Descriptions of single-winner voting systems

Warren D. Smith*
WDSmith@fastmail.fm

July 12, 2006

Abstract — **We describe over 40 single-winner voting systems, and elucidate their most important properties.**

Contents

Previous surveys of voting systems may be found in [6] [7] [16] [18] [23] [39] [43] [47] [45] [46] [55] [57] [58] [62] [65] [68] [75]. Although the bulk of our content is not original, we have several new voting systems, new theorems about properties, and new examples and remarks, and the first-ever discussion of which pairwise totals and margins are attainable; and nobody has previously collected all this material in one place before.

*21 Shore Oaks Drive, Stony Brook NY 11790.

1 Notation

In all descriptions we shall assume there are N candidates and V voters, and the goal is to select a single winning candidate and to do so in $\le$ polynomial(N, V) steps. (NP-hard [5] election algorithms will not be considered here.[1]) Unless otherwise stated, all ties are broken randomly. Schulze [66] suggested breaking ties by choosing a random ballot and breaking the tie in the manner that voter wanted (which, if votes are full preference orderings, always suffices).

In all the voting systems we list, each "vote" is one of the following.

§4 Ignored. (Example: random winner.)
§5 The name of a single candidate. (Examples: "plurality" and "bullet" voting.)
§6 A "preference ordering" among the candidates, that is, a permutation of the set $\{1, 2, 3, \dots, N\}$. (Examples: Borda, Condorcet, Dodgson, IRV, etc.)
§7 A real N-vector, that is, N real numbers. (Sometimes these real numbers x are required to satisfy some condition, e.g. in "range voting" we demand $0 \le x \le 1$ and in "approval voting" we demand $x \in \{0, 1\}$.)
§8 In "sarvo-range voting" each vote is *both* a real N-vector of numbers x satisfying $0 \le x \le 1$ *and* one additional bit (the "strategy bit").

Anybody trying to employ one of these systems in a real election would have to resolve many nasty "real-world" details (such as what to do if voters refuse to rank all the candidates in Condorcet's system) which we ignore here – we merely aim to describe the simplest possible variant of each system.[2]

In the systems based on real N-vectors, we shall often make use of the sum-vector $\vec{s}$.

In the systems based on preference orderings, if all candidates are ignored, except for two (A and B) in each preference ordering, then the result would be a 2-candidate election between A and B, with some winning margin

$$M_{AB} \;=\; \#\text{votes for } A - \#\text{votes for } B \tag{1}$$

which would be positive if A wins and negative if B wins. We shall often make use of this $N \times N$ antisymmetric matrix M. Another useful $N \times N$ matrix is U, where U_{kj} is the number of voters who prefer k to j. Then $M = U - U^T$.

If $M_{WA} > 0$ for *all* $A \ne W$ then W is a "Condorcet-Winner." (A better name might be "beats-all-winner.") Condorcet-Winners need not exist (either because of a tie, or because there can be a "preference cycle"), but if one does exist, it plainly is unique.

#voters	their vote
4	$A > B > C$
3	$B > C > A$
2	$C > A > B$

Figure 1.1. 9-voter example of a Condorcet cycle: for each candidate, some other is preferred by a majority. ▲

We single out the following "prototypical" systems for special attention: Plurality, Borda, Condorcet-Least-Reversal, IRV, Clarke-Tideman-Tullock, and Range.

2 Which vote-count and margins matrices are possible?

This problem is fundamental to voting theory, and also had not previously been posed and solved, so we do that here. Readers who want to skip the math and plunge directly into descriptions of voting systems should skip to section 4.

In a V-voter election, the vote-count matrix U obeys $U_{AB} + U_{BA} = V$ for $A \ne B$ (since we have required full rank orderings as votes, i.e. with no omissions allowed), $U_{AB} \ge 0$, and $U_{AA} = 0$.

But not every matrix obeying these conditions is actually achieveable in an election. Only matrices that arise as weighted sums of actual single-vote-representing matrices are achievable, where the "weights" are the number of voters who cast that vote. In other words, *a U-matrix obeying $U_{AB} + U_{BA} = V$ is achievable if and only if the following integer program has a solution:*

$$\underbrace{\sum_{\pi \in S_N} w_\pi Q_\pi = U,}_{\binom{N}{2}\text{ equalities}} \qquad \underbrace{w_\pi \ge 0}_{N!\text{ inequalities}} \quad . \tag{2}$$

Here there are "$\binom{N}{2}$ equalities" since only the upper triangle of U matters and it contains $\binom{N}{2}$ entries (the lower triangle is determined automatically by antisymmetry). Here Q_π is the single-vote-representing matrix got by permuting both the rows and the columns of the upper-triangular $N \times N$ matrix with all-1's above the diagonal and 0's elsewhere, by the N-permutation $\pi \in S_N$ (and S_N denotes the group of $N!$

[1] H.P.Young in 1975 suggested minimizing (over choice of W) the number of voters whose entire preference-order votes have to be ignored, in order to cause W to be the Condorcet-Winner. We point out that this is this is NP-hard by an easy reduction from 3-dimensional matching [29]. Kemeny suggested finding the rank-ordering of the candidates which was "closest" to the rank-orderings provided by the voters in the sense that it had the minimum number of pairwise reversals. This too is NP-hard (it is essentially a traveling salesman problem). J.Rothe and H.Spakowski have indeed claimed to have shown that it is $P_{\|}^{NP}$-complete to determine Young and Kemeny winners. Slater's (1961) version of Condorcet voting, in which one first finds an *ordering* of the candidates compatible with the M-matrix – or if no ordering is compatible with it, we find the one compatible with an altered version of M requiring the minimum number of sign reversals of its elements – and a related version in which we minimize the *sum* of the absolute values of the elements in M_{AB} whose sign is incompatible with the ordering, both are NP-hard. (See also footnote 4.) The "Max-Tourneys" ranked-ballot voting method is probably also NP-hard: Consider all of the meaningfully different ways of seeding a (fixed balanced binary tree) single elimination tournament among the N candidates. The candidate that wins the most such tournaments (counting pairwise ties in the logical manner, e.g. half a win for a 2-way tie) wins. ‘

[2] In preference-order-based systems, the emerging consensus seems to be that unranked candidates should be ranked co-equally and beneath all ranked candidates. Usually when some algorithm designed for the full-ranking case compares two "victory margins" M_{AB}, it is a good idea to replace that comparison by a lexicographic comparison of 2-tuples (U_{AB}, M_{AB}).

permutations of N items):[3]

$$Q_\pi \stackrel{\text{def}}{=} P_\pi \begin{pmatrix} 0 & 1 & 1 & \cdots & 1 & 1 \\ 0 & 0 & 1 & 1 & \cdots & 1 \\ 0 & 0 & 0 & 1 & \cdots & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & 0 & \cdots & 0 & 0 \end{pmatrix} P_\pi^T \tag{3}$$

and w_π are the (non-negative integer) weights. In other words, w_π is the number of voters with vote π and Q_π is the U-matrix resulting from the single vote π. This is a linear program with $N!$ non-negative integer *variables* w_π obeying $\binom{N}{2}$ *equality constraints*. For many purposes (i.e. if you do not care about an overall re-scaling) we may ignore the integrality constraints, thus instead getting a *rational* solution, since then a genuine integer solution – albeit to a *rescaled problem* with more voters – may always be got by multiplying everything by the least common denominator.

Linear programs are usually phrased as an optimization problem, rather than merely an existence problem. We can do that too by asking for the w_π's which maximize

$$\sum_{\pi \in S_N} w_\pi c_\pi \tag{4}$$

for any particular $N!$ constants c_π we desire. For example if we choose the c_π's to form a sufficiently rapidly exponentially decreasing positive sequence then this will find the lexicographically minimum vote-set which achieves our U-matrix. The given U-matrix is achieveable (up to rescaling) if and only if this linear program has a solution (which, if it exists, will necessarily be bounded). For the sole purpose of deciding whether suitable w_π *exist*, however, optimization is irrelevant, and we may regain that irrelevancy by setting $c_\pi \equiv 0$.

The *dual* form [20] of this linear program has $\binom{N}{2}$ variables Y_{AB} (of unrestricted sign) for $1 \le A < B \le N$, now obeying $N!$ *in*equality constraints:

$$\sum_{1 \le A < B \le N} Y_{AB} \, (Q_\pi)_{AB} \ \ge \ c_\pi \tag{5}$$

and with our above formulation of the primal problem as a *max*imization problem, the dual problem is to *min*imize

$$\sum_{1 \le A < B \le N} Y_{AB} U_{AB} \tag{6}$$

by choice of the Y_{AB}'s. It follows from the duality theorem of linear programming [20] that our given U-matrix is achieveable (up to rescaling) if and only if this dual linear program has a *finite* optimal solution (no matter what the c_π's are), and is unachieveable if and only if this dual problem has an *unbounded* optimal solution.

How difficult are the primal and the dual problems? We know the primal feasibility problem is in NP because any point inside a convex d-polytope is a convex combination of at most $d+1$ of the vertices. Therefore in the primal LP, U may be taken to be a convex combination of at most $\binom{N}{2}+1$ of the Q_π, and we may specify the solution by specifying the $\binom{N}{2}+1$ permutations π and positive weights w_π. We suspect that the problem is in fact NP-complete when N is allowed to become large, because (1) the problem of detecting a violated constraint in the dual problem is NP-complete (cf. footnote 1 re Slater voting), and (2) the "optimal digraph ordering" problem (find an ordering of the vertices of a directed graph, which minimizes the number of arcs, or more generally the weight-sum of arcs, that point backwards)[4] is NP-complete [61][35][53][29].

Three examples: The reader may enjoy confirming that the following U-matrix

$$\begin{pmatrix} * & 1 & 14 \\ 17 & * & 2 \\ 4 & 16 & * \end{pmatrix} \tag{7}$$

is unachieveable in a 3-candidate elections with 18 voters, but the U-matrix in EQ 26 is achieveable and the M-matrix in EQ 21 may easily be chosen to make it achieveable. ▲

One may similarly write the corresponding difficult integer linear program that defines the achievable victory-*margins* matrices M (obeying $M_{AB} + M_{BA} = 0$):

$$\underbrace{M = \sum_{\pi \in S_N} w_\pi M_\pi,}_{\binom{N}{2} \text{ equalities}} \quad \underbrace{w_\pi \ge 0}_{N! \text{ inequalities}} \quad . \tag{8}$$

Here M_π is the matrix got by permuting both the rows and the columns of the anti-symmetric $N \times N$ matrix with all $+1$'s above the diagonal and all -1's below it (and 0s on the diagonal) by the N-permutation π:

$$M_\pi \stackrel{\text{def}}{=} P_\pi \begin{pmatrix} 0 & +1 & +1 & \cdots & +1 & +1 \\ -1 & 0 & +1 & +1 & \cdots & +1 \\ -1 & -1 & 0 & +1 & \cdots & +1 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ -1 & -1 & \cdots & -1 & 0 & +1 \\ -1 & -1 & \cdots & -1 & -1 & 0 \end{pmatrix} P_\pi^T \tag{9}$$

and w_π are non-negative integer weights. This again is a linear program with $N!$ non-negative integer variables w_π obeying $\binom{N}{2}$ equality constraints. Its dual form is

$$\underbrace{\sum_{1 \le A < B \le N} Y_{AB} \, (M_\pi)_{AB} \ \ge \ c_\pi}_{N! \text{ inequalities}} \tag{10}$$

[3]Here P_π denotes the permutation matrix corresponding to the permutation π. A permutation matrix is a matrix all of whose entries are 0 except for exactly one nonzero entry in each row and in each column; all of these nonzero entries being 1. The effect of multiplying an N-vector by an $N \times N$ is to permute its entries, hence the name. The inverse P^{-1} of a permutation matrix P is the same thing as its transpose P^T.

[4] This has also been called the "linear ordering problem." It also is equivalent both to the problem of finding the maximum-edge-cardinality acyclic subgraph of a digraph, and to the "minimum feedback arc set" problem of finding the minimum set of arcs whose deletion from a digraph leaves it acyclic, and to this problem: given an 0-1 (or more generally integer) square matrix M, find a permutation of its rows and columns (using the same permutation for both) which maximizes the sum of the entries in its upper triangle.

Our given M-matrix is achieveable (up to scaling) if and only if the primal problem has a real solution, which happens if and only if the dual problem has a bounded optimal solution.

3 Some combinatorics: the number of kinds of votes

If permissible "votes" are rank-orderings of N candidates, then there are $N!$ possible votes: $0! = 1! = 1$, $2! = 2$, $3! = 6$, $4! = 24$, $5! = 120$, etc. The factorial function obeys the recurrence $N! = (N-1)! \cdot N$ and arises from the generating function $\sum_{N\ge 0} x^N/N! = e^x$.

If the "votes" instead are rank-orderings of N candidates *with equalities allowed* (e.g. a possible vote is "$A > B = C = D > E = F > G$"), then there are $E(N)$ possible votes where $E(0) = E(1) = 1$, $E(2) = 3$, $E(3) = 13$, $E(4) = 75$, $E(5) = 541$, and we may calculate other $E(N)$ from

$$E(N) = \sum_{k\ge 1} \frac{(k-1)^N}{2^k} = \left(x\frac{\mathrm{d}}{\mathrm{d}x}\right)^n \left(\frac{1}{2-x}\right)_{x=1} \tag{11}$$

$$= \sum_{k=0}^{N} \binom{N}{k} E(N-k) \sim \frac{N!}{2\ln(2)^{N+1}}$$

or the exponential generating function (EGF)

$$\sum_{N\ge 0} E(N)\frac{x^N}{N!} = \frac{1}{2-\exp(x)}. \tag{12}$$

Amazingly, $E(N)$ is equal to the nearest integer to $N!/(2\ln(2)^{N+1})$ for $N = 0, 1, \dots, 16$ (and we have asymptoticity as $N \to \infty$) *but* equality fails when $17 \le N$:

$$E(17) = 130370767029135901 \ne 130370767029135900. \tag{13}$$

If "votes" are rank-orderings of (≥ 2)-element *subsets*[5] of the N candidates (with the remaining candidates left unranked) then the number $S(N)$ of possible votes obeys $S(2) = 2$, $S(3) = 12$, $S(4) = 60$, $S(5) = 320$, and

$$S(N) = \sum_{k=0}^{N-2} \frac{N!}{k!} = \lfloor N!e - N - 1 - \frac{1}{N+1} \rfloor \sim N!e \tag{14}$$

where $e \approx 2.71828$.

One could also consider combining both of the preceding generalizations of "$N!$" by permitting a vote to be a ranking, containing optional equalities, of any (≥ 0)-element subset of the candidates. The number of possible votes then would be $\sum_{k=0}^{N} \binom{N}{k} E(k)$. The EGF for this quantity is

$$\frac{\exp(x)}{2-\exp(x)} = 1 + \frac{2}{1!}x + \frac{6}{2!}x^2 + \frac{26}{3!}x^3 + \frac{150}{4!}x^4 + \frac{1082}{5!}x^5 + \dots \tag{15}$$

The Nth term of this sequence, amazingly, is the nearest integer to $N!/\ln(2)^(N+1)$ for $N \le 15$ and this also is asymptotically correct as $N \to \infty$, but the equality *fails* for $N = 16$.

Another stunning formula for the number of N-candidate ranked ballots with truncation and ranking-equalities both permitted is $\sum_{k\ge 1} k^N/2^k$.

If a vote is a *subset* of the N candidates, then the number of possible votes (i.e. subsets) is 2^N. If a vote is an N-letter word from an L-letter alphabet, then the number of possible votes is L^N.

If a vote is a *partial ordering* of the N candidates (example: "$A > B$, $B > C$, $B > D$"; with all implications of transitivity, such as $A > C$, being assumed), then [9] the number $P(N)$ of inequivalent possible votes obeys $P(0) = P(1) = 1$, $P(2) = 3$, $P(3) = 19$, $P(4) = 219$, $P(5) = 4231$, $P(6) = 130023$, $P(7) = 6129859$, $P(8) = 431723379$, $P(9) = 44511042511$. No closed formula for $P(N)$ is known. However, a remarkable asymptotic formula is known:[6]

$$P(N) = \left[1 + O(\frac{1}{N})\right] \exp\left(\frac{\ln 2}{4}N^2 + \frac{3\ln 2}{2}N\right) \left(\frac{2\sqrt{2}}{\pi N}\right)^{1/2} C_{N \bmod 2} \tag{16}$$

where

$$C_m = \sum_{k=-\infty}^{\infty} 2^{-(m/2+k)^2} \approx \begin{cases} 2.1289312505 & \text{if } m = 0 \\ 2.1289368272 & \text{if } m = 1 \end{cases}. \tag{17}$$

Finally, one could allow votes to be partial orders *with equalities allowed.* (Example: "$A > B = C$, $D > E$, $D > F$.") The number $Q(N)$ of inequivalent votes[7] is then

$$Q(N) = \sum_{k=1}^{N} S(N,k)P(k) \sim P(N) \tag{18}$$

where $S(N,k)$ is the Stirling number of the second kind [34] generated by

$$\sum_{n\ge k} S(n,k)\frac{x^n}{n!} = \frac{1}{k!}(e^x - 1)^k. \tag{19}$$

The $Q(N)$ sequence begins

$$1, 4, 29, 355, 6942, 209527, 9535241, 642779354, 63260289423.$$

4 Systems that ignore the votes

4.1 Random winner

▶ Ignore the voters and just select a winner from the candidates at random (all equally likely).

4.2 Optimum winner

▶ You magically read the mind of each voter and determine exactly how much that voter would benefit from the election of candidate n (for each n and each voter). (How to measure "how much"? Well, one could use "dollars," but the right

[5] "≥ 2" since we are only interested in nontrivial votes.

[6] This formula is the combined result of work by D.J.Kleitman, B.L.Rothschild, and J.L.Davison; see [21]. Its agreement with the known exact $P(N)$ counts when $N \le 16$, is poor. $P(N)$ is the number of N-element "labeled posets," or N-node "labeled (acyclic) transitive digraphs."

[7] This count has also been called the number of inequivalent "quasi-orders" of N items, or the number of "finite topologies" with n labeled elements.

utility-units really would be more like "total lifetime happiness." Benefits can be negative or positive depending on the candidate and the voter.) The candidate maximizing the total sum (over all voters) of benefit, wins. This would be the best possible voting system, if it were achievable.

4.3 Worst winner

► Same as optimum winner, except the candidate *min*imizing the benefit-sum is the winner. This would be the worst possible voting system, if it were achievable.

The importance of both this pessimal and the preceding optimal system is that they *are* achieveable inside computer simulations of elections, i.e. involving artificial voters.

5 Systems in which each vote is the name of a single candidate

5.1 Plurality

► Each voter names a single candidate as his vote. Winner is the one named the most times. This system, despite its many flaws, is the world's most-used voting system.

Among the most serious of Plurality's flaws are the "vote splitting" and "cloning" pathologies. These can easily allow honest plurality voters to elect a candidate that the majority of them rank dead last.

In the example of table 1.1, but with the final two votes changed to $C > B > A$ (Borda 1781) A is the Plurality winner but is the Condorcet-*loser*. B is the Condorcet-Winner. (This also shows how sincere *approval voting*, §7.7, can elect the Condorcet-loser.)

In South Korea's 1987 presidential election, two liberals (Kim Dae Jung and Kim Young Sam) faced the heir (Roh Tae Woo) of a military dictatorship. The liberals together got 55.2% of the plurality votes but split their supporters, so Roh won with 36.6%, then claimed a mandate to continue repressive policies. After their defeat at the next election years later, the militarist party's leaders were convicted of treason for ordering the tragic shooting of pro-democracy demonstrators.

It is often strategically wise to vote for somebody other than one's favorite to avoid "wasting" one's plurality vote. (E.g. see table 6.5.) The strategic decision is to vote for the least-worst among the two "frontrunners." Over time this tends to cause two-party domination and the morbidity of "third parties," a phenomenon that has been called "Duverger's law."

5.2 Anti-plurality (sometimes called "Bullet voting"; perhaps better would be "veto-voting")

► Winner is the one named the *fewest* times.

Anti-plurality behaves extremely badly in the presence of strategic voters: they will always try to eliminate their favorite's greatest perceived rival, with the results that all the pre-race favorites will be eliminated and a "dark horse" will always win.

5.3 GR1: Gibbard's Random Dictator (1977) [33]

► Select a ballot at random (all voters equally likely) and make the winner be the candidate named in that one vote. (Ignore all other ballots.)

5.4 P+R: Plurality with genuine runoff

► Really two elections one after the other. The first round is a plurality election with the two candidates named the most often being the only two candidates in the second and final round. (The second round is often avoided if the first round produces a winner with $> 50\%$ of the votes.)

A system very much like this is used to elect the president of France. (Compare §6.2.) Since having two elections is usually very much less desirable than having one election, the P+I system below should be preferred.

6 Systems in which each vote is a preference ordering of the candidates

6.1 GR2: Gibbard Random pair (1977) [33]

► Each vote is a preference order among the candidates. Select 2 candidates at random (all pairs equally likely), then perform a 2-candidate election among them by ignoring the other $N - 2$ candidates in each preference ordering.

Both this and preceding GR1 system were introduced by Allan Gibbard, who proved they were essentially the only two voting systems in which each voter was motivated to express his *honest* opinion as his preference-ordering vote. However, as Gibbard remarks, they "leave too much to chance" to be acceptable in practice.

6.2 P+I: Plurality with instant runoff

► Each vote is a preference order among the candidates. In the first round we find the two candidates who are top-ranked most often. In the second and final "runoff" round, we perform a 2-candidate election among *just those two* by ignoring all the other candidates in every preference ordering.

In both P+I and P+R, it can be strategically desirable to vote dishonestly. For example, it can be optimal to vote for a Horrible candidate in round 1. That could cause the final round to be Your-Favorite vs. Horrible, at which point Your-Favorite will win. An honest vote for Your-Favorite, on the other hand, could cause the final round to be Your-Favorite vs. Somebody-More-Popular. Thus, honestly top-ranking Your-Favorite can cause him to lose, and it can be better for your cause to vote dishonestly or not vote at all.

6.3 Nauru [60]

Nauru Island, a Pacific atoll near the equator (pop.$\approx 10,000$), adopted the following voting method in 1971.

► Each vote is a preference order among the candidates. The kth-ranked candidate ($k = 1, \ldots, N$) is awarded a score of

$1/k$, the scores are summed, and the candidate with the greatest total score wins.

This is a typical **"weighted positional"** (WP) system with weights $w_k = 1/k$. In a WP system, each candidate gets a score of w_k if he is kth ranked in some vote, and the candidate with the greatest score-sum wins. (For some fixed real constants $w_1 \geq w_2 \geq \cdots \geq w_N$, not all equal.)[8]

6.4 Borda count (1781) [64][65][84]

Jean-Charles le Borda (1733-1799) was a French mathematician and nautical astronomer. It appears, however, that "Borda's" system actually had been published earlier by Nicholas of Cusa ($\approx$ 1400-1464).

► Each vote is a preference order among the candidates. We award points to the candidates as follows: If a candidate is Kth-ranked in a vote, he gets $N-K$ points ($1 \leq K \leq N$). The candidate with the most points wins.

Another way to view it: Borda is equivalent to the aggregate sum of all $\binom{N}{2}$ pairwise elections, assuming voters vote in each pairwise election in a manner compatible with their Borda vote; a candidate W's Borda score is $\sum_A M_{WA}$ (up to a rescaling, which does not matter).

Another: this is weighted positional with weights $w_k = N-k$.

The Borda system has several optimality properties (see footnote 14 and [65]) which seem of little relevance to the real world.

What unfortunately seems of much greater relevance to the real world is the fact that Borda behaves very badly in the presence of strategic voters.[9] For example, in a 3-way contest between 3 strong candidates and 10 mediocrities, strategic voters tend to rank their favorite candidate (usually one of the Big Three) "top" and his two Big Rivals "bottom" because honestly ranking all three of them near the top would give that voter's vote only $\approx 10\%$ of the "discriminatory strength" attainable through this sort of exaggeration. The trouble is: if nearly all voters adopt this strategy, then the 3 "strong" candidates are guaranteed to each get a below-average Borda score so that a "dark horse" non-entity is guaranteed to win. [10] A large number of voting systems based on preference-ordering ballots suffer from this devastating pathology, so it is convenient to give it a name: the "**DH3**" (Dark Horse wins 3^+-way race in the presence of strategic voters) pathology. We shall discuss it further in table 6.4.

Observe that Borda's response to DH3 is far worse than either plurality, approval (§7.7) or IRV (§6.16) voting, each of which would – quite sanely – elect one of the Big Three.

Another – even more commonly occurring – example of Borda's poor response to strategic voting is this. Consider a **3-candidate election** in which two of the candidates are (according to pre-election polls) extremely likely to win. Strategic Borda voters always will vote the maximum for the better, and the minimum for the worse, among these two frontrunners, forcing them to rank the remaining candidate middle (regardless of that voter's opinion of him). Just as in Plurality voting, this forces one of the two pre-election poll frontrunners to win, even if 90% of the voters honestly consider the non-frontrunner to be the best.[11] In contrast, under approval (§7.7) or range voting that truly-best candidate would be elected, by either strategic or honest voters, by a large margin. (Thus we again might expect two-party domination, with third parties having no chance, under Borda voting.)

Since Donald Saari has written a book [65] championing Borda voting it is worthwhile to criticize his results. Two of the reasons Saari likes Borda voting are: (1) He considered the class of Weighted Positional Voting (WPV) systems. He showed *Borda (with honest voters) is the only fair WPV system also satisfying "reversal symmetry."*[12] Saari then considered the "dictionary" mapping the voters's candidate preference orders (i.e. V permutations of the N candidates) into the N-permutation output by the voting system. (2) The more entries such a dictionary could have, the more Saari considered a voting system "paradoxical." Saari showed that *Borda is the uniquely least paradoxical WPV system.*

Let me counterargue.

1. In [68] I've worked with COAF voting systems, a highly general class. Because WPV systems are merely an infinitesimally tiny subclass of COAF systems there is no reason to care about optimizing over them if we can instead optimize over COAF.

2. In particular, range voting (§7.7) *also* obeys fairness and reversal symmetry. This would not be possible (by Saari's theorem) if honest range voting were a WPV system – but we evade Saari's theorem by working in the wider class of COAF systems.

3. It is wrong to force the input to the voting system to be V preference permutations. Really, the true input is V real utility N-vectors. Saari, and every WPV system, ignore (and prevent the voter from honestly expressing) the fact that a voter cares more about making A beat B if $U_A - U_B = 999$, than he cares about making B beat C if $U_B - U_C = 0.01$. (Where U_n is the utility of candidate n.)

[8] In votes by sportswriters for Major League Baseball's MVP award, a positional voting system with weights $(14, 9, 8, 7, 6, 5, 4, 3, 2, 1, 0, 0, \ldots, 0)$ is employed.

[9] A frustrated Borda is alleged to have exclaimed, in reply to criticisms of this sort: "My method is only intended for honest men!"

[10] This in fact is exactly what happened at a certain recruitment committee meeting I once attended. This scenario is quite common in practice. A lessened form of this pathology can occur with only *two* strong candidates – if they are close to being tied, then the most above-average among the mediocrities will win. The Pacific Island Republic of Kiribati adopted Borda voting and immediately suffered from exactly this kind of problem as the two "most popular" candidates were eliminated and two dark horses who did not campaign and were not recognized as serious contenders were nearly elected ([60] p.368).

[11] Admittedly, such a candidate would probably be high ranked in the polls. I am simply trying to dramatize the fact that, in 3-candidate Borda or plurality elections, the two candidates with the most advertising and loudest propaganda have essentially 100% chance of being elected purely because this loudness causes rational voters to *think* they have a high chance of being elected, totally independent of those two candidates's actual or perceived virtues.

[12] For another characterization of Borda, see [84].

4. Nobody cares about rank-ordering the losers! We care about finding the winner. (Well, there are uses for a voting system outputting a full ordering, but they are of secondary importance.) So Saari's "dictionary" is dominated by irrelevancy.

5. Saari ignores the reality that voters are rational – instead modeling them as imbeciles who always vote "honestly," no matter how tactically stupid that is.

Saari's paradox theorem is beautiful, but do not be deluded into thinking it tells us much about how to build a good voting system. It doesn't.

#voters	their vote	with no D
3	$A > B > C > D$	$A > B > C$
2	$B > C > D > A$	$B > C > A$
2	$C > D > A > B$	$C > A > B$

Figure 6.1. 7-voter Borda example by Paul Johnson.

Totals: $\mathbf{C = 13}$, $B = 12$, $A = 11$, $D = 6$.

But now suppose it is revealed that D – top ranked by *nobody* and far in last place – was a criminal noncitizen and hence ineligible to run. Remove D from all votes (right) and the totals become

Totals: $\mathbf{A = 8}$, $B = 7$, $C = 6$.

The removal of the "irrelevant" candidate D completely reverses the election results! ▲

ER-Borda (Borda with *e*qual *r*ankings permitted) seems very superior to plain Borda since, e.g, it is immune to candidate-cloning *provided* the voters rank the clones equally. (If they wish to express slight preferences among the clones, then Borda remains horribly vulnerable to cloning.)

#voters	their vote	with no D
51	$A > B$	$A > B_1 > B_2 > B_3$
49	$B > A$	$B_1 > B_2 > B_3 > A$

Figure 6.2. Cloning in Borda. Obviously, A wins the election at left 51-to-49. But if the Bs are cloned into B_1, B_2 and B_2 in decreasing order of attractiveness (the clones are not exactly identical), then B_1 wins easily with 249. (A only gets 153.) The B party assures victory by simply entering a large number of near-identical candidates into the race. ▲

ER-Borda highly resembles Range Voting (§7.7) but differs from it in that the score-range has *variable* size (n with n candidates) whereas with, e.g. 0-99 integer-score range voting it would have *fixed* size (100). Variable size seems to be the worse choice because it leaves ER-Borda vulnerable to removal and addition of "irrelevant" candidates as in table 6.1; range voting is immune to that.

The Borda count has been used by the Associated Press and ESPN/USA Today for US college basketball and football polls; in these applications Borda might work well because they are probably largely free of strategic voting.

6.5 Condorcet least-reversal system (1785) [86]

The Marquis de Condorcet (1743-1794) was among the earliest and best investigators of voting systems. Ironically, he died in a prison cell where he had been thrown as a plausible enemy of the French revolution. Condorcet observed that Borda's system can elect a unique winner *other* than the Condorcet-Winner, even when a Condorcet-Winner exists. More seriously, *every* weighted-positional scoring system can simultaneously elect a unique winner *other* than the Condorcet-Winner, even when a Condorcet-Winner exists, as shown in figure 6.3.[13]

#voters	their vote
5	$A > B > C$
4	$B > C > A$
2	$B > A > C$
2	$C > A > B$

Figure 6.3. 13-voter Condorcet vs WP example. A is the unique Condorcet-Winner. However, the total numbers of (top-ranked, mid-ranked, bottom-ranked) votes garnered are respectively $A : (5, 4, 4)$, $B : (6, 5, 2)$, and $C : (2, 4, 7)$ so that in any weighted-positional score-sum system, no matter what weights $w_1 \geq w_2 \geq w_3$ (not all equal) are employed, B would be the unique winner.[14] ▲

We also remark that most weighted-positional systems[15] will elect B instead of the Majority-Winner A in the following situation: $50\% + \epsilon$ of the votes are $A > B > C$, and $50\% - \epsilon$ are $B > C > A$ for sufficiently small $\epsilon > 0$.

Condorcet therefore disparaged all weighted-positional score-sum systems and proposed the following system which avoids that perceived flaw.

► The winner is the candidate W minimizing $\sum_{n=1}^{N} \text{neg}(M_{Wn})$, where

$$\text{neg}(x) \stackrel{\text{def}}{=} \begin{cases} -x & \text{if } x \leq 0 \\ 0 & \text{if } x \geq 0. \end{cases} \qquad (20)$$

In other words, the candidate W who would win every pairwise election wins, but if there is no such "Condorcet-Winner" then the candidate is chosen who would win every pairwise election after first *reversing* the minimum possible number of pairwise-comparison subvotes.

[13] A weaker version of our example was stated by Condorcet.

[14] Gehrlein, Fishburn, and van Newenhizen [27][30][56] showed that the probability (assuming each vote is a random N-permutation) that Borda elects a non-Condorcet-Winner (conditioned on a CW existing) is about 10% if $N = 3$ and there are a large number of voters. They showed that Borda *minimizes* this conditional probability over all weighted-positional scoring systems, for each $N \geq 3$; thus the corresponding conditional probability for *plurality* voting is larger, namely $\approx 22\%$ in the $N = 3$ case. Borda also is the unique weighted-positional scoring system that cannot elect a Condorcet (or Smith) *loser* ([31], [65] p.192, known to E.J.Nanson in 1882). Finally, Merlin et al. [51] argued that the probability of a situation such as the one in table 6.3 in which *every* weighted-positional scoring system simultaneously elects some unique winner other than the CW is about 1.808% if $N = 3$.

[15] The exceptions are plurality with optional tie-breaking by another WP system.

Put another way: the candidate with the minimum sum of defeat-margins (where "defeat-margin"=0 in the case of a victory), is elected.

Unfortunately, although Condorcet's principle may overcome some deficiencies of the Borda system, it automatically ensures vulnerability to strategic manipulation in precisely the same two scenarios (DH3 and the 3-candidate election with two perceived "frontrunners") which we used to criticize Borda:

#voters	their vote
51%	$A > C > B$
49%	$B > C > A$

#voters	their vote
x_1	$A > D > B > C$
x_2	$A > D > C > B$
y_1	$B > D > C > A$
y_2	$B > D > A > C$
z_1	$C > D > A > B$
z_2	$C > D > B > A$

Figure 6.4. Damaging effects of strategic voting. (left) Two perceived "frontrunners" A and B run in a **3-way race** versus C. Strategic voters, in order not to "waste their vote," always rank their more-favored frontrunner "top" and their less-favored frontrunner "bottom," forcing C to be ranked middle regardless of the voters' actual honest opinions of C. Even if 90% of the voters honestly think C is the best candidate, he has no chance with this kind of voters in any voting system obeying Condorcet's principle, because this strategic exaggeration makes A the Condorcet-Winner. (A also wins in the Borda, Arrow-Raynaud, and IRV systems, but C could win convincingly under Approval or Range voting.)

(right) The *minimal* **DH3** example, i.e. with only a *single* "dark horse" candidate D. Assume A, B, C are three excellent candidates and D is a mediocrity. All voters honestly regard all three of $\{A, B, C\}$ to be far superior to D, but their opinions are split concerning the ordering within $\{A, B, C\}$. Therefore, each voter strategically ranks his favorite top, and his two top rivals artificially "last" (exaggerating to get more "discriminating power"). The result is the scenario here with $x_1+x_2 \approx y_1+y_2 \approx z_1+z_2$. Then the uniquely-worst candidate D becomes the Condorcet-Winner and wins the election under *any* voting system that elects Condorcet-Winners or Smith-Set members! (However, Plurality, IRV, Approval, and Range would elect one of $\{A, B, C\}$.) This is a common scenario in practice and illustrates the vulnerability of systems obeying Condorcet's principle to strategic manipulation. ▲

In both examples in table 6.4, it is important to realize that, in either the Borda or Condorcet sum-of-defeats system, each voter *is* acting rationally in his own self-interest – his vote is one that maximizes the expected election result for him, given his prior conviction that the alphabetically-last candidate has essentially no chance of victory. For example, if in the first scenario some voter had honestly voted $C > A > B$, then that would, in the event of a Condorcet cycle as in table 1.1, increase A's margin of pairwise defeat versus C and hence possibly help B to win. For a completely concrete instantiation of that, see table 6.5.

#voters	their vote
8	$B > C > A$
6	$C > A > B$
5	$A > B > C$

Figure 6.5. Favorite Betrayal, or how dishonest exaggeration can pay. In this 19-voter example there is a Condorcet cycle, and the winner is B under either Plurality, Borda=Dabagh, P+I, Black, Schulze-Beatpath, IRV, Loring, BTR-IRV, Coombs, River, Maxtree, Tideman Ranked Pairs, Improved-Dodgson, Simpson-Kramer, Nanson, Rouse, Raynaud, Arrow-Raynaud, Condorcet Least Reversal, Woodall DAC, Sarvo-Plurality, Sinkhorn, or Keener-eigenvector.[16]

But if the 6 $C > A > B$ voters insincerely switch to $A > C > B$ ("betraying their favorite" C) then A becomes the winner under all these voting systems (and is the Condorcet-winner), which in their view is a better election result. ▲

This favorite-betrayal example is very important because, once voters understand that exaggerating their stances on the apparent-frontrunners can be necessary to prevent the enemy frontrunner from winning, strategic voting is guaranteed, often causing "third parties" to tend to die out (since the strategic voters will not "waste their vote" on third-party candidates like C whom they perceive as having "no chance of winning"). IRV proponents have sometimes falsely stated that IRV (§6.16) eliminates the "wasted vote" phenomenon.[17] Table 6.5 was a counterexample, both for IRV, and for most other ranked-ballot systems too. However, *approval* (§7.7) and *range* voting (§7.7) *are* arguably immune to this pathology, in the sense that dishonesty is never strategically useful in a (≤ 3)-candidate approval election.

Eventually it was realized [54][59][85][67] that *every* single-winner voting method based on preference-orderings which satisfies Condorcet's principle (that it elects the Condorcet-Winner whenever one exists) must exhibit

No-show paradoxes in which adjoining an additional set of identical votes, all favoring A over B, causes A to lose and B to win (so that these voters would have been better off not voting).

Consistency paradoxes in which two disjoint subsets of votes, each of which by itself elects A, when combined elect B.

[16] The Smith and Fishburn sets are $\{A, B, C\}$ and this example can demonstrate favorite-betrayal for them too, if, e.g. the tie-breaking method is random choice and the utility of A is higher than average. Similarly this demonstrates favorite-betrayal for Copeland and TMR.

[17] Many false claims about IRV may be found on the "Center for Voting and Democracy" web site, e.g. to quote Steven Hill of the CVD: "The instant runoff ensures the election of the candidate preferred by most voters. [False, see table 6.15.] It eliminates the problem of spoiler candidates knocking off major candidates. [False, see table 6.5.] It frees communities of voters from splitting their vote among their own candidates. [We discuss this at the end of §6.16.] It promotes coalition-building and more positive campaigning. [Perhaps too vague to justify, and the CVD gives no historical evidence for it.]" The CVD is a biased IRV-advocacy organization which pretends to provide unbiased information, but in fact refuses to correct errors of this nature.

For a combined example of both sorts of paradox simultaneously in the Condorcet least-reversal system, consider

$$M = \begin{pmatrix} * & -5 & H & G \\ 5 & * & -3 & -3 \\ -H & 3 & * & -x \\ -G & 3 & x & * \end{pmatrix} \quad (21)$$

where x, G, H are any numbers obeying $|x| \ll G, H$ and where the candidates in order are A, B, C, D. Here A is the winner (5 reversals required). But if we add 4 $A > B > C > D$ ballots then B becomes the Condorcet-winner. (This also shows failure of "add-top.")

6.6 Black's system (1958) [6]

Embarrassingly, Condorcet's least-reversal system can, if there are ≥ 4 candidates, elect a Condorcet *loser*! In the 7-voter scenario with U-matrix[18]

$$U = \begin{pmatrix} * & 7 & 0 & 4 \\ 0 & * & 7 & 4 \\ 7 & 0 & * & 4 \\ 3 & 3 & 3 & * \end{pmatrix}, \quad (22)$$

the last candidate is the Condorcet-loser, i.e. a loser of *every* pairwise battle. It changes to a Condorcet-Winner by making 3 preference reversals, while other candidates need at least 4 preference reversals to become a Condorcet-Winner. ▲

Duncan Black therefore proposed the following, in an effort to get a system which (1) was monotonic (as is Condorcet's original system); (2) elects the Condorcet-Winner whenever one exists; (3) never elects a Condorcet-loser (if there is more than one candidate):

► If there is a Condorcet-Winner, he wins. Otherwise, employ the Borda count system.

This "fix," however, evidently does not cure all the deficiencies of Condorcet Least-Reversal voting; the example of EQ 21 still applies to Black voting.

6.7 ID: Improvement of Dodgson's system

Charles Lutwidge Dodgson (1832-1898) was best known under his pseudonym Lewis Carroll, under which he wrote famous children's books such as *Alice in Wonderland* and verse such as *The hunting of the snark*. However, he was also a mathematician. He invented a very interesting way to evaluate determinants in 1876, and discussed several kinds of voting systems in pamphlets [23] also published in 1876. In his first pamphlet Dodgson recommended Condorcet Least-Reversal (which he apparently reinvented), but he later abandoned it.

The method Dodgson recommended in his final pamphlet is this. Each vote is a preference order among the candidates. The winner is the candidate W who would win every pairwise election. But if there is no such "Condorcet-Winner" then take the candidate who would win every pairwise election after performing the minimum possible total number of adjacent interchanges inside the preference orderings. (If $A > B > C$ is a preference ordering then $A > C > B$ is 1 adjacent-interchange away and $C > A > B$ is two adjacent interchanges away.)

This system is algorithmically infeasible: Hemaspaandra et al [38] proved that the problem of determining the Dodgson election winner from the votes is "Θ_2^p-complete," that is, complete among all problems soluble by a Turing machine in polynomial time where the Turing machine is allowed, in each step, to pose a problem to an NP-oracle, and the answers to all the questions then come back in *exactly one batch* from that oracle.

But there is a feasible variant: We now wish to point out for the first time, however, that a voting method very similar to Dodgson's[19] *is* algorithmically feasible. It is this new method which we shall, therefore, call "Dodgson's" method.

► The method is this. Each vote is a preference ordering among the N candidates and may therefore be regarded as N *points*, one for each candidate, located at the locations 0, 1, 2,..., $N-1$ on the *real line*. Let $f(x)$ be some arbitary but fixed increasing real-valued function of one real argument, with the properties that $f(x) = -f(-x)$, $f(0) = 0$, and $f(x)$ is (non-strictly) concave-∩ for $x > 0$. If we restrict attention to just *two* candidates A and B, then each vote in which A's location is a distance x to the right of B's location may be thought of as contributing "score" $f(x)$ to A in his pairwise battle versus B. Let the total score by which A then beats B (which would be negative if B beats A) be called D_{AB}.

If some candidate W beats every other (i.e. if $D_{WA} \geq 0$ for all A), then W wins.[20] Otherwise, consider the minimum total *distance* ℓ_W that all points besides W on the real line have to *move* in order to cause W to beat (or at worst be tied with) each other candidate. The candidate W with the minimum ℓ_W is the winner.

Note: if $f(x) \equiv x$ then this is just the Borda count system. If $f(x) = \text{sign}(x) + \epsilon\text{sign}(x)\sqrt{|x|}$ in the limit $\epsilon \to 0^+$, then it becomes Condorcet's least-reversal method. Other choices of $f(x)$ would lead to other voting systems intermediate between the two. *We shall, for concreteness, by default use* $f(x) \stackrel{\text{def}}{=} \sqrt{|x|}\,sign(x)$ *– the "geometric mean" of Borda and Condorcet – throughout this paper whenever we speak of "Dodgson's" method.*

The reason this voting method is algorithmically feasible is that here is a polynomial time "greedy" algorithm for evaluating ℓ_W:

procedure Greedy-ℓ_W-Evaluation
1: Start with $\ell = 0$.
2: Find, among all the candidates A that currently beat W, and among all votes, the location of a point A that is nearest to W. (Prefer locations to the right of W, if there

[18]Actually, this particular U-matrix is not achieveable – which is a good illustration of the fact that not every V- or M-matrix is achieveable – but by adding a suitably large constant to all off-diagonal entries, it becomes achieveable and the same scenario happens. This same example also works against the Simpson-Kramer system.

[19]And indeed, conceivably actually the one he had in mind – Dodgson's methods are described "by example" in such an informal manner that it is hard to be sure exactly what he meant.

[20]Note, for general f this no longer forces a Condorcet-Winner to be elected, although with the choice $f(x) = \text{sign}(x) + \epsilon\text{sign}(x)\sqrt{|x|}$ in the limit $\epsilon \to 0^+$, it does.

is a distance-tie between one on W's left and one on W's right.)

3: Move that A-point a distance 1 to the left, and $\ell \leftarrow \ell + 1$.

4: Go back to step 2 until W beats (or at worst is tied with) every other candidate.

5: Output ℓ.

The candidate W with the least ℓ_W wins.

Why this works: Essentially, each distance-1 movement causes the maximum possible decrease in $\sum_A \text{neg}(D_{WA})$. This is due to the concave-$\cap$ nature of $f(x)$.

Another way to look at it: The fundamental structural reason why our small modification in the definition of the Dodgson system renders it polynomial time instead of Θ_2^p-complete, is that in the modified system, the ℓ_W's may be formulated as the solution of a *convex programming problem* [36][80]. In Dodgson's original formulation, the requirement to move candidates *one past* others by pairwise interchanges introduced nonconvexity.

Runtime: If a brute-force exhaustive search is made in step 2, then its runtime will be $O(VN)$ steps, the runtime of the algorithm to evaluate ℓ_W will be $O(V^2N^2)$ steps, and the runtime of the full algorithm to find the winner (best W) will be $O(V^2N^3)$ steps. If the search instead is done by using a simple data structure consisting of $O(N)$ linked lists, then the data structure may be built in $O(VN)$ steps and searched in $O(N)$ steps. Hence the runtime of the algorithm to evaluate ℓ_W will be $O(VN^2)$ steps, and the runtime of the full algorithm to find the winner will be $O(VN^3)$ steps. By doing "several of A's hops in one go" (as is often possible) further speedup can be achieved in large-distance cases.

#voters	their vote
4	$A > B > C$
3	$B > C > A$

Figure 6.6. Here Improved-Dodgson fails to elect the Condorcet-Winner, and indeed Majority-Winner, A; here B wins because $4.243 \approx 3\sqrt{2} > 4$. (But apparently I.D. is incapable of electing a Condorcet-Loser.) ▲

#voters	their vote
3	$A > D > B > C$
3	$A > D > C > B$
4	$B > C > A > D$
5	$D > B > C > A$

Figure 6.7. 15-voter Improved-Dodgson "add-top" failure. In this situation, A and B are tied for the win. (A beats D since $10 > 5\sqrt{3}$ and beats C since $3\sqrt{3}+3\sqrt{2} > 9$ but B needs to move distance 6 for A to beat it. B beats A and C pairwise but D needs to move distance 6 for B to beat it. The reason D and C cannot win is that D cannot beat A without motion of distance 8, and C cannot beat B without motion of distance 9.) If we now add 5 $B > D > A > C$ voters then D becomes the clear winner (no motion required), violating "add-top." ▲

6.8 TMR: Topmost Median Rank

► The candidate W with the topmost median rank wins.

Annoyingly, TMR often yields a tie. An improvement of it (*improved TMR*, or ITMR) which yields a tie less often is as follows. Suppose a candidate is ranked $\le m-1$ by a fraction $f < 1/2$ of the voters and $\ge m+1$ by a fraction $g < 1/2$ of the voters. Then regard that candidate's "median rank" as the *real number* $m + (f - 1/2)/(f - g)$ instead of the integer m.

6.9 A.H.Copeland's system (1951) [15]

► The candidate W maximizing $\sum_A \text{sign}M_{WA}$, wins. (In other words, W is the one with the most "pairwise victories" where a tie counts as half a victory.)

In most voting methods, if the number V of voters is made very large while the number N of candidates is held fixed, the probability of a tie approaches zero. However, that is not the case with Copeland and (unimproved) TMR; in both there is still a large probability of a tie even with huge V. For example in table 6.5, both TMR and Copeland declare the election a 3-way tie. Alex Small therefore proposed the following improvement: eliminate all the candidates except for the co-equal Copeland winners. Then re-run Copeland on those candidates only (using the vote preference orderings with all eliminated candidates erased from them). This elimination and re-Copeland process is continued until the winner-set ceases shrinking. Although Small's variant reduces the probability of a tie, that probability *still* is bounded above zero even with V random votes when $V \to \infty$. (Small then admits defeat and breaks any *still*-remaining ties randomly.)

Alternatively, we could employ any other voting system to break the Copeland ties. Copeland is essentially the method used to determine the winner of round-robin chess tournaments.

6.10 Schulze's beatpath system (1997) [66]

A "beatpath" from candidate A to candidate B is a path of candidates A-X-Y-$\cdots$-Z-B, and its "strength" is the minimum margin of victory of each candidate in the path over the one immediately to his right. Thus the strength of this beatpath would be $\min\{M_{AX}, M_{XY}, \dots, M_{ZB}\}$. Let S_{AB} denote the maximum strength among all beatpaths from A to B.

► Markus Schulze proved that at least one candidate W must exist such that $S_{WA} \ge S_{AW}$ for every A (and indeed proved that the relation $S_{AB} > S_{BA}$ is a transitive relation among candidates A, B). Such candidates are "potential winners."

Schulze observed that if there are no pairwise ties and there are no pairwise defeats of equal strength (both are generically true if the number of voters is very large and their votes contain independent randomness), then the Schulze winner is unique.

In the event of non-uniqueness, Schulze proposed the following randomized tie-breaking procedure for choosing a winner from among the potential winners: Order the potential winners compatibly with the ordering on a random ballot.

Schulze presented a simple $O(N^3)$-step algorithm to compute all the S_{AB}:

procedure Schulze-beatpath-strengths[21]

for $i = 1, \dots, N,\ \ j = 1, \dots, N$ **do**
  $S_{ij} \leftarrow M_{ij}$;
end for
for $i = 1, \dots, N,\ \ j = 1, \dots, N,\ \ k = 1, \dots, N$ **do**
  if $i \neq j$ and $j \neq k$ and $i \neq k$ **then**
    $m \leftarrow \min\{S_{ji},\, S_{ik}\}$;
    $S_{jk} \leftarrow \min\{S_{jk}, m\}$;
  end if
end for

Schulze's method has been adopted by various large software development organizations (e.g. Debian Linux, Board of Software in the Public Interest) for their votes.

S.Eppley pointed out that, in the absence of a Condorcet-Winner (due to a Condorcet cycle as in table 1.1) somebody could always complain that the election winner W would have lost to somebody else (A) by majority vote in a head-to-head election. However, suppose any such complaint could be rebutted by turning the complainer's own logic against him, i.e. by noting that A would also have been beaten by somebody else B, who in turn was beaten by somebody else C,..., who in turn was beaten by W with all of these defeats being *even stronger* than the $A > W$ defeat that inspired the complaint – then those complainers could be squelched.

In the absence of ties, Schulze's Beatpath system enjoys precisely this "immunity to complaints" property. So do River, Tideman's Ranked Pairs system, and our new maxtree system (§6.13), but somehow Schulze seems the most natural embodiment of the idea.

Another way to look at Schulze's method (and these other two methods) is that they are attempts to make Condorcet's least-reversal method (§6.3) less vulnerable to strategic manipulation.[22] They provably accomplish this in the sense that all three of these methods are "immune to candidate cloning," i.e, replacing a candidate with several "clones" will not affect the election results (as opposed to in the plurality system, where the clones "split the vote," and the Borda system, where they hugely increase each others chances via "teaming"). Thus in the DH3 scenario of §6.3-6.3, the *number* of mediocre "dark horse" candidates will not affect the situation (unlike in Condorcet's original method, where more dark horses make it worse).[23] However, since even a single dark horse suffices (table 6.4) this improvement is insufficient to cure the DH3 problem.

#voters	their vote
12	$D > E > C > A > B > F$
10	$F > A > B > C > D > E$
8	$E > C > D > B > F > A$
6	$F > A > B > D > E > C$
4	$A > B > C > D > E > F$
4	$C > D > B > F > E > A$
4	$F > E > D > B > C > A$
4	$A > E > B > F > C > D$
2	$A > E > F > B > C > D$
2	$B > F > A > C > D > E$
2	$A > B > F > D > E > C$
2	$C > D > B > E > F > A$

Figure 6.8. Schulze-voting "no-show" and add-top paradoxes. In this situation by Schulze [66], the U-matrix and path-strength matrix S are below and the Schulze winner evidently (in view of the first row and column of S) is A.

$$\begin{pmatrix} - & 40 & 30 & 30 & 30 & 24 \\ 20 & - & 34 & 30 & 30 & 38 \\ 30 & 26 & - & 36 & 22 & 30 \\ 30 & 30 & 24 & - & 42 & 30 \\ 30 & 30 & 38 & 18 & - & 32 \\ 36 & 22 & 30 & 30 & 28 & - \end{pmatrix}, \quad \begin{pmatrix} - & 20 & 8 & 8 & 8 & 16 \\ 12 & - & 8 & 8 & 8 & 16 \\ 4 & 4 & - & 12 & 12 & 4 \\ 4 & 4 & 16 & - & 24 & 4 \\ 4 & 4 & 16 & 12 & - & 4 \\ 12 & 12 & 8 & 8 & 8 & - \end{pmatrix}$$

But after adding 3 $A > E > F > C > B > D$ voters, the unique Schulze winner changes from their top-ranked candidate A to their bottom-ranked candidate D! Those voters evidently would have been better off had they not shown up. The new V- and S-matrices are

$$\begin{pmatrix} - & 43 & 33 & 33 & 33 & 27 \\ 20 & - & 34 & 33 & 30 & 38 \\ 30 & 29 & - & 39 & 22 & 30 \\ 30 & 30 & 24 & - & 42 & 30 \\ 30 & 33 & 41 & 21 & - & 35 \\ 36 & 25 & 33 & 33 & 28 & - \end{pmatrix}, \quad \begin{pmatrix} - & 23 & 5 & 5 & 5 & 13 \\ 9 & - & 5 & 5 & 5 & 13 \\ 7 & 7 & - & 15 & 15 & 7 \\ 7 & 7 & 19 & - & 21 & 7 \\ 7 & 7 & 19 & 15 & - & 7 \\ 9 & 9 & 5 & 5 & 5 & - \end{pmatrix}$$

respectively, and D is the winner in view of the 4th row and column of S. ▲

6.11 N.Tideman's "ranked pairs" (1989) = S.Eppley's "maximum affirmed majorities" [88]

► Find[24] the candidate pair AB with the largest pairwise margin of victory M_{AB} and "lock it in" by drawing an arrow from A to B. We proceed through all victories in decreasing-magnitude order, "locking them in" if so doing does not create a directed cycle in the directed graph we are drawing. The

[21] In **for** statements describing multiple loops, the variable stated last is for the innermost loop.

[22] One could attempt to "improve" upon Schulze's method by computing the maximum flow from A to B in the directed graph with arc-capacities. This is "better" because it considers *all* beatpaths from A to B, not just the strongest one. However, such a "maxflow voting system" would be highly vulnerable to candidate-cloning, unlike Schulze's system.

[23] Also more generally, by cloning a candidate who pairwise-beats the Condircet Least Reversal winner, more reversals are required, which can prevent him from still winning.

[24] Eppley's and Tideman's procedures actually differ, but if all voters provide full tie-free rankings of all candidates – which is the only case we shall consider here – then they are the same.

root of the resulting directed-graph (the only candidate with no arrows pointing to him) then is the winner.

Note: Tideman (in the absence of pairwise ties) always produces a full ordering of the N candidates. (Proof: If any two A, B were not ordered at the end, then Tideman would not be prevented from adding a directed arc $A \to B$ or $B \to A$, since this addition could not cause a directed cycle.)

#voters	their vote
12	$A > C > E > B > D$
10	$E > B > D > A > C$
9	$D > C > B > E > A$
7	$B > A > D > C > E$
5	$E > A > C > B > D$
4	$E > D > C > B > A$
3	$D > A > B > C > E$
3	$A > C > B > E > D$
3	$E > D > C > A > B$
2	$C > E > B > D > A$
1	$D > E > B > A > C$
1	$E > C > B > D > A$

Figure 6.9. Tideman-voting drastic "no-show" and add-top paradoxes. (Example by M.Schulze.) The Tideman ranking is $A > C > B > E > D$ and hence A wins. The V- and M-matrices (where we place *minus signs* above rather than before numbers) are

$$U = \begin{pmatrix} * & 26 & 41 & 27 & 25 \\ 34 & * & 21 & 40 & 22 \\ 19 & 39 & * & 23 & 36 \\ 33 & 20 & 37 & * & 20 \\ 35 & 38 & 24 & 40 & * \end{pmatrix}, \quad M = \begin{pmatrix} 0 & \overline{8} & 22 & \overline{6} & \overline{10} \\ 8 & 0 & \overline{18} & 20 & 16 \\ \overline{22} & 18 & 0 & \overline{14} & 12 \\ 6 & \overline{20} & 14 & 0 & \overline{20} \\ 10 & \overline{16} & \overline{12} & 20 & 0 \end{pmatrix}$$

If we add 3 $A > D > B > C > E$ voters then

$$U = \begin{pmatrix} * & 29 & 44 & 30 & 28 \\ 34 & * & 24 & 40 & 25 \\ 19 & 39 & * & 23 & 39 \\ 33 & 23 & 40 & * & 23 \\ 35 & 38 & 24 & 40 & * \end{pmatrix}, \quad M = \begin{pmatrix} 0 & \overline{5} & 25 & \overline{3} & \overline{7} \\ 5 & 0 & \overline{18} & 17 & \overline{13} \\ \overline{25} & 18 & 0 & \overline{17} & 15 \\ 3 & \overline{17} & 17 & 0 & \overline{17} \\ 7 & 13 & \overline{15} & 17 & 0 \end{pmatrix}$$

Now the Tideman ranking is either $E > A > D > C > B$ or $E > D > A > C > B$, depending on tie-breaks, and either way E (the bottom-ranked choice of the extra voters!) wins. ▲

6.12 "River" system by Jobst Heitzig (2004)

► Find the candidate pair AB with the largest pairwise margin[25] of victory M_{AB} and "lock it in" by drawing an arrow from A to B. We proceed through all victories in decreasing-magnitude order, "locking them in" if so doing creates neither a directed cycle nor a "branching" (that is, two arrows XB and YB with the same destination-node B) in the directed graph we are drawing. The root of the resulting directed-tree then is the winner.

Note: the no-branching condition forces each node to have only one "parent" and hence forces the directed acyclic graph we are drawing to be a *tree* with all arc-arrows directed away from the root.

According to Schulze, River obeys the following property (while both Schulze's own method, and Tideman's, fail it as he demonstrates [66] with two examples): if a new candidate Z is added, and every voter ranks $A > Z$ for some-already running candidate A, then there is zero probability of the election result changing.

#voters	their vote
6	$B > A > D > C$
6	$D > C > A > B$
6	$C > A > B > D$
5	$D > B > C > A$
4	$A > D > B > C$
4	$B > C > A > D$
3	$A > B > D > C$
2	$C > B > A > D$

Figure 6.10. River-voting "no-show," "add-top," and "later harm" paradoxes. (Example by M.Schulze.) River's winner is A and the tree-diagram is $B \leftarrow A \to D \to C$.

$$U = \begin{pmatrix} * & 19 & 13 & 25 \\ 17 & * & 22 & 21 \\ 23 & 14 & * & 12 \\ 11 & 15 & 24 & * \end{pmatrix}, \quad M = \begin{pmatrix} 0 & 2 & \overline{4} & 14 \\ \overline{2} & 0 & 8 & 6 \\ 4 & \overline{8} & 0 & \overline{12} \\ \overline{14} & \overline{6} & 12 & 0 \end{pmatrix}$$

If we add 3 $A > B > C > D$ voters then the matrices are

$$U = \begin{pmatrix} * & 22 & 16 & 28 \\ 17 & * & 25 & 24 \\ 23 & 14 & * & 15 \\ 11 & 15 & 24 & * \end{pmatrix}, \quad M = \begin{pmatrix} 0 & 5 & \overline{7} & 17 \\ \overline{5} & 0 & 11 & 9 \\ 7 & \overline{11} & 0 & \overline{9} \\ \overline{17} & \overline{9} & 9 & 0 \end{pmatrix}$$

Then the tree-diagram is $B \to C \to A \to D$ and B is the winner. (This also illustrates a failure of "later-no-harm" since the 3 votes could have been $A > D > C > B$ in which case A still would have won. Thus the "later" decisions about B, C, D "harmed" A.)[26] ▲

There is an algorithm to compute the River winner in $O(VN + N^2)$ steps, i.e. in $O(N^2)$ steps from the M-matrix:

Each node B knows its best potential parent (N^2 steps to get that knowledge initially) i.e. the node A maximizing M_{AB}. Each tree knows its member-nodes (and its root-node) so we may avoid considering descendants as "potential parents." We now proceed as follows. Each stage we examine the current root nodes to find the one with the best potential parent (maximizing M_{AB} over all current-root-nodes B and suitable parents A). We then adjoin the arc $A \to B$ and combine A and B's two trees into one tree, with root the same as A's

[25]Actually, Heitzig prefers to use U_{AB} rather than M_{AB}, but the two are equivalent if we require voters to provide full preference orderings as votes (and we do).

[26]Tideman's ranked pairs method also violates add-top in 4-candidate cases, although not in this one; and River can also be made to violate participation in as maximally-drastic a manner as Tideman in figure 6.9.

old tree's root. These stages continue until only a single tree remains, and its root is the winner. (Initially we have N one-node trees.) Each stage requires $O(N)$ time, so the total time is $O(N^2)$.

Heitzig pointed out that River obeys "monotonicity," i.e. increasing the rank of A in some set of ballots (without changing anything else) cannot harm A. The algorithm makes that clear: Since A's pairwise victory margins over others cannot decrease, A's chances of being the best potential parent of anybody cannot decrease; meanwhile A's pairwise defeats cannot worsen, hence every other candidate, if it is a best potential parent of anybody, is less likely to be the best potential parent of A rather than somebody else first. Consequently A's chances of being the root of the tree cannot decrease. Q.E.D.

6.13 Max-tree system

Heitzig's River system might perhaps better be renamed the "greedy tree system." Compare it with J.B.Kruskal's famous algorithm for finding the "minimum spanning tree" in an undirected graph with real edge-lengths, i.e.

1. sort the edges in increasing-cost order,
2. go through the list picking each edge that would not create a cycle in the graph formed by the already-picked edges.

In other words, the *greedy* tree, in an *un*directed graph with edge-lengths, is in fact the *optimum* tree. (Of course, by sorting in *de*creasing order, or by using negated edge lengths, Kruskal similarly would obtain the *max*imum sum-of-edge-lengths spanning tree.) But that greedy=optimal property is not true in the case of directed graphs, as is shown by the digraph with nodes $ABCD$ and the following arc-cost matrix:

$$\begin{pmatrix} * & 3 & 2 & 1 \\ * & * & 4 & * \\ * & * & * & 7 \\ * & 10 & * & * \end{pmatrix}. \tag{23}$$

The min-cost directed tree is $A \to B \to C \to D$ with cost $3+4+7=14$, while the greedy-min tree is $A \to D \to B \to C$ with cost $1+10+5=16$.

This all suggests that the following might be a better voting system than Heitzig's River greedy-tree system:

► In the complete N-node directed graph with arc $A \to B$ having value M_{AB}, find the directed spanning tree with all arcs directed away from the root, with maximum sum of arc-values. Its root is the election winner.

(Other variants are possible, i.e. the goal of maximizing the *sum* may be replaced with the goal of maximimizing the sum of pth powers for some p, or of maximizing the *minimum* arc-value in[27] the tree [32], and we could restrict to trees with root-outvalence 1, in which case we would get not only a winner, but also a "second-place finisher.")

A polynomial-time algorithm for finding the max-directed-tree is known. It was originally invented by Jack Edmonds [25] and independently by Chu and Liu [12], and is explained in [40] and [44]. It was sped up to run in $O(E + N \lg N)$ steps for a digraph with E arcs and N nodes (in either the case where the root is pre-specified, or not) [28]. In our case $E = N^2$ so the total runtime for maxtree voting is $O(N^2)$ steps once the M matrix is known (and this bound is also valid for both of the variant algorithms we mentioned). This is the same time bound as for the greedy tree system, although we admit the algorithm is substantially more complicated to describe.

procedure Chu-Liu-Edmonds original $O(N^3)$-time algorithm to find min-summed-cost outward-directed tree with specified root-node r in a directed graph with arc-costs.

1: Discard arcs with destination r (if any).
2: For each non-root node η, select the min-cost arc with destination η.
3: If this set of $N-1$ arcs is cycle-free (tree) we are done.
4: Otherwise, contract each directed-cycle into a supernode k, modifying the arc-costs via

$$C_{ab}^{(\mathrm{new})} = C_{ab} - C_{\mathrm{pred}(b),b} + \min_j C_{\mathrm{pred}(j),j}$$

where a is a node not on the cycle, b is a node on it, and $\mathrm{pred}(j) \to j$ are the directed arcs on the cycle.
5: For each supernode, select the entering arc which has the least modified cost; replace the arc which enters the same *real* node with the new selected arc.
6: Go back to step 2 with the contracted graph and the modified arc costs.

The key idea behind the algorithm is to find the replacing arc(s) with the minimum extra cost to eliminate any cycle(s); the given cost-modification equation incorporates the associated extra cost. (If it is desired to leave the root-node unspecified, that may be accomplished by adjoining one artificial extra root with huge-cost edges from it to every genuine node.)

In the election of figure 6.10, the River and Max-tree systems produce identical trees, showing that max-tree also violates add-top and later-no-harm.

We shall now sketch proofs that maxtree voting obeys (1) monotonicity, (2) is unaffected by candidate "cloning," and (3) obeys Eppley's "immunity to complaints" property mentioned in §6.10.

(1) If some node A's pairwise victory-margins increase and defeat margins decrease, then in the Edmonds algorithm step 2, that can only increase A's chances of finding a child and decrease A's chances of being a child, i.e. A has a larger (or same) chance of being root. If we do some supernode-contractions and cost-modifications, then observe from the modified-cost formula that arcs entering A become (if anything) disfavored in comparison to those entering any other member of the same cycle forming that supernode. Again this can only decrease A's chances of being a child, i.e. a non-root. Q.E.D.

(2) There are two ways candidate-clones might affect the winner and the maxtree structure and we have to rule them both out. First of all, the maxtree could the the same as it was

[27] Markus Schulze points out: When you maximize the minimum arc-value, then you get the Schulze voting method (§6.10): The strongest paths from the Schulze winner to every other candidate form that tree where the minimum arc-value is maximized.

before the clones, but with extra duplicated arcs for those clones, i.e. an arc BC is now duplicated to have an additional arc BC' where C' is a clone of C, etc. This is fine since the winner is not affected. Second, interclone-arcs could be involved in the tree. If interclone arc never are involved in a directed cycle this is again fine since the tree structure is again unaffected. Now suppose interclone arcs form a directed cycle. We contract it into a supernode-of-clones, and then the modified costs leading to that supernode just get constants added to them, where these constants *do not depend* on that arc's start-node, only on its destination-node. This kind of modification cannot affect tree-shape (as can be seen from step 2 of the Edmonds algorithm). Q.E.D.

(3) Any arc entering the root R from a "better" winner B will form a directed cycle when combined with the tree-path of directed arcs $R \to \cdots \to B$. If we adjoint the new arc, and delete the least-margin victory in this cycle, then we get a higher-sum tree, contradicting it being a maxtree, *unless* B is "immune to complaints" because the $B \to R$ arc is the least-margin victory in that cycle. Q.E.D.

Because of these properties, it seems to us that (unless the greater complexity of the algorithm is regarded as an insuperable obstacle) maxtree voting should be adopted instead of River voting; we get the same properties but a better tree and hence presumably, in some statistical sense, a better election winner.

6.14 Keener's eigenvector system (1993) [41][24]

► The U-matrix has an eigenvector $\vec{x}$ with a positive real eigenvalue which is, in fact, the largest-modulus eigenvalue. (By a well known theorem of Perron & Frobenius. This eigenvector is most conveniently found via the "power method.") It consists entirely of nonnegative real numbers. The winner corresponds to the largest entry of this eigenvector. (Note: $\vec{x}$ also provides a natural *ordering* of the candidates.)

This may also be thought of in the following way. Consider the following Markov Chain on the N candidates. You are sitting on candidate A. You select a random voter v and a random candidate B. If v prefers B to A then you jump to candidate B, otherwise you stay with A. Keep doing this process forever. You tend to sit on better candidates for a larger fraction of the time than poor ones. The fraction of the time you will be at candidate n is x_n.

???ACTUALLY NOT. TO REALLY MAKE THAT MARKOV-DESCRIPTION TRUE YOU NEED TO ADD TO THE DIAGONAL ENTRIES OF THE MATRIX, ENOUGH TO MAKE ALL THE COLUMN SUMS BE EQUAL TO THE #VOTERS. SEE MY OTHER PAPER ON "SINKHORN" FOR THE CORRECT DOPE. THIS AND THE NEXT SECTION HAVE TO BE FIXED/UPDATED BASED ON THAT OTHER PAPER.

Example: In the situation of table 6.5, the U-matrix is the left-hand matrix below:

$$\begin{pmatrix} 0 & 11 & 5 \\ 8 & 0 & 13 \\ 14 & 6 & 0 \end{pmatrix} \qquad \begin{pmatrix} 0 & 11 & 11 \\ 8 & 0 & 13 \\ 8 & 6 & 0 \end{pmatrix} \tag{24}$$

and alteration of the 6 $C > A > B$ votes to $A > C > B$ changes it to the right-hand matrix (in which A is a Condorcet winner). The eigenvectors are respectively

$$(0.301, \underbrace{0.361}_{B \text{ wins}}, 0.338) \text{ and } (\underbrace{0.372}_{A \text{ wins}}, 0.353, 0.275) \tag{25}$$

with respective eigenvalues 18.841 and 18.551. ▲

Example: When a Condorcet-winner exists, Keener does not necessarily elect it. In the 4-candidate election among $\{A, B, C, D\}$ with the following U-matrix,

$$U = \begin{pmatrix} 0 & 11 & 11 & 11 \\ 9 & 0 & 20 & 20 \\ 9 & 0 & 0 & 20 \\ 9 & 0 & 0 & 0 \end{pmatrix} \tag{26}$$

the Condorcet-Winner is A but B is the Keener winner because the Frobenius eigenvector is $(0.309, 0.372, 0.205, 0.113)$ and B's entry 0.372 is maximum. ▲

However, Keener is incapable of electing a Condorcet-loser.

Example: On the left is a U-matrix (e.g. B beats A pairwise 19 votes to 1 vote):

$$\begin{pmatrix} 0 & 19 & 19 & 17 \\ 1 & 0 & 17 & 2 \\ 1 & 3 & 0 & 2 \\ 3 & 18 & 18 & 0 \end{pmatrix} \qquad \begin{pmatrix} 0 & 20 & 20 & 18 \\ 1 & 0 & 18 & 3 \\ 1 & 3 & 0 & 3 \\ 3 & 18 & 18 & 0 \end{pmatrix} \tag{27}$$

On the right is the U-matrix after adding an extra vote $A > B > C > D$. The Frobenius eigenvector of the left matrix is $(0.494, 0.134, 0.081, 0.292)$ and of the right matrix is $(0.485, 0.145, 0.088, 0.282)$. In both cases A is the winner, but A's eigenvector entry has diminished in the second case, which is somewhat "paradoxical." Because A still wins, this is not a failure of "add-top" for Keener eigenvector voting, but it is a failure in a weaker sense. ▲

Does Keener eigenvector voting obey "monotonicity"? A computer search of millions of random U-matrices failed to find any example of a monotonicity failure (even in this extremely weak form: switching an adjacent preference $A > B$ to $B > A$ in a vote causes A's eigenvector entry to increase or B's to decrease). It also failed to find any example of a (genuine) add-top failure (or a "no-show paradox") in Keener. That suggests that such failures are either impossible, or else so rare that they are not worth worrying about in practice.

The following theorem goes in the direction of, but seems inadequate to prove, the desired result.

Theorem 1 (Markov chain monotonicity). *Let $p(X, Y)$ and $q(X, Y)$ be irreducible Markov chain transition matrices on the same finite state space. Suppose there exist three distinct states A, B, C such that*

$$p(A, B) < q(A, B), \quad p(A, C) > q(A, C) \tag{28}$$

and that $p(X, Y) = q(X, Y)$ for all other entries (X, Y). Let $\pi_p(\cdot)$ and $\pi_q(\cdot)$ be the stationary distributions. Then $\pi_q(B) > \pi_p(B)$.

Proof (by David Aldous): The key realization is that the expected time between returns to B is $1/\pi(B)$, because if it were anything else, then by the law of large numbers, after a very long time the number of returns would, with probability$\to 1$, tend to a different number than it should, forcing an incorrect occupancy probability $\pi(B)$.

So it suffices to prove that the expected return time to B is smaller in the q-chain. This is fairly obvious. There are two kinds of return paths to B, those that go through A and those that do not. The latter kind are not affected.

The former kind return faster to B after the alteration since it takes 1 step from A to reach B directly, but ≥ 2 steps to reach B through C, and the probability of going directly to B increases, whereas the probability of going to C decreases (and all else stays the same) so the net effect on expected return time to B starting at A is

$$(1-c)\,\mathrm{d}p \quad \text{for some} \quad c \geq 2 \tag{29}$$

which is negative. Q.E.D.

Unfortunately the power of this kind of reasoning is very limited. For example, the following question appears still to be open: suppose we increase $p(A,B)$ and decrease $p(B,A)$ (again using some 3rd and/or 4th states to compensate). Then: does $\pi(B)-\pi(A)$ increase?

Keener is not "immune to clones" as the following example demonstrates. The U-matrix on the left

$$\begin{pmatrix} 0 & 2 & 1 \\ 1 & 0 & 2 \\ 2 & 1 & 0 \end{pmatrix} \qquad \begin{pmatrix} 0 & 2 & 1 & 1 & 1 \\ 1 & 0 & 2 & 2 & 2 \\ 2 & 1 & 0 & 2 & 1 \\ 2 & 1 & 1 & 0 & 2 \\ 2 & 1 & 2 & 1 & 0 \end{pmatrix} \tag{30}$$

represents a Condorcet cycle situation with the three candidates A, B, C tied. It changes to the matrix on the right upon splitting C into three cloned candidates C_1, C_2, C_3 themselves forming a tied Condorcet cycle. The eigenvectors are

$$(1,1,1)/3 \text{ and } (0.177,\, 0.229,\, 0.198, 0.198, 0.198) \tag{31}$$

respectively so that the cloning of C breaks the tie[28] and causes B to become the winner. However, somehow this does not seem very serious because in 1-on-1 contests, cloning the candidates does not seem to affect Keener.

This eigenvector scheme was originally invented by J.P.Keener [41] who indeed observed more generally that it could be used to "rank football teams." Indeed, Keener has the advantage that not only does it output a *ranking* of the candidates, it produces real numbers *quantifying* their strengths. Dwork et al [24] then independently re-invented the same idea, without noting Keener's work. Both Keener and Dwork et al also considered several other natural Markov chains and/or matrices and explained how they too could be used for voting and ranking.

6.15 Sinkhorn voting (2005)

A criticism of Keener eigenvector voting is: why is it based on the Perron-Frobenius eigenvector of some matrix A, instead of A^T???

Sinkhorn voting is a new kind of voting I invented which avoids that criticism. It relies on the theorem [22] that, for any $N\times N$ matrix A with positive entries there exist a unique[29] pair (R,C) of diagonal $N\times N$ matrices such that RAC is doubly stochastic. A $N\times N$ matrix X is "doubly stochastic" if $X_{ij}\geq 0$ and $\sum_{i=1}^{N} X_{ij} = \sum_{j=1}^{N} X_{ij} = 1$. In the case where all the entries are positive, Sinkhorn proved that this (R,C) pair of row and column scalings may be got by the "Sinkhorn iteration" (which he proved converges), which consists of repeatedly and alternately scaling the rows and columns of A to make those row and column sums equal 1.

► Sinkhornize $U+J$ (where J is the "all 1" matrix $J_{ab}=1$) and rank the candidates in the same order as the ordering of the entries (all of which automatically are positive and real) of the resulting diagonal matrix CR^{-1}.

Like Keener voting, Sinkhorn voting also has an interpretation in terms of a Markov chain. The ij entry of the matrix RAC (since it is doubly stochastic) is precisely the transition probability $j\to i$ for a Markov chain whose unique stationary distribution is *uniform.* The iith entries of R and C may be thought of as win and loss "amplification factors" required to cause candidate i to have the same stationary probability as all the other candidates, in a Markov chain of the same structure as U.

6.16 Simpson-Kramer min-max system [42]

► The candidate W with the minimum $\max_{A\neq W} M_{AW}$ wins. (Mentioned in [46]. Also mentioned on p.104 of [1] as "G.Köhler's dual method" for the single-winner special case.)

In other words: each candidate's score is his greatest margin of defeat in a pairwise contest. (Nonpositive if undefeated.) The candidate with the lowest score wins.

6.17 IRV (Instant Runoff Voting) [77][39]

This is thee single-winner special case of the Hare/Droop STV (Single transferable vote) system. Thomas Hare (1806-1891) was an English solicitor who was involved in the theory and advocacy of election methods.

► The candidate with the fewest top-rankings is eliminated. We then erase that candidate from all preference orderings and continue on. Once $N-1$ candidates have been eliminated, the remaining one wins.

#voters	their vote
7	$A>B>C>D$
6	$B>A>C>D$
5	$C>B>A>D$
3	$D>C>B>A$

Figure 6.11. 21-voter IRV monotonicity-failure example [7]. First round counts of top-rank votes: $A=7$, $B=6$, $C=5$, $D=3$, so eliminate D. Second round counts: $A=7$, $B=6$, $C=8$, so eliminate B. Final round: A wins over

[28] And of course, by slightly perturbing this example we can cause the original scenario to have no ties and to have anybody we want as winner.
[29] Up to an arbitrary scaling $R\to kR$, $C\to C/k$.

C, 13-to-8. (This is despite the fact that B is a Condorcet-Winner.)

But if the last 3 voters lift A from bottom to top (changing vote $D > C > B > A$ to $A > D > C > B$) then the election instead proceeds thus: First round counts: $A = 10$, $B = 6$, $C = 5$, $D = 0$, so eliminate D. Second round: eliminate C. Final round: B wins over A, 11-to-10. (B is still the Condorcet-winner.)

Also if the last 3 voters had ranked D first but refused to say more (i.e. refused to provide their 2nd, 3rd, and 4th choices), then B would have won (which those voters prefer over A). That illustrates the fact that in IRV, voters can be motivated to refuse to rank-order some of the candidates, thus defeating IRV's purpose of garnering ordering information from the voters. ▲

Table 6.11 is a severe example of "non-monotonicity" in IRV showing that ranking somebody top instead of bottom can cause them to *lose*!

Further examples may be contrived to cause IRV to exhibit all sorts of bizarre behavior, Table 6.12 gives a fairly plausible example constructed by Bart Ingles demonstrating both a "no-show paradox" and a monotonicity violation in an IRV election.

#voters	their vote
900	$A > C > B$
300	$C > A > B$
300	$C > B > A$
575	$B > C > A$

Figure 6.12. Plausible nonmonotonic-IRV example by Bart Ingles: the "centrist" candidate C is supported by voters whose second choice is evenly split among the two "extreme" candidates A and B. B is eliminated and then C wins by 1175 votes to A's 900. But if an additional 50 "absentee" votes $B > C > A$ appear, then C (although still the Condorcet-Winner by a huge margin) is eliminated and A wins by 1200 votes to B's 925. These additional voters would have been better off not voting (a "no show paradox") since their votes caused their most-hated candidate to be elected. If these additional 50 votes were changed to $A > B > C$ then B is eliminated and C wins 1175 votes to A's 950. This is a monotonicity violation: Altering the 50 votes to rank A top and C bottom caused A to lose and C to win! (This example also applies to P+I voting, since P+I and IRV are the same thing in elections with ≤ 3 candidates.) ▲

Despite such examples, in practice IRV tends to have one significant advantage: human voters empirically usually find IRV *difficult to manipulate drastically*, either by strategically lying in their votes or by manipulating the nomination process [10][11]. This is not just a human perception: it is also supported by objective evidence.

system	% manipulable
IRV	23-51
Plurality	49-100
Borda	83-100
Coombs	51-100

Figure 6.13. Manipulability. The percentages of "manipulable" elections with either 21 or 1000 voters, with either "impartial" voters (all vote-types random, independent, and equally likely) or with votes based on a randomized "spatial model." Summarizes data from a large Monte-Carlo study [10] under 4 kinds of voting, with "manipulability" determined by linear programming. IRV also tends to require larger coalitions of colluding manipulative voters than the other 3 systems. (Chamberlin [10] did not study Approval (§7.7) and Range voting but if he had, he probably would be found them even more manipulation-resistant than IRV.) ▲

Another interesting property obeyed by IRV (pointed out by James Green-Armytage) is

> **Dominant mutual third:** If a more than a third of the voters rank (in any order) the members of a subset S of candidates above all others, and all the members of S pairwise beat all the non-members; then the winner must come from S.

That is because the IRV election will, after eliminations and vote-transfers, reduce to a 2-man contest between an S-member and somebody else. (This is a weakened form of the Condorcet or Smith Set property.)

#voters	their vote
339	$(\overline{A, B, C, \dots, Y})z(\overline{a, b, c, \dots, y})$
171	$(\overline{a, b, c, \dots, y})(\overline{A, B, C, \dots, Y})z$
490	z $(\overline{A, B, C, \dots, Y, a, b, c \dots, y})$

Figure 6.14. 1000-voter example IRV election illustrating "dominant mutual third." The overlines denote random permutations (with different randomness for each voter among the 339, etc); the votes are candidate lists in descending order of preference. In this situation if the "random" orderings actually yield all their orderings *exactly* equally frequently so we get *exact ties*, and if we are using "nondiscriminatory" IRV rules where when we eliminate a candidate, we simultaneously eliminate all the ones he's tied with (as opposed to any arbitrary tie break) simultaneously, then the election proceeds thus:

1. eliminate $a, b, c, \dots, y$
2. eliminate $A, B, C, \dots, Y$
3. z wins, contradicting the dominant mutual third property.

But if the randomness merely causes *inexact* ties (or if all ties are randomly *broken*), then $a, \dots, y$ and $A, \dots, Y$ candidates keep getting eliminated until it is down to just three remaining: one from $a, \dots, y$, one from $A, \dots, Y$, versus z, and then the $a, \dots, y$ candidate loses, and then in the final round with the $A, \dots, Y$ candidate pitted against z, he wins and z loses. This confirms the claim of dominant mutual third, while illustrating the need for an *asterisk* about ties. ▲

IRV is "immune to clones" (9 of §9), exhibits "later-no-harm" (9 of §9), and is one of the few preference-ranking voting systems unhurt by the devastating DH3 pathology of §6.3. Further theoretical support for this perception of manipulation-resistance was provided by Bartholdi and Orlin's proof [4] that it is NP-hard [29] to determine how to change your vote in IRV in order to change the winner.

While this was a cute idea, do not be deluded into thinking such NP-hardness results necessarily have anything to do with building a good voting system. This is because none of these NP-hardness results hold in the limit $V \to \infty$ (large number of voters) with the number N of candidates held fixed (nor even with $N = O(\log V)$). Indeed, in this limit (which is apparently the one relevant for elections in which humans participate!) these computational tasks are easy, i.e. linear or even very sublinear time (or, more generally, in P). Also, even without a bound on N, it is usually still easy to think of a dishonest vote which seems to be more strategic than the true honest vote, despite difficulty in finding the *optimal* way to be dishonest. Thus, dishonesty is not necessarily prevented or discouraged by NP-hardness. Furthermore, the major political parties will have access to computers and hence will be capable of determining (by exhaustive search), and advertising, optimal voter strategies.

So what would be more important than NP-hardness (a concept that concerns worst-case input) would be some kind of *average*-case hardness result, and a result *not* about optimal strategy but rather about strategies significantly improving on honesty. But no such result is known.

My personal feeling is that it usually is easy to think of a probably-productive way to be dishonest in IRV, but it is usually hard to devise scenarios in which pervasive strategic dishonesty of this type causes *drastic* pathologies.

One final remark about IRV: It is "immune to vote splitting" in the following (fairly weak) sense. Suppose the electorate (and the candidates) consists of various disjoint "camps," where every voter in camp j ranks all type-j candidates ahead of all non-j candidates. Then, regardless of how the camps "split their votes" among their-type candidates, *if* there exists a camp containing $> 50\%$ of the voters, then *some* candidate from that camp is guaranteed to win the election.[30] This is not a terribly strong property and it is not unique to IRV. For example, it is obeyed by every voting system preprocessed using the "Smith set" (§6.28), and (since their winner is a Smith-set member) by Schulze beatpath, Heitzig River, Tideman ranked pairs, Copeland, Nanson, and Raynaud.[31]

(Most weighted positional systems, e.g. Borda, disobey it, so in that sense IRV is superior to them.)

"Approval voting" (§7.7) also obeys this property provided camp members approve of candidates from their camp and disapprove of all others, and indeed AV then obeys the stronger property that a candidate from the most-numerous camp then must win.

6.18 Loring's and Cretney's Condorcet-IRV systems

Robert Loring liked the empirical resistance of IRV to strategic manipulation, but did not like the fact that IRV can fail to elect Condorcet-Winners, as illustrated in table 6.12. Indeed much more dramatic examples are possible, see table 6.15.

#voters	their vote	simplified
50	$A > B > C > D > E$	$A > B$
51	$B > A > C > D > E$	B
100	$C > D > B > E > A$	$C > D$
53	$D > E > C > B > A$	D
49	$E > D > C > B > A$	$E > D$

Figure 6.15. IRV election example by Mike Ossipoff. The centrist candidate C is the favorite of far more voters than anybody else, and not only is the Condorcet-Winner, but in fact would win a head-to-head election versus any other candidate by approximately *a 2:1 margin*.[32] Hence almost every ranked-ballot voting system would elect C. But IRV elects D! (IRV only examines the preference relations in the "simplified" votes, and ignores the others.) This scenario is quite likely to arise in practice thanks to the prevalence of 1-dimensional "spatial voting." ▲

Therefore, Loring suggested this – hopefully superior – method:

► If there is a Condorcet-Winner, he wins. Otherwise, use IRV.

Meanwhile Blake Cretney suggested this one:

► Continue IRV-type eliminations until a Condorcet-winner exists among the remaining candidates, then it wins.

Yet another idea of this ilk would be, each round, to eliminate every candidate not in the "Smith set" (§6.28), and if that did nothing, then eliminate whoever has the fewest top-rank votes.

Loring is immune to the DH3 pathology of §6.3 *except* in the rarer and less damaging case when there is a Condorcet-Winner among the "mediocrities" (but: this unfortunately is forced if there is only a *single* mediocrity, as in table 6.4, or only *two* of them). Cretney, however, is not immune.

Loring and Cretney's systems still are subject to essentially all the bizarre pathologies that IRV suffers from (because we may simply add one extra candidate, who has no chance of winning and whose sole purpose is to prevent a Condorcet-Winner from existing, to any nasty IRV example) and still enjoys much of IRV's empirical manipulation-resistance.

6.19 BTR-IRV: Bottom Two Runoff IRV

The proponents of this scheme[33] suggest pronouncing BTR as "better." BTR-IRV is the same as standard IRV, except that

[30] Proof sketch: at some point a top-camp candidate will have $> 50\%$ of the top-rank votes (after transfers) and hence can never be eliminated.] *But* if no camp contains $> 50\%$ of the voters, then it is not the case that a candidate from the most numerous camp must win. E.g, camp A could have 35% and camps B and C 32% and 33% of the voters, but a type-C candidate could win if many of the B-voters transfer their allegiance to him after all B-candidates and appropriate other candidates are eliminated – and whether this happens *does* depend on how the B-voters [and A- and C-voters] split their votes among the various A and C candidates.

[31] It also is obeyed by Woodall-DAC (even though Woodall-DAC does not obey the Smith-set property and can fail to elect a Condorcet winner). Proof: That is because that camp A's candidates are a $> 50\%$-acquiesced set, whereas any set S containing some non-A candidate and avoiding some A-camp candidate, is $< 50\%$-acquiesced. Therefore as Woodall goes down the sets it will refuse to use S. Q.E.D.

[32] And the reader may enjoy this easy exercise: construct a scenario in which the Condorcet Winner beats each other candidate pairwise by a 99:1 margin, but nevertheless is eliminated in the very first round of IRV voting.

[33] BTR-IRV was invented by Rob LeGrand and also the BTR idea was suggested to me for use in multiwinner voting systems by Jan Kok.

when a candidate is eliminated, he is chosen by an "instant runoff" between the two candidates top-ranked by the fewest voters. This runoff examines *all* of the ballots to determine which of the bottom two candidates is least preferred by the voters.

BTR-IRV is actually a Condorcet method, because it will choose the Condorcet Winner, if there is one, as the winner of the election. BTR-IRV thus provides the advantages of Condorcet in IRV "packaging." Unfortunately BTR-IRV suffers from the DH3 pathology of table 6.4, suffers "add-top failure" as in table 6.16, and suffers "later harm" (table 9.2), whereas plain IRV has none of those deficiencies.[34]

#voters	their vote
3	$A > D > B > C$
3	$A > D > C > B$
4	$B > C > A > D$
5	$D > B > C > A$
4	$C > A > B > D$
6	$A > C > B > D$

Figure 6.16. BTR-IRV add-top failure. With the 19 voters above the line, A wins (C, B, and D are eliminated in that order). Adding the 6 voters below the line, all of whom top-rank A, causes C to become the winner (and Condorcet-winner). ▲

6.20 Coombs STV system (1954) [14]

Clyde H. Coombs (1912-1988) was an American psychologist.

► The candidate with the most bottom-rankings is eliminated. We then erase that candidate from all preference orderings and continue on. Once $N-1$ candidates have been eliminated, the remaining one wins.

This is a very bad system in the presence of strategic voters, because they will rank their least-liked among the favorites artificially "last." This will cause all the favorites to be eliminated in early rounds, causing the winners to consist entirely of unknown "dark horse" candidates.

#voters	their vote
4	$A > B > C$
4	$B > C > A$
5	$C > A > B$
2	$C > B > A$

Figure 6.17. 15-voter Coombs "no-show paradox" example. B wins 8 to 7 after A is eliminated. If the two $C > B > A$ voters had not shown up, then B would have been eliminated whereupon their first choice, C, would have won by 9 to 4. ▲

(i)

#voters	their vote
2	$B > A > C > D$
3	$D > B > A > C$
4	$C > A > D > B$

(ii)

the new vote	elim order	winner
$D > B > C > A$	B, D, A	C
$D > A > B > C$	B, D, C	A
$D > A > C > B$	B, D, C	A
$D > C > A > B$	B, D, A	C
$D > C > B > A$	B, D, A	C

Figure 6.18. 9-voter Coombs "favorite betrayal" example.

(i) Coombs eliminates B, then D, then A so C wins. If the 3 $D > B > A > C$ voters change their vote to $B > A > C > D$ ("maximally betraying" D) then D, C, A are successively eliminated, so B wins, an outcome those voters prefer.

(ii) If those 3 voters had instead changed their votes to anything else with D still top-ranked, then a worse candidate (from their point of view) would have been elected, hence this betrayal of D was strategically *essential*. ▲

6.21 Nanson-Baldwin elimination (1882) [55]

Edward John Nanson (1850-1936) was a high honors graduate in math from Trinity College Cambridge, and later became a professor of mathematics at the University of Melbourne in 1875.

► The candidate with the least Borda count is eliminated.[35] We then erase that candidate from all preference orderings (then recalculate the Borda scores, which may be done from the M-matrix with one row and column removed, as in §6.3) and continue on. Once $N-1$ candidates have been eliminated, the remaining one wins.

Nanson's point was that this always elects a Condorcet-Winner if one exists (and similarly can never elect a Condorcet-loser), because of the lemma that a Condorcet-Winner must have above-average Borda score.

Nanson unfortunately does not share IRV's resistance to strategic manipulation: it exhibits a DH3 pathology like Borda (§6.3) in the presence of strategic voters – a race between 3 favorites and 10 mediocrities can easily feature early elimination of all three favorites and consequent guaranteed election of a "dark horse."

#voters	their vote
8	$A > B > C$
5	$C > A > B$
5	$B > C > A$
2	$C > B > A$
2	$A > C > B$

[34] For definitions of such properties as "later harm" see §9. Since these disadvantages may outweigh its advantages, it is not clear BTR-IRV is better than ordinary IRV, despite the pronunciation tip! Still, BTR-IRV seems superior to Loring and Cretney.

[35] In Nanson's original variant, which has the same properties but is inequivalent, we save time by *simultaneously* eliminating *all* candidates with below-average Borda counts. The Baldwin one-at-a-time elimination variant was adopted in 1926 by the University of Melbourne for its internal elections. Rob LeGrand suggests that Nanson's original method is superior to Baldwin's (at least with honest voters) because more of its elimination decisions are made on the basis of more information. At any stage in either method, if a candidate is recognized to be a Condorcet winner, then we may immediately elect him and terminate the procedure, which also saves time. We also remark that instead of eliminating the candidate with least Borda score, equivalently Nanson may be viewed as eliminating the candidate with "lowest center of mass" position inside the ranked ballots.

Figure 6.19. 20+2-voter Nanson example. (Due to P.C.Fishburn and/or J.T.Richelson.)

Among the first 20 voters the Borda scores are $A = 21$, $B = 20$, $C = 19$ so C is eliminated whereupon A wins.

If the 2 $C > B > A$ voters change to $C > A > B$ then $A = 19$, $B = 18$, $C = 19$ so B is eliminated whereupon C wins by 12 to 8 over A. This is a "monotonicity" violation: a change in favor of A caused A to lose.

Without that change, consider what happens if the 2 additional $A > C > B$ voters show up. Then $A = 25$, $B = 20$, $C = 21$ so B is eliminated whereupon C wins. This is an "add-top" violation: addition of A-top voters caused A to lose. ▲

On the other hand, Nanson has the significant algorithmic and communication advantage over most elimination systems (e.g. IRV, BTR-IRV, and Coombs) that it may be implemented efficiently in parallel by computation of subtotals of the M matrix in subdistricts, with *only* transmission *of those subtotals* to the central tabulating agency, being required. This is because the Nanson-winner may be computed solely from the M matrix. (The same "M-only" trick works for Rouse, Raynaud, Arrow-Raynaud, Schulze, Tideman, etc.) Incidentally, by updating Borda scores after eliminations rather than recomputing them, this may be done in $O(N)$ time per update, i.e. $O(N^2)$ steps in total.

6.22 Rouse's elimination method

► Rouse is like Nanson-Baldwin but with an extra level of recursion: it successively pseudo-eliminates the candidate with the highest Borda score until one is left, then it genuinely-eliminates that one from the original list; this step is repeated until a single candidate is left. (There are again two – Nanson and Baldwin – variants.)

This was proposed by web-site designer Michael Rouse in 2001. (But this still does not cure Nanson's DH3 pathology, and I find it very annoying to use Rouse manually, because of the large amount of work for the tallier.)

#voters	their vote
5	$C > A > B$
4	$A > B > C$
2	$B > C > A$
2	$A > C > B$

Figure 6.20. 11+2-voter Rouse "add-top" failure example. Among the first 11 voters the Borda scores are $A = 13$, $B = 8$, $C = 12$ so we pseudo-eliminate A, whereupon C is genuinely eliminated. Then A wins over B.

Now consider what happens if the 2 additional $A > C > B$ voters show up. Then the Borda scores are $A = 17$, $B = 8$, $C = 14$ so we pseudo-eliminate A, whereupon B is genuinely eliminated. Then C wins over A. This is an "add-top" violation: addition of A-top voters caused A to lose.

(This also illustrates "later harm": if the 2 $A > C > B$ voters had changed their votes to $A > B > C$ then C would be genuinely eliminated and A would win. Thus these voters' "later" preference of $C > B$ "harmed" A.) ▲

6.23 H.Raynaud's elimination method

► The candidate who suffered the largest-margin pairwise defeat (B such that M_{AB} is the maximum entry of the M matrix) is eliminated. We then erase that candidate from all preference orderings (this may be accomplished by erasing the Bth row and column of the M matrix) and continue on. Once $N - 1$ candidates have been eliminated, the remaining one wins.

#voters	their vote
7	$A > B > C$
7	$B > C > A$
6	$C > A > B$

Figure 6.21. Example of "add-top failure" in Raynaud (by Chris Benham) The defeats and their margins are $A > B$ 13-7 (6), $B > C$ 14-6 (8), $C > A$ 13-7 (6). Hence C is eliminated whereupon A wins. However, if we add three $A > C > B$ ballots the winner changes to C (because the directions of the defeats are unchanged, but now B has the worst defeat and hence is first-eliminated). ▲

6.24 Arrow-Raynaud pairwise elimination

► Eliminate the "least-convincing victor" candidate A associated with $\min_A \max_B M_{AB}$. Continue eliminating until only one candidate remains: the winner. (Mentioned in [1] p.105; this is the single-winner special case.)

In other words: each candidate's score is the largest margin by which he wins a pairwise contest. (Nonpositive if never won.) We repeatedly eliminate the candidate with the smallest score.

One glaring defect of Arrow-Raynaud is that it can eliminate a Condorcet-Winner – even one who is a Super-Majority-Winner, i.e. whom 99% of voters rank top – in the very first round!

#voters	their vote
$N-1$	$W > A_1 > A_2 > \cdots > A_{N-2} > A_{N-1}$
$N-1$	$W > A_2 > A_3 > \cdots > A_{N-1} > A_1$
$N-1$	$W > A_3 > A_4 > \cdots > A_1 > A_2$
...	
$N-1$	$W > A_{N-1} > A_1 > \cdots > A_{N-3} > A_{N-2}$
$1+\epsilon$	$A_1 > A_2 > \cdots > A_{N-2} > A_{N-1} > W$
$1+\epsilon$	$A_2 > A_3 > \cdots > A_{N-1} > A_1 > W$
$1+\epsilon$	$A_3 > A_4 > \cdots > A_1 > A_2 > W$
...	
$1+\epsilon$	$A_{N-1} > A_1 > \cdots > A_{N-3} > A_{N-2} > W$

Figure 6.22. W is ranked top by a fraction $(N-1)/(N+[N-1]\epsilon)$ of the voters where $\epsilon \to 0^+$. But he is eliminated in Arrow-Raynaud's first round. (Also, essentially, demonstrates failure of "later-no-harm.") ▲

Arrow-Raynaud has a few compensating virtues of greater resistance than Condorcet methods to the strategic vulnerabilities in table 6.4. (One of these virtues – shared by Simpson-Kramer – is: adding votes which rank A top, to a scenario in which A uniquely wins, cannot prevent A from winning; and adding votes ranking B bottom, to a scenario in which B did not win, cannot cause B to win.) But they seem nowhere near sufficient compensation.

#voters	their vote
3	$A > D > B > C$
3	$A > D > C > B$
4	$B > C > A > D$
5	$D > B > C > A$

Figure 6.23. 15-voter "no-show paradox" example for Arrow-Raynaud voting. In the first round, C's biggest victory (over A) is only by 3, which is a smaller margin than everybody else's biggest pairwise victory (A over D by 5, B over C by 9, D over B by 7). Hence C is eliminated. In the second round, B's biggest victory (over A by 3) is smaller than everybody else's, so B is eliminated. In the final round, A wins by 10 to 5 over D.

Now add 4 extra votes $C > A > B > D$ to get a new 19-voter situation. In the first round, D's biggest victory (over C or B) is only by 3, which is a smaller margin than everybody else's biggest pairwise victory (A over D by 9, B over C by 5, C over A by 7). Hence D is eliminated. In the second round, A's biggest victory (over B by 1) is narrower than everybody else's, so A is eliminated. In the final round, B wins 12 to 7 over C.

Thus the 4 extra votes, despite ranking $A > B$, caused A to lose and B to win. ▲

6.25 Bucklin

► Bucklin tries to find a majority for some candidate by counting only the first-place votes. If no candidate has more than $V/2$ votes, all second-place votes are added to the count, then third-place, etc., until the candidate with the most votes has a number of votes exceeding $V/2$; he is the winner. At most $1 + \lfloor N/2 \rfloor$ rounds are required.

Bucklin was used in 7 US states starting in 1912 in votes for important offices, including gubernatorial races. It is mentioned in [19].

Bucklin resembles approval voting (§7.7) in the sense that if the last Bucklin round is the kth, then the same result would be obtained by a Bucklin election and an Approval election in which each voter "approved" all ($\le k$)-ranked candidates.

Bucklin also resembles plurality voting in some ways, but improves upon it in others – for example it is not capable of electing a Majority-Loser, although Plurality can. Bucklin suffers from many of the same "vote splitting" problems as does plain plurality: lack of "clone immunity," failure to always elect Condorcet-winners (table 6.3), and capability of electing Condorcet-losers (table 6.24).

#voters	their vote
4	$A > L > B > C$
4	$B > L > C > A$
4	$C > L > A > B$
3	$A > B > C > L$
3	$B > C > A > L$
3	$C > A > B > L$

Figure 6.24. Bucklin can elect Condorcet-loser. 21-voter example. First round: A, B, C tied with 7 votes. Second round: The Condorcet-loser L has 12 votes, exceeding the 21/2 required for election, while A, B, C each only have 10 votes. ▲

Favorite betrayal:[36] In table 1.1 the unique Bucklin winner is B. If the two $C > A > B$ voters had insincerely voted $A > C > B$ ("betraying their favorite" C) then the unique Bucklin winner would have been A, which from their point of view would have been an improvement.

Truncation: In the same example, if the 5 $A > B > C$ voters had truncated their votes by just saying "A" then A would have won. (I.e.: a voter's second-choice vote often can help to defeat his first-choice.)

The fact that in Bucklin, it can often be strategically advantageous to refuse to provide a second-choice vote – resulting in plurality-style voting and thwarting Bucklin's goal of causing voters to express full preference orderings – was the main reason that Bucklin was abandoned, eventually by all US states.

A 2-round Bucklin system (essentially) remains in use in the London Mayor's election (where it is called the "supplementary vote").

Tables 6.25 and 6.26 demonstrate some other Bucklin pathologies by Kevin Venzke and Markus Schulze respectively.

#voters	their vote
5	$A > B > C$
4	$B > C > A$
2	$C > A > B$

Figure 6.25. Bucklin "no-show paradox." For the first 9 votes, A is a majority favorite and wins in Bucklin round 1. The last 2 voters would be best off not showing up since their effect is to remove the majority favorite and cause their most-hated candidate B to win in round 2. ▲

#voters	their vote	#voters	their vote
19	$A > C > B$	4	$A > B > C$
20	$B > C > A$	5	$B > C > A$
1	$C > A > B$	6	$C > A > B$
1	$C > B > A$	—	—
1	$B > A > C$	4	$A > B > C$
1	$A > B > C$	5	$C > A > B$

Figure 6.26. Bucklin reversal and subdistrict-consistency paradoxes.
(left) C is the unique Bucklin winner. When all individual preferences are reversed, it still is.

(right) C is the unique Bucklin winner for either the first 15 or the last 9 votes, but for the combined 24-vote set, A is the unique Bucklin winner. ▲

[36] A more common kind of favorite betrayal in Bucklin voting is the same story as in plain Plurality voting – if one's Favorite has essentially no chance of winning, and we are convinced the winner is going to be either A or B by a first-round majority, then one's best strategic move is to abandon Favorite and vote for the best among $\{A, B\}$.

6.26 Woodall's DAC [83]

In 1997, Douglas Woodall introduced his Descending Acquiescing Coalitions (DAC) method, which probably takes the brass ring as the most complicated ranked-ballot voting method seriously proposed so far.

► A voter "acquiesces" to a set of candidates if he does not rank any candidate outside the set higher than any inside the set. (Every voter acquiesces to the full candidate-set.)

Sort all possible sets from most acquiescing voters to fewest. Going down the list, disqualify every candidate not found in each set (i.e. take set intersections) unless that would disqualify all remaining candidates (i.e. would result in the empty set). Continue until only one candidate is not disqualified; he is the winner.

Although it might seem that DAC requires exponential runtime because there are 2^N possible candidate-subsets, in fact there is a polynomial-time algorithm because the number of nonempty subsets that anybody acquiesces to,[37] is at most VN. We simply find all of them and store them in a suitable data structure such as a trie or hash table, as we do so computing the voter-count acquiescing to each one. With a slight amount of cleverness, this requires only $O(VN)$ steps. The sorting then may be done in $O(VN \lg V)$ steps and then the set-intersection requires $O(N)$ steps per intersection, i.e. at most $O(VN^2)$ steps total. If the number of candidates N is smaller than the number of bits in a computer word, then each set-intersection may be done via a word-wide bitwise-AND in only 1 step. Also, often only $O(V+N)$ set-intersections are required. In either case the runtime bound could be decreased to $O(VN \lg V)$.

Woodall proved DAC satisfies "participation,"[38] "majority-winner," most (but not all) monotonicity properties, and "later-no-help," although it fails "later-no-harm."

#voters	their vote
4	$A > B > D > C$
4	$B > C > D > A$
4	$C > A > D > B$
1	$D > B > A > C$
1	$D > C > B > A$
1	$D > A > C > B$

Figure 6.27. Woodall-DAC can elect a Condorcet-loser. In this 15-voter example, 5 voters acquiesce to each 3-set containing the Condorcet-Loser D (beating out all other 3-sets, 2-sets, and singleton sets, all with at most 4 acquiescers); therefore D wins. ▲

In the DH3 example in figure 6.4, if $x_1 = x_2 = 10$, $y_1 = z_2 = 11$, $y_2 = z_1 = 8$ then Woodall-DAC will elect the "dark horse" D. This proves Woodall can "fail the DH3 test," but it also is capable of passing that test, for example if $x_1 = x_2 = 10$, $y_1 = y_2 = z_1 = z_2 = 9$ then it elects A and spurns the Condorcet-winner D.

Woodall-DAC enjoys "clone-immunity." That is because if, in each candidate subset, a candidate is replaced by the full set of his clones, then that set has exactly the same acquiescence counts as it had before cloning, and the only way the procedure can delete a clone from a set (unless it is the full set of candidates) is to delete somebody else at the same time.

6.27 SPl: Sarvo-plurality (2004) [71]

► For each k ($1 \le k \le N$) let V_k be the number of voters who top-rank candidate k. As usual we also let U_{kj} be the number of voters who prefer k to j. Define

$$E_k \stackrel{\text{def}}{=} \frac{V_k}{V} \sum_{j \neq k} \left(\frac{V_j}{V - V_k} + \frac{V_j}{V - V_j} \right) U_{kj}. \tag{32}$$

The candidate W with the maximum E_W wins.

Favorite-Betrayal Example: In the situation of table 6.5, B wins because $E_B \approx 5.51$ is greater than $E_A \approx 3.76$ and $E_C \approx 3.28$. If the 6 $C > A > B$ voters insincerely switch to $A > C > B$ ("betraying their favorite" C) then A becomes the winner since $E_A = 11$ is greater than $E_B = 8$ and $E_C = 0$.

The "sarvo" methods [71] are designed to try to be immune to strategic voting, but as this example shows, that attempt is not completely successful.

6.28 Smith set

► The "Smith set" (introduced in the early 1970s? But Dodgson knew of it a century before?) is the smallest nonempty set of candidates such that each candidate A in the Smith set beats each candidate B not in it: $M_{AB} > 0$. This always is a nonempty set. If a Condorcet-Winner exists, then the Smith set is the singleton set consisting of that winner.

By starting with the Smith set, and then using some other voting procedure to decide among the Smith candidates (if there are more than one), we can get numerous interesting Smith-hybrid voting systems. Such hybrids will always elect a Condorcet-Winner if one exists. The Smith set can be *every* candidate (the same is true of the Fishburn and Banks sets below[39]) hence has little if any use as a stand-alone voting system.

Forest Simmons pointed out this interesting

Lemma 2. *Any rank-order-based voting method featuring "immunity to second place complaints" (I2PC; the candidate X who would win with the same votes but with the winner W removed from the election, should not pairwise-beat W) automatically must elect a Smith set member as winner.*

Proof: (by induction on the number of candidates, and ignoring the possibility of pairwise ties):

If there are fewer than two candidates there is nothing to prove.

If there are just two candidates, then I2PC means that the winner beats the only other candidate, and so is in the Smith set.

[37] Assuming voters provide full tie-free orderings of the candidates as their votes. The Woodall DAC procedure would in fact permit a vote merely to be a partial ordering of the candidates, in which case a voter could acquiesce to as many as 2^N subsets and I do not know of a subexponential algorithm.

[38] For definitions of such properties as "participation" (IP) see §9. Proof sketch: The whole effect of adding an extra ballot acquiescing to a set S is to increase the acquiescence counts for some sets X with $X \subseteq S$ or $S \subset X$. This cannot decrease S's chances of victory.

[39] Make one big Condorcet cycle.

If there are $n \geq 3$ candidates, then (by induction) withdrawal of W gives the win to a member X of the Smith set of the reduced election. But by I2PC, W beats X, so W has a "beat-path" to every candidate (except self), and so W cannot be a member of the complement of the Smith set and hence must be a member of the Smith set. Q.E.D.

6.29 Fishburn set [26]

If $M_{ZA} > 0$ implies $M_{ZB} > 0$, and for some X, $M_{AX} \geq 0$ and $M_{XB} > 0$, then we say that A "Fishburn dominates" B. In other words, A Fishburn dominates B if every candidate who pairwise beats A also beats B, and A beats or ties some candidate who beats B.

► The candidates that are not Fishburn dominated are the "Fishburn set." The Fishburn set is always nonempty and is always included inside the Smith set. If a Condorcet-Winner exists, then the Fishburn set is the singleton set consisting of that winner.

By starting with the Fishburn set, and then using some other voting procedure to decide among the Fishburn candidates (if there are more than one), we can get numerous interesting Fishburn-hybrid voting systems. These hybrids will have the advantage that they will always elect a Condorcet-Winner if one exists.

6.30 Banks set [3][52]

► Define "$B \gg A$" if and only if B pairwise defeats A and B wins among all C that pairwise defeat A. The "Banks set" is characterized as the set of candidates B which fulfil $B \gg A$ for at least one A.

Equivalently, the Banks set consists of the Condorcet-Winners of maximal candidate-subsets. That is, if C is the Condorcet-Winner within some candidate-subset X (with at least one element besides C), but there is no Condorcet-Winner in any larger candidate-subset Y (i.e. with $X \subset Y$), then C is in the Banks set.

If a Condorcet-Winner exists, then the Banks set is the singleton set consisting of that winner.

The Banks set is nonempty and contained in the Smith set.

By starting with the Banks set, and then using some other voting procedure to decide among the Banks candidates (if there are more than one), we can get numerous interesting Banks-hybrid voting systems. Such hybrids will always elect a Condorcet-Winner if one exists.

7 Systems in which each vote is a real N-vector

7.1 Dabagh "vote and a half" (1934) [19]

► Each voter awards 1 point to one candidate (presumably his favorite) and half a point to another (presumably his second choice). The candidate with the most points wins.

7.2 "Vote-for-and-against" (2004)

► Each voter awards 1 point to one candidate (presumably his favorite) and *negative* 1 point to another (presumably his most-hated). The candidate with the greatest point sum wins.

(Both Dabagh and for-and-against may be viewed as weighted positional systems. In the 3-candidate case Dabagh, for-and-against, and Borda are equivalent.)

Although the strategic thinking in this system is similar to Plurality ("must vote among the two frontrunners to avoid wasting my vote") and hence might be presumed to lead to 2-party domination – there is an interesting twist: One of the two "frontrunners" is always going to get a negative number of votes! Hence in a close race where a third-party candidate had a number of votes exceeding the margin between the perceived "frontrunners" (such as Bush vs. Gore in the US 2000 Presidential election, where Nader's votes exceeded the Bush-Gore margin) the third-party candidate could be elected. So it is an interesting question just what effects for-and-against voting would have on 2-party domination.

For-and-against voting occupies a unique place in the WP systems: it is the only one with $w_{\text{top}} = -w_{\text{bottom}}$ which is immune (in the absence of a multi-way tie) to the DH3 pathology of §6.3.

7.3 Signed: G.A.W.Boehm's "negative voting" (1976)

We hereby rename this "signed voting."

► Each voter awards *either* 1 point to one candidate (presumably his favorite) *or* negative 1 point to to one candidate (presumably his most-hated). The candidate with the greatest number of points wins.

7.4 Cumulative voting (continuum version)

► Each vote is a real N-vector every entry of which is nonnegative with all of the entries summing to 1. For example a legal vote would be $(0.4, 0.3, 0, 0.3)$ in a 4-candidate election, since $0.4+0.3+0.3 = 1$. The winner is the candidate corresponding to the maximum entry in the sum-vector $\vec{s}$.

(C.L.Dodgson mentioned an integer version of this method, under the name "method of marks," in his first pamphlet. It is also mentioned in [43].)

Cumulative voting is not immune to candidate "cloning," as table 7.1 demonstrates.

#voters	their vote
2	$(0,1) \to (0,0,1)$
3	$(1,0) \to (1/2,1/2,0)$

Figure 7.1. The first candidate A wins the election by 3 votes to 2 over B. But after cloning, A_1 and A_2 each get 3/2 votes and lose to B. ▲

7.5 "Asset voting" (W.D.Smith 2004) [69]

► Same as the above cumulative voting system, except for choosing the winner. After we compute the sum-vector $\vec{s}$, we regard each s_n as the amount of an "asset" now owned by candidate n. The candidates now negotiate; any subset of them may redistribute their assets among themselves. After all negotiations and redistributions end, the candidate with the most assets wins. (Originally proposed as a multiwinner system, in which case the k topmost asset-rich candidates win, if we want k winners. Variants involve placing constraints on the negotiation procedure.) This is the only voting system mentioned here in which the *candidates* play an active role.

Tideman & Tullock's 1976 "perfect" scheme for voting-with-money [79]

In principle, the ultimate voting scheme would be "honest utility voting" in which each voter states the "utility" (measured in some common, agreed-upon units) of each possible candidate for him, and then the candidate with the greatest utility (summed over all of society) is elected.

Unfortunately, honest utility voting seems unachievable in practice, since (a) there are no common agreed-upon units, and (b) just one dishonest-strategic voter, by making some vast exaggeration, could control the election. (However, in certain unusual scenarios, such as where the voters are not dishonest humans, but rather honest *robots*, and in which utility *is* easily measured, this voting method would be best possible.)

Therefore (the common thinking before the 1970s was) honest utility voting is a pipe dream – an idealization of no practical interest. However in the 1970s several people realized, some independently, that, at least in a mathematical idealization of voters as rational economic money-maximizing animals, such a "perfect" voting system actually *is* achievable!

The initial idea (1961) was due to W.Vickrey [81] in a different context: auctioneering. Imagine some moderate number of bidders (e.g. 10-20) want to buy some expensive object.

Vickrey second-price auction protocol:

1: Each bidder privately estimates the true worth of the object, to him, in dollars.
2: Each bidder submits that secret estimate, in a sealed envelope, as his bid.
3: All the envelopes are opened. The winner is the one who submitted the greatest bid, *but* now he only *pays* the amount specified by the *second*-highest bidder. [*Modification suggested by Jan Kok:* if the top bid is T and the second-top bid S, then we can make the payment be $S + (T - S)X$ for any constant or randomly chosen real X with $0 \le X < 1$.]

Vickrey argued that in this scenario, *there is no strategic motivation for bidders to be dishonest in their bids*, and plenty of motivation for them to be honest.[40] This is also true in Jan Kok's modified version of Vickrey. Note that in this scheme it is important that the other bids be *secret*, since a bidder who knew the topmost external bid could *then* have motivation to be dishonest and bid 1 cent below it to minimize the profit of that enemy bidder. But it also is important that the bids ultimately be *revealed* since otherwise the auctioneer could cheat by pretending the second-top bid was much higher than it really was.

▶ Clarke-Groves-Tideman-Tullock public choice protocol:

1: Each voter as his (secret ballot) vote submits his private estimate of the true worth of each candidate (or election alternative) to him, in dollars.
2: The alternative (or candidate) with the greatest amount of money voted for it/him, wins.
3: We now re-examine all the votes, but now with the name of its voter-author unveiled on each. Suppose that vote "made a difference," i.e. caused the election result to differ from what it would have been with that vote removed. Namely, suppose without that vote-bid B, alternative Y would have won, but with it alternative Z wins. Suppose the minimum amount of money that vote could have been replaced by that still would have caused Z to win, is $M_Z > 0$. Then that voter now must pay a *fee* (the "Clarke tax") equal to M_Z. [Modified version: the payment is $M_Z + (B - M_Z)X$ for any constant or randomly chosen real X with $0 \le X < 1$.]

The Clarke-Groves-Tideman-Tullock **central claim** [13][37][79] is that there is no strategic motivation for voters to be dishonest in their votes, and plenty of motivation for them to be honest.[41] And if the votes indeed are honest, then this voting system therefore is "perfect" and always elects the best choice, as measured in (perceived) dollars, for all of society.[42] And this is also true for the Jan Kok modified version.

voter	Y	Z
voter #1	5	0
voter #2	5	0
voter #3	3	0
voter #4	0	4
voter #5	0	7

Figure 7.2. A Clarke-Tideman-Tullock 2-candidate election. Y wins by \$13 to \$11. Voters #1 and #2 then pay Clarke tax \$3 each (because without voter #1, Z would have won by \$3), and voter #3 pays \$1. The other voters pay nothing. ▲

Unique features. This is the only system (of those we survey) claiming to result in "perfection" and "voter honesty." It also is the only one that directly involves *money* (and therefore it cannot be used in abstract scenarios in which the "voters" actually are not money-owning economic entities).

Additional work. For some reason, interest in this scheme apparently died after 1984 and it was almost entirely forgotten. (Later note: actually it was not forgotten; a remarkable book by Bailey [2] uses it as a key ingredient.) One of the last works (besides Bailey) on the matter was by Tideman [76] who conducted experiments on actual use of the scheme. Specifically, Tideman paid several college fraternities to do their decisions by his process instead of plurality voting for about 1 year. Tideman's experiments failed to reach a convincing conclusion. The fraternity members (polled afterward) preferred plain voting to Tideman's method, 108-to-56, because of increased administrative work! But they preferred Tideman (112-to-42) *if* the issue-proposer was willing to pay a

[40] Bid too high? Risk paying more than it is worth. Bid too low? Risk not acquiring the object at cost maximally below its true worth to you.
[41] Bid too high? Risk paying more than the election $Y \to Z$ result-change you caused, actually was worth to you. Bid too low? Risk not acquiring the election result you want, even though you could have done so at cost below its true worth to you.
[42] Tideman and Tullock felt obliged to point out that nevertheless their method "would not cure cancer."

administration-fee to get his issue considered by the Tideman method – *but* that in fact never happened during the subsequent year.

Tideman's process yielded different decisions than plain voting in about 10% of the fraternity decisions. This led to about 2.25% extra perceived dollar benefit. However the "Clarke taxes" paid were larger than that (namely 3.04%) so in some sense it *wasn't* worth it. However, frat houses are small – and in larger elections the Clarke taxes would be comparatively negligible. Finally, there may have been collusions or other effects (see criticism list below) which really caused CTT voting to perform even worse, and if so Tideman had no way to know it.

Criticisms. Unfortunately there are several reasons that Clarke-Tideman-Tullock, in practical use in large elections, would in fact fall considerably short of perfection.

1. In the large elections typical in governments, the probability is extremely small that any one vote will "make a difference." (In all the ≈ 2000 US presidential and senatorial elections so far, there has *never* been a case where any single vote has affected the outcome.) This *contrasts* mightily with the situation in most auctions (typically there are 10-20 bidders so each has a reasonable chance – 1/10 or 1/20 – of winning) and in most small fraternity-house elections. And that contrast, as we shall see next paragraph, does matter.

2. In auctioneering and fraternity houses, the problem of bidders not paying, is a minor problem. But in nationwide elections in the *very rare* instances when payment was required, there might be a severe nonpayment problem. Even if all voters were honest, some would die during the election. And if only one election in a million required the typical voter to pay, that might encourage a culture of dishonest exaggerated votes, followed on the rare occasions every 3000 years (!?) where payment actually *was* required, by a culture of tax-evasion. How could voters be forced to pay? Throwing the election the other way in response to nonpayment would not work because the other side also might not pay! This all could lead to a nightmare scenario giving a new meaning to the term "election fraud."

The problem here is the *combination* of 1 and 2. To defeat it we could require each voter to submit his payment in "escrow" (to be refunded in most cases). That would avoid the nonpayment problem. However, it would cause a new problem: the cost in time, hassle, and interrupted investments to place one's money in an escrow account – a cost greatly and unfairly varying from person to person – would vastly exceed the actual expected value of the Clarke tax. (If the Clarke tax must be paid 1 time in 40,000 then even missing out on one day of interest even at only 1% annually, would exceed the expected value of the Clarke tax.) Furthermore, the escrow account set-ups and necessary financial motion might defeat the goal of *secret* voting. (And, as with Vickrey auctions, secrecy is *essential* to prove the CTT central claim on which everything rests.) Therefore, in practice the economic rationale of this sort of voting would be *dominated* by these distortionary effects and the CTT goal of a voting system in which undistorted true utility dominates would not be achieved at all.[43]

The point of this has been that certain "negligibly small" effects such as the cost of moving money and cost of voting, while they *are* negligibly small in the case of auction of an expensive object, are in fact the *most* important thing in the voting case. Margolis [48] made similar points. He observed that in a typical election situation the expected Clarke tax on a typical voter would be 10^{-5} of one penny, and hence expressed doubt that voters would bother to work out and use their optimum-utility votes. The transport, money-motion, and other transactin costs experienced by a typical voter would dwarf this and hence, rather than being "negligible" would in fact "dominate" the thinking of rational voters. The two problems of auctineering and voting might seem almost equivalent except that some numbers are changed, but the problem is that these numbers change so *vastly* that the approximations of certain effects as negligible, become completely invalid.

3. Most people would *not* agree utility is the same thing as money – even though (economist) Tullock may think it is! If alternative A leads to a rich man dying while alternative B kills five poor men, then CTT voting would choose alternative A. More generally this system might do whatever the rich and fanatical want, and might exhibit a systematic bias against poor people.

4. The Clarke taxes would in fact be paid into government funds which would then be used to reduce (ordinary) taxes. Hence in reality the true amount voters would effectively pay, would differ from the Clarke tax. That in practice would distort the system away from the ideal of employing "perfect utilities." (One nightmare: consider a candidate whose platform was "if elected, I will refund all Clarke taxes.")

5. Collusions: Suppose "Nixon" bids some enormous amount of money and "Agnew" also does, to make the Nixon-Agnew ticket win. Now, each of {Nixon, Agnew} alone did not change the election result because their bids were each vastly huger than everybody else in the country combined, so Nixon could argue "I would have been elected anyway thanks to Agnew's bid" therefore Nixon pays zero Clarke tax. Similarly Agnew also pays zero and both get elected for free. This pathology could reduce everything to 2-man teams of colluders each trying to "name the largest number they can" – which would be a ridiculous dysfunctional state of affairs.

However, this Nixon-Agnew criticism largely falls to the ground if *escrowed* vote moneys are demanded. That makes ridiculously large bids impossible. But that, as we said, leads to other difficulties. (Another related to #3 is: what if my utility exceeds the amount of money I am able to raise in cash form?[44])

[43] To minimize these effects, the government could give each voter a "free donation" of a certain small amount of money into his escrow account, provided he voted, and could pay interest on the escrow accounts. These moves would attempt to null-out transaction costs. However, there would unavoidably be remaining inequities and mismatches in that nulling-out attempt, which would usually still dominate the true expected utility estimates made by, and true "costs of voting" for, each voter.

[44] Consider abortion rights. Jill in losing her abortion rights, is losing a lot of expected dollars. The ratio of abortions to live births was approximately 32:100 in the 1990s USA. Suppose it costs $100,000 to birth and raise a child but only $300 to get an abortion (1990s dollars). Suppose Jill is going to have 3 children and hence (on average) one abortion. That means $100,000 of expected utility is at stake for her if there is a vote on a law that would make it impossible for her to abort. (Actually she could still try to get an abortion in a foreign country, or try to abandon

Here is another (new) idea to reduce the collusion problem: randomization. Suppose every vote-bid X is automatically *replaced*, before doing the election, with a *random number* in the interval $[0, 2X]$ with *mean* X. One possible probability distribution would be the two-mass one with either 0 or $2X$ chosen by a coin flip.[45]

This idea has the disadvantage that it adds extra "noise" to the election and that it can stick unlucky voters with a higher than-expected Clarke tax bill (although they could compensate by underbidding by a constant factor). But its advantage – and point – is that it causes the Nixon-Agnew collusion we just described to become *very risky*. Randomization causes your collusion-teammates to become effectively "untrustworthy" even if they actually were totally loyal.

Incidentally, Margolis [49] tried to argue using differential equations that the CTT scheme was the essentially *unique* one with its properties. Margolis unfortunately did not actually state an explicit theorem with proof, but suffice it to say that *Margolis was wrong* in the sense that we have just exhibited two different kinds (Jan Kok's X-generalization and our randomization generalization), constituting an infinite number of, ways to generalize CTT. So no, CTT is not unique at all, but yes, it is unique if we restrict to deterministic voting schemes in which voter payments are made as small as possible.

6. H. Margolis [48][50] offered another criticism: which is that voter "altruism" would invalidate CTT. That is, in many cases the amount of money at stake for each voter is small so that voter is willing to be altruistic and lose money to further what he considers to be a good cause. That seems to me to be an excellent and completely valid criticism: the only reason almost anybody votes at all, is "altruism" since it is totally economically irrational to vote considering the cost in time and transport to vote versus the low probability of that vote having an election-swinging effect. So altruism (aka economic irrationality) is not "negligible"; it is the "dominant" effect in voting! Given any such huge distortionary effect on utilities, CTT's claims of "perfection" fall to the ground.

In my view these criticisms are very serious and probably render the Clarke-Groves-Tideman-Tullock voting system practically useless (or at least uncompetitive enough to remove it from consideration) for large governmental elections.[46] CTT also is useless for (or at least dubious for) small elections such as in Tideman's frat house study (since for them the Clarke taxes exceed the benefit of switching to the system). However, for elections of *intermediate* size (500-50000 voters?) and in which utility *does* closely correspond with money (stockholder elections in corporations, with Clarke taxes donated to charity?) this system may make excellent sense. It might also make sense if future techno/political developments alter the nature of "money" so that large anonymous and undetectable monetary transfers become easy and cheap.

7.6 Median Rating

► Each alternative is given a score (for example from 0 to 100) by each voter. The alternative with the highest median score wins.

This method is ludicrous in the presence of strategic voters, since they will rate everybody either 0 or 100. Then each median therefore will be either 0 or 100, usually causing a vast number of ties, in which case the voting system has not actually accomplished anything.

7.7 Approval Voting [8]

Approval voting was invented by 6 political scientists[47] independently during 1968-1978, but they all were beaten to the punch by an amateur – artist, astronomer, and inventor Guy Ottewell.

► Each vote is a real N-vector every entry of which is in the 2-element set $\{0, 1\}$. For example a legal vote would be $(1, 0, 0, 1, 1)$ in a 5-candidate election. The winner is the candidate corresponding to the maximum entry in the sum-vector $\vec{s}$.

#voters	their vote
1	$A > B > L > C$
1	$B > C > L > A$
1	$C > A > L > B$

Figure 7.3. Approval voting can elect a Condorcet loser. All voters approve of their top 3 candidates. The Condorcet loser L is elected since he is the unique candidate with unanimous approval.[48] ▲

7.8 Range voting [68]

► Each vote is a real N-vector every entry of which is in the real interval $[0, 1]$. For example a legal vote would be $(0.4, 0.3, 0, 0.7, 1)$ in a 5-candidate election. The winner is the candidate corresponding to the maximum entry in the sum-vector $\vec{s}$.

This system in used by various organizations on the internet to rate the quality of movies, recipes, etc. A system very much like it has been used to select gold medal winners in Olympic events.

Range voting is *more expressive* than almost every voting system considered so far – permitting not only expression of preferences, but also of intensity of preferences – and also one of

her child for adoption, so it is less than \$100,000, but we shall ignore that.) But more utility is at stake the younger Jill is. However, the younger Jill is, the less likely she has \$100,000 to vote with! Now on the other side are the moralists who believe abortion is murder. From their view, it is worth all the money some fetus would ever be able to earn, to avoid being murdered. However... fetuses don't have any money and can't vote. So it appears both sides in this issue are going to regard the CTT voting system as unfairly biased against them and invalid. How do CTT-proponents answer them?

[45] The problem of making everybody confident the randomizations were genuine, independent, and unbiased, is a solvable cryptographic problem, see chapter ???.

[46] Also, CTT voting is unconstitutional in the USA under the 24th amendment.

[47] John Kellett, Kenneth Mott, Robert J. Weber, Steven J. Brams, Peter C. Fishburn, and Richard A. Morin. A system very much like AV is used by the United Nations Security Council to vote for Secretary General.

[48] However, in this situation, L, despite being the Condorcet loser, nevertheless can be the best winner for society. For example if the utilities of the candidates to each voter are 9, 8, 7, and 0, then electing L would cause the greatest benefit to society.

the few systems considered here in which votes have *continuum* freedom.

#voters	their vote	#voters	their vote
999	$A > B \gg C$	49	$A \gg B > C$
1	$C > B \gg A$	48	$C \gg B > A$
		3	$B > A \gg C$

Figure 7.4. Some examples where intensity of preference matters.
(left) Saari's anti-approval example. Approval voting would elect B, but the societally-best (and Range and Borda and Condorcet) winner is A.

(right) Anti-Condorcet example. The Condorcet-winner is B, but A is the societally-best (and Approval and Range) winner. ▲

Approval and Range voting are uniquely invulnerable to manipulation by strategic voters, in this weak sense: If a voter knows the exact totals of all the other votes, then that voter can cast a strategically-optimal approval vote which happens to also be honest in the sense that it is a *limit* of vectors $\vec{x}$ such that the $>$, $<$, and $=$-relations among $\vec{x}$-entries agree with that voter's honest preference relations among candidates.

Also, in a 3-candidate election,[49] even with only *in*exact probabilistic knowledge of the total of the (large number of) other votes, one may find a strategically-optimal and limit-sense-honest Range and Approval vote. Call the two frontrunners (most likely to win) A and B. Let their a priori probabilities of election be P_A and P_B respectively. Award the better one 1 and the worse one 0. Now award C vote 1 if $U_C > P_A U_A + P_B U_B$ where U_K denotes your personal utility for the election of candidate K.

In contrast, *dis*honest voting – even in 3-candidate elections – often is strategically optimal in the IRV (table 6.12) and Condorcet least-reversal[50] systems (in which it usually pays to rank your favorite's most-feared rival artificially "last" even if you do not honestly regard him as the worst of the three).

7.9 Condorcet-range voting

The following method was suggested to me by Clay Schoentrup, a Libertarian from Portland Oregon. It is more complicated than plain range voting, but it has the advantage (in some eyes) that it elects Condorect winners when they exist.

▶ Each vote is a real N-vector every entry of which is in the real interval $[0, 1]$. Ignore the actual numerical values of the scores for the candidates, only taking account of the $<$ and $>$ and $=$ relationships amongst them. If there is a Condorcet ("beats all") winner, elect him. Otherwise, elect the range-voting winner, with the greatest sum-of-scores.

Condorcet-range voting is monotonic in the sense that raising some candidate X's score in your vote (while leaving the others fixed) cannot decrease X's chances of winning. (Hint for proof: As a lemma, prove making an adjacent interchange of $X > Y$ to $Y > X$ in one vote by altering the vote's score for Y, cannot hurt Y's winning chances.)

#voters	their vote
3	$A = 1$, $B = 0.3$, $C = 0$, $D = 0.5$
3	$A = 1$, $B = 0$, $C = 0.3$, $D = 0.5$
4	$A = 0.1$, $B = 1$, $C = 0.2$, $D = 0$
5	$A = 0$, $B = 0.9$, $C = 0.5$, $D = 1$
6	$A = 0.1$, $B = 1.0$, $C = 0$, $D = 0.2$

Figure 7.5. "No show paradox" in Condorcet-Range voting. In the situation above the line, there is no Condorcet winner so that B wins (since the totals are $A = 2.1$, $B = 2.2$, $C = 1.0$, $D = 2.0$). But if we add the 6 voters below the line, all of whom top-rank the current winner B, then D becomes a Condorcet winner and wins. Those 6 voters were better off "not showing up" (and if we add them one by one, the first individual voter to cause the winner-change, was better off not showing up). ▲

#voters	their vote
8	$A = 0$, $B = 1$, $C = 0.5$
6	$A = 0.5$, $B = 0$, $C = 1$
5	$A = 1$, $B = 0.5$, $C = 0$

Figure 7.6. Favorite-betrayal in Condorcet-Range voting. (Based on table 6.5.) In this situation, there is no Condorcet winner so that B wins (totals $A = 8$, $B = 10.5$, $C = 10$). But if the 6 voters in the middle line change their vote to $A = 1$, $B = 0$, $C = 0.5$ ("betraying their favorite C") then the "lesser evil" A becomes the Condorcet winner and wins. ▲

8 ???Sarvo-Range voting [71]

▶ Each vote is a real N-vector every entry of which is in the real interval $[0, 1]$, *plus* a single additional "strategy bit."

We now randomize the order of the V voters (all $V!$ possible ordering equally likely) and start computing the sum of the vectors, one vector at a time, in that order, *except* that before we incorporate each vector-vote $\vec{v}$ into the vector-sum, *if its strategy bit is 1*, we first *transform* it as follows.

WRONG – NEED TO REDO???

procedure Strategic-transform

Reorder the coordinates i of the vector-vote $\vec{v}$ so that the current (partial) vector sum $\vec{s}$ is in decreasing order, i.e. the candidates are in decreasing order of performance-so-far in the election. (Break ties randomly.)
for $i = 1, \dots, N$ **do**
  $A \leftarrow ((i-1)A + v_i)/i$;
  if $v_i > A$ or $A = v_i = v_{i+1} = v_{i+2} = \cdots = v_{j-1} > v_j$ **then**
    $w_i \leftarrow 1$;
  else if $v_i < A$ or $A = v_i = v_{i+1} = \cdots = v_{j-1} < v_j$ **then**
    $w_i \leftarrow 0$;
  else
    $w_i \leftarrow$random value;
  end if

[49] Or a "trichotomous" election in which each voter regards all candidates as falling into one of only three equivalence classes.

[50] Note: many other systems, such as Tideman ranked pairs, Schulze, and Raynaud, are equivalent to Condorcet least-reversal in the 3-candidate case.

11: **end for**
12: Un-reorder the coordinates of $\vec{w}$ and output it.

(If a voter's strategy bit is 0 then we simply use his untransformed vote $\vec{w} = \vec{v}$.) We compute the full sum $\vec{s}$ of all the (perhaps transformed) votes $\vec{w}$. This $\vec{s}$ will be a function of the random order among the voters that we used (as well as among the random choices inside `Strategic-transform`).

We take the *average* of $\vec{s}$ over all $V!$ possible voter-orders (or, what is more feasible, we average over a vast number of random orders and keep going until the "noise" seems sufficiently small) and the winner is the candidate corresponding to the maximum entry in the averaged $\vec{s}$.

Note: Sarvo-range voting is exactly the same as range voting if there are only honest voters. It only differs in the presence of strategic voters and is intended to equip range voting with increased resistance to strategic manipulation.

9 Properties

There are numerous desirable-sounding properties which a voting system might or might not have. Here are some of the more important ones:

Anonymity: means that all voters are treated equally.
Neutrality: means that all candidates are treated equally.
Homogeneity: means duplicating each vote leaves election result the same.

All the systems we have described are Anonymous, Neutral, and Homogeneous.

CW: If a Condorcet-Winner (who would beat every other candidate in a one-on-one plurality election compatible with the original votes) exists, he is elected. A related but stronger demand would be that the winner must always belong to the Smith Set.

Condorcet himself may have had a slightly different definition in mind, which I shall call CW′: replace the words "plurality election compatible with the original votes" with "election of the original type, with the same votes with all candidates but these two omitted." The two definitions are equivalent when applied to virtually every rank-order-ballot based method, but differ for, e.g, range voting. Range voting disobeys CW but obeys CW′.

IP: Incentive to Participate. Adjoining an additional set of identical votes, all favoring A over B, cannot cause A to lose or B to win.

AT: Add-Top: Adjoining an additional set of identical votes, all ranking A top, cannot harm candidate A.

LN: Later-no-harm [83]: Adding a later preference to a ballot should not harm any candidate already listed. Satisfying LN is highly related to, but not exactly the same as (cf. figure tab:IRVantiex), removing the incentive for voters to "truncate" their preference-ballots. In trying to change from Plurality to a ranked-ballot method, LN is an excellent selling-point to plurality-minded voters. All methods that obey LN automatically also obey AT, but not vice versa.

SD: Subdistrict Consistency. If two disjoint subsets of votes, each by itself uniquely elects A, then the combined set uniquely elects A also.

MN: Monotonicity. Increasing the rank of A in some set of identical votes (while leaving the relative ranks of the other candidates unchanged) cannot decrease A's probability of winning. (To prove a method monotonic, it suffices to show that changing just one ballot by a single adjacent-interchange [$A > B$ becomes $B > A$] cannot hurt B's and cannot help A's chances of election. This idea makes it easy to see, for example, that Improved-Dodgson, Bucklin, Fishburn-set and Banks-set obey MN. In some other cases, e.g. Simpson-Kramer, one may prove MN by reasoning directly from the voting method's definition.)

CI: Clone-Immune [78]. Adding "clones" of a candidate C (whose relative rankings to other candidates, in all votes, are the same as C's) will not alter victory probabilities.[51] Warning: "clones" of A can be viewed as slightly better or worse than A by voters – and are, if we are discussing a preference-ranking-based system, even one allowing ranking-equalities. (One could also discuss "$\mathrm{CI}_=$" for ranking-based systems allowign ranking equalities where we *demand* than all clones be given exactlyequal rankings.) But if we are discussing real-vector-based voting systems, we require all clones to receive arbitrarily close votes.

GU: Generically Un-tied. If V voters participate, each of whom chooses his vote independently from some fixed nontrivial probability distribution, then as $V \to \infty$ while the number N of candidates remains fixed, the probability of a tied election approaches 0, i.e. the probability a single-winner is got, approaches 100%.

CL: Incapable of electing a Condorcet-Loser. Related, but weaker, would be the Majority Loser condition that a candidate uniquely ranked bottom on a majority of ballots, should be incapable of winning.

UD: Unanimous Domination. (Also called Pareto.) If at least one voter ranks B over A and no voters rank A over B, then A must not be elected. (A related, but weaker, property would be the Unanimous Winner property that anybody whom all voters rank best, must win.)

EP: Efficiently Parallelizable. When $V \gg N \geq 2$ there is an efficient way to perform the election in which each precinct only sends some kind of "subtotal" to the central tabulating agency, i.e. a much smaller amount of information than sending every vote cast in that precinct individually. And all the necessary communication is one-way. Note: strictly speaking we have no *proofs* of the nonexistence of such algorithms, so that all our assertions of the falsity of EP are, in fact, merely (highly plausible) conjectures. One *could* prove, e.g, the weaker claim that IRV and Coombs election results are not computable purely from the U-matrix ([63] p.463).

MJ: Majority. A candidate uniquely top-ranked by a majority of ballots, is elected. Related, but stronger, is the "Mutual Majority Criterion": If there is a majority of

[51] More precisely: If there is a subset S of alternatives such that no voter ranks any alternative outside S between any alternatives in S, then the election outcome must not change if a strict subset of S is deleted from the votes and from the set of nominees. By "not change" we mean the same winner wins if the clones were of a non-winner, but if the clones were of the winner, then one of the clones still wins.

voters which rank a *set* of candidates above all others, then one of those candidates must win.

A related condition would be "Majority Defense" which says that a majority is always capable of electing anyone they choose (or blocking the election of anyone). Range Voting fails Majority but satisfies Majority Defense.

FO: Definition of voting system is readily and naturally generalizable to make it output a full ordering of the candidates, rather than merely a single winner. (Actually, any system outputting single winners may be made to output an ordering by eliminating the winner and then re-running the system to get the "second place finisher," etc. However this is not natural because it ranks the 2nd, 3rd,... place winners without utilizing information about the 1st-place winner. Also this requires much more work than simply doing everything only once.)

AFB: Avoidance of "Favorite Betrayal."[52] Voters never have incentive to dishonestly rank someone over their favorite.

DH3: Avoids returning devastatingly poor results in the DH3 scenario of §6.3 and table 6.4.

Honest versus Strategic voting. Although in the literature properties of voting systems are usually only analysed under the assumption that the voters are honest, i.e. in terms of the votes cast, they could also be examined for *strategic* voters in terms of the honest votes they did *not* cast.

The results sometimes differ. For example, range voting with honest voters can elect a Majority Loser, if 51% of the voters considered him bottom-ranked by only a slight amount, while 49% consider him top-ranked by a large gap. But with strategic range-voters, the ML cannot be elected[53] (Proof sketch: The ML will get a minimum vote from over 50% of the voters. But the more popular among the two pre-election-poll frontrunners will get maximum votes from over 50% of strategic voters.)

Similarly range voting with honest voters can never elect a candidate B who is ranked below some other (A) by 100% of the voters; but range voting with strategic voters *can* uniquely elect B [68].[54]

With honest voters, the Condorcet least-reversal system will elect a Condorcet-Winner if one exists – but with strategic voters, that Condorcet-Winner might no longer be one and thus could avoid being elected. (And in the DH3 scenario of table 6.4 the strategic voting *creates* a Condorcet-Winner who with honest voting would be a unanimous-loser.)

#voters	their vote
51	$A, B > C$
49	$B, C > A$

Figure 9.1. An example by Marc LeBlanc of how Arrow's I.I.A. property can be both valid and invalid under Approval voting. Under Approval voting, the first group votes for B as a (successful) defense against the perceived threat of C, and B wins. If C drops out of the race, then the same voters would have instead voted $A > B$ and A would have won. Thus C's decision to run, or not, affected the result of the A versus B battle in an Approval election with strategic voters acting in response to partial information, whereas, with honest voters, C's candidacy would have been irrelevant. ▲

Implications among & theorems about these properties.

Additive-vector methods automatically satisfy SD, MN, IP, AT, and ME, but automatically fail CW [67][85][87] and LN (except if the "later-ranked" candidates have vector-entry 0, as for approval and plurality voting, in which case LN is true).

SD-satisfying methods are the same thing as "composed" weighted-positional scoring systems (i.e. the ties in a WP system are broken by use of another, and so on[55]) [67][85][87][72]. Hence (assuming non-negative weights) they automatically satisfy MN, IP, AT, and ME.

Every weighted-positional score-sum system (with honest voters) is non-immune to clones. The same is true for elimination systems whose rounds are based on WP systems (e.g. Nanson, Rouse). Elimination systems based purely on inequalities among elements of[56] the M-matrix, and in which one's clones cannot "protect" one from being eliminated, are automatically clone-immune because once all of somebody's clones have been eliminated, the method proceeds as before. Thus Tideman Ranked pairs, River, and Raynaud are clone-immune.

But Simpson-Kramer is not clone immune since we can make three clones of its winner forming a Condorcet cycle, who defeat one another by huge margins. That prevents any of them from winning, even if one otherwise would have. Similarly Arrow-Raynaud is not clone-immune.

Systems based solely on the M- and/or U-matrices, or solely on the sum of vote-vectors, automatically satisfy EP. Also TMR satisfies EP since for each candidate we may accumulate the counts of voters ranking him kth, for all k, then use these counts to compute median rankings. (This rank-count matrix may also be used to compute scores under any weighted-positional system. However, *repeated* TMR, i.e. breaking the ties in TMR by using TMR again within the tied-subset until a fixpoint is reached, is *not* possible with this approach and presumably does not obey EP.)

A voting system whose input is preference orderings and whose output is a "winner" subset "respects majority" if whenever a majority of voters say X is their favorite, then X is

[52] Advocated by Mike Ossipoff in numerous web postings.

[53] This is assuming the strategic range voters believe that ML is one of the two "frontrunners" one of whom is regarded as almost certain to win. Oppositely, plurality voting with *honest* voters can elect ML, but with strategic voters cannot! This is an exceptional case where strategic-dishonest voters actually produce a better result for society than honest ones.

[54] Similarly plurality voting with honest voters will elect a Majority Winner (whom 99% agree is the best candidate) but with strategic voters may elect somebody else; with Borda voting the story is precisely the opposite in both cases. Also, plurality with honest voters can never elect a unanimously-dominated candidate, but can with strategic voters.

[55] Equivalently, we have only a single WP system, but whose weights are taken from a larger field than the reals which contains a hierarchy of infinitesimal quantities.

[56] But not linear or nonlinear combinations of elements of

elected, and "respects anti-majority" if whenever a majority of voters say X is worst, then X is not chosen. (Thus the plurality system respects majority but Anti-Plurality and Borda do not.) Woeginger [82] notes that no "weighted positional scoring system" can simultaneously respect majority and anti-majority (for a proof consider the example we stated immediately after table 6.3). But many multiround methods such as Schulze, P+I, and IRV do.

The failure of most of the methods that fail AFB may be demonstrated by table 6.5. (And Bucklin fails since if A and B are tied, with 50% of the other top-rank votes, it almost certainly would be foolish for you to rank some other candidate top; ER-Bucklin, i.e Bucklin with rank equalities permitted, satisfies AFB.) In all the cases I know where AFB can be proven (besides Approval and Range voting, where it is obvious) the proof strategy is: you show that *if* demoting favorite F in order to cause X to win over Y is a valid strategy, *then* raising F and X to co-equal top rank in that same planned votes, also works (but does not betray F). All known methods obeying AFB allow equal rankings in ballots.

Methods satisfying CL automatically satisfy Majority Loser.

Methods satisfying "immunity to complaints" (mentioned in §6.10) automatically satisfy CW, the Smith set property, and UD.

Elimination methods based on point-scoring systems in each round (e.g. P+I, Nanson-Baldwin, and IRV), or, even more generally, whose eliminations are governed by any system that eliminates Condorcet-Losers (e.g. Rouse and Arrow-Raynaud) automatically fail MN (by [67] thm. 2, reproven and extended near the end of [72]; see also [83]). Hence, in view of the characterization of SD-satisfying systems in [67][85][87], they also fail SD.

The status of most of our elimination methods with respect to MA, AT, and IP have been considered by [63][58], although the reader is warned that many of the examples of Richelson [63] depend on simultaneous eliminations in case of ties, which is not the usual definition of these voting methods. For example, the Coombs and IRV variants with simultaneous elimination can easily be made to elect a Condorcet-Loser by simply eliminating everybody else (all tied for most last-place rankings or fewest first-place rankings, respectively) in the first round; but with one-at-a-time elimination, neither Coombs nor IRV, nor any other elimination method of their ilk, can elect a Condorcet-Loser L because if so, eventually it would come down to L versus some other candidate, whereupon L would be eliminated.

Methods satisfying CW automatically satisfy MJ but automatically fail IP [54][59], ME (see Schulze's proof of this in [72]), SD [67][83][85][87], and DH3 (table 6.4). (Thus Woodall's DAC method, because it satisfies IP, fails CW; and Nanson's method, because it satisfies CW, fails IP, SD, and DH3.) A large class of CW-satisfying methods fail AT [83].

Jobst Heitzig points out that all anonymous neutral methods satisfying CW (and also all WP methods other than plurality) automatically fail LN because of table 9.2.

#voters	their vote
1	A>B>C
1	B>C>A
1	C>A>B

Figure 9.2. Later harm. All three candidates are tied and must be elected with probability 1/3 each. (Or by a slight perturbation of this example we may make A be the unique winner.) But C becomes both the Condorcet-Winner and the winner in any WP system with weights $w_1 \ge w_2 > w_3$, when the first voter switches to $A > C > B$; thus A is "harmed" by these later choices. (This also shows many other methods, e.g. Improved-Dodgson and Keener, fail LN too.) ▲

Methods satisfying Smith-Set automatically satisfy CW, CL, and Mutual Majority.

Methods satisfying either IP or LN automatically satisfy AT (and hence methods failing AT must fail IP and LN).

Theorem 3 (Condorcet methods exhibit Favorite Betrayal). *"Favorite betrayal" is sometimes "strategically forced" in every Condorcet voting method (whether rank-equalities are allowed or not) [We shall only consider "Condorcet voting methods" which determine winners purely from the pairwise-results matrix, and which are "anonymous" and "neutral."] The theorem also is true of all WP systems with weights $w_1 \ge w_2 > w_3$; the same proof shows that.*

Proof (Kevin Venzke[57]). As a lemma: if "favorite betrayal" is never strategically necessary then it must be the case that increasing the votes for A over B in the pairwise matrix can never increase the probability that the winner comes from $\{A, B\}$; that is, it must not move the win from some other candidate C to A. The lemma is true because if sometimes it were possible to move the win from C to A thusly, then a $B > A > C$ voter would have incentive to reverse B and A in his ranking (and such a voter wlog would exist by adding a constant number of each of the 6 kinds of voters if necessary - this will always not be a problem wherever we will use the lemma below - and note equal ranking would be inadequate).

#voters	their vote
3	$A = B > C$
3	$C = A > B$
3	$B = C > A$
2	$A > C > B$
2	$B > A > C$
2	$C > B > A$

Consider the above 15-voter scenario. (If equalities are disallowed, then double every vote and replace equalities $X = Y$ by $X > Y$ in one copy and $Y > X$ in the other.)

There is an $A > C > B > A$ cycle, and the scenario is "symmetrical." Hence an anonymous and neutral method must elect each candidate with 1/3 probability.

Now suppose some $A = B$ voters change their vote to $A > B$. This would turn A into the Condorcet winner, who would have to win with 100% probability since it is a Condorcet method.

But the probability that the winner comes from $\{A, B\}$ has increased from 2/3 to 1, so the property of the lemma is violated. Q.E.D.

[57]This proof is based on one by Venzke in a June 2005 Electorama web post.

Relative importances of these properties. The importance of a property is somewhat subjective and depends on the application one has in mind. Thus, for political applications (in which election ties induce crises) GU failure is very serious. EP failure similarly may be enough to simply disqualify a voting system.

Although electing a Majority Loser (or failing to elect a Majority Winner, or a "reversal symmetry" scenario such as in figure 6.26) may be embarrassing, these are rare in practice in most systems besides Arrow-Raynaud. The DH3 pathology, on the other hand, *both* has a very devastating effect (pessimal winner elected) *and* is common in practice in all the systems that fail DH3 (albeit to a diminished extent in Loring's system). So I regard a DH3 failure as extremely serious. However, if for some reason we knew that strategic-dishonest voters were not going to exist, then DH3 would be irrelevant.

		scheme	CW	GU	MN	SD	IP	AT	LN	CI	AFB	CL	UD	EP	MJ	FO	DH3	si	expr. fr.	runtime	#
I	1	OptWin	V^A	Y	Y	A	A	A	F	Y	Y	F_M	Y	—	F	Y	Y	—	∞^N	NV	10
	2	WorstWin	F	Y	F	Y	F	F	F	Y	F	F_M	F	—	F	Y	F	—	∞^N	NV	4
II	3	RandCand	F	Y	F	F	Y	Y	Y	F	Y	F_M	F	Y	F	Y	Y	0	0	1	8
	4	RandDict*	F	Y	Y	F	Y	Y	Y	Y	Y	F_M	Y*	Y	F	(Y)	Y	1	N	V	11
	5	RandPair*	F	Y	Y	F	Y	Y	Y	Y	Y	Y	Y	Y	F	Y	Y	3	$N!$	VN	12
III	6	Plurality	V^A	Y	A	A	A	A	Y	F	F	F	Y	Y	Y	Y	Y	1	N	$V+N$	11
	7	AntiPlur	V^A	Y	A	A	A	A	Y	F	Y	F	F	Y	F	Y	F	1	N	$V+N$	9
IV	8	P+I	F	Y	E	F	F	Y	Y	F	F	Y	Y	Y	Y	F	Y	4	$N!$	VN	8
	9	Nauru	V^A	Y	A	A	A	A	F	F	F	F	Y	Y	F	Y	Y	4	$N!$	VN	9
	10	Borda	V^A	Y	A	A	A	A	F	F	F	Y	Y	Y	F	Y	F	4	$N!$	VN	9
	11	CondrctLR	Y_S	Y	Y	F	F	F	F	F	F	F_M	Y	Y	Y_M	Y	F	6	$N!$	$VN+N^2$	7
	12	Black	Y_S	Y	Y	F	F	F	F	F	F	Y	Y	Y	Y	Y	F	6	$N!$	$VN+N^2$	8
	13	ImprvDod	F	Y	Y	F	F	F	F	F	F	Y?	Y	F	F	Y	F	8	$N!$	VN^3	4-5
	14	SchulzeBP	Y	Y	Y	F	F	F	F	Y	F	Y	Y	Y	Y	Y	F	8	$N!$	$VN+N^3$	9
	15	TidemanRP	Y	Y	Y	F	F	F	F	Y	F	Y	Y*	Y	Y	Y	F	6	$N!$	$VN+N^4$	9
	16	River	Y	Y	Y	F	F	F	F	Y	F	Y	Y*	Y	Y	F	F	7	$N!$	$VN+N^2$	8
	17	Maxtree	Y	Y	Y	F	F	F	F	Y	F	Y	Y*	Y	Y	F	F	9	$N!$	$VN+N^2$	8
	18	TMR	F	F	Y	F	F	F	F	F	F	F	Y	Y	Y	Y	F	4	$N!$	$VN+N\lg N$	5
	19	Copeland	Y	F	Y	F	F	F	F	F	F	Y	Y	Y	Y	F	F	4	$N!$	$VN+N^2$	6
	20	Keener	F	Y	Y	F	F	Y?	F	F	F	Y	Y	Y	Y	Y	F	8	$N!$	$VN^2\lvert\log\epsilon\rvert$	7-8
	2??	Sinkhorn	F	Y	Y	F	F	Y?	F	F	F	Y	Y	Y	Y	Y	F	8	$N!$	$VN^2\lvert\log\epsilon\rvert$	7-8
	21	SimpsonKr	Y_S	Y	Y	F	F	Y	F	F	F	F_M	Y	Y	Y_M	Y	F	6	$N!$	$VN+N^2$	8
	22	IRV	F	Y	E	F	F	Y	Y	Y*	F	Y	Y	F	Y	Y	Y	4	$N!$	$VN+N^2$	9
	23	LoringIR	Y	Y	E	F	F	F	F	Y*	F	Y	Y	F	Y	F	(F)	5	$N!$	$VN+N^2$	6
	24	CretneyIR	Y	Y	E	F	F	F	F	Y*	F	Y	Y	F	Y	F	F	5	$N!$	$VN+N^2$	6
	??	BTR-IRV	Y	Y	E	F	F	F	F	Y*	F	Y	Y	F	Y	Y	F	5	$N!$	$VN+N^2$	7
	25	Coombs	F	Y	E	F	F	F	F	F	F	Y	Y	F	F	Y	F	4	$N!$	$VN+N^2$	4
	26	NansonB	Y_S	Y	E	F	F	F	F	F	F	Y	Y	Y	Y	Y	F	5	$N!$	$VN+N^2$	7
	27	Rouse	Y	Y	F	F	F	F	F	F	F	Y	Y	Y	Y	Y	F	6	$N!$	$VN+N^3$	7
	28	Raynaud	Y	Y	F	F	F	F	F	Y	F	Y	Y	Y	Y	Y	F	6	$N!$	$VN+N^2$	9
	29	ArrowRayn	F	Y	F	F	F	Y	F	F	F	Y	Y	Y	F	Y	F	6	$N!$	$VN+N^2$	6
	30	Bucklin	F	Y	Y	F	F	Y	F	F	F	F	Y	Y	Y	Y	F	4	$N!$	$VN+N^2$	7
	31	WoodallDAC	F	Y	F	F	Y	F	F	Y	F	F	Y	F	Y	F	(F)	9	$N!$	VN^2	5
	32	Sarvo-Plur	F	Y	Y	F	Y	Y	Y	F	F	Y	Y	Y	F	Y	Y	7	$N!$	$VN+N^2$	10
	33	SmithSet†	Y	F	Y	F	F	Y	Y	Y	F	Y	Y	Y	Y	F	F	4	$N!$	$VN+N^2$	9
	34	Fishburn†	Y	F	Y	F	F	F	F	Y	F	Y	Y	Y	Y	F	F	5	$N!$	$VN+N^3$	7
	35	BanksSet†	Y	F	Y	F	F	F	F	Y	F	Y	Y	Y	Y	F	F	5	$N!$	$VN+N^2$	7
V	36	Dabagh	V^A	Y	A	A	A	A	F	F	F	F_M	Y	Y	F	Y	F	2	$(N-1)N$	V	8
	37	for-&-against	V^A	Y	A	A	A	A	F	F	F	F_M	Y	Y	F	Y	Y*	2	$(N-1)N$	V	9
	38	Signed	V^A	Y	A	A	A	A	F	F	F	F_M	Y*	Y	F	Y	Y	2	$2N$	V	9
	39	Approval	(V^A)	Y	A	A	A	A	(F)	Y	Y	(F_M)	Y*	Y	F	Y	Y	2	2^N	VN	11
VI	40	Cumulatv	V^A	Y	A	A	A	A	F	F	F	F_M	Y	Y	F	Y	Y	2	∞^{N-1}	VN	9
	41	MedianRk	F	F	Y	F	F	F	F	F	Y	F	Y*	F	Y	F	F*	4	∞^N	VN	3
	42	Range	V^A	Y	A	A	A	A	F	Y	Y	F_M	Y	Y	F	Y	Y	3	∞^N	VN	11
VII	43	Asset‡	V^A	Y	Y	Y?	Y?	A	F	Y	Y	F_M	Y†	Y	F_R	(Y)	Y	3	∞^{N-1}	VN	8-10

Figure 9.3. Important properties of our voting systems (*=non-deterministic system, †=returns a possibly non-singleton set of winners). [1][7][26][58][59][62] [63][66][67][83][85][87][88]. Please notify me of any errors found in this table.

The horizontal lines separate major categories of voting system: (I) non-implementable systems of only theoretical importance, (II) non-deterministic systems [rarely if ever recommended for practical use], (III) minimally expressive (but simple) systems where a vote is just the name of a candidate, (IV) more-expressive systems where a vote is a preference-ordering among the N candidates, (V) systems where a vote is a real N-vector, (VI) same but now with *continuum* voter freedom hence maximal expressivity, (VII) Asset is an unconventional "voting system" only useable if the candidates are sentient beings [denoted by ‡]. *Two*-letter properties pertain to honest and *three*-letter ones to strategic voting.

The "si" column rates the subjective simplicity of the system on an 0-9 scale with 0=simplest. The "expr. fr." column counts how many possible kinds of votes there are in each system (measure of "expressive freedom"), the "#" column counts the number of properties the method obeys, and the "runtime" column gives a formula such that the runtime (on a "real RAM" computational model if $V \geq 2$ and $N \geq 1$) to perform the election is O(the formula).

Key: Y=obeys that property. (*=if no ties. Y_S=fails the stronger Smith Set property; Y_M=fails the stronger Mutual

Majority property; Y† means Asset voting will obey UD if the candidates=negotiators have the *same* unanimous preferences as the electorate.) F=fails to obey property. (F_M if fails the stronger Majority Loser property, F_R if "Y" for strategic voters, F* means that MedianRk and for-and-against will, in the DH3 scenario, award everybody 0 so the choice will be made by tiebreaking.) A=additive method, hence automatically obeys MN, IP, SD, AT, AFB and fails CW and also fails LN except for a few special cases. (Approval Voting can fail to elect a Condorcet-Winner. *But* there always *exists* a way to choose the approval votes in a manner compatible with each voter's honest preference ordering, so that a member of the Smith Set will be elected – hence the parentheses. Also, CW and CL failures under Approval and some other systems have the "virtue" of perhaps being undetectable since preference orderings are not deducible from Approval ballots.) E=elimination method based on point-scoring, hence fails MN and SD. The fact that the best-possible voting method OptWin fails LN, MJ, CW, Smith Set, CL, and Majority Loser suggests that those six properties are not actually desirable ones for a voting system to have. ▲

10 Which system is the best?

If we uncritically accept all the properties in table 9.3 as "desirable" and crudely weight them all equally, then Plurality, Approval, and Range voting are the best deterministic systems tabulated, each with 11 properties satisfied. The other practically-useable undominated deterministic methods satisfying ≥ 9 properties (if we give some methods the benefit of the doubt in some doubtful cases) are Sarvo-Plurality, Asset, Raynaud, IRV, Schulze beatpaths, Tideman Ranked Pairs, Nauru, and Borda.

Excluding more complicated, slower, or less-expressive methods whenever there are two competitors with otherwise similar properties, leaves us with **Range** and **Approval** in the top tier, and Asset, IRV, Schulze, Raynaud, Nauru, and Borda in the second tier.[58] It might seem from this that it is going to be difficult to say with confidence that some one method is the "best," and rather, we should expect different methods to be best in different situations.

But my very large computer-aided study [68] concluded that **range voting** was clearly, and robustly, the best system among those compared, across a large variety of different situations, when judged by a statistical yardstick called "Bayesian regret." However, when I did that study, I was not aware of all of the voting systems described here.

Recently, I rewrote my voting-simulator software, with the aim of including all the systems described here (some of which, incidentally, are new inventions). The preliminary conclusion is that range voting still clearly has the lowest Bayesian regrets, *except* that

1. asset voting can sometimes outperform it,
2. sarvo-range voting apparently *always* outperforms it in the presence of any nonzero fraction of strategic voters. (If all voters are honest, then sarvo-range and range voting are identical.)

Asset voting is different from all the other voting systems in that it requires the candidates to be sentient beings capable of having preferences about, and negotiating with, each other – as opposed to just abstract choices. It is probably not a better choice than range voting for single-winner elections both because of this and also because it can exhibit disturbingly higher Bayesian regrets than range voting in certain kinds of realistic situations, whereas in the realistic situations where it seems superior to range voting, the improvement is not huge. Asset voting was originally invented with the idea that it would be a good *multiwinner* system, and for that purpose it may indeed be very good. (There is also a multiwinner "reweighted" version of range voting [70]; it also looks good.)

Sarvo-range is an excellent system. It was devised by combining ideas of L.F.Cranor [16][17] about "declared strategy voting" with my own [68] understanding of optimal range-voting strategy. It was specifically designed to beat the "world champion" (range voting)'s Bayesian regret scores, and experiments [71] show that it succeeded. (The paper [71] will indeed describe other more sophisticated sarvo-range variants too.) However, sarvo-range voting is much more complicated to describe and use than range voting, and therefore may not be acceptable in much of the real world.

What real world voters think: Smith and Greene [73] polled random voters in New York State and found that they *prefer plurality* to range voting by 70 to 45 (with 8 "don't know"s), because the latter is considered "too complicated."[59] Similarly Quintal [73] polled random voters near Philadelphia and found plurality preferred to approval 296 to 238 (with 122 "don't knows"). But range and (especially) approval voting are simpler than almost all the other reasonable voting methods we've described! Smith, Quintal, and Greene therefore concluded that efforts of devising mathematically elegant but complex voting systems were "mere mental masturbation" incapable of real world success and advocated "Keep It Simple Stupid."

11 Acknowledgements

I thank Kevin Venzke for his helpful comments.

References

[1] Kenneth J. Arrow & Hervé Raynaud: Social choice and multicriterion decision making, MIT Press 1986 (Cambridge MA).

[2] Martin J. Bailey (with editing by Nicolaus Tideman): Constitution for a future country, Palgrave 2001.

[3] J.Banks: Sophisticated voting outcomes and agenda control, Social Choice & Welfare 2 (1985) 295-306.

[58] I have ordered these in roughly decreasing order of subjective quality based on my current perception of the magnitudes and importances of the property violations.

[59] One might hope they would change their minds if they each had thought about the subject for *more* than 1 minute – e.g. the Irish have shown in two referenda that they want to keep their present Hare/Droop-STV (multiwinner) system and not switch back to plurality, despite the fact that probably 80% of Irish voters would be unable to actually describe how their system works – but clearly, in the USA the force is against voting reformers.

[4] J.J.Bartholdi, J.A. Orlin: Single transferable vote resists strategic voting, Social Choice and Welfare 8 (1991) 341-354.

[5] J.Bartholdi III, C.Tovey, M.Trick: Voting schemes for which it can be difficult to tell who won the election, Social Choice & Welfare 6,2 (1989) 157-165.

[6] Duncan Black: The theory of committees and elections, Cambridge Univ. Press 1958.

[7] S.J. Brams: Voting procedures, ch. 30 pp.1050-1090 in vol.2 of Handbook of Game theory, (ed. R.Aumann and S.Hart) Elsevier Science 1992. Updated version: Brams & Fishburn pp.173-236=ch.4 in Handbook of Social Choice and Welfare, ed. K.J.Arrow A.K.Sen., K.Suzumura, Elsevier 2002.

[8] S.J.Brams & P.C.Fishburn: Approval Voting, American Political Science Review 72,3 (September 1978) 831-847; also book with same title, Birkhäuser Boston 1983.

[9] G. Brinkmann & B. D. McKay: Posets on up to 16 points, Order 19 (2002) 147-179.

[10] J.R.Chamberlin: An investigation into the effective manipulability of four voting systems, Behavioral Science 30 (1985) 195-203.

[11] John R. Chamberlin, Jerry L. Cohen, Clyde H. Coombs: Social Choice Observed: Five Presidential Elections of the American Psychological Association, J. of Politics 46,2 (1984) 479-502.

[12] Y.J. Chu & T.H. Liu: On the shortest arborescence of a directed graph, Scien. Sinica 14 (1965) 1396-1400.

[13] Edward H. Clarke: Demand-revelation and the provision of public goods, Ballinger Cambridge MA 1980. (2nd ed. 2000.)

[14] C.H. Coombs: Social choice and the strength of preference, in *Decision processes* (R.M.Thrall, C.H. Coombs, R.L. Davis eds.) Wiley NY 1954.

[15] A.H. Copeland: A reasonable social welfare function. Mimeo, University of Michigan, 1951.

[16] Lorrie F. Cranor: Declared-Strategy Voting: An Instrument for Group Decision-Making, PhD. dissertation, Washington University, St. Louis, December 1996. Available from `http://lorrie.cranor.org/dsv.html`.

[17] L.F.Cranor & R.K.Cytron: Towards an Information-Neutral Voting Scheme That Does Not Leave Too Much To Chance, Presented at the Midwest Political Science Association 54th Annual Meeting, April 18-20, 1996.

[18] Blake Cretney: elction-method research web site at `http://condorcet.org/emr/criteria.shtml`.

[19] Thomas S. Dabagh: The vote-and-a-half system of majority choice by plurality vote, Amer. Polit. Sci. Review 28,1 (1934) 93-101.

[20] G.B. Dantzig: Linear Programming and Extensions, Princeton University Press, Princeton, NJ, 1963.

[21] J.L.Davison: Asymptotic enumeration of partial orders, Combinatorics, graph theory, and computing, Proc. 17th Southeast. Conf., Boca Raton/Fl. 1986 = Congressus Numerantium 53 (1986) 277-286.

[22] D.J. Djokovic: Note on nonnegative matrices, Proc. Amer. Math'l. Soc. 25,1 (1970) 80-82.

[23] C.L.Dodgson: The Pamphlets of Lewis Carroll, vol. 3, The political pamphlets and letters of Charles Lutwidge Dodgson and related pieces, (compiled & edited by Francine F. Abeles) Lewis Carroll Society of North America, NY, distributed by Univ. Press of Virginia, Charlottesville & London 2001.

[24] Cynthia Dwork, Ravi Kumar, Moni Naor, D.Sivakumar: Rank Aggregation for the web, 10th World Wide Web conference, Hong Kong (2001) 613-622. `www.wisdom.weizmann.ac.il/~naor/PAPERS/rank_www10.html`

[25] J. Edmonds: Optimum branchings, J. Res. Natl. Bureau Standards 71B (1967) 233-240.

[26] P.C.Fishburn: Condorcet social choice functions, SIAM J. Appl. Math. 33 (1977) 469-489.

[27] P.C. Fishburn & W.V. Gehrlein: Majority Efficiency For Simple Voting Procedures: Summary and Interpretation, Theory Decision 4 (1982) 141-153.

[28] H.N. Gabow, Z. Galil, T. Spencer, R.E. Tarjan: Efficient algorithms for finding minimum spanning trees in undirected and directed graphs, Combinatorica 6,2 (1986) 109-122.

[29] M.R.Garey & D.S.Johnson: Computers and intractibility, a guide to the theory of NP-completeness, Freeman 1979.

[30] W.V. Gehrlein & P.C. Fishburn: Coincidnce problems for simple majorities and positional voting rules, Social Sci. Res. 7 (1978) 272-283.

[31] W.V. Gehrlein & P.C. Fishburn: Borda's rule, positional voting, and Condorcet's simple majority rule, Public Choice 26 (1979) 79-88.

[32] Leonidas Georgiadis: Arborescence optimization problems solvable by Edmonds' algorithm, Theoretical Computer Science 301, 1-3 (2003) 427-437

[33] Allan Gibbard: Manipulation of voting schemes, Econometrica 41,4 (1973) 587-602; Manipulation of schemes that mix voting with chance, Econometrica 45,3 (1977) 665-681.

[34] Ronald L. Graham, Donald E. Knuth, Oren Patashnik: Concrete Mathematics, Second Edition, Addison-Wesley, 1994.

[35] M. Grötschel, M. Jünger, G. Reinelt: A cutting plane algorithm for the linear ordering problem, Operations Research 32,6 (Nov.-Dec. 1984) 1195-1220.

[36] M.Grötschel, L.Lovasz, A.Schrijver: Geometric algorithms and combinatorial optimization, Springer 1993 (2nd corrected ed.).

[37] T. Groves & J. Ledyard: Optimal allocation of public goods, a solution to the free rider problem, Econometrica 45 (1977) 783-809.

[38] Edith Hemaspaandra, Lane A. Hemaspaandra, Jörg Rothe: Exact analysis of Dodgson elections: Lewis Carroll's 1876 voting system is complete for parallel access to NP, J. ACM 44,4 (1997) 806-825.

[39] Clarence G. Hoag & George H. Hallett: Proportional Representation, Macmillan, New York 1926; Johnson reprint corp. 1965.

[40] R.M. Karp: A simple derivation of Edmonds' algorithm for optimum branchings, Networks 1 (1971) 265-272.

[41] James P. Keener: The Perron-Frobenius Theorem and the Ranking of Football Teams, SIAM Review 35,1 (March 1993) 80-93.

[42] G.H. Kramer: A dynamical model of political equilibrium, J. Economic Th. 16 (1977) 310-334. P.B. Simpson: On defining areas of voter choice, Quarterly J. Econ. 83 (1969) 478-490.

[43] Enid Lakeman: How democracies vote, Faber & Faber, London 1955 (3rd ed). (There is also a 1959 edition retitled "Voting in democracies" with J.D.Lambert added as a coauthor, but I have not seen it.)

[44] E.L. Lawler: Combinatorial optimization: networks and matroids, Holt, Rinehart and Winston, New York, 1976; Dovwer reprint 2001.

[45] Rob LeGrand: ranked-ballot election methods web site at `http://cec.wustl.edu/~rhl1/rbvote/`.

[46] Jonathan Levin & Barry Nalebuff: An introduction to vote counting schemes, J. Economic Perspectives 9,1 (1995) 3-26.

[47] W.J.M. Mackenzie: Free elections, George Allen & Unwin Ltd. London 1958.

[48] H. Margolis: A thought experiment on demand-revealing mechanisms, Public Choice 38 (1982) 87-91.

[49] H. Margolis: A note on demand-revealing, Public Choice 40 (1983) 217-225.

[50] H.Margolis: Selfishness, altruism, and rationality, Cambridge Univ. Press 1982.

[51] Vincent Merlin, Monica Taturu, Fabrice Valognes: On the likelihood of Condorcet's profiles, Soc. Choice & Welfare 19,1 (2002) 193-206.

[52] N.R. Miller, B. Grofman, S.L. Feld: The structure of the Banks set, Public Choice 66 (1990) 243-251.

[53] Sounaka Mishra & Kripasindhu Sikdar: On approximability of linear ordering and related NP-optimization problems on graphs Discrete Applied Mathematics 136, 2-3 (Feb. 2004) 249-269.

[54] Hervé Moulin: Condorcet's Principle Implies the No Show Paradox, J. Economic Theory 45 (1988) 53-64.

[55] Edward J. Nanson: Methods of election, Royal Society of Victoria (Australia) 1882. Reprinted in I.McLean & R.B.Urken (eds. & translators) *Classics of Social Choice*, Univ. of Michigan press 1995. This book contains pioneer works in Social Choice by Borda, Condorcet, and Nanson.

[56] J. van Newenhizen: The Borda method is most likely to respect the Condorcet principle, Economic Theory 2 (1992) 69-83.

[57] Richard G. Niemi & William H. Riker: The choice of voting systems, Scientific American 234,6 (June 1976) 21-27.

[58] Hannu Nurmi: Voting procedures: a summary analysis, British J. Polit. Sci. 13,2 (1983) 181-208.

[59] Joaquín Pérez: The Strong No Show Paradoxes are a common flaw in Condorcet voting correspondences, Social Choice & Welfare 18 (2001) 601-616.

[60] Benjamin Reilly: Social Choice in the South Seas: Electoral Innovation and the Borda Count in the Pacific Island Countries, Int'l Political Science Review 23,4 (2002) 355-372.

[61] G. Reinelt: The Linear Ordering Problem: Algorithms and Applications, Heldermann Verlag, 1985.

[62] J.T. Richelson: A comparative analysis of social choice functions I, II, III: a summary, Behavioral Science 24 (1979) 355.

[63] J.T. Richelson: Runoff point systems, Public Choice 35 (1980) 457-468.

[64] D.G. Saari: A chaotic exploration of aggregation paradoxes, SIAM Review 37,1 (1995) 37-52.

[65] D.G.Saari: Geometry of voting, Springer-Verlag, Berlin 1994.

[66] Markus Schulze: A New Monotonic and Clone-Independent Single-Winner Election Method, Voting Matters 17 (Oct. 2003) 9-19. (Later expanded version available from Schulze.)

[67] J.H.Smith: Aggregation of preferences with variable electorate, Econometrica 41,6 (1973) 1027-1041.

[68] Warren D. Smith: Range voting, #56 at `http://math.temple.edu/~wds/homepage/works.html`.

[69] Warren D. Smith: "Asset voting" scheme for multiwinner elections, #77 at `http://math.temple.edu/~wds/homepage/works.html`.

[70] Warren D. Smith: Reweighted range voting – new multiwinner voting method, #78 at `http://math.temple.edu/~wds/homepage/works.html`.

[71] Warren D. Smith: Manuscript – work in progress on sarvo-range voting, 2004.

[72] Warren D. Smith: The voting impossibilities of Arrow, Gibbard & Satterthwaite, and Young, #79 at `http://math.temple.edu/~wds/homepage/works.html`.

[73] W.D. Smith, J.N. Quintal, D.S. Greene: What if the 2004 US presidential election had been held using Range or Approval voting? #82 at `http://math.temple.edu/~wds/homepage/works.html`.

[74] Jon C. Sonstelie & Paul R Portney: Truth or consequences: cost revelation and regulation, J. Policy Analysis & Management 2,2 (1983) 280-295.

[75] P.D. Straffin Jr.: Topics in the theory of voting, Boston Birkhäuser 1980.

[76] T.N. Tideman: An experiment in the demand-revealing process, Public Choice 41 (1983) 387-401.

[77] Nicolaus Tideman: The Single transferable Vote, J. Economic Perspectives 9,1 (1995) 27-38. (Special issue on voting methods.)

[78] T. Nicolaus Tideman: Independence of Clones as a Criterion for Voting Rules, Soc. Choice & Welfare 4 (1987) 185-206.

[79] T.N.Tideman & G. Tullock: A new and superior process for making public choices, J. of Political Economy 84 (1976) 1145-1159.

[80] Pravin M. Vaidya: A New Algorithm for Minimizing Convex Functions over Convex Sets. Mathematical Programming 73 (1996) 291-341.

[81] W.Vickrey: Counterspeculation, auctions, and competitive sealed tenders, J.Finance 16 (1961) 1-17.

[82] Gerhard J. Woeginger: A note on scoring rules that respect majority in choice and elimination, Math'l Social Sci. 46,3 (2003) 347-354.

[83] Douglas R. Woodall: Monotonicity of single seat preferential election rules, Discrete Applied Maths. 77,1 (1997) 81-98.

[84] H.P.Young: An axiomatization of Borda's rule, J. Economic Theory 9 (1974) 43-52.

[85] H.P. Young: Social choice scoring functions, SIAM J. Appl. Math. 28,4 (1975) 824-838. Else:

[86] H.P.Young: Condorcet's theory of voting, American Political Science Review 82 (1988) 1231-1244.

[87] H.P.Young & A. Levenglick: A consistent extension of Condorcet's election principle, SIAM J. Appl. Math. 35,2 (1978) 285-300.

[88] B. T. Zavist & N. Tideman: Complete independence of clones in the ranked pairs rule, Social Choice and Welfare 6 (1989) 167-173.