2022

An Exploration of Voting with Partial Orders

Mason Acevedo
Harvey Mudd College

An Exploration of Voting with Partial Orders

Mason Acevedo

Michael E. Orrison, Advisor

Darryl H. Yong, Reader

HARVEY MUDD COLLEGE

Department of Mathematics

May, 2022

Abstract

In this thesis, we discuss existing ideas and voting systems in social choice theory. Specifically, we focus on the Kemeny rule and the Borda count. Then, we begin trying to understand generalizations of these voting systems in a setting where voters can submit partial rankings on their ballot, instead of complete rankings.

Contents

Acknowledgments

I'd like to thank several folks in the HMC Math Department. Firstly, Professor Orrison, for teaching me math and patiently guiding me through this project. I'd like to thank Professors Jacobsen, Yong, and Karp for showing me how to be a human being first, and a mathematician second.

Additionally, I'd like to thank some friends for their support, encouragement, and belief in me, especially when I doubted myself. In no particular order: Thank you, Kailee Lin. Thank you, Katie Partington. Thank you, Hannah Larson. Thank you, Addissyn House. Thank you, Katheryn Wang. Thank you, Ruth Alemu Mekonnen. Thank you, Devon Tao. Thank you, Lars Vlahakis. Thank you, Athena Li. Thank you, Amit Harlev. Thank you, Lian Morales. Thank you, Will Gilroy.

Chapter 1

Introduction

1.1 Why Study Voting?

Human beings are social creatures. We often have to make collective decisions, despite differences in opinion, and decide on one alternative among many. In other words, we often have to *vote.*

In everyday discussions surrounding voting, we often take for granted that putting a decision up to a vote is often "the fair thing to do." We implicitly assume that "putting it up to a vote" is a method which accurately captures the collective desire of the group. However, voting is often not so simple. To illustrate one of the less obvious complexities inherent to voting, let us look at a specific example.

Imagine a group of 17 people who are voting amongst 3 candidates, who we will name a, b, and c. Suppose that 7 people think that a is the best candidate, 4 people think that b is the best candidate, and 6 people think that c is the best candidate. The following table summarizes the results:

	Results
a	7 votes
c	6 votes
b	4 votes

Now, take a moment to imagine a slightly different, but still reasonable voting system. Instead of each person simply voting for their favorite candidate, each person submits a full ranking which describes their preferences. So, in addition to saying their favorite candidate, everyone will also say who their 2nd and 3rd favorite candidates are.

In this case, we can represent the opinions of our voters in the following table:

abc	acb	bac	bca	cab	cba
5	2	1	3	1	5

The above table is read as follows: The 5 in the abc column means that 5 people put a above b, and b above c. The 2 in the acb column means that 2 people ranked a above c, and c above b, etc.

Note that this voting body could be the same as the one we mentioned before: 7 people still think a is the best candidate, 4 people still think b is the best candidate, and 6 people still think c is the best candidate.

Now, how do we run an election with this additional information? Well, one thing to do would be to assign points to candidates as follows:

- Give 2 points to a candidate whenever a voter puts them as their favorite candidate.
- Give 1 point to a candidate whenever a voter puts them as their middle candidate.
- Give 0 points to a candidate whenever a voter puts them as their least favorite candidate.

Now, if we use this voter information, we can run the election with the above points-based system. We summarize the results below.

	Results
b	18 points
c	17 points
a	16 points

Wow! Comparing this to our first election, we got a completely different result! b wins instead of a!

Let us contemplate what we just observed. We took a group of people, and we had them vote under two different voting systems which both seemed fair and reasonable. But we got two completely different results! How do we make sense of this?

Well, to start, take another look at the table with our voters' preferences and notice that candidate a gets 7 first place votes, 2 second place votes, and 8 third place votes. In other words, candidate a is polarizing. This strategy works fine in our first system, but in our second system, having a lot of last place votes means you lose out on a lot of points.

On the other hand, candidate b only has 4 first place votes, but they have 10 second place votes and 3 last place votes. In short, b is not as fiercely loved as a, but b does have broad approval.

Hopefully, we've made some sense of that surprising result. Nevertheless, this observation—that the same people voting under different systems can obtain extremely different outcomes—is what motivates much of voting theory. In what scenarios will two voting systems agree or disagree? What are the benefits and drawbacks of those different voting systems? Is there a "best" voting system that we should always use?

1.2 A Brief History of Voting Theory, and Some Terminology

The ideas discussed in the previous section are hardly new. As early as 1770, mathematician Jean-Charles de Borda proposed the second system we looked at, in which voters submit full-rankings and candidates received points according to how highly they're ranked (Borda (1784)). This voting method has since been named after him, and is called the *Borda count.* Despite the Borda count being named after Borda, it's worth noting that Nicholas of Cusa came up with this voting system in 1435, and it has been independently developed several times (Emerson (2016)).

While we're naming voting systems, the first voting system we looked at, in which everyone simply votes once for their favorite candidate, is called *plurality voting.* Furthermore, we will refer to the set of preferences of a group of voters as a *voter profile,* or sometimes just a *profile* for short. Note that a table similar to the one we saw previously is one way to represent a voter profile, but there are other ways to represent this information. Now, back to some history.

In 1785, mathematician Marie Jean Antoine Nicolas de Caritat, Marquis de Condorcet (whom we'll now refer to as simply Condorcet) published a paper in which he criticized the Borda count for failing to have a certain "fairness" property (Condorcet (1785)).

Specifically, given a voter profile, there may exist a candidate who wins in a head-to-head race against all other candidates, but is not selected by as the winner by the Borda count. To illustrate this, consider the following profile:

abc	acb	bac	bca	cab	cba
3	0	0	2	0	0

Notice that, among the 5 voters, 3 of them prefer a to b, and 3 of them prefer a to c. Thus, if a were to run in a two-candidate election against b, then a would win 3-2. Similarly, if a were to run in a two-candidate election against c, then a would win 3-2. Such a candidate who wins a head-to-head race against all other candidates is called a *Condorcet winner*, so a is the Condorcet winner of this election. One could argue that a Condorcet winner should *always* win an election. After all, one might ask candidate a how they're feeling about the election, and a might respond "Well, I win against every other candidate, so I think I should win."

With this in mind, let's see what the election results are if we run the Borda count:

	Results
b	7 points
a	6 points
c	2 points

This is troubling! The Borda count has failed to elect a, the Condorcet winner, and has instead elected b instead! We can make sense of this result by realizing that, yes, 3 of the 5 voters prefer a to b, **but** the other 2 voters preferred b to *a with greater intensity.* Notice that the 3 voters who prefer a to b have b as the middle candidate, while the 2 voters who prefer b to a have a ranked as the last candidate. So, when it comes to the Borda count, it's not just about being preferred, but it's about *how much* a candidate is preferred.

In this specific example, the Borda count failed to elect the Condorcet winner, but in other scenarios, the Borda count is totally capable of electing the Condorcet winner—it's just not guaranteed.

Furthermore, note that some voter profiles do not have a Condorcet winner. As an example, consider the following profile:

abc	acb	bac	bca	cab	cba
1	0	0	1	1	0

This profile has 3 voters. Two of them prefer a to b. Two of them prefer b to c. One might expect from the transitive property that the voters also prefer a to c. But actually, 2 of the 3 voters prefer c to a, the opposite! It seems that our voter profile as a collective is saying that $a > b > c > a$. For this reason, this voter profile is called *cyclic.* Notice that this occurs in spite of the fact that *no individual preference is cyclic*. That is, if a voter prefers x to y and y to z, that voter also prefers x to z.

This observation that a profile can have cyclic preferences even when no individual has cyclic preferences is called *Condorcet's paradox*, and is one of the ways that a voter profile may fail to have a Condorcet winner.

Condorcet felt strongly that if a Condorcet winner exists, they should always win. Thus, he deemed that the Borda count was flawed. However, Condorcet did not specify an election system that does have this property. At this point, it is again useful to define a small piece of terminology: We say that an election system is *Condorcet*, or satisfies the *Condorcet Criterion* if that system *always* elects the Condorcet winner when such a candidate exists. With this defined, we turn our attention to a Condorcet voting system.

In 1959, John G. Kemeny published a paper in which he presented voting as a distance minimization problem, instead of as a point-assigment problem (Kemeny (1959)). As far as the author knows, Kemeny was among the first to use this approach.

In Kemeny's paper, he defines a *metric* on the set of rankings, giving us a notion of distance between rankings. Then, loosely speaking, the consensus ranking in an election is the ranking which is, on average, closest to whichever rankings have been voted for. Working in this framework, Kemeny identified two possible procedures for identifying consensus rankings. He did not specify one procedure as being more desirable than another.

A subtle, but important distinction in Kemeny's work is that this distance approach gives a natural way to assign scores to arbitrary permutations of the candidates, rather than assigning scores to individual candidates themselves. Contrasting this with the Borda count, it is not immediately obvious how one would use the Borda count to assign points to arbitrary permutations of the candidates.

Building off Kemeny's 1959 work, H.P. Young and A. Levenglick found that only one of the two voting systems Kemeny put forward was a *consistent* method. This system came to be known as the Kemeny-Young rule, (or the Kemeny rule) and it can be characterized as the *unique* voting procedure which is Condorcet, consistent, and neutral (Young and Levenglick (1978)). We have defined what it means for a voting system to be *Condorcet,* but not consistent or neutral, so we define these terms now.

A voting system is said to be *consistent* if, whenever a group of profiles all independently elect some candidate X, then combining those profiles together into one aggregate profile and running the election on that aggregate profile also results in X winning the election. Instant-runoff voting does not obey this property, because there are certain groups of profiles that, when aggregated together, do not elect the same candidate that each individual group did. Finally, a voting system is *neutral* if permuting the labels of the candidates in an election causes the election results to be permuted in the same way. We can interpret this as saying that no candidate receives special treatment. Every voting system we have discussed so far is neutral.

Hopefully, the reader has some appreciation of the result that the Kemeny rule is the unique voting procedure which is consistent, Condorcet, and neutral. At this point, one might wonder: Is this proof that the Kemeny rule is the *best* voting system?

As nice as the Kemeny-Young rule is, it's not perfect. One of its shortcomings is that it's extremely computationally expensive to compute for elections with many candidates. On a deeper level, it fails to have certain fundamental properties that we might reasonably expect our voting systems to have, in a similar way that the Borda count failed to meet the Condorcet Criterion. However, these flaws are not unique to the Kemeny Rule and Borda count. In fact, economist Kenneth Arrow proved that a set of three "fairness" criteria, each similar in flavor to criteria we briefly defined above, could not coexist in any voting system (Arrow (1950)). This result has come to be known as Arrow's Impossibility Theorem, and other similar results have been proven since.

As fascinating as it is, a deep discussion of these kinds of "fairness" properties is beyond the purview of this thesis. To learn more, simply type "Arrow's Impossibility Theorem" into a search engine.

Now, we turn our attention back to topics that are more directly relevant to this thesis.

1.3 Voting With Partial Orders

All of the voting systems we've considered so far are systems in which voters submit a complete ranking of the candidates. That is, no voter submits a ballot with a tie or multiple ties. Submitting a ballot with ties may seem like an edge case, a niche feature that isn't necessary, but we argue that allowing ties enables individuals to express more complex, nuanced opinions on their ballots.

For example, suppose that you and a group of friends are going to eat at a restaurant, and you are trying to choose between restaurants a, b, and c. You might really like restaurant a, but you've never been to b or c, in which case you might want your vote to say "I prefer restaurant a to both b and c, but b and c are incomparable." For another example, suppose you're trying to pick a movie to watch, and you have 4 movies to choose from—two sci-fi and two horror. It may be the case that within sci-fi, you prefer movie a to b, and within horror, you prefer movie x to y, but you don't care whether you watch sci-fi or horror. For one last example, suppose a pollster is polling people about a large set of political candidates. People might not know certain candidates and thus can't compare them to others. Alternatively, someone might decide that they want to move on with their day and only give the pollster an incomplete set of information. In these cases, a ballot that can represent preferences with many ties is not just a nice feature, but is essential.

More formally, we're interested in voting systems where voters submit some *partial ordering* on the candidates, rather than a *total* or *complete* order. Luckily for us, in 2014, Cullinan et al. proposed a generalization of the Borda count to this new setting where voters submit partially ordered ballots (Cullinan et al. (2014)). We will refer to this generalization as the partial Borda count for the rest of this thesis. There are a couple reasons that the partial Borda count can reasonably be thought of as a generalization of the existing Borda count, rather than an entirely distinct voting system. Firstly, if every voter decides that their ballot contains no ties, and they end up submitting a complete ordering on the candidates, then the partial Borda count always gives the same result as the original Borda count. Furthermore, the traditional Borda count is characterized as the unique voting procedure that accepts totally ordered ballots and is consistent, faithful, neutral, and has the cancellation property. Similarly, Cullinan et al. proved that the proposed partial order version of the Borda count was the unique voting procedure that accepts partially ordered ballots and is consistent, faithful, neutral, and has the cancellation property (Cullinan et al. (2014)).

As far as the author can tell, there is not an analogous poset version of the Kemeny rule that has been characterized by its "fairness" properties. That is, while Young and Levenglick showed that the Kemeny rule is the unique voting procedure which is neutral, consistent, and Condorcet, there is no analogous result for voting systems in which voters can submit partially ordered preferences.

However, similar to the origins of the Kemeny rule, authors have proposed a notion of distance between partial orders (Bogart (1973); Cook et al. (1986)).

While the work of these authors is closely related to this thesis, there are subtle differences. Firstly, the metric introduced in Bogart's 1973 paper does not explicitly discuss using the metric for voting or social choice purposes. Cook et al.'s 1986 paper uses a slightly different definition for what a partial order is—namely, they make a distinction between two candidates being "tied" versus simply "incomparable" whereas we do not.

1.4 Where We're Going

At this point, we hope the reader has a reasonable grasp on some foundational ideas in social choice theory. Now, we'd like to summarize the contents in the rest of this document. Chapter 2 discusses the Borda count, and the Kemeny rule in more detail. Chapter 3 defines and briefly discusses partial orders. Chapter 4 discusses the partial Borda count in more detail, as well as one possible extension of the Kemeny rule to the partial order setting. Chapter 4 also discusses some relationships between the partial Borda count and poset Kemeny rule. Chapter 5 details some computational work that I have done as part of this thesis, and various avenues for research related to voting with partial orders.

Chapter 2

The Borda Count and the Kemeny Rule

2.1 The Borda Count

We introduced the Borda count in the introduction. Here, we will elaborate on it and briefly discuss some of its properties.

The Borda count is a voting procedure which assigns points to candidates based on how highly they are ranked by each voter. Specifically, in an election on n candidates, the Borda count gives $n-1$ points to a candidate when it is a voter's first choice, $n-2$ points when it is a voter's second choice, $n-3$ points when it is a voter's third choice, etc. The Borda count is perhaps the simplest or most natural thing to do that incorporates the entirety of a voter's complete ranking.

Note that this distribution of points is somewhat arbitrary—the resulting ranking would be the same if we instead assigned k points to someone's last choice, $k+1$ points to the second to last, $k+2$ points to the third to last, etc., and finally $k+n$ points to the first choice candidate, for *any* choice of k.

More formally, when we run the Borda count, we get a choice of a *weighting vector*, which we will denote $\vec{w}$. A weighting vector is simply a vector which encodes the number of points assigned to each candidate. In the classic example, our weighting vector $\vec{w}$ is $\vec{w} = [n-1, n-2, \ldots, 2, 1, 0]$. But we just observed that, all weighting vectors of the form $\vec{w} = [k+n-1, \ldots, k+2, k+1, k]$ for some $k \in \mathbb{R}$ will result in the same election outcome.

Similarly, if we instead scale our classic weighting vector by some positive number, we still don't change the results. That is to say, a weighting vector of the form $\vec{w} = [a(n-1), a(n-2) \ldots, 2a, a, 0]$ also defines the same election system as the classic weighting vector when $a > 0$.

A voting system that can be understood as a point-assignment system with some weighting vector $\vec{w}$ is called a *positional* voting system. So far, we have only discussed positional voting systems which always give the same result as the Borda count, but there *are* weighting vectors that result in legitimately distinct voting procedures. For example, the island nation Nauru uses the weighting vector $w = [1, \frac{1}{2}, \frac{1}{3}, \ldots \frac{1}{n}]$. (Fraenkel and Grofman (2014)). Furthermore, now that we've introduced the weighting vector, we can notice that plurality voting can be thought of as a positional voting system with weighting vector $\vec{w} = [1, 0, 0, \ldots, 0]$.

2.1.1 The Borda Count as a Linear Transformation

As the subsection title suggests, we can also encode the Borda count as a matrix-vector multiplication. Let us see how to do this.

Suppose we want to run an election on three candidates, a, b, and c. Recall the introduction, which featured such an election. Here is a voter profile we used:

abc	acb	bac	bca	cab	cba
5	2	1	3	1	5

Now, we can encode this information in a vector $v \in \mathbb{R}^6$ as follows:

$$v = \begin{pmatrix} 5 \\ 2 \\ 1 \\ 3 \\ 1 \\ 5 \end{pmatrix}.$$

Note that this vector has 6 entries because there are $3! = 6$ possible rankings on three candidates. In general, with an election on n candidates, a profile vector will have $n!$ entries. Furthermore, note that this is an entirely different kind of vector than the weighting vectors discussed in the previous section. The weighting vectors are used to define voting systems, whereas the above vector encodes information about a particular voter profile.

There exists a 3×6 matrix which, when multiplied against this vector, will give us the results of our election. Specifically, the columns of this matrix will be indexed by the 6 different rankings, and the rows will be indexed by the candidates. This, in general, this will be a $n \times n!$ matrix. Let's look at the 3x6 matrix, which we'll denote B:

$$B = \begin{array}{c} \\ \\ \\ \end{array}\begin{array}{cc} \begin{array}{cccccc} abc & acb & bac & bca & cab & cba \end{array} & \\ \begin{pmatrix} 2 & 2 & 1 & 0 & 1 & 0 \\ 1 & 0 & 2 & 2 & 0 & 1 \\ 0 & 1 & 0 & 1 & 2 & 2 \end{pmatrix} & \begin{array}{c} a \\ b \\ c \end{array} \end{array}$$

To create this matrix, we first decide on a weighting vector. We will use $\vec{w} = [2, 1, 0]$. Then, we look at a column, which corresponds to a certain ranking, and a row, which corresponds to a certain candidate, and ask how many points this candidate will receive for that ranking. For example, we write a 1 in column bac, row a, because candidate a gets 1 point when a voter has bac on their ballot.

Then, multiplying this matrix against our profile vector v, we have:

$$Bv = \begin{pmatrix} 2 & 2 & 1 & 0 & 1 & 0 \\ 1 & 0 & 2 & 2 & 0 & 1 \\ 0 & 1 & 0 & 1 & 2 & 2 \end{pmatrix} \begin{pmatrix} 5 \\ 2 \\ 1 \\ 3 \\ 1 \\ 5 \end{pmatrix} = \begin{pmatrix} 16 \\ 18 \\ 17 \end{pmatrix}.$$

The rows of our matrix are indexed by candidates, so the rows of this resulting vector are indexed by candidates as well. So, we can read this resulting vector as saying candidate a scored 16 points, candidate b scored 18, and candidate c scored 17.

Now that we've encoded the Borda count as a linear transformation, we suddenly can use linear algebra to study it. Immediately, we see that B is a function whose domain is $\mathbb{R}^6$, but whose codomain is $\mathbb{R}^3$. Thus, the Borda count has a 3-dimensional nullspace, i.e, there is a 3-dimensional subspace of $\mathbb{R}^6$ that is mapped to the zero vector.

2.2 The Kemeny Rule

2.2.1 The Kemeny Rule as a Point Distribution System

The Kemeny rule is a voting procedure which uses the same information as the Borda count, but in a very different way. Namely, it assigns points to *rankings* of candidates rather than the candidates themselves. Specifically, it looks at how any given ranking of the candidates aligns with the rankings that its presented with in the ballots. Let us see how to compute it.

Suppose we have the following profile:

abc	acb	bac	bca	cab	cba
2	6	4	5	1	4

Then, to assign points to some ranking, say, abc, we count how many pairwise comparisons it agrees with in the voting body. The ranking abc has $a > b$, $b > c$, and $a > c$, so let's see how these opinions align with the voting body.

- There are 9 voters who agree that a is better than b, so abc gets 9 points for that.
- There are 11 voters who agree that b is better than c, so abc gets 11 points for that.
- Finally, there are 12 voters who agree that a is better than c, so abc gets 12 points for that.

In total, the ranking abc gets 9 + 11 + 12 = 32 points.

For another example, let us compute the number of points cba gets:

- There are 11 voters who agree that c is better than b, so cba gets 9 points for that.
- There are 10 voters who agree that c is better than a, so cba gets 10 points for that.
- Finally, there are 13 voters who agree that b is better than a, so cba gets 13 points for that.

In total, ranking cba gets 11 + 10 + 13 = 34 points.

Doing this process for all 6 rankings, we get:

Ranking	Score
abc	32
acb	32
bac	36
bca	34
cab	30
cba	34

Thus, bac is the winning *ranking.* If we wanted to choose a winning candidate from here, we would choose candidate b.

It's worth noting that as the number of candidates increases, the number of computations required to run the Kemeny rule increases drastically. To see this, simply note that the Kemeny rule assigns some number of points to every single ranking of the n candidates—there are $n!$ such rankings, so the Kemeny rule requires, at minimum, $n!$ operations to compute. We did not run into the same difficulty with the Borda count.

It's worth noting that we only run into this $n!$ issue as the number of candidates increases—if we fix n at a relatively small number of candidates, the number of voters can grow considerably without incurring these same runtime problems.

2.2.2 The Kemeny Rule as a Linear Transformation

Similarly to the Borda count, we can encode a voter profile as a vector, and run an election as a matrix-vector multiplication. Consider the following voting profile:

abc	acb	bac	bca	cab	cba
2	6	4	5	1	4

Then our vector v is

$$v = \begin{pmatrix} 2 \\ 6 \\ 4 \\ 5 \\ 1 \\ 4 \end{pmatrix}.$$

However, constructing a Kemeny matrix K is a little different than constructing our Borda matrix. Specifically, the Borda count assigned points to candidates, so the matrix's rows were indexed by the candidates. However, the Kemeny rule assigns points to rankings, so the Kemeny matrix's rows will be indexed by rankings. This matrix is:

$$K = \begin{array}{c c} \begin{array}{cccccc} abc & acb & bac & bca & cab & cba \end{array} & \\ \begin{pmatrix} 3 & 2 & 2 & 1 & 1 & 0 \\ 2 & 3 & 1 & 0 & 2 & 1 \\ 2 & 1 & 3 & 2 & 0 & 1 \\ 1 & 0 & 2 & 3 & 1 & 2 \\ 1 & 2 & 0 & 1 & 3 & 2 \\ 0 & 1 & 1 & 2 & 2 & 3 \end{pmatrix} & \begin{array}{c} abc \\ acb \\ bac \\ bca \\ cab \\ cba \end{array} \end{array}.$$

To fill in the entries for this matrix, look at the row ranking and the column ranking, and ask for the number of pairwise preferences that they agree on. For example, abc and acb both agree that $a > b$, and $a > c$, but they disagree on the placement of b and c. So out of the 3 possible pairwise opinions to have, they agree on 2. So we put a 2 in the column indexed by abc and the row indexed by acb. We also put a 2 in the row indexed by abc and the column indxed by acb. This explains why the matrix is symmetric—for each pair of rankings, there are 2 places in the matrix that compare those rankings, and those places will have the same entry.

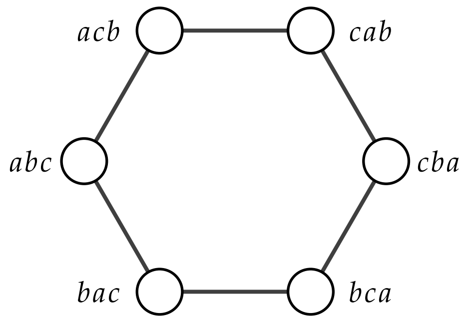
Then, computing the matrix-vector multiplication, we get:

$$Kv = \begin{pmatrix} 3 & 2 & 2 & 1 & 1 & 0 \\ 2 & 3 & 1 & 0 & 2 & 1 \\ 2 & 1 & 3 & 2 & 0 & 1 \\ 1 & 0 & 2 & 3 & 1 & 2 \\ 1 & 2 & 0 & 1 & 3 & 2 \\ 0 & 1 & 1 & 2 & 2 & 3 \end{pmatrix} \begin{pmatrix} 2 \\ 6 \\ 4 \\ 5 \\ 1 \\ 4 \end{pmatrix} = \begin{pmatrix} 32 \\ 32 \\ 36 \\ 34 \\ 30 \\ 34 \end{pmatrix}.$$

Finally, 36 is the highest entry in the resulting vector, and that entry corresponds to the ranking bac, so bac is our winning ranking with 36 points. Note that this matches exactly with our computation in the previous section.

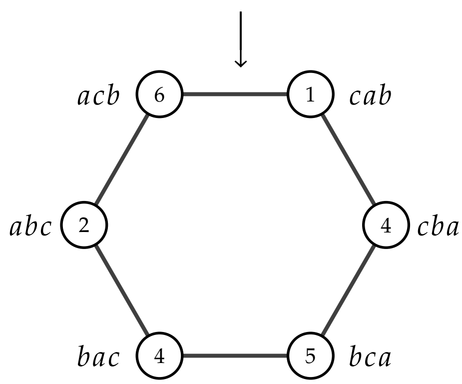
2.2.3 The Kemeny Rule as a Distance Minimization Problem

Now, we present another way to think of the the Kemeny rule that has a very different flavor. Given n candidates, we begin by constructing the *permutohedron.* This permutahedron is a specific type of graph, defined as follows: The vertices are the $n!$ permutations of the n candidates, and two vertices are adjacent if and only if they differ by one adjacent transposition. Constructing the permutohedron on 3 elements, we have:



Then we can assign each vertex a weight, which is equal to the number of voters who submitted that vertex's ranking as their ballot. So continuing with one of our previous voter profiles, we have:

abc	*acb*	*bac*	*bca*	*cab*	*cba*
2	6	4	5	1	4

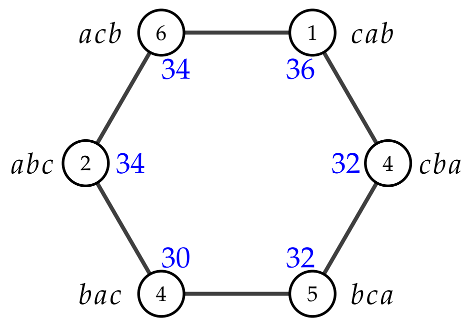


Now that we've assigned weights to vertices, we can ask the question: Which vertex minimizes the "weighted distance" to all the other vertices?

More formally, we can assign each vertex a score as follows: Let $S(v)$ denote the score given to vertex v. Furthermore, let $p(v)$ be the weight assigned to vertex v, i.e., the number of voters who chose the permutation corresponding to that vertex on their ballot. Finally, given two vertices v, w let $d(v, w)$ be the distance between them, where d is the usual metric on a graph. Then, $S(v)$ is given by:

$$S(v) = \sum_{w \in V(G)} d(v, w)p(w).$$

Then compute $S(v)$ for all vertices $v \in V(G)$. Finally, the vertex with minimum $S(v)$ is the winner of the Kemeny rule election. For convenience, we will show $S(v)$ in blue near every vertex:



Thus, bac is our winning ranking, because it has the lowest score.

While this perspective on the Kemeny rule may feel very different from our previous two perspectives, it actually works quite nicely with our understanding of the Kemeny rule as a linear transformation. We can think of constructing the matrix K by looking at pairwise agreement between rankings, but we can *also* construct it in a different way using this graph perspective. Specifically, we can construct a *distance matrix* D, where the rows and columns are indexed by rankings, and the (i, j) entry is simply the distance between the ranking i and ranking j. Constructing this distance matrix, we have:

$$D = \begin{array}{cc} \begin{array}{cccccc} abc & acb & bac & bca & cab & cba \end{array} & \\ \begin{pmatrix} 0 & 1 & 1 & 2 & 2 & 3 \\ 1 & 0 & 2 & 3 & 1 & 2 \\ 1 & 2 & 0 & 1 & 3 & 2 \\ 2 & 3 & 1 & 0 & 2 & 1 \\ 2 & 1 & 3 & 2 & 0 & 1 \\ 3 & 2 & 2 & 1 & 1 & 0 \end{pmatrix} & \begin{array}{c} abc \\ acb \\ bac \\ bca \\ cab \\ cba \end{array} \end{array}.$$

Then to get our Kemeny Matrix K, we simply take an entry a_{ij} and replace it with $3 - a_{ij}$. Thus we have:

$$3J_6 - D = 3J_6 - \begin{pmatrix} 0 & 1 & 1 & 2 & 2 & 3 \\ 1 & 0 & 2 & 3 & 1 & 2 \\ 1 & 2 & 0 & 1 & 3 & 2 \\ 2 & 3 & 1 & 0 & 2 & 1 \\ 2 & 1 & 3 & 2 & 0 & 1 \\ 3 & 2 & 2 & 1 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 3 & 2 & 2 & 1 & 1 & 0 \\ 2 & 3 & 1 & 0 & 2 & 1 \\ 2 & 1 & 3 & 2 & 0 & 1 \\ 1 & 0 & 2 & 3 & 1 & 2 \\ 1 & 2 & 0 & 1 & 3 & 2 \\ 0 & 1 & 1 & 2 & 2 & 3 \end{pmatrix} = K.$$

(J_6 denotes the 6×6 matrix that has a 1 for every entry.)

This may seem surprising, but we can actually make sense of this: we originally filled in the entries of the K matrix by looking at how much two rankings *agreed* with each other. But to fill in the distance matrix D, we instead look at how much they *disagree* with each other.

Furthermore, instead of running an election by computing Kv and looking for the maximum score, we can instead compute Dv and look for the *minimum* score.

Now that we know how the Kemeny rule works, (in 3 different ways!) we can discuss the differences between the Kemeny rule and the Borda count.

2.2.4 Comparing the Kemeny Rule and the Borda Count

At this point, one could argue that the Kemeny rule is significantly more complicated than the Borda count. While the Borda count involves a relatively straightforward points system that most people will understand relatively quickly, computing the Kemeny rule involves the somewhat unintuitive idea of assigning points to *rankings* instead of *candidates.* So, why bother with this more complicated technique? What do we get by doing all this?

Well, for one thing, recall from the introduction that the Kemeny rule is *Condorcet*, while the Borda count is not. Recall the following profile from the introduction:

abc	acb	bac	bca	cab	cba
3	0	0	2	0	0

The Borda count, with weighting vector $w = [2, 1, 0]$ outputs

b	7
a	6
c	2

Thus, b is the winner of this election, even though a is the Condorcet winner. (a beats b 3-2, and a beats c 3-2.)

Now, using this same profile in a Kemeny style election instead, we have:

Ranking	Score
abc	11
acb	6
bac	10
bca	9
cab	5
cba	4

Thus, abc is the winning ranking, and if we want to select an individual candidate, this makes a our winner, which does match our expectation that the Kemeny rule is Condorcet. We do not prove here that the Kemeny rule always selects the Condorcet winner. For that proof, see Kemeny (1959).

However, this does not necessarily mean that the Kemeny rule is perfect. Recall that the Kemeny rule is computationally expensive to run. Even more specifically, the Borda count runs in polynomial time, while the Kemeny rule does not.

While the Kemeny rule and Borda count are dissimilar in these ways, there are other ways in which they are quite similar.

One perspective that relates the Borda count and the Kemeny rule is that both the Borda matrix B and the Kemeny matrix K can be factored as follows:

$$B = QP$$
$$K = P^T P$$

where P is defined as

$$P = \begin{array}{c} \\ a>b \\ b>a \\ a>c \\ c>a \\ b>c \\ c>b \end{array} \begin{array}{c} \begin{array}{cccccc} abc & acb & bac & bca & cab & cba \end{array} \\ \begin{pmatrix} 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 1 \end{pmatrix} \end{array}.$$

This matrix simply records information about the pairwise preferences given in each ranking (Crisman and Orrison (2017)). So, taking the transpose of this matrix and multiplying it against itself, we get:

$$P^T P = \begin{array}{c c} & \begin{array}{cccccc} abc & acb & bac & bca & cab & cba \end{array} \\ \begin{array}{c} abc \\ acb \\ bac \\ bca \\ cab \\ cba \end{array} & \begin{pmatrix} 3 & 2 & 2 & 1 & 1 & 0 \\ 2 & 3 & 1 & 0 & 2 & 1 \\ 2 & 1 & 3 & 2 & 0 & 1 \\ 1 & 0 & 2 & 3 & 1 & 2 \\ 1 & 2 & 0 & 1 & 3 & 2 \\ 0 & 1 & 1 & 2 & 2 & 3 \end{pmatrix} \end{array} = K.$$

Now, we can factor the matrix B in a similar way. Namely, we have $B = QP$, where

$$Q = \begin{array}{c c} & \begin{array}{cccccc} a>b & b>a & a>c & c>a & b>c & c>b \end{array} \\ \begin{array}{c} a \\ b \\ c \end{array} & \begin{pmatrix} 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 \end{pmatrix} \end{array}.$$

To construct Q, we have rows indexed by candidates, and columns indexed by pairwise comparisons of candidates. A candidate gets a 1 in a certain column if it "wins" that comparison, and a 0 otherwise.

Then, calculating QP, we have:

$$QP = \begin{array}{c c} \begin{array}{cccccc} abc & acb & bac & bca & cab & cba \end{array} & \\ \begin{pmatrix} 2 & 2 & 1 & 0 & 1 & 0 \\ 1 & 0 & 2 & 2 & 0 & 1 \\ 0 & 1 & 0 & 1 & 2 & 2 \end{pmatrix} & \begin{array}{c} a \\ b \\ c \end{array} \end{array} = B.$$

The fact that we can write $QP = B$ and $P^T P = K$, where P is the same matrix in both equations, means that the Kemeny rule and the Borda count are deeply related. Namely, we can think about both the Borda count and Kemeny rule as compositions of two different linear transformations, but the first transformation is the same! In other words, they're both looking at the pairwise information obtained by multiplying the profile by P, then they proceed to do different things with that information.

Chapter 3

What's a Poset?

3.1 Defining a Partial Order

The definitions in this section are from the textbook *Mathematics: A Discrete Introduction* by Edward A. Scheinerman (Scheinerman (2012)).

Before giving the formal definition of a partial order, we informally discuss the idea that the definition tries to encapsulate.

Informally, a partially ordered set (or poset for short) is a set of elements in which there is some generalized notion of *order.* Partial orders give us a precise way to think about what it means to say that some elements of a set are greater than others. However, partial orders also allow for the possibility that two elements have incomparable order. For voting purposes, this gives voters the freedom to rank some candidates as better than others, or to declare some candidates as tied, or to not making a decision about some pairs of candidates. Let's look at the definition.

Definition: A partial order is an ordered pair $P = (X, R)$, where X is a non-empty set and R is a relation on X that satisfies the following properties:

- R is reflexive: $\forall x \in X,\ xRx$
- R is antisymmetric: $\forall x, y \in X$, if xRy and yRx then $x = y$
- R is transitive: $\forall x, y, z \in X$ if xRy and yRz , then we have xRz.

The antisymmetric requirement ensures that we only run into a scenario where x is 'greater' than y and also y is 'greater' than x when $x = y$, which also seems reasonable. Requiring transitivity means that whenever we have x greater y, and y greater than z, we also have x greater z. With this definition in mind, let us turn to some examples.

3.2 Examples of Partial Orders

- Given the set of real numbers, there exists the familiar "less than or equal to" relation. This forms a partial order.
- Given the set of positive integers, there exists the "divides" relation. The "divides" relation is a partial order. Note that we can use the familiar "less than or equal to" relation as well.
- Starting with a set X, consider it's power set $\mathcal{P}(X)$. One can define a relation R on $\mathcal{P}(X)$ by stipulating ARB if and only if $A \subseteq B$. Then, R, the "is a subset of" relation, is a partial order.

For a moment, let us consider the 'divides' relation on the positive integers and verify that it is indeed a partial order:

- Every positive integer divides itself, so "divides" is reflexive.
- If a positive integer n divides another positive integer m, it is definitely not the case that m divides n (unless $n = m$), so it's anti-symmetric.
- If n divides m and m divides q, then n divides q, so it's transitive.

Great! The "divides" relation on the set of positive integers is indeed a partial order. Before we continue, we will define a piece of notation: Given a set X, and a partial order R on X, if $(x, y) \in R$, we will write $x \preceq y$. If $(x, y) \in R$, and $x \neq y$, we will write $x \prec y$.

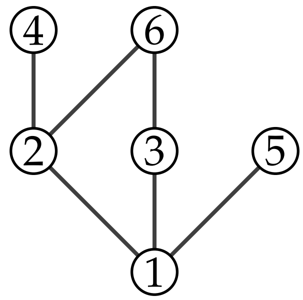
3.3 Hasse Diagrams

Hasse Diagrams are pictures that show how the elements in a partial order relate to each other. Let us see some examples of posets and their corresponding Hasse diagrams.

Consider the "divides" relation on the set $X = \{1, 2, 3, 4, 5, 6\}$. This relation consists of the following ordered pairs:

$$R = \{(1,1),(1,2),(1,3),(1,4),(1,5),(1,6),(2,2),\\ (2,4),(2,6),(3,3),(3,6),(4,4),(5,5),(6,6)\}.$$

Looking at this large set of ordered pairs does not give immediate insight as to how exactly the poset is structured. We can instead represent this poset as a Hasse diagram:

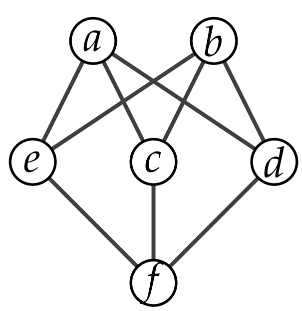


To read this, notice that 1 is directly below 2, 3, and 5, because 1 divides 2, 3 and 5. Furthermore, we see that 1 is below 3, and 3 is below 6, and partial orders must have the transitive property, so we may conclude 1 is below 6 as well. By similar reasoning, we can conclude 1 is below 4. If we wanted to, we could connect 1 to 6 or connect 1 to 4, but we don't need to—requiring transitivity makes such a line redundant.

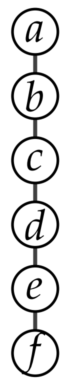
Furthermore, we could connect every element to itself to indicate that every number divides itself, but we don't need to. Also note that 5 does not divide any of the other elements, so there is no line attaching 5 to some element above it. Similarly, 2 does not divide 3 and 3 does not divide 2, so there is no line between them. In situations like this, we say that 2 and 3 are *incomparable.*

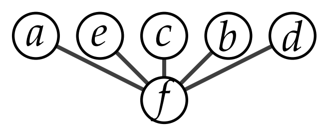
Finally, a common misunderstanding of Hasse diagrams involves looking for "paths" between elements. For instance, one can walk from 5, down to 1, up to 3, up to 6. Thus, one might be tempted to conclude that we have $5 \preceq 6$ in this diagram. However, this is not how Hasse digrams are read—if there is no path between two elements that involves going **strictly** up or down, then those two elements are incomparable.

For another example, consider the following Hasse diagram of a partial order on the elements set $\{a, b, c, d, e, f\}$:



Hopefully, these Hasse diagrams give an immediate sense of the structure of a poset. Furthermore, they give an immediate sense of whether the poset is particularly wide or tall, which gives a rough idea for what proportion of the elements are comparable. For example, consider the following two posets on the letters a-f :



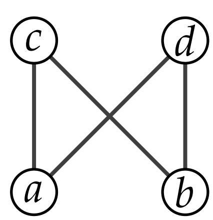


One can immediately see that the left poset is tall but not wide. Loosely speaking, we can see this and expect that if we were to randomly grab a pair of elements, one will likely be smaller the other. The right poset is wide but not tall, meaning if we randomly grab a pair of elements, we might expect them to be incomparable.

3.4 Weak Orders

A weak order on a set X is a partial order in which *transitivity of incomparability* holds. That is, for all $x, y, z \in X$, if x is incomparable to y, and y is incomparable to z, then x is incomparable to z.

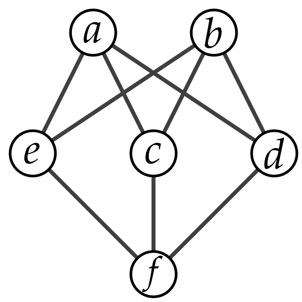
Informally, we can think of weak orders as posets in which the elements fall into "buckets", where elements within the bucket are all incomparable, and certain buckets are better than others. To illustrate what we mean, consider the following partial order:



.

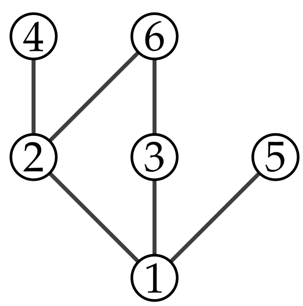
In this poset, transitivity of incomparability does indeed hold. However, we can more easily verify this is a weak order by finding "buckets." We can put c and d in the "good" bucket, and a and b in the "bad" bucket, and say everything in the good bucket is better than everything in the bad bucket.

For another example of a weak order, recall a poset from the previous section:



In this poset, we could reasonably say that a and b are in a "top tier", then e, c, and d are in a "middle tier" and f is by itself in the bottom tier.

It may be tempting to see these examples and come to the conclusion that *all* posets can be interpreted as some set of "buckets". However, this is *not* the case. For example, recall the poset we got from the "divides" relation on $\{1, 2, 3, 4, 5, 6\}$:



Firstly, notice that in this poset, transitivity of incomparability does not hold. 4 is incomparable to 3, and 3 is incomparable to 2, but 4 and 2 are comparable.

Furthermore, this poset is fundamentally different from the previous two examples, in that *there is no way* to put the elements into tiers or buckets that respects the original poset. One might be tempted to say that we can put 4 and 6 in the top tier, 2, 3, and 5 in the middle tier, and 1 by itself in the bottom tier. However, such a "bucket-ing" does not respect the original poset. In this "bucket-ing" scheme, having 6 in the top tier, and 5 in the middle tier would imply that $5 \leq 6$, but the original poset has 5 and 6 incomparable. Thus, such a "bucket-ing" scheme does not respect the original poset. More generally, because the original poset is not a weak order, there is no such bucket-ing scheme that respects the original order.

Chapter 4

Voting With Posets

At this point, we hope you're eager to see how partial orders and voting come together. That's what this chapter (and arguably, this entire thesis) is dedicated to!

Specifically, what happens when we allow our voters to submit partial orders as their ballots instead of total orders? By allowing this, we've just opened up an entire new world of voting procedures that begs lots of questions.

Our previous voting systems, the Borda count and the Kemeny rule, do not accept partially ordered ballots, so we'll discuss generalizations of them that do.

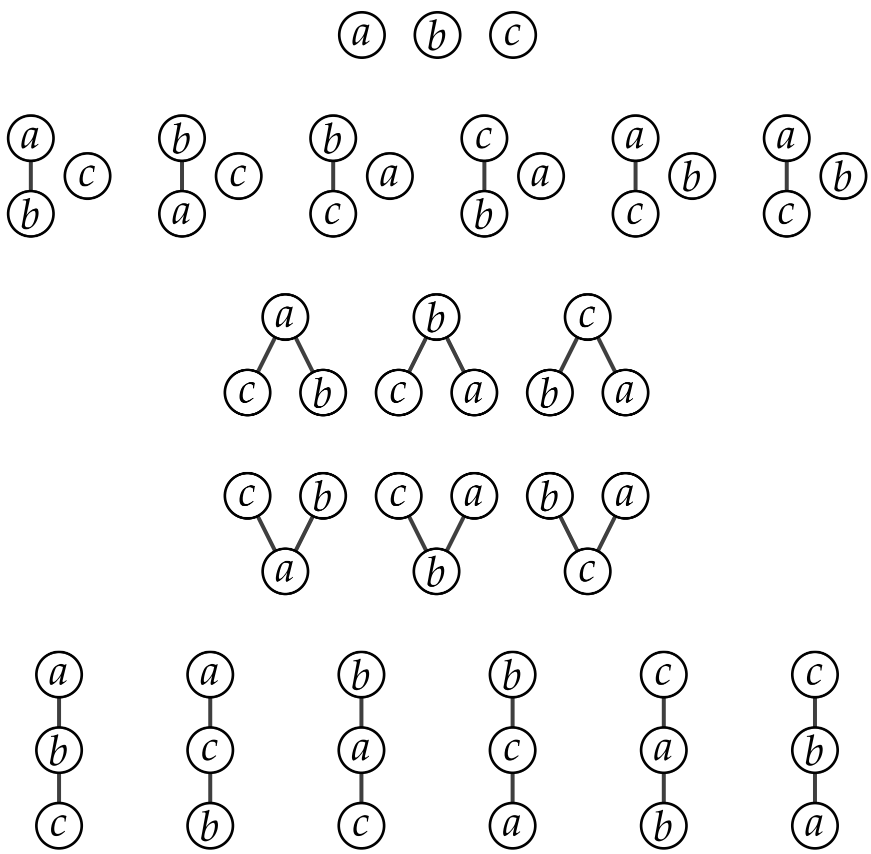
Even once we've defined these generalizations of the Borda count and Kemeny rule, we have all sorts of new questions! What kinds of new, unexpected behavior can we observe in the poset world that wasn't previously possible in the complete ranking world? How can we be sure that these generalizations are indeed the correct generalizations?

While these questions about the voting systems are exciting and interesting, this first section of this chapter doesn't begin by addressing these questions. This first section is dedicated to understanding voting *profiles.*

Why? Well, when we looked at voting systems for total rankings, we know exactly what everyone's ballot looked like—a 1st choice candidate, a 2nd choice candidate, a 3rd choice candidate, and so on, all the way down. But with posets, the *ballots themselves have structure.* This additional complexity is crucial for informing the way we think about partial voting systems.

4.1 Voter Profiles with Posets

To begin, consider an election on 3 candidates: a, b, and c. Voters will be allowed to submit partial orders on their ballots. There are 19 possible partial orders on that they might submit. It's not obvious why there are 19 of these—that's just how it works out. Enumerating all the possible partial orders on n elements is quite difficult, as discussed in Monteiro et al. (2017). Here are those 19 posets:

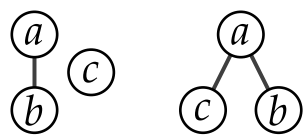


This may seem a little unwieldy. How do we begin making sense of the posets above? Well, as a start, notice that each row above contains an *isomorphism class* of posets—the actual structure of the partial order is the same, and we've simply relabeled the nodes in order to change between posets. So there are 19 possible posets, but only 5 distinct posets up to isomorphism.

However, this is not the only way to organize our 19 possible ballots. Instead, we can arrange them into a graph by defining a notion of *adjacency* between two partial orders.

Specifically, we will say that two posets are adjacent to each other if and only if their underlying relations differ by exactly one ordered pair.

For example, consider the following two posets:



.

The underlying relation for the poset on the left is

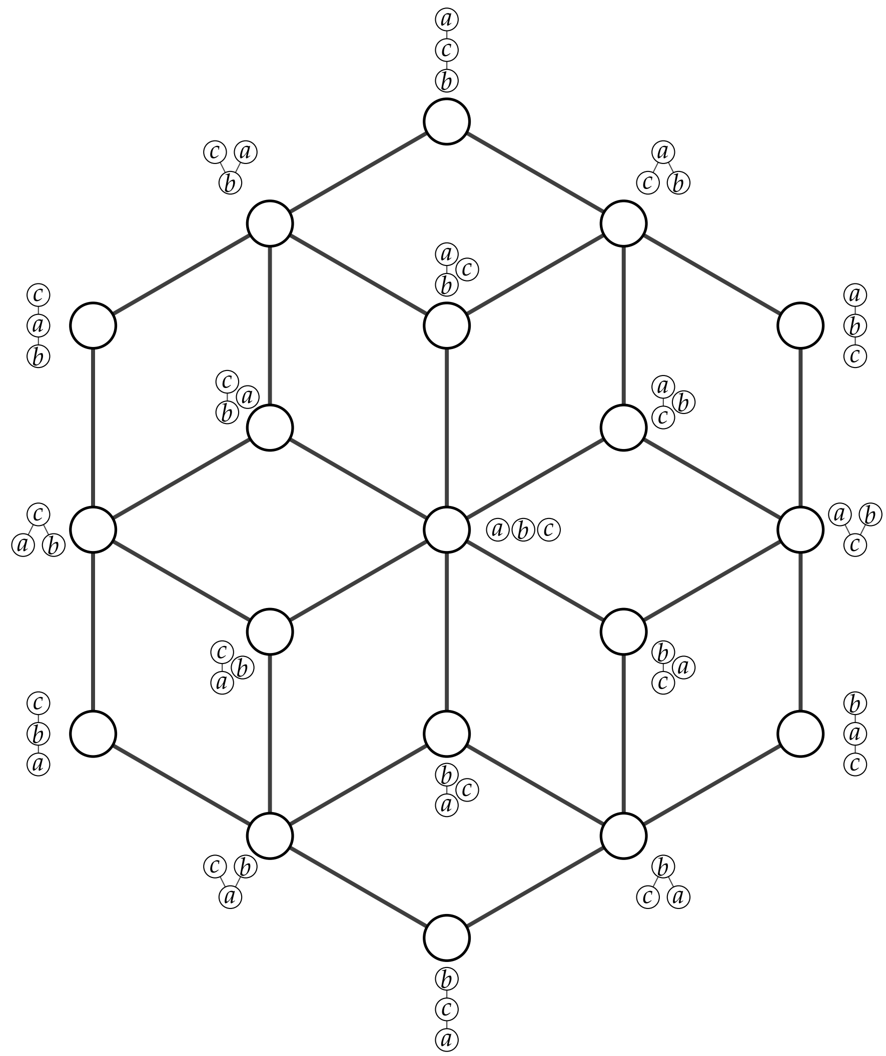
$$R = \{(a, a), (b, b), (c, c), (b, a)\}.$$

The underlying relation for the poset on the right is

$$R = \{(a, a), (b, b), (c, c), (b, a), (c, a)\}.$$

These two posets are similar orderings on the candidates. They both place a higher than b, they both think that b and c are incomparable, and their only disagreement is that the left poset considers a and c incomparable, while the right poset places a is higher than c. Thus, we can think of moving from the left poset to the right poset as *adding one piece of information,* namely adding that $c \prec a$. Similarly, we can think of moving from the right poset to the left poset as *removing one piece of information,* namely removing the information that $c \prec a$. So, we can think of these two posets as being *adjacent* to each other.

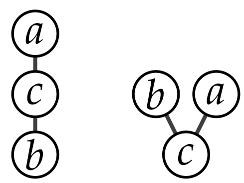
Equipped with a notion of adjacency, we can define a graph which is analogous to a permutohedron. Specifically, the vertices of this graph will be posets, and there will be an edge between two posets/vertices if and only if they are adjacent by the above definition. Doing so, we obtain:



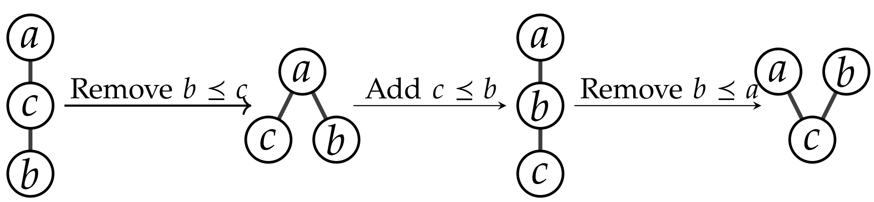
This graph, to the author's knowledge, first appeared in a 1973 paper called *Preference Structures I: Distances Between Transitive Preference Relations*, written by Kenneth Bogart (Bogart (1973)).

Now look at this! We've created a graph in which all of our posets appear as vertices. As a sanity check, recall the permutohedron on 3 elements from Chapter 2. Notice that the distances in the above graph respect the distances on the original permutohedron. For example, abc and cba remain on opposite sides of the graph, at a maximum distance apart. Similarly, the complete rankings which are closest to abc are acb and bac. This is a nice reassurance that this graph, which we will now refer to as the *Bogart graph*, is indeed a generalization of the complete-ranking situation that preserves the structure we're interested in.

Furthermore, this graph gives us a very natural notion of distance between posets! Namely, we take the distance between two posets to be the length of the shortest path between them. Note that there may be multiple paths with minimal length. Furthermore, we have a nice interpretation of this graph distance. As an example, consider the following two posets in the graph:



The distance between these two posets is 3, but we can interpret this 3 nicely without the graph. Namely, we ask the question: How many pieces of information do we have to remove or add from the underlying relations to get from one poset to another? To get from the left poset to the right poset, we can remove $b \leq c$, add $c \leq b$, and remove $a \leq c$. This is what that would look like:



Because it takes 3 steps to get from the far left poset to the far right poset, they are distance 3 away.

Now that we have established this graph, and this notion of distance between partial orders, we can discuss a generalization of the Kemeny rule to the partial order setting.

4.2 The Partial Kemeny Rule

In this section, we define a generalization of the Kemeny rule. Similarly to the full-ranking case, we can think of this partial Kemeny rule in three different ways, namely:

- A distance optimization problem on a graph.
- A system which assigns points to posets, and selects the poset(s) with the most points as the winner(s).
- A linear transformation.

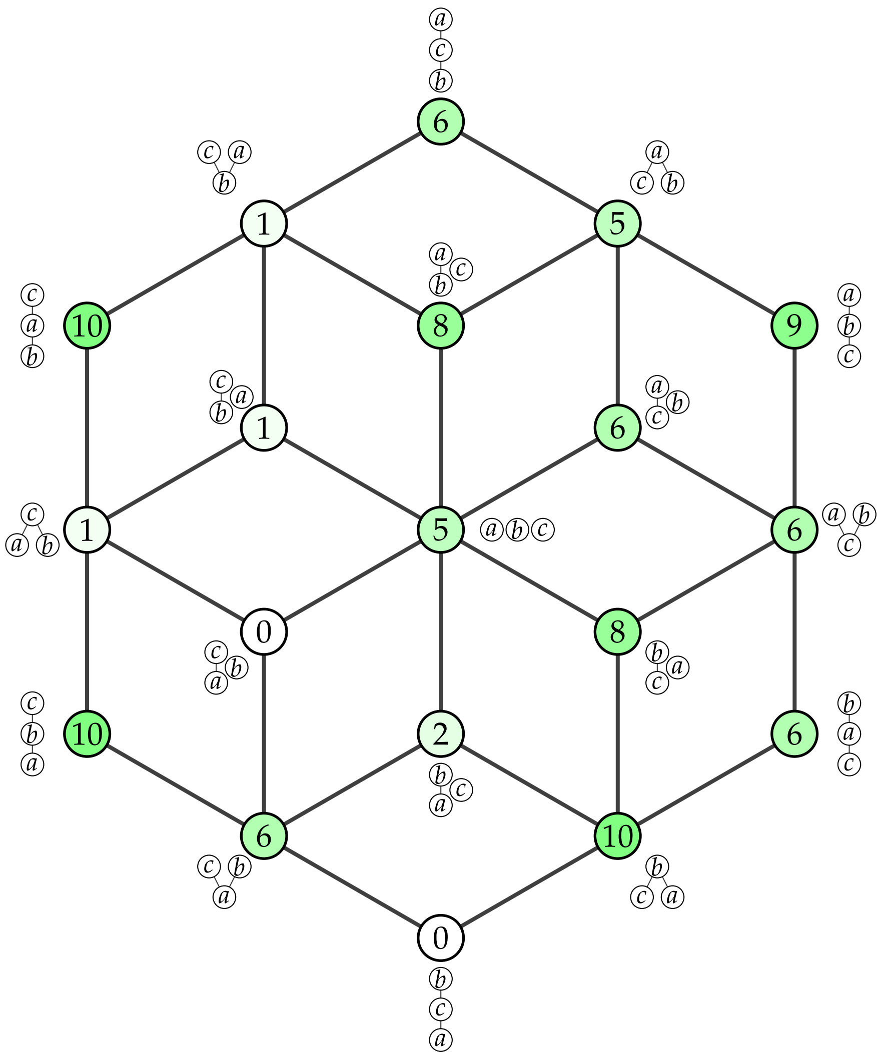
Throughout the rest of this document, we will find that being able to view a voting system from multiple perspectives often yields great insight. So we call attention to these three distinct perspectives.

4.2.1 The Kemeny Rule as a Distance Optimization Problem

In Chapter 2, we saw that the Kemeny rule could be thought of as a distance minimization problem on a permutohedron. Specifically, we made a graph in which each vertex was a possible ballot that a voter might submit. Then, we assigned a weight to each vertex equal to the number of people that voted for it. Then, for every vertex, we computed the sum of the weighted distances to all the other vertices. Finally, the Kemeny rule winner was the vertex with the minimum such weighted sum.

Now that we have an analagous graph for all the partial orders on 3 candidates, we can do the exact same thing here!

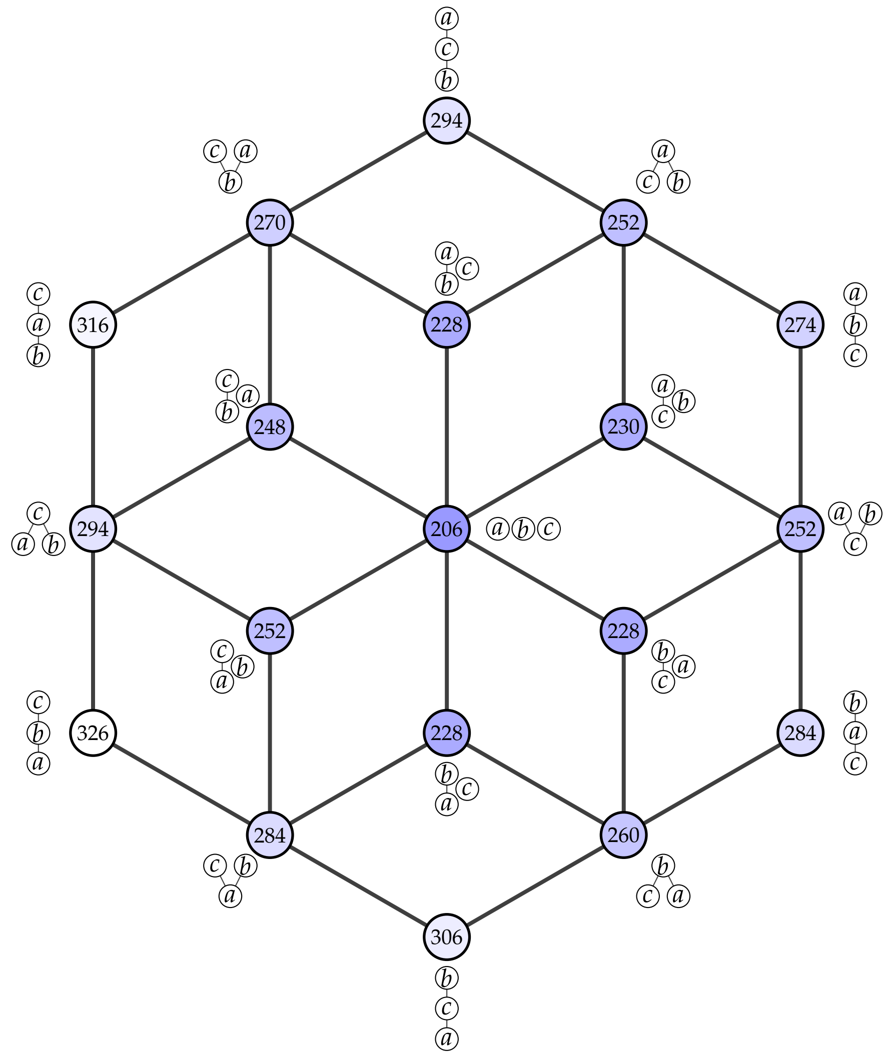
So, suppose that we have the following voter profile:



Note that we have colored the above vertices according to their weights. Higher weights are in darker greens. Then, given this arrangement of weights on the posets, we can compute, for each poset, a weighted sum. To do this, we first identify the vertex associated with that poset, which we will call v. Then, let $p(v)$ be the weight assigned to vertex v. Furthermore, given two vertices v, w let $d(v, w)$ be the distance between them. Then, $S(v)$ is given by:

$$S(v) = \sum_{w \in V(G)} d(v, w) p(w)$$

Note that this is the exact same equation that we used for looking at the Kemeny rule for complete rankings. Now we compute $S(v)$ for each vertex, and we get the following:



We have colored the above graph in blue instead of green to emphasize that this is the result of an election, and not a profile. Furthermore, note that the colors here are "inverted" from the last example—we have colored smaller numbers in darker blue, such that the winning poset is the darkest blue.

4.2.2 The Kemeny Rule as a Linear Transformation

Similarly to the complete ranking case, we can take a voter profile, encode it as a vector, then multiply that vector by a matrix to run our Kemeny rule election.

In the complete ranking case, there were 6 possible rankings on 3 candidates, so our profile was encoded as a vector $v \in \mathbb{R}^6$. In the poset case, however, there are 19 possible posets, so now v lives in $\mathbb{R}^{19}$ instead. Similarly, instead of a 6×6 matrix, we will make a 19×19 matrix.

With the Bogart graph from the previous section, we are equipped to make a distance matrix D. This matrix will have rows and columns indexed by posets. Given two posets, p and q, then the entry in the row corresponding to p and the column corresponding to q is simply the distance between p and q in the above graph.

Once D is constructed, computing Dv will result in another vector whose entries are the weighted sums of distances that we defined in our distance-minimization interpretation of the Kemeny rule. Finally, we simply find the minimum entry of Dv, and that will determine the winning poset.

4.2.3 The Kemeny Rule as a Point-Assigning System

Generalizing the point-assignment idea is less obvious than generalizing the distance minimization idea.

To see why, recall that in the full ranking case, a ranking r that features $a < b$ receives 1 point for every voter who thinks $a < b$, and 0 points for $b < a$. Thus, it seems natural that a poset p featuring $a \prec b$ receives 1 point for every voter who thinks $a \prec b$ and 0 points for every voter who thinks $b \prec a$. However, how many points should p receive when a voter thinks that a and b are incomparable? 0? 1? some number between them? Furthermore, if a poset has a incomparable to b, should it receive points when a voter thinks $a \prec b$ or $b \prec a$?

To answer these questions, recall that in the complete ranking case, the point-assignment system and the graph distance were very closely related. Specifically, if a ranking r_1 was a distance d from r_2, then r_1 received $3 - d$ points from r_2. Thus, we should create a point-assignment system that satisfies a similar relationship. With this in mind, we now give the point-assignment system.

Suppose we are given a poset p. To compute that poset's score, we consider every possible pair of candidates, a and b. In the voting body, there are a certain number of people who've submitted posets in which $a \prec b$, a certain number who've submitted posets in which $b \prec a$, and a certain number who've submitted posets in which a and b are incomparable. Then, our poset p is given points as follows:

- If $a \prec b$ in p, then p gets 1 point for every voter who thinks that $a \prec b$, and $\frac{1}{2}$ a point for every voter who thinks that a is incomparable to b.
- If a is incomparable to b in p, then p gets $\frac{1}{2}$ a point for every voter who thinks that $a \prec b$, $\frac{1}{2}$ a point for every voter who thinks that $b \prec a$, and 1 point for every voter who thinks that a is incomparable to b.
- If $b \prec a$ in p, then p gets 1 point for every voter who thinks that $b \prec a$, and $\frac{1}{2}$ a point for every voter who thinks that a is incomparable to b.

We do this for every pair of candidates in the partial order, and at the end, we have given a certain number of points to each of the 19 posets on 3 elements. Then the winning poset is the poset with the highest number of points.

Note that if every voter submits a poset which is a complete ranking, and we limit ourselves to looking at the points that complete rankings receive, this point assignment system simply reverts back to the point assignment system we had for the complete ranking version of the Kemeny rule.

4.3 The Partial Borda Count

Now that we've defined a generaliztion of the Kemeny rule that accepts partially ordered ballots, we do the same for the Borda count. Note that this definition comes from Cullinan et al. (2014).

4.3.1 Defining the Partial Borda Count

Recall that the Borda count for complete rankings was a *positional* voting system, a system in which candidates receive points according to their position on each voter's ballot. Thus, we'd like to assign points to candidates according to how highly they're ranked.

Let A be a set of candidates. Then, given a poset, we can define the *down set* of a candidate $a \in A$. The down set is defined as follows:

$$\text{down}(a) = \{b \in A : b \prec a\}.$$

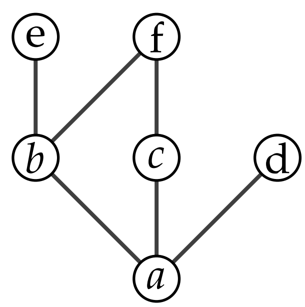
Similarly, we can define the *incomparable set* as

$$\text{incomparable}(a) = \{b \in A : b \text{ is incomparable to } a\}.$$

Then, the number of points candidate a_i receives is given by

$$\text{Points Received } = 2 \cdot \|\text{down}(a_i)\| + \|\text{incomp}(a_i)\|.$$

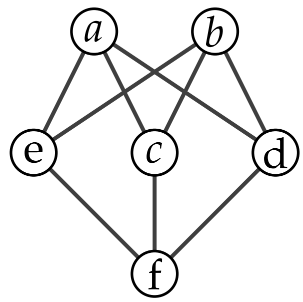
In short, a candidate gets 2 points for every candidate it beats, and 1 point for every candidate it's incomparable to. Then we tally all the points, and the candidate with the most points wins. Let's see some examples:



If someone were to submit this poset, then:

- Candidate a receives 0 points.
- Candidate b receives 4 points: 2 for being above a, and 2 for being incomparable to c and d.
- Candidate c receives 5 points: 2 for being above a, and 3 for being incomparable to b, d, and e.
- Candidate d receives 6 points: 2 for being above a, and 4 for being incomparable to b, c, e, and f.
- Candidate e receives 7 points: 4 for being above a and b, then 3 points for being incomparable to c, d, and f.
- Candidate f receives 8 points: 6 for being above a, b, and c, then 2 points for being incomparable to e and d.

Suppose another voter submitted the following poset on these same candidates:



- Candidate a receives 9 points: 8 for being above e, c, d, and f, and 1 for being incomparable to b.
- Candidate b receives 9 points: 8 for being above e, c, d, and f, and 1 for being incomparable to a.
- Candidate c receives 4 points: 2 for being above f, and 2 for being incomparable to d and e.
- Candidate d receives 4 points: 2 for being above f, and 2 for being incomparable to c and e.
- Candidate e receives 4 points: 2 for being above f, and 2 for being incomparable to c and d.
- Candidate f receives 0 points, because it's not above or incomparable to any other candidates.

Adding up the totals, we get:

- Candidate a finishes with $0 + 9 = 9$ points.
- Candidate b finishes with $4 + 9 = 13$ points.
- Candidate c finishes with $5 + 4 = 9$ points.
- Candidate d finishes with $6 + 4 = 10$ points.
- Candidate e finishes with $7 + 4 = 11$ points.
- Candidate f finishes with $8 + 0 = 8$ points.

Now that we've assigned points to candidates, we simply rank our candidates according to how many points they've received, and that's our election!

4.3.2 The Partial Borda Count as a Linear Transformation

Similarly to the complete-ranking case, and similarly to the Kemeny rule, we can run a Borda count election via a matrix vector multiplication.

Once again, we construct a vector v with 19 entries, tallying the number of people who voted for each poset. When we are running an election with three candidates, there are 19 possible posets that a voter might choose.

Then, to run a Borda election, we can define a matrix B, then compute Bv, and we will end up with a 3x1 vector that tells us the number of points a, b, and c each received. Specifically, B is a matrix with 19 columns and 3 rows. We can think of the columns as being indexed by posets, and we can think of the rows as being indexed by candidates. Each entry in the matrix corresponds to one poset and one candidate. In particular, that entry is given by the number of points that candidate receives in that poset. To demonstrate, we can fill in some of that matrix:

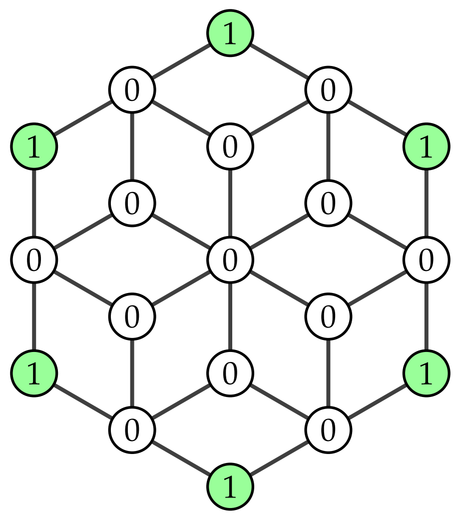
$$B = \begin{array}{c} \\ a \\ b \\ c \end{array} \begin{array}{c} \begin{array}{ccccccc} p_1 & p_2 & p_3 & p_4 & p_5 & \dots & p_{19} \end{array} \\ \begin{pmatrix} 2 & 3 & 3 & 2 & 1 & \dots & 3 \\ 2 & 1 & 2 & 3 & 3 & \dots & 0 \\ 2 & 2 & 1 & 1 & 2 & \dots & 3 \end{pmatrix} \end{array}$$

Note that in the above matrix, we have assigned indices to the posets. The index assigned to a particular poset doesn't matter, as long as one is consistent with that index throughout.

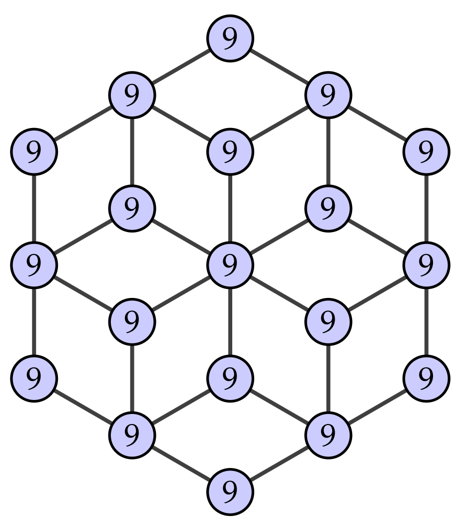
With the matrix B defined in this way, we can compute Bv. This will give us a 3x1 vector where the first entry is the number of points that a receives, the second entry is the number of points that b receives, and the third entry is the number of points that c receives.

4.4 The Null Space of the Partial Kemeny Rule and Graph Symmetry

One approach to understanding the partial Kemeny rule is to understand voter profiles that can be addded to other voter profiles that don't change election outcomes. As an example of such a profile, consider the following:



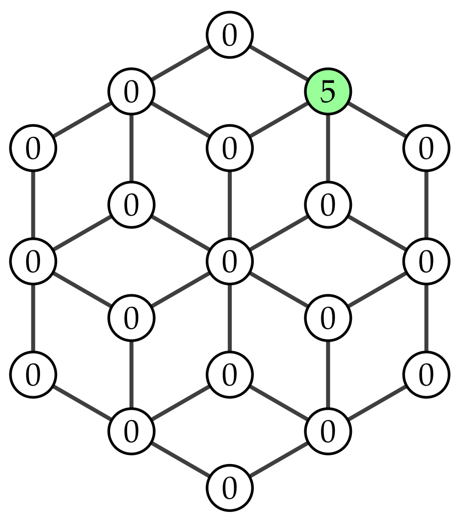
The resulting point distribution is:



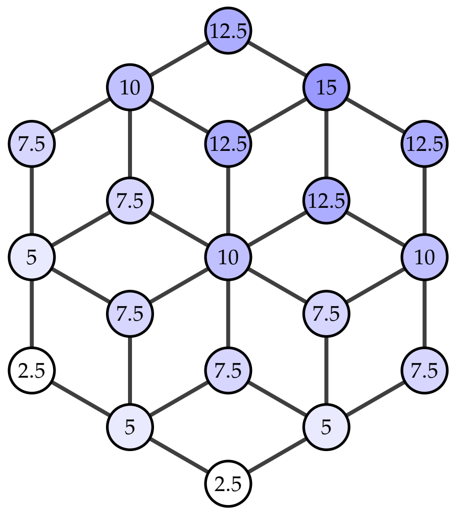
We can attempt to intuitively rationalize this. First, note that if we were to limit our attention to the complete ranking case, exactly one vote for every complete ranking would result in a 6-way tie between all the complete rankings. This result respects that, which is a nice sanity check, and makes this result less surprising.

While this profile is interesting by itself, it's also interesting to combine it with other profiles. Suppose that we name the above profile v. Furthermore, suppose we have some other arbitrary profile w. Then, because the Kemeny rule can be thought of as a linear transformation, we have $K(v+w) = Kv+Kw$. Practically speaking, this means that when we add v to a profile then run the Kemeny rule election, the effect is the same running the Kemeny rule election, then adding adding Kv. But Kv consists of a constant function on the vertices, so Kv has no effect.

As an example, consider a different profile:

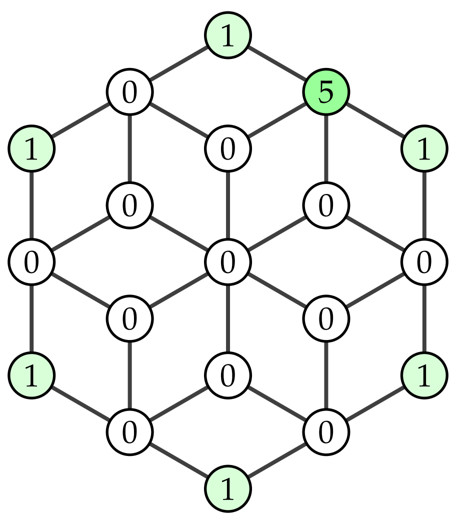


With this profile, we can compute the number of points each poset would receive in a Kemeny rule election. Note that the poset with the highest number of points will be the winner. The resulting point distribution is:

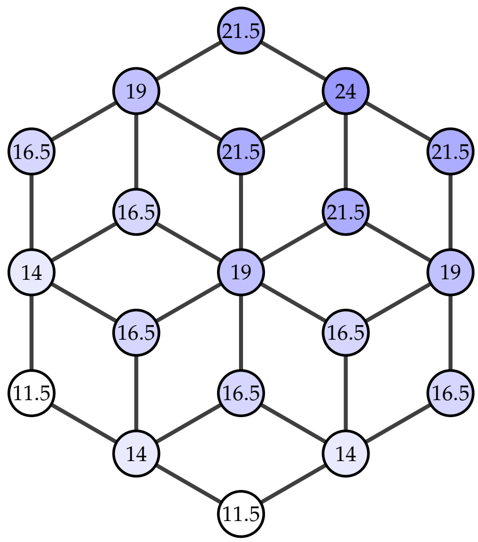


This is unsurprising—when all our voters vote for the same poset, that poset wins the resulting election, with other posets scoring fewer points according to how far away they are.

If we combine the two profiles we've looked at in this section, we get:



Then, running the Kemeny Rule, we have:

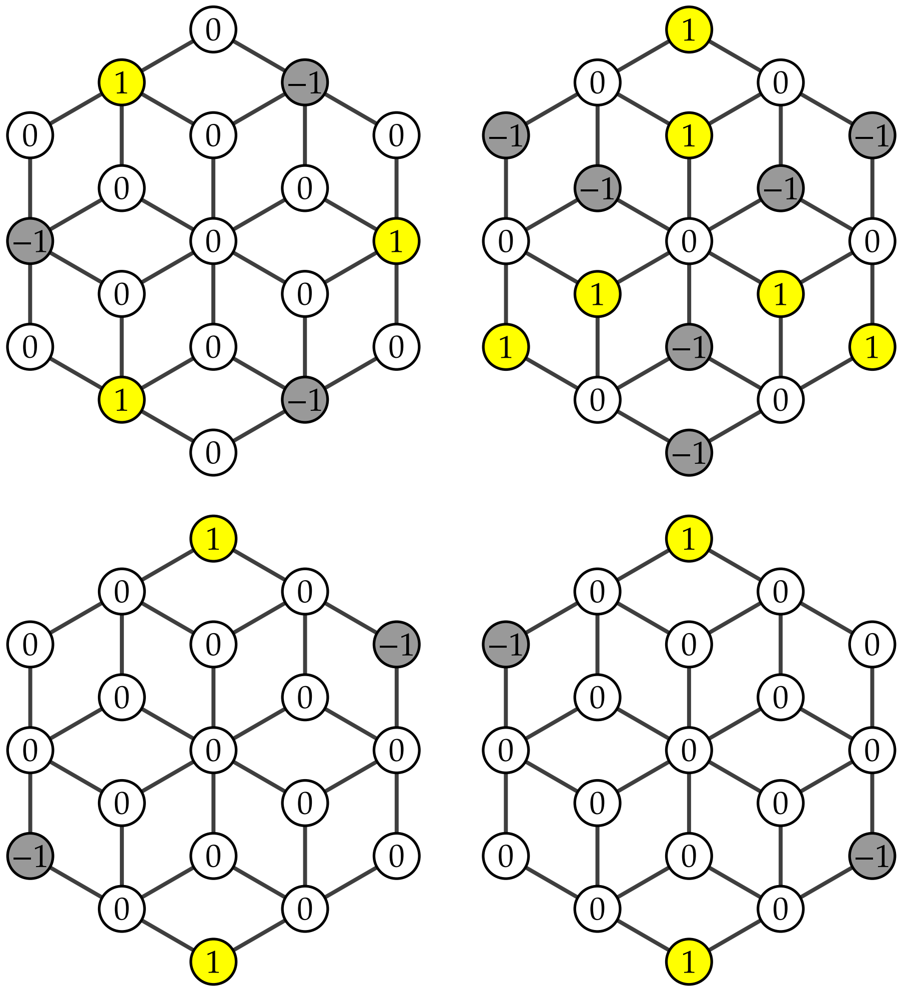


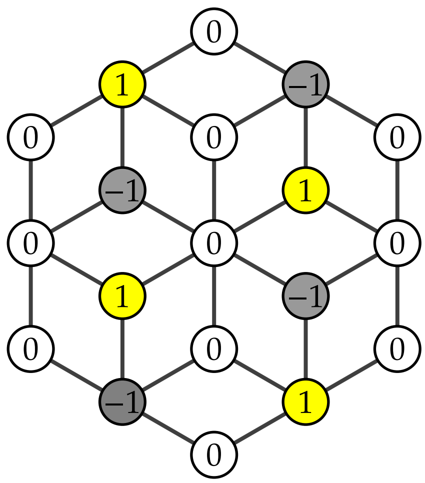
Wow! Notice that this is almost the same as our previous results, except we've increased the score of every poset by 9.

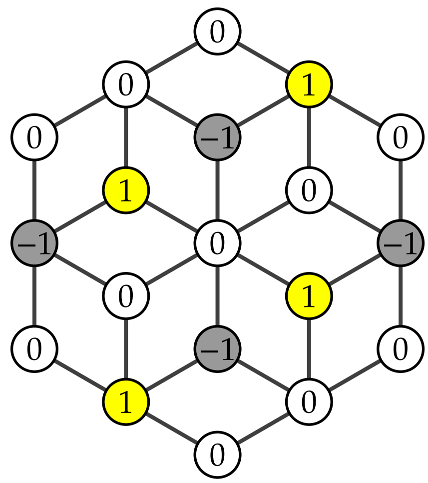
There are other ways to have a profile which doesn't influence the outcome of the election. Namely, we can look at profiles in which some posets receive a *negative* number of votes. This doesn't exactly have a clean real-world interpretation that the author can think of, but this type of profile is useful for linear algebraic considerations, and is difficult to avoid. Such profiles allow for the possibility that a vector is contained in the null space of the linear transformation.

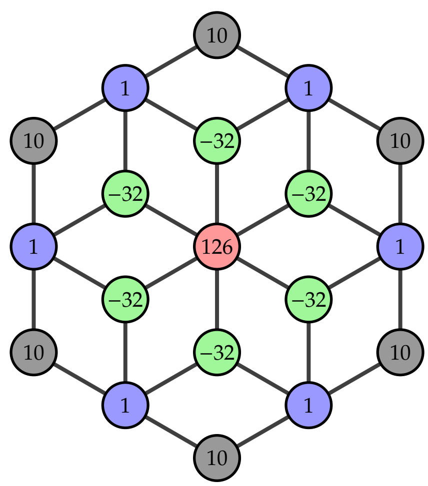
Recall that we can think of the Kemeny Rule as a linear transformation $K : \mathbb{R}^{19} \to \mathbb{R}^{19}$. We can encode K as a matrix and use standard linear algebra software to find the dimension of the kernel of the matrix. Doing so, we get 7. However, if we use standard linear algebra software to find a basis for this nullspace, we get a sequence of vectors which does form a valid basis, but is nigh impossible to interpret or understand.

However, over the course of this thesis project, we have found a different basis that is easier to understand. Specifically, the vectors that make up this basis can be encoded as voter profiles, and those profiles can then be encoded as vertex-weighted graphs. The next pages feature pictures of those graphs.


0
1 −1
0 0 0
0 0
−1 0 1
0 0
0 0 0
1 −1
0
1
0 0
−1 1 −1
−1 −1
0 0 0
1 1
1 −1 1
0 0
−1
1
0 0
0 0 −1
0 0
0 0 0
0 0
−1 0 0
0 0
1
1
0 0
−1 0 0
0 0
0 0 0
0 0
0 0 −1
0 0
1







The 7 graphs above were found by hand by hunting for symmetry in the Bogart graph.

4.5 Numerical Experiments

So far, we haven't discussed much about what voter profiles might look like. Now, we (briefly) turn our attention to this issue.

It would be nice if we could simply look at all possible voter profiles for some number of candidates, and see what every voting system does. However, this is simply not feasible. Even with the 3 candidate case, and only 10 voters, there are $19^{10} \approx 6.13 \times 10^{12}$ possible profiles.

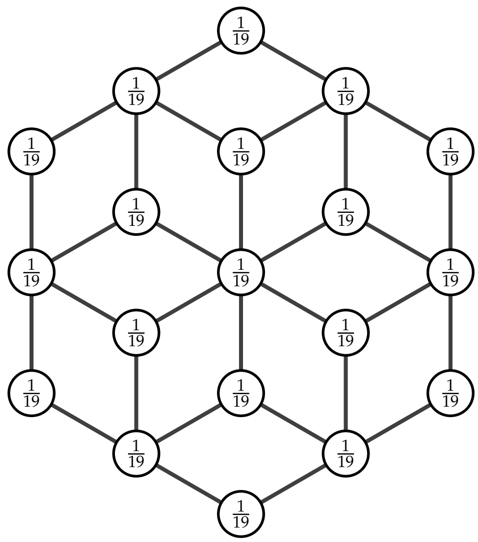
Instead, we often use some probabilistic model to simulate the behavior of a voting body, and run many such simulations.

But this begs the question—how do we simulate the behavior of a voting body? Well, we usually begin by choosing some model of voter behavior, then using that model to create a probability distribution on the posets.

While these simulations will prove to be insightful and interesting, they are largely exploratory, and we do not have specific, proven results to share. An interesting area of future work would be to prove probabilistic results that can explain the data we are about to see.

4.5.1 Impartial Culture Model

Perhaps the simplest model for voter behavior is the Impartial Culture (IC) model, in which voters are equally likely to choose any of the options in front of them. While this model has been criticized for being unrealistic, it is simple (Lehtinen and Kuorikoski (2007)). In our case, this would mean that among the 19 possible posets that a voter might submit, a voter has a $\frac{1}{19}$ probability of choosing any of them. This isn't particularly hard to visualize, but we will look at more complex profiles shortly, so visualizing these models on the Bogart graph will be helpful. We have:



Now, with this model established, we can ask some questions. For example: If we construct a voter profile at random using the above IC model, what is the probability that the Kemeny rule returns the center poset, the 3 element antichain, as the winner? How does this probability change as the number of voters increases?

Immediately, we know that if there's only one voter, the only way for the 3 element antichain to win is for that voter to select it. Thus, $p = \frac{1}{19}$ when there's 1 voter.

However, for larger numbers of voters, the author does not know of a simple expression for p. Instead, we can run some simulations to get an idea of the answer. So, we run 10,000 simulations with a fixed number of voters, and calculate the proportion in which the 3 element antichain wins.

Number of Voters	Percentage in which antichain won
1	5.26%
5	27.97%
15	67.13%
45	97.14%
100	99.96%

Looking at this table, we have very strong evidence that as the number of voters approaches infinity, the probability that the antichain wins the election approaches 1. This also matches an intuitive understanding for how the partial Kemeny rule works. If voters choose vertices on the Bogart graph at random, then it makes sense that the vertex that is, on average, closest to all the others, is the vertex in the center.

How does this behavior compare to that of the Borda count? Does the Borda count also often select the antichain as the winner? Let us find out:

Number of Voters	Percentage in which antichain won
1	5.26%
5	2.07%
15	0.62%
45	.25%
100	.13%

Fascinating! The Borda count also starts at $\frac{1}{19}$, for the exact same reasoning as the Kemeny rule: there's a $\frac{1}{19}$ chance of choosing the antichain, which is the only way for the antichain to be the winning poset. However, we see that the percentage decreases instead of increases. We offer an explanation, without proof, for why this is. The Borda count assigns points to candidates, rather than searching for a poset that is closest to all the others. While it's likely that the candidates earn a similar number of points in these models, it's very unlikely that all three candidates earn the exact same number of points. Thus, the percentage of simulations in which the antichain wins is very small.

In summary, we have found that, for IC profiles, the Kemeny rule is likely to choose the antichain, whereas the Borda count is not. It is unclear what we should make of this observation. On the one hand, we could see this and argue that the Kemeny rule is extremely indecisive—it almost always throws its hands up and chooses the antichain. On the other hand, we could argue that the Borda count is too-willing to "force the issue" and just choose *something* even when the societal preference is clearly ambivalent or unclear.

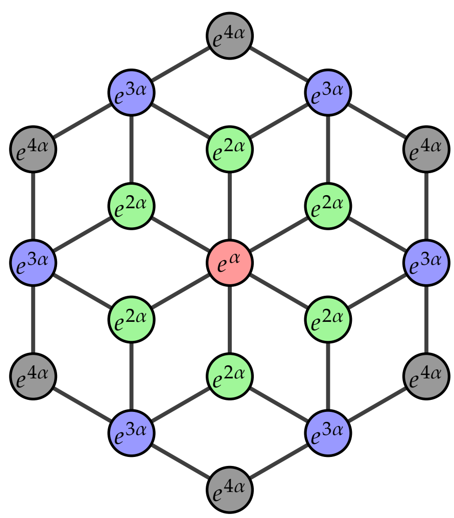
An interesting area of future work would be to explicitly find the probability of the antichain winning as a function of the number of voters, in either election system.

4.5.2 Objective Truth Model

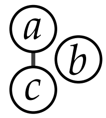
Another model for voter behavior is an *Objective Truth Model*. Such a model was first introduced by Condorcet in 1785—specifically, he modeled a group of jurors as independent agents trying to make the correct decision. For our purposes, we will specify some poset as being objective *best* partial ordering of the candidates. Then,voters learn about the candidates to try and make an informed decision about how to rank them. However, the real world has noise, misinformation, and complexity, so not every voter makes the "correct" choice.

Now, one might reasonably object to the validity of this model by arguing that real-world elections often don't have an objective truth, a "best" ranking of the candidates. This is a very reasonable point. However, it is still an interesting way to create structured voter profiles. Furthermore, recall that the Kemeny rule can be used in contexts outside of voting and elections—for example, consensus rankings of web pages, or artificial intelligence systems in which there is an objectively correct answer, but different software agents disagree.

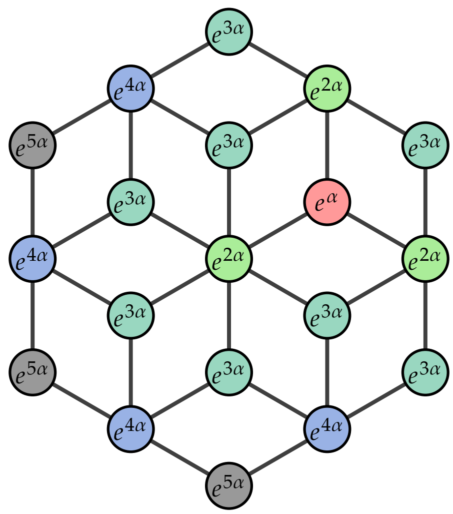
Translating this idea into a probabilistic model, we'd like some number of voters to choose the objectively correct choice, and then fewer voters to choose other posets as those posets get farther from the objective truth. So, we pick some $\alpha < 0$ and we assign a weight of e^{α} to the correct choice, a weight of $e^{2\alpha}$ to posets that are one step away, a weight of $e^{3\alpha}$ to posets that are two steps away, etc. Then, once we've assigned all the weights, we normalize them such that they sum to 1, and we have our weights. Visualizing this kid of model on our Bogart graph, we have:



Note that this is not the only "objective truth model" that we can define on the Bogart graph—in the above model, we have taken the central poset as the objective truth, the 3-element antichain. However, we can pick some other *non-central* poset as the "objective truth." For example, pick one of the posets that's not a weak order as the objective truth, such as:



Making this poset the "objective truth", we get:



Notice that α here serves as a metric for how "sharp" our distribution is. For more negative values of α, the objective truth poset has a higher probability of being chosen, and every other poset has a lower probability. As values of α get closer and closer to zero, we get closer and closer to an IC model.

With this kind of model, we can ask: How far away are the Kemeny rule and Borda count outcome from each other? Once again, we provide some experimental evidence that addresses this question, but do not provide rigorous proof.

In the following tables, we run 10,000 simulations. In each simulation, we find the winning poset according to the Kemeny rule, the winnning poset according to the Borda count, and find the distance between them.

4.5.3 Objective Truth in the Center

Number of Voters	α	Avg. Distance Between Borda and Kemeny
1	-1/2	0.88
5	-1/2	2.16
15	-1/2	2.74
45	-1/2	2.91
100	-1/2	2.94

To construct this table, we have run simulations in which the center poset is the objective truth, and has the highest weight. This evidence seems to suggest that, given a fixed value of α, increasing the number of voters causes the average distance between the Borda count and Kemeny rule to approach 3. We suspect the following explanation for this: as the number of voters increases, the Kemeny rule is more and more likely to select the 3 element antichain as the winning poset. However, the Borda count almost always selects a complete ranking. Once again, this is because the Borda count adds up points for each candidate, and is is unlikely that two candidates score the *exact same* number of points. Thus, ties and incomparable elements are very rare.

Instead of varying the number of voters while holding α constant, we will vary the α while holding the number of voters constant.

Number of Voters	α	Avg. Distance Between Borda and Kemeny
5	0	1.70
5	-1/2	2.15
5	-1	2.41
5	-2	2.44
5	-4	1.12
5	-8	0.03

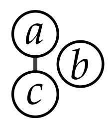
We observe some very interesting behavior here! It seems that the average distance increases, hits some maximum, and then decreases. Why?

We hypothesize the following explanation: $\alpha = 0$ corresponds to the IC model. So, as α becomes more negative, we move towards an Objective Truth model that behaves as the previous examples did. As voters move closer to the central poset, the Kemeny rule becomes increasingly likely to select the central poset. Meanwhile, it remains unlikely for the Borda count to give the same number of points to two different candidates, so the

Borda count is choosing posets towards the outside of the Bogart graph. However, as α increases more, our objective truth model features a sharper and sharper peak at the central poset. By the time we reach $\alpha = 8$, there is a 99.8% chance that a voter chooses the central poset. So, our model is very likely to give a profile where all 5 voters choose that same central poset. In these profiles, the Borda count and the Kemeny rule agree. Thus, we see the average distance between them decrease again.

4.5.4 Objective Truth on a Weak Order

For this set of simulations, we used models in which the poset



was used for the objective truth. With this in mind, we run the same numerical experiments that we did in the last section.

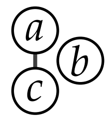
Number of Voters	α	Avg. Distance Between Borda and Kemeny
1	-1/2	0.81
5	-1/2	1.89
15	-1/2	2.40
45	-1/2	2.63
100	-1/2	2.73

This is a very similar result to when the antichain was the objective truth. Now, we fix the number of voters and let α vary.

Number of Voters	α	Avg. Distance Between Borda and Kemeny
5	0	1.71
5	1/2	1.89
5	1	1.91
5	2	1.95
5	4	1.99

This is interesting! When we made this same table with the antichain as our objective truth poset, we got a totally different result. Why are things different here?

Well, in this example, as we increase α, we are creating an objective truth model with a steeper and sharper peak. Thus, for high values of α, many of our voter profiles are 5 voters who all vote for the poset



.

This is where the key difference comes in. Imagine a voter profile in which exactly one voter submits this poset as their ballot. Then, a receives 3 points, b receives 2 points, and c receives 1 point. Thus, the Borda count would say that the poset $a > b > c$ is the best. But the poset $a > b > c$ is a distance two away from the above poset, which explains why the distances are approaching 2 in the table above.

We have just observed an interesting property of the Borda count. Because the Borda count assigns points to candidates, it is forced to return a weak order. This is even true when it's given a partial order that's not a weak order.

At this point, we could continue on and look at objective truth models in which we feature other posets as the objective truth. However, there is not much more interesting behavior to glean out of these same experiments applied to those models. The results are largely the same as those we saw when the central poset was the objective truth.

Chapter 5

Conclusion and Future Work

In this thesis, we've discussed the Borda count and Kemeny rule in situations where voters submit complete orders *or* partial orders. Then we briefly looked at symmetry in the Bogart graph, and used it to find a more "understandable" basis for the null space of the Kemeny rule. Finally, we ran some simulations to get a sense of what the partial Borda count and partial Kemeny rule do with certain kinds of profiles.

At this point, we are equipped to understand many directions for possible future research. We discuss some here, in no particular order.

5.1 The Partial Borda Count as an SRSF

In an n candidate election, the Borda count looks at how many people voted for each of the $n!$ rankings, then uses that information to distribute points to the n candidates. However, there are voting systems which instead distribute points *back to those* $n!$ *rankings.* We call such voting systems, ones that assign points to rankings rather than individual candidates, Simple Ranking Scoring Functions, or SRSFs for short. The Kemeny rule is an SRSF.

The Borda count by itself is *not* an SRSF, but there is a relatively straightforward way to make it an SRSF. In fact, a 2009 paper by Conitzer et al. (Conitzer et al. (2009)) demonstrates a method for taking *any* positional voting system and making it into an SRSF. We discuss this method here. It will require some notation.

Let a be a candidate, and let v be a vote, which is a specific permutation of the candidates. Then, let t be a function such that $t(v, a)$ is the number of points that candidate a gets for vote v. Now, we'd like to define a function

$s(v, r)$, which will be the number of points that a ranking r receives for a vote v. Given r, let $r(1)$ be the 1st ranked candidate in the permutation r, let $r(2)$ be the 2nd ranked candidate in the permutation r, etc. We can write $r = (r(1), r(2), \ldots, r(n))$. Now, we are ready to define $s(v, r)$:

$$s(v, r) = \sum_{i=1}^{n}(n - i)t(v, r(i)).$$

At this point, notice that $s(v, r)$ tells us how many points a ranking r receives from a single vote v. To compute how many points a ranking v receives in the election, we simply sum $s(v, r)$ over all votes. So, letting V be the set of all votes, the score that a particular ranking r receives is given by

$$\sum_{v \in V} s(v, r) = \sum_{v \in V} \sum_{i=1}^{n}(n - i)t(v, r(i)).$$

This is non-trivial to parse, so we will provide an example computation. Suppose we are looking at an election on 3 candidates, and we'd like to know the number of points that the ranking abc receives when 1 person votes for bac, where $t(v, a)$ is defined using the Borda count. Computing this, we have

$$\begin{aligned} s(bac, abc) &= \sum_{i=1}^{3}(n - i)t(v, r(i)) \\ &= 2t(bac, a) + 1t(bac, b) + 0t(bac, c) \\ &= 2(1) + 1(2) + 0(0) \\ &= 4. \end{aligned}$$

For another example, let us compute the number of points that ranking bca receives from vote acb:

$$\begin{aligned} s(acb, bca) &= \sum_{i=1}^{3}(n - i)t(v, r(i)) \\ &= 2t(acb, b) + 1t(acb, c) + 0t(acb, a) \\ &= 2(0) + 1(1) + 0(2) \\ &= 1. \end{aligned}$$

Note that in the second to last line in both this computation and the one above, the numbers in the parenthesis are permutations of $\{0,1,2\}$. The way to maximize this, and thus to maximize the score, is to have $2(2)+1(1)+0(0)$. Thus, a ranking will receive the most points from a vote with that ranking.

Hopefully, the reader feels comfortable computing the number of points a ranking receives from some other ranking. However, we are not limited to making this computation with a specific formula. Instead, we can model this computation as a matrix-vector multiplication as well!

Specifically, given a profile vector v, encoded in the same way we saw before, we compute $(B^T B)v$ instead of Bv. Let us examine this:

$$B^T B = \begin{pmatrix} 2 & 1 & 0 \\ 2 & 0 & 1 \\ 1 & 2 & 0 \\ 0 & 2 & 1 \\ 1 & 0 & 2 \\ 0 & 1 & 2 \end{pmatrix} \begin{pmatrix} 2 & 2 & 1 & 0 & 1 & 0 \\ 1 & 0 & 2 & 2 & 0 & 1 \\ 0 & 1 & 0 & 1 & 2 & 2 \end{pmatrix} = \begin{pmatrix} 5 & 4 & 4 & 2 & 2 & 1 \\ 4 & 5 & 2 & 1 & 4 & 2 \\ 4 & 2 & 5 & 4 & 1 & 2 \\ 2 & 1 & 4 & 5 & 2 & 4 \\ 2 & 4 & 1 & 2 & 5 & 4 \\ 1 & 2 & 2 & 4 & 4 & 5 \end{pmatrix}.$$

The resulting matrix is 6×6. Note that, by keeping track of how these columns and rows are being multiplied, both the rows *and* the columns are indexed by permutations on the 3 candidates now, rather than having rows indexed by candidates and columns indexed by permutations. Then, recall our profile vector $\vec{v}$ from earlier, $\vec{v} = (5,2,1,3,1,5)$. Computing $B^T B\vec{v}$, we have

$$B^T B\vec{v} = \begin{pmatrix} 50 \\ 49 \\ 52 \\ 53 \\ 50 \\ 52 \end{pmatrix}.$$

Recall that the entries in this vector are indexed by permutations of candidates in a specific order. Thus, we can read this as saying abc receives 50 points, acb receives 49, bac receives 52, etc. Notice that bca receives 53 points, the most of any permutation. This is good, because this aligns with the previous times we've run this election with the Borda count: we saw that candidate b won, c was in 2nd, and a was in 3rd. Thus, we expect bca to receive the most points, and that expectation is met. Note that this entire $B^T B$ construction is simply a nice way to encode the previous page's technique for making an SRSF out of a positional scoring function. Non-Borda positional voting systems are not the focus of this thesis, but in general, given some positional scoring matrix A, $B^T A$ will be the matrix that computes scores for rankings, rather than candidates.

By making the Borda count into an SRSF, we now have a matrix $B^T B$ which is symmetric and real-valued. Thus, we know that the profile space has an orthogonal basis of eigenvectors of $B^T B$. We will use this later to compare the Borda count to the Kemeny rule. Now, we turn our attention to the Kemeny rule.

With all of this established, we ask: Is it possible to do the same for the partial Borda count? How is this done? Does this technique generalize to other poset voting systems that assign points to candidates?

We conjecture that if the partial Borda count is encoded as a matrix B, then $B^T B$ is a matrix that encodes a partial Borda SPSF—Simple Poset Scoring Function.

5.2 Characterizing the Partial Kemeny Rule

Recall from the introduction that, when voters are restricted to submitting complete rankings, the Kemeny rule is the unique voting system which is consistent, Condorcet, and neutral.

Does this characterization remain true when we look at the partial version of the Kemeny rule that we have focused on in this thesis? That is, when voters are allowed to submit partial orders, is it the case that the partial Kemeny rule is the unique voting system which is consistent, Condorcet, and neutral?

It is not exactly obvious whether this is true. One could argue that the Kemeny rule is initially defined as a distance minimization procedure on a permutohedron. Then, we simply take that permutohedron, add more vertices in the middle of the graph, and run the same procedure. Thus, it seems reasonable that the Kemeny rule is consistent, Condorcet, and neutral. However, whether the Kemeny rule remains as the *unique* social choice function with this property is less obvious. It seems plausible that, by introducing the additional complexity of partially ordered ballots, there suddenly exist other voting systems which are consistent, Condorcet, and neutral.

Recall that in addition to our distance minimization understanding of the

Kemeny rule, we also conceived of the Kemeny rule as a voting procedure that assigns points to posets based on how much they agree with the posets in the voter profile. Specifically, we wrote:

Suppose we are given a poset p. To compute that poset's score, we consider every possible pair of candidates, a and b. In the voting body, there are a certain number of people who've submitted posets in which $a \prec b$, a certain number who've submitted posets in which $b \prec a$, and a certain number who've submitted posets in which a and b are incomparable. Then, our poset p is given points as follows:

- If $a \prec b$ in p, then p gets 1 point for every voter who thinks that $a \prec b$, and $\frac{1}{2}$ a point for every voter who thinks that a is incomparable to b.
- If a is incomparable to b in p, then p gets $\frac{1}{2}$ a point for every voter who thinks that $a \prec b$, $\frac{1}{2}$ a point for every voter who thinks that $b \prec a$, and 1 point for every voter who thinks that a is incomparable to b.
- If $b \prec a$ in p, then p gets 1 point for every voter who thinks that $b \prec a$, and $\frac{1}{2}$ a point for every voter who thinks that a is incomparable to b.

Now, we conjecture that by changing the point assignments, we can achieve an entire spectrum of voting procedures which are consistent, Condorcet, and neutral. Instead of assigning 1, $\frac{1}{2}$, or 0 points to p according to how much p agrees with a voter's pairwise preference, we could instead assign 1, t, or 0 points to p, where $0 \leq t < 1$.

We give a loose justification for why such voting systems would remain Condorcet, and neutral. Firstly, neutrality holds because no candidate receives preferential treatment. Then, we might expect the Condorcet criterion to hold because the Condorcet criterion is concerned with pairwise head-to-head elections, and these systems seem like they should always assign more points to a candidate that wins a head-to-head election against another candidate. Finally, we don't have a strong justification for why these systems would remain consistent. This is just a hope that seems plausible.

5.3 Generalizing the Eigenstory

We now present another way to compare the Kemeny rule and Borda count. Specifically, we ask: How often, and in what situations, do the Kemeny rule and the Borda count output the same result? When they are different, how far apart are those results? What if one wants to use the Kemeny rule, but settles for the Borda count as an approximation which is faster to compute?

It turns out others have asked these same questions, and have began addressing them in the complete ranking case. Specifically, Eric Sibony proved a result which puts an upper bound on the distance between the Kemeny rule result and the Borda count result (Sibony (2014)).

In Sibony's 2014 paper, Eric Sibony was able to find an upper bound on the distance between the result of a Kemeny rule election and a Borda election.

To do so, Sibony uses tools from linear algebra. Specifically, he analyzes the general versions of the matrices K and B that we have defined here—the K and B we have defined in this thesis are the matrices for the 3 candidate case, whereas Sibony considers these matrices used for elections on n candidates, for arbitrary n.

Sibony uses spectral decomposition to analyze these matrices. Specifically, he explicitly finds the eigenspaces of K. These eigenspaces can be thought of as follows: An all-ones space (corresponding to voter profiles where an equal number of people vote for every choice), a Condorcet space, and a Borda space. So, the Kemeny rule can be thought of as a procedure which takes a profile vector, looks at it's projection into these spaces, scales those projections, then adds the results back together.

The Borda count can be thought of in a similar manner; it takes a profile vector, looks at the profile's projection into *those same eigenspaces*, scales them, and adds them back together. However, the Borda count and the Kemeny rule scale those spaces differently—the Borda count scales the *Condorcet* component of the profile down to 0, meaning that there is an entire subspace of the profile that the Borda count maps to the zero vector.

Finally, Sibony uses this understanding of the Kemeny rule and the Borda count to prove an upper bound on the distance between the Borda count result and the Kemeny rule result.

For a more rigorous, thorough treatment of these ideas, see Sibony (2014).

Might we be able to make similar statements in this new setting with partial orders? Sibony's result relies on decomposing the profile space into eigenspaces, and analyzing those eigenspaces. In particular, Sibony's result utilizes the fact that, in the complete ranking case, the Kemeny rule and Borda count both break the profile space into the same subspaces, compute a profile's projection into those subspaces, scale those projections by different amounts, and add them together at the end.

Do the Borda count and the Kemeny rule share this relationship in the partial order setting? When we open the door to partial orders, we suddenly have many, many, many more possible ballots, and so our space of voter profiles increases drastically in dimension. Perhaps this increase in dimension complicates matters. Perhaps the generalizations of the Borda count and Kemeny rule that we've defined here are not actually the voting systems for which this story works out as we'd like.

This seems like a promising direction for future research.

5.4 Final Words

Thank you for reading!

Bibliography

Arrow, Kenneth J. 1950. A difficulty in the concept of social welfare. *Journal of Political Economy* 58(4):328–346. URL http://www.jstor.org/stable/1828886.

Bogart, Kenneth P. 1973. Preference structures. I. Distances between transitive preference relations. *J Mathematical Sociology* 3:49–67. doi:10.1080/0022250x.1973.9989823. URL https://doi-org.ccl.idm.oclc.org/10.1080/0022250x.1973.9989823.

———. 1975. Preference structures. II. Distances between asymmetric relations. *SIAM J Appl Math* 29(2):254–262. doi:10.1137/0129023. URL https://doi-org.ccl.idm.oclc.org/10.1137/0129023.

Borda, J-C de. 1784. Mémoire sur les élections au scrutin (memoir on elections by ballot). *Histoire de l'Académie royale des sciences avec les mémoires de mathématique et de physique tirez des registres de cette Académie* 657—-665. URL https://gallica.bnf.fr/ark:/12148/bpt6k35800/f787.item#.

Condorcet, marquis de. 1785. Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix. *Paris: Imprimerie Royale*, URL https://gallica.bnf.fr/ark:/12148/bpt6k417181.

Conitzer, Vincent, Andrew Davenport, and Jayant Kalagnanam. 2006. Improved bounds for computing Kemeny rankings. In *AAAI*, vol. 6, 620–626.

Conitzer, Vincent, Matthew Rognlie, and Lirong Xia. 2009. Preference functions that score rankings and maximum likelihood estimation. In *Proceedings of the 21st International Joint Conference on Artificial Intelligence*, 109–115. IJCAI'09, San Francisco, CA, USA: Morgan Kaufmann Publishers Inc.

Cook, Wade D., Moshe Kress, and Lawrence M. Seiford. 1986. Information and preference in partial orders: A bimatrix representation. *Psychometrika* 51(2):197–207. URL https://doi.org/10.1007/BF02293980.

Crisman, Karl-Dieter, and Michael E. Orrison. 2017. Representation theory of the symmetric group in voting theory and game theory. In *Algebraic and geometric methods in discrete mathematics*, *Contemp. Math.*, vol. 685, 97–115. Amer. Math. Soc., Providence, RI. doi:10.1090/conm/685. URL https://doi.org/10.1090/conm/685.

Cullinan, John, Samuel K. Hsiao, and David Polett. 2014. A Borda count for partially ordered ballots. *Soc Choice Welf* 42(4):913–926. doi: 10.1007/s00355-013-0751-1. URL https://doi-org.ccl.idm.oclc.org/10.1007/s00355-013-0751-1.

Emerson, P. 2016. *From Majority Rule to Inclusive Politics*. Springer International Publishing. URL https://books.google.com/books?id=6txlCwAAQBAJ.

Fraenkel, Jon, and Bernard Grofman. 2014. The Borda count and its real-world alternatives: Comparing scoring rules in Nauru and Slovenia. *Australian Journal of Political Science* 49(2):186–205. doi:10.1080/10361146.2014.900530. URL https://doi.org/10.1080/10361146.2014.900530, https://doi.org/10.1080/10361146.2014.900530.

Kemeny, John G. 1959. Mathematics without numbers. *Daedalus* 88(4):577–591. URL http://www.jstor.org/stable/20026529.

Lehtinen, Aki, and Jaakko Kuorikoski. 2007. Unrealistic assumptions in rational choice theory. *Philosophy of the Social Sciences* 37(2):115–138. doi: 10.1177/0048393107299684. URL https://doi.org/10.1177/0048393107299684.

Monteiro, Luiz F., Sonia Savini, and Ignacio Viglizzo. 2017. Hasse diagrams of posets with up to 7 elements, and the number of posets with 10 elements, without the use of computer programs. doi:10.48550/ARXIV.1710.10343. URL https://arxiv.org/abs/1710.10343.

Saari, Donald G., and Vincent R. Merlin. 2000. A geometric examination of Kemeny's rule. *Soc Choice Welf* 17(3):403–438. doi:10.1007/s003550050171. URL https://doi-org.ccl.idm.oclc.org/10.1007/s003550050171.

Scheinerman, E.A. 2012. *Mathematics: A Discrete Introduction*. Cengage Learning. URL https://books.google.com/books?id=DZBHGD2sEYwC.

Sibony, Eric. 2014. Borda count approximation of Kemeny's rule and pairwise voting inconsistencies. In *Proceedings of the NIPS 2014 Workshop on Analysis of Rank Data.*

Young, H. P., and A. Levenglick. 1978. A consistent extension of Condorcet's election principle. *SIAM J Appl Math* 35(2):285–300. doi:10.1137/0135023. URL https://doi.org/10.1137/0135023.