Version 1.5

Making Voting Equitable (2/21/2022)

Prepared by:

Joseph Malkevitch
Department of Mathematics and Computer Studies
York College (CUNY)
Jamaica, New York

email:

malkevitch@york.cuny.edu

web page:

http://york.cuny.edu/~malk/

Introduction

The crux of democratic life is the electoral process. In the United States, we elect the President, members of Congress, and numerous state and local officials. We, as Americans, take pride in our elections. However, before the collapse of the Soviet Union they had elections, too, as did an emerging Nazi Germany. Elections are not the whole story behind democracy. What makes elections fair and equitable? Obviously one needs a system where a wide variety of individuals can vote and run for office, where voters can vote easily and freely, and their votes can be counted accurately. But is this enough?

Consider, for example, the results of the following two presidential elections:

Nixon, Richard	31,785,480
Humphrey, Hubert	31,275,166
Wallace, George	9,906,473

(minor party candidates omitted: election of 1968)

Bush, George	39,102,343
Clinton, William	44,908,254
Perot, Ross	19,741,065

(minor party candidates omitted: election of 1992).

Even more dramatic has been the election of 2000, where the winner of the popular vote Albert Gore, lost the election for president to George W. Bush in the electoral college. Though there are on going disputes about the counting of the ballots, the official "certified" votes for the three major candidates are indicated below:

George W. Bush	50, 456, 141 (47.87%)
Albert Gore	50, 996, 039 (48.38%)
Ralph Nader	2, 878, 157 (2.73%)

(minor party candidates eliminated.)

In 2016, once again the person who got more popular votes was not the person who becamse president.

Donald J. Trump	Republican	304	62,984,828
Hillary R. Clinton	Democratic	227	65,853,514
Gary Johnson	Libertarian	0	4,489,341
Jill Stein	Green	0	1,457,218
Evan McMullin	Independent	0	731,991

Note that more than 6 million votes were cast for candidates other than the two front runners.

What is remarkable about these three elections is the unusual presence, for America, of candidates from three parties, all of whom received significant numbers of votes. In neither of these elections was the person who won elected with a *majority*, that is, with more than half of the total votes cast. Instead, only a *plurality* winner, the person with the largest number of votes, was elected. Is it ideal in a democracy for someone to be elected with only a plurality? The fact that in many elections in America, a run-off election between two candidates is required when there is no majority candidate shows the "malaise" with plurality voting. However, run-off elections are inefficient. First, they are costly, and often the second phase of such elections results in dramatically fewer people voting than in the original election, which calls into question the fairness and wisdom of this procedure altogether. Second, especially when the election has many candidates who split the vote in a complex way, run-offs may eliminate a candidate who has less than one of the two largest first-place tallies but who is very popular with nearly all the voters as indicated by the fact that the eliminated candidate is among their top choices. Against this backdrop of discomfort with dealing with deciding the winner of multiple candidate elections, we will proceed. Before doing this, it is worthwhile to broaden the context of the discussion, as is typical when mathematical thinking comes into play.

Decision making between alternatives is what is involved here, but such behavior is not confined merely to elections. Members of committees must decide which action is to go forward as that recommended by the committee. The people who decide on the prizes at a beauty contest or skating contest must produce a ranking of the participants. The winner of the Cy Young award, best picture of the year, and best play must be decided upon. Sometimes we require a single winner, sometimes several seats are to be filled and several "winners" must be chosen, and sometimes a ranking, from best to worst, of all the alternatives being evaluated is necessary.

Elections and rankings

If we are to make a careful study of the process by which elections are conducted and rankings are made, with a mind to doing this in as fair and equitable a way as possible, we will need to examine the component parts of such a system.

Take a typical case: What should be served at the school club luncheon: pizza, Chinese takeout, or salads?

Typical of voting or ranking situations, we have a collection of alternatives to chose from: $A_1,...,A_n$. These alternatives may be the candidates for president in an election, the food choices for the school luncheon, the "hunks" in a male pulchritude contest, or different policies that might be pursued by the Federal Reserve Board. We are especially interested in the case when the number of these alternatives n is 3 or more. The reason for this is that if there are only two candidates, there is a popular and attractive democratic election method, called *majority voting,* to use: each voter casts his/her vote for his/her favorite choice, and as long as there is an odd number of voters, one of the two candidates must have a majority. (If the number of voters is even, there might be a tie, and the election method must include a scheme for breaking ties. More be will said about this later.) The candidate who receives a majority wins. Note that merely because this election method is an attractive one does not mean that there might not be some even better method of conducting elections in the case of two candidates. However, the mathematician Kenneth May showed that under some simple assumptions concerning desirable fairness conditions, majority voting is the only election method of choosing between two candidates that obeys the fairness rules.

In addition to alternatives to choose from, we need people to do the choosing. This may seem a simple matter in theory, but in practice it is, in fact, quite complicated. In the context of presidential elections, should there be an age restriction on who can vote? Should felons be allowed to vote? It may seem hard to believe today, but at one time blacks and women were not allowed to vote in the United States. Generally, such questions are administered through a registration system, where those who think they meet the eligibility requirements to vote are asked to prove that they meet the requirements. Thus, a new voter must produce proof of age, proof of citizenship, etc. in order to be allowed to vote. Historically, various political parties have made it hard for people to register to vote for various reasons. (Even today, the idea that one should be allowed to register to vote at the time one applies for or renews a driver's license is controversial.) In the context of deciding on food for the school lunch, should only club members be polled or their guests as well? Though there are critical issues of fairness and equity in this regard, we shall not specifically pursue some of the interesting questions here. (Other questions come up in regard to allowing those eligible to vote to vote physically. Here I refer to such issues as the hours the polls are open, absentee ballot arrangements, and access to the polling place for people, who, say, must use a wheelchair, etc.)

Having agreed on the alternatives available and on who are the people who are allowed to reach the decision, it must be decided how the people who do the voting are to convey their feelings about the alternatives to those who conduct the election. This is done using a ballot.

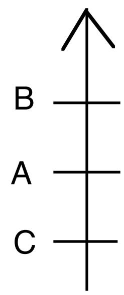
Ballot types

If you yourself have ever voted, you have been given a ballot. The ballot that nearly everyone is familiar with has a list of the alternatives which are to be considered. If there is one seat to be filled, you are instructed to vote for only one alternative. If there is more than one seat to be filled, you may be instructed to vote for as many candidates as there are seats to be filled. In any case, we will refer to these types of ballots as: Standard ballot (choose one) and standard ballot (choose k (where k is the number of seats to be filled)). The counting of the ballots raises issues concerning what to do if a person does not follow the instructions for filling out the ballot correctly. Ballot rules must be precise, but the rules can have subtle effects. For example, if k seats are to be filled, but the rules allow voters not to vote for candidates for all the positions (i.e. to vote for fewer than k candidates), what, if any, are the strategic advantages to a single voter or a group of voters, if they decide to vote for fewer than the k candidates they are allowed to vote for? (Voting behavior of this sort may arise when someone wants to see some particular candidate elected so strongly, that the person would forgo voting for any other candidates rather than have these additional votes defeat the one candidate they felt so strongly about! This example shows that there are strategic or "game theory" issues involved in voting situations. More about this will appear below, where the consequences of strategic or "sophisticated" voting (i.e. not voting in accordance with one's preferences regarding the alternatives) is studied in comparison with sincere voting.)

What is done with the standard ballots after the individual voters use them? They are counted in accordance with a predetermined election decision method and a winner (or winners) selected based on the election decision method.

If the standard ballot is the only kind of ballot that you have ever used, it is likely that you are thinking that our discussion here is rather extended for such a simple matter. Based on the standard ballot one is essentially limited to plurality voting or in the case of run-off systems, a determination that plurality voting has failed to elect a winner and a second round of elections must be conducted.

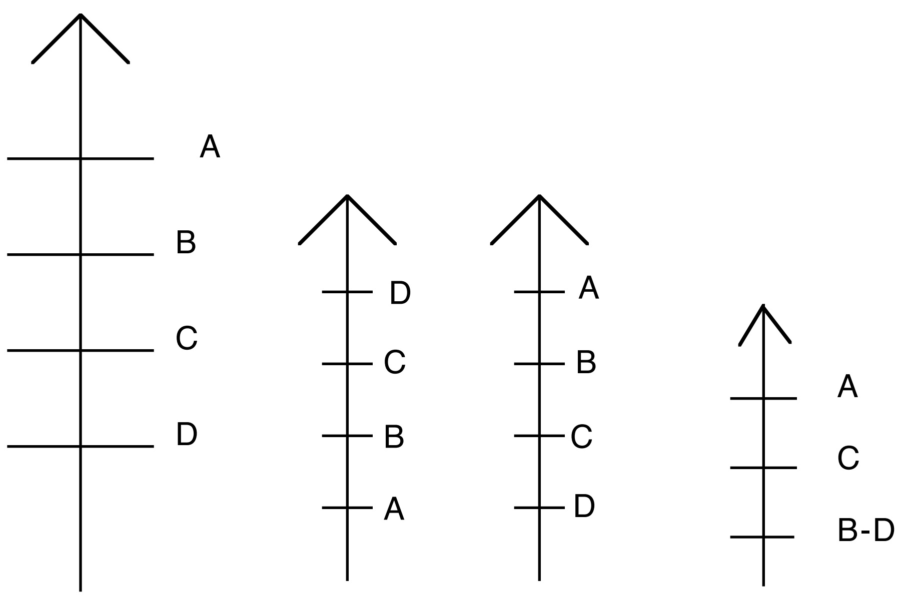
However, there are other kinds of ballots. Perhaps you have had a ballot where you were asked to rank the alternatives: 1 for the alternative you liked most, 2 for the alternative you liked next most, etc. The result of using such a ballot, called an *ordinal ballot,* can be represented geometrically. Thus, the result of a voter's ranking three alternatives A, B, and C so that A is ranked 2, B is ranked 1 and C is ranked 3 can be represented:



vote: 1

Although rare compared with the use of the standard ballot, there are numerous situations where such ballots are in use. Examples include many professional society elections and situations which involve rankings in entertainment or sports.

Here are some examples of ordinal ballots that might be produced in an election. Another name for such an ordinal ballot is a *preference schedule*.



voter: u v w x

Note that with an ordinal ballot people with the same first place choice of candidate can express differences of opinions concerning the other candidates. This can not be done with the standard ballot. Furthermore, the ordinal ballot allows a voter to express the fact that the voter is indifferent between two or more candidates. Hence, the ordinal ballot provides more information about how the individual voters view the candidates than can be done with a standard ballot.

Stop and Explore:

a. How many different types of ordinal ballot can a voter chose from if the voter ranks all the candidates and is not indifferent between any of the candidates and there are three candidates? (Can you give a formula when there are n candidates?)

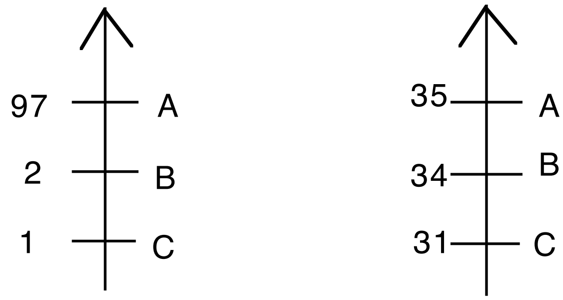
b. How many ordinal ballots are there when there are three candidates but indifference between two or more candidates is allowed?

c. Can you find a formula for the number of different preference schedules there are for $n>3$ candidates if a voter is indifferent between more than two of the candidates?

d. How many preference schedules are there when there are 4 alternatives and each voter has no indifference between any of the alternatives or a voter is indifferent between exactly two of the alternatives?

The use of a ballot of this type raises many different issues. First, are people capable of producing such rankings, especially in situations where there are many alternatives (say, as many as 10) to be considered? Second, must one rank all the candidates, even if one would rather not? (The reasons one might choose not to rank all the candidates might range from not knowing all of them and, hence, not being willing to rank alternatives one is not familiar with, to strategic considerations where one fears that by ranking more than just a few top choices, one somehow might be contributing to the selection of an alternative that one would rather not see go forward. Below, unless otherwise stated we will assume that a valid ordinal ballot requires that all the alternatives be ranked.

In addition to asking a person to rank the candidates, there are other possibilities. For example, one might give each voter a certain number of points, say 100 points, and allow the voter to distribute these points in any way among the alternatives to be chosen. Thus, two different people might distribute their 100 points for A, B, C, as shown:



Although these two ballots produce the same ordinal rankings, they indicate a seemingly dramatic difference in the way the two people view the three candidates. A ballot of this kind is called a *cardinal ballot*. It seems to allow voters, to some extent, the ability to express the intensity of preferences they have. It is clearly of interest to try to see if the method by which an election is decided can use the "information" in the difference between these two ballots.

Ordinal preference and cardinal preference ballots do not exhaust all of the possibilities that imaginative thinkers have come up with. For example, relatively recently there has been an explosion of new ideas about voting. Two such methods are called *approval voting* and *negative voting*. In approval voting the ballot consists of a list of all the candidates, and the voters indicate any of these individuals they are willing to see serve. (Note that the approval ballot does not require ranking of the candidates.) In the ballot for "yes-no" voting, the voters are allowed to vote "yes" or "no" but not both for each candidate. Another possibility would be to allow a voter to place "no" next to the name of some candidates and then rank the remaining candidates. Next we will consider how the use of the ordinal ballot broadens the choice of election decision methods.

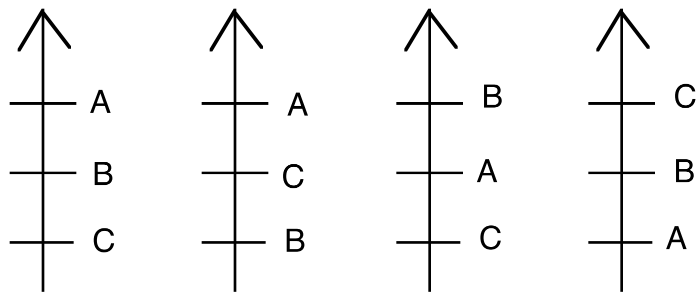
Election decision methods based on the ordinal ballot

Not surprisingly, having voters use an ordinal ballot makes it possible to use a variety of election decision methods that would not be possible to use if voters only completed a standard ballot. Note, however, that the plurality method can still be used.

1. Plurality voting

The number of first place votes for each candidate is counted and the winner of the election is the alternative with the largest number of (first place) votes. In essence, when plurality voting is employed, any benefit of using the more elaborate ordinal ballot rather than the standard ballot is obviated. However, it makes clear that what one can do with ordinal ballots certainly includes what one was able to do with the standard ballot.

Example:



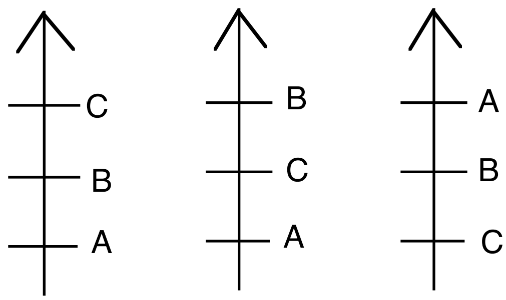
votes: 5 2 8 9

Since A gets 7, B gets 8 and C gets 9 first-place votes, the winner using plurality voting is C. Note, however, that C is the last-place choice of 13 of the 24 voters.

2. Run-off elections

Since in some elections no candidate will get a majority of the votes cast, if one uses ordinal ballots, one can, in effect, conduct a run-off election without the expense of asking voters to return to the polls. What is done is that all candidates except the two getting the largest number of first-place votes are eliminated and a rematch is made between the two candidates who got the largest number of first-place votes. Note that this can be done using the ordinal ballot by just seeing which candidate is above the other on the ballot, without regard to the alternative(s) eliminated.

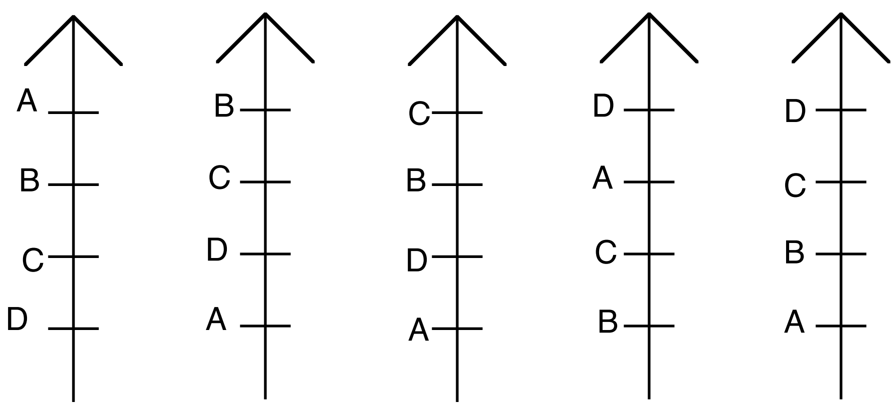
Example:



votes: 13 14 16

In this election no candidate gets a majority of the votes cast. Hence, a run-off is conducted. Since C got the smallest number of first-place votes, C is eliminated. In the next round of the procedure, those voters unable to vote for C would vote for the next highest person on the voter's preference schedule. Thus, the 13 voters in this example who ranked C first would vote for B in the run-off. The winner of the election would be B by a vote of 27 votes to 16. Note that candidate A would have been the winner had the plurality method been used.

Example:



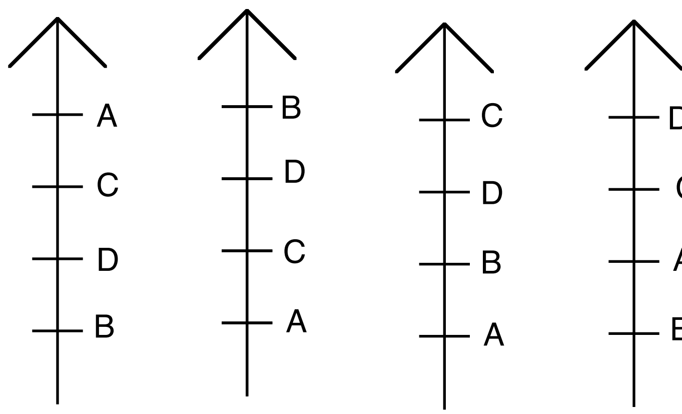
votes: 12 10 9 2 4

Initially, A gets 12 votes, B gets 10 votes, C gets 9 votes and D gets 6 votes. The two lowest vote getters, C and D, are eliminated and a run-off election between A and B is held. In the run-off A gets 14 votes while B gets 23 votes. Thus, B is the run-off winner.

3. Sequential run-off elections

In elections with many candidates the procedure which eliminates all the candidates other than the two with the highest number of first-place votes seems to run the risk of eliminating a "popular" choice too early in the voting. One way to meet this objection would be to eliminate one candidate at a time, until only two candidates remain. In this context the method which results will be called *sequential run-off*; however, a related idea (the *single transferable vote*) is used to achieve "proportional representation" when more than one candidate is to be selected from a field where many candidates are seeking office.

Example:



votes: 18 14 12 9

Although A has the largest number of first-place votes, no candidate gets a majority. Thus, under the sequential run-off method one first eliminates the candidate who got the lowest number of first-place votes. This candidate is D. In the next round, the 9 voters who voted for D initially would vote for C instead. In the next round B has the fewest first place votes and would be eliminated. In the final round, when A and C are the only remaining candidates, A gets 18 votes and C gets 35. Thus, C emerges as the winner. The summary of first place votes for each round is shown below:

Round 1: A = 18 B = 14 C = 12 D = 9

Round 2: A = 18 B = 14 C = 21

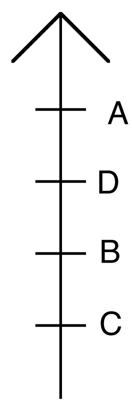
Round 3: A = 18 C = 35.

Note that in a run-off election, rather than a sequential run-off election, both

C and D would have been eliminated in the first round. In the final round between A and B, A would have beaten B by a vote of 27 to 26. Candidate A would have also been the winner if the plurality method had been used.

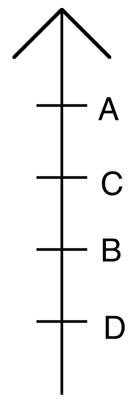
4. Borda Count

The idea behind the method of Jean-Charles de Borda was to give credit to the different alternatives according to how high up on the ballots of the individual voters the alternatives were listed. Thus, in the *Borda Count* highly ranked alternatives are assigned more points than lower ranked alternatives. In the standard version of the Borda Count, as it has come to be called, if there are n alternatives, the number of points assigned to an alternative depends on the number of alternatives that are below it. Thus, an alternative placed alone at the top of a voter's ranking would get (n-1) points. To illustrate the method the point counts for the ballots below are computed:



votes: 1

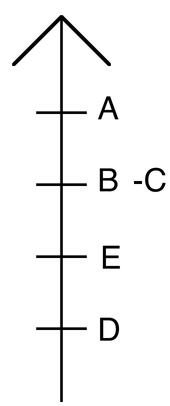
The number of points for A is 3, for D is 2, for B is 1, and for C is 0.



votes: 8

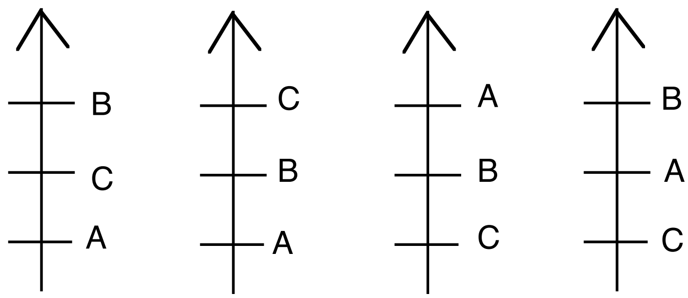
The number of points for A is 3 times 8 or 24, for C is 2 times 8 or 16, for B is 1 times 8 or 8, and for D is 0 times 8 or 0.

As illustrated below, the Borda Count can be applied even when voters are indifferent among some of the candidates.



A gets 4 points, B gets 2 points, C gets 2 points, E gets 1 point, and D gets 0 points.

Example:



votes: 12 16 19 5

We can compute the Borda Count for each of the candidates in the election above:

A: 12(0) + 16(0) + 19(2) + 5(1) = 43

B: 12(2) + 16(1) + 19(1) + 5(2) = 69

C: 12(1) + 16(2) + 19(0) + 5(0) = 44

Since candidate B has the largest number of points, B would be declared the winner.

What is intuitively appealing about the Borda Count is that it does not operate by taking into account only first-place votes for a candidate. Rather it chooses an alternative which is highest on the average (as measured by the mean) of the alternatives to be chosen. More precisely, if we compute the Borda Count of an alternative ß divided by the total number of voters we obtain the average (mean) number of alternatives below ß.

Additional insight into the Borda Count (as well as the method of Condorcet that is to be discussed below) can be obtained by constructing the *pair-preference matrix* associated with an election. A *matrix* is nothing more than a table which consists of rows and columns. The rows and columns of the pair-preference matrix are labeled with the alternatives being voted on. The entry in the row labeled X and the column labeled Y is the number of voters who prefer alternative X to Y in the election being considered. In the election above, for example, since A was preferred to B by 19 voters, the entry in row A and column B is 19. The pair-preference matrix for the election shown above is given below:

	A	B	C	Row sum

$$\begin{array}{c} A \\ B \\ C \end{array} \begin{bmatrix} \text{---} & 19 & 24 \\ 33 & \text{---} & 36 \\ 28 & 16 & \text{---} \end{bmatrix} \begin{array}{c} 43 \\ 69 \\ 44 \end{array}$$

Note that the sum of the (X,Y) and (Y,X) entries in the matrix add up to the total number of voters and the row sums in the matrix are exactly the values of the Borda Count.

<u>Stop and Explore</u>:

a. Compute the column sums for the pair-preference matrix above. Can you give some intuitive meaning to these numbers?

b. How might the column sums be used to construct an election decision method?

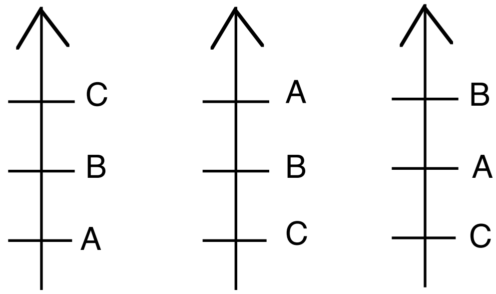
c. Given a 3-row and 3-column matrix (blank entries on the diagonal) with non-negative entries, when is it possible to construct a collection of preference schedules involving three alternatives such that the given matrix is the pair-preference matrix for the preference schedules constructed? When will the matrix represent a unique collection of preference schedules?

d. Can you prove that the row sums of the pair-preference matrix are the Borda Counts for the alternatives?

5. Condorcet's method

Condorcet's method requires that the winner of an election be a candidate who can beat all the other candidates in a two-way race.

Example:



votes: 12 16 13

B beats A by 25 votes to 16 and B beats C by 29 votes to 12. Hence, B is the Condorcet winner.

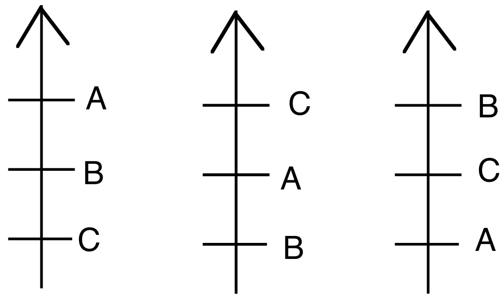
It is instructive here also to construct the pair-preferences matrix:

$$
\begin{array}{c c c}
 & \begin{array}{ccc} A & B & C \end{array} & \\
\begin{array}{c} A \\ B \\ C \end{array} & \begin{bmatrix} \text{---} & 16 & 29 \\ 25 & \text{---} & 29 \\ 12 & 12 & \text{---} \end{bmatrix} & \begin{array}{c} 45 \\ 54 \\ 24 \end{array}
\end{array}
$$

The fact that B is the Condorcet winner can been seen from the fact that the (B, A) entry (i.e. the row B, column A entry) exceeds the (A, B) entry and the (B,C) entry exceeds the (C, B) entry. In this example, the Borda Count winner is also B since the largest row-sum for the pair-preference matrix is attained in row B.

A remarkable property of Condorcet's approach, which Condorcet himself was aware of, is that the method (unlike the Borda method) does not always produce a winner in every election. Even though examples to show this are not hard to produce, it surprises most people the first time they are made aware of this phenomenon.

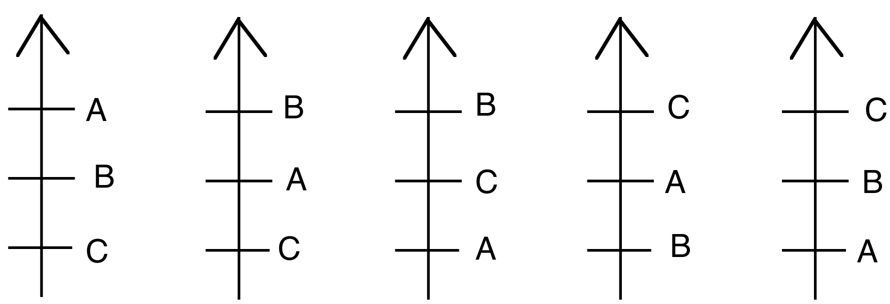
Here is a simple example which involves three alternatives:



votes: 1 1 1

Note that in a two-way race A beats B two votes to one, B beats C two votes to one, and C beats A two votes to one. This phenomenon is known as either Condorcet's paradox or the voting paradox. Note that the preference schedule headed by B can be obtained from the one headed by A by moving the C on the bottom of the one headed by A to the top. Similarly, the preference schedule headed by B can be obtained from the one headed by B.

Condorcet's original example is also of interest, because it shows that it is not necessary for there to be either the highly symmetric situation in the example above or solely the cyclic structure present in the preference schedules themselves. His example is given below:

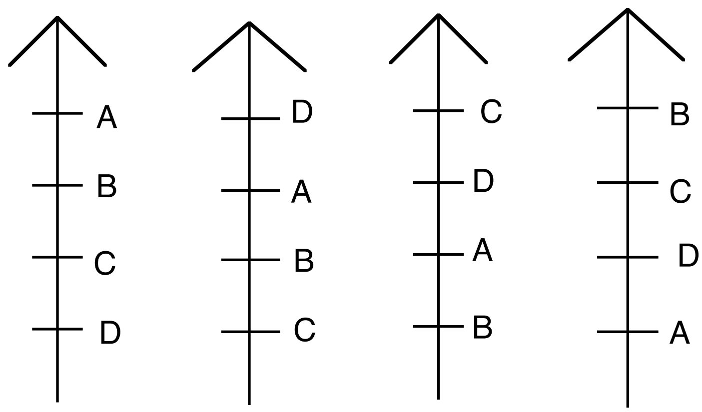


votes: 23 2 17 10 8

<u>Stop and Explore</u>:

a. If one discards any of the blocks of votes represented by the 5 preference schedules above in turn, how does this affect the results of applying the Condorcet Method?

It is easy to generalize the first example to any number of alternatives greater than 3. The situation for four alternatives is shown below:



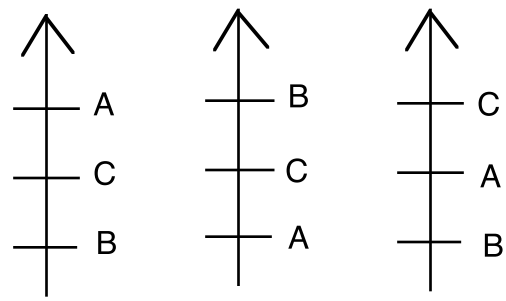
votes: 1 1 1 1

The reason why this failure of the Condorcet method always to give a winner seems surprising is that in many situations we are used to reasoning that if A is "better" than B and B is "better" than C, then A is "better" than C. Relationships which obey this property are called *transitive* relationships, and most of the relationships we commonly deal with in mathematics obey this rule. Thus, equality of numbers is a transitive relationship and parallelism of lines is a transitive relationship. However, as the example here shows, not all relationships are transitive. The existence of examples such as the one above means that an arbitrary collection of voter preference schedules does not guarantee a winner for the election. Many scholars have offered suggestions for what to do if there is no Condorcet winner to modify this approach to produce a winner. For example, one can construct a hybrid method which chooses the Condorcet winner if there is one, but if there is none, the Borda Count is used. Another approach is to count the number of two-way races that each candidate can win and declare the winner of the election to be that candidate who wins the largest number of two-way races. (A more detailed discussion of ways to deal with the fact that there are elections for which the Condorcet method gives no winner appears later.)

6. Medial voting

Medial voting gives a point to each candidate who is above the median level for each voter. If there is an odd number (2n+1) of alternatives and no indifference about alternatives on the part of the voters, this means a point is given to each alternative ranked (n+1) or above. For an even number (2n) of alternatives, again with no voter indifference, this means giving a point to each alternative ranked n or above. The winner is the alternative with the highest number of points.

Example:



votes: 17 14 12

Since there are three alternatives (an odd number), a point is given to the top two candidates on each preference schedule. This results in the following number of points for each candidate:

A: 17 + 0 + 12 = 39

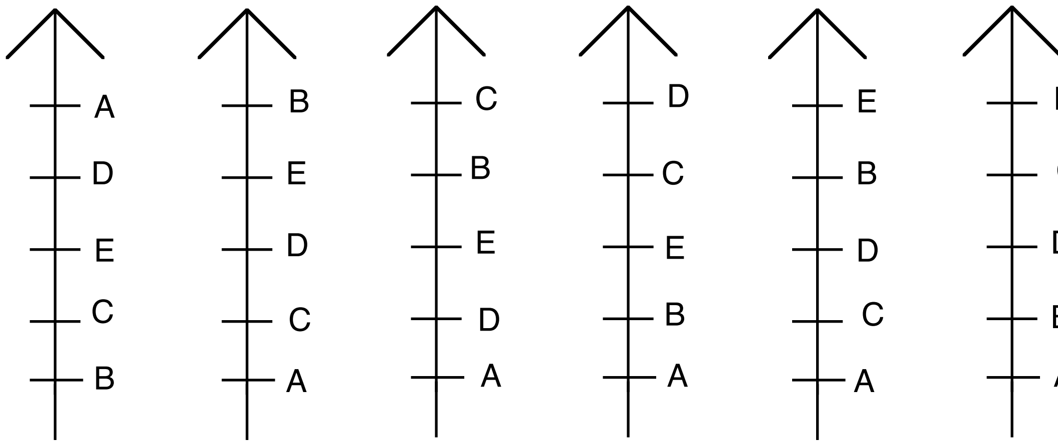
B: 0 + 14 + 0 = 14

C: 17 + 14 + 12 = 43

C would be the winner. Note that C would also be the Condorcet and Borda Count winner for this election.

To summarize the discussion of these methods, in table form below are the results of applying the first 5 of these methods on the example below.

Example:



Votes: 18 12 10 9 4 2

Table: (Winners in the rounds and over-all winners are denoted with bold type.)

Method 1 (Plurality Voting)

A = 18 B = 12 C = 10 D = 9 E = 6

Method 2 (Run-off)

Round 1 **A** =18 B = 12 C = 10 D = 9 E = 6

Round 2 A = 18 **B** = 37

Method 3 (Sequential Run-off)

Round 1 **A** = 18 B = 12 C = 10 D = 9 E = 6

Round 2 **A** = 18 B = 16 C = 12 D = 9

Round 3 A = 18 B = 16 **C** = 21

Round 4 A = 18 **C** = 37

Method 4 (Borda Count)

A = 4(18) + 0(12) + 0(10) + 0(9) + 0(3) + 0(2) = 72

B = 0(18) + 4(12) + 3(10) + 1(9) + 3(4) + 1(2) = 101

C = 1(18) + 1(12) + 4(10) + 3(9) + 1(4) + 3(2) = 107

D = 3(18) + 2(12) + 1(10) + 4(9) + 2(4) + 2(2) = 136

E = 2(18) + 3(12) + 2(10) + 2(9) + 4(4) + 4(2) = 134

Method 5 (Condorcet)

A versus E	A = 18	**E** = 37
B versus E	A = 22	**E** = 33
C versus E	A = 19	**E** = 36
D versus E	A = 18	**E** = 37

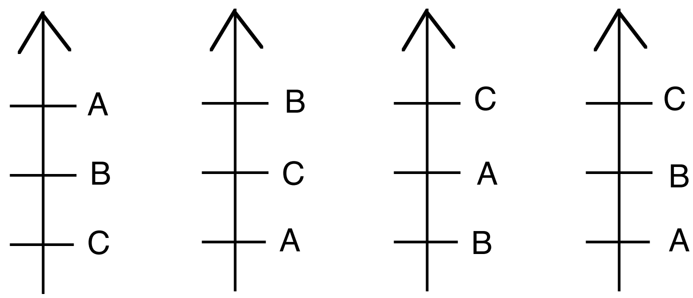
The remarkable reality is that the winner of the election differs for each of five different methods! (The sixth method we discussed, medial voting, would result in E (the Condorcet winner) as the winner. It is an enjoyable challenge to find an election involving 6 candidates for which all six methods yield different winners.) This example shows the independence of these methods from each other in general circumstances (i.e. one can not prove theorems of the kind: if alternative A is a run-off winner, then alternative A is the Borda count winner). It also calls into question the complacency most Americans feel with the election processes that we participate in. We feel that having gone freely to the polls and voted, the will of the people has been achieved. In fact, this example shows that the outcome of a democratic choice procedure depends on the election procedure used. Which of the election methods produces an outcome which comes closest to the "intent" or "desire" of the voters is not at all clear!

<u>Stop and Explore</u>:

a. For methods which are sequential, one can, as has been done up to now, eliminate at each stage on the basis of smallest first-place vote, or on the basis of who would lose in a two-way race between the two lowest first place vote getters at the given stage. Examine whether or not these alternative approaches can lead to different winners when sequential run-off is used or a method based on sequential elimination via the Borda Count (see Nanson's method below) is used.

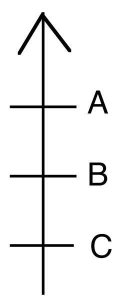
Although in our discussion above we have emphasized the context of selecting one winning alternative for society, each of the methods we have discussed can be extended to the context of selecting a ranking of the alternatives for society.

Example:

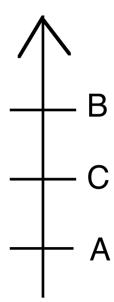


votes: 20 18 5 12

The plurality winner is A but the plurality ranking would be:

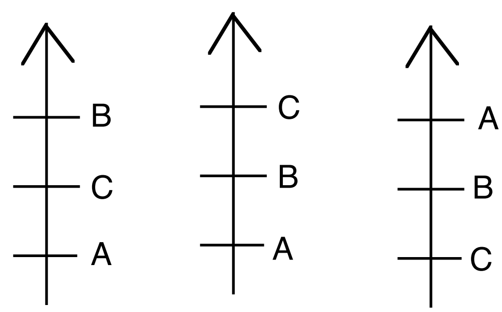


The Borda Count winner would be B (68 points), while the Borda ranking would be:



In the methods we have discussed so far we have paid considerable attention to the first-place votes that a candidate received. However, there are other approaches. For example, one can design a candidate elimination scheme which depends not on considering the smallest number of first-place votes but considering the largest number of last-place votes. (For convenience assume that there are an odd number of voters.) Thus, one can eliminate the alternatives one at a time based on the largest number of last-place votes, checking at each stage that no candidate currently has a majority. (When a candidate at some stage has a majority, that candidate is declared to be the winner.) Eventually one gets to a stage where there are two candidates left and one will have a majority. (In the case of a tie for number of last-place votes, some tie breaking scheme must be adopted.)

Example:



votes: 12 16 17

Notice that no candidate got a majority of first-place votes. In this example, first A is eliminated because A has 28 last place votes and C has 17 last-place votes. After A is eliminated, B gets a majority of the votes against C and would be declared the winner. Note that for this election B happens to be the Condorcet winner. However, it is possible to produce examples where this method does not select the Condorcet winner. The method we have just described is known as *Coombs' Method*. It is named for Clyde Coombs, an American psychologist who argued that it was a useful procedure for deciding winners in highly confrontational situations.

Stop and Explore

a. Try to construct an example where Coombs' Method does not select the Condorcet winner although there is one.

b. Which candidate is the Coombs' winner for the election on page 18?

c. Modify Coombs' Method so that between the two candidates with the largest number of last place votes, the one eliminated is the one of the two who loses in a two-way race between them. Can this result in a different winner from the usual Coombs' winner?

d. Show by an example that it is necessary at each stage of Coombs' Method to check if some candidate has a majority, since otherwise a candidate other than the Coombs' winner will win the election.

These ideas clearly raise the issue of which of these methods is the fairest and most equitable. There are many potential fairness principles or rules (axioms) that one might like an election decision method to obey. For example, there is the *Majority Principle*:

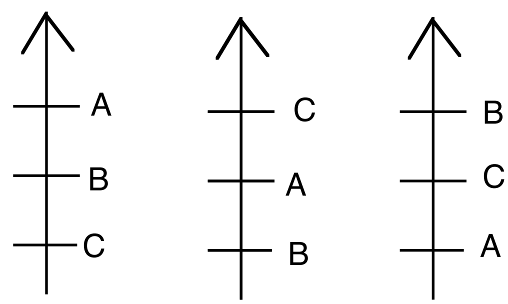
(Majority Principle:) If there is an alternative which is preferred by a majority of the voters, the election decision method should select this alternative.

It is possible to study which election decision methods obey the fairness and equity principles (axioms) which we would like to see obeyed. In fact, this approach will be considered below. However, first we wish to come back to the issue raised above concerning the consequences of voters' not voting their true feelings or preferences about the alternatives in the hope of achieving a strategic advantage.

Sincere and sophisticated voting

When the democratic election process is considered, it would be nice to believe that the wisest thing that each voter can do is vote for the alternatives in accordance with the voter's preferences for the alternatives. When a voter behaves in this manner, he or she is using *sincere voting*. However, whether or not the voter has "solid" information about how other voters might vote, a voter may decide by some reasoning process to vote in some manner other than what he/she truly feels will result in a better outcome for him/her. Thinking in this manner is called *strategic voting* or *sophisticated voting*. Though surely this kind of thinking has been used by voters over and over again, the modern formal study of this notion began in 1969 with work of Robin Fahrquarson. The example below is essentially due to him, though no doubt similar ideas had independently occurred earlier.

Imagine that there are three voters in a committee. One committee member, u, serves as a chairperson. Not only does u cast a vote in the way the other two members of the committee do, but she breaks ties in the event that no alternative has a majority. Suppose that the other two voters are names v and w and that the three alternatives are A, B, and C. Below are shown the preferences schedules of the three voters.



voter: u v w

We will now assume that each voter is aware of the preference schedule of each of the other voters and that each is capable of the reasoning that another could make of having such information. (For example, u knows v's preference schedule and knows that v knows this, etc.) However, the players are not able to enter into coalitions and agreements with each other but must act on his/her own.

What should the chair, u, do? Voter u can do no better than vote for alternative A, her first choice. The reason is that if v and w vote for different alternatives, then A gets to break the tie. (Remember that v and w are acting strategically and may not vote for their first choices if they think that this is their best response to a strategic choice on the part of u!) In this case u's choice will be the one that wins, and since A is her first choice, she should vote for A. On the other hand, if v and w vote for the same alternative, then u will be outvoted and there is no loss in voting for alternative A. Hence, u will vote for alternative A, since this is the best she can do regardless of what the other voters do.

Clearly, it can never pay for v or w to vote for his/her last choice alternative. Now, how will v vote in light of the fact that v can be sure that a wise u will vote for A? In analyzing what to do, v will have to consider what w might do, knowing that w will not vote for A since this is w's last choice. If v votes for his/her second choice A, then if w votes for B, v will get his/her second choice A, while if w votes for C, v will still get his/her second choice A. This follows from the fact that since u votes for A, when v does too, there is a majority for A regardless of what w does. On the other hand, if v votes for C, then if w votes for B there will be a tie and u, breaking the tie, will result in A's being chosen. However, if v votes for C and w votes for C, then v gets his/her first choice. Thus, whatever w does, v can do no worse than getting A and might get C by voting for C. Hence, it makes sense for v to vote for C.

If we look at w, w knows that u will vote for A and that v will vote for C. Thus, if w votes for B, the outcome will be A, since if v votes for C, u will break the 3-way tie by using her chairperson's tie-breaking privilege. W would get his/her worst outcome. If w votes for C, then since v did also, the result will be C. Thus, it makes sense for w to vote for C. Hence, we can conclude that u votes for A, v votes for C and w votes for C, with the result that C wins.

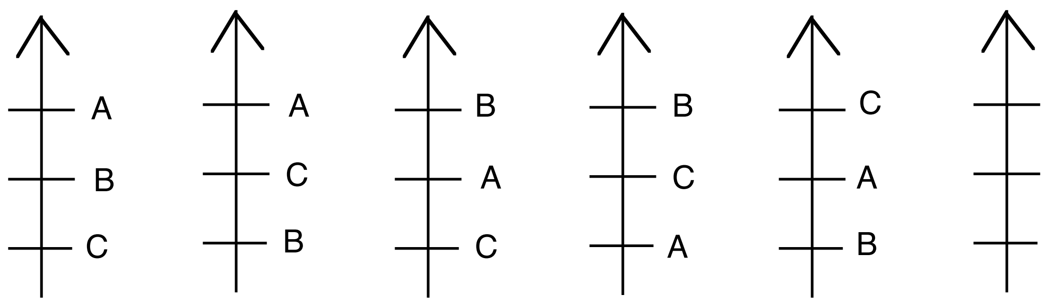
The net result of this contorted reasoning is that the chairperson, who a priori seemed to have an advantage due to being able to break ties, winds up with her worst choice (rather than A or B) because the other voters vote in a sophisticated manner!

This convoluted analysis certainly suggests the desirability of using an election method that would avoid voters' trying to take advantage of sophisticated voting. What makes sophisticated voting possible is the information about the preference schedules of the other voters. This kind of information is made possible by the elaborate polls that are being conducted prior to elections in many democracies. If voters give honest preference schedules to pollsters and these become available to all other voters, then sophisticated voting becomes a possibility. Thus, we see that polls have a complicating effect on the operation of a democratic voting system because they can encourage individuals not to vote in a way which may represent their true feelings. In US presidential elections, exit polls (i.e. surveys of voters after they vote, which ask how the voter voted in the voting booth) on the East Coast can be used to make computer predictions for the winner of the presidency. Since voters on the West Coast can continue to vote after voting ends on the East Coast due to the time change (3 hours), exit polls can have an effect on the election by perhaps discouraging West Coast voters going out to vote. This results in more influence on the outcome of an election by East Coast than West Coast voters, which is not fair. On the other hand, restricting the rights of the press or television networks to investigate voter sentiments also has repercussions. Ideally, it would be nice if a method of conducting elections could be found which avoided such problems. Whether or not this is possible will be examined later. First we will look at an important issue which raises complications in what we have discussed above.

Dealing with ties

Up until now, to keep the discussion as simple as possible we have purposely avoid the troubling issue of what to do when election decision methods result in ties. (What we are concerned with here is not that a particular voter may be indifferent, that is, have ties in preference between alternatives that are being voted for, but ties that occur between or among alternatives when an election decision method is carried out for a particular set of ballots.) Ties can arise due to a variety of reasons. Sometimes there are very strong symmetries in the pattern with which voters vote which results in a tie as due to the symmetry involved. In other cases ties seem to arise as a numerical coincidence. Some examples will clarify the issue.

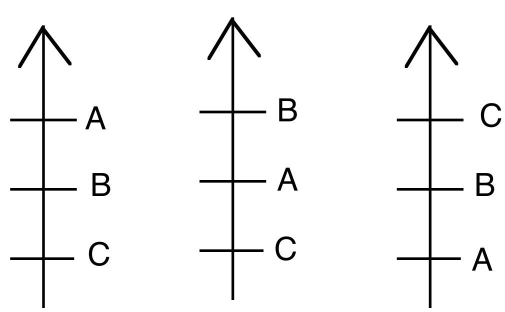
Example:



votes: 2 2 2 2 2 2

The plurality method results in: A: 4 votes, B: 4 votes and C: 4 votes, while the Borda Count gives each candidate 12 points. Since equal numbers of voters select each of the total of 6 preference schedules possible due to the symmetry of the situation, it is not surprising that a three-way tie results for each of these methods.

Example:



votes: 10 8 2

The Borda Count for A is 2(10) + 1(8) + 0(2) = 28, for B is 1(10) + 2(8) + 1(2) = 28, and for C is 0(10) + 0(8) + 2(2) = 4. Thus, there is a tie between A and B. Who should win? This tie seems to have nothing to due with symmetry.

There are a variety of approaches to dealing with ties. One approach is to single out some voter to break ties. One might have some voter designated as a chairperson, and if there is a tie, then this person gets to break it. Another approach is to say that when there is a tie, instead of trying to pick a single winner for society, one should pick a collection of winners for society. In some contexts this approach makes a lot of sense. If there is a monetary prize being offered and the voting for a prize winner results in a tie, one might choose to divide the prize equally between the tied winners. In other contexts, such as electing a candidate for public office, where it is not feasible for "sharing" to occur, we need another way out.

One way to proceed with tie breaking is to agree in advance on an ordering of the alternatives and break ties when they occur by using this pre-decided ordering. For example, if one is electing candidates for public office, one could agree to break ties by using the alphabetical order of the candidates' names (and if these are the same, use first names). Alternatively, one could break ties by using "reverse" alphabetical order. Another approach to breaking ties is to use a fair randomization device. Thus, for a tie between two candidates a fair coin could be tossed.

The reason ties cause problems for discussion of the fairness of methods is that one wants election decision methods to obey a fairness rule which says that the decision should not take into account the name of a voter in arriving at a winner for society. Thus, we would certainly want our election decision method to obey the *Anonymity Principle*:

The names of the voters should play no role in an election decision method.

Comment:

The Anonymity Principle means that if two voters exchange the ballots that they cast, this should not affect the winner of an election.

Yet this is exactly the kind of thing that a tie-breaking rule which involves a chairperson does. Some voters are more equal than others!

Another fairness rule that makes sense is that an election decision method should not depend on the names of the alternatives. We will refer to this fairness idea as the *Neutrality Principle*:

An election decision procedure should not make use of the names of the alternatives.

Comment:

If we exchange two alternatives' names and implement this change on all of the preference schedules that the voters produce, then the outcome of the election should change accordingly. (Thus, if we exchange alternative A_i and A_j and before the change A_i won the election then after the exchange A_j should win the election, while if neither won before the exchange, then neither should win after the exchange.)

Note that using some scheme based on alphabetical order to break ties would violate the neutrality principle.

Suffice it to say that despite the technicalities that one is lead to in dealing with breaking ties, this problem can be dealt with in a sensible manner. Perhaps equally important, in situations where large numbers of voters are involved, the chances of such ties' occurring is very small and can, in fact, be disregarded as a practical matter. (Students of elections have probably noted that in "close" elections, when recounts are called for, there is invariably a change in the original versus the final figures. This suggests other considerations than those of mathematics come into play here.)

In passing it turns out the principles of anonymity and neutrality together with the *principle of monotonicity* (more support for a candidate can only help the candidate's cause) are what are necessary as the assumptions for an interesting theorem of Kenneth May:

Theorem (May, 1952)

There is only one election decision method which involves two candidates (e.g. X and Y) and which obeys the Anonymity, Neutrality, and Monotonicity Principles. This method is majority rule, that is, alternative X is chosen rather than Y if and only if X receives at least as many votes at Y.

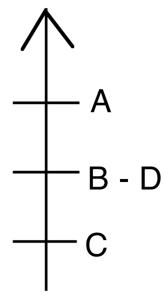
May's Theorem stands in dramatic contrast to what the situation is when there are three or more candidates to choose from.

Next we will return to the problem of determining those election methods which are fair when there are three or more alternatives to chose from.

Arrow's Theorem

In pioneering work in the 1950's Kenneth Arrow, the mathematical economist, turned the question we have been raising implicitly above on its head. Up to now we have tried to develop methods (i.e. sequential run-off, Borda Count, Condorcet, etc.) which seemed defensible as fair or equitable methods from some point of view or other. What Arrow did was to ask what fairness conditions we would like a good method to satisfy. What will be referred to below as an election decision method is usually called a *function* by mathematicians. (The intuitive idea behind a function is that of an association scheme which for any "legal" input assigns a unique (legal) output. In our context the legal inputs are the preference schedules (indifference not permitted) of the voters and the output is a preference schedule with ties allowed. Thus, we are dealing with how to chose for society on the basis of what the individual voters feel.) As an economist, Arrow was interested, for example, in how a group of socialist planners might reach a target decision for society on the basis of the group's individual views, in addition to deciding the winner of an election. Before actually stating Arrow's extraordinary result, we need a bit of background.

We will first assume that there are three or more alternatives that are acted upon by the voters, and that each voter has produced a ballot which ranks all the alternatives that are to be considered. (Arrow's Theorem has been extended in a variety of ways; the presentation here is chosen not for giving the most general result known but for the ease of presentation.)
Next we will assume that if there are $n > 2$ alternatives to choose from, then each voter ranks all of these alternatives with no ties (or indifference) between two or more candidates. Finally, we will assume that rather than selecting a single winner for the voting situation from among the alternatives on behalf of the "society" formed by the individuals who are voting, we will choose a ranking for society. Shown below is an example of a ranking such as might be produced for society if there were 4 alternatives to chose from:



This would be interpreted to mean that given how the individuals who voted felt, that society has ranked alternative A highest, is indifferent between alternatives B and D, but ranks these two alternatives above C.

Now we will examine the fairness rules of the kind that Arrow examined. The presentation below reflects improvements that have been made in presenting Arrow's ideas since they were first introduced.

Condition 1: (Decisiveness)

The election decision method used must be prepared to take into account any possible pattern of preference schedules that the voters produce, and use based on these preference schedules decides a ranking for society.

Comment:

This condition means that after the votes are cast, the decision procedure can not look at some ballots and say, for example, that they are "irrational" and choose not to count some of the ballots. The decision method should not depend on looking at the particular way the voters voted as part of the decision procedure. The procedure is fixed in advance and will apply regardless of the particular voter inputs. In particular, the Condorcet method does not satisfy this condition since there are elections for which there is no Condorcet winner, and thus the method does produce a ranking for society in such a case.

Condition 2: (The *Pareto Condition*)

If all the voters vote for the same preference schedule, then this should be the ranking for society.

Comment:

The election decision procedure should take the unanimous voice of the voters into account and not base its decision on other input or information. This condition rules out using a decision method based on an "oracle" external to the system. There are many variants of this condition, one of which we will consider later.

Condition 3: (*Independence of Irrelevant Alternatives* (often written, IIA))

Suppose that for the given set of alternatives and voters we have two different election preference patterns, election L and election M. Now assume that exactly the same voters rank alternative I over alternative J in election L and election M. When these conditions hold, the election decision method should produce a social choice ranking R_L for election L and a social choice ranking R_M for election M where:

a. I is above J in both R_L and R_M

or

b. J is above I in both R_L and R_M

or

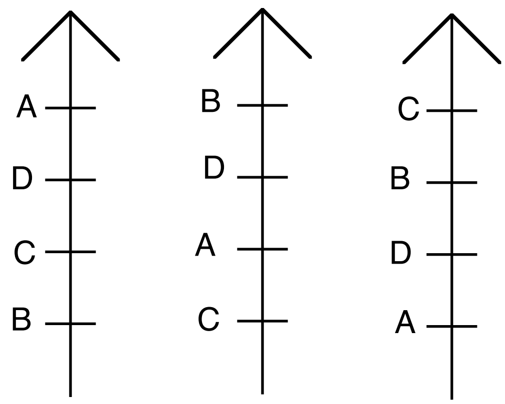
c. I and J are tied in both R_L and R_M.

Comment:

IIA (*Independence of Irrelevant Alternatives*) is a consistency condition on an election decision method which restricts what can be the outcomes when two related elections L and M are presented for adjudication. The condition states that society's treatment of any two alternatives I and J should depend only on the way I and J are viewed by the voters and not also on the way the voters rank other alternatives. In particular, this condition rules out the Borda Count where the number of points I and M get depends on the other alternatives involved. Here is an example which illustrates this:

Example:

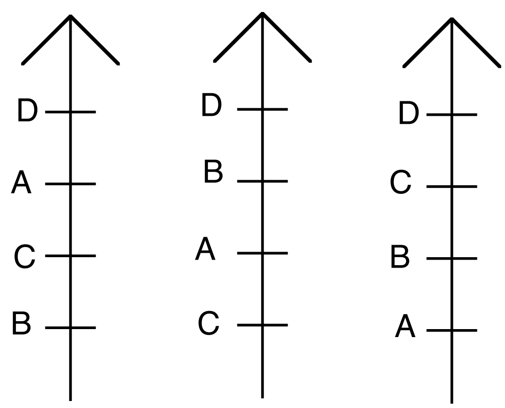
Election L:



votes: 14 10 9

The Borda Count for B is 48, while the Borda Count for C is 41. Thus, in election L the Borda Count ranks B above C.

Next consider the election M:



votes: 14 10 9

Note that the relative positions of B and C have not been changed, only the position of D has been altered relative to the other candidates. However, now the Borda Count for B is 29 and the Borda Count for C is 32. Thus, in election M the Borda Count ranks C over B. The elections L and M illustrate that the Borda Count can violate the condition of Independence of Irrelevant Alternatives.

Note that Condition 3 applies to any pairs of alternatives I and J, not just to two particular alternatives.

Condition 4: (*Monotonicity*)

Suppose that for election L the election decision method ranks alternative I over alternative J. Now suppose election M is identical to election L except that some voters who listed J over I on the voters' preference schedule now lists I over J. When these conditions hold, the election decision method for M should still list alternative I over alternative J.

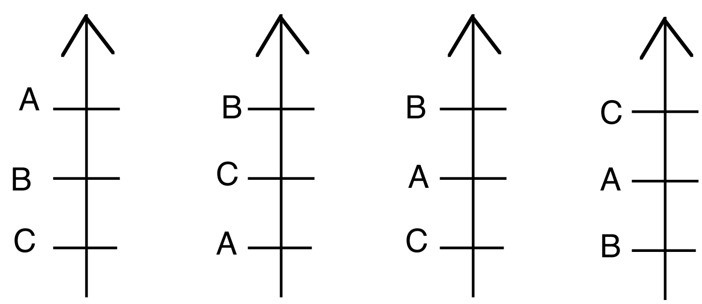
Comments:

As with condition 3, this condition requires internal consistency for a fair method when two related elections are presented to it. If a voter changes his/her mind and ranks an alternative I more highly than J in an election which already resulted in society's choosing I over J in the election before the change of mind occurred, then the method should still choose I over J after the change.

Example: (H. Moulin)

This example shows that the run-off method does not obey the Monotonicity Principle.

Consider first the election L:

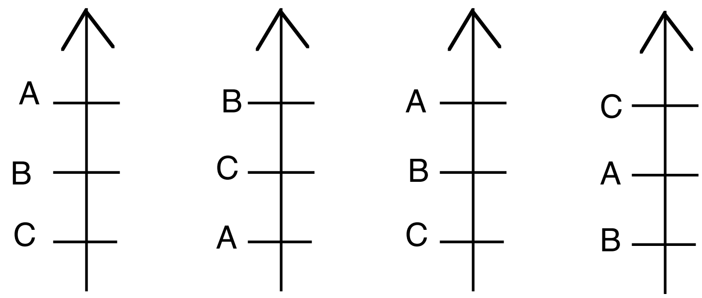


votes: 6 4 2 5

Since C gets the fewest number of first-place votes, alternative C is eliminated and in the run-off phase of the decision process, A beats B by 11 votes to C. Hence, A is the winner.

Now consider election M, which differs from L in that the preference schedule which was voted for by two voters has been altered so that now alternative A is preferred to alternative B. Note C is still rated last by these two voters.

Election M:



votes: 6 4 2 5

In this election, B gets the fewest first-place votes and is, therefore, eliminated. In the run-off phase of the election between A and C, C beats A by 9 to 8. Hence, the additional support that A received costs A the election!

Condition 5: (No *dictator*)

No voter should be a dictator.

Comments:

A voter w is a *dictator* if for all elections, the election decision method chooses for society the ranking which w voted for.

The remarkable result that Arrow demonstrated is the following:

Theorem (Kenneth Arrow):

There does not exist any election decision procedure where the voters rank three or more alternatives which obeys all of the conditions 1-5! (In fact, one can show that any election decision method which satisfies the first four conditions must be a dictatorship.)

This extraordinary result shows that no matter what election method might be used, some fairness rules do not hold for the method. This does not mean that we should be content with existing methods, merely that we can not have it all. This result has had dramatic effects on the philosophical community (by showing democracy's inherent limitations), political science community, and economics community. For this extraordinary piece of scholarship, as well as other major contributions to the field of Economics, Kenneth Arrow was awarded the Nobel Memorial Prize in Economics in 1972.

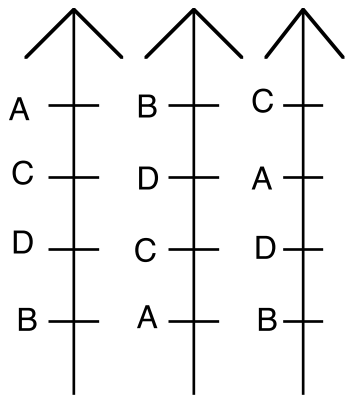
Stop and Explore:

a. Show that the Coombs' Method can violate the Condition 4 (Monotonicity) in Arrow's Theorem.

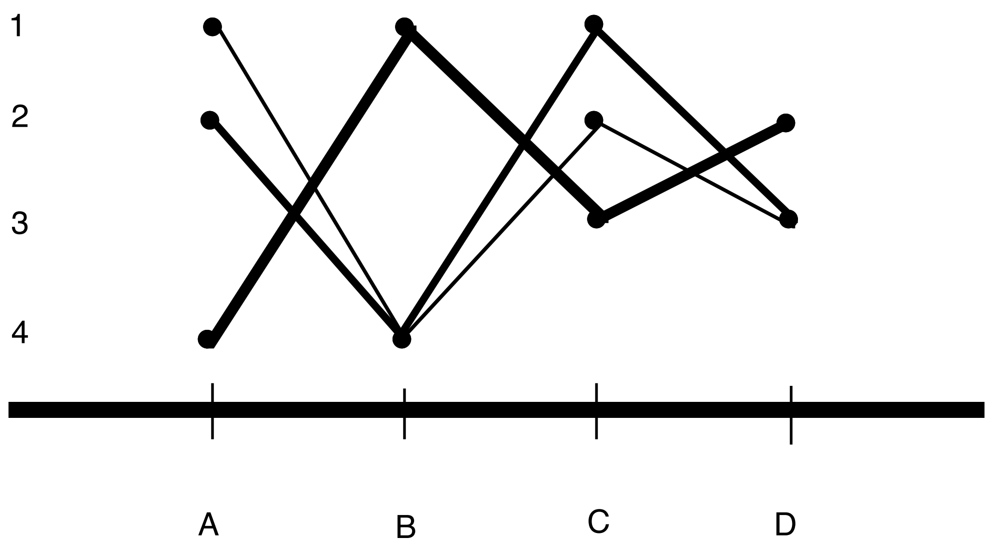
b. Which condition(s) of Arrow's theorem are violated by the plurality method?

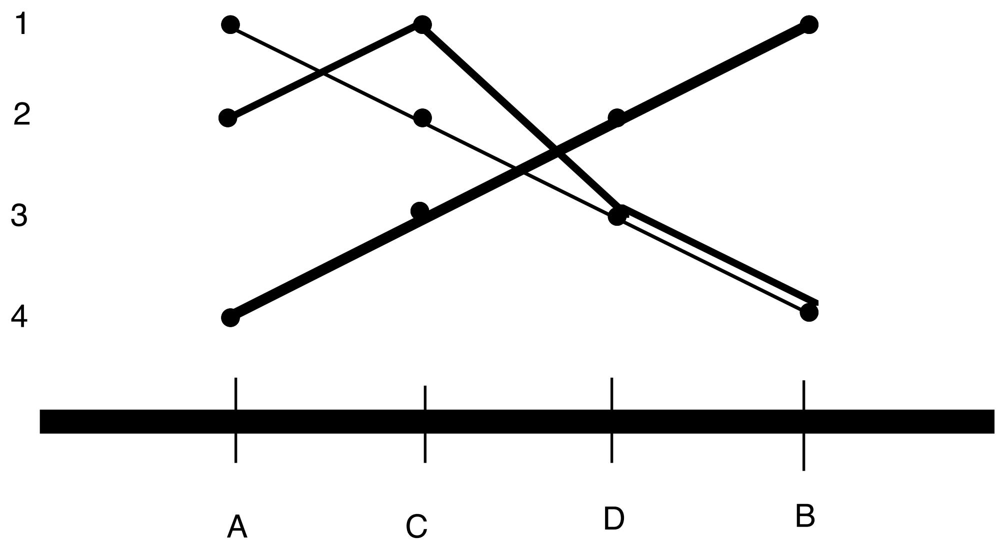
Many attempts have been made to find some way that the negative implications of Arrow's theorem can be avoided. One interesting approach has been to assume that the citizens who vote, though they may have many different opinions, have certain common views which cut down on the number of different potential elections that will have to be decided by any election decision method. For example, suppose that all the alternatives have been placed on a scale from "left" to "right." With respect to this left-right scale of the alternatives one can plot or draw a graph of any particular preference schedule. This is illustrated for two different scales and three different preference schedules. Corresponding preference schedules in the two graphs are shown by lines of the same thickness. In the first graph each of the preference schedules has a graph which does not strictly rise as one moves to the right or which rises and then falls as one moves to the right. A graph with these properties is known as single-peaked. Intuitively, when this occurs, we can think of the voter's having some optimal way of assessing the alternatives on the scale.

Example:



votes: 1 1 1





For the second scale each of the preference schedules is single-peaked when drawn. When this is true, Duncan Black proved a remarkable theorem:

Theorem (Duncan Black)

If all the alternatives to be voted on can be ordered from left to right so that respect to this scale all the voter preference schedules are single peaked then there is an alternative which can beat all the others in a two-way race. In other words, the Condorcet method always yields a winner in this case.

You can check for yourself that there is a Condorcet winner for the election above, namely C, as guaranteed by Black's Theorem.

It is natural to ask how likely it is that a society will have sufficient homogeneity of opinion that its citizens will produce preference schedules which are single-peaked with respect to some left-right scale. The number of preference schedules that voters might produce is n! when there are n alternatives. It is not difficult to see that the maximal number of preference schedules that can be single-peaked on any scale when there are n alternatives is 2^{n-1}. For large numbers of alternatives n! is much larger than 2^{n-1}. Thus, it seems unlikely that a society would produce such a set of preference schedules. Even when there are relatively few alternatives, the conditions that guarantee that a collection of preference schedules is single peaked seem unlikely to be met. (For example, for 4 alternatives the voters would have to vote for a maximum of 8 of the 24 possible preference schedules, and these 8 schedules have a very stringent structure.) As interesting as the line of research opened up by Duncan Black's work is, it is unlikely that it would ever be possible to be assured that Condorcet's method would always yield a winner because the members of the society who were voting had single-peaked preference schedules on some left-right scale.

Manipulation of voting methods

Ideally, one would like to have a voting system for which the voters voted their true preferences, rather than voting "insincerely" to obtain a strategic objective. We discussed this phenomenon above in our look at sincere and sophisticated voting. Lest you believe that this is not an issue, remember how common this phenomenon is with plurality voting when there are three or more candidates. How often have you heard people claim that you should not vote for a third party candidate, since to do so is to throw one's vote away? Anyone who heeds this advice is voting insincerely to help the chance of their second choice's being elected to avoid having their third favorite choice elected. A natural question to ask is whether or not one way to improve on plurality voting would be to change to a system which would not make it possible for people to improve their outcomes in the election by voting insincerely. Note that it might be possible to improve on plurality voting in this regard, even though any other method, like plurality voting, is still subject to the implications of Arrow's Theorem. Inspired by Arrow's remarkable result, other scholars addressed this problem: To what extent is it possible to design an election decision method which prevents sophisticated voting from occurring?

Independently and at approximately the same time this question was answered by A. Gibbard (1973) and M. Satterthwaite (1975). Consider the following fairness condition:

Condition 1: (*Citizen sovereignty*)

An election method obeys *citizen sovereignty* if, given any ranking of the alternatives E, there is some collection of voter preferences, which, when the election method is applied to these voter preferences, results in the ranking E being chosen for society.

Comments:

Condition 1 states that by suitably arranging the way the voters cast their votes, the election method has the flexibility of resulting in any ordering of the alternatives. Not to allow this condition would mean that there would be some ordering of the alternatives that could not be achieved no matter how the voters voted.

Theorem: (Gibbard and Satterthwaite)

Given a collection of alternatives (not necessarily a finite one) but having at least 3 elements which is ranked by a (finite) number of voters. Suppose that the election system obeys Condition 1 (Citizen Sovereignty); then the election system is not affected by sophisticated voting if and only if the system is dictatorial.

In other words, what this theorem says is that only the election system we call dictatorship will fail to be manipulated when voters vote in a sophisticated (insincere) manner! Hence, all democratic systems are subject to manipulation by insincere voting.

From a mathematical point of view it is interesting to note that Arrow's Theorem and the Gibbard/Satterthwaite Theorem are "logically equivalent." This means that each of these theorems can be used to prove the other.

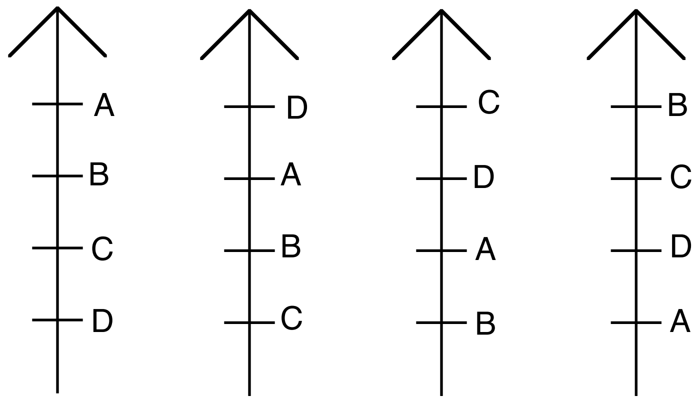
Even though one may be depressed by the implications of the Arrow and Gibbard/Satterthwaite Theorems, this does not mean that all voting methods are equally poor. The theorems do, however, mean that no voting method can have all the features that someone committed to democracy might desire.

More on Condorcet's Method

Although Condorcet's method does not always choose a winner, many people feel that when it does produce a winner, there is a compelling reason to use it. Supporters of the Condorcet method wonder what more can one ask for than that the candidate who wins an election would be the majority winner in a two-way race against any other candidate? Isn't that what democracy is about, majority rule? Other people feel that electing a Condorcet winner who polarizes the electorate may be worse than electing some other candidate who is the favorite of few but who is popular with lots of voters. (This issue will be addressed in the next section below in detail.)

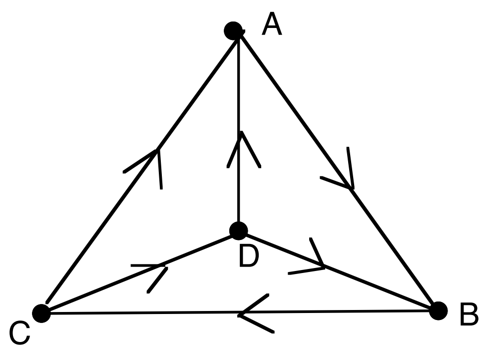
Is there some way that Condorcet's method can be extended in a natural way when there is no candidate who is the winner by Condorcet's method? To help see how this might be done, first we will consider what are the consequences when voters have preference schedules which do not produce a Condorcet winner.

Consider, for example, what happens when there are voters who rank 4 alternatives as in the election situation below:



votes: 10 13 14 16

The results of conducting a two-way race between every pair of candidates are summarized in the diagram below (called a *directed graph* or *digraph,* for short). An arrow from one candidate to another indicates that the candidate to which the arrow is pointed is beaten by the other candidate in a two-way race. In what follows, we will refer to a digraph of this kind as the *two-way race digraph*. In graph theory, a digraph which arises by having exactly one directed edge between every pair of vertices joining n points is called a *tournament digraph* since such a digraph provides a convenient way of constructing a way of representing the outcomes of a round robin tournament, one in which each pair of players plays exactly one match. Questions about how to rank the players (or teams) based on the results of a chess, tennis, or ping-pong tournament are related to the problem faced in deciding how to rank alternatives that have been voted upon.



For the two-way race digraph shown above it is easily seen that there is no

Condorcet winner, since no alternative beats the others in a two-way race. However, something perhaps even stranger occurs here. One can find a sequence of alternatives, namely A, B, C, D, A where A beats B in a two-way race, B beats C in a two-way race, C beats D in a two-way race, and D beats A in a two-way race!! This phenomenon is sometimes known as the *voters' paradox*, the *sportswriters' paradox*, or *Condorcet's paradox*. (The reason for the phrase sportswriters' paradox is that nearly every sports season, some sportswriter finds an example of a sport where the lowest ranked team in the league won a majority of games against some team, which in turn won a majority of games against some other team, ..., which in turn won a majority of games against the league champion!) The name which I will give to this phenomenon, when there is a "directed cycle" which includes all the alternatives, is that the voter preferences result in *cyclical majorities*. (It can be shown that if there is no Condorcet winner, then although there may not be a cycle through all of the alternatives, there must be some cycle of shorter length.)

<u>Stop and Explore</u>

a. Construct an example with 4 alternatives which shows that there may be no Condorcet winner, yet no cyclic majorities either.

Consider what would happen if the preferences shown in the election above, which display the phenomenon of cyclical majorities, is the way that members of a legislature feel about alternatives that have arisen as bills to chose from in conjunction with, say, welfare reform. You can think of one of the alternatives as maintaining the status quo (what is done now) and the others as being changes from the status quo. Many legislatures decide on laws by a system of voting which requires pairwise votes on alternatives using a prearranged list of the order in which the votes between the pairs of alternatives are taken. This is sometimes referred to as *agenda voting*. An agenda of pairwise votes is scheduled and the alternative that emerges as the winner is what becomes the law of the land.

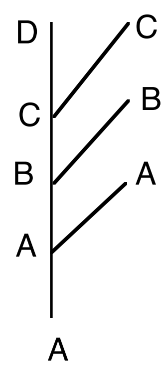
Theorem:

If agenda voting is used to pick a single alternative as winner when the voters have preference schedules that display cyclical majorities, then by suitably choosing an agenda, any of the alternatives can emerge as winner!

Comments:

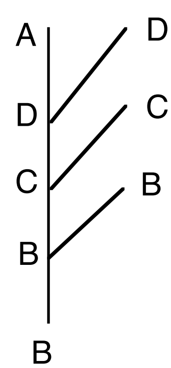
When voters have preferences which display cyclical majorities, any alternative can wind up being chosen by choosing an appropriate agenda. This means that when this phenomenon occurs in a legislature, what becomes law is less an issue of what is "right" than the skill at procedural manipulation of the people who run the legislative body or by accident!

Let us illustrate how this strange phenomenon works. Suppose you wanted A to emerge as the winner in the situation above. You would then arrange for the votes to be taken in the order:

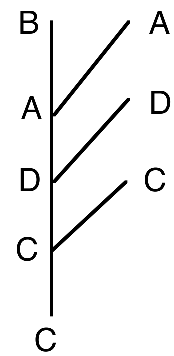


This diagram is read as follows: First, hold a vote between D and C. C will win. Second, hold a vote between C and B. B will win. Third, hold a vote between B and A. Hence, A, listed at the bottom of the diagram, emerges as the winner.

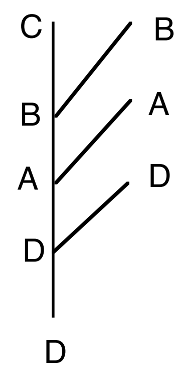
On the other hand, if you wanted B to emerge as the winner, you would arrange for the votes to be taken in the order:



Similarly, if you wanted C to win:



Similarly, if you wanted D to win:

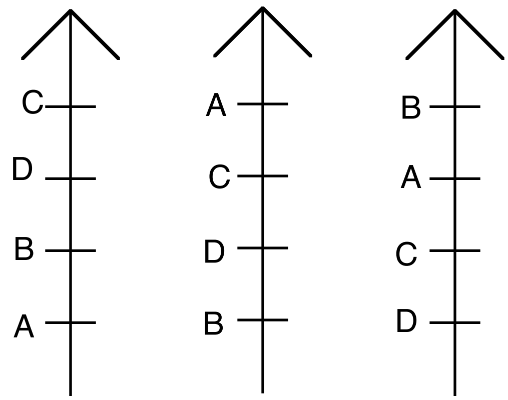


Thus, we see that the procedure to schedule the order in which the votes are to be taken is more important than the vote itself, since that procedure controls what will emerge as law!

Political scientists are divided on to what extent and how commonly this phenomenon occurs in actual legislative voting, but examples have been given which suggest that it has happened sometimes, even for the United States House of Representatives! Anyone who is committed to democracy can not help but be made nervous by examples such as the one above and the real world examples that are waiting in the wings.

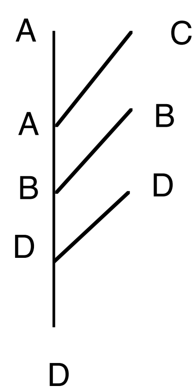
While on the subject of agenda voting, consider the following example:

Example:



votes: 1 1 1

We will use the following agenda: A versus C, winner B, winner versus D.



As illustrated in the diagram above, the result of this agenda is that D emerges as the winner. However, C is preferred to D by all of the voters! This example illustrates that agenda voting can violate the following fairness condition, which is a (weaker) variant of the Pareto condition that we used in Arrow's Theorem:

(*Pareto Condition* II)

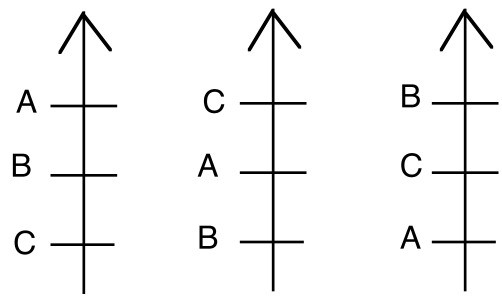
If all the voters rank alternative X over alternative Y, then the election decision method should rank X over Y.

Returning to how to modify Condorcet's Method to obey the first Condition of Arrow's Theorem (i.e. guarantee that for all elections there is a winner), a variety of ideas have been proposed. Clearly, one wants a method by which, when there is a Condorcet winner, that candidate should be selected by these "extensions" of Condorcet's Method. A method which always yields a winner and which selects the Condorcet winner when there is one is called a *Condorcet consistent method*.

Condorcet consistent methods

a. (*Black's Method*) If there is no Condorcet winner, use the Borda Count to decide the election's winner.

Example:



votes: 8 12 9

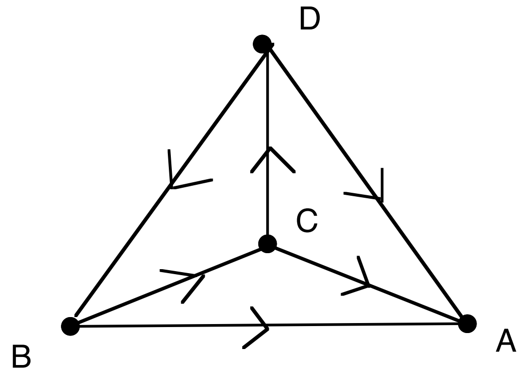
Since A can beat B in a two-way race, B can beat C in a two-way race, and C can beat A in a two-way race there is no Condorcet winner. Hence, we use the Borda count to decide the election. A gets 16 + 12 + 0 = 28, B gets 8 + 0 + 18 = 26, and C gets 0 + 24 + 9 = 33 points. Thus, C is the Borda Count winner and also the winner under Black's method.

Related to the idea of Condorcet Consistency is (Smith's condition):

If the alternatives can be divided into two groups S (superior) and I (inferior), where every alternative in S can beat every alternative in I in a two-way race, then the election method should not select any alternative in I.

Example:

Consider the two-way race digraph below which arose for an election:



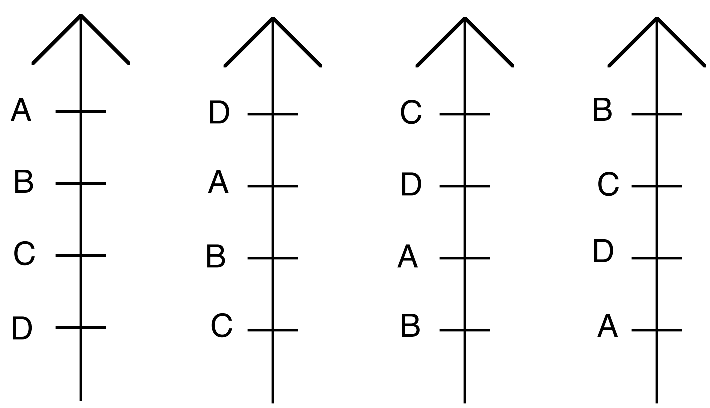
Although there is no Condorcet winner for this election, in order for an election decision method to obey Smith's Condition, the method could not choose alternative A as the winner. This is because B, C, and D can all beat A in a two-way race.

Stop and explore

a. Construct an election which gives rise to this two-way race digraph.

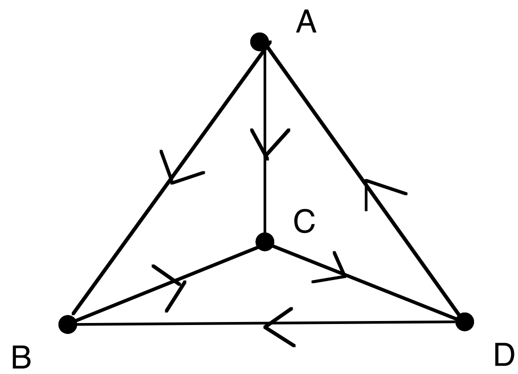
b. (*Copeland's Method*) Count the number of two-way races an alternative wins and subtract from this the number of two-way races that an alternative loses. Whichever alternative gets the largest value for this procedure wins. (This may not be a decisive method in that it may often produce ties, and, hence, no unique winner.)

Example:



votes: 18 13 11 7

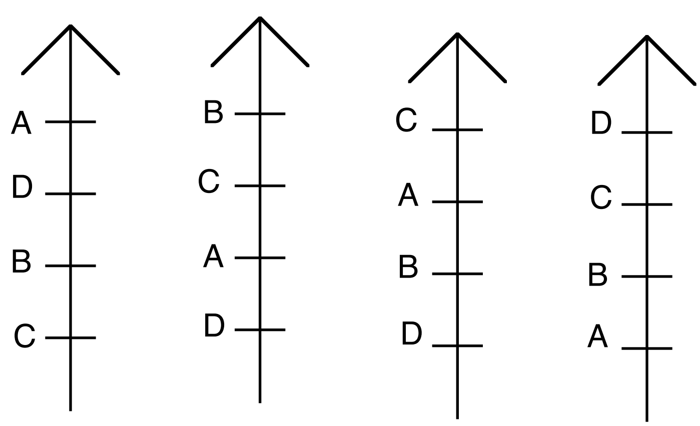
The associated two-way race digraph is:



From this digraph we can see that A wins two races and loses 1, B wins one race and loses 2, C wins one race and loses 1 and D wins two races and loses 1. Thus, using the Copeland method there would be a tie between A and D.

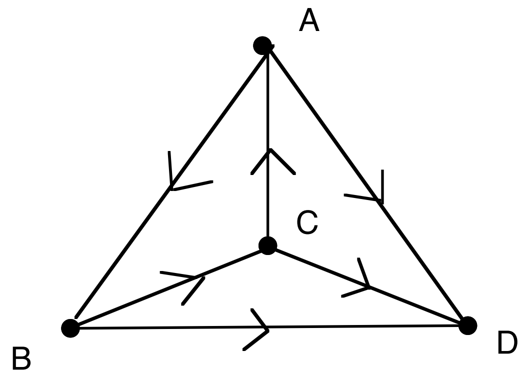
c. (*Simpson's Method*) For each alternative A compute s(A, Z), for every other alternative Z, s(A,Z) being the number of voters who prefer A to Z. The Simpson Value of A would be the smallest value obtained for s(A,Z) as Z ranges over all alternatives other than A. The Simpson winner is the alternative achieving the highest value of s(A,Z). If the pair-preference matrix has been computed, the Simpson winner can be found by first computing the minimum entry in each row of this matrix. Now the winning alternative(s) is found by finding the maximum value of the row minima.

Example:



votes: 18 12 10 1

The associated two-way race digraph is:



$$
\begin{array}{c}
\\
A \\
B \\
C \\
D
\end{array}
\begin{array}{c}
\begin{array}{cccc} A & B & C & D \end{array} \\
\begin{bmatrix}
\text{----} & 28 & 18 & 40 \\
13 & \text{----} & 30 & 22 \\
23 & 11 & \text{----} & 22 \\
1 & 19 & 19 & \text{----}
\end{bmatrix}
\end{array}
$$

To find the Simpson winner, first compute the minimum in each row of this matrix:

Row A: 18   Row B: 13   Row C: 11   Row D: 1.

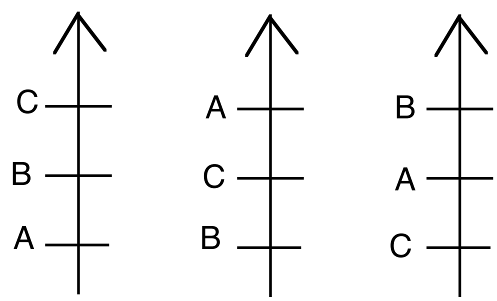
The maximum of these numbers is 18 in A. Hence, A is the Simpson winner. Note that if the Copeland method is applied to this election, the result is a three way tie between A, B, and C since each of these alternatives can win 2 two-way races and lose 1 two-race, while D loses all two-way races.

The difference between Copeland's and Simpson's approaches is that Copeland's method places importance on the number of two-way races one wins and not by how much you might win each two-way race, while Simpson's method weighs by how much you beat the other candidates. If there is a Condorcet winner, then either Copeland's or Simpson's method will select this winner.

d. (*Nanson's Method*) Use the Borda Count as a sequential elimination method.

(Nanson proved that if there is a Condorcet winner, this method will result in that alternative's being chosen!)

Example:

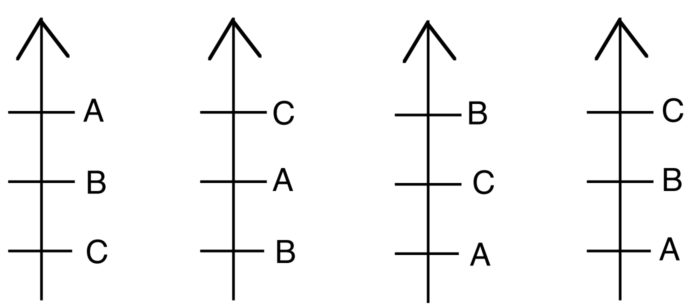


votes: 9 14 10

To apply Nanson's method we first compute the Borda Count for each of the candidates. A gets 0 + 28 + 10 = 38, B gets 9 + 0 + 20 = 29 points, while C gets 18 + 14 + 0 = 32 points. Since B got the fewest points, B is eliminated and a run-off is held (via the Borda Count, though when there are only two candidates this is equivalent to the usual plurality/majority method). The run-off between A and C results in a victory for A. Thus, A is the Nanson winner. Note that this election has no Condorcet winner.

All of these ideas are appealing to some extent but, of course, all are subject to Arrow's Theorem or disobeying other reasonable fairness criteria. By way of illustration, consider the following example due to Peter Fishburn, which shows that Nanson's method is not monotonic: more favorable treatment for a candidate can hurt the candidate!

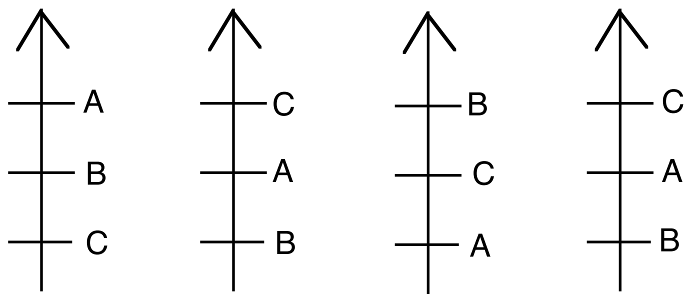
Example (Peter Fishburn)



votes: 8 5 5 2

Applying the Borda Count (2 points for a first place, 1 point for a second place, 0 points for a third place) gives A 21 points, B 20 points, and C 19 points. C is eliminated and in the subsequent run-off, A beats B 13 to 7.

Now consider the modified election where the last two voters have changed their minds and ranked A higher than B (rather than B higher than A). Intuition suggests that this should only help A's cause.



votes: 8 5 5 2

The Borda Count now gives A 23 points, B 18 points, and C 19 points. As a result, this time candidate B is eliminated. In the run-off between candidate C and candidate A, C wins 12-8! Thus, Nanson's appealing method suffers from the unpleasant defect that more support for a candidate can change a win to a loss!

Due to the seeming appeal of electing a Condorcet Method winner when there is one, many ideas related to Condorcet's approach have emerged. For example, one fairness rule that one might like an election decision to obey is that the winner of an election never be a candidate who would lose to every other candidate in a two-way race:

Condorcet Loser Condition

If alternative A would lose to every other candidate in a two-way race, then an election decision method should not choose A as its winner.

The run-off, sequential run-off, and Borda Count can elect a Condorcet loser. However, plurality voting can not choose a Condorcet loser. Yet the plurality winner may lose to any other candidate in a two-way race! Discussions of this kind show the subtle relations between the methods and desirable fairness conditions for these methods.

Stop and explore:

a. Can you construct an example of an election where the Borda Count elects a Condorcet loser?

b. Can you explain why the plurality method can not elect a Condorcet loser?

c. Can you construct an example of an election where the run-off method elects a Condorcet loser?

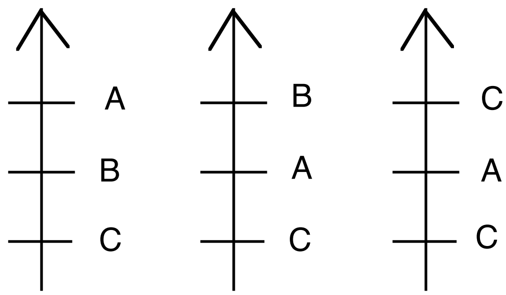
d. Can you construct an example of an election where the sequential run-off method elects a Condorcet loser?

(e). *Carroll's Method*

This method was developed by Lewis Carroll, the famous author of Alice in Wonderland. Somewhat less well known is that C.L. Dodgson - Carroll was only a pen name - was a mathematician at Cambridge University. He was among the pioneers in using mathematics to understand elections. Carroll developed the method which is described below, using modern terminology, some of it chosen to honor him.

Consider an election in which each voter produces a preference schedule where no ties among candidates are allowed. For each candidate, define the Carroll number for candidate i as the minimum number of switches between adjacent candidates in the preference schedules which is necessary for candidate i to become the winner of the (revised) election using Condorcet's Method. The winner of the election under Carroll's Method is the candidate whose Carroll number is as small as possible. (There may be several candidates with smallest Carroll number.) When the smallest Carroll number is zero, the candidate who achieves this number must be the unique Condorcet winner, so this method is Condorcet consistent.

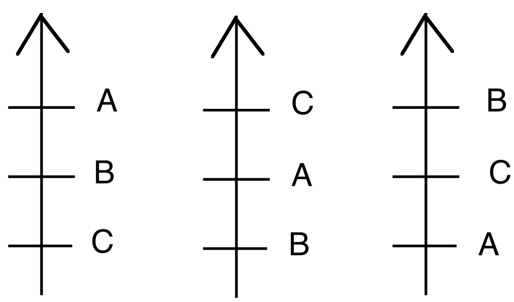
Example:



votes: 7 5 3

In this example the Carroll score of A is 0, the Carroll score of B is 3, and the Carroll score of C is 10. Hence, A is the winner, who, of course, is also the Condorcet winner.

Example:



votes: 4 3 2

In this example, no candidate has Carroll number 0, and the winner would be A, who has the lowest Carroll number, namely 1. The Carroll numbers of B and C are 3 and 2, respectively.

Stop and Explore

a. Who is the Borda Count winner for the two elections above?

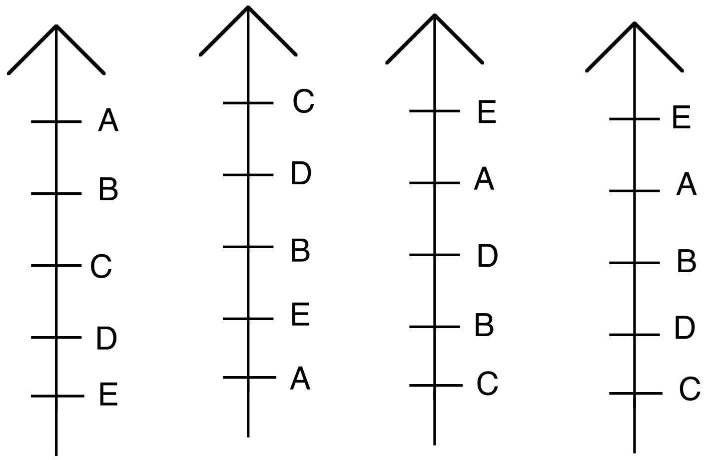
b. Will the Borda Count winner always be the same as the Carroll winner?

This method, unfortunately, introduces a new wrinkle into our analysis of election methods. This new wrinkle involves the "computational complexity" of computing the winner of the election. In this context, computational complexity refers to the amount of work that must be done to compute the winner. Even when there are many voters and many candidates, the amount of work to decide the winner of a large election using the plurality method or the Borda Count is not extraordinarily large. However, it has been shown that the Carroll method is a computationally difficult problem to solve. In other words as the size of an election that must be decided by the Carroll method grows, the amount of computation necessary to solve the problem appears to increase rapidly. (More precisely, no known method of deciding the winner relatively quickly for large problems is known, and it is thought unlikely that any "fast" method will be found.) Using very fast computers, it is still feasible to deal with the question for small numbers of candidates and large numbers of voters, but for large numbers of candidates and voters no computer, no matter how fast, could decide the winner of the election. Hence, no matter how appealing this method may be in principle, it would be difficult to implement in practice.

Contrasting the Borda Count and methods which select a Condorcet winner

In light of Arrow's Theorem and the flaws with plurality voting and methods based on run-off elections (either the traditional run-off or sequential run-off methods), it is natural to look to either the Borda Count or the Condorcet Method as the election method of choice. Among the problems with the Condorcet Method is that it does not choose a winner in all elections, while although the Borda Count always chooses a winner, it does not always choose the Condorcet winner when there is one. (Note, supporters of the Borda Count are not always apologetic over this fact.) One natural approach, as explored above, was to use a method which, if there is a Condorcet winner chooses this alternative but which does yield a winner in the case that no Condorcet winner exists. Methods which do this are Copeland's Method, Simpson's Method, Nanson's Method, and Carroll's Method. As we saw we might wish to eliminate Nanson's Method from consideration because it violates the Monotonicity Condition. However, Copeland's Method obeys the monotonicity condition. How does Copeland's Method stack up against the Borda Count?

Example: (Philip Straffin, Jr.)

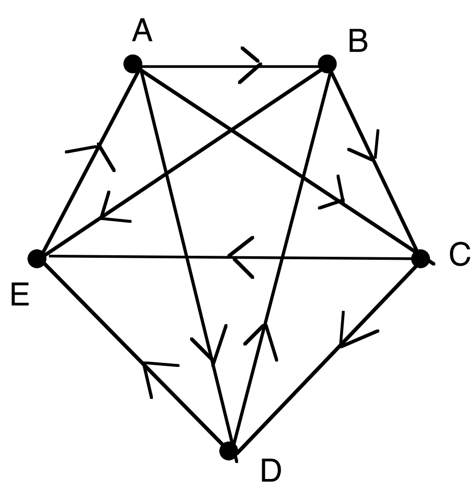


votes: 1 4 1 3

The Borda Count for this election gives:

A: 16 B: 18 C: 18 D: 18 E: 20

Thus, E is the Borda Count winner. Now we will construct the two-way race digraph for the same election. It is shown below:



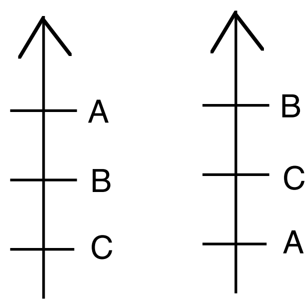
As show in the digraph, A can beat 3 alternatives in two-way races (loses one), B can beat 2 alternatives in two-way races (loses two), C can beat 2 alternatives in two-way races (loses two), D can beat 2 alternatives in two-way races (loses two), and E can beat 3 alternatives in two-way races (loses one).

Since the Copeland score for an alternative is obtained by taking the difference between the number of two-way races that an alternative wins and the number of two-way races it loses, the Copeland scores of the alternatives are:

A: 2 B: 0 C: 0 D: 0 E: -2.

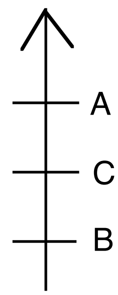
Thus, candidate A is the Copeland winner. In this example, the alternatives with the highest and lowest Borda Counts, E and A, come in last and first, respectively, using the Copeland Method! Yet, when we make a direct comparison of E and A, 8 of the 9 voters prefer E to A! Examples such as this perhaps take the steam out of people who love the Copeland Method. On the other hand, consider the following example.

Example:



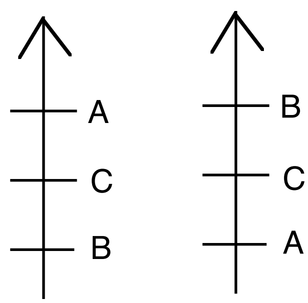
votes: 3 2

The Borda Count gives: A: 6 points, B: 7 points, and C: 4 points. Thus, B is the Borda Count winner. However, A can beat B and C in two-way races, and, thus, is the Condorcet winner (and Copeland winner). However, more dramatically, A gets a majority of the votes cast. Therefore, this example shows that the Borda Count can violate the majority principle (i.e. if there is a candidate who gets a majority of the votes cast, that candidate should win). If the three voters who prefer candidate A know that the Borda Count method is being used, they can use sophisticated voting. If the voters supporting A now vote for the preference schedule below:



votes: 3

we now have the election below:



votes: 3 2

In this election the Borda Count winner is A, and no strategic action of the voters who prefer candidate B will counteract the sophisticated action of the voters who prefer A.

In a general way, voters who know that the Borda Count is to be used and prefer candidate X yet view candidate Y as the greatest threat to candidate X, can minimize the chance Y will beat X. This is accomplished by placing X at the top of the preference schedule they vote for and Y at the bottom, regardless of what they truly feel about candidate Y. When this "sophisticated voting" approach to dealing with the use of the Borda Count was called to Borda's attention, he is reputed to have replied "My scheme is only intended for honest men."

When we developed the Borda Count, we chose a particular method of assigning points to the alternatives in doing the computation. If we assign points in some other way, might we get a method which was even better than the Borda Count? The Borda Count is one of a family of voting methods which are known as *scoring rule* voting methods. The idea behind scoring rule methods is the selection of a sequence of weights, points, or "scores" which are assigned to alternatives depending on the rank the alternative has on a voter's preference schedule. We will assume that the voting method is specified by the selection of a non-decreasing sequence of scores:

$$s_0 \leq s_1 \leq \ldots \leq s_{n-1} \text{ (where } s_0 < s_{n-1} \text{ and n is the number of alternatives).}$$

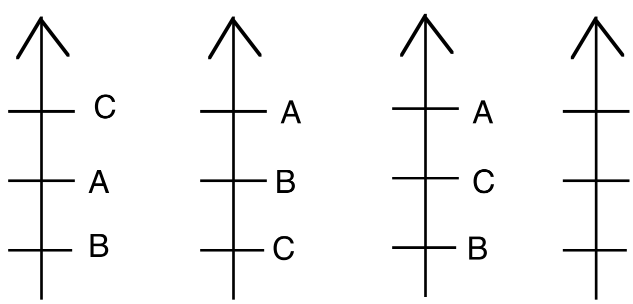
The significance of the scores is that a last-place choice is given s_0 points, a next-to-last-place choice is given s_1 points, etc. Note that we do not allow all the scores to be the same, since if all the scores were equal, then no account would be being taken of how high or low on the preference schedule a voter ranked a candidate. Scoring rule methods include not only the Borda Count but also the plurality method. For plurality, one assigns the same number as scores for $s_0,\ldots,s_{n-2}$ and a larger number than these for s_{n-1}. This interpretation of plurality voting as a scoring rule method suggests another voting method, known as the *anti-plurality,* where voters pick the candidate they like the least and the candidate who gets the smallest number of votes wins. We saw that the Borda Count may not chose a Condorcet winner when there is one. The follow theorem, proved by Peter Fishburn, shows that this is true for all scoring rule voting methods.

Theorem (Fishburn):

There exist elections where the Condorcet winner is never elected by any scoring rule voting method.

Proof:

Case (i) (The scores of the score sequence are strictly larger than each other):



votes: 3 2 1 1

In this example, C would be the Condorcet winner. Now let us consider who would win if a score sequence is being employed. Using s_0, s_1, and s_2 (where $s_0 < s_1 < s_2$) to denote the scores, we can write down the scores for the three candidates:

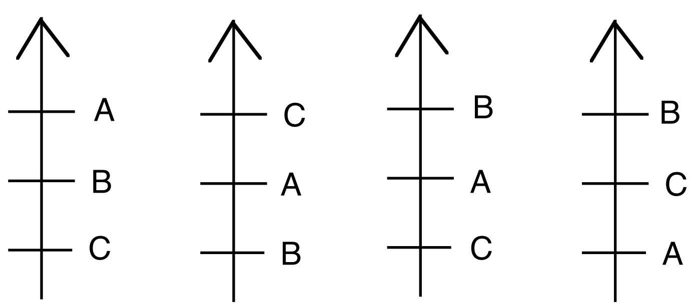
A: $3s_1 + 2s_2 + 1s_2 + s_0 = 3s_2 + 3s_1 + s_0$

B: $3s_0 + 2s_1 + 1s_0 + 1s_2 = 1s_2 + 2s_1 + 4s_0$

C: $3s_2 + 2s_0 + 1s_1 + 1s_1 = 3s_2 + 2s_1 + 2s_0$

Due to the ordering of the scores these values mean that the score for A is larger than either that of B or C and that A will be the victor, despite the fact that C is the Condorcet winner.

Case (ii) (The scores need not be strictly increasing):



votes: 6 3 4 4

In this example A is the Condorcet winner. Now let us see who would win if a score sequence were being used. For convenience, and without the argument's being less rigorous, we can assume that s_0 is zero. The other two scores will then satisfy $0 \le s_1 \le s_2$ where $s_2 > 0$. (In this we allow the possibility that some scores are the same but we do not allow them all to be equal, since to do otherwise would violate the definition of a scoring rule.) Now let us compute the scores for all the candidates:

A: $6s_2 + 3s_1 + 4s_1 + 4s_0 = 6s_2 + 7s_1$ (remember $s_0 = 0$)

B: $6s_1 + 3s_0 + 4s_2 + 4s_2 = 8s_2 + 6s_1$

C: $6s_0 + 3s_2 + 4s_0 + 4s_1 = 3s_2 + 4s_1$

Clearly, under the assumptions about the scores, B beats C. Does B beat A? We can rewrite B's total number of points as: $8(s_2 - s_1) + 14s_1$. Now $8(s_2 - s_1) + 14s_1 > 6(s_2 - s_1) + 13s_1 = 6s_2 + 7s_1$, which is the total number of points A gets. Thus, B gets more points than either A or C and wins, despite the fact that A is the Condorcet winner.

Fishburn's Theorem shows that fans of scoring rule methods, in particular the Borda Count, must be resigned to the fact that scoring rule methods can not always guarantee the election of a Condorcet winner when there is one.

We can get an even clearer understanding of the trade-off between methods which elect a Condorcet winner when there is one which uses scoring rule methods. This is because it is possible to give an exact description of a simple set of fairness rules that scoring rule methods satisfy. This remarkable accomplishment is the work of H. Peyton Young.

In order to explain Young's result first we will describe two fairness principles, one which applies when we are using an election decision method which picks a single winner for society (no ties allowed); the other applies when we are using an election decision method which picks a set of winners for society (ties are allowed):

(*Disjoint groups principle, no ties*); Suppose two groups of voters G_1 and G_2 which have no voters in common (i.e. G_1 and G_2 are disjoint) must choose from the same set of alternatives X. Suppose that G_1 chooses A (from X) as its winner and that G_2 also chooses A (from X) as its winner. Now assume that groups G_1 and G_2 are combined into one group, then the combined group should also choose alternative X.

(*Disjoint groups principle, ties allowed*): Suppose two groups of voters G_1 and G_2 which have no voters in common (i.e. G_1 and G_2 are disjoint) must choose from the same set of alternatives X. Suppose that G_1 chooses A_1 (a subset of X) and that G_2 chooses A_2 (a subset of X). Now assume that groups G_1 and G_2 are combined into one group. If A_1 and A_2 have any alternatives in common, then the combined group should select these as being equally good outcomes.

Theorem (H. Peyton Young, 1977)

i. Any scoring rule method which when there is a tie between alternatives with the highest score chooses the subset with highest score satisfies the disjoint groups principle, ties allowed. If the way that ties are broken is that a specific fixed ordering of the alternatives X is given and the ties are broken as in the ordering, then the corresponding scoring rules obey the disjoint group principle, no ties allowed.

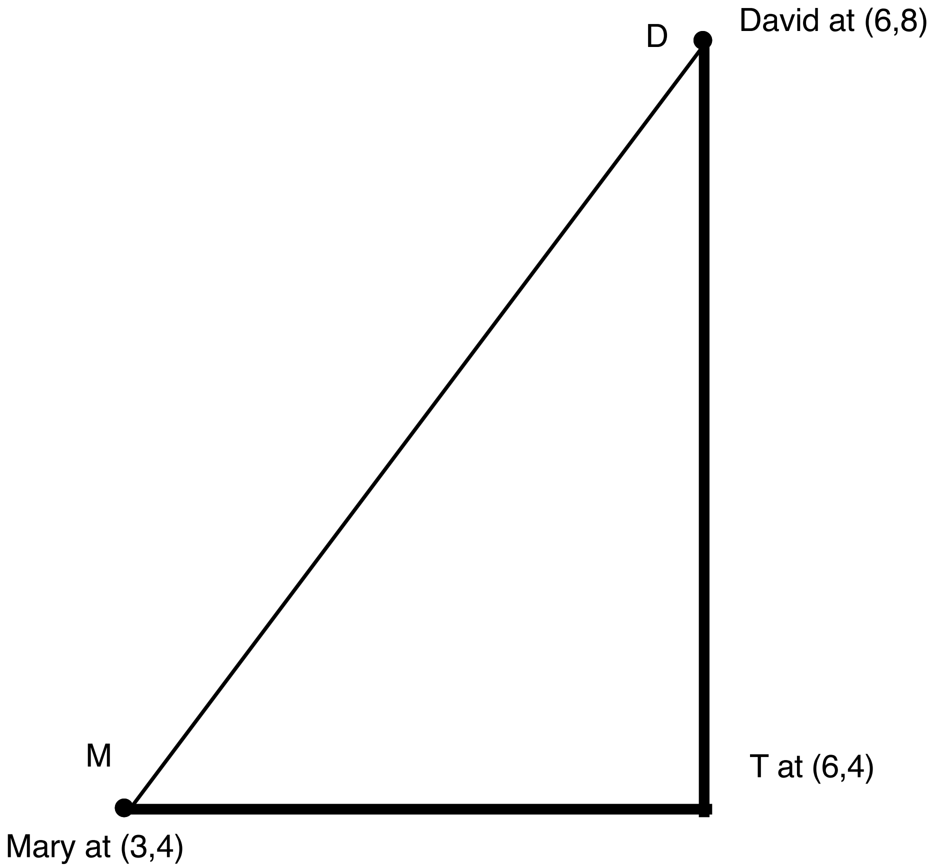
ii. There is no election decision procedure without ties which chooses a Condorcet winner and obeys the disjoint group principle, ties allowed. There is no election decision method with ties which chooses a Condorcet winner and which obeys the disjoint group principle, ties allowed.

Young's theorem places in sharp relief one dramatic contrast between the Borda Count and methods which select a Condorcet winner.

Finding a society choice close to that of a group of individuals choice

One natural way to try to decide on a society ranking based on the preference schedules of a group of individuals is to choose for society a ranking which is "as close as possible" to the schedules produced by the individuals. However, how do we tell how close or far apart two rankings are to each other? For that matter, how do we tell when any objects, two cars or two people or two molecules are close or far apart from each other? This may seem like a strange question since, although we often ask how close or far apart two places are, we are not accustomed to ask how far apart two cars, molecules, or preference schedules are. So let us begin with the familiar. How do we tell how far apart two locations are?

To tell how far apart two locations are we measure the distance between them. However, even for distance between two locations there is more here than might meet the eye. For example, how far apart are the two friends Mary and David, who live at 3rd Street and 4th Avenue and 6th Street and 8th Avenue respectively? If you look at the diagram showing these two locations below, remembering that Mary and David live in a city, how far apart would you say their homes are?



As the crow flies, the distance between where Mary and David live can be computed using the Pythagorean Theorem. The distance from Mary's home to David's is the length of the line MD, the side opposite the right angle (called the *hypotenuse*) in the right angle triangle MDT. The Pythagorean Theorem states that in a right angle triangle, the square of the length of the hypotenuse is equal to the sum of the lengths of the squares of the other two sides. Hence, in this situation, since MT has length 3 and TD had length 4, the length of MD is 5. (Note that $5^2 = 3^2 + 4^2$.) The distance between two points in a flat plane, measured as the crow flies, is called *Euclidean distance*. The problem in this situation is that since neither Mary or David is a crow, they can not get between each other's houses by traveling 5 units. Assuming the city is laid out in a grid of streets as implied by the names of the locations they live at, each would have to travel 7 units to the other's home. This distance is often referred to as the *taxicab* distance between the two points to distinguish it from the more familiar Euclidean distance, since this is how far a taxi would have to travel to get between the two homes. One route that would achieve this distance in walking (or taking a taxi) to David's house would be for Mary to go from M to T (3 units) and then from T to D (4 units).

Stop and Explore

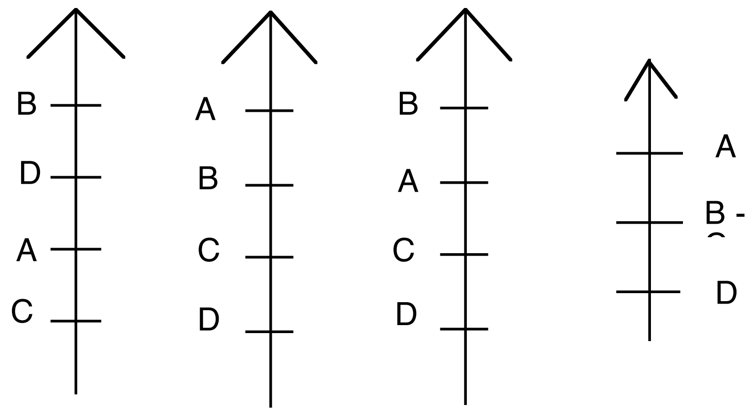
a. If the town where Mary and David live actually has a 3rd, 4th, 5th and 6th Street, and a 4th, 5th, 6th, 7th, and 8th Avenue, how many different routes could David find that have length 7 which would get him to Mary's House.

The distinction between Euclidean and taxicab distance shows that one can measure distance between locations in more than one way. The natural distance between two locations, as this example shows, depends on circumstances. Sometimes Euclidean distance is the right measure and sometimes taxicab distance is the right measure.

In fact, from an abstract point of view, distance is nothing more than a way of assigning a number to a pair of objects so that certain rules are obeyed. The rules that the number assigned should obey are: 1. The distance between two objects is always a positive number or zero. 2. The distance between two objects is zero only if the objects are the same. 3. (Symmetry) The distance from object A to B is the same as the distance from B to A. (In some applications one can "omit" this condition on distance). 4. (Triangle inequality) Given three objects, the distance from A to B plus the distance from B to C is greater than the distance from A to C.

Any scheme that obeys these conditions (axioms) can be called a *distance*. Thus, we can now contemplate computing the distance between two insulin molecules (to see if chimpanzees are closer to baboons or to gorillas by this measure), between two people (to see if I am more closely related to one cousin than another), or between preference schedules (to see which might be a better choice for society based on the preferences of a group of individuals).

For example, how might we determine the distance between a pair of the following preference schedules?



voter u v w x

In comparing the preference schedules of voters v and w they seem more in agreement than the preference schedules of voters u and v. Let us concentrate on two particular preference schedules, those of voters u and v. For some pairs of alternatives (e.g. B and C) u and v agree in preferring one of the alternatives over the other. In other cases (e.g. D and A) u and v disagree as to which is the better alternative. The way we will find the distance between two preference schedules is by adding up the numbers which show how the preference schedules differ on each pair of alternatives. Here are the details. Suppose that P and Q are two different preference schedules which rank the alternatives of set X, and let A and B be any two alternatives in X. First, we define r(A,B) as follows:

r(A,B) = r(B,A) = 2 if on preference schedules P and Q one ranks A over B and on the other B ranks over A

1 if on preference schedules P and Q one ranks A over B or B over A and on the other A and B are tied

0 if on preference schedules P and Q the alternatives A B are ranked in the same way (i.e. both P and Q rank A over B or both rank B over A.)

Note that if the preference schedules show no ties (indifference) between alternatives, then r for any pair will be either 0 or 2.

Next, to compute d(P,Q):

d(P,Q) = sum of r(A,B) for all choice of pairs {A,B} (in X)

Thus, for example, if X contains 4 alternatives, the distance between two preference schedules requires that r be computed for 6 pairs.

Example:

1. Compute the distance between the preference schedules of voter u and v above. r(A,B) = 2; r(A,C) = 0; r(A,D) = 2; r(B,C) = 0; r(B,D) = 0; r(C,D) = 2. The distance between the preference schedules for u and v equals the sum of these six numbers and is 2 + 0 + 2 + 0 + 0 + 2 = 6.

2. Compute the distance between the preference schedules of voter v and w above. r(A,B) = 2; r(A,C) = 0; r(A,D) = 0; r(C,D) = 0; r(B,D) = 0; r (C,D) = 0. The distance between the preference schedules for v and w equals the sum of these six numbers and is 2 + 0 + 0 + 0 + 0 + 0 = 2.

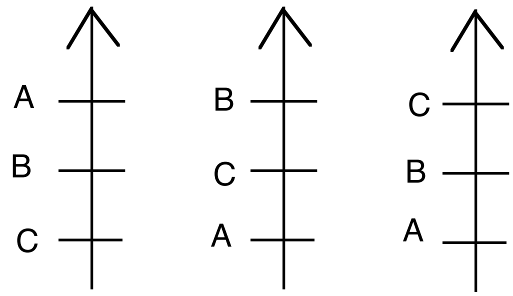
<u>Stop and Explore</u>

a. Compute the distance between the preference schedules of voters u and w above.

b. Compute the distance between the preference schedules of voters u and x above. Note that here, voter x is indifferent between some of the alternatives.

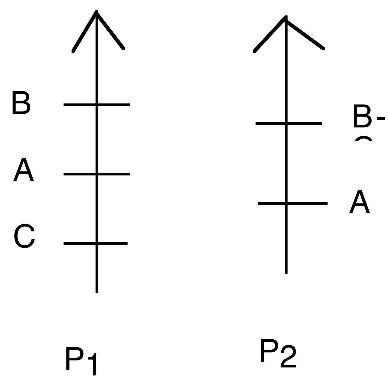
The ideas we have been developing here were the work of the mathematician John Kemeny. He showed that subject to some reasonable conditions that one might like to have the distance between preference schedules to obey, there was only one way to define distance between preference schedules, and this is what has been described above. The next step that Kemeny took was to develop an election decision method based on the idea of the distance between preference schedules. Here are the basic ideas. Suppose that L is the election consisting of a collection of preference schedules for a group of voters who are choosing from the ranking k different alternatives. We can now measure the distance between a particular possible choice P that might be made for society (ties allowed) and each of the preference schedules produced by the voters in L, and add up all of these distances. We can now choose among the many possible choices of P for society, that choice for society which minimizes the sum of the distances. Although this is an intriguing idea, there are several complications. First, the number of choices of rankings for society goes up quickly. Even for 3 alternatives, there are 13 choices for society. Second, there is no guarantee that the method will produce a unique ranking.

Example:



votes: 1 1 1

Consider the rankings below, which are two of the 13 that might be chosen for society, denoted here by P_1 and P_2:



First, we compute the distance between P_1 and the leftmost ranking in the election. Then we must obtain a distance for the middle and right hand rankings of the election. Finally, we sum these three numbers to obtain the distance we need for P_1 and the election as a whole. It turns out that this distance is 8. If we compute the distance between P_2 and the election as a whole, the result is only 7. Thus, we would conclude that in this case P_2 is closer to the votes of the individuals that make up society than is P_1.

Stop and Explore

a. Compute the distance between the other 11 possible rankings for society and on the basis of these calculations decide which preference schedule is the closest one to the votes of the three individuals involved.

As you see, the computation required is rather great, though in principle one can write a computer program which could do this, which could handle the situation for small numbers of alternatives.

Now we will continue by looking into majority rule itself, the core of what many find appealing about the Condorcet approach.

Is Majority Rule fair?

In many democracies the process of carrying out an election is going on in many districts simultaneously. For example, in the United States every two years the whole House of Representatives, consisting of 435 members, is elected. Each of these elections is conducted currently by plurality in each of 435 districts. The 435 districts fall within the different states and the number of districts in each state is arrived at by using the Huntington-Hill Method of apportionment (see [2]). In the United States, which historically has been dominated by two parties at any given time (though the two parties involved have changed with time), each of the two parties typically runs a candidate in each of these districts. The candidates who run may be men, women, black, white, gay or heterosexual, etc. Sometimes, people are interested in characteristics of the candidates that go beyond mere party identification. For example, there are times or periods where interest in electing women candidates or black candidates can take on importance. Let us see how the plurality system used in America can affect this desire.

For example, imagine a hypothetical state in which race is important, which has 5 districts and in which the racial data of the population of the district is given.

District 1: White 125,000; Black 117,000

District 2: White 132,000: Black 113,000

District 3: White 118,000; Black 110,000

District 4: White 123,000; Black 115,000

District 5: White 116,000: Black 112,000

Totals: White: 614,000   Black: 567,000

Furthermore, imagine that party A tends to support white candidates and party B tends to support black candidates. Suppose that party A runs only white candidates and party B runs only black candidates. If voters vote only on the basis of race, 5 white candidates will be elected, all from party A! In the winner-take-all environment characteristic of American democracy even though nearly 50% of the population of this state is black, a black representative is never elected. Some people feel that this phenomenon, that large minorities feel that they do not have effective representation with regard to some characteristic, sows the seeds of long-term social unrest. One way to resolve a problem of this kind is to "gerrymander" districts so that one or more districts with a black majority would be drawn. *Gerrymandering* is the practice of drawing districts, often with irregular shapes, with the goal of achieving a particular political objective. However, the Supreme Court has recently decided a series of rulings which restricts the rights of state legislatures from drawing districts to achieve specific racial goals. (However, the Supreme Court has never ruled that it is not legal to draw districts in a way that favor a particular political party!) To deal with the problem that occurs when election methods result in skewed situations, a movement has evolved that urges that elections take place in an environment where *proportional representation* can occur. The idea in this context would be that instead of having 5 separate districts with a separate election in each district, one district be created and 5 candidates be selected for the district. Now the goal would be to select a method which "allows" proportional representation.

Proportional Representation

Unfortunately, the term proportional representation has come to be used for two different kinds of situations. First, when seats in a legislature are to be filled and voters vote for parties rather than individual candidates, one wants the seats in the legislature to be assigned in proportion to the vote the parties received. Second, when many candidates are running, with different characteristics (i.e. party, gender, race, etc.) for a block of seats in a legislature, one might hope that the voters who consider these characteristics important, can elect representatives with these characteristics in proportion with these characteristics among the voters. Here we will be concerned with the second interpretation of proportional representation.

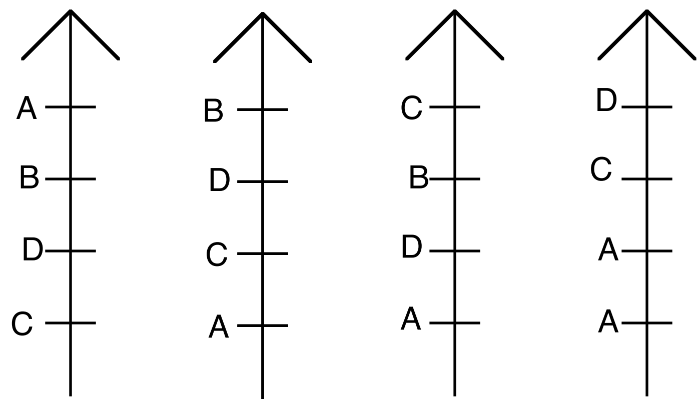
Imagine that h seats are to be filled by a group of candidates who are voted for "at large" by m voters. We will assume that the voters use an ordinal ballot to rank all of the h candidates, with no indifference among candidates allowed. (In fact, it is known that there can be a strategy when the election decision method described below is used in "truncating one's ballot" (e.g. not voting for all h candidates). However, for ease of explaining the basic ideas, this complication will be avoided here.) We will describe a method know as the *Hare Method* or the *Single Transferable Vote* (STV), which seems to some extent to work as a proportional representation system. (The method of *Cumulative Voting*, mentioned below, allows one to try to achieve similar goals.)

The idea behind STV is that if a candidate gets a certain minimum quota of votes that the candidate is elected. If a candidate gets more than the minimum quota of votes, such extra votes will be transferred to other candidates so that no voter's vote is "wasted." (In principle excess votes should be transferred proportionally to lower place choices.) If at a particular stage in the vote count no candidate has received the requisite vote to fill a seat not already filled, then the candidate with the lowest vote count at this stage is eliminated and his/her votes transferred to the candidate who is next highest on the preference schedules of the ballots involved.

In practice STV is complicated by the fact that there are several different ways that the minimum quota necessary for election can be computed (e.g. Hare Quota, Droop Quota, or Imperiali Quota, etc.) Furthermore, if a candidate has more than the quota for election, the results of the election can depend on which particular "excess" ballots above the quota are transferred. Here we will consider only the Hare quota and work an example to show the spirit of the method without getting into "practical" details.

Note that the reason why this method can work as a proportional representation system is that if a group of voters who do not constitute a majority with respect to some characteristic (e.g. race or gender) vote as a "block," then if the size of the group exceeds the quota they can guarantee the election of at least one candidate whom they desire to see elected.

Example:



votes: 45 34 27 14

Suppose there are two seats to be filled. In this case, since there are 120 votes total and two seats to be filled, the Hare quota is (120/(2+1)) + 1 = 40 + 1 = 41.

Next consider the number of first place votes for each candidate:

A = 45 B = 34 C = 27 D = 14.

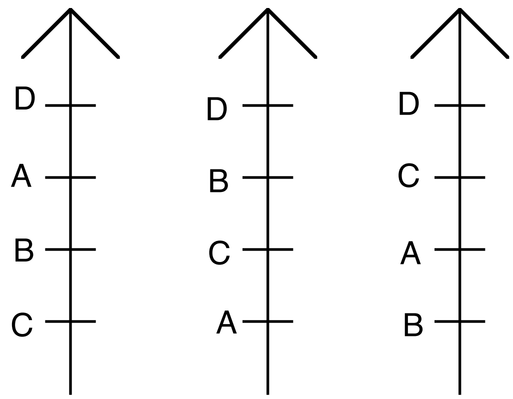
Since A's vote exceeds the quota required to be elected, A gets one of the two seats. However, A's 45 votes exceeded the quota by 4, so 4 of A's votes are transferred to the next highest candidate on the preference schedules involved. (In this example, there is only one kind.) Since all voters who voted for A ranked B second, 4 votes are transferred to B. At this stage we have:

A = 41 (elected) B = 38 C = 27 D = 14.

Since after the transfer no new candidate has a quota, the candidate with the lowest number of votes at this stage is eliminated. This would be candidate D and since the people who voted for D rank C second, 14 ballots are transferred to C. Since C now has 27 + 14 = 41 votes, where 41 is the quota, C is declared elected. Thus, A and C are the two candidates who win seats. Note that if only one seat had been up for grabs, it would have gone to C, not to A.

There are a variety of theoretical difficulties with the Single Transferable. One problem is that if the rules of the election allow a voter not to vote for all the choices available, this can make a difference in the results of the election. The example which follows is due to S. Brams.

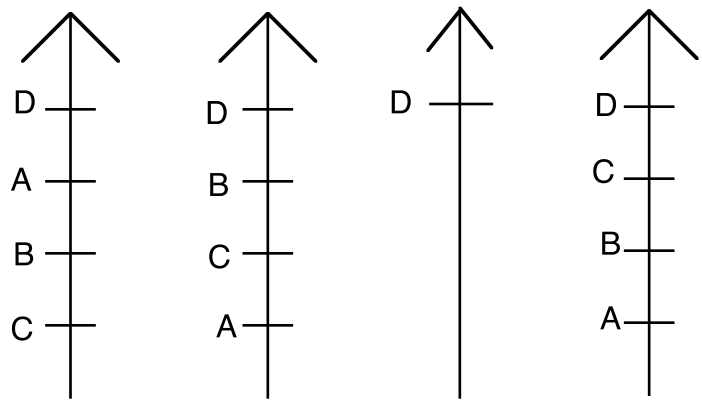
Example (S. Brams):



votes: 6 6 5

Suppose there are two seats to be filled. Then the quota would be 6. On the first round D gets 17 first-place votes and wins one of the two seats. In the next round, the 11 votes above the quota are transferred to second-place candidates in the ration of 6 to 6 to 5. This means that 3.9 votes are transferred to A, 3.9 votes are transferred to B, and 3.2 votes are transferred to C. Since C would have the fewest votes, C's 3.2 votes would be transferred to A, who now would have 3.9 + 3.2 = 7.1 votes and would, thus, be elected to the second seat.

Compare the results of the election above with what would have happened if 2 of the six voters who had B as a second choice had voted only for D instead. The election involved is shown below:



votes: 6 4 2 5

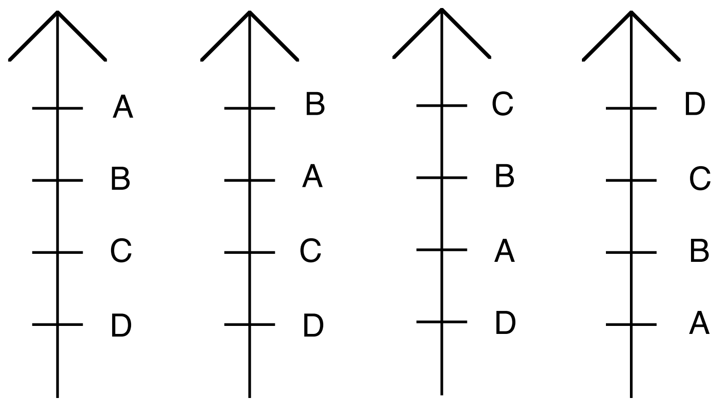
Again, in the first round, 17 votes are cast for D and D is elected to one of the two seats. Now it is necessary to transfer 11 votes in the proportions: 6 to 4 to 2 to 5. This means that 3.9 votes are transferred to A, 2.6 to B and 3.2 to C. Note that no transfer can be made for the voters who did not indicate a second place choice. No candidate receives the needed quota on the basis of this transfer. Hence, since B got the fewest votes in this round, B is eliminated and the 2.6 votes for B are transferred to the next highest candidate, namely, C. After this transfer no candidate is elected. Thus, the candidate with the next lowest number of votes, A (with 3.9 votes) is eliminated. The votes are now transferred to the remaining candidate C, who is elected. Hence, the winners this time are D and C.

This example shows that (sincere) truncation of one's ballot can affect who is the winner if the STV method is used.

It is possible to design an example which does not depend on the transferring of surplus votes.

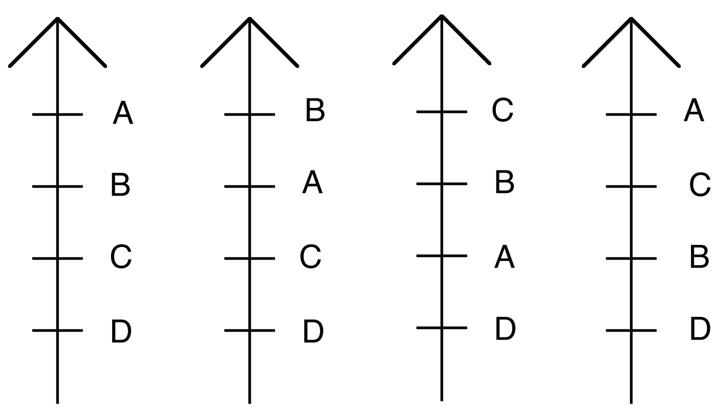
Perhaps even more disturbing than this phenomenon is that STV does not obey the monotonicity principle.

Example (S. Brams)



votes: 7 6 5 3

Since no candidate has a majority (11 votes), the lowest vote getter D is eliminated. In the next round C gets the votes that went to D in the previous round but still no candidate has a majority. Hence, B is eliminated. In the final round vote between A and C the winner is A, by 13 to 8. Now consider what happens if the three voters who ranked D first change their ranking by interchanging A and D, leaving the positions relative to the other candidates the same. The new election is shown below:



votes: 7 6 5 3

Now, in the first round C is eliminated. In the election that results between A and B, B wins 11 to 10! Thus, increased support for A results in B's winning instead of A. This feature of the Single Transferable Vote (sequential run-off) is rather disconcerting!

One example of the success of the STV in obtaining proportional representation was its use by voters in New York City to elect members of the City Council. In 1947 the STV resulted in a member of the Communist Party's (a legal party at the time) being elected to the City Council. Not only did the council not allow the elected person to serve but also immediately changed the system. At least in this case, STV seemed to have worked too well for some!

Other examples of where STV is used, for varying reasons not always related to proportional representation, are in Australia, Ireland, and New York City local school board elections.

Below is another approach to protecting minority interests.

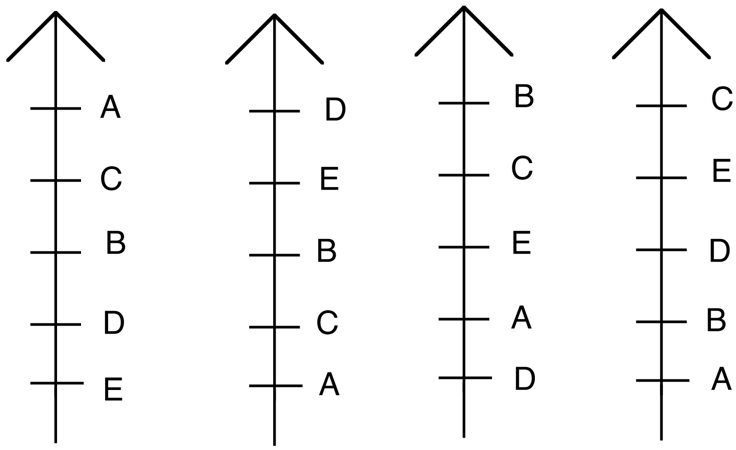
Veto voting systems

In order to protect the rights of minorities in recent years some new approaches to voting have been explored. One of these involves voting by a pattern of vetoes.

In this conception we are given a collection of m alternatives and a collection of n voters and sequence S of length (m-1) whose entries are the names of the voters. The method proceeds by looking at the name of the voter first listed in S and allowing that voter to veto or eliminate any candidate (not already eliminated) from consideration. Now one goes on to the next entry in S and that voter vetoes another alternative. This procedure continues until the last entry in the sequence S, at which point all but one of the candidates (alternatives) has been eliminated.

Here is an example to illustrate the system:

Example:



votes: 1 1 1 1

If the voters are allowed to exercise a veto in the order from left to right above, then sincere behavior results in C winning, since the first voter eliminates E, the second A, and the third D. Thus, when a vote is taken by the fourth voter, the choice is between B and C and this voter prefers C.

<u>Stop and Explore</u>

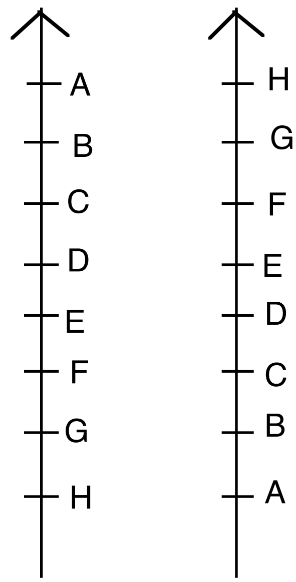
a. Determine the winner using the veto method if the order of voting is from right to left instead of from left to right, but each voter votes sincerely.

b. Who is the winner of the election when insincere voting occurs?

In illustrating this voting method we have assumed that the voters voted sincerely. What would happen if they voted in a sophisticated way? It has been shown by D. Mueller and H. Moulin that the winner using sophisticated voting can sometimes be obtained in a particularly easy manner.

In the case where there are only two players and S lists the players in alternating fashion, the sincere and sophisticated outcomes of voting by veto coincide. Furthermore, it does not matter in which order S alternatingly lists the two players; the result will be the same.

The idea of voting by veto can be adapted to try to deal with the issue of majority and minority interests in a voting system. Here is an example due to Hervé Moulin. Imagine there are two groups with the preference schedules below and votes of 60% and 40% by the groups involved on the 8 alternatives.



votes: 60% 40%

The Condorcet winner in this election would be alternative A, which is the first choice of 60 percent of the electorate but the last choice of 40 percent of the electorate. Having the majority get its way in this case may sow the seeds of tension, perhaps even revolution (to try to make the point, perhaps, too dramatically). However, if a "shared" or "proportional" point of view is adopted and the majority is given 4 vetoes and the minority 3 vetoes in a vote by veto system with alternating vetoes, then the majority will veto H, G, F, and E in that order, and the minority will veto A, B, and C in that order. The winner D may be suitably acceptable to both groups, thereby avoiding tension or revolution. In fact, D may be better for society.

When there is no Condorcet winner, yet there are two groups which constitute a majority and a minority set of positions, there is a question of political philosophy to be explored. Should one choose a social decision rule which gives the majority full power for making decisions but thereby sets up the roots of explosive confrontation, or should one choose a social decision rule which shares power with the minority in the interests of greater stability?

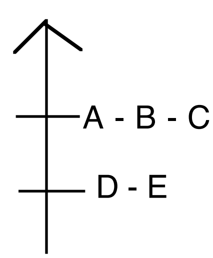
Some argue that when cyclic majorities exist, majority rule is good protection for minorities, since in the near future ever-changing ruling majorities will give them some say, while others argue that cyclic majorities strike at the very heart of social consensus and that voting methods which are stable and share power with the minority should be sought. Debates of this kind are at the very core of what goes on in democratic societies.

Cumulative voting

Cumulative voting is a method which was developed to help protect the interests of minority groups. Typically it works when there are a group of candidates who are running at large for a collection of seats in a legislature. Each voter is given a certain number of points, say 100, which can be distributed in any way among the candidates that the voter wishes. The votes for each candidate are totaled and the seats are assigned in order of their vote totals. By casting all their votes for a single candidate, a large enough group can guarantee the election of that candidate, even though the group forms a minority within the total group of voters. Cumulative voting has received considerable renewed interest as a way of achieving proportional representation, especially in light of recent Supreme Court decisions which involve redrawing district lines to achieve racial representation goals.

Approval Voting

A surprisingly simple voting system which has been discovered independently by several people recently is approval voting. In this system, voters vote for any of the candidates that they are willing to have serve. The system can be used in a situation where a single winner is to be chosen or where there are several winners to be chosen. Note that the approval voting system does not explicitly use an ordinal ballot. It is not clear whether or not when a person uses an approval ballot and votes, for example, say for A, B, and C from a list that also includes D and E, that if the voter had been presented with an ordinal ballot, the ranking that the voter would have produced would have been:

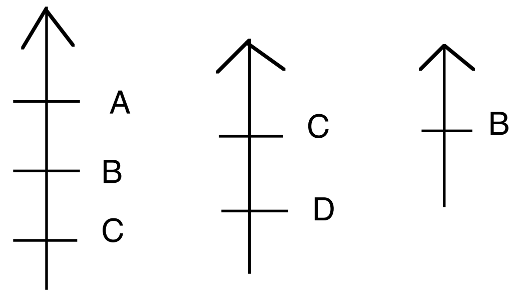


Approval voting has an appealing simplicity but some have criticized it for possibly resulting in the election of "bland" and centrist candidates, since they will receive the approval of large numbers of people. Only if approval voting is adopted in a variety of election situations can it emerge whether or not this criticism is valid.

Practical considerations in voting

Certain of the assumptions we have made here which make it easier to think systematically about elections, unfortunately, can not always be assumed to hold in practice. Consider some of the complications that can actually occur. For example, in our discussion of the ordinal ballot we have assumed that given a large collection of alternatives to rank, so that the ordinal ballot can be used properly, voters can always do this in a consistent manner. By consistent, we mean that if a voter prefers A to B, B to C, and C to D, say, the voter does not prefer D to A! Given a relationship R (e.g. equal to, parallel to, prefers to, etc.), where when a has the relationship R to b, and b has the relationship R to c, then a has the relationship R to c, we say that the relationship R is *transitive*. (Often, one writes aRb instead of a has the relationship R to b). A person would be viewed as having inconsistent views if when he/she expressed preferences, the preference relation did not exhibit transitivity. However, when the number of choices is large, people sometimes do make intransitive choices when they rank only pairs. For this reason, in designing elections with ordinal ballots, it may not be reasonable, if the number of alternatives is large, to force the voters to rank all the alternatives. On the other hand, when all the alternatives are not ranked, either because the voter chose not to or was unable to carry out such a ranking, it creates problems for some of the voting methods we have been discussing here. Consider what to do if voters produce the following ballots in an election.

Example:



votes: 12 8 9

In this election various voters have chosen not to rank all the candidates, though they had the right to do so. In terms of first place votes A would get 12, B would get 9, C would get 8, and D would get none. However, can one say whether or not a majority of voters preferred A to B or C to D? How is one to decide who would be the Borda Count winner for this election or the Condorcet winner? On the other hand voters might resent being told that they must rank all the candidates in order for their votes to be counted.

It is important to note in the election above that the 9 voters who voted for candidate B may have done so for a variety of different reasons. Some of the voters may not have known much about the other candidates and not listed them on their ballot. Some of the voters might have been voting "strategically" in the hope that due to the particular way the ballots were being counted that truncation of their ballots would help B get elected. In light of this, it may not be reasonable to assume that the 9 voters were indifferent between candidates A and C.

Another practical consideration to take into account is acceptability to the public. If a method is viewed as being "overly complex" or one that can not be easily understood, it may not win acceptance by the public, even when it is provably a better method than other methods being considered. Finally, there is the issue of computational complexity. Any method that is chosen must work, given the state of computer support for assisting in deciding the winner of the election.

What voting systems are used?

We have seen that there is a staggering variety of ideas and proposals about how elections might be conducted. Are any of these systems actually used? In fact, there are remarkably many methods that either have been used at one time or are currently being employed. Although plurality voting, sometimes in conjunction with run-off, is the most widespread procedure, other methods which have been used are indicated below:

Single-transferable vote: (Hare method):

a. Irish National Elections

b. Australian elections

c. New York City School Board elections

Approval voting

Numerous professional societies have adopted approval voting as the method for selecting officers such as president.

Cumulative voting

Cumulative voting was used in elections in Illinois for a considerable period. In the situation in Illinois each district had seats to be filled and voters could cast 3 votes. Intriguingly, some parties failed to run as many candidates as there were seats, choosing instead to run fewer. This stemmed from a concern that if candidates ran for all the seats, other parties which ran fewer candidates might have an advantage. (J. Sawyer and D. MacRae Jr., Game Theory and Cumulative Voting in Illinois: 1902-1954, Amer. Pol. Sci. Review, 56 (1962) 937f.; Blair, G., Cumulative Voting: An Effective Electoral Device in Illinois Politics, Studies in Social Sciences, U. of Illinois Press, 45 (1960)). Cumulative voting has also been used in some corporate settings.

Interestingly, the Borda Count has not been used in political elections, though it has been used in various other voting situations.

Why have these different approaches and methods to conduct elections evolved? In some cases the different systems evolved to deal with problems which arose in certain countries or localities. For example, in the United States there are two chambers to the legislature one based on population and one based on geography. This electoral system was an adaptation to the fact that initially the colonies that formed the United States had very different populations and very different economies; the needs and interests of the states that came together had to be included.

In Great Britain there are also two houses to the legislature, the House of Commons and the House of Lords. Here again the structure reflects British history. As British democracy changed from one where nobles fought against arbitrary decisions of a King (or Queen) to one where "common people" fought against the arbitrary decisions of the nobles, the House of Lords has lost power to the House of Commons. At this point in British history (since 1911 and even more since 1949) the House of Lords can delay legislation passed by the House of Commons but can not prevent it from becoming the law of the land. Not surprisingly, the ancestor of the House of Lords predates that of the House of Commons.

Given the institutions we see at any given time, we can attempt to study how effective these institutions are in meeting the equity and fairness needs of the institution at that particular time, or we can study on some absolute scale how fair and equitable a particular system is.

Summary

In this section we have tried to show the richness of using an analytical point of view to discuss the phenomenon of elections and rankings. Perhaps the most important consequence of this point of view is to show the complexity of something to which most people do not give a second thought. By airing these complexities we can try to improve the institutions in our society which are concerned with democracy and fair and equitable elections and rankings.

<u>Project</u>

a. Investigate the ranking and/or decision-making procedures used to decide the elections in your state, city, club, school, etc. Look into the procedures used for selecting the winner of various awards, grants, sports honors, etc. Can you suggest improvements in the method currently used to find a winner or produce a ranking? Can you determine unusual features of the situation you are studying that have not been considered in models of elections or rankings in the past?

References:

Arrow, K., Social Values and Individual Choice, (2nd. ed)., John Wiley, New York, 1963.

Aumann, R. and S. Hart, (eds.), Handbook of Game Theory, Volume 1, North-Holland, New York, 1992.

Aumann, R. and S. Hart, (eds.), Handbook of Game Theory, Volume 2, North-Holland, New York, 1994.

Balinski, M. and H. Young, Fair Representation, Yale U. Press, New Haven, 1982.

Brams, S. and P. Fishburn, Approval Voting, Birkhäuser-Boston, Cambridge, 1983.

McLean, I. and A. Urken, Classics of Social Choice, Michigan U. Press, Ann Arbor, 1995.

Moulin, H., Axioms of Cooperative Decision Making, Cambridge U. Press, New York, 1988.

Saari, D., Geometry of Voting, Springer-Verlag, New York, 1994.

Saari, D., Basic Geometry of Voting, Springer-Verlag, New York, 1995.

Straffin, P., Survey of Voting Methods, Birkhäuser-Boston, Boston, 1980.

Taylor, A., Mathematics and Politics, Springer-Verlag, New York, 1995.

Acknowledgements

I would like to thank Professors Hope Young and John Drobnicki of the York College (CUNY) Library for their assistance in obtaining many books and articles which helped enrich my understanding of the topics discussed in this essay.

I would also like to thank my wife, Nina Malkevitch, for reading many of the earlier drafts of this manuscript and making many suggestions for improvement of its clarity.

This work was prepared with partial support from the National Science Foundation (Grant Number: DUE 9555401) to the Long Island Consortium for Interconnected Learning (administered by SUNY at Stony Brook).