Chapter 1 Introduction

1.1 Motivation

This dissertation focuses on voting as a means of preference aggregation. Specifically, empirically testing various properties of voting rules and theoretically analyzing how much information it takes to make tampering with an election computationally hard.

Democratic societies and fair minded individuals have used voting procedures to come to group decisions since at least 508 B.C.E. [107], and, since at least 105 A.D., there have been concerns over the honesty and security of voting [92]. The disconcerting and ever present threat is that one person within the group, a coalition within the group, some outside actor(s), or the individuals counting the ballots, would, could, or do express undue influence on the result of the vote. These threats are so significant that many laws and methods of voting have been devised to intentionally dissuade individuals from attempting to tamper with a vote.

A prime example of a society constructing a voting procedure to prevent tampering is the selection process for the Doge of Venice. The city-state of Venice (in present day Italy) began selecting its leaders in the following, somewhat convoluted, manner in 1172 A.D. The procedure remained mostly unchanged for over 600 years (about 75 iterations), until the fall of the Venetian Republic in 1797. The election proceeds in 10 rounds over the course of several days. Each round created a college, either by lottery or by election, for the next round. In the first round, every member of the Great Council age 30 or more (and only one member per family) convened in a college, and 30 of them were selected by lottery for the next round. Round 2 saw these 30 reduced to 9, as selected by lottery. In the third round, the 9 elected 40 for representation in the next round and each of the 40 had to be approved by at least 7 of the 9 members. The fourth round saw the college of 40 narrowed to 12 by lottery draw. The fifth round had the 12 elect a college of 25, each requiring 9 of the 12 votes. In the sixth round the 25 was reduced to 9 by lottery and the seventh round had these 9 elect a college of 45, each by a 7 of 9 majority. In the eighth round the 45 were again pared down to 11 by lottery. The 11, in the ninth round, elected a final college of 41 with each member of the college requiring a 9 of 11 majority. The tenth and final college of 41, with a majority vote of at least 25 of the 41, elected the Doge of Venice [21].

While this procedure seems insane by modern standards, a comprehensive study of its security properties by Mowbray and Gollmann show that it was extremely robust to tampering [94]. It turns out that, with so many lotteries and rounds, it becomes very hard to manipulate or bribe voters; one never knows who exactly will be voting in the next round. This alternating of the type of selection round provides representation opportunities to minority candidates (egalitarianism) while still ensuring that more popular candidates (majoritarian) have a higher probability of winning.

This dissertation falls in the area of computational social choice (ComSoc), an emerging and rapidly evolving subfield of artificial intelligence. My work in ComSoc is focused on how groups of agents make collective decisions. social choice, an established research field at the intersection of mathematics and political science, has long studied the implications of group decisions in human systems.

The growth of multi-agent systems in computer science has created many situations where individual agents, be they robotic, software, or human, need to come together to make a group decision. These decisions can take the form of a recommendation on a shopping website, rankings of search results from the web, or coordinated robot behavior. This growth and proliferation of multi-agent systems research in the AI community had made it necessary to closely investigate how agents can work together and make group decisions [119]. ComSoc and social choice are related by two main bridges: bringing a computational perspective to decision systems already in use and/or studied by social choice, and bringing systems and processes developed through years of social choice research to bear on multi-agent systems. Social choice has broad application in computer science including: multi-agent systems, intelligent systems, and human-computer interaction [23].

This dissertation contains both theoretical and empirical work. The main focus of the theoretical work is on questions of manipulation and bribery in social choice settings. The study of manipulation in social choice is about security. The central question of this dissertation is: how much information does it take to make tampering with an election computationally hard? To this end, I investigate the bribery and manipulation problems under two information assumptions: uncertain information and structured information. Voting rules are subject to multiple forms of attack, and the classical and most current literature studies these issues in a perfect information world: every agent knows exactly how every other agent will vote. I feel that this model is lacking since manipulation is trivially easy for many common voting rules under perfect information. Typically, agents have uncertain or probabilistic information. Pollsters have an idea of how you will vote, just as we have an expectation of what our friends will want to eat for dinner. It turns out that the question of security is much different when we take uncertain information into account. Another way to frame manipulation is in terms of resource allocation. Consider the process of electing or gaining support for a particular alternative. If we have some resources to achieve a goal, canvassers in elections or concessions with friends about where to eat next week, then how can we best distribute our influence or resource in order to achieve consensus?

1.2 Main Contributions and Related Publications

This dissertation is supported by several publications with partially disjoint groups of researchers. Most of the chapters detail work that was done jointly and, therefore, I use "we" when describing this work. While I have had a principle role in defining, performing, and writing the work detailed in these chapters, I would not have been able to do anything (and wouldn't be writing this dissertation) if it wasn't for the help of my coauthors. I am deeply indebted to my coauthors on work that has directly supported this dissertation: Daniel

Binkele-Raible, Gábor Erdélyi, Henning Fernau, Judy Goldsmith, Andrew Klapper, Maria Silvia Pini, Francesca Rossi, Jörg Rothe, and K. Brent Venable.

My work has been directly supported by NSF-EAGER grant CCF-1049360: Changing Minds, Changing Probabilities, NSF-ITR grant 0325063: Decision-Theoretic Planning with Constraints, NSF IIS-1107011: IJCAI 2011 Doctoral Consortium and International Experience, and the 2010 Northern Kentucky Alumni Association Fellowship. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the National Science Foundation or the Northern Kentucky Alumni Association.

Bribery and Manipulation with Uncertain Information

In the work on bribery under uncertain information we proposed a novel and flexible model to represent majority voting with uncertain information. Given a set of voters, their individual preferences represented as probability distributions over a set of issues; their prices for changing their preferences; and a budget, we classified the complexity of finding efficient bribery schemes. We showed that, depending on the particular combination of evaluation and bribery models chosen, the problems range in complexity from polynomial time to NP-complete. This difference reveals that modeling choices can have significant effects on the complexity of calculating efficient bribery schemes. This work has resulted in two joint publications.

- Daniel Binkele-Raible,Gábor Erdélyi, Henning Fernau, Judy Goldsmith, Nicholas Mattei and Jörg Rothe, "The Complexity of Probabilistic Lobbying." Technical Report arXiv:0906.4431v3 [cs.CC], *ACM Computing Research Repository (CoRR)*, June 2009. Revised, February 2011.

- Gábor Erdélyi, Henning Fernau, Judy Goldsmith, Nicholas Mattei, Daniel Raible and Jörg Rothe, "The Complexity of Probabilistic Lobbying." *Proc. 1st International Conference on Algorithmic Decision Theory (ADT-09)*, October, 2009.

In addition to the work on multiple referenda, sports tournaments represent a domain where it is natural to express winners and losers in terms of probabilities of outcomes. We use this as a motivating example to study the related problems of manipulation in elimination and round-robin tournaments. These results can also be mapped to their corresponding voting rules. Given a set of teams and their probabilities of each possible win, along with prices for decreasing their competitive output (purposefully losing or underperforming in a match), we classified the complexity of finding efficient bribery schemes for three common types of sports tournaments. The evaluation complexity of these problems range over a variety of complexity classes from the easy to the very hard. Our results show that in some cases the added uncertainty increases the complexity of manipulating sports tournaments while in other cases it does not. While this increase in complexity is not uniform across all tournament types, the change shows strong evidence that reasoning in domains with uncertain data leads to an increase in reasoning complexity. This work has resulted in one joint conference publication.

- Nicholas Mattei, Judy Goldsmith, and Andrew Klapper, "On the Complexity of Bribery and Manipulation in Tournaments with Uncertain Information." *Proc. 25th Intl. Florida Artificial Intelligence Research Society Conference (FLAIRS 2012)*, June 2012.

Bribery and Manipulation in Combinatorial Domains

When looking at voting it is often natural to express group decision problems as the combination of a sequence of decisions. This method is used in many settings, from the United States Congress (specifically, votes for amendments to a bill) to a group of friends deciding what appetizer, main course, desert, and wine should be served for a group meal. In all these cases, agents express preferences and vote on parts of the overall decision to be taken. Agents may also have dependent preferences within this construction: the choice of wine may depend on the choice of main course for the meal. We consider a scenario where agents use the CP-net formalism ("Conditional Preference" or *"Ceteris Paribus"* networks) which allows us to compactly model these conditional dependencies [15]. We investigated the computational complexity of bribery and manipulation schemes in combinatorial voting domains where voters' preferences are expressed as CP-nets. To do this, we generalized the traditional bribery problem to encompass these domains and found that, for most of the combinations of these parameters, bribery in this domain is computationally easy. This indicates that either CP-net preferences lead to highly manipulable aggregation schemes, or that we have over-constrained the problem. As CP-nets have become more ingrained in preference and social choice research we hope to continue this line of research into the security of CP-nets. This work was supported by one conference publication and one invited paper.

- Nicholas Mattei, Maria Silvia Pini, Francesca Rossi, K. Brent Venable, "Bribery in Voting Over Combinatorial Domains Is Easy." *Proc. 11th International Conference on Autonomous Agents and Multiagent Systems (AAMAS-12, short paper)*, June 2012.

- Nicholas Mattei, Maria Silvia Pini, Francesca Rossi, K. Brent Venable, "Bribery in Voting Over Combinatorial Domains Is Easy." *12th International Symposium on Artificial Intelligence and Mathematics (ISAIM-12), Special Session on Computational Social Choice*, January 2012.

Empirical Analysis of Voting Rules and Election Paradoxes

In addition to the theoretical work, I investigated the behavior of voting rules and occurrences of voting paradoxes using empirical data. To facilitate an empirical study, I mined a large dataset of "elections" from real preference data. This dataset has several million individual elections with tens to tens of thousands of voters. This represents orders of magnitude more election data than previously available, and I used it to analyze the behavior of voting rules. Analysis of this dataset has provided useful insight into voting methods – including the surprising conclusion that, in contrast to much of the theoretical work, voting rules declare the same winner a majority of the time. I recently began separate research collaborations with colleagues in political science and psychology in order to more extensively study the data from the Netflix Prize. This work is supported by one conference publication.

- Nicholas Mattei, "Empirical Evaluation of Voting Rules with Strictly Ordered Preference Data" *Proc. 2nd International Conference on Algorithmic Decision Theory (ADT-11)*, October, 2011.

1.3 Structure of the Dissertation

This dissertation attempts to walk the reader through the study of how access to different types of information changes the reasoning complexity of various security problems in social choice systems and how we can test these computational properties.

Chapter 2, Preliminaries: In Section 2.1 I survey the mathematical and algorithmic fundamentals that one would need to understand the discussion in later chapters. This section is focused for readers not from computer science or mathematics. I expect the reader of this section to be a mathematically competent individual from another discipline.

In Section 2.2 I provide a comprehensive background of social choice and preference aggregation. I give an overview of the field of Computational Social Choice; survey research in voting systems including bribery and manipulation; and detail the specifics of several voting rules. I frame the work presented in this dissertation in the overall stream of research about bribery and manipulation in ComSoc.

Chapter 3, Bribery and Manipulation with Uncertain Information: In Section 3.1

I detail the work on bribery in majority elections when the outside actor has access to uncertain information. This section details the development of a new model of reasoning in voting domains with uncertain information and presents a complete analysis of the complexity of reasoning in the uncertain information setting.

In Section 3.2 I detail the work on bribery in sports tournaments when the outside actor has access to uncertain information. While the discussion deals primarily with sports tournaments, the model can be extended to certain voting rules which have the same structure as sports tournaments. This section details the extension of the model presented in Section 3.1 to the domain of sports tournaments and presents a complexity analysis of reasoning in this extended model.

Chapter 4, Bribery and Manipulation in Combinatorial Domains: In this chapter I detail the work on bribery in elections where each decision is described by the combination of a set of smaller decisions. I present an overview of CP-nets and their use in structured preference representation. I extend the traditional bribery problem and apply it to combinatorial domains. I then present complete complexity analysis of these new models of reasoning.

Chapter 5, Empirical Analysis of Voting Rules and Election Paradoxes: In this chapter I detail the work on empirically verifying some of the underlying assumptions in social choice research. I survey existing data sets and their properties and then identify and study a novel set of data that approximates real data from real elections. This chapter looks at why so few empirical studies occur in ComSoc and begins to close the gap between the theoretical and empirical.

Chapter 6, Conclusions and Future Directions: In the final chapter we look back at the research presented in this dissertation and attempt to gain a better perspective on the role that the type of information has in reasoning about election systems. We also detail directions future research that will help us better understand how to protect the way we choose.