2009

Range voting is resistant to control

Curtis Menton

Range Voting is Resistant to Control
Master of Science Thesis

Curtis Menton

Rochester Institute of Technology
Golisano College of Computing and Information Science
Department of Computer Science
102 Lomb Memorial Drive
Rochester, NY 14623
USA

December 14, 2009

Signatures

Dr. Edith Hemaspaandra — Date
Committee Chair

Dr. Christopher Homan — Date
Reader

Dr. Ivona Bezakova — Date
Observer

Contents

Abstract

Social choice theory is concerned with developing and evaluating voting systems, both for the use of political and organizational elections and for use as decision making process for multiagent systems. Particularly in the context of multiagent systems, computational resistance to various types of control has become a desired property of a voting system. Though manipulative actions may always be possible, strong computational barriers to efficient control can give us sufficient confidence in the integrity of an election.

Range Voting is a natural extension of approval voting that is resistant to a large number of cases of control. In particular, the variant Normalized Range Voting has among the largest number of control resistances among natural voting systems.

Chapter 1

Introduction

Many of the key results in voting theory show that all voting systems are flawed in some way. Arrow's impossibility theorem states that that in any election with more than two candidates any voting system will disobey at least one of several reasonable and natural criteria [1]. The Gibbard-Satterthwaite and Duggan-Schwartz theorems show that all reasonable voting systems are susceptible to strategic voting, where a voter votes counter to their true preferences in order to achieve a better outcome [22] [27] [12]. A dishonest election organizer might always be able to subtly alter the election to achieve their desired end. Thus, much of the subsequent study of voting systems has been directed towards finding the best compromises and most reasonable, if imperfect solutions.

The concept of election control represents cases where the authority conducting the election attempts to alter the outcome by changing the structure of the election. The study of control of elections was initiated by Bartholdi, Tovey, and Trick [3], who also introduced a novel defense against it. Even if control is possible, it may be computationally very difficult to find an ideal plan. The standard tools of complexity theory can be brought to bear on the problem and help to restore confidence in election systems. In many cases, a control problem can be shown to be NP-hard and therefore very unlikely to be solvable in polynomial time. We may be able to accept theoretical vulnerability to control if computational difficulty would make it essentially impossible for any computationally limited attacker.

Since the initial work of Bartholdi et al. a number of voting systems have been studied with an eye towards computational resistance. Several systems have been found with a high number of resistances [13] [23] [17], although some of them are not sufficiently natural for practical use, remain vulnerable to some of the cases of control, or have other technical flaws. Thus it is still desirable to search for natural and robust voting systems with high degrees of control resistance.

Range voting is a voting system with an alternate voter preference representation that allows a voter to score their level of approval of each candidate [28]. Though primarily lauded for its expressiveness and good behavior with several candidates, it also has an increased degree of resistance to control over closely related systems. We will also introduce a variant of range voting which has the highest degree of resistance to control among natural voting systems.

The rest of this thesis will provide background on complexity theory and voting theory and then present original results on the computational properties of range voting and normalized range voting. Chapter two provides a brief overview of complexity theory as pertains to this work. Chapter three provides some background on voting theory and computational social choice, with the systems under study described in sections 3.10 and 3.11. Chapter four contains original results pertaining to these two systems. A reader with a reasonable background in complexity theory and computational social choice should be able to manage by reading the sections on these voting systems and the results, while others may benefit from reading the previous chapters as well.

Chapter 2

Complexity Theory

One of the key topics in computer science is analyzing the computational difficulty of various problems.

Though general algorithms may have a wide variety of types of output, in the context of complexity theory, we are concerned primarily with *decision problems*, those where the output is just yes or no. Though this may seem like a severe limitation, it is not much of an issue, as most problems can easily be phrased in this way. For example, instead of computing the size of the largest clique in a graph, we can ask if there is a clique of size $\geq k$. Given an algorithm for that problem, the actual size could be extracted through binary search in a reasonable number of steps.

Modern computers are too complicated for mathematical analysis and so the model of choice is the Turing machine, an abstract machine consisting of a rewritable tape for input and memory and a finite set of defined transitions. Though Turing machines are generally more difficult to program compared to our modern computers, the model is simpler and easier to define formally our notions of algorithms, running time, and other key concepts with.

Furthermore, it is equivalent in power to any other reasonable computational model, including modern computers and programming languages. So while the important theorems and classes are formally defined in terms of Turing machines, much work can be done more informally with the use of a more a familiar computational model.

Problems are organized into complexity classes, which define their requirements of computation and performance characteristics. One of the simplest and most important is the class P, which contains all problems which are computable in polynomial time, or where the runtime is bounded by some polynomial on the size of the input. An algorithm would not normally be considered to be very fast if it runs in time $O(n^{100})$, but in practice most problems with polynomial-time algorithms are not especially high in degree, and so we frequently use this as a convenient analogue for tractability.

Also, if a problem has a polynomial-time algorithm for any reasonable Turing-equivalent computer model, there will be a polynomial algorithm for all Turing-equivalent models. This conveniently allows us to work in the more familiar territory of modern computers while achieving results that apply across all types of computers, including Turing machines. This also solidifies the usefulness of P for this analysis, as membership in P is independent of the messy details of the computation.

The class NP consists of problems that are computable on a non-deterministic Turing machine (NTM) in polynomial time. This model is equivalent to a Turing machine that is allowed to guess what path to take whenever it reaches a branch, or one that can simultaneously run a number of paths at the same time. For instance, an NTM for the max clique problem could guess for each vertex whether to try to include that vertex in the clique it is searching for. An algorithm on an NTM can be naively simulated on a deterministic TM with an exponential slowdown. However, that is not necessarily the best that can be done.

One can show a upper bound on the complexity of a problem by simply providing an algorithm for it. Providing an lower bound is somewhat more difficult. Just because no one has yet successfully designed a polynomial-time algorithm for a problem does not mean it does not exist. However the technique of problem reductions is useful for this purpose.

Formally, a problem L is many-one polynomial-time reducible to M (written $L \leq_m^p M$) if there is a polynomial-time function f such that for every instance x of L, x is a yes-instance for L if and only if $f(x)$ is a yes instance for M. We can then use M to solve L with only a polynomial amount of additional work for the bookkeeping. Consequently, L can be no harder than M. Also, it shows that we can embed L inside of M, and so M is at least as hard as L. It will be helpful to define a few more classes.

Definition 1. *A language L is NP-hard if for every $L' \in$NP, $L' \leq_m^p L$*

Definition 2. *A language L is NP-complete if it is NP-hard and additionally it is in NP*

The NP-complete problems are thus in a sense the hardest problems in NP, and that seems like a good goal if we want to prove a problem's difficulty. Showing that a problem is in NP is usually easy, but showing that each problem in an infinite class can be reduced to one sounds a bit imposing. Fortunately, most of the work has been done for us. The Cook-Levin theorem, independently discovered by Stephen Cook and Leonid Levin, showed that the boolean satisfiability problem is NP-complete, and so such a problem does actually exist [9]. From this point, rather than show that every possible problem in NP can reduce to our problem, we can just reduce from one of the thousands of known NP-complete problems, a much more reasonable process.

Now, of course, since membership in P is our de facto measure of tractability, we want to know whether the problems in NP-complete are actually distinct from P. NP-complete is distinct from P unless the entirely of NP collapses down to P. It remains one of the most important open problems in computer science whether P=NP. An enormous amount of results are contingent on the way this lies.

For instance, the problem of factoring integers lies in NP and thus would have a polynomial time algorithm if P=NP. This problem is at the foundation of the RSA algorithm, a public key cryptosystems that is essential for secure communication and commerce on the internet. If P=NP, this algorithm would be effectively broken, and any archived encrypted communications could potentially be cracked. Every time someone sends their credit card information over the internet, they are depending on P not being equal to NP, whether they know it or not. If P=NP countless problems of practical significance such as image recognition, language processing, and industrial optimization would suddenly be much easier. Other scientific disciplines would be affected as well, as fundamental problems in economics, biology, physics, and mathematics are known to be NP-complete [19].

Most computer scientists suspect that P$\neq$NP though it continues to elude proof [25]. Such is the import of the problem that it is included as one of the Millennium Prize problems by the Clay Mathematics Institute, putting a million dollar bounty on a proof either way. Still, while the problem remains open, membership in NP-complete is considered to be very strong evidence for the intractably of a problem, and it is at least an indication that a fast, exact algorithm will not soon be forthcoming.

Chapter 3

Voting Theory

3.1 Elections and Voter Representation

The standard model of an election is as a tuple $E = (C, V)$ where C is the set of candidates and V is the set of voters, each identified by their preferences. The voters preferences are traditionally represented as a strict ordering of the candidates from most preferred to least preferred. This is the primary model dating back to the work of Arrow [1]. This model forces the voter's preferences to be rational, that is, they are transitive and acyclic. Occasionally irrational voter preferences are considered, where voters express preferences as a potentially cyclic relation over C [17]. Other than any sort of ranked ballot the most common voter model is a $0, 1$ vector over the candidates denoting approval or disapproval, most prominently used by approval voting.

3.2 Condorcet's Criterion

The Marquis de Condorcet developed a voting system in 18th century France based around the concept of pairwise comparisons of the candidates. The idea is that the winner should be a candidate that is pairwise preferred to every other candidate in the election, with such a candidate known as the Condorcet winner. However, such a candidate will not always exists. Even when voter's preference relations are guaranteed to be rational, the combined societal preferences may be irrational and contain cycles. This is known as the "Condorcet paradox." Still, some voting systems are designed to select as the winner the Condorcet winner if it exists, and are described as "Condorcet methods" or "Condorcet compatible."

For instance, consider the following election among candidates a, b, and c, and with the voters shown below.

# Voters	Preferences
1	$a > b > c$
1	$b > c > a$
1	$c > a > b$

Two out of three prefer a to b, b to c, and c to a. Thus there is a cycle in the aggregate preference relation and this election admits no clear winner.

3.3 Arrow's Theorem

Arrow's theorem implies that all election systems are flawed in some way. His election model requires voter preferences be provided as strict orderings and a voting system must generate an aggregate preference ordering of the candidates rather than just a single winner. Given an election with at least three candidates and following this model, no voting system can satisfy all of the following:

Nondictatorship No one voter always decides the election.

Citizen sovereignty All aggregate preferences are possible.

Monotonicity If a voter increases their ranking of a candidate, that candidate does not do any worse.

Independence of irrelevant alternatives (IIA) Aggregate preferences between two candidates should only depend on the relative preferences between those two candidates. Changes to the candidate set or changes to the rankings of other candidates should not change the result of an election.

The first three criteria are reasonably straightforward but the fourth deserves a bit more explanation. IIA captures that a preference for one alternative over another should depend on any other alternatives. Yet many voting systems violate this. For example, plurality does not satisfy IIA. Consider the following election with $C = \{a, b, c\}$ and V as follows.

# Voters	Preferences
5	$a > b > c$
4	$b > a > c$
2	$c > b > a$

Here, plurality will rank the candidates $a > b > c$ and thus a will be the winner. Consider, however, what happens if c, the third place candidate, drops out. The voters will now be as follows.

# Voters	Preferences
5	$a > b$
6	$b > a$

The aggregate preferences flip and b changes to become the winner of the election, an arguably undesirable result.

Even the more obvious criteria can fail in otherwise reasonable voting systems. For instance, the system single transferable vote fails monotonicity [11]. The system is defined as follows. Voters vote by specifying their entire list of preferences from first to last. If any candidate receives the most first place votes, they are the winner of the election. If there is no majority winner, the candidate with the fewest first place votes is eliminated and the process repeats. This system is used in political elections around the world. Its use is advocated in the United States by the organization FairVote under the name "instant runoff voting" [14].

Consider the following example election over the candidates $\{a, b, c\}$

# Voters	Preferences
6	$a > b > c$
5	$b > a > c$
4	$c > b > a$
2	$c > a > b$

Initially, b is the first candidate eliminated. The votes of the second class of voter shift to a and a wins the election. However, consider if the fourth class of voters all change their votes to rank a first and the following election occurs instead:

# Voters	Preferences
6	$a > b > c$
5	$b > a > c$
4	$c > b > a$
2	$a > c > b$

Now, c will be eliminated in the first round. The votes of the third class of voters will shift to b and b will win the election. By raising their preference for a, the voters caused a to lose the election, and so STV is not monotonic.

3.4 Gibbard-Satterthwaite and Duggan-Schwartz Theorems

The Gibbard-Satterthwaite theorem [22] [27] and later the more general Duggan-Schwartz theorem [12] show that any reasonable voting system is vulnerable to strategic voting. Barring dictatorships and voting systems with certain uses of randomness, there will always be cases where a voter will receive a more beneficial result by distorting their true preferences.

Consider the following election over candidates $\{a, b, c\}$, using the plurality voting system.

# Voters	Preferences
5	$a > b > c$
4	$b > c > a$
2	$c > b > a$

Initially, a will be the winner. However, note that if the two voters of the third type instead vote as $b > c > a$, b will instead be the winner, a better result according to their true preferences.

3.5 Voting Systems

Whatever the preference representation, a voting system aggregates the voter preferences to determine a winner. Note that the preference representation includes much more information than many real-world ballots. However voting systems are not required to use all of the information on the ballot.

3.5.1 Scoring Protocols

A large class of voting systems qualify as *scoring protocols*. A scoring protocol is a vector $(\alpha_1, \alpha_2, \ldots, \alpha_m)$ of natural numbers in non-increasing order, where each number denotes the amount of points awarded to the candidate placing at that position on a particular ballot. The final ranking and winner are determined by summing up the points awarded by every ballot. For instance, consider the scoring protocol $(3, 1, 0)$ applied to the follow election:

# Voters	Preferences
3	$a > b > c$
2	$b > c > a$
1	$c > b > a$

a receives $3 * 3$ points from the first type of voters and none from the rest for 9 total. b receives $2 * 3$ points from the first class, $3 * 2$ from the second, and $2 * 1$ from the third for 14 total. c receives none from the first voters, but then $2 * 2$ from the second and $3 * 1$ from the third for 7 points total. Thus the final preference ordering is $b > a > c$.

As a scoring protocol is a fixed-length vector, families of scoring protocols exists, which are sets of scoring protocols which use the same scoring technique over all sizes of elections. Many common voting systems can be represented as families of scoring protocols.

Plurality

Plurality is the most familiar real world election system. It can be represented as a family of scoring protocols as $(1, 0, \dots, 0)$.

Veto

Veto is in a sense the inverse of plurality, where each voter must reject exactly one candidate. As a family of scoring protocols it is $(1, 1, \dots, 1, 0)$.

k-Approval

k-approval is a family of systems where voters must select and approve exactly k candidates. As a family of scoring protocols this is $(\underbrace{1, \dots, 1}_{k}, \underbrace{0, \dots, 0}_{m-k})$.

Borda

Borda count awards points in steadily decreasing order from the most preferred down. It can be represented as $(m-1, m-2, \ldots, 1, 0)$ as a family of scoring protocols.

3.5.2 Condorcet Compatible

Another class of voting systems follow Condorcet's ideas by basing the election on pairwise comparisons of the candidates and meeting the Condorcet criteria.

Dodgson Dodgson elections were introduced by Charles Dodgson, better known by his pen name Lewis Carrol.

The winner of the election is the candidate with the lowest Dodgson score, where the Dodgson score is the number of swaps of adjacent candidates in a voter's preferences are required to make a candidate the Condorcet winner. Naturally, if there is a Condorcet winner, this candidate requires zero swaps and so they will win the election.

Consider the following election among candidates $\{a, b, c, d, e, f\}$.

# Voters	Preferences
10	$a > d > b > e > c > f$
10	$c > a > e > b > d > f$
10	$b > f > c > a > d > e$

This election has no Condorcet winner as $a, b,$ and c form a cycle with a preferred to b, b preferred to c, and c preferred to a. However, we can make a the Condorcet winner by making just one preference swap, resulting in, for instance, the following election.

# Voters	Preferences
10	$a > d > b > e > c > f$
9	$c > a > e > b > d > f$
1	$a > c > e > b > d > f$
10	$b > f > c > a > d > e$

The other two near-Condorcet candidates will require at least two swaps to become Condorcet winners, as the candidates d, e, f, though they are not close to winning the election, are in the way of the candidates they need to supersede.

Though this seems to be a reasonable voting system, it is actually NP-hard to calculate the Dodgson score of a candidate, severely limiting its practicality [2]. Approximation algorithms have been developed for finding the [8], though this does not lead to the kind of confidence we usually seek in a voting system.

Llull-Copeland Llull-Copeland voting is a Condorcet-compatible voting system with polynomial-time winner determination. The system was described by Copeland in 1951 [10], though it was recently discovered to have been first described by the 13th century mystic Ramon Llull [17] who proposed it as a method for electing figures in the church.

The system works by selecting the candidate that has the most pairwise victories over other candidates. Variations on the system are possible by changing how many points are awarded for a tie. Clearly, this is a Condorcet-compatible system, as a Condorcet winner will have the maximum number of pairwise victories possible.

3.6 Control

Control represents the efforts of a centralized authority, the chair of an election to alter the structure of the election in order to affect its outcome.

This involves changing either the candidate or voter sets or partitioning either into subelections. In real world political elections, this corresponds to voter fraud and voter suppression, back-room dealings with potential candidates, and gerrymandering and similar manipulations. In the context of multiagent systems, it is related to any efforts by a system designer or administrator to alter the results by changing the parameters of the system.

More formally, for the purposes of the complexity theoretic analysis of the control problems, we will analyze the cases of control in the form of decision problems. That is, we will define a problem where the goal is to find whether in a particular election a certain case of control can succeed in its goals.

The goal is to classify a voting system as vulnerable, resistant, or immune to each of the various cases of control, with these terms initially defined by Bartholdi et al. [3] and widely adopted since. It is helpful now to define these notions precisely.

Vulnerability A voting system is *vulnerable* to a case of control if that action has potential to affect the result of an election, and the associated decision problem can be solved in polynomial time; that is, it is in P. This has a very good practical correspondence with real world efficiency of the problem, and thus the case of control is computationally easy.

Resistance A voting system is *resistant* to a case of control if that action has potential to affect the result of an election, and the associated decision problem is NP-hard. The idea of NP-hardness has a long and storied history, but for the current purposes, it suffices to say that such problems are very unlikely to have efficient solutions, barring a major shift in our understanding of computer science.

Immunity A voting system is *immune* to a case of control if that action cannot affect the result of the election. This is obviously a desirable notion but it is generally harder to come by, and many immunities are incompatible with very basic and reasonable properties of voting systems [13].

The control cases of Bartholdi et al. were all *constructive*, that is, the control is directed towards making a distinguished candidate the winner. In some cases, a malicious chair could conceivably want above all to prevent a particular candidate from winning the election, regardless of who else wins. This idea was introduced by Conitzer et al. as *destructive* manipulation and later by Hemaspaandra et al. in the context of control [24]. Though this may seem to be a less desirable goal, it may be feasible in some cases where constructive control is not and thus is it also worth studying.

Among the cases of control are control by adding or deleting either voters or candidates. In the case of adding voters or candidates, the new participants must be chosen from a set rather than arbitrarily created. While this type of control is not necessarily thus limited, the decision problems are defined as having a limit on the number of voters or candidates that can be added or deleted. In the candidate cases, the distinguished candidate must be in the original candidate set. In the cases of destructive control by deleting candidates, the distinguished candidate cannot be among those deleted as that would trivially solve the problem.

The various cases of control by partition are not quite straightforward and deserve a little explanation. In any control by partition problem, initial subelections are performed with segments of the voter and candidate sets and a final election is performed with the candidates that survive these subelections.

In control by partition of voters, the voter set is partitioned into two subsets and an subelections are run with each (with the original candidate set). The candidates that survive each subelection face off to find the final winner of the election.

Control by partition of candidates has two major variants. In one variant, *control by partition*, one set of candidates is separated off from the rest for an initial subelection. Whatever candidates survive this election then rejoin the rest of the candidates for the final election with the entire voter set. In the other variant, *control by run-off partition*, the candidate set is partitioned into two sets and each set conducts an initial subelection. The candidates that survive each of these elections then are brought together for the final election with the entire voter set.

There is an additional variation in the tiebreaking rule that is chosen in the subelections. In the case of a tie, either all of the top scoring candidates are promoted to the final election, or none of them are. These two cases are called *ties-promote* and *ties-eliminate*. Notably, in the second case, an election can fail to elect any candidate. Though these may seem like subtle differences, many voting systems will resist one of the cases while being vulnerable to another. It is probably sufficient to skim this section and read a few definitions to get the flavor, and refer to it later to clarify a specific definition.

Control by Adding Candidates

Given An election $E = (C, V)$, a distinguished candidate $w \in C$, a spoiler candidate set D, and $k \in \mathbf{N}$

Question (Constructive Is it possible to make w the winner of an election $(C \cup D', V)$ with some $D' \subseteq D$ where $|D'| \leq k$?

Question (Destructive) Is it possible to make w not the winner of an election $(C \cup D', V)$ with some $D' \subseteq D$ where $|D'| \leq k$?

Control by Deleting Candidates

Given An election $E = (C, V)$, a distinguished candidate $w \in C$, and $k \in \mathbf{N}$

Question (Constructive) Is it possible to make w the winner of an election $(C - C', V)$ with some $C' \subseteq C$ where $|C'| \geq k$?

Question (Destructive) Is it possible to make w not the winner of an election $(C - C', V)$ with some $C' \subseteq (C - \{w\})$ where $|C'| \leq k$?

Control by Adding Voters

Given An election $E = (C, V)$, a distinguished candidates $w \in C$, an additional voter set U, and $k \in \mathbf{N}$

Question (Constructive) Is it possible to make w the winner of an election $(C, V \cup U')$ for some $U' \subseteq U$ where $|U'| \leq k$?

Question (Destructive) Is it possible to make w not the winner of an election $(C, V \cup U')$ for some $U' \subseteq U$ where $|U'| \leq k$?

Control by Deleting Voters

Given An election $E = (C, V)$, a distinguished candidates $w \in C$, and $k \in \mathbf{N}$

Question (Constructive) Is it possible to make w the winner of an election $(C, V - V')$ for some $V' \subseteq V$ where $|V'| \leq k$?

Question (Destructive) Is it possible to make w not the winner of an election $(C, V - V')$ for some $V' \subseteq V$ where $|V'| \leq k$?

Control by Partition of Candidates

Given An election $E = (C, V)$ and a distinguished candidates $w \in C$

Question (Constructive) Is there a partition C_1, C_2 of C such that w is the final winner of the election $(D \cup C_2, V)$, where D is the set of candidates surviving the initial subelection (C_1, V)?

Question (Destructive) Is there a partition C_1, C_2 of C such that w is not the final winner of the election $(D \cup C_2, V)$, where D is the set of candidates surviving the subelection (C_1, V)?

Control by Runoff Partition of Candidates

Given An election $E = (C, V)$ and a distinguished candidates $w \in C$

Question (Constructive) Is there a partition C_1, C_2 of C such that w is the final winner of the election $(D_1 \cup D_2, V)$, where D_1 and D_2 are the sets of surviving candidates from the subelections (C_1, V) and (C_2, V)?

Question (Destructive) Is there a partition C_1, C_2 of C such that w is the final winner of the election $(D_1 \cup D_2, V)$, where D_1 and D_2 are the sets of surviving candidates from the subelections (C_1, V) and (C_2, V)?

Control by Partition of Voters

Given An election $E = (C, V)$ and a distinguished candidates $w \in C$

Question (Constructive) Is there a partition V_1, V_2 of V such that w is the final winner of the election $(D_1 \cup D_2, V)$ where D_1 and D_2 are the sets of surviving candidates from the subelections (C, V_1) and (C, V_2)?

Question (Destructive) Is there a partition V_1, V_2 of V such that w is not the final winner of the election $(D_1 \cup D_2, V)$ where D_1 and D_2 are the sets of surviving candidates from the subelections (C, V_1) and (C, V_2)?

3.7 Caveats

It is important to note that NP-hardness is a worst case notion and an NP-hard problem could still be easy in many cases, or it could be easy to find a good approximation. Recent research has shown that many NP-hard control problems are easy when voter's preferences are arranged in the common single-peaked model [18]. Work by Friedgut et al. has shown that random manipulations of the voters in an election are reasonably likely to be effective in any reasonable election system [20]. Other research has studied the application of approximation algorithms to manipulative problems in voting systems, which can potentially find a solution within a fixed bound of the ideal solution, even if the ideal solution is out of reach [15] [7]. Therefore this work is a good first step in providing confidence in the integrity of range voting, but it is not the final word.

3.8 Other Voting Systems and Control

Most natural systems have at least a few gaps with regard to control resistance. Several systems have been designed and considered specifically for their high number of resistances. Hemaspaandra et al. developed a hybrid election system that is resistant to all standard types of control [23]. Though the construction is highly unnatural, it resolved the open problem of whether it is possible for a election system to possess all of the resistances.

Copeland voting is a Condorcet-compatible voting system that ranks candidates based on their number of victories in head to head contests among the other candidates. Faliszewski et al. found Copeland voting to be resistant to every case of constructive control and to have a good number of resistances to destructive control [17]. This was also a useful result, as it established that natural voting systems could hold at least every constructive resistance. See table 3.1 for their results.

The system sincere-strategy preference-based approval voting, introduced by Brams and Sanver [6] and adapted to deal with control by Erdélyi et al. [13], has been shown to have a large number of resistances, and it is a fairly natural system. Brams and Sanver originally introduced SP-AV to discuss outcomes under various possible voter strategies and to integrate approval style ballots with the classical ranked order voter preferences [6]. Their system represents the voter's preferences as a ranked ordering of the candidates along with an approval threshold, where a voter approves of every candidate ranked at least as high as that threshold. In addition, there is the restriction that each voter must both approve of and reject at least one candidate with any other ballot considered inadmissible. This creates problems when dealing with candidate control, as a voter's ballot can be transformed from admissible to inadmissible if their only approved candidate is deleted, for instance. Erdélyi et al. suggested to add an additional step of vote coercion to the system. If a voter supplies an inadmissible vote, either accepting or rejecting all the candidates, their approval threshold is shifted by one candidate in the appropriate direction. This allows every voter to have an admissible vote in the end, and it also gives the system interesting properties with regard to control. The system in effect captures the resistance of both approval voting and plurality. The results are listed in table 3.2. However, Baumeister et al. took issue with the system as presented [4]. The vote coercion step is not properly a part of Brams and Sanver's original system. Also, the rule makes an arbitrary decision about where to set the new threshold when coercing a vote. Baumeister et al. propose a subtly different system that more explicitly includes the coercion step. The system remains somewhat unsatisfying, as its very effect is to force a distinction between candidates that a voter may have honestly ranked the same. Therefore it is still desirable to continue the search for a natural system with a similar or better level of resistance.

Table 3.1: Control Results for Plurality and Copeland [17]

Control by	Tie Model	Plurality		Copeland	
		C	D	C	D
Adding candidates		R	R	R	V
Deleting candidates		R	R	R	V
Partition of Candidates	TE	R	R	R	V
	TP	R	R	R	V
Run-off Partition of Candidates	TE	R	R	R	V
		R	R	R	V
Adding Voters		V	V	R	R
Deleting Voters		V	V	R	R
Partition of Voters	TE	V	V	R	R
	TP	R	R	R	R

3.9 Manipulation and Bribery

Manipulation and bribery are the other classes of manipulative actions that are studied in the context of computational social choice. The manipulation problem studies attempts of a voter coalition to alter the outcome of an election by strategically coordinating their votes. As with control problems we can define constructive and destructive versions of this problem.

Table 3.2: Control Results for Approval and Sincere-Strategy Preference-Based Approval Voting (SP-AV) [13]

Control by	Tie Model	Approval		SP-AV	
		C	D	C	D
Adding candidates		I	V	R	R
Deleting candidates		V	I	R	R
Partition of Candidates	TE	V	I	R	R
	TP	I	I	R	R
Run-off Partition of Candidates	TE	V	I	R	R
	TP	I	I	R	R
Adding Voters		R	V	R	V
Deleting Voters		R	V	R	V
Partition of Voters	TE	R	V	R	V
	TP	R	V	R	R

Manipulation

Given An election $E = (C, V)$, a distinguished candidate $w \in C$, and an additional collection of manipulators V' with votes as yet unassigned

Question (Constructive) Is it possible assign the votes of V' in a way to make w the winner of the election $(C, V \cup V')$?

Question (Destructive) Is it possible assign the votes of V' in a way to make w not the winner of the election $(C, V \cup V')$?

Note that this differs slightly from the traditional social choice definition as in the Gibbard-Satterthwaite theorem, which instead is concerned with voters who already have preferences changing their vote as to get a better outcome by these preferences. For the rest of this thesis, "manipulation" will refer to the computational social choice definition given above.

The bribery problem is concerned with an outside manipulator with limited resources attempting to change the outcome of an election by buying the votes of some limited set of voters.

Bribery

Given An election $E = (C, V)$, a distinguished candidate $w \in C$, and $k \in \mathbf{N}$

Question (Constructive) Is it possible to make w the winner of the election by changing the votes of up to k voters?

Question (Destructive) Is it possible to make w not the winner of the election by changing the votes of up to k voters?

Manipulation is easy in many voting systems while bribery is more frequently difficult. Bribery is usually at least as hard as manipulation for a given voting system. We can think of the bribery problem of first finding an ideal set of voters to alter, and then running the manipulation problem on them to assign their votes. However this is not always the case, and unnatural voting systems do exist where the bribery problem is easy ($\in$ P) and the manipulation problem is hard (NP-hard) [16].

3.10 Range Voting

Range voting (RV) is a voting system with an alternate voter representation that allows voters to express their degree of approval in each candidate. We will describe a k-range election as $E = (C, V)$ where C is the set of candidates with $|C| = m$, and V is the set of voters with $|V| = n$ and for a voter $v \in V$, $v \in (0, 1, \ldots, k)^m$. Each voter expresses their preferences by giving a score for each candidate. The parameter k sets the highest score a voter is allowed to give a candidate. The winner of the election is the candidate with the highest cumulative score among all voters.

Example The following is an example of a 2-range election of the candidates $\{a, b, c\}$. a will be the winner with a total of 14 points.

# Voters	a	b	c
5	2	0	1
6	0	2	0
4	1	2	0

Though range voting is sometimes described allowing scores over a real interval such as $[0, 1]$ [28], this paper will deal with the more limited integral version for its practicality of implementation and to avoid issues with the size of representation. Our primary concern is to study the difficulty of decision problems relating to the system and allowing scores of unbounded size would greatly complicate that analysis. Note that any bounded size and precision real number representation would be equivalent to an integral representation, so this version will be just as expressive as a rational representation or any other which would be suitable for computational analysis.

Arrow's theorem was formulated with the traditional voter preference models of a strict ordering. Since range voting uses a different model, it is not bound by that result and, though subject to interpretation, achieves all of the normally impossible criteria [28] [26].

To demonstrate, let us revisit the earlier example that showed violation of IIA, which is typically the hardest criteria to achieve. Let us formulate this as a 1-range election, and assume that each voter only gives any points to their top candidate.

# Voters	a	b	c
5	1	0	0
4	0	1	0
2	0	0	1

Again, a wins this initial election. Now, if we remove the last place candidate c:

# Voters	a	b
5	1	0
4	0	1
2	0	0

The c voters are left not awarding any points to anybody, which is a perfectly legal vote, and perfectly rational, if one does feel no distinction between the candidates. Consequently the original result stands and a remains the winner.

3.10.1 Other Similar Voting Systems

Approval Voting Approval voting is a voting system where voter preferences are represented by approving or disapproving of each candidate individually [5].

Range voting can be seen as a straightforward extension of approval voting where voter's preferences are cast over a $\{0, \dots, k\}$ vector rather than a $\{0, 1\}$ vector. Consequently we can express an approval election as a 1-range election. This is quite useful for the purpose of analyzing the computational properties of range voting and finding problems to which it is resistant. Since approval is just a special case of range voting, any problem relating to approval which is NP-hard will automatically be NP-hard for range voting as well.

Utility Based Voting The model utility based voting allows voters to allocate some (generally limited) number of points among the candidates according to their preferences [15]. Utility is a term from economics that quantifies the value that one receives from a given outcome. These systems allow voters to express their personal utility for each of the candidates and use that information to select the winner. Parameters to the system are the total number of points a voter can allocate and the maximum number of points that can be allocated to a given candidate. Also, in a *free-form* election, voters are not required to allocate all their points. This model is flexible enough to describe such systems as plurality, approval, n-approval, and also range voting. A k-range election with m candidates is equivalent to a free-form utility based election allowing up to k points per candidate and mk points total for each voter.

Utilitarian Voting Range voting is part of a class of voting systems termed utilitarian voting by Hillinger [26]. This encompasses any system that allows the voters to independently score each candidate according to their personal utility derived from that candidate winning. This includes approval, range voting, and evaluative voting, which allows voters to score candidates as $-1, 0$, or 1. Such systems have the potential to satisfy conditions, such as the conditions of Arrow's impossibility theorem, which are out of reach for any voting system with the traditional ordered ballot [26].

3.11 Normalized Range Voting

A rational voter seeking to maximize their impact in an election would always give their most preferred candidate the highest score possible (k) and their least preferred candidate the lowest score possible (0). The system Normalized Range Voting (NRV) captures this and also gives the system more interesting behavior under several types of centralized control.

In this system each voter specifies their preferences as in standard range voting. However, as part of the score aggregation, the system normalizes each vote to the rational range $[0, k]$. Formally, for a voter v and their maximum and minimum scores a and b, their each score s for a candidate is changed to $\frac{k(s-b)}{a-b}$. If $a = b$, a voter shows no preference among the candidates and this vote will not be counted. The system does not make an effort to coerce such an unconcerned vote into one that distinguishes between the candidates.

The relationship between RV and NRV is closely analogous to the relationship between approval voting and SP-AV. The normalization step ends up removing several cases of control immunity, but it introduces more complex behavior on alterations of the candidate set that gain back a greater number of control resistances.

Unlike RV, NRV unambiguously fails the criteria independence of irrelevant alternatives. Consider a 2-NRV election with $C = \{a, b, c\}$ and V below.

# Voters	a	b	c
7	2	0	0
4	0	2	0
4	0	1	2

a will win this election with a score of 14, with 12 and 8 for their rivals b and c. However, consider the election with the same voters but with the candidate c removed.

# Voters	a	b
7	2	0
4	0	2
4	0	1

At first, a appears to still be winning the election. However the normalization step will scale up the votes from the third group of voters to give b 16 points in total, making b the winner of the election.

While this seems to be a negative against this system, this complex, shifting behavior on the changing of the candidates is exactly what allows us to achieve a large number of control resistances for NRV over RV.

Chapter 4

Results

This chapter will consist of technical proofs of immunity/susceptibility as well as proofs of complexity using the standard technique of many-one problem reductions. The idea is to show that we can convert an instance of a known difficult problem into an instance of the problem under study and thus our new problem is at least as hard as one known to be difficult. Table 4.1 summaries these results as well as comparing them to the resistances possessed by approval voting and SP-AV. The results for RV and NRV are original to this thesis.

4.1 Immunity and Susceptibility

Before analyzing resistance, it is necessary to examine whether it is in fact possible to alter an election through that type of control, that is, whether the voting system is *susceptible* or *immune* to that case of control.

Several of the control cases are linked in terms in susceptibility, as one may be the inverse of the other, or just a slightly more elaborate version. This simplifies the matter of achieving the full set of results as fewer cases actually have to be proved. Furthermore several susceptibility results fol-

Table 4.1: Control Results for Approval, Sincere-Strategy Preference-Based Approval Voting, Range Voting, Normalized Range Voting [13]

Control by	Tie Model	Approval		SPAV		RV		NRV	
		C	D	C	D	C	D	C	D
Adding candidates		I	V	R	R	I	V	R	R
Deleting candidates		V	I	R	R	V	I	R	R
Partition of Candidates	TE	V	I	R	R	V	I	R	R
	TP	I	I	R	R	I	I	R	R
Run-off Partition of Candidates	TE	V	I	R	R	V	I	R	R
	TP	I	I	R	R	I	I	R	R
Adding Voters		R	V	R	V	R	V	R	V
Deleting Voters		R	V	R	V	R	V	R	V
Partition of Voters	TE	R	V	R	V	R	V	R	?
	TP	R	V	R	R	R	V	R	R

low from simple properties of the voting system.

One simple property, the unique version of the Weak Axiom of Revealed Preference (or Unique-WARP) implies several control immunities.

Definition 3. *In a voting system satisfying Unique-WARP, a unique winner among a collection of candidates will remain the winner among any subcollection of which they are a part [3].*

Any voting system obeying this property will be immune to constructive control by adding candidates, destructive control by deleting candidates, and destructive control by partition or runoff partition of candidates in both tie handling models [24]

Theorem 4.1.1. *RV is immune to constructive control by adding candidates, destructive control by deleting candidates, and destructive control by partition or runoff partition of candidates.*

Proof RV can easily be seen to satisfy Unique-WARP. The unique winner among a collection of candidates has the highest sum score among all the voters. In any subcollection they will still have the highest score and they will still be the winner. Thus RV achieves the immunities. □

Theorem 4.1.2. *RV is immune to constructive control by partition and runoff partition in the ties-promote model.*

Proof In RV, any candidate that is in first place (possibly tied) will be in first in any subset of the candidates of which they are a part. Therefore they will survive any initial subelection in the ties-promote model. Consequently no candidate that is not already the unique winner can be made to be the unique winner through partition of candidates in this model. □

Hemaspaandra et al. [24] defined the notion of a *voiced* voting system, where an election with exactly one candidate will result in that candidate as the winner. For every voiced voting system, the following is true: it is susceptible to constructive control by deleting candidates, destructive control by adding candidates, and if it is susceptible to destructive control by partition of voters, it is susceptible to destructive control by deleting voters.

Theorem 4.1.3. *RV and NRV are susceptible to constructive control by deleting candidates, destructive control by adding candidates, constructive and destructive control by partition of voters, and destructive control by deleting voters.*

Proof Consider the following election over the candidates $\{a, b, c\}$, either a 2-range or 2-normalized-range election:

# Voters	a	b	c
2	2	0	0
4	1	2	0
3	1	0	2

The initial winner of this election is a. However, we can affect this election through partition of voters to make b the winner, as follows.

# Voters	a	b	c
1	2	0	0
4	1	2	0

# Voters	a	b	c
1	2	0	0
3	1	0	2

b wins the first subelection and c wins the second, with b coming ahead in the final election between the two. We have both prevented a from winning the election and made b the winner of the election. Thus RV and NRV are susceptible to constructive and destructive control by partition of voters in either tie handling model. This, plus the fact that both of these systems are voiced, imply the other cases. $\square$

Notably, NRV does not satisfy Unique-WARP and does not posses the immunities that RV does. We can show it is in fact susceptible to these cases of control.

Theorem 4.1.4. *NRV is susceptible to destructive control by deleting candidates and constructive control by adding candidates.*

Proof Consider the follow 2-normalized-range election over candidates $\{a, b, c\}$.

# Voters	a	b	c
7	2	0	0
4	0	2	0
4	0	1	2

The winner of the original election will be a. However, with candidate c deleted, the votes of the third class of voters will be normalized to allot 2 points to candidate b, and b will become the winner of the election. Therefore NRV is susceptible to destructive control by deleting candidates. Susceptibility to the other case follows from this [24]. $\square$

Theorem 4.1.5. *RV and NRV are susceptible constructive and destructive control by adding and deleting voters.*

Proof Consider the following election over $\{a, b\}$ (interpreted as either a 1-range or 1-normalized-range election).

# Voters	a	b
5	1	0
4	0	1

By deleting two or more of the first class of voter we will make a no longer the winner (destructive control by deleting voters) as well as making b the new winner (constructive control by deleting voters). These cases imply the other two. □.

Theorem 4.1.6. *RV is susceptible to constructive control by partition and runoff partition of candidates in the ties-eliminate model.*

Consider the following 1-range election over $\{a, b, c\}$.

# Voters	a	b	c
1	1	0	0
2	0	1	0
2	0	0	1

Initially, this election has no unique winner, though candidates b and c are tied for first. However, by partitioning the candidate set into $\{a\}, \{b, c\}$, the outcome is different. b and c will tie in their subelection and be eliminated. If this action is interpreted as control by runoff partition, a will win their single candidate election. In either case, a is the sole remaining candidate and becomes the winner of the election. □

Theorem 4.1.7. *NRV is susceptible to constructive and destructive control by partition and runoff partition of candidates in both tie handling models.*

Consider the following 2-normalized-range election over candidates $\{a, b, c\}$.

# Voters	a	b	c
7	2	0	0
4	0	2	0
4	0	1	2

The initial winner of this election is a. Partitioning the candidates into $\{a\}, \{b, c\}$ will change the outcome of the election. If this action is interpreted as runoff partition a will win their initial subelection, or otherwise $\{a\}$ is the remaining set of candidates. b will win their subelection and go on to win the final election over a after the scores are normalized. No ties occur so this will work in either tie handling model. We both prevent a from becoming the winner and cause b to become the winner, so NRV is susceptible to both constructive and destructive control by partition and runoff partition of candidates in both tie handling models.

4.2 Vulnerability Results

The majority of the problems to which RV and NRV are vulnerable have trivial algorithms which are only slight variations of those used for approval voting, which can be seen in the work of Hemaspaandra et al [24]. The following problems are sufficiently different and interesting to merit detailed inclusion.

4.2.1 Destructive Control by Partition of Voters

Theorem 4.2.1. *k-RV is vulnerable to destructive control by partition of voters in either tie handling model, for any fixed k.*

A dynamic programming algorithm, a variation of one used to solve PARTITION in pseudo-polynomial time [21] can be used to solve this problem. The idea of the algorithm is that if the distinguished candidate w

is the original winner, a partition must cause w to lose to one candidate in the first partition and another candidate in the second to make them lose the election. The algorithm loops over pairs of candidates and checks if a successful partition exists for that pair of candidates.

Given an instance $(C, V), w$ of the control problem with $|C| = m$ and $|V| = n$, first perform the initial check if w is the winner, and if not return true immediately. Then perform the following procedure for every pair of candidates in $C - \{w\}$.

Given w and the two other candidates c_1, c_2 from C, construct a $(2*m*k+1) \times (2*m*k+1) \times (n+1)$ three-dimensional array, with the first two indices ranging from $-2*m*k$ to $2*m*k$ and with the third index ranging from 0 to n. The cell at position (x, y, z) records whether a partition of voters in V_1, V_2 can be made using the first z voters (in some fixed order) such that x is the score of c_1 in V_1 minus the score of w in V_1, and y is the score of c_2 in V_2 minus the score of w in V_2. The cell $(0, 0, 0)$ will be marked as reachable and all cells will initially be marked as unreachable. Reachable cells will also record whether this allocation came from placing the last voter in V_1 or V_2, in order to aid in retrieving the actual allocation. A successful partition of the voters will exist causing w to lose the election if there is some reachable cell at an index (x, y, n) where $x, y > 0$ in the ties-promote model, or $x, y \geq 0$ in the ties-eliminate model.

The array can be filled in as follows. Given a reachable cell (x, y, z), fill in cell $(x + d_1, y, z + 1)$ with $(T, 1)$, denoting that that cell is reachable with voter $z + 1$ placed in V_1, where d_1 is the difference of the scores of c_1 and w according to voter $z + 1$. Fill in cell $(x, y + d_2, z + 1)$ with $(T, 2)$, denoting that that cell is reachable with voter $z + 1$ placed in V_2, where d_2 is the difference of the scores of c_2 and w according to voter $z + 1$. Also, do not rewrite over the contents of a cell if it has already been marked.

By proceeding in this way all reachable partitions will eventually be marked and a successful partition, if it exists, can be derived from work-ing backwards down the table from the cell at the top level by checking the contents of the cell and the preferences of the appropriate voter. If a successful partition is found for any this of candidates, exit and return true. Otherwise continue through to the next pair of candidates, eventually returning false if no successful partition is ever found.

This algorithm will run in time $O(m^4k^2n)$. Since k is a fixed constant and m and n are polynomial in the size of the input, the algorithm runs in polynomial time. Therefore RV is vulnerable to these control problems. $\square$

4.3 Resistance Results

4.3.1 Generalization of Resistance Results

The proofs here will refer necessarily to specific RV and NRV elections with a particular scoring range k. However we want to be able to show resistance for any arbitrary k.

Theorem 4.3.1. *If RV or NRV exhibits resistance to a case of control for a particular scoring range k, it will exhibit that resistance for any range $l \geq k$*

We can reduce an instance of a control problem for a k-range election to a l-range election for any $l \geq k$ by simply modifying the theoretical maximum score without changing anything else about the election. Since the bound does not affect the votes except by setting the maximum score, this will not change the election or its behavior under control.

The same construction will work for NRV though a bit more explanation is required. We can reduce any control problem for a k-normalized-range election to a l-normalized-range election for any $l \geq k$ by simply changing the maximum score and not changing the voters or candidates. In this case, the election is substantially changed as the increased bound will affect the normalization step of the scoring process. Note in a normalized range election, all votes that rank any candidates differently are normalized to $\frac{k(s-b)}{a-b}$ where a and b are the maximum and minimum scores allotted to a candidate by that particular voter. Thus in our new l-normalized-range election, all votes will instead be changed to $\frac{l(s-b)}{a-b}$, a change of $\frac{l}{k}$. All votes and thus all cumulative candidate scores will increase uniformly and whoever won the original election will win the new election as well. However, in cases of candidate control, the minimum and maximum scores for a particular voter can change as well and thus further affect the election. Consider a vote with the initial minimum and maximum scores of a and b and with new scores of a' and b' after changes to the candidate set due to control. If a and b and a' and b' are both distinct, this voter's scores will change from $\frac{k(s-b)}{a-b}$ to $\frac{k(s-b)}{a'-b'}$, a factor of $\frac{a-b}{a'-b'}$. In the new election, the voter's scores will change from $\frac{l(s-b)}{a-b}$ to $\frac{l(s-b)}{a'-b'}$. If either a and b or a' and b' are equal, that vote will be ignored in both the original k-normalized-range election and the l-normalized-range election. Therefore the score for any candidate from that voter will go from 0 to 0 and thus will also change at a factor of $\frac{l}{k}$. Thus in every case all scores and also the aggregate scores will change at a ratio of $\frac{l}{k}$ and so the election will proceed in the same way.

4.3.2 Results Derived From Approval

Due to RV's great similarity with approval voting, many results relating to approval trivially apply to RV and NRV.

Theorem 4.3.2. *If approval voting is resistant to a case of control, RV and NRV will also be resistant for any scoring range.*

This is easy to show. We can reduce from an instance of any approval control problem by simply considering the election a 1-range election or a 1-normalized-range election. A 1-range election is exactly equivalent so this will trivially work. For the NRV election, though this does technically include the normalization step which can modify the election, when the score range is 1, no normalization is actually performed, so again this election is equivalent to the original approval election. These results will also generalize to k-RV and k-NRV any $k \geq 1$ as previously described.

Theorem 4.3.3. *1-RV and 1-NRV are resistant to the following cases of control for: constructive control by adding voters, constructive control by deleting voters, and constructive control by the partition of voters in the ties-promote and ties-eliminate models.*

All of these resistances are derived from reductions from approval as described above, and the fact that approval is resistant to these cases of control [24]. □

4.3.3 Adding/Deleting Candidates

Theorem 4.3.4. *2-NRV is resistant to constructive and destructive control by adding or deleting candidates.*

This proof is based on a similar proof relating in SP-AV by Erdélyi et al. [13], which itself was based on proofs relating to plurality by Bartholdi et al [3] and Hemaspaandra et al. [24].

We will reduce from an instance of the hitting set problem, defined as follows [21].

Given: A collection $\mathcal{S}$ of subsets of a set B, $k \in \mathbf{Z}^+$

Question: Does B contain a hitting set B' of size k or less that contains at least one element from every $S \in \mathcal{S}$?

Given a hitting set instance $(B, (\mathcal{S}), \|)$ with $|B| = n$ and $|\int| = m$ we will construct a 2-range election. The candidate set C will consist of $B \cup \{c, w\}$. The voter set V will be as follows:

- $2m(k+1)+4n$ voters have a score of 2 for c, and a score of 0 for all other candidates.
- $3m(k+1)+2k+1$ voters have a score of 2 for w, and a score of 0 for all other candidates.
- For each $b \in B$, 4 voters have a score of 2 for b, a score of 1 for w, and a score of 0 for all other candidates.
- For each $S_i \in \mathcal{S}$, $2(k+1)$ voters have a score of 2 for b, for each $b \in S_i$, a score of 1 for c, and a score of 0 for all other candidates.

This will lead to scores in $(\{c, w\}, V)$ as follows:

Candidate	Score
c	$8m(k+1)+8n$
w	$6m(k+1)+8n+4k+2$

The candidate c will win with a margin of $2m(k+1)-4k-2$.

Additionally the scores in $(\{c, w\} \cup B, V)$ will be as follows:

Candidate	Score
c	$6m(k+1)+8n$
w	$6m(k+1)+4n+4k+2$
$b \in B$	$\leq 8+4m(k+1)$

Here, c will win with a margin of $4n-4k-2$, which will be positive as long as $k < n$.

w will be the winner of $(\{c, w\} \cup B', V)$ if and only if $B' \subseteq B$ is a hitting set of size $\leq k$.

Proof ($\rightarrow$) The candidate w loses 4 points for each $b \in B'$ chosen, of which there are no more than k. c loses $2(k+1)$ points for each S_i hit. There will be m such sets if B' is a hitting set, so they lose $2m(k+1)$ points total.

Candidate	Score
c	$6m(k+1)+8n$
w	$\geq 6m(k+1)+8n+2$
$b \in B'$	$\leq 8+4m(k+1)$

w will end up with an advantage of at least 2 points and therefore w will be the winner of the election.

Proof (←) No $B' \subseteq B$ which is not a hitting set of size $\leq k$ will make w the winner by adding those candidates. If B' is a hitting set but $|B'| > k$, c will have $6m(k+1)+8n$ points and w will have $\leq 6m(k+1)+8n$. If B' is not a hitting set c will have $\geq 6m(k+1)+8n+2(k+1)$ points and w will have $\leq 6m(k+1)+8n+2k+2$ points. In either case w will not be the unique winner.

This construction can be used to create cases of constructive and destructive control by adding candidates and destructive control by adding candidates. $((C, V), (m-k), c)$ is such an instance of destructive control by deleting candidates. $((\{c, w\}, V), k, c)$ is an instance of destructive control by adding candidates, and $((\{c, w\}, V), k, w)$ is an instance of constructive control by adding candidates. □

As in Erdélyi et al. this reduction is not sufficient to show resistance to constructive control by deleting candidates as c and w are the only candidates with a shot at winning and deleting c will instantly make w the winner. The remaining case can be handled by the following reduction.

We will again reduce from an instance of hitting set $(B, \mathcal{S}, k)$.

The candidate set C will consist of $B \cup \{w\}$, the set B from the instance of hitting set together with an additional candidate.

The voter set V will be constructed as follows:

- $n+k$ voters have a score of 2 for b for every $b \in B$, and a score of 0 for w.

- $3 + 2mk$ voters have a score of 2 for w and a score of 0 for all other candidates.
- For each $S \in \mathcal{S}$, $4k + 1$ voters have a score of 2 for s for every $s \in S$, a score of 1 for each candidate in $B - S$, and a score of 0 for w.
- For each $S \in \mathcal{S}$, $4k+1$ voters have a score of 2 for b for every $b \in B-S$, a score of 2 for w, and a score of 1 for s for every $s \in S$
- For each $b \in B$, $2n - k$ voters have a score of 2 for b, a score of 1 for w, and a score of 0 for every other candidate.

w will be the winner of $(B' \cup \{w\}, V)$ if and only if B' is a hitting set over $\mathcal{S}$ of size no more than k or if $B' = \emptyset$.

Proof (→): Assume B' is a hitting set and $|B'| = k$ Each $b \in B'$ will receive $12mk + 4n - 2k + 4$ points. w will receive $8mk + 6 + 4mk + 4(n - k) + 2k = 12mk + 4n - 2k + 6$ points. Therefore w will be the winner in the election.

Proof (←): First assume B' is a hitting set but $|B'| = l > k$. Since B' is a hitting set, every $b \in B'$ will receive exactly $12mk + 4n - 2k + 4$ points. w will receive $4mk + 6 + 8mk + 4(n - l) + 2l = 12mk + 4n - 2l + 4$ points. $score(b) - score(w) = -2k + 2l - 2$ which is non-negative since $l > k$. Therefore w will lose the election.

Next consider the case where $|B'| = l \leq k$ but B' is not a hitting set. Therefore every $b \in B'$ will have a score $\geq 12nk + 4m - 2k + 4 + 4k$ as they will gain an extra $4k$ points from one set of group 3 voters. w will have the score $12nk + 4m - 2l + 6$ Therefore $score(b) - score(w) = 2k + 2l - 2$ which is non-negative and w will again lose the election.

An instance of hitting set $(B, \mathcal{S}, k)$ can thus be reduced to finding whether w can be made the winner of (C, V) as above by deleting $m - k$ candidates.
□

4.3.4 Destructive Control by Partition of Voters

Theorem 4.3.5. *2-NRV is resistant to destructive control by partition of voters in the ties-promote model.*

We will reduce from restricted hitting set. Restricted hitting set is an NP-complete hitting set variant introduced by Hemaspaandra et al. with additional restrictions on the sizes of the sets in an instance [24]. The version as used here has a slightly stronger bound which is necessary due to the somewhat larger numbers required in this proof.

Given: A collection $\mathcal{S}$ of subsets of a set B, $k \in \mathbf{Z}^+$, with $|\mathcal{S}| = m$, $|B| = n$, and the additional restriction that $m(k+1) + 3 \leq n - k$

Question: Does B contain a hitting set B' of size k or less that contains at least one element from every $S \in \mathcal{S}$?

Given an instance of restricted hitting set $(B, \mathcal{S}, k)$ with $|B| = n$ and $|\mathcal{S}| = m$, create a 2-normalized-range election with $C = B \cup \{w, c\}$ and V as follows.

- $2m(k+1) + 4n$ voters have a score of 2 for c, and a score of 0 for every other candidate.
- $3m(k+1) + 2k$ voters have a score of 2 for w, and a score of 0 for every other candidate.
- For each $b \in B$, 4 voters have a score of 2 for b, a score of 1 for w, and a score of 0 for every other candidate.
- For each $S_i \in \mathcal{S}$, $2(k+1)$ voters have a score of 2 for b, a score of 1 for c, and a score of 0 for every other candidate.
- For each $b \in B$, 1 voter has a score of 2 for b and a score of 0 for every other candidate.

c can be made to lose (C, V) through partition of voters if and only if there is a hitting set of size $\leq k$ over $\mathcal{S}$ in B.

Proof ($\rightarrow$) Given an appropriate hitting set B', partition V into sets V_1, V_2. Let V_1 contains a voter from the final group corresponding to every $b \in B'$ and one voter from the second group (allotting just 2 points to w) and let $V_2 = V - V_1$. After the initial subelections, we will be left with w, c, and the candidates B' corresponding to the hitting set, and w will win this election (see the reduction to adding/deleting candidates for the details of that proof).

Proof ($\leftarrow$) If there is no hitting set $B' \in B$ of size $\leq k$, c cannot be made to lose the election through partition of voters. For any actions attempting to control the election by forcing the final candidate set, see the previous reduction to adding/deleting candidates. As for other efforts concentrated at more typically partitioning the voters, among the initial candidates, c has as high of a score as any two other candidates, so they must at least tie in at least one of the subelections and so they will always make it to the final election. The scores of the candidates in the initial election follow.

Candidate	Score
c	$6m(k+1) + 8n$
w	$6m(k+1) + 4k + 4n$
b	$\leq 4m(k+1) + 10$

c's score minus the next two highest scores will therefore be at least $4(n-k) - 4m(k+1) - 10$. However, due to our use of restricted hitting set, we have that $m(k+1) + 3 \leq n - k$, and so this is at least 2. Therefore the only way to defeat c is to face them against a hitting set as described. $\square$

4.3.5 Partition of Candidates

Theorem 4.3.6. *4-NRV is resistant to constructive control by partition and runoff partition of candidates*

We will reduce from control by deletion of candidates in NRV. We will show the reduction to constructive control by run-off partition, though the other partition case quite similar.

Given an r-NRV election[1] (C, V) with $|C| = m$ and $|V| = n$ the distinguished candidate $w \in C$, and a deletion limit $k \in \mathbf{N}$, construct a $2r$-NRV election (C', V') as follows. $C' = C \cup \{a, b\}$, where a and b are additional auxiliary candidates. V' will consist of the original voter set V in addition to the following:

- For each $c \in C$, $2n$ voters have a score of $2r$ for c, a score of r for a, and a score of 0 for every other candidate.
- For each $c \in C^*$, $3nm$ voters have a score of $2r$ for c and a score of 0 for every other candidate.
- $2nm$ voters have a score of $2r$ for w, a score of r for a, and a score of 0 for every other candidate.
- nm voters have a score of $2r$ for w and a score of 0 for every other candidate.
- $(m - k - 1)n$ voters have a score of $2r$ for all $c \in C$ and a score of 0 for every other candidate.
- $2n + 1$ voters have a score of $2r$ for a and a score of 0 for every other candidate.

[1]Note that since k-NRV is resistant to deletion of candidates for $k \geq 2$, this reduction shows resistance for $k \geq 4$

- $3n+3nm+(m-k-1)n+2$ voters have a score of $2r$ for b and a score of 0 for every other candidate.

Let $s_0(c)$ be the score of candidate c among the original voters V. Note for any candidate $s_0(c) \leq nr$.

The following are the scores for the candidates in (C', V') and various relevant subelections thereof:

(C', V')

Candidate	Score
a	$4nmr + 4nr + 2r$
b	$6nr + 6nmr + 2(m-k-1)nr + 4r$
$c \in C^*$	$4nr + 6nmr + 2(m-k-1)nr + 2s_0(c)$
w	$4nr + 6nmr + 2(m-k-1)nr + 2s_0(w)$

The winner in this case will be b.

$(\{a, w\}, V')$

Candidate	Score
a	$4nr + 6nmr + 2$
w	$4nr + 6nmr + 2(m-k-1)nr + 2s_0(w)$

The winner of this election will be w.

$(\{w\} \cup D, V')$ **where** $D \subseteq C^*, |D| = l$

Candidate	Score
$c \in D$	$4nr + 6nmr + 2(m-k-1)nr + 2s_0(c)$
w	$4nr + 6nmr + 2(m-k-1)nr + 2s_0(w)$

The winner will again be whatever candidate is the winner over the original voter set V.

$(\{a,b\} \cup D, V')$ **where** $D \subseteq C^*$**,** $|D| = l$

Candidate	Score
a	$4nr + 6nmr + 2(m-l)nr + 2r$
b	$6nr + 6nmr + 2(m-k-1)nr + 4r$
$c \in D$	$4nr + 6nmr + 2(m-k-1)nr + 2s_0(c)$

In this election, a will be the winner whenever $l \leq k$. Otherwise the winner will be b.

$(\{b,w\}, V')$

Candidate	Score
b	$6nr + 6nmr + 2(m-k-1)nr + 4r$
w	$4nr + 6nmr + 2(m-k-1)nr + 2s_0(w)$

In this case, b is the clear winner.

$(\{a,w\} \cup D, V')$ **where** $D \subseteq C^*, |D| = l$

Candidate	Score
$c \in D$	$4nr + 6nmr + 2(m-k-1)nr + 2s_0(c)$
a	$4nr + 4nmr + 2(m-l)nr + 2$
w	$4nr + 6nmr + 2(m-k-1)nr + 2s_0(w)$

The winner will again be whatever candidate is the winner over the original voter set V.

w can be made the winner of (C, V) through deleting candidates if and only if they can be made the winner of (C', V') through partition or runoff partition of candidates.

Proof (→) Suppose w can be made the winner of (C, V) through deleting $\leq k$ candidates. Let D be the set of candidates which were deleted in the deletion problem. Partition the candidates into the subelections $(D \cup$

$\{a, b\}, V')$ and $(C - D, V')$. a will win the first subelection as shown above. w will win the second subelection, as it must if it is capable of winning with the candidates in D deleted. The final election will then come down to w and a, and as we see above, w will come out the victor. Alternately, in the non-runoff partition case, let the initial subelection be $(D\ cup\{a, b\}, V')$, which a will win. The final election will come down to $(\{a, w\} \cup C - D, V')$, which w will win.

Proof (←) Suppose w can be made the winner of the election (C', V') through control by runoff partition of candidates. It must be that this occurs through a partition of the form $(\{a.b\} \cup D, \{w\} \cup (C^* - D))$ with $D \subseteq C^*, |D| \leq k$. b will always beat w, so they cannot face each other in either the initial or final elections. The only candidate capable of beating b is a when not in an election with w and when accompanied by no more than k other candidates from C. w must also be able to defeat the remaining $m - k$ candidates. In the final election w will then face a and win, or in the partition case a and $C - D$. Consequently w can also be made the winner of (C, V) by deleting k candidates.

The preceding construction will shows that NRV is resistant to constructive cases of partition of candidates. However it is not sufficient for the destructive cases, as a winning candidate in the original election (C, V) will not actually win in (C', V'). Therefore we will present a new construction to handle the destructive cases.

Theorem 4.3.7. *2-NRV is resistant to destructive control by partition and runoff partition of candidates.*

We can reduce the Hitting set problem to the problem of destructive control by partition of candidates. Let $(B, \mathcal{S}, k)$ be an instance of hitting set where $B = \{b_1, b_2, \ldots, b_n\}$, $\mathcal{S} = \{S_1, S_2, \ldots, S_m\}, S_i \subseteq B$, and $k \in \mathbf{Z}^+, k \leq n$.

We will construct a 2-range election based on this instance. The candidate set C will consist of $B \cup \{w\}$. The voter set V will be as follows.

- For each $S \in \mathcal{S}$, $4(k+1)$ voters have a score of 2 for each $b \in S$ and a score 1 for w
- For each $S \in \mathcal{S}$, $4(k+1)$ voters have a score of 2 for each $b \in B, b \notin S$, and a score of 0 for every other candidate.
- For each $b \in B$, 4 voters have a score of 2 for b, a score of 1 for each $b' \in B, b' \neq b$, and a score of 0 for w
- $2(k+1)m + 4n - 2k + 1$ voters have a score 2 for w and a score of 0 for every other candidate.

Again, we will first consider the outcome of several forms of subelections of this election.

(C, V)

Candidate	Score
w	$8(k+1)m + 8n - 4k + 2$
$b \in B$	$8(k+1)m + 4n + 4$

w will win this election for any $k \leq n$.

$(\{w\} \cup D, V)$**,** D **is a hitting set,** $|D| = l$

Candidate	Score
w	$8(k+1)m + 8n - 4k + 2$
$b \in B$	$8(k+1)m + 8n - 4l + 4$

If $l \leq k$, every $b \in B$ will tie for first with w as the clear loser. Otherwise w will be the winner.

$(\{w\} \cup D, V)$**, D is not a hitting set, $|D| = l$**

Candidate	Score
w	$8(k+1)m + 8n - 4k + 2 + 4(k+1)$
$b \in B$	$8(k+1)m + 8n - 4l + 4$

w will win this election.

There is a hitting set $B' \subset B$ where $|B'| \leq k$ if and only if w can be made to lose the election through partition or run-off partition of candidates.

Proof ($\rightarrow$) If there is a hitting set $B' \subset B$ of size $\leq k$, w will can be made to lose the election through control by partition or runoff partition of candidates with the partitions $\{w\} \cup B', B - B'$ w will lose the initial subelection $(\{w\} \cup B')$ as shown above, and will therefore lose the entire election.

Proof ($\leftarrow$) If there is no such hitting set, w cannot be made to lose the election through partition or runoff partition of candidates. As shown above, w will win any election $(\{w\} \cup B', V)$ where $|B'| > k$ opponents, or where their opponents do not comprise a hitting set. The one special case is when $k = n$, where w will lose the original election but in this case there is a hitting set by default, so this problem should also always accept. Therefore w will win against any subset of B, so they will win both any initial subelection and the final election, and so there is no partition to make them lose. □

4.3.6 Missing Case

One case of control for NRV, destructive control by partition of voters in the ties-eliminate model, remains open. The polynomial time algorithm previously described for this case of control for RV will not work, as a candidate can potentially lose an election a different way than that algorithm can find. Still, even with this case open NRV has at least as many control resistances as any other natural system.

4.4 Manipulation

RV is vulnerable to manipulation. The simple and transparent scoring rule in range voting makes the manipulation problem very easy. To make their preferred candidate the winner, the set of manipulators can do no better than to give that candidate the maximum score possible and to give every other candidate the minimum score possible (0). If this is enough to make the distinguished candidate win it is possible and otherwise not. Since manipulation does not involve changes to the candidate set, the same algorithm will work for NRV as well as RV. Thus manipulation is clearly in $\mathbf{P}$ for both RV and NRV. $\square$

4.5 Bribery

RV is resistant to bribery. This result trivially follows from approval voting, which is already known to hold this resistance [15]. As before, we can just reduce from an instance of approval bribery to a 1-range or 1-normalized-range election which is exactly equivalent to the approval election, and so RV and NRV hold this resistance as well. $\square$

Chapter 5

Future Work

This work leaves open a number of questions. Clearly, it is desirable to resolve the one missing case of NRV and find whether it is the uniquely most resistant natural voting system to control. Regardless, NRV still does come short of resistance to all cases of control, so some other natural system could still best it. Just as useful would be results about the conditions that are required for a voting system to hold various resistances. It may still be that natural systems are incapable of holding every control resistance simultaneously. Any useful results here would first require a formalization of what exactly a natural voting system is. Most desirable would be a reasonable set of conditions which could be shown to be incompatible with holding all resistances simultaneously, à la Arrow's Theorem.

Other useful work would be to analyze methods for sidestepping the worst-case difficulty of the control problems here. For instance, the use approximation algorithms as by Baumeister et al. [7], or analysis of the problems with a restricted preference model [18].

Chapter 6

Acknowledgments

This work has been supported in part by NSF grant IIS-0713061.

I would also like to thank Edith Hemaspaandra and Andrew Lin for their assistance in preparing this thesis.

Bibliography

[1] K. Arrow. A difficulty in the concept of social welfare. *J. Political Economy*, 58(4):328–346, 1950.

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

[3] J. Bartholdi, C. Tovey, and M. Trick. How hard is it to control an election. In *Mathematical and Computer Modeling*, pages 27–40, 1992.

[4] D. Baumeister, G Erdélyi, E. Hemaspaandra, L. Hemaspaandra, and J. Rothe. Computational aspects of approval voting.

[5] S. Brams and P. Fishburn. *Approval Voting*. Springer, 2007.

[6] S. Brams and M. Sanver. Critical strategies under approval voting: Who gets ruled in and ruled out. *Electoral Studies*, 25(2):287–305, 2006.

[7] E. Brelsford, P. Faliszewski, E. Hemaspaandra, H. Schnoor, and I. Schnoor. Approximability of manipulating elections. In D. Fox and Carla P. Gomes, editors, *AAAI*, pages 44–49. AAAI Press, 2008.

[8] I. Caragiannis, J. Covey, M. Feldman, C. Homan, C. Kaklamanis, N. Karanikolas, A. Procaccia, and J. Rosenschein. On the approximability of dodgson and young elections. In *SODA '09: Proceedings of the Nineteenth Annual ACM -SIAM Symposium on Discrete Algorithms*,

pages 1058–1067, Philadelphia, PA, USA, 2009. Society for Industrial and Applied Mathematics.

[9] S. Cook. The complexity of theorem-proving procedures. In *STOC '71: Proceedings of the third annual ACM symposium on Theory of computing*, pages 151–158, New York, NY, USA, 1971. ACM.

[10] A. Copeland. A "reasonable" social welfare function. Seminar on Mathematics in Social Sciences, University of Michigan, 1951.

[11] G. Doron and R. Kronck. Single transferable vote: an example of a perverse social choice function. *American Journal of Political Science*, 21(2):303–311, 1977.

[12] J. Duggan and T. Schwartz. *Strategic manipulability is inescapable: Gibbard-Satterthwaite without resoluteness*. Division of the Humanities and Social Sciences, California Institute of Technology, 1992.

[13] G. Erdélyi, M. Nowak, and J. Rothe. Sincere-strategy preference-based approval voting broadly resists control. In *MFCS '08: Proceedings of the 33rd international symposium on Mathematical Foundations of Computer Science*, pages 311–322, Berlin, Heidelberg, 2008. Springer-Verlag.

[14] FairVote. The irv america program. `http://www.fairvote.org/irv/?page=20`.

[15] P. Faliszewski. *Manipulation of Elections: Algorithms and Infeasibility Results*. PhD thesis, University of Rochester, 2008.

[16] P. Faliszewski, E. Hemaspaandra, and L. Hemaspaandra. The complexity of bribery in elections. In *AAAI*. AAAI Press, 2006.

[17] P. Faliszewski, E. Hemaspaandra, L. Hemaspaandra, and J. Rothe. Llull and copeland voting computationally resist bribery and constructive control. *Journal of AI Research*, 35:275–341, 2009.

[18] P. Faliszewski, E. Hemaspaandra, L. Hemaspaandra, and J. Rothe. The shield that never was: societies with single-peaked preferences are more open to manipulation and control. In *TARK '09: Proceedings of the 12th Conference on Theoretical Aspects of Rationality and Knowledge*, pages 118–127, New York, NY, USA, 2009. ACM.

[19] L. Fortnow. The status of the P versus NP problem. *Commun. ACM*, 52(9):78–86, 2009.

[20] E. Friedgut, G. Kalai, and N. Nisan. Elections can be manipulated often. In *FOCS '08: Proceedings of the 2008 49th Annual IEEE Symposium on Foundations of Computer Science*, pages 243–249, Washington, DC, USA, 2008. IEEE Computer Society.

[21] M. Garey and D. Johnson. *Computers and Intractability: A Guide to the Theory of NP-Completeness*. W. H. Freeman, 1979.

[22] A. Gibbard. Manipulation of voting schemes: A general result. *Econometrica*, 41(4):587–601, 1973.

[23] E. Hemaspaandra, L. Hemaspaandra, and J. Rothe. Hybrid elections broaden complexity-theoretic resistance to control, 2006.

[24] E. Hemaspaandra, L. Hemaspaandra, and J. Rothe. Anyone but him: The complexity of precluding an alternative. *Artif. Intell.*, 171(5-6):255–285, 2007.

[25] L. Hemaspaandra. Sigact news complexity theory column 36. *SIGACT News*, 33(2):34–47, 2002.

[26] C. Hillinger. The case for utilitarian voting. Discussion Papers in Economics 653, University of Munich, Department of Economics, May 2005.

[27] M. Satterthwaite. Strategy-proofness and Arrow's conditions: Existence and correspondence theorems for voting procedures and social welfare functions. *Journal of Economic Theory*, 10(2):187 – 217, 1975.

[28] W. Smith. Range voting. `http://www.math.temple.edu/~wds/homepage/rangevote.pdf`, November 2000.