LEWIS CARROLL, VOTING, AND THE TAXICAB METRIC

THOMAS C. RATLIFF

> "The following paper has been written and printed in great haste, as it was only on the night of Friday the 12th that it occurred to me to investigate the subject, which proved to be much more complicated than I had expected."
> – Charles Dodgson, December 18, 1873 [3]

So begins the first of three pamphlets written by Charles Dodgson (aka Lewis Carroll) for his faculty colleagues at Christ Church, Oxford, between 1873 and 1876. Dodgson was quite disturbed by several decisions that had been made by the faculty at the college, especially the selection of students to Christ Church and the choice of the redesign of the college belfry. Since Dodgson did not agree with the decisions, he concluded that the problem was not a lack of support for his positions but must lie with the decision procedure used! For a wonderful account of the way in which these events shaped Dogson's work in elections, see Duncan Black's discussion in *Theory of Committees and Elections* [2] .

In his first pamphlet, Dodgson lays out the trouble with many of the standard voting methods, such as plurality, plurality with elimination, the method of marks, etc. By his third pamphlet, *A Method of Taking Votes on More Than Two Issues* [3], Dodgson displays an amazing insight into many of the fundamental issues that have shaped modern social choice theory. In this pamphlet, Dodgson uses a specific criterion to critique other voting methods, and although he does not explicitly propose it as a method to determine the winner, this criterion has come to be known as Dodgson's method. Since this method is computationally expensive, several techniques have been proposed for approximating the Dodgson winner. The focus of this paper is to show how to interpret Dodgson's method as an almost taxicab metric which helps us understand how different approximations can give different results and which gives us further insight into Dodgson's method.

The Condorcet winner and the Dodgson winner

While there are many ways to pick the winner of an election, one of the most intuitively appealing criteria is to choose the candidate who beats every other candidate in a head-to-head, or pairwise, election. For example, consider the three candidate election, or *voting profile*, given in Figure 1. Here, candidate A defeats B in a head-to-head election by a margin of 2, A defeats C head-to-head by a margin of 2, and B defeats C head-to-head by a margin of 18. The directed graph in Figure 1 provides an easy way to visualize these pairwise outcomes. Since A defeats every other candidate in a head-to-head election, we call A the *Condorcet* winner, after the 18th century French mathematician, the Marquis de Condorcet.

Date: March 5, 2010.

# Voters	Ranking
10	$A > B > C$
1	$A > C > B$
5	$C > A > B$
0	$C > B > A$
9	$B > C > A$
5	$B > A > C$

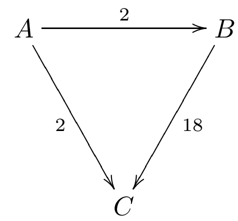


FIGURE 1. A is the Condorcet winner

However, a Condorcet winner may *not* exist. For example, consider Figure 2. In this case, we have a *cycle* among the four candidates. If we take the Condorcet criterion as the gold standard for a winner, how can we break this cycle to pick the candidate who is closest, under some measure, to the Condorcet winner? One option is to pick the candidate whose margins of loss in the pairwise elections are the smallest. In this case, both A and B lose a single pairwise election by the same margin of one. Dodgson's intuition was to look at the preferences of the voters and determine the candidate who requires the fewest *adjacency* switches of candidates in the voting profile in order to become a Condorcet winner.

Using this as our measure, we see that B will become the Condorcet winner if one voter with preference * performs a single adjacency switch between the two top-ranked candidates, changing their preference from $A > B > C > D$ to $B > A > C > D$. A also loses a single pairwise election by a margin of 1 to D, but notice that there is no voter who prefers D to A that has $D > A$ *adjacent* in their preference. Thus, we would require two adjacency switches for A to become the Condorcet winner. For example, the voter with preference † must perform two adjacency switches to change from $D > B > A > C$ to $A > D > B > C$ and reverse their $D > A$ preference. We can easily see that both C and D would require more than one adjacency switch by examining the margins of their pairwise losses. Therefore, B requires the fewest adjacency switches and is the *Dodgson* winner.

# Voters	Ranking
3	$A > B > C > D^*$
1	$D > B > A > C^\dagger$
1	$D > C > A > B$
1	$B > D > C > A$
1	$C > D > B > A$

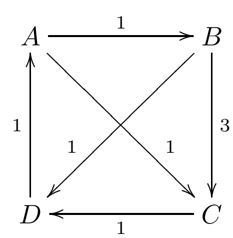


FIGURE 2. No Condorcet winner, B is Dodgson winner

In this example, all four candidates are in the cycle. However, there may be a *top cycle* wherein each candidate is preferred to every candidate outside the cycle. For example, consider Figure 3. Here candidate D is excluded from the top cycle, since it loses elections to all three candidates in the cycle. Dodgson was well aware of this possibility. He would restrict consideration to the candidates A, B, and C in the top cycle and then select the candidate who requires the fewest number of adjacency switches to become the Condorcet winner.

# Voters	Ranking
5	$A > B > C > D$
6	$D > C > A > B$
5	$C > A > B > D$
5	$D > B > C > A$
4	$B > C > A > D^*$
4	$D > A > B > C$
1	$C > D > A > B^\dagger$
1	$B > A > C > D$
1	$B > D > A > C^\dagger$
1	$C > A > B > D$
1	$A > D > B > C^\dagger$
1	$C > B > A > D$

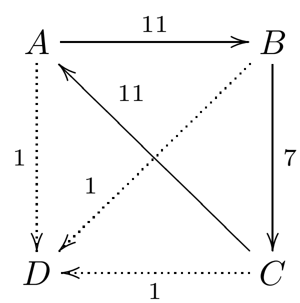


FIGURE 3. D is excluded from the top cycle, C is the Dodgson winner

A unique top cycle always exists in an election without a Condorcet winner. (We follow the argument of Miller in [6].) For any election, a set of candidates $\mathcal{S}$ is *undominated* if every candidate in $\mathcal{S}$ defeats every candidate not in $\mathcal{S}$. Note that a cycle cannot contain candidates from both an undominated set and its complement. An undominated set is *minimal* if no proper subset is undominated. For example, if there is a Condorcet winner, then it constitutes a minimal undominated set. Also notice that an undominated sets always exist since the set of *all* candidates is undominated. (There are no candidates to defeat.) Finally, every election has a unique minimal undominated set. To see this, suppose that $\mathcal{S}_1$ and $\mathcal{S}_2$ are distinct minimal undominated sets. Then neither is a subset of the other, by minimality, and there must exist candidates $A_1 \in \mathcal{S}_1 \setminus \mathcal{S}_2$ and $A_2 \in \mathcal{S}_2 \setminus \mathcal{S}_1$. Since $\mathcal{S}_1$ is undominated, we have that A_1 defeats A_2, and since $\mathcal{S}_2$ is undominated, we have that A_2 defeats A_1, a contradiction. Thus, the minimal undominated set is unique.

If there is no Condorcet winner, then we argue that the minimal undominated set $\mathcal{S}$ forms the top cycle. Since there is no Condorcet winner, every candidate, and in particular every candidate in $\mathcal{S}$, is part of a cycle. As noted above, these cycles must consist entirely of candidates in $\mathcal{S}$ or entirely of candidates not in $\mathcal{S}$. We now need to argue that there is a cycle that contains *every* candidate in $\mathcal{S}$. Pick any cycle in $\mathcal{S}$ and expand it to include as many candidates as possible. If the new cycle does not include every candidate in $\mathcal{S}$, then either every candidate in the cycle defeats every candidate not in the cycle, or else every candidate in the cycle is defeated by every candidate not in the cycle. In either case, we have a proper subset of $\mathcal{S}$ that is an undominated set, contradicting that $\mathcal{S}$ is the minimal undominated set. Thus, there is always a top cycle in any election that does not have a Condorcet winner.

We now state Dodgson's method precisely.

(1) If there is a Condorcet winner, then that is the Dodgson winner.
(2) If not, there is a top cycle. For each candidate in the top cycle, determine the number of adjacency switches in voters' preferences that are necessary

to make the candidate the Condorcet winner. The candidate in the top cycle with the fewest required switches is the Dodgson winner.

Therefore, C would be the Dodgson winner of the election in Figure 3 requiring the four voters with preference * to make one adjacency switch each from $B > C > A > D$ to $C > B > A > D$. Note that D requires only three adjacency switches (one from each voter with preference †) to become the Condorcet winner, but D is excluded from consideration for the Dodgson winner since it is not in the top cycle.

On the surface, Dodgson's method seems like a very plausible extension of the eminently reasonable Condorcet criterion. However, as we will see, one must always be wary of unintended consequences especially when attempting to generalize a voting system.

Difficulties in Calculaton

The examples we have given so far have been fairly simple, with a small number of voters and a small number of candidates. It has not been difficult to determine the Dodgson winner by computing the pairwise outcomes and then examining the individual voters' preferences to determine the number of adjacency switches that are required. Unfortunately, this task becomes much more difficult as the number of candidates and voters increases. In fact, Barth, et al [1] show that determining the Dodgson winner is an NP-hard problem.

In a paper that examines many voting methods, Tideman [11] proposes an approximation for the Dodgson winner based solely on the sum of the margins of loss in the pairwise elections that a candidate loses: The candidate with the smallest sum is used to approximate the Dodgson winner. Table 1 gives the Tideman scores for the examples from Figure 2 and Figure 3. Notice that the Tideman winner for these examples is not the Dodgson winner.

For Figure 2

Candidate	Tideman Score
A	1
B	1
C	4
D	2
A and B tie	

For Figure 3

Candidate	Tideman Score
A	11
B	11
C	7
D	3
D is winner	

Table 1. Tideman Scores

However, in most cases the Tideman approximation does give the Dodgson winner. McCabe-Dansted, et al [5] show that the probability that the Tideman winner corresponds to the Dodgson winner approaches 1 as the number of voters approaches infinity. However, they propose a second approximation for the Dodgson winner, which they call Dodgson Quick, that is very similar to the Tideman approximation but is surprisingly a better predictor of the Dodgson winner.

Let n_{AB} denote the number of voters who prefer candidate A to candidate B in a head-to-head comparison, a notation propopsed in [5]. Then the **advantage of A over B** is

$$\text{adv}(A, B) = \max(0, n_{AB} - n_{BA})$$

In other words, the advantage of A over B looks at the head-to-head comparison of A vs B and is 0 if B defeats A or is the margin of victory if A defeats B. Then the **Tideman score** of a candidate A is

$$T(A) = \sum_{B \neq A} \operatorname{adv}(B, A),$$

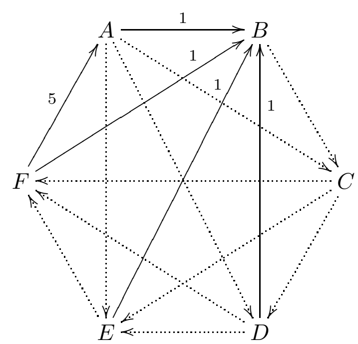
and the **Tideman winner** is the candidate with the lowest Tideman score. Similarly, the **Dodgson Quick score** of a candidate A is

$$DQ(A) = \sum_{B \neq A} \left\lceil \frac{\operatorname{adv}(B, A)}{2} \right\rceil,$$

and the **Dodgson Quick winner** is the candidate with lowest DQ score.

The Tideman score is the total margin of loss for a candidate in all head-to-head elections, while the DQ score slightly adjusts each margin by dividing by 2 and rounding up. To see how this adjustment can make a difference, consider Figure 4. All six candidates here are in the top cycle, but we can limit consideration to A and B, since all other candidates lose multiple pairwise elections by margins of 19, as indicated by the dotted arrows in the diagram.

# Voters	Ranking
19	$A > B > C > D > E > F^\dagger$
12	$F > A > B > C > D > E^*$
12	$E > D > C > B > F > A$
9	$B > A > C > D > E > F$
9	$F > E > D > C > B > A$
10	$F > B > A > C > D > E^\dagger$
10	$E > D > C > A > F > B$
10	$E > B > A > C > D > F^\dagger$
10	$F > D > C > A > E > B$
10	$D > B > A > C > E > F^\dagger$
10	$F > E > C > A > D > B$



Each dotted line has margin 19

FIGURE 4.

Table 2 shows the Tideman and DQ calculations for A and B: B is the Tideman winner; but A is the DQ winner. Further, A is the Dodgson winner since it requires three adjacency switches from three voters of type *, whereas B requires four adjacency switches, one from a voter of each type †. What is going on here? In the next section, we will use a little geometry to explain why these methods can differ.

Note that there are two obvious ways in which *both* of these approximations may fail. First, the Tideman and DQ winners may not be in the top cycle, as in Figure 3. Second, the margins of the pairwise elections may not capture all information about the number of adjacency switches required. In Figure 2 both A and B lose a single pairwise election by a margin of one, but B requires fewer adjacency switches to become the Condorcet winner.

For A	
adv(B, A)	0
adv(C, A)	0
adv(D, A)	0
adv(E, A)	0
adv(F, A)	5

For B	
adv(A, B)	1
adv(C, B)	0
adv(D, B)	1
adv(E, B)	1
adv(F, B)	1

$$T(A) = 0 + 0 + 0 + 0 + 5 = 5$$
$$T(B) = 1 + 0 + 1 + 1 + 1 = 4$$

$$DQ(A) = 0 + 0 + 0 + 0 + \left\lceil \frac{5}{2} \right\rceil = 3$$
$$DQ(B) = \left\lceil \frac{1}{2} \right\rceil + 0 + \left\lceil \frac{1}{2} \right\rceil + \left\lceil \frac{1}{2} \right\rceil + \left\lceil \frac{1}{2} \right\rceil = 4$$

TABLE 2. Tideman and Dodgson Quick calculations for Figure 4

Figure 4 also illustrates a very disturbing property of Dodgson's method (an observation, I believe, originally due to Fishburn [4]). Suppose that each voter brings four friends with exactly the same preference as they have. This has the effect of multiplying by five the number of voters with each ranking and increasing all of the pairwise margins by a factor of five. Repeating the calculations of Table 2 for the new profile gives

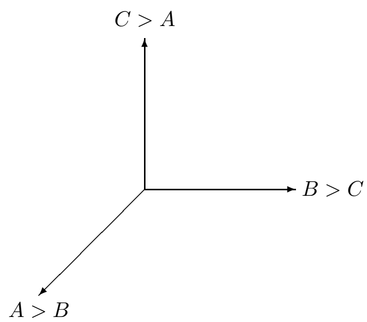
$$T(A) = 0 + 0 + 0 + 0 + 25 = 25$$
$$T(B) = 5 + 0 + 5 + 5 + 5 = 20$$

$$DQ(A) = 0 + 0 + 0 + 0 + \left\lceil \frac{25}{2} \right\rceil = 13$$
$$DQ(B) = \left\lceil \frac{5}{2} \right\rceil + 0 + \left\lceil \frac{5}{2} \right\rceil + \left\lceil \frac{5}{2} \right\rceil + \left\lceil \frac{5}{2} \right\rceil = 12$$

Now B is the Tideman, DQ, *and* Dodgson winner! Merely scaling the profile by a factor of five, gives a different Dodgson outcome. This is entirely unsatisfactory on many levels: Dodgson's method depends on not only the distribution of voters' preferences but also on the absolute number of voters with each preference. This is not, of course, a desirable property for a voting method. It demonstrates that unintended consequences are possible when attempting to extend to situations where it does not apply a reasonable sounding standard like the Condorcet criterion.

DODGSON AND THE TAXICAB METRIC

We use the geometric model developed by Saari ([9], [10]) to explain why the Tideman and DQ winners may differ. In an election with three candidates, a profile specifies the number of voters who prefer each of the six rankings of the candidates. If we fix a particular order for the six rankings, then each election specifies a point in the *profile space* $\mathbb{R}^6$. For example, the election with three candidates in Figure 1 corresponds to the point $(10, 1, 5, 0, 9, 5) \in \mathbb{R}^6$ using the given order for the six rankings. Further, if we pick a specific order for the three pairwise elections, then our profile also defines a point in the *pairwise space* $\mathbb{R}^3$. For example, if we use the ordering $(A > B,\ B > C,\ C > A)$, then the point in pairwise space corresponding to Figure 1 is $(2, 18, -2)$ where the -2 in the $C > A$ coordinate indicates that A defeats C by a margin of two. Notice that the signs of the coordinate values $(+, +, -)$ completely specify the pairwise outcomes $A > B$, $B > C$ and $A > C$, which gives the transitive outcome $A > B > C$. Thus, each octant in $\mathbb{R}^3$ defines a unique pairwise outcome, including all six transitive outcomes with a Condorcet winner, as well as the two cyclic outcomes (see Figure 5). This same framework extends to elections with n candidates: the profile space is $\mathbb{R}^{n!}$; the pairwise space is $\mathbb{R}^{\binom{n}{2}}$ which contains $2^{\binom{n}{2}}$ *orthants*, the higher dimensional analog of quadrants and octants.



Octant	Pairwise outcomes
$(+, +, +)$	$A > B > C > A$
$(+, +, -)$	$A > B > C$
$(+, -, +)$	$C > A > B$
$(+, -, -)$	$A > C > B$
$(-, +, +)$	$B > C > A$
$(-, +, -)$	$B > A > C$
$(-, -, +)$	$C > B > A$
$(-, -, -)$	$C > B > A > C$

FIGURE 5. The three candidate profile space $\mathbb{R}^3$

In an n-candidate election there are $n \cdot 2^{\binom{n-1}{2}}$ orthants with a Condorcet winner: there are n choices of candidates for the Condorcet winner (which determines the outcome for $n - 1$ pairwise elections), and then we can pick any outcome for the remaining $\binom{n-1}{2}$ pairwise elections. Note that with more than three candidates, the number of orthants corresponding to a Condorcet winner is larger than the number of orthants with a complete transitive outcome. That is, we can have a Condorcet winner and still have a cycle among some, or all, of the remaining candidates. For example, with four candidates we could have A as the Condorcet winner but have a cycle among B, C, and D.

Table 3 illustrates how, as the number of candidates increases, the number of orthants without a Condorcet winner dominates the number of orthants with a Condorcet winner. This implies that as the number of candidates increases, it is less likely that we will have a Condorcet winner, and Dodgson's method will require calculating the number of adjacency switches. We should be explicit that this conclusion assumes an *impartial culture*, or that all outcomes in pairwise space are equally likely. Whether or not this assumptions applies in practice has been a source of considerable debate in the social choice community.

We now reframe Tideman's approximation in this geometric framework. If the image of a profile in pairwise space lands in an orthant without a Condorcet winner, then Tideman's approximation is equivalent to finding the orthant with a Condorcet winner that is the shortest l_1 distance from the pairwise point. The l_1 metric, also known as the *taxicab*, or *Manhattan*, metric measures the distance between two points by summing the absolute value of the differences between the coordinates. For example, the taxicab distance between $(1, -3)$ and $(2, 1)$ is $|1-2|+|-3-1| = 5$

# candidates	Profile space	Pairwise space	# orthants with Condorcet winner	# orthants without Condorcet winner
3	$\mathbb{R}^6$	$\mathbb{R}^3$	6	2
4	$\mathbb{R}^{24}$	$\mathbb{R}^6$	32	32
5	$\mathbb{R}^{120}$	$\mathbb{R}^{10}$	320	704
n	$\mathbb{R}^{n!}$	$\mathbb{R}^{\binom{n}{2}}$	$n \cdot 2^{\binom{n-1}{2}}$	$2^{\binom{n}{2}} - n \cdot 2^{\binom{n-1}{2}}$

TABLE 3. Dimensions for profile and pairwise spaces

(as in Figure 6). We think of this as the shortest driving distance between the two points where we are only allowed to travel along streets that run east-west or north-south.

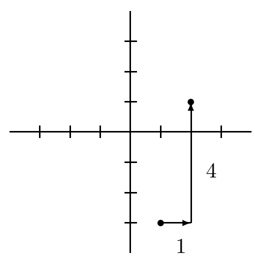


FIGURE 6. The taxicab distance between $(1, -3)$ and $(2, 1)$ is 5

Let us examine the effect on a point in pairwise space of making an adjacency switch in the profile. If we switch a single voter in Figure 1 from $B > C > A$ to $C > B > A$, then we have changed the $B > C$ margin from 18 to 16 since we have reduced B's vote by one and increased C's vote by one, and our point in pairwise space changes from $(2, 18, -2)$ to $(2, 16, -2)$. Thus, an adjacency switch corresponds to moving in steps of size two along a coordinate axis as in the taxicab metric.

If we apply the taxicab metric to Figure 4, the closest orthant with a Condorcet winner comes from moving a distance of one along each of the four axes corresponding to the B vs A, B vs D, B vs E, and B vs F elections. However, this does not capture all of the information required for Dodgson's method. We must move in steps of two along each of these coordinate axes, which causes us not just to touch the coordinate planes of the orthant with a Condorcet winner, but to break through each plane and reach a point in the orthant with integer coordinates. Thus, there is an added fixed cost for each pairwise election that must be reversed. We can also view this as adding a fixed cost to turning a corner in the taxicab metric.

This is how we can understand the advantage of the Dodgson Quick method over Tideman's approximation. By dividing each margin by two and rounding up, the additional cost of turning corners is taken into account. That is, the *number* of pairwise comparisons changed has an effect in Dodgson's method, which DQ detects, whereas this does not come into play in the taxicab approximation. In addition, this explains Fishburn's observation: When the point in pairwise space is fairly close to the coordinate planes, the extra fixed cost of turning corners, or breaking through the coordinate planes, can be large compared to the taxicab distance. However, once we scale further out, this fixed cost is small in comparison to the total distance, yielding a different winner.

Although we will not go into the details here, this geometric framework can be used to create examples that demonstrate many of the pathologies of Dodgson's method. For example, the Dodgson winner can appear at any position in the ranking obtained from the Borda Count, the plurality method or any other positional voting procedure [8], and the Dodgson winner can appear at any position in the ranking obtained from Kemeny's rule [7]. The moral is that we should always be very careful when attempting to extend a desirable property to settings where it does not initially apply.

Summary

The Dodgson winner seems very intuitive and reasonable: When a Condorcet winner does not exist, pick the candidate that is closest, under some measure, to being a Condorcet winner. However, Dodgson's method is very computationally intensive. By placing the method in a geometric framework, we can understand how several approximations to the Dodgson winner can identify different candidates as the winner. Further, this framework provides intuition about some unexpected properties of Dodgson's method.

References

[1] J. Bartholdi III, C. A. Tovey, and M. A. Trick, Voting schemes for which it can be difficult to tell who won the election, *Soc. Choice Welf.* **6** (1989) 157–165.

[2] D. Black, R. A. Newing, I. McLean, A. McMillan, and B. L. Monroe, *The Theory of Committees and Elections*, 2nd ed., Kluwer Academic Publishers, Boston, 1998.

[3] C. L. Dodgson and F. F. Abeles, *The pamphlets of Lewis Carroll*, vol. 3, *The Political Pamphlets and Letters of Charles Lutwidge Dodgson and Related Pieces: a Mathematical Approach*, Lewis Carroll Society of North America, New York, 2001.

[4] P. C. Fishburn, Condorcet social choice functions, *SIAM J. Appl. Math.* **33** (1977) 469–489.

[5] J. C. McCabe-Dansted, G. Pritchard, and A. Slinko, Approximability of Dodgson's rule, *Soc. Choice Welf.* **31** (2008) 311–330.

[6] N. R. Miller, Graph-theoretical approaches to the theory of voting, *American Journal of Political Science* **21** (1977) 769–803.

[7] T. C. Ratliff, A comparison of Dodgson's method and Kemeny's rule, *Soc. Choice Welf.* **18** (2001) 79–89.

[8] ———, A comparison of Dodgson's method and the Borda count, *Econom. Theory* **20** (2002) 357–372.

[9] D. G. Saari, *Basic Geometry of Voting*, Springer-Verlag, New York, 1995.

[10] ———, Mathematical structure of voting paradoxes: I. pairwise votes, *Econom. Theory* **15** (2000) 1–53.

[11] T. N. Tideman, Independence of clones as a criterion for voting rules, *Soc. Choice Welf.* **4** (1987) 185–206.

Department of Mathematics and Computer Science, Wheaton College, Norton, Massachusetts, 02766

E-mail address: `tratliff@wheatoncollege.edu`