dScore(B)>2; dScore(C)=2; dScore(D)>2. So, A, C are Dodgson winners.

8.2.4 Simpson or Maxmin voting system

Let $N(A, B)$ be the number of votes that prefer A to B in a 1-1 contest. For the election in table 6 we have: $N(A, B) = N(B, A) = N(A, D) = N(D, A) = 3$, $N(A, C) = N(C, D) = N(D, B) = 4$, $N(B, D) = N(C, A) = N(D, C) = 2$, $N(B, C) = 1$, $N(C, B) = 5$. The Simpson score (sScore) for a candidate X is defined by:

$$\text{sScore}(X) = \min_{X \neq Y} N(X, Y)$$

. The sScore for each candidate is: sScore(A)=3, sScore(B)=1, sScore(C)=2, sScore(D)=2. The winner(s) are candidates with the maximum Simpson scores. In this case A.

Theorem 5. *A Condorcet winner is always a Simpson winner.*

8.2.5 Young voting system

The Young score (yScore) of a candidate is the minimum number of votes that must be deleted to make the candidate a weak Condorcet winner. A similar strong Young score (syScore) can be defined for a Condorcet winner. The Young winner is the candidate with the minimum yScore or syScore. For the example in table 6 we get: yScore(A)=0, yScore(B)=4, yScore(C)=2, yScore(D)=2. So, A is the Young winner.

8.2.6 Kemeny voting system

Kemeny's system allows partial ordering in votes (that is indifference), for example $A > B, C > D$ means indifferent between B and C. Define the distance between two votes for a pair of candidates as follows. Let $v_1, v_2 \in V$ then

$$d_{v_1,v_2}(X,Y) = \begin{cases} 0 & if v_1, v_2 \text{ agree on ordering of } X, Y \\ 2 & if v_1, v_2 \text{ strictly disagree on ordering of } X, Y \\ 1 & \text{otherwise} \end{cases}$$

Agreement means $X > Y$ or $Y > X$ or $X = Y$ in both v_1 and v_2; strict disagreement means $X > Y$ in v_1 and $Y > X$ in v_2 or vice versa; otherwise $X = Y$ in v_1 or v_2 and either $X > Y$ or $Y > X$ in the other. Define

$$d(v_1, v_2) = \sum_{X,Y \in C} d_{v_1,v_2}(X,Y)$$

The kScore for a vote is defined as:

$$kScore(v_1) = \sum_{v_2 \in V} d(v_1, v_2)$$

Note that the above is calculating a score for each vote (not candidate). The Kemeny winners are the candidates who are ranked first in the votes with minimum kScore.

Dodgson, Young and Kemeny winners are hard to compute.

8.2.7 Schwarz sets, Schulze voting system

Schwarz sets are useful in characterizing some voting systems. We first define a Schwartz component.

Definition 10 (Schwartz component, Schwartz set)**.** *A* `Schwartz component` *$S \subseteq C$, $S \neq \varnothing$ is such that for every $X \in S$, S is unbeaten by any candidate in $C - S$ and no non-$\varnothing$ proper subset of S satisfies the previous property.*

A `Schwartz set` *is a union of Schwartz components.*

Any VS that elects candidates from the Schwartz set satisfies the Schwartz criterion.

The Schwartz set satisfies the following properties which are stated in the form of lemmas. They follow almost directly from the definition.

Lemma 1. *A Schwartz is always non-empty.*

Lemma 2. *Any two distinct Schwartz components are disjoint.*

Lemma 3. *If there is a Condorcet winner then it is the only member of the Schwartz set.*

Lemma 4. *If a Schwartz set component contains only one member it is a weak Condorcet winner.*

Lemma 5. *If a Schwartz component has multiple members then they are members of a top cycle.*

Lemma 6. *Any two candidates in different Schwartz components are pairwise tied in 1-1 contests.*

We now describe the Schulze VS a widely used system in small communities, especially for electing leaders of open source communities like Debian, FSF and others. It is also used by some smaller political parties to elect leaders. The Schulze VS satisfies the Schwartz criterion.

To find the winner in a Schulze VS we first construct a labelled digraph for $\mathcal{E}$. The node set is C and the edge between two nodes X, Y is labelled with a weight equal to $N(X,Y)$ - that is the number of votes that prefer X to Y. We then find the `strongest path` between all pairs of nodes. The strongest path $p(X,Y)$ is defined as:

$$p(X,Y) \equiv \max\{w | \text{w is the weight of the minimum weight edge of a path between X and Y}\}$$

To get the strongest path value we first collect the weakest links in all paths between X and Y and then pick the maximum out of these values. The winner of $\mathcal{E}$ is the candidate X such that:

$p(X,Y) \geq p(Y,X)\ \forall Y \in C - \{X\}$

We will now find the Schulze winner for the election in table 6. The digraph for this is shown in figure 8. This has been converted into a matrix with entries $N(X,Y)$ in table 9. The strongest path values between pairs of nodes is shown in table 10 - these can be calculated from table 9 or from the digraph in figure 8. It is clear from table 10 that A is the Schulze winner.

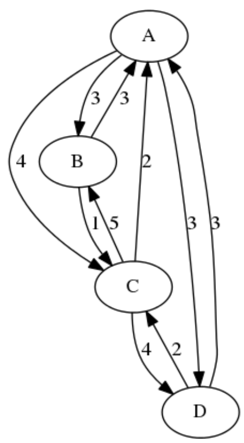

Figure 8: The digraph for finding Schulze winner.

	A	B	C	D
A	-	3	4	3
B	3	-	1	2
C	2	5	-	4
D	3	4	2	-

Table 9: Matrix representation of digraph in figure 8.

	A	B	C	D
A	-	(4)	(4)	(4)
B	3	-	3	3
C	3	5	-	4
D	3	4	3	-

Table 10: Strongest path matrix of digraph in figure 8.

8.3 Approval voting

In approval voting the vote is not a linear order on the candidate set C. So, it is fundamentally different from the previous system where the vote was a linear ordering of candidates.

In its simplest form an approval vote divides C into an approved set and a not approved set. More complex forms of an approval vote allow approval on a scale from 0 to k - this is often called `k-range voting`.

A simple way to represent an approval vote is as a vector of length $|C|$ where it is assumed that the candidates are placed in some sequence $c_1, \ldots, c_m$ and the vote vector $\vec{v} = (v_1, \ldots, v_m)$ gives the intensity of approval such that v_i is the approval intensity for candidate c_i. For k-range voting the intensity is often normalized so that the intensity score for candidate c_i in a vote v is given by $k \times \frac{v_i - v_{min}}{v_{max} - v_{min}}$ where v_{max}, v_{min} are the maximum and minimum intensity values in the vote v. If (v_{max} equals v_{min}) the vote is dropped because it cannot affect any candidate either positively or negatively.

The winner in an approval system is the candidate(s) with the maximum approval score calculated by adding the normalized approval intensities in each vote in V for each candidate. Approval voting is manipulable and it is easy to vote strategically. For this reason in approval voting we are concerned with 'sincere' votes. These are votes that are consistent with a voters real ordinal choices which means that if a voter approves some candidate X then the voter also approves all Y such that $Y > X$. Suppose the ordering of a voter is $A > B > C > D$. Then all the following votes will be considered sincere: A; A,B; A,B,C; A,B,C,D. It is clear that a sincere approval voting system will not satisfy the Condorcet criterion.

8.4 Multiple round voting systems

Some voting systems use multiple rounds. In most such systems each round tries to either pick a proper subset of candidates or eliminate some candidates. Consequently, in finitely many rounds we are left with the winner(s).

8.4.1 Plurality with run-off

This is a two round voting system. Again we assume that a vote is a linear order on the set of candidates.

Round 1:

- Pick the top two with the most top preferences for round two.
- If there are any ties break ties to pick only two.
- If any candidate has a majority (greater than 50%) of the votes cast she is a winner else go to round 2.

Round 2:

- From each vote cancel all other candidates except the two winners in round 1.
- Recount the top preferences for the two candidates.
- If there is a tie break the tie for a winner else the candidate with the most top preferences is the winner.

Example: Consider the election in table 6 with $C = \{A, B, C, D\}$.

Votes (V)	Round 1	Votes after Round 1	Round 2
B D A C	A-3	A C	A-4
A D C B	B-1	A C	C-2
C D B A	C-2	C A	
A C D B	D-0	A C	
A C B D		A C	
C D B A		C A	

Table 11: Plurality with run-off for table 6 election. A is the winner.

Another example of plurality with run-off is the French President's election. In round 1 voters just give a single preference which is used to decide the top two candidates. In round 2 there is a run-off between the top two where voters vote again and give a single preference. The winner in round 2 is the candidate who gets a majority of the votes.

8.4.2 Single tranferable voting system

This is a multiple round system in which some candidates are eliminated in each round.

- **repeat**
- Count first preferences for all candidates.
- If one candidate has a majority she is the winner.
- Delete the last candidate(s) by removing them from all the votes.
- **until** (winner found or there is a tie; break ties)

Example: Consider the election in table 6 with $C = \{A, B, C, D\}$.

Votes (V)	Round 1	Votes round 2 (D-)	Round 2	Votes round 3 (B-)	Round 3
B D A C	A-3	B A C	A-3	A C	A-4
A D C B	B-1	A C B	C-2	A C	C-2
C D B A	C-2	C B A	B-1	C A	
A C D B	D-0	A C B		A C	
A C B D		A C B		A C	
C D B A		C B A		C A	

Table 12: Single transferable voting system for table 6 election. A is the winner.

Single transferable voting is used in multiple countries and also to elect the President of India.

8.4.3 Bucklin voting system

This is a multi-stage elimination system that considers progressively lower preferences in each stage till a winner is found. Define $M = \lfloor \frac{|V|}{2} \rfloor + 1$.

- stage=1
- **step 1:** Give 1 point to each candidate who is present in preferences from 1 to stage. Put all candidates with points greater than equal to M in bucket B and go to step 2 if B is not empty.
- stage++
- repeat from step 1
- **step 2:** The candidate with maximum points in bucket B is the winner.
- If there is a tie either declare all those in B as winners or use a tie breaking rule to get a single winner.

Example: Consider the election in table 6 with $C = \{A, B, C, D\}$. The value of M is $\lfloor \frac{6}{2} \rfloor + 1 = 4$.

Votes (V)	stage=1	stage=2
B D A C	A-3	A-3
A D C B	B-1	B-1
C D B A	C-2	C-4
A C D B	D-0	D-4
A C B D		
C D B A		

Table 13: Bucklin for table 6 election. C, D are the winners.

Unlike other systems this system has picked a different set of winners. Surprisingly D is a winner. But it is not unreasonable given that D has 4 second preference votes. Bucklin is also a fairly widely used system.

8.4.4 Borda multiple rounds system

There are several multiple round Borda systems. The simplest one is where in each round the candidate with the least Borda score is eliminated and then the process is repeated with the remaining candidates after calculating new Borda scores. This is repeated until we get a winner(s). Note that eliminating a candidate means removing the candidate from all votes.

Another variant uses the average Borda score as a threshold and eliminates all those who are on or below the threshold but is otherwise similar to the one that eliminates only the last candidate.

Example: We repeat the average Borda threshold system for the election in table 6.

	Votes V						
Candidates	BDAC	ADCB	CDBA	ACDB	ACBD	CBDA	Borda score
A	1	3	0	3	3	0	10
B	3	0	1	0	1	1	6
C	0	1	3	2	2	3	11
D	2	2	2	1	0	2	9
Avg=6, eliminate B							
	DAC	ADC	CDA	ACD	ACD	CDA	Borda score
A	1	2	0	2	2	0	7
C	0	0	2	1	1	2	6
D	2	1	1	0	0	1	5
Avg=6, eliminate C and D. Winner is A.							

Table 14: Borda with average for table 6 election. A is the winner.

Notice that the orginal single stage Borda winner is C.

8.5 Hybrid systems

A hybrid system uses two or more voting systems usually in multiple steps to decide winners. There are many hybrid systems. We will look at two.

8.5.1 Black voting system

The debate between whether Condorcet or Borda is better still continues. Black's system is a hybrid of both in the simplest possible way. If a Condorcet winner exists then she is also a Black winner, if not all Borda winners are Black winners.

8.5.2 Fallback voting system

This system assumes that a vote has both approvals and linear orders. Each vote is split between candidates who have been approved and those who are not. Then the Bucklin system is applied using only the approved candidates on each vote. If a Bucklin winner then she is also a fallback winner otherwise all approval winners are fallback winners.

Example: Consider the election in table 15 which has four candidates A, B, C, D. The vertical line in the linear ordering separates the approved (on left) from the not approved (right of line). The value of M for Bucklin is 4.

Votes order and approval	A|BCD	CDB|A	|ABCD	BD|AC	A|BCD	ACB|D

Table 15: A fallback vote. Left of vertical are approved.

Stage	A	B	C	D
1	3	1	1	0
2	3	1	2	2
3	3	3	2	2
No Bucklin winner, find approval winners				
Approvals	3	3	2	2
A, B are approval winners				

Table 16: A fallback example.

There is a fair amount of consensus that a Condorcet winner, if one exists, is a legitimate winner. However, a Condorcet winner does not always exist. And the probability that one exists goes down with increasing m (number of candidates) and increasing n (number of voters).

Different voting systems often produce different winners and it is hard to say which voting system is to be preferred. To study this question more systematically we will define intuitive properties that we would like a voting system to satisfy. The assumption is that the more properties a voting system satisfies the more the possibility that it is better than another voting system that satsifies fewer properties.

9 Desirable properties of voting systems

There are many commonsense, intuitive properties that we would like voting systems to have. Below we describe the most common such properties and then create a table that summarizes information on which voting systems have which properties. In what follows, unless otherwise noted, we assume that a vote is a linear order on the set of candidates or alternatives.

9.1 One-person-one vote or reduced preferences

The one-person-one-vote property of a voting system (VS) refers to a lack of discrimination in who casts a vote but more specifically, in the context of social choice, between votes. This means votes with the same preferences can be aggregated and the result depends only on such aggregated numbers. This is often called `reduced preferences`. For example, in an election with three candidates A, B, C if the individual preferences of 6 voters are:
v1-ABC, v2-BAC, v3-ABC, v4-CAB, v5-BAC, v6-ACB
then the aggregated values are ABC-2, BAC-2, ACB-1, CAB-1. This reduction delinks the individual voter from the preference so every voter's vote has the same worth.

9.2 Majority criterion

This comes in multiple flavours. A VS satisfies the majority criterion when if a candidate has a majority of the votes (that is $\geq$50%+1) as determined by the nature of the vote (e.g. a candidate is preferred to another by a majority of voters - called ranked majority criterion) the candidate is a winner. A related criterion is the simple majority criterion where the winner has at least 50% of the votes. Plurality obviously satisfies the majority criterion. Borda does not. Consider an election with the following preference profile: ABC-3, BCA-3, CAB-5. The Borda scores are: A-11, B-9, C-13 so C wins. But B is preferred to C by 6 voters.

9.3 Pareto consistency

A VS is Pareto consistent when: Given a preference profile V if $C > D$ in all the votes in V then $C > D$ in the final outcome. This is directly relevant for a social welfare function (SWF). For a social choice function what it means is that D can never be the winner for the given preference profile. A related property is Pareto efficiency. A VS is Pareto efficient if a candidate who is ranked first in every vote is the winner. All the VSs discussed earlier are Pareto consistent.

9.4 Condorcet criterion

A VS satisfies the Condorcet criterion or is Condorcet consistent if a candidate who wins 1-1 contests with all other candidates is also the VS winner or equivalently whenever a Condorcet winner exists the candidate is also a VS winner. Plurality does not satisfy the Condorcet criterion. For example consider the profile: CAB-5, BAC-3, ABC-3. The plurality winner is C (5 first preferences compared to 3 for B and 3 for C) but A beats both B and C in 1-1 contests and is the Condorcet winner. Borda also does not satisfy the Condorcet criterion.

9.5 Consistency or separability or convexity

A VS is consistent if when V is partitioned into $V_1, V_2, \ldots, V_k$, $k \geq 2$ the candidates who win in each partition also win with preference profile V - that is when V is not partitioned.

This is similar to Simpsons paradox in probability theory. For example, plurality with run-off is not consistent. Consider the election: V: ABC-417, BAC-143, CBA-224, ACB-82, BCA-357, CAB-285. Split V into V1, V2:
V1: ABC-160, BAC-143, CBA-285
V2: ABC-257, ACB-82, CBA-39, CAB-285, BCA-357

Plurality winner of V1-C,A, runoff winner A; plurality winner of V2-B,A, runoff winner A; plurality winner of V-C,B, runoff winner B.

Several other systems are also not consistent.

9.6 Resoluteness

A resolute VS is one that chooses only a single winner. A VS that either has no winner or chooses more than one is not resolute. A VS that normally chooses more than one winner can be made resolute by incorporating a tie breaking rule. Condorcet, Borda are not resolute.

9.7 Homogeneity

Given an election $\mathcal{E} = (C, V)$ define $\mathcal{E}' = (C, V^k)$, $k \geq 1$ where each vote in V is duplicated k times in V^k. Then a VS is homogeneous if $\forall V$ and $k \geq 1$, $f(\mathcal{E}) = f(\mathcal{E}')$ where f is a social choice or social welfare function.

9.8 Independence of clones

A clone of a candidate C is a candidate who is right next C (either before or after) in every vote. A VS is independent of clones if it is not possible for a candidate who is not winning originally to win by introducing clone(s). Plurality fails to have this property. For example, consider an election where the vote profile is: A-13, B-23. After introducting a clone C for B the vote splits and becomes: A-13, B12, C-11. This is a common manipulation in elections that use plurality especially in India.

9.9 Anonymity

For anonymity the result of an election should not depend on what order votes are cast. It should depend only on the aggregate information of votes cast. So, all vote profile permutations are identically processed.

9.10 Strategy proofness

A VS is strategy proof if no voter benefits by giving a false/untruthful preference or vote. If a VS is not strategy proof it is manipulable. While this is a desirable property it is hard to design voting systems that have it. For example: consider Borda's system with the following profile:
v1, v2 - ABCD
v3 - DCBA is actual preference but votes BDCA
Voter v3 does not want candidate A (last choice) to win. In Borda's VS A would have won with 6 points if v3 had voted DCBA. But by voting BDCA v3 makes B the winner and prevents A from winning.

9.11 Neutrality

A neutral VS treats all candidates equally. This means if the preferences of two candidates are swapped on all votes then the candidates positions on the output ranking should also be swapped. All VSs discussed earlier are neutral.

9.12 Monotonicity

A VS is monotonic if the following holds:
If X is a winner and X's preference ranking on some votes is improved while all other rankings remain unchanged then X still remains the winner.

A related property is strong monotonicity. A VS is strongly monotonic if a winner X in a profile V is also a winner in a profile V' where X's position improves in some votes while those who were above X or below X can be reordered freely but they cannot cross from below to above or vice versa. Clearly, strong monotonicity imples monotonicity. All scoring based systems are monotonic. Dodgson is not monotonic.

9.13 Participation criterion

If a candidate favoured by a voter gains by the voter not voting then the VS is said to have a no-show paradox. A VS for which the no-show paradox does not occur satisfies the participation criterion.

A more general version of the property is: an additional vote should not produce a winner X who is ranked lower than the original winners in the vote implying thereby the voter is better off not voting. This is also a form of manipulative or strategic voting. All scoring protocols satisfy the participation criterion. Those that violate the participation criterion are STV, plurality with runoff amongst others.

9.14 Twin paradox and welcome twin criterion

Consider a tournament VS with the schedule $(A|B)|C$ and a tie-break rule that picks the earliest alphabet as the winner. Let the preference profile be:
1 - CBA
2 - ABC
1 - CAB
2 - BCA

So, in $A|B$, A wins by tie-break since tied at 3:3 and in $A|C$ C wins 4:2 and C is the winner. Now suppose we clone the first vote so we have 2 - CBA. This means B wins 4:3 and in $B|C$ B again wins 4:3 and B is the winner. So, the candidate favoured by the voter voting CBA, namely C, loses when a twin's vote (the same vote) is introduced.

The welcome twin criterion is satisfied by a VS if cloning a voter and adding it to V does not do harm to their preferred candidates.

9.15 Citizens sovereignity

A VS satisfies the citizens sovereignity criterion if the outcome of an $\mathcal{E}$ is completely decided only by the preference profile and nothing else. This implies that for every candidate $X \in C$ there is a voting profile V such that $f(V) = X$ or equivalently X is at the top of $\rho(C)$ society's ranking if f is a SWF. A simpler way to describe this criterion is to say that the function f is surjective. All voting systems discussed satisfy this criterion.

9.16 Independence of irrelevant alternatives (IIA)

This is a seemingly intuitive, but quite problematic requirement that has been debated extensively in the literature. The basic requirement is that the outcome of an election between any two candidates should depend only on the relative position of the two candidates in the preference profile. All other candidates and their relative positions in the linear order should be irrelevant to society's preference for these two candidates.

This condition is more relevant when we are considering a social welfare function. It implies that the outcome or more precisely society's relative position in the linear order of the two candidates in question, say X, Y, should not be affected if new candidates are inserted into the vote or existing candidates (barring X, Y) are removed from the vote or the relative order of the other candidates is shuffled. This, at first sight, seems like a reasonable requirement. All it demands is that in any $\mathcal{E}$ if we pick any two candidates then the outcome, that is their relative position in society's linear order should not change between two voting profiles V and V′ where the relative positions in the linear order of X and Y in each vote in the two profiles does not change. However, none of the voting systems discussed in section 8 that always produce a winner satisfies this condition and trying to simultaneously satisfy IIA with some other conditions leads to impossibility results which implies that the condition is more complex than it appears.

9.17 Non-dictatorship criterion

A VS satisfies the non-dictatorship criterion if there does not exist a voter whose preference profile determines the social linear order of an election for all preference profiles. The election is completely decided by a single voter's linear order. Such a voter if one exists is called the dictator since the voting preferences of this voter completely decides the election outcome in the sense that if this voter changes her preferences society also changes its preferences in the same way.

The non-dictatorship is a property of the social choice or social welfare function and not of a particular voting profile in the domain of f. Also, note that the domain of f is assumed to contain all logically possible preference profiles, that is it is a universal domain.

Let i be some voter, n the number of voters, V a voting profile, X, Y any two candidates, $>$ the individual voter's preference relation and $\underset{s}{>}$ the social preference relation. Then formally the non-dictatorship condition can be stated as ($\ni$ means such that):

$$\nexists i \in 1..n \ni \forall V[(X > Y) \implies (X \underset{s}{>} Y)]$$

If a dictator exists then there exists a voter whose individual preference determines the social preference for any pair of candidates for all voter profiles V. Another important consequence of the definition is that a dictator, when one exists, is unique.

	Voting System							
Property	Plurality	Borda	Copeland	Dodgson	Young	Bucklin	STV	Schulze
Anonymity	✓	✓	✓	✓	✓	✓	✓	✓
Neutrality	✓	✓	✓	✓	✓	✓	✓	✓
Non-dictatorship	✓	✓	✓	✓	✓	✓	✓	✓
Pareto consistency	✓	✓	✓	✓	✓	✓	✓	✓
Citizen's sovereignity	✓	✓	✓	✓	✓	✓	✓	✓
Majority criterion	✓	✗	✓	✓	✓	✓	✓	✓
Condorcet criterion	✗	✗	✓	✓	✓	✗	✗	✓
Monotonicity	✓	✓	✓	✗	✓	✓	✗	✓
Homogeneity	✓	✓	✓	✗	✗	✓	✓	✓
Indp. Of Irr. Alt	✗	✗	✗	✗	✗	✗	✗	✗
Indp. of clones	✗	✗	✗	✗	✗	✗	✓	✓
Strategy proof	✗	✗	✗	✗	✗	✗	✗	✗
Particip. criterion	✓	✓	✗	✗	✗	✗	✗	✗
Consistency	✓	✓	✗	✗	✗	✗	✗	✗

Table 17: Summary table of voting systems and their properties.

10 Impossibility results

A large number of impossibility results have been disovered starting with Arrow's famous result in 1951. In this section we state and prove Arrow's impossibility theorem[2] and state a few well known other impossibility theorems.

10.1 Arrow's theorem

Let $\mathcal{E}$ be an election where $C = \{c_1, \ldots, c_m\}$, $m > 2$ is a set of candidates, $V = \{v_1, \ldots, v_n\}$ is a preference profile where each v_i, $i \in 1..n$, is a linear order on C, basically the ith voter's linear order. The order relation of the individual voter is denoted by $>$ ($c_i > c_j$ means c_i is preferred to c_j). We also have a social welfare function $f : \{V | \mathcal{E} = (C, V) \text{ is an election}\} \rightarrow \rho(C)$ which maps a preference profile V to a linear order on C - ρ is an ordering or ranking function. The social linear order relation is denoted by $\underset{s}{>}$. Function f aggregates the votes or individual linear orders and computes a ranking/linear order on all the candidates in C. f is assumed to have universal domain. That is any logically possible preference profile is a valid point of f's domain. Note also that $>$ and $\underset{s}{>}$ are transitive relations since they are linear orders.

Theorem 6 (Arrow's theorem)**.** *Given linear orders $>$ and $\underset{s}{>}$ (as above) there does not exist an SWF f, that simultaneously satisfies a) Pareto consistency b) IIA c) non-dictatorship.*

Another way to say it is: Any SWF f that satisfies Pareto consistency and IIA must be dictatorial.

The proof argues that SWF f that satisfies IIA and Pareto consistency will have a dictator. We do this by proving 4 lemmas. In what follows we use voter and vote interchangeably for stylistic reasons to refer to a preference in preference profile. We assume that f satisfies Pareto consistency and IIA.

Lemma 7. *Let $c \in C$. Then any V in which c is either ranked first or last in every $v \in V$ will be ranked either first or last in the social order.*

If $c \in C$ is such that $\forall v \in V[\forall i(c > c_i, \text{ or } c_i > c) \implies \forall i(c \underset{s}{>} c_i \text{ or } c_i \underset{s}{>} c)]$.

Proof. The proof is by contradiction. Assume c is not first or last in the social ranking. Then $\exists c_i, c_j \in C$ s.t. $c_i \underset{s}{>} c \underset{s}{>} c_j$ and $c_i > c_j$ in some votes in V. Now move c_j s.t. $c_j > c_i$ in every $v \in V$. Let this preference profile be V'. Since c is either first or last in every vote the relative rank of the pair c, c_i and c, c_j remains the same in V' and V. That is in votes where c is first we continue to have $c > c_i, c > c_j$ and votes where it is last we continue to have $c_i > c, c_j > c$. By IIA the relative ranks of the pairs c, c_i, c, c_j should not change in the social ranking $f(V')$ that is $c_i \underset{s}{>} c$, $c \underset{s}{>} c_j$ continues to hold. But in V', $c_j > c_i$ in every $v \in V'$ so $c_j \underset{s}{>} c_i$ by Pareto consistency. By transitivity we have $c_j > c$ which is a contradiction. So our assumption is incorrect and c must indeed be either first or last in the social ranking. □

Lemma 8. *There exists a voter profile $V^\star$ in which there exists a voter $v^\star$ who is a pivotal voter. A pivotal voter is one who can change the social ranking of c from bottom to top by changing the ranking of c from bottom to top in $v^\star$.*

Proof. Start with a preference profile V such that c is at the bottom of each $v \in V$, c is an arbitrarily chosen candidate. Then c is also at the bottom in the social ranking by Pareto consistency. Assume we order all voters/votes in V in some arbitrary order from 1 to n. Starting with v_1 we shift c from the bottom to the top leaving everything else unchanged and repeat this successively for $v_2, v_3, \ldots, v_n$. Again, due to Pareto consistency and lemma 7 there must exist some $k \in 1..n$ when c shifts from bottom to top in the social ranking exactly when voter v_k shifts c from bottom to top in her vote. This v_k is our pivotal voter $v^\star$. □

Let V_1 be the preference profile just before v_k or $v^\star$ switches c from bottom to top and V_2 be the preference profile immediately after $v_\star$ has switched c to the top.

[2]Adapted from John Geanakopolos, Three brief proofs of Arrow's theorem, Economic theory, 26, 211-215, 2005.

Lemma 9. *The pivotal voter $v^\star$ is a dictator for any pair c_i, c_j different from c. Without loss of generality assume that $c_j > c_i$.*

Proof. Construct preference profile V_3 from V_2 by moving c_i immediately before c in $v^\star$ and for all other $v \in V_2$ rearrange relative ranking of c_i, c_j arbitrarily without moving c which is either first or last. So now we have $c_i > c > c_j$ in $v^\star$.

For V_1 we have social ranking $c_i \underset{s}{>} c$ and $c_j \underset{s}{>} c$ since c is at the bottom of the social ranking. In V_2 we have social ranking $c \underset{s}{>} c_i$ and $c \underset{s}{>} c_j$ since c is now at the top. In V_3 the relative ranking of the pair c_i, c is exactly as for V_1 since c_i has been moved immediately before c. So for V_3, by IIA relative social ranking is $c_i \underset{s}{>} c$. For the pair c_j, c the relative social ranking is exactly as it was for V_2. By transitivity for V_3 we get the relative social ranking $c_i \underset{s}{>} c \underset{s}{>} c_j$. So, voter $v^\star$ has been able to change the relative social ranking of the pair c_i, c_j by making a change in her vote. Note that other votes are allowed to change relative rankings arbitrarily. So, the change is only due to the change in $v^\star$ and $v^\star$ is a dictator. □

Lemma 10. *$v^\star$ is also a dictator for any pair c, c_i.*

Proof. Choose c' that is different from c and c_i. Now repeat with c' everything we did with c. That is start with a preference profile, say V' that has c' at the bottom of each $v \in V'$ etc. Then we will have a pivotal voter, say $v^{\star\star}$, who is a potentially new dictator for any pair c_i, c_j different from c'. In particular $v^{\star\star}$ is a dictator for the pair c_i, c. But note that for preference profiles V_1, V_2 we have already shown that $v^\star$ is a dictator and completely determines society's relative ordering for c_i, c. So, $v^{\star\star}$ must be the same as $v^\star$ because for V_1, V_2 the dictator must be unique. Therefore, $v^{\star\star} = v^\star$. □

Arrow's theorem also holds when preferences are not strict which means individuals are allowed to be indifferent between candidates.

10.2 Comments on Arrow's theorem

Since the preference based VSs we have discussed in section 8 satisfy non-dictatorship and Pareto consistency none of them satisfy the IIA criterion. Arrow's theorem requires that the domain of the SWF is all possible preference profiles. There have been attempts to constrain the domain. Human preference profiles tend to be single peaked. That is a large number of voters prefer one or more attributes to be at certain levels and the numbers drop of on either side of the preferred levels. A voter's preference for candidates reflects to what extent the candidates possess the required attributes at the desired levels. So preference profiles are possibly single peaked in real elections. Black has shown that if voter preferences are peaked along a single dimension then all Arrow conditions are satisfied by the majority rule.

Another point is that Arrow assumes ordinal preferences the degree or intensity to which one prefers one to the other (that is cardinality) is not important. So, a preference rating of 100, 10, 1 will be treated the same way as 100,99,98.Votes are assumed to be independent. In real elections both these conditions are not true.

10.3 Other impossibility theorems

We state a few other impossibility theorems below.

Theorem 7 (Gibbard-Satterthwaite). *Given an election with preference based votes $\mathcal{E} = (C, V)$ if $|C| > 2$ then there does not exist a VS that simultaneously satisfies:*

a) Non-dictatorship.

b) Resoluteness.

c) *Citizen's sovereignity.*

d) *Strategy proofness.*

Resoluteness implies it is a social choice situation but that is not essential. The theorem can be generalized to the social correspondence function.

Theorem 8 (Muller-Satterthwaite)**.** *Given an election with preference based votes* $\mathcal{E} = (C, V)$ *if* $|C| > 2$ *then there does not exist a VS that simultaneously satisfies:*

a) *Non-dictatorship.*

b) *Resoluteness.*

c) *Citizen's sovereignity.*

d) *Strong monotonicity.*

Theorem 9 (Moulin I)**.** *Given an election* $\mathcal{E} = (C, V)$*:*

1) *If* $|C| < 4$ *there exists a VS that satisfies the participation criterion and Condorcet criterion.*

2) *If* $|C| \geq 4$ *and* $|V| \geq 25$ *then there is no VS that simultaneously satisfies the participation and Condorcet criteria.*

Theorem 10 (Moulin II)**.** *Given an election* $\mathcal{E} = (C, V)$ *with* $|C| > 3$ *and* $|V| \geq 25 + \frac{|C|(|C|-1)}{2}$ *there is no VS that simultaneously satisfies the Condorcet criterion and the welcome twin criterion.*