How should we count the votes?

Bruce P. Conrad

January 16, 2008

Were the Iowa caucuses undemocratic? Many politicians, pundits, and reporters thought so in the weeks leading up to the January 3, 2008 event. Polls couldn't predict the outcome, frustrating reporters. In the case of the Democrats, the ballot wasn't secret. To participate a voter had to take a few hours off on a Thursday evening. Turnout statewide—although unusually strong for this event—was only 17% of registered voters.

The critics prefer a primary election. A mathematician cannot comment on the cited weaknesses of the caucuses, but elections with more than two candidates are a thorny problem hat has been the subject of study for centuries. There is still no consensus among the experts.

To see the problem clearly, consider this example. Fifteen judges are rating figure skaters A, B, and C. Six judges say the performance of A was best, B second; five say C was best, B second; and the remaining four say B was best, C second. Who gets the gold medal, and who the silver? The **plurality** system gives first place to the competitor who gets ranked first by the greatest number of judges; second place goes to the one who gets the next to the most first ranking votes, and so on. Plurality is used to decide most elections in the United States. For the figure skaters, the plurality result is $A \succ C \succ B$.

Because no skater is first-ranked by a majority, the judges might try a **runoff** election. Since B has the fewest first-place votes, she is eliminated in the first round, and ranks last. The four judges who favored B transfer their votes to their second choice, C. But then C beats A, 9–6 for first place, resulting in the ranking $C \succ A \succ B$.

The International Skating Union (ISU) has rules to decide competitions. Skaters are ranked according to the number of skaters that a majority of the judges rank below them. Thus, A is not ranked above any skater by a majority, because 9 judges place her last. B is ranked above A by those 9 judges, and is ranked above C by 10. Thus B gets credit for being ranked ahead of two other skaters. C is ranked ahead of A by a majority of judges, but is not ranked ahead of B by a majority. The official result would therefore be: B, gold; C, silver; and A, bronze.

Depending on the method used to count the votes, any one of the skaters could win the gold medal. Thus, there is reason to worry about how a vote can find the best of more than two alternatives. Perhaps caucusing, with interaction and discussion between voters instead of a rigid algorithm, has some merit after all.

Election algorithms

An election involves with a set $\mathcal{C}$ of candidates, and a set $\mathcal{V}$ of voters. Each voter has made a **preference list** of the candidates in the order that he or she prefers. We will use the notation $U > V$ to mean that U is preferred to V on an individual preference list, and $U \succ V$ to mean that the election algorithm in question places U ahead of V. The ranking $>$ on each preference list must be transitive, complete, and without ties. That means that for two candidates Y and Z either $Y > Z$ or $Z > Y$ (but not both), and $(X > T)\&(T > Z)$ implies $X > Z$. For example, a Democrat who participated in an Iowa caucus might have preferred Obama, with second choice Clinton, third choice Edwards, fourth choice Richardson, fifth choice Biden, sixth choice Dodd, seventh choice Kucinich, and last choice Gravel. Her preference list would be $O > C > E > R > B > D > K > G$.

Let $\mathcal{P}$ be the set of all orderings of $\mathcal{C}$; that is, the set of all possible preference lists. A **voting profile** is defined to be a function $p : \mathcal{V} \longrightarrow \mathcal{P}$ that associates each voter with his or her preference list. An **election algorithm** is a rule that, when given a profile, determines the winner of the election. If we call the algorithm E, then the input would be a voting profile p and the output would be the winning candidate, or in case of a tie, a set of winning candidates. In the figure skating contest, the output of the election algorithm should be an ordering of the candidates—that is, an element of $\mathcal{P}$. In case of a tie, the output would be a subset of $\mathcal{P}$.

The most undemocratic election algorithm is **dictatorship**. A particular voter, D, is the dictator. The winner is the dictator's choice. For example, if the dictator is scoring a gymnastics contest, then for any profile p, $E(p) = p(D)$, the dictator's personal ranking. The other judges' votes are not used.

We have seen the **plurality** algorithm. Its obvious flaw is that it ignores information that may be relevant. With 5 candidates, it is possible that the winner may have only 30% of the first-place votes, and it could be that the voters who did not rank the winner first might have ranked him last.

In an effort to look below the first-place rankings, many runoff algorithms have been invented. These always involve repeated elections, eliminating at each stage one or more candidates. With "Instant runoff voting," or **IRV**, each voter enters his or her entire preference list on a ballot. The runoff elections can be conducted by the computer, so it is unnecessary to have repeated trips to the polls. Also, voters can't change their priorities at each stage. With IRV, as soon as a candidate receives a majority of the first place votes, then he or she wins. Until a winner is chosen, the candidate with the fewest first-place votes at any stage is eliminated, and all voters who ranked that candidate first would have their votes transferred to their next choice.

The runoff algorithms are not **monotone**: in other words a voter can cause a winning candidate to lose by changing her preference list in that candidate's favor. Here is a profile with that property:

$$\begin{array}{lcl} 6 \text{ voters} & \mapsto & X > Y > Z \\ 5 \text{ voters} & \mapsto & Y > X > Z \\ 6 \text{ voters} & \mapsto & Z > Y > X \end{array}$$

With this profile, X wins an IRV election. But suppose two of the voters with preference $Z > Y > X$ switch to $X > Z > Y$. They thought they were helping X, but they cause Y to win, and X to lose!

Ramon Llull (1232–1315) wrote the first known study of multicandidate elections. In the 1780's, elections became a topic of interest in the French Academy of Sciences. One academician, Jean Charles de Borda (1733–1799) who was an applied mathematician and naval officer best known for his contributions to fluid mechanics, developed an algorithm that is now known as the **Borda count**. For a candidate X, find, for each opponent Y_i the number n_i of voters who prefer X to Y_i. The Borda count for X is then the sum of the numbers n_i. The algorithm then ranks the candidates in decreasing order of Borda count. For the figure skating competition example, A has two opponents below her on 6 preference lists: her Borda count is 12. B has one opponent below her on 11 preference lists, and two below her on 4 preference lists, so her Borda count is $11 + 2 \times 4 = 19$. Similarly, C has a Borda count of 14. Thus the skaters would be ranked $B \succ C \succ A$, which agrees with the ISU result.

The marquis de Condorcet (1743–1794) is another French academician who is remembered for his contributions to the theory of elections. A mathematician specializing in probability theory, Condorcet objected to Borda's algorithm and promoted his own. We will say that a candidate C who is preferred by a majority of the voters to every other candidate Y_i is a

Condorcet winner. For example, in the figure skating example, B is a Condorcet winner. According to the ISU rules, a Condorcet winner will win the gold medal, because he or she will be ranked higher than any single opponent in a head to head comparison by a majority of the judges.

Condorcet's criticism of Borda's algorithm was that it did not guarantee victory for the Condorcet winner. In the following profile, the Condorcet ranking is $P \succ Q \succ R$ while the Borda ranking is $Q \succ P \succ R$.

$$\begin{array}{lcl} 6 \text{ voters} & \mapsto & P > Q > R \\ 3 \text{ voters} & \mapsto & Q > P > R \\ 2 \text{ voters} & \mapsto & Q > R > P \end{array}$$

Condorcet winners can and sometimes doalso lose elections decided by plurality or runoffs. The figure skating example shows this, and here is a real-world example: In the 2007 French presidential election, the top candidates in the first round were Nicolas Sarcozy, François Bayrou, and Ségolène Royal. Bayrou was eliminated in the first round, and Sarcozy went on to win the presidency in the runoff by a margin of 53–47%. Pre-election polling indicated that Bayrou could defeat Royal, 58–42%, and that he could defeat Sarcozy, 55–45%, in head-to-head contests. If these polls are to be believed (and of course polling is notoriously unreliable), Bayrou was the Condorcet winner. The polling data can also be used to estimate the Borda count. The table below shows percentage of the electorate that preferred the candidate of a particular row over an opponent in the column. The column on the right is a sum of the percentages in the rows, and is the Borda count, expressed as a percentage of the number of voters.

	Bayrou	Royal	Sarcozy	Borda count
Bayrou	———	58%	55%	113%
Royal	42%	———	47%	89%
Sarcozy	45%	53%	———	98%

Based on polling data, it seems that Bayrou was also the Borda count winner.

Condorcet's idea has a complication—which he addressed. It can happen that there is no Condorcet winner at all. Here is a profile that demonstrates this.

$$\begin{array}{lcl} 5 \text{ voters} & \mapsto & U > V > W \\ 4 \text{ voters} & \mapsto & V > W > U \\ 2 \text{ voters} & \mapsto & V > U > W \\ 6 \text{ voters} & \mapsto & W > U > V \end{array}$$

When we match two candidates, the one who loses is not a Condorcet winner, because to qualify one must win all matches. In this case, you can see that $U \succ V$ (this means a majority prefers U to V), so V is a Condorcet non-winner. Also $V \succ W$ by an 11–6 margin. However, if we match U against W, thinking that U will win, we are in for a surprise, because $W \succ U$, 10–7. We have discovered a **Condorcet cycle**, $U \succ V \succ W \succ U$. These Condorcet cycles are the source of much of the confusion associated with multicandidate elections.

If a Condorcet cycle occurs in the judges' ranking for a figure skating competition, the skaters involved in the cycle would be tied. The ISU rules specify that when this happens, the Borda count should be used to break the tie. In the $\{U, V, W\}$ profile, the Borda counts are U, 18; V, 17; and W, 16. The ISU ranking is therefore $U \succ V \succ W$.

Have the skaters figured it out? Should we just elect politicians by the ISU rules? Before we close the book on this subject, let's consider two additional criteria that seem reasonable for an election algorithm to satisfy.

The Theorem of Arrow

Here is an unfair algorithm, called the **agenda algorithm**, because it uses an ordered list of the candidates, called the agenda.

If the agenda is $C_1, \ldots, C_n$, then the algorithm matches C_1 against C_2, the winner of that match takes on C_3, and so on. When the end of the agenda is reached there will be one survivor, who is declared the winner. This procedure will always select a Condorcet winner if there is one, since he or she is the only candidate who can't been eliminated.

When the profile has no Condorcet winner, it is advantageous to be last on the agenda—then you only have to win one contest. This makes the agenda algorithm unfair, but it has another problem, because it fails a test that was devised by Vilfredo Pareto (1855–1923): *If the voters unanimously prefer candidate C_1 to candidate C_2 and C_1 is not a winner, then C_2 must not win, either.*

Here is an example that shows the agenda algorithm fails Pareto's test. Three princes, A, B and C, are holding an election to decide who should be monarch. Prince B announces that his dog, D, is also a candidate. These are the preference lists.

$$\begin{array}{ll} \text{Prince } A \mapsto & A > B > D > C \\ \text{Prince } B \mapsto & B > D > C > A \\ \text{Prince } C \mapsto & C > A > B > D \end{array}$$

Notice that each prince prefers B to D, and that with the agenda A, B, C, D, the dog wins. The dog's advantage was being last on the agenda. Any one of the princes could also be the winner—just rearrange the agenda. This algorithm is only used by legislative committees, and the ability to set the agenda is one of the perquisites of the chairperson. The other algorithms that we have discussed, even dictatorship, pass the Pareto test.

Independence from irrelevant alternatives (IIA). A voter who submits a preference list that does not reflect his or her true priorities is said to be voting insincerely. An algorithm that does not have the following property, called IIA, will offer opportunities to vote insincerely. *Given that A is a winner, and B an non-winner then B cannot become a winner and A a non-winner as a result of changes in one or more voters' preference lists, as long as no voter reverses the relative positions of A and B.*

Suppose that in the first round of the 2007 election in France, those who supported Royal had preference lists Royal > Bayrou > Sarcozy. If these voters voted insincerely by moving Royal to second place and voting for Bayrou, then Bayrou would have won. No one would have switched their preference between Bayrou and Sarcozy, so the runoff method is not IIA.

The plurality algorithm is also not IIA. Suppose an election involving a Democrat D, a Republican R and a third-party candidate T, with the following profile:

$$\begin{array}{rcl} 23\% \text{ of voters} & \mapsto & D > T > R \\ 20\% \text{ of voters} & \mapsto & D > R > T \\ 40\% \text{ of voters} & \mapsto & R > D > T \\ 5\% \text{ of voters} & \mapsto & R > T > D \\ 12\% \text{ of voters} & \mapsto & T > D > R \end{array}$$

If everyone votes sincerely, the Republican will win with 45% of the vote. The third-party voters realize the contest between the Democrat and Republican is close, and their candidate has no chance. If more than one-sixth of them vote insincerely for the Democrat, they will cause the Democrat to win, and the Republican to lose. No one has changed their relative preference between R and D, so plurality fails the IIA test.

In work that reawakened interest in the election decision problem, the economist Kenneth Arrow proved the following theorem. For a short proof (actually three of them), see [3].

Theorem (Arrow, 1951)
There is no election algorithm—other than dictatorship—that is applicable

to elections with more than two candidates and satisfies the Pareto and IIA conditions.

Arrow's theorem raises questions abut using any election algorithm based on voter preference lists.

Range voting is an alternative to profile-based election algorithms. Let N be a positive integer; usually 9 or 99 is used. A **range ballot** is a list in which a voter has marked each candidate's name with a number from 0 to N to indicate a preference or with a blank to indicate no opinion. The numbers don't have to be distinct.

The winner is the candidate who has the greatest average score. In computing the average, blank votes are ignored. Thus a 0 is a vote against a candidate because it affects the candidate's average, but a blank vote is just an abstention. Because a candidate who is completely unknown might get a high score on one ballot (his own) and blank on everyone else's, there is a quorum rule: The *total* score of the winner must be at least half of the greatest total score received by any candidate.

Range voting is relatively untested in general elections, and opponents of this method caution that there would be insincere voting, in which most voters would rate their most favored candidates with N and rate all others 0. If this behavior were prevalent, range voting would produce a result identical to the plurality algorithm. Proponents of range voting point out that the supporters of a third party candidate would not have to forsake their favorite: they could rate her with N and give the major candidate whom they find least offensive a lower score. There is an organization that promotes range voting, rangevoting.org.

Approval voting is an alternative that is simpler than range voting. Each voter is asked to check the names of the candidates whom he or she approves on the ballot. A voter may approve as many candidates as he or she thinks is appropriate. The candidate with the largest number of approvals wins. Approval voting is not the same as range voting with $N = 1$, because there is no way to indicate abstention, as one can in range voting. The principal objection to approval voting is that voters will approve their favorite candidates, and not approve any other candidates. The winner would be the plurality winner. However, a quorum condition could address this problem by requiring that the winner must have the approval of more than half of the voters. If a quorum is not achieved by any candidate, the election would be repeated. S. Brams and P. Fishburn have written a book [1] advocating approval voting.

By avoiding voter preference lists, range and approval voting offer voters the chance to indicate which candidates are acceptable. A preference list does not have that information, since it doesn't say where the voter would draw a line. Nevertheless, it would be unscientific to predict that either range or approval voting will solve our electoral problems without doing some experiments first.

Among election algorithms based on profiles, the favorites among mathematicians are the Borda count and an algorithm originally described, albeit not very clearly, by Borda's rival, the marquis. Perhaps the strangest part of this story is that the controversy between Condorcet and Borda is still unsettled after more than 200 years. Meanwhile, the most prevalent election algorithms, plurality and various sorts of runoff methods, although known to be inferior, are still the most widely used.

The Borda count belongs to a class of election algorithms called **positional** algorithms, which are constructed as follows. Suppose that there are c candidates, and let $\vec{s} = (s_1, s_2, \ldots, s_c) \in \mathbf{R}^c$ be a fixed vector such that $s_1 \geq s_2 \geq \cdots \geq s_c \geq 0$. We call $\vec{s}$ the **scoring vector**. For each candidate C_i the score would be the sum $s_1 \times$ the number of voters ranking C_i first, plus $s_2 \times$ the number of voters ranking C_i second, and so on, until $s_c \times$ the number of voters ranking C_i last. Candidates are ranked by score in descending order. The scoring vector $\vec{s} = (1, 0, 0, \ldots, 0)$ gives C_i the number of voters who ranked him first as score, and in this case the ranking is the same as the plurality method. If $\vec{s} = \vec{b_c} = (c-1, c-2, \ldots, 0)$ a candidate's score is his or her Borda count.

D. Saari showed in [4] that for most scoring vectors, positional algorithms can exhibit very strange behavior. Unless the scoring vector is carefully chosen, it is possible to specify orderings for all of the subsets of $\mathcal{C}$, and then find a voting profile such that the ranking of each subset is the specified ordering. For example, suppose that $\mathcal{C} = \{A, B, C, D\}$. We could specify, for each of the 11 subsets of $\mathcal{C}$ that have more than one element, a randomly chosen ordering, something like $A \succ D \succ B \succ C$, $C \succ B \succ D$, $D \succ B \succ A$, $C \succ A$, etc. Then there exists a voting profile that would give the first ordering; yet if A dropped out, the second ordering would be the result; if C dropped out, the result would the the third ordering; and in A v. C head-to-head, C would win. In particular this applies to the plurality algorithm. Saari has written a book [4], aptly titled *Chaotic elections.*

Saari found that the Borda count scoring vector $\vec{b_c}$ was an exception. His theorem applies to "most" scoring vectors, but not to $\vec{b_c}$. Although still subject to Arrow's theorem, the Borda count is a relatively well-behaved election algorithm.

We have seen an example where the majority of the voters had the same preference list, $P > Q > R$; yet the Borda ranking was $Q \succ P \succ R$. This brings to mind Condorcet's algorithm, which, with eight other algorithms that satisfy the Condorcet condition, are reviewed in [2].

If there are n candidates, then there are $n(n-1)/2$ pairs of candidates, and thus $2^{n(n-1)/2}$ possible sets of pairwise orderings, or **opinions**. An opinion, $C_i \succ C_j$, is a statement that in a majority of the preference lists, voters preferred C_i to C_j. The **margin** of the opinion $C_i \succ C_j$ is defined to be the number of voters who had $C_i > C_j$ on their preference lists, minus the number who had $C_j > C_i$. For $n = 3$ we have 3 pairings and 8 possible sets of opinions. Since the number of total rankings is $3! = 6$, two of the sets of opinions cannot be consistent with a total ordering: these are Condorcet cycles.

Condorcet's assumption is that the purpose of voting is to find the true ranking of the candidates. If there is a Condorcet cycle, some voters must have erred, and the correct ranking is the one that can be achieved by changing the fewest voter preferences. If a Condorcet cycle occurs in a three-candidate election, delete the opinion with the least margin. If we apply this method to the cycle involving the candidates U, V, and W, (see above) the following list occurs (the margin appears above each relation).

$$U \overset{5}{\succ} V \overset{9}{\succ} W \overset{3}{\succ} U$$

and hence the Condorcet ranking is the same as the Borda ranking, $U \succ V \succ W$. Notice that in deleting the opinion $W \succ U$ we have in effect reversed it, because it is now implied by our result, $U \succ V \succ W$.

Condorcet's algorithm requires a modification if there are four or more candidates. P. Young [8] suggested that some of the opinions should be *reversed* to eliminate cycles and produce a complete and transitive ranking of the candidates. The sum of the margins of the reversed opinions must be as small as possible. This algorithm has also been described by John Kemeny, and is listed under his name in [2]. This Condorcet-Kemeny (CK) algorithm gives the total ordering of the candidates that is "closest" to the profile, in the sense that it can be made with the fewest reversals of priority on voter preference lists. It does, however, involve some vote changing.

In [7], P. Young proved that as a ranking algorithm, CK is the only one satisfying Condorcet's criterion with the following additional properties: It treats candidates equally and it treats voters equally; it passes Pareto's test; it is **consistent** in the sense that if the voters are divided into two disjoint subsets and both give the same ranking by this algorithm, then the combined set of voters will also give this ranking; and it is **locally IIA** (LIIA). Suppose that an election algorithm has obtained a ranking of candidates. An **interval** in that ranking is a subset $\mathcal{I} \subset \mathcal{C}$ such that if $C_1, C_3 \in \mathcal{I}$ and $C_1 \succ C_2 \succ C_3$ then $C_2 \in \mathcal{I}$. The algorithm is LIIA if the ranking within any interval will remain unchanged if candidates who are not in the interval drop out. This is not the case with the Borda count, and Young points out that that this makes the Borda algorithm vulnerable to manipulation, by the introduction of low-ranked candidates on the ballot.

The CK algorithm has been compared with the Borda count by Saari and Merlin and in [6], using geometric methods. They note that Borda counts ignore Condorcet cycles, and of course the Borda ranking is achieved without switching any voter preferences. Their analysis includes a geometric explanation of the differences between the algorithms. Here is a sample theorem from [6], which makes a precise statement of how much the Borda count and CK rankings can differ. *Suppose that two rankings of $\mathcal{C}$ are given. Then there is a voting profile such that the first ranking is by the Borda count, and the second is the CK ranking, if and only if the first and last candidates in one ranking are not reversed in order in the other.*

What should we do?

There are points in favor of deciding elections by Approval or Range voting, or by the Borda or CK algorithms. We have little experience with any of these, but all have advantages over the methods that are now widely used. In the political arena, it will be very difficult to implement any of these methods, simply because politicians who don't want to be faced with a new system. Another hurdle that applies to all except approval voting is the added complexity of the ballot. If I were dictator for a moment, I would require some experiments to see what works. Probably we can't choose a better election algorithm by voting—there are too many candidates.

References

[1] Brams, Steven and Peter C. Fishburn, *Approval voting,* second edition (2007), Springer-Verlag New York.

[2] Fishburn, Peter C., "Condorcet social choice functions," *SIAM Journal on Applied Mathematics* **33** (1977) pp 469–489.

[3] Geanakoplos, John, “Three Brief Proofs of Arrow’s Impossibility Theorem” (June 2001). Yale Cowles Foundation Discussion Paper No. 1123RRR. Available at SSRN: http://ssrn.com/abstract=275510

[4] Saari, Donald G., “A dictionary for voting paradoxes,” *Journal of Economic Theory* **48** (1989) pp 443–475.

[5] ——— *Chaotic Elections! A Mathematician Looks at Voting.* American Math. Society, 2001.

[6] ——— and Vincent R. Merlin, “A geometric examination of Kemeny’s rule,” *Soc Choice Welfare* **17** (2000) pp 403–438.

[7] Young, Peyton, “Optimal voting rules,” *The Journal of Economic Perspectives*, **9** (1995), pp. 51-64.

[8] ———, “Condorcet’s theory of voting” *The American Political Science Review*, **82** (1988), pp. 1231-1244.