5-1-2010

Collective belief models for representing consensus and divergence in communities of Bayesian decision-makers

Kshanti Greene

Kshanti A. Greene
Candidate

Computer Science
Department

This dissertation is approved, and it is acceptable in quality and form for publication:

Approved by the Dissertation Committee:

______________________________, Chairperson

____

____

____

____

____

____

____

____

Collective belief models for representing consensus and divergence in communities of Bayesian decision-makers

by

Kshanti A. Greene

M.S. Computer Science, University of New Mexico, 2004
B.S. Computer Science, University of Pittsburgh, 1997

DISSERTATION

Submitted in Partial Fulfillment of the
Requirements for the Degree of

Doctor of Philosophy
Computer Science

The University of New Mexico

Albuquerque, New Mexico

May, 2010

Dedication

For Thomas, for joining me on this roller-coaster ride and for his endless support.

Acknowledgments

Many people have contributed to the ideas and inspiration that resulted in this dissertation. I would first like to thank my advisor, Dr. George Luger, who gave me just the right balance of guidance and encouragement to keep me focused and freedom to allow me to explore new areas. Dr. Joe Kniss has been a significant force in the direction of this work, and helped to spark my research into game theory. I also would like to thank him for having high expectations and for always encouraging me to do my best work. My committee, including Drs. Luger and Kniss as well as Drs. Melanie Moses, Tim Ross and Carl Stern, has provided excellent direction. Their complementary expertise and interests have formed a strong foundation for my work.

Three individuals with a shared vision for social decision-making and collective problem-solving have had a significant influence on my research. Dr. Antonio Sanfilippo at the Pacific Northwest National Laboratories introduced me to a community of researchers through his workshops in technosocial predictive analytics. Dr. Pietro Michelucci at DARPA has opened up endless possibilities to continue this research with his Massively Distributed Problem Solving program. Finally, Dr. Carl Stern's passion for democracy has helped me envision the long-term potential my research. I would also like to thank Dr. Stern and Ken and Kay Blemel at Management Sciences Inc. for their flexibility while I finished my dissertation.

I acknowledge Dr. Mohammed El-Genk at UNM for his financial support in my early graduate work through the NASA Space Grant, and the Kraft family for their support through the Kraft Fellowship. Also thanks to the sponsors of the 2009 AAAI Doctoral Consortium and my mentor Dr. Sara Sood for her insightful guidance.

I would also like to thank fellow graduate students in Dr. Luger's "Loopy Group" for their ideas and interesting intellectual discussions. In particular, Roshan Rammohan, Seth Moore, Yamel Torres Rodriguez, Chayan Chakrabarti, Ben Gordon and Dan Pless have all contributed in some way to this work.

Finally, I would like to thank my family for their support and encouragement, in particular my mom Deborah Spurlock, who has always believed in me, and my father Jim Auster. My grandfather, Dr. Harry Greene, formerly of Yale University, has inspired me to pursue a PhD. I would also like to thank Lily for being an excellent four-legged companion and for helping me with the hard math.

Collective belief models for representing consensus and divergence in communities of Bayesian decision-makers

by

Kshanti A. Greene

ABSTRACT OF DISSERTATION

Submitted in Partial Fulfillment of the
Requirements for the Degree of

Doctor of Philosophy
Computer Science

The University of New Mexico

Albuquerque, New Mexico

May, 2010

Collective belief models for representing consensus and divergence in communities of Bayesian decision-makers

by

Kshanti A. Greene

M.S. Computer Science, University of New Mexico, 2004
B.S. Computer Science, University of Pittsburgh, 1997

Ph.D., Computer Science, University of New Mexico, 2010

Abstract

Bayesian belief aggregation is the process of forming a consensus model from the probabilistic beliefs of multiple individuals. Preference aggregation attempts to find an optimal solution for a population considering each individual's beliefs, desires and objectives. Belief and preference aggregation approaches that form a single consensus average away any diversity in a population. In the process they may fail to uphold a set of mathematical properties for rational aggregation defined by social choice theorists. This dissertation introduces a new aggregation approach that maintains the diversity of a population and allows the competitive aspects of a situation to emerge, enabling game theoretic analysis in large populations of decision-makers. Each individual's beliefs and preferences are represented by a Bayesian network. Based on the result of inference on the networks, a population is separated into *collectives* whose members agree on the relatively likelihood or desirability of the possible outcomes of a situation. An aggregate for each collective can then be computed such that the aggregate upholds the rationality properties. Game theoretic analysis is then applied using "super-agents" that represent each collective as the game players. In this manner, the set of Pareto optimal and Nash equilibrium solutions can be found, even in situations that cause single consensus models to return non-Pareto or otherwise "irrational" solutions.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Any community or organization that is interested in making decisions based on the opinions of its members would benefit from the ability to form computational decision models from the combined input of a potentially diverse population. A number of research areas have addressed this goal. *Social choice theory*, which combines economics, voting and statistical theories, has investigated the mathematical behavior of combining preferences to form a single consensus, or *social choice* [1, 19, 56, 31, 5]. *Bayesian belief aggregation* is the process of combining probability estimates from multiple individuals to form a single Bayesian belief model representing all individuals' contributions [52, 40, 48, 41, 53]. *Collaborative filtering* attempts to target information to individuals based on the opinions of similar individuals [24]. *Dempster-Shafer theory* extends probability theory with the concept of plausibility and uses *combining rules* to aggregate the belief estimates of multiple experts [58]. Most of these techniques aim to find a single "consensus" model or solution that combines everyone's beliefs and preferences, no matter how divergent those beliefs may be.

The research described in this dissertation crosses boundaries between computer science and the social sciences. The research is relevant to social choice theory, Bayesian reasoning, and other research areas that are confronted with the task of combining divergent information or conflicting opinions. The potential of applying the approach to policy-making and political science is demonstrated through examples.

1.1 Motivation

The goal of this research is to combine the diverse beliefs and preferences of many individuals to form graphical models for democratic decision-making. Computational models for social decision-making have the potential to change the policy and decision-making paradigm in a society. For example, there is promise in leveraging the rapid spread of information in popular social-networking tools to elicit opinions that form computational models for collective decision-making. These collective intelligence models can communicate the diverse ideas, beliefs and preferences of individual stakeholders to decision-makers so that community and business representatives can form policy that best serves their constituency. In this manner, these models will enable individuals to have a direct influence in the social, economic, and political decisions that affect them. In addition, they can enable people to visualize how their actions affect their own circumstances and their environment, taking into consideration the simultaneous actions and goals of other community members. Finally, models of collective belief will help increase understanding of human behavior and the origin of personal and public opinion.

Existing motivations for forming a consensus model from multiple individuals fit into two categories. First, one may be interested in building the most *accurate* model of a domain or situation of interest from a set of experts or sensors with heterogeneous but potentially overlapping specializations. This situation would be typical of an expert system or data fusion [28, 62], which combines the output from a multiple input streams to improve situation understanding. In this case the goal of the aggregation method may be to maximize the quality and reliability of the information it receives and try to resolve any conflicts. Information from reliable experts (or agents) may be given higher weight while contributions from less reliable sources may be discounted.

Second, one may be interested in combining the opinions of the individuals to form a solution that is *representative* of the potentially divergent beliefs of contributors. In this case, each person's opinion is considered equivalent, and no judgement about the accuracy or quality of provided information is made. Instead, one may be interested in how accurately the model *represents* the beliefs of the population. In other words, have all the opinions been given equal consideration when forming the conclusion and does the conclusion distinguish the significant opinions of the population? Even when some judgement of quality is made on beliefs, one may be interested in capturing a consensus model that maintains the diversity of the elicited and inferred beliefs. This motivation has been addressed in social choice theory and collaborative filtering techniques. Social choice theory is interested in finding one solution for everyone, whereas collaborative filtering is interested in finding a personalized solution for each individual.

The approach discussed in this dissertation combines the motivations of social choice theory with the techniques of Bayesian networks. My goal is to build representative models of a population similar to social choice theory. However, instead of considering opinions in isolation, entire models are built that incorporate the reasoning behind each contributor's beliefs. In particular, the decision models in the described approach consider the uncertainties, context, and other factors that influence a decision. These models will enable one to make predictions and decisions based on the state of the environment and will help decision-makers understand the factors that contribute to the outcomes.

1.2 Challenges

The research discussed in this dissertation forms a mathematically sound approach towards achieving *social* decision-making. Before models can be developed that societies can use to make decisions and form policy, several theoretical, practical and philosophical issues must be overcome that have inhibited progress in belief and preference aggregation. Research in areas such as social choice theory and Bayesian belief aggregation reached their peak in previous decades, partly because a number of "impossibility" theorems exposed the limitations of finding a consensus model. This dissertation revisits these theories because the advent of social networking tools and techniques such as *crowdsourcing* [8] and *folksonomy* [29] enable many people to contribute to collaboratively solve a problem or classify information, yet these techniques still lack consistent, theoretically sound methods to form consensus models from diverse input.

Social choice and Bayesian theorists have stated that it is not possible to combine, or *aggregate* arbitrary beliefs or preferences to build a consensus model that conforms to a set of mathematical principles of rationality for shared preference[1, 19, 56, 52]. Voting theory, the study of combining votes or preferences to select an outcome, has found that combining preferences using a rank order of the options can result in irrational behavior. In particular, the economist Kenneth Arrow developed a theorem that states that there is no mathematically sound way to aggregate a number of arbitrary votes when there are three or more options to choose from [1]. In summary, the social choice function breaks a transitive assumption that states if A is preferred to B by a majority of voters, and B is preferred to C by a majority, then A should be preferred to C. A simple example shows that transitivity cannot hold unless there is a dictatorship (one person's vote is the rule).

In Bayesian reasoning, *probabilities* are used to represent an individual's (or group's) belief in the likelihood of an event given the factors that influence the event, while *utilities* indicate the value of an outcome given the inherent uncertainty. Early Bayesian theorists showed that aggregating probabilistic beliefs and utilities can result in a *social choice* solution that is preferred by no one [31, 56]. In other words, the solution is not *Pareto optimal*. A solution is Pareto optimal if there is no other solution that provides a higher utility for an individual without another individual having a lower utility [10]. A decade later, researchers in Bayesian networks showed that opinion pool functions, used to aggregate probability distributions, fail under relatively mild Bayesian assumptions [52]. They show that even when all contributors agree on the structure of the network, that the aggregation methods create dependencies between random variables that did not exist in the original network.

Complexity is always an issue when dealing with Bayesian networks and graphical models. Using a well-known exact inference algorithm called *variable elimination* on a single network of n nodes, inference is exponential in the network's induced width $w*(d)$ given a processing order d of the nodes [15]. If one is interested in aggregating many individuals, one must also consider the aysmptotic behavior of combining multiple probability distributions to form consensus models. In the worst case, one would need to run inference on each individual's network, and then aggregate the results.

In addition to the theoretical limitations for Bayesian belief aggregation, existing aggregation techniques do not capture the diversity of beliefs, motivations and preferences that can be found in a representative sampling of a population. These approaches attempt to force a cooperative solution by forming one consensus that "averages away" the differences. However, many decision problems have individuals and groups with opposing goals, therefore forcing consensus does not accurately represent the situation. As opinions become more divergent, the average becomes less representative of the original beliefs, and information loss increases.

1.3 Objectives

The following objectives summarize my hypotheses and the primary goals of the research endeavors described in this dissertation.

1.3.1 Separation of Beliefs Into Representative Groups

If one were to elicit beliefs and preferences about a topic of interest from an arbitrary population of individuals, it is likely that the beliefs will be divergent. The described approach discovers individuals that share similar beliefs and groups them accordingly. To this end, I expected that:

- In some situations, groups will have consensuses that are in opposition to other groups and that these opposing beliefs will not be accurately represented by the single consensus approach
- The outcomes for some groups may be in direct opposition to the optimal outcome found using a single consensus approach. In other words, the "optimal" result is the worst case scenario for some individuals
- There will be situations in which the single consensus approach can generate a "social choice" solution that is not preferred by any of the individuals in a group
- The developed approach will group individuals in such a way that their group consensus more accurately reflects the beliefs of the group members than the single consensus approach. In other words, the new models will be more *representative* of a population than the single consensus models

This objective is discussed further in Chapter 5.

1.3.2 Overcome Theoretical Limitations of Belief and Preference Aggregation

Several impossibility theorems, discussed in Chapter 3, describe the limitations of using belief and preference aggregation to form a single consensus model. The objective is to partition a population into subgroups such that the aggregate of each of the subgroups will uphold the mathematical principles of rationality defined by social choice theorists [1, 19, 56, 31]. This will require defining the extent to which individuals in a subgroup must agree such that their aggregate upholds rational behavior. This objective is discussed in Sections 4.3 and 5.5.

1.3.3 Enable decision and game theoretic analysis

Situations in which a population has opposing goals are not well represented using a single consensus approach. An approach that allows the competition within the population to emerge may represent these situations better. The objective is to enable game theoretic analysis to be applied to the set of solutions formed from groups of individuals that meet objective 1.3.2. The set of Pareto optimal solutions is first extracted from the set of solutions, even in situations for which the single consensus approach cannot. When appropriate, for instance in strategic situations, additional game theoretic analysis may be applied, including Nash equilibrium and minimax solutions that cannot be done using a single consensus solution. This objective is discussed in Chapter 6.

1.3.4 Efficient discovery of collectives

An algorithm that automatically discovers the collectives within a population based on the population's beliefs and preferences will be developed. The algorithm will involve three steps: inference on Bayesian belief networks, identification of the collectives and aggregation of the collective beliefs. A goal is to reduce the asymptotic complexity of a brute force, exact algorithm, while maintaining the integrity of the collectives. An understanding of the issues involved in designing algorithms for collective discovery and belief aggregation will be applicable to engineering effective algorithms. The successes and challenges relating to these objectives are discussed in Section 7.

1.4 Summary of Contributions

This dissertation describes a belief aggregation approach, called *collective belief aggregation*, that combines the probabilistic beliefs of many individuals using Bayesian decision networks to form *collectives* such that the aggregate of each collective, or *collective belief* has rational behavior according to a set of properties defined by social choice theorists. Game theoretic analysis is then applied to the set of collectives to enable social decision-making. This dissertation will:

1. Show that partitioning a population into subgroups of individuals with similar beliefs results in belief aggregation that reduces information loss and creates a more representative consensus model (Chapter 5).

2. Define characteristics for subgroups that will enable rationally consistent belief and preference aggregation according to mathematical principles of rationality defined by social choice theorists (Sections 4.3 and Chapter 5).

3. Demonstrate how game theoretic analysis can be applied to a set of consensus models, including finding the set of Pareto optimal solutions for situations in which single consensus models fail to generate rational results (Chapter 6).

4. Extend multi-agent influence diagrams with the CBA approach, enabling one to apply further game theoretic analysis including finding the Nash equilibrium solutions and minimax/maximin strategies (Section 6.4).

5. Introduce two algorithms to discover the collectives in a population. The first algorithm will be an exact algorithm, guaranteeing that the collectives are rational according to social choice theory. The second algorithm is an approximation, attempting to reduce the runtime of the exact algorithm without rationality guarantees (Chapter 7).

1.5 Outline of Dissertation

This dissertation is organized as follows:

- Chapter 2 discusses existing techniques for combining opinions and preferences.
- Chapter 3 continues this discussion with a focus on the theoretical foundations for the described research.
- Chapter 4 defines notation and several key concepts that are used throughout the dissertation.
- Chapter 5 presents the *collective belief aggregation* approach.
- Chapter 6 demonstrates how game theoretic analysis can be applied to a large population using *collectives*.
- Chapter 7 describes and compares three algorithms for collective belief aggregation.
- Chapter 8 summarizes this dissertation's contributions and lists several possible research directions.

Chapter 2

Related Work

The research described in this dissertation deals with combining *beliefs* and preferences to form consensus models. Beliefs in this case refer specifically to an individual or agent's perception of the uncertainty in the world. Beliefs will be given in the form of probabilities or *likelihoods*. Preferences refer to the value an agent places on an outcome. This value could be given as a preference order that indicates the relative preference between a set of options, or it could be given as a numeric value for each option. Typically, each individual's beliefs and preferences will be represented by a Bayesian belief network or Bayesian decision network. This section discusses these foundational computational models as well as existing techniques for combining preferences, belief and evidence. I also provide a brief introduction to game theory, which deals specifically with the behavior of individuals or groups acting competitively or strategically.

2.1 Bayesian Networks

This research is based on a framework that is well-studied in Artificial Intelligence. Bayesian networks, also known as belief networks, are a form of graphical model that integrate the concepts of graph theory and probabilistic reasoning [50, 32, 44]. These networks define dependencies (and independencies) between random variables that can represent causality, implication or correlation. In a typical Bayesian network, random variables are represented by nodes and conditional relationships are represented by directed edges between the nodes. A variable is *conditioned on* all of its parents, described by the expression $P(X|Pa_x)$ where Pa_x is the set of parents of X. The joint *posterior* probability of a network is the product of local distributions at each node and is defined by the equation:

$$P(X_1, \ldots, X_n) = \prod_{i=1}^{n} P(X_i|Pa_{X_i}) \tag{2.1}$$

Where n is the number of nodes (variables) in the network. If a node has no parent, then its probability is its *a priori* or *prior* probability, or simply $P(X_i)$. In discrete Bayesian networks, the distribution at each node is represented by a conditional probability table (CPT) that defines the probability of each possible value of a variable given all possible values of each of its parents. The size of this CPT is dependent on the number of parents and the number of values each variable can take on. For binary variables, the size of a CPT is 2^{m+1} for m parents. Figure 2.1 shows a simple Bayesian network in which the grass being wet can be caused by either rain or the sprinkler. The conditional probability of grass wet is shown by the bottom table representing the expression P(Grass Wet | Sprinkler, Rain).

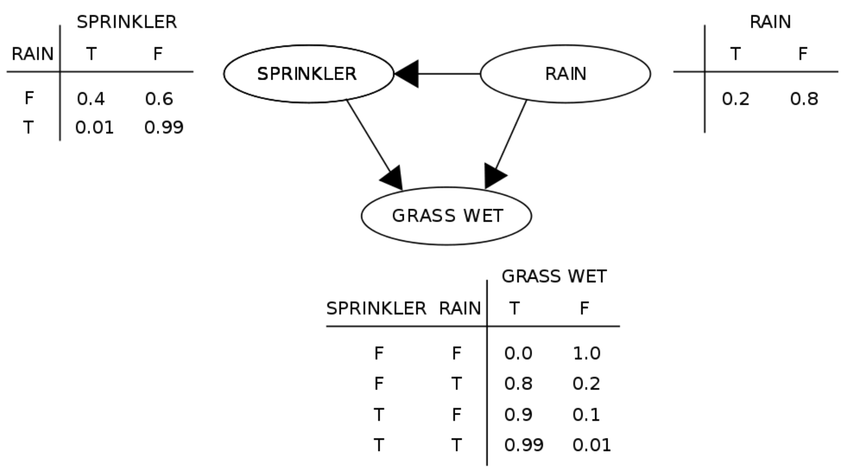


Figure 2.1: A simple Bayesian network with conditional probability tables at each node.

Inference on Bayesian networks is the process of determining the posterior probability of a variable given all of its parents and any observations that have been made about the variables in the network [68, 15, 14, 50]. A Bayesian network provides the structure of the relationships between variables in an uncertain environment, but in order to determine the probability of a given outcome, for example a *query* on a particular variable, one must perform inference on the network. The prior and conditional probabilities are propagated throughout the network and each variable updates its posterior probability based on any messages it receives from its parents and children. It then sends its posterior probability to its neighbors.

Several inference algorithms exist for Bayesian networks. Bucket or variable elimination sums out the effect of each variable from a list of factors, representing each "family" in the network [68, 15]. A family is composed of a node and its parents. The clique tree or junction tree propagation algorithm converts a Bayesian network into an intermediate structure in which families of nodes are represented by a single node [32, 68]. The variable elimination and clique tree algorithms are considered *exact* inference because they compute the posterior probability of a variable exactly.

Other algorithms exist for *approximate* inference, in which some reduction in runtime may be achieved, but the algorithm is not guaranteed to return the exact probability of a variable [50, 16]. Judea Pearl's loopy belief propagation algorithm [50] uses a very simple message passing concept, in which each node passes its current probability estimate to its neighbors, which then update their probability based on all their incoming messages. Over multiple cycles, the network will converge upon the actual posterior probabilities for each node. However convergence is not guaranteed to occur.

Bayesian networks can be extended to address decision problems using *influence diagrams* [30, 59], also known as Bayesian decision networks. In addition to nodes representing random variables (or *chance* nodes), influence diagrams contain *decision* nodes, representing a set of decision alternatives; and *utility* nodes, representing the value or risk associated with a possible outcome. Influence diagrams efficiently represent the uncertainty involved in real-world decision problems. Inference on Bayesian decision networks involves incrementally removing nodes [59] or converting the decision network to a Bayesian network [42, 35].

Bayesian Belief Aggregation

Belief aggregation is the process of combining probability estimates to form a *consensus* model from multiple human or software agents. Matzkevich and Abramson [40] cited two different approaches to belief aggregation. The first was called *posterior compromise*, which combines the beliefs after the network and probabilities have been defined and a query has been made. In other words, one would query separate networks and then combine the result. The authors introduced their alternative approach called *prior compromise* that instead found a consensus network *before* inference was done to determine the result of a query. This approach would involve fusing together networks that may also have different structure, called topological fusion. Once networks were fused, they combined the beliefs on local relationships using an approach called *family aggregation* [52].

Early researchers developed various *opinion pool* functions whose output was a numeric result of the combination of a number of inputs. An opinion pool function is a mathematical function to form a single aggregate value from multiple beliefs. Mathematically, $P_0 = f(P_1, P_2, \ldots, P_n)$ where each P_i is the probability estimation from the i^{th} contributor given N contributors. P_0 is the *consensus* estimation. The two most commonly used opinion pools are the linear opinion pool (LinOP) and the logarithmic opinion pool (LogOP). If the world is composed of m possible binary events LinOP is a weighted arithmetic mean with the following formula:

$$P_0(x) = \sum_{i=1}^{N} \alpha_i P_i(x) \tag{2.2}$$

where α_i is a non-negative weight assigned to each of the N contributors and $\sum_{i=1}^{N} \alpha_i = 1.0$. LogOP is a weighted geometric mean with the following formula:

$$P_0(w_j) = \frac{\prod_{i=1}^{N} [P_i(w_j)]^{\alpha_i}}{\sum_{k=1}^{2^m} \prod_{i=1}^{N} [P_i(w_k)]^{\alpha_i}} \tag{2.3}$$

Where w_j and w_k are each one of 2^m possible events given m states of the world [52].

Given the potential for opinion pool functions to form consensus models by aggregating multiple beliefs, Pennock and Wellman investigated whether consensus models maintain the structural integrity of models that formed them. In [52] they answered this in the negative,

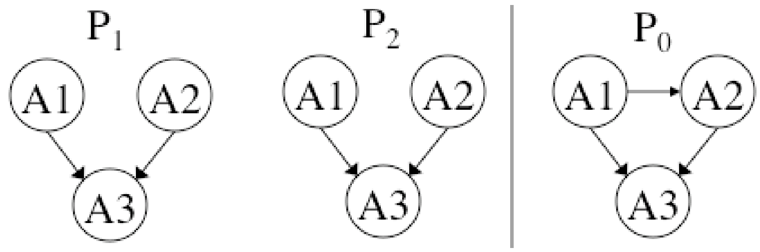


Figure 2.2: Opinion pool function results in a new dependency between $A1$ and $A2$.

showing that even when agents are in agreement on the structure of a model, opinion pool functions do not yield the same structure. They prove that it is not possible to maintain consistent structures using an opinion pool. The authors list a number of properties which must hold after aggregation in order for the structure to be maintained. These properties include independence preservation properties.

The authors show that with LinOP, independence is not preserved in the most basic case of two independent events A and B because $P_0(A)P_0(B) \neq P_0(AB)$ when $P(A)P(B) = P(AB)$ for individual agents. Using LogOp, they show that even when all contributors agree on the structure of the network, that the aggregation methods create dependencies that did not exist in the original topology. For example, given the Bayesian network shown in Figure 2.2 (from [52]), if two people supply beliefs on the probabilities of the same network, the aggregate result using LogOP creates a dependency $P(A_2|A_1)$ that did not exist in the original network. In other words, the aggregate defies the rules of independence, causing a dependency to arise between the two events.

Pennock and Wellman [52] show that no aggregation function can satisfy all of the consistency properties in a Bayesian network. However, they prove that these assumptions *can* hold with a Markov network because of a distinction between Bayesian and Markov networks. In particular, Markov networks include the axiom *strong union*, which states that if $P(A_j|A_k) = P(A_j)$ then $P(A_j|WA_k) = P(A_j|W)$ for all $W \subset Z$. When strong union is included, then independence is preserved in a consensus network. Therefore, the authors show that by converting a BN to an MN, one will maintain consistent consensus structures.

Pennock and Wellman also introduced the market-based belief elicitation and aggregation approach which is intended to move individuals towards consensus. This approach requires that individuals back up their beliefs by buying and selling *stocks* that indicate their confidence in an event occurring [51, 53]. The consensus value is determined by the resulting stock price. While this approach may improve accuracy when all agents have the same risk tolerance, this case is highly unlikely. In general a market based approach increases the subjectivity of the result as each individual has unequal desire to make a bet on their beliefs. In addition, the assumption that beliefs will move towards consensus by placing a financial risk on the humans ignores other non-market based factors that contribute to people's beliefs. In particular, one's background and experience play a strong role in political, religious and other subjective beliefs. Despite these limitations, a number of other researchers have followed in the competitive market-based approach to belief aggregation [48, 45, 41].

The Bayesian belief aggregation approach discussed in this dissertation does not attempt to force consensus, but instead enables the representation of diverse beliefs. Situations in which opinions will be polarized, such as in politics and policy-making are of particular interest. These fundamental beliefs are not typically modified by market-based forces, but rather they are defined by the diverse background and experiences of each individual.

2.2 Social Choice Theory

Social choice theory, also called social welfare theory, is a branch of research that has involved researchers in voting theory, economics and statistics. Social choice theory analyzes the manner in which one can determine a *social choice*, or collective decision based on the opinions and preferences of a group of individuals. The area of research was launched by economist Kenneth Arrow's when he introduced his rationality properties for combining preferences and theorems on the limitations of finding a social choice [1]. Many researchers followed to analyze and expand upon his findings in deterministic and Bayesian environments [31, 57, 22, 19, 56]. In summary, theorists discovered that there is no single social choice function that conforms to a set of mathematical principles for combining beliefs and preferences in general. The findings of these authors are discussed in more detail in Chapter 3, as well as their relevance to the research described in this dissertation.

Some researchers and theorists have introduced social choice methods that relax a particular constraint or demonstrate particular situations in which the rationality properties will hold [57]. Duncan Black introduced a resolution when a condition called *single peakedness* occurs [5]. If a set of alternatives can be ordered in a manner such that there is a logical range between them the alternatives can be ordered on a single line from one end of the range to the other. Given this order of alternatives, single peakedness will occur if an individual's most preferred alternative is placed highest on a the y axis of a graph, and all other alternatives fall away from the peak in a monotonically decreasing manner. If single-peakedness holds for all individuals, then a rational aggregate can be found. However, this is a structure of the alternatives that does not hold in general.

Other approaches include relaxing the requirement that the social choice function results in a complete ordering over all the options [56]. In other words, each individual need not have a complete preference order. Another option is to separate conflicting opinions into a "neutral position." Specifically, the authors discuss a method to group individuals into those who have a strict preference between options and those who have conflicting beliefs [56].

Some researchers suggest rejecting the Pareto condition when the beliefs are in conflict [54]. For instance, [23] demonstrates a situation in which two individuals are deciding whether to enter a duel that will be fatal for one individual. The Pareto choice is to duel, because they both believe the other individual will die. However, in reality one of their beliefs is incorrect. While in this dramatic situation, conflicting beliefs mean that some individual is incorrect, it also means that one would require an oracle to make a well-informed decision. If individuals are acting under their own control and opinions then their tendency may be to make a decision that maintains those beliefs. In many situations, beliefs cannot always be tested. For instance, one's belief about the existence or non existence of God and a vote based on this belief cannot be objectively analyzed and dismissed.

Maes and Faber introduced an approach in which the social choice function is determined probabilistically [39]. Specifically, the preference orders for a set of options is derived from probabilistic inference. The final preference order is then determined using an update rule similar in concept to Naive Bayes classification [64]. The approach described in this dissertation also derives the preference order from probabilistic inference, but then forms collectives and applies game theoretic analysis to the collectives, as opposed to computing a likelihood over the possible preference orders.

2.3 Game Theory

Game theory is the mathematical study of interacting agents, each with the self-interested goal of improving their own situation [10, 37]. The value of a situation is determined by the utility for each player. This is equivalent to the utility that is represented by Bayesian decision networks [59]. Game theory describes many types of solutions and strategies that are considered rational behavior in competitive and strategic environments. For instance, a Pareto optimal solution is one such that no other solution is preferred by all players in the game. A Nash equilibrium solution occurs if all players have taken on a strategy that maximizes their own utility, given the strategies of the other players.

Game theory is in some ways in opposition to social choice theory. While social choice theory attempts to find a solution that is best for the whole population, game theory analyzes the quality of solutions for individuals or cooperative groups within a population. However, it does enable one to objectively analyze a competitive situation and determine if a solution is minimally acceptable by a population, meaning it is Pareto optimal. In many cases there will be more than one Pareto solution. Social choice theory shows that in an attempt to find a single best solution, it is possible that the solution found will not even be minimally acceptable. Many situations for which social choice theory has been applied, such as a community selecting a leader, have more in common with non-cooperative situations.

Game theoretic analysis is typically applied in situations in which a player represents a single individual in a game. Coalition game theory analyzes the behavior of individuals who form coalitions, in which the members must decide how to share a payoff [37]. Coalitions involve members forming agreements between them that improve the situation for all members, even if they have competing goals [4]. Ortiz [47] introduced the concept of graphical games for representing games with many players. To my knowledge, no approach has grouped individuals according to their beliefs and preferences before applying game theory.

2.4 Collaborative Decision-making and Prediction

Other techniques that are used to combine beliefs and evidence to form a prediction are now summarized.

2.4.1 Collaborative filtering

Collaborative filtering is the process of making predictions from many individual opinions on a given topic. Collaborative filtering algorithms are typically used to make a recommendation to individuals by suggesting other items that individuals with similar interests have enjoyed. Web *etailers* such as Netflix and Amazon use collaborative filtering in this manner based on ratings supplied by customers and purchase history. Collaborative filtering is not interested in finding a solution for the whole population, as in social choice theory, but forms a customized solution based on personal preferences and the preferences of similar individuals. Recent techniques include matrix factorization such as singular value decomposition (SVD) and combining a number of algorithms into a hybrid approach. These techniques are used by one of the leaders of the Netflix prize competition [3].

While the approach described in this dissertation is similar to collaborative filtering in that a prediction is based on community input versus singular input, it advances existing techniques by building entire models from community input instead of simply providing a prediction on one item or one event. Collaborative filtering maintains the diversity of a population, but predictions are based on shallow beliefs with no context or reasoning to back them up. In contrast, the intention of the described approach is to elicit opinions that ask contributors to consider the reasoning behind their beliefs, not just isolated beliefs without context. This enables predictions at a much finer level of granularity, and allows one to build complex causal and contextual models that can provide recommendations as well as help decision-makers understand the possible outcomes as well as the factors that contribute to the outcomes.

2.4.2 Dempster-Shafer Theory

Dempster-Shafer theory is a branch of probabilistic reasoning that extends the theory of probability with the concept of plausibility [60, 17, 58, 67]. The theory was created to address situations in which one is not certain about the likelihood of an event. It was intended to distinguish probabilities that are certain (such as an unweighted coin), from probabilities that are given because the situation is unknown. Belief is similar to Bayesian networks in that it is an estimate of the likelihood of an event. Plausibility is a second measure indicating an upper bound on the probability. An individual will typically supply both their belief (as a lower bound) and the plausibility as an upper bound.

Dempster-Shafer theory can be used to combine evidence from multiple individuals. Combining multiple high-confidence belief estimates will result in a stronger confidence in an event occurring [58]. However, aggregating opposing beliefs will result in zero probability [67]. Therefore, Dempster-Shafer is not appropriate for handling diverging beliefs and preferences.

2.4.3 Data and Sensor Fusion

Data fusion is an area of study dedicated to combining information from multiple sources to produce a *situation awareness* of an environment [28]. Sensor fusion is essentially the same concept as data fusion, however specifying that the data comes from a sensor network. Data fusion typically occurs at multiple levels of abstraction, beginning from low-level fusion that attempts to determine the objects in an environment; to high level fusion that attempts to determine the intention of the entities in the environment [62]. Data fusion often uses techniques discussed in this section including Bayesian networks and Dempster-Shafer theory [34]. Kalman filters are mathematical approach used for low-level fusion that incorporate a representation of the noise in an environment to aggregate multiple noisy pieces of information [33].

Chapter 3

Theoretical Foundation

This chapter summarizes several decades of theoretical research related to combining beliefs and preferences of multiple individuals to form a consensus solution. I first discuss the "rationality" principles for aggregation as they were defined by social choice theorists. I then summarizes the impossibility theorems developed by theorists that show that no single consensus approach combining an arbitrary set of beliefs and preferences can uphold all the principles at once. The first theorem is related to qualitative preferences in a deterministic environment. The remainder of the theorems address quantitative beliefs and preferences in a Bayesian environment, in which there is uncertainty associated with events of interest. Finally, this chapter discusses some additional observations on the challenges of belief aggregation that extend the findings of prior theorists.

The reader may ask; "why does the author care about rationality?" To clarify, this dissertation does not attempt to address the highly controversial and philosophical argument about whether humans act rationally. Specifically, do humans take actions that maximize their own (or their community's) expected utility? This discussion is left to the fields of psychology, philosophy and the other social sciences. There are many excellent dissertations on the perceived irrationality and cognitive bias in humans [66, 12, 38, 21, 11, 6]. However, a motivation of this work is the belief that if we could fully understand all the factors that go into a human's thought processes and behaviors, we may discover that humans do in fact act rationally according to their own "internal" models. As stated succinctly by Maes and

Faber, "it is, to a large extent, our professional duty to see to it that decision making is at the same time (a) rational and (b) sensitive to the perception of the individuals at large." [38]

This dissertation presents a precise definition of rationality, introduced by social choice theorists for the purpose of combining beliefs and preferences [1, 2, 19, 31, 56]. The goal is to find one or more social choice solutions that uphold a set of principles for rational aggregation based on the beliefs and preferences of a group of individuals. No judgment is made about the correctness of the individuals' beliefs. Nor are any assumptions made that the individuals or group would act consistently with their beliefs and preferences.

This chapter discusses the principles of rationality that have been defined by social choice theorists, and summarizes the challenges researchers and theorists have come across in their attempts to form aggregate models that uphold these principles. The remainder of the dissertation addresses an approach to enabling belief and preference aggregation for decision-making that upholds these pre-defined properties for rational aggregation.

3.1 Rationality Principles for Aggregation

The *collective belief aggregation* approach splits a population of individuals with divergent beliefs into groups of similar beliefs. The objective is to define a splitting function such that the aggregate of beliefs within each group will have rational behavior according to a set of mathematical principles of rationality with respect to combining the preferences, beliefs and utilities of multiple individuals, as defined by theorists in the fields of voting, game and probability theories. This section describes these principles. Well-known counter examples that demonstrate the failure of aggregation methods are then summarized.

3.1.1 Preference Relations

The notation in table 3.1 is used to describe the *pairwise* preference relationships that describe the preference ranking between two alternative options x and y [1]. A combination of these relations will form a preference order over a set of three or more options.

x, y:	alternative options
i, j:	individuals
P:	a preference relation representing *strict preference*
I:	a preference relation representing *indifference*
R:	a preference relation representing *either* strict preference or indifference
xPy:	a group prefers x to y
xIy:	a group is indifferent to x and y
xRy:	a group prefers x to y or is indifferent to them
xP_iy:	an individual i strictly prefers x to y
xI_iy:	an individual i is indifferent to x and y (doesn't prefer either)

Table 3.1: Pairwise preference relations

The following example illustrates the preference order relations in Table 3.1. Suppose person A prefers vanilla ice cream to chocolate, and chocolate to strawberry. Their pairwise preference orders would be vP_ac and cP_as. Putting them together for all flavors results in vP_acP_as. Suppose another individual B prefers vanilla to chocolate, but is indifferent between chocolate and strawberry. B's preference order could be vP_bcI_bs or vP_bsI_bc. Since both individuals prefer vanilla to chocolate, we could make a generalization that states $vPcRs$. However, we could not state that $vPsRc$ because A strictly prefers chocolate to strawberry.

The personal relation xR_iy was included in the Arrow's definition, however xR_iy is not considered in this discussion. An assumption is made that a person will either be indifferent to two options x and y or prefer one option over the other. While in reality individuals' preferences may vary over time, the described approach considers the answer at a single point in time, given a finite set of factors for that selection. If a person sometimes picks x and sometimes picks y, then they can safely be considered to be indifferent to the two. Or they could provide only their preference for x and y at the time of elicitation.

3.1.2 Arrow's Axioms

The properties in Figure 3.1 were introduced by economist Kenneth Arrow in 1950 [1], and later adapted by Arrow and other researchers [1, 2, 19, 56]. These social choice theorists determined that these properties must hold for belief and preference aggregation to be considered *rational*. Specifically, any attempt to combine the beliefs and preferences of multiple individuals to form a *social choice* should adhere to these properties. Some of the properties are mathematical in nature, for instance transitivity is a fundamental requirement for order relations on natural numbers. The Pareto optimal requirement attempts to maximize expected utility. The completeness property implies that an order must be returned from any social choice function and that the order includes all options given. The remainder of the properties serve to enable the fair consideration of all options and opinions— restricting no options, orderings or individual beliefs.

Given a society of interest, S:

1. *Universal Domain (UDP)*: All preference orders are allowed
2. *Completeness (CP)*: Social choice function returns an order that includes all relevant alternatives
3. *Transitivity (TP)*: if S prefers A to B and B to C then S prefers A to C (also replace "prefers" to "is indifferent to")
4. *Pareto optimality (should be at least weekly Pareto optimal)*:
   (a) Weak Pareto principle (WP). For all x and y, if xP_iy for all i, then xPy:
   (b) Strong Pareto principle (SP). For all x and y, if xR_iy for all i, and xP_iy for some i, then xPy:
5. *Independence of irrelevant alternatives: (IIP)* A society's preference order over a subset of options should be the same as the order over the whole set of options.
6. *Non-dictatorship and non-imposition (NDIP).* There is no dictator. Individual i is a dictator if, $\forall x$ and $\forall y$, $xP_iy \rightarrow xPy$. Non-imposition means that no order has been pre-determined for any individual.

Figure 3.1: Rationality properties defined by Kenneth Arrow [1, 2, 19]

3.1.3 Bayesian Rationality Principles

Hylland and Zeckhauser [31] added two properties for Bayesian inference:

1. *Unanimity: (UP)* If each individual i has the same probability estimate P_i, then the aggregate P_o should equal P_i.
2. *Bayesian group decision: (BGP)* The aggregate is a function of only the probability estimates (P) and utilities (U) supplied by the group of interest.

3.1.4 Rational Social Choice

An aggregation technique that upholds the properties defined in Figure 3.1 and Section 3.1.3 will be referred to as a *rational social choice (RSC)* in this dissertation. This is defined formally in Section 4.3, Definition 6.

3.2 Impossibility Theorems

3.2.1 Arrow's Impossibility Theorems

The remainder of this chapter discusses the *impossibility* theorems developed by social choice theorists that show no social choice function conforms to all of the properties in Figure 3.1. Rational social choice is first discussed in the context of individuals and groups considering their preferences deterministically and symbolically. In other words they supply a precise preference order over a set of options that is not based on the uncertainty of outcomes or their perceived risk. A simple example [63] demonstrates a situation in which a social choice function fails to conform to the transitivity principle. Table 3.2 shows three individual's preferences for ice cream, where a rank of 1 indicates that the flavor is the top choice for the individual. When attempting to find a consensus order that combines all the individuals given their ice cream preferences, one compares each pair of flavors in the following manner using a majority vote:

Individual	Vanilla	Chocolate	Strawberry
X	1	2	3
Y	2	3	1
Z	3	1	2

Table 3.2: Table showing the preference orders for ice cream for three individuals. A rank of 1 indicates the most preferred flavor.

- Two out of the three individuals prefer vanilla (v) to chocolate (c), so vPc.
- Two out of the three individuals prefer chocolate to strawberry (s), so cPs.
- Given the above preferences, if vPc and cPs, then vPs should hold (according to transitivity (TP)). However, it can be seen that the majority actually prefers strawberry to vanilla (sPv).

Thus, the transitivity principles cannot hold without relaxing another principle.

3.2.2 Bayesian Rational Social Choice

Rational social choice in the presence of uncertainty is now discussed. In this case, probabilities are used to represent an individual's (or group's) belief in the likelihood of an event given its causal factors, while utilities indicate the value of an outcome given the inherent uncertainty. It is natural that the properties defined above for rationality in the non-deterministic case also hold in probabilistic belief aggregation.

Hylland and Zeckhauser show that aggregating probabilistic beliefs and utilities separately can result in an aggregate that breaks the Pareto condition [31]. The following example demonstrates their results using a Bayesian decision network. In the network in Figure 3.2 there is a decision node A that has two possible options (a_1 and a_2), a binary variable X and a utility node U that is dependent on A and X. $U(A|X)$ represents the utility of a decision option in A given the state of X.

Suppose there are two individuals who each supply their beliefs about $U(A|X)$ and $P(X)$ (the probability of X). Individual 1's probability distribution is $P_1(X = T) = 0.25, P_1(X =$

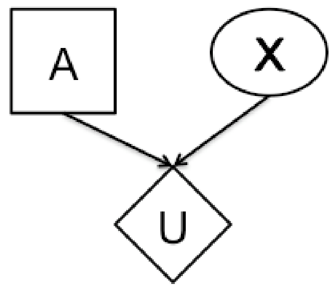


Figure 3.2: Bayesian decision network for a simple decision involving two possible actions in A, a binary variable X and utilities U.

$F) = 0.75$. Individual 2's probability distribution is $P(X = T) = 0.75, P(X = F) = 0.25$. Table 3.3 shows the conditional utility of A (rows), given X (columns) for the first and second individuals. For example, individual 1 believes that utility of option a_1 given that X is false is 1.0, (represented by $U_1(A = a_1|X = F)$). The aggregate of $U_1(A|X)$ and $U_2(A|X)$ using an arithmetic mean would generate the values in Table 3.4 for $U_o(A|X)$.

U_1		X	
		F	T
A	a_1	1.0	-1.2
	a_2	0.0	0.0

U_2		X	
		F	T
A	a_1	-1.2	1.0
	a_2	0.0	0.0

Table 3.3: Conditional utilities $U_1(A|X)$ and $U_2(A|X)$ for two individuals.

		X	
		F	T
A	a_1	-.1	-.1
	a_2	0.0	0.0

Table 3.4: The aggregate of the conditional utilities in Table 3.3.

The average probability, $P_0(X) = 0.5$. The *expected* utility, $U_0(A) = U_0(A|X)P(X)$, can be found by marginalizing over the values of X, giving the following values:

$U_0(A = a_1|X) = -.1 * 0.5 - .1 * 0.5 = -.1$

$U_0(A = a_2|X) = 0$

In this situation it appears that the aggregate prefers a_2 over a_1. However, it is evident that each individual actually prefers a_1 to a_2:

First individual:

$U_1(A = a_1|X) = 1.0 * 0.75 - 1.2 * 0.25 = 0.45$

$U_1(A = a_2|X) = 0 * 0.75 + 0 * 0.25 = 0.0$

Second individual:

$U_2(A = a_1|X) = -1.2 * 0.25 + 1.0 * 0.75 = 0.45$

$U_2(A = a_2|X) = 0 * 0.25 + 0 * 0.75 = 0.0$

Thus, the aggregation result breaks the Pareto condition. Seidenfeld, et al., [56] extend these findings to reiterate that there is no Bayesian social choice function that always returns a Pareto optimal result when probabilities and utilities are aggregated separately. They show that a small variation in belief can cause a situation in which no range of average expected utility results in a shared preference. Nehring [46] extends the results of [31, 56], showing that aggregation of separable factors (for example independent events) in the computation of expected utility can result in a non-Pareto solution.

3.3 Additional Observations and Distinctions

This section discusses some additional situations that arise from belief and preference aggregation, based on the author's observations. First, a distinction is made between prior and posterior aggregation.

3.3.1 Prior versus Posterior Aggregation

The results described in Section 3.2.2 refer to separate aggregation of probabilities and utilities before computing expected utility. This is equivalent to prior aggregation, discussed in [40]. In contrast, posterior aggregation means that each individuals' expected utility is computed first, and the aggregate is a function of these expected utilities. Formally, given two individuals, prior aggregation finds:

$$U_0 = f(g(U_1(A|X), U_2(A|X)), g(P_1(X), P_2(X))) \tag{3.1}$$

Where g is an aggregation function and f finds the expected utility. In contrast, posterior aggregation finds:

$$U_0 = g(f(U_1(A|X), P_1(X)), f(U_2(A|X), P_2(X))) \tag{3.2}$$

Aside from the aggregation formula, other differences exist between prior and posterior aggregation. Prior aggregation attempts to form a consensus representation of a belief model, including all of the variables in the model. Given a consensus model, one could query the consensus distribution of any node in the network. In contrast, posterior aggregation is only interested in the posterior value of a given node, ignoring the possible variation in dependencies represented by the node's network. In other words, individuals may agree on an outcome but disagree on the factors involved in the outcome. The determination of which aggregation approach is "best" may depend on the desired goals of a set of decision makers. If one would like to determine which decision option to select, or which probabilistic outcome is more likely, one could use posterior aggregation. If one wanted to maintain a model from which to make queries about many variables, then prior aggregation may be most appropriate. A potential benefit of prior aggregation is that complexity improvements may be possible. In posterior aggregation, one assumes that inference is run on each of m networks, representing the beliefs of m individuals in a population. In prior aggregation, the number of consensus networks that inference is performed on is potentially less than m.

3.3.2 Dictatorship with Bayesian Beliefs

I now demonstrate an additional failure of Bayesian belief aggregation. An interesting property of probabilistic beliefs is that a single individual can skew aggregation such that the result sways his direction. Given this phenomenon, it is possible for a person to observe the beliefs of others and intentionally set his beliefs to override the majority. Thus, I have extended theoretical findings described previously to show that in a Bayesian situation, it is possible for the non-dictatorship (ND) principle to be broken.

Consider a Bayesian treatment of the ice cream preference example used to describe Arrow's theorems. Suppose we have a group of three individuals who are going to a birthday party. They are trying to decide which flavor of ice cream to bring given the flavor of cake that will be served at the party. The possible ice cream and cake flavors are chocolate, vanilla and strawberry. The individuals agree that each flavor of cake is equally likely ($33.3\bar{3}\%$), but they have different opinions on the quality of combinations of cake and ice cream. In addition, one person really likes chocolate so he plans on skewing the model so he can get chocolate ice cream no matter what the others prefer. He also hates vanilla and plans to skew the model so his preference order over all ice cream flavors wins. Imagine that this person is compiling the model, so he can see the other individuals' utility values before supplying his own.

The following table indicates the utilities of the ice cream flavor given the cake flavor of two individuals. Both individuals happen to have identical opinions and therefore this table is an average of their utilities. In the table, rows = cake and columns = ice cream. Cell(i, j) represent the utility of ice cream flavor j given cake flavor i. While utility selection can be somewhat arbitrary in that there are no bounds set on the values supplied in general, in this case utilities are restricted to the range $[0, 20]$.

	Chocolate	Vanilla	Strawberry
Chocolate	0	10	5
Vanilla	10	0	10
Strawberry	5	10	3

The probability distribution for cake flavor is:

[Chocolate= $.33\bar{3}$; Vanilla=$.33\bar{3}$; Strawberry=$.33\bar{3}$]

Given the opinions of these two individuals, it can be seen that vanilla wins and strawberry is second, with the overall social preference order of $vPsPc$:

Utility of chocolate ice cream:

$$0.33\bar{3} * 0 + 0.33\bar{3} * 10 + 0.33\bar{3} * 5 = 5.0$$

Utility of vanilla ice cream:

$$0.33\bar{3} * 10 + 0.33\bar{3} * 0 + 0.33\bar{3} * 10 = 6.67$$

Utility of strawberry ice cream:

$0.33\bar{3} * 5 + 0.33\bar{3} * 10 + 0.33\bar{3} * 3 = 6.0$

Now imagine that the third individual sees these results, and thinks to himself "Hmm.., now I only need to set my utilities so that chocolate wins and vanilla loses'.' He then selects the following utilities:

	Chocolate	Vanilla	Strawberry
Chocolate	20	0	0
Vanilla	20	0	5
Strawberry	20	0	0

The next utility table is the mean of the three individuals:

	Chocolate	Vanilla	Strawberry
Chocolate	6.67	6.67	3.33
Vanilla	13.33	0	8.33
Strawberry	10	6.67	2.0

The result of adding his utilities to the mix causes chocolate to win by a landslide and vanilla to be scorned, yielding the social preference order $cPsPv$:

Utility of chocolate ice cream:

$0.333 * 6.67 + 0.333 * 13.3 + 0.333 * 10 = 9.98$

Utility of vanilla ice cream:

$0.333 * 6.67 + 0.333 * 0 + 0.333 * 6.67 = 4.44$

Utility of strawberry ice cream:

$0.333 * 3.33 + 0.333 * 8.83 + 0.333 * 2.0 = 4.72$

As defined in Fig. 3.1, an individual i is a dictator if $\forall x, y : xP_iy \rightarrow xPy$. It can be seen that this is the case in this scenario, as the pairwise preferences of the two individuals have been reversed. The pairwise preferences for the first two individuals are: vPs, sPc, vPc. The pairwise preferences for the society are identical to the the dictator: sPv, cPs, cPv. The same outcome is observed using posterior aggregation, showing a situation in which posterior aggregation breaks one of the principles defined in Figure 3.1:

Utility of chocolate ice cream:

$(5.0 * 2 + 20)/3 = 10.0$

Utility of vanilla ice cream:

$(6.67 * 2)/3 = 4.45$

Utility of strawberry ice cream:

$(6.0 * 2 + 1.67)/3 = 4.56$

Thus, this example extends the findings by Bayesian theorists to demonstrate that Bayesian belief aggregation can break the non-dictatorship principle. I conclude that there is no Bayesian social choice that maintains the social choice principles, even when using posterior aggregation.

3.3.3 Inconsistency between the individual and the social choice

Many of the examples in this chapter describe discrepancies between the social optimum, derived from the average expected utility, and the individual optima, derived from each individual's expected utility. The following example demonstrates a situation in which an outcome predicted by the average expected utility is not the one that is preferred by the majority of the population. Suppose we have three individuals and two options, A and B. The expected utility for the options for each individual are shown in the following table.

	1	2	3	Votes
A	0.8	1.0	0.8	1
B	1.0	0.5	1.0	2

The majority of the individuals would select B based on their expected utilities. However, the average expected utility would select A. The following table shows the arithmetic and geometric means for the above expected utilities:

	Arithmetic mean	Geometric mean
A	.867	.83
B	.86	.79

Technically this result does not break the principles of rationality defined in Figure 3.1 since not every individual prefers B to A. However it does demonstrate an unexpected effect of comparing individual preferences with the social choice. Based on the individual expected utilities, the majority would choose one option, while the average expected utility would choose another. In this case one may ask which is the appropriate option to select? This phenomenon may be related to the game theoretic concept of the *price of anarchy*, which is the difference in the utility if everyone acts for the social good, versus if everyone acts for their own selfish good [49]. Section 6.5 discusses the inconsistency in more detail.

Chapter 4

Formal Definitions

This chapter introduces notation and definitions that will be used throughout the dissertation. The building-blocks that form the foundation for the collective belief aggregation approach will be defined and validated. The reader should refer to this chapter when notation and definitions are referenced by other chapters.

4.1 Notation

P	A population of m individuals
i, j, q, r, s	individuals in P
P, Q	Belief estimates in the form of probability distributions
$P(X)$	A probability distribution for a variable X
$P'(X)$	A posterior probability distribution for a variable X
$U(X)$	A utility for a variable X
$P(Y\|X)$	A conditional probability for a variable Y given the value of variable X
$U(Y\|X)$	A conditional utility for a variable Y given the value of variable X
$EU(X)$	The expected utility for variable X
$B(*)$	A belief where B could be any of P, U, EU and * could be a variable or a conditional relationship
$B_i(*)$	A belief belonging to an individual or specified group i
$B_0(*)$	A consensus (mean) probability distribution for a group
C	A set of k subgroups such that each element of a population is in exactly one subgroup
C_j	A single subgroup (cluster or *collective*) in a population
T	A *partition* of a population into collectives where $T = C$ if C is a set of collectives
ϕ_j	A *collective belief* for a collective j
Φ	A set of collective beliefs for a partition T
O	A set of r options
R	A *rank order* over a set of options
X_i, X_j	variables (nodes) in a Bayesian network containing n nodes
Pa_i	The parents of node X_i
$P(Pa_i)$	The belief in the probability of X_i's parents
$P(X_i\|Pa_i)$	The belief in the conditional probability of X_i given its parents
PT_q	A probability distribution (prior, conditional or joint) representing an individual or group q's beliefs
F_i	The family of a node X_i in a network, containing a node and its parents

Table 4.1: This notation is used throughout the dissertation. The terms in *italics* are defined later in this chapter.

4.2 Kullback-Leibler (K-L) divergence measure

The *Kullback-Leibler divergence measure*, also known as *relative entropy*, measures the difference between two probability distributions P and Q. Specifically, it measures the number of extra bits it would take to encode samples from P using samples from Q [36]. The K-L divergence measure has been used to compare the parameters in multiple Bayesian models [65], as it will be used to compare consensus models in this dissertation. Given a "target" or "actual" distribution P and an estimate Q, the K-L divergence of Q from P over n data points is:

$$KL(P||Q) = \sum_{i=1}^{n} P(i) log \frac{P(i)}{Q(i)} \tag{4.1}$$

A number of examples in this dissertation will measure the K-L divergence of a consensus distribution formed from a partition of a population from the distribution of all individuals' beliefs. The distributions were computed from a histogram of n bins over the range of values provided by the population, which could be probability estimates or utilities. K-L divergence may be measured for one or more variable's distributions. Each individual will provide probability or utility values $B(X_i)$ for each variable X_i in a set of beliefs X. Each belief estimate will fit into one of the n bins. $P_{Ii}(b)$ is the probability that an individual in a population provided a belief value for variable X_i that fits into bin b. For each subgroup C_j and each variable X_i, the mean likelihood, $B_{0j}(X_i)$ is defined in eq. 4.2.

$$B_{0j}(X_i) = \frac{1}{|C_j|} \sum_{l=1}^{|C_j|} B_l(X_i) \tag{4.2}$$

$P_{Ci}(b)$ is the probability that an individual's belief for variable X_i is in bin b assuming that each individual assigned to subgroup C_j provides the belief $B_{0j}(X_i)$. Given a set of beliefs X and the set C of k subgroups, the K-L divergence between P_I, the distribution of the individuals, and P_C, the distribution of the set of subgroups is computed as follows:

$$KL(P_I(X_i)||P_C(X_i)) = \sum_{b=1}^{n} P_{Ii}(b) log \frac{P_{Ii}(b)}{P_{Ci}(b)} \tag{4.3}$$

$$KL(P_I||P_C) = \sum_{i=1}^{|X|} KL(P_I(X_i)||P_C(X_i)) \tag{4.4}$$

4.3 Collectives and Collective Beliefs

The following definitions form the foundation for the collective belief aggregation approach. The first definition is based on a concept from social choice theory, variously called a "preference profile" [19], "preference ranking" [56], "preference pattern" [1] and "preference order" [20].

Definition 1. *Rank Order: A rank order* R *is a partial order* $\rho_1 \leq \rho_2 \leq ... \leq \rho_r$ *over a set of options* O *containing* r *options* $\mathrm{o}_1..\mathrm{o}_\mathrm{r}$*, where* ρ_i *is some* o_j*. A rank order* R *over* O *is an order such an item* ρ_i *is preferred to (or indifferent to)* ρ_j *if and only if* ρ_i *is before* ρ_j *in* R*.*

For example, given $O = \{a, b, c\}$, if b is preferred or indifferent to c (bRc), and c is preferred or indifferent to a (cRa), then the rank order over O is $R = b \leq c \leq a$.

The term "collective" is often used in the study of society and social behavior and has many different sociological definitions. The Wikipedia definition states that "a collective is a group of people who share or are motivated by at least one common issue or interest, or work together on a specific project(s) to achieve a common objective." (wikipedia.org 12/26/09) The Merriam Webster definition states that a collective is "marked by similarity among or with the members of a group" and "involving all members of a group as distinct from its individuals" (m-w.com 12/26/09). The second definition implies that there is some generalization that is "distinct" from the specifics of its individuals. Thefreedictionary.com (3/7/10) includes the definition "forming a whole or aggregate."

The goal is to provide a more rigorous definition of a collective that captures the implications of these English language definitions in a mathematical or set theoretic form. Since such a definition has not been found in the literature, the following definition has been created. A collective is a group such that a specific generalization of the group holds for all members of the group. In set theory, this is simply a subset of a set, with the generalization being the property that defines the subset.

Definition 2. *Collective: A collective* C *w.r.t. a property* A *is a subset of a population* P *(*$\mathrm{C} \subseteq \mathrm{P}$*) s.t.* A *holds for all members of* C*. If* A *holds for an individual* $\mathrm{p} \subseteq \mathrm{P}$ *then* p *is a member of* C*. A null set* $\emptyset$ *with respect to property* A *indicates that the generalization does not hold for any member of* P*.*

A collective is not a "hive mind" and its members do not necessarily agree on all issues. Typically a common interest or set of interests defines a collective. Likewise, the set theoretic definition of collective has a single common property. In theory A could be a non-empty set of common properties.

Arrow's theorem shows that there is no rational social choice that can hold for an arbitrary group and their preferences [1, 19, 56]. In other words, there is no generalization that can be made about an arbitrary population based only on their preferences. However, if a group of individuals happens to have an identical rank order of their preferences, then we can make a generalization about that group based on this rank order. If the options are taken from a discrete set of values, and an individual i's rank order R_i is identical to another individual j's rank order R_j then a generalization R_0 can be made about the individuals i and j, such that $R_0 = R_i = R_j$. The following example illustrates this: if $O = \{A, B, C\}$ and $R_i = BCA$ and $R_j = BCA$, then clearly $R_0 = BCA$. Thus, if a group of individuals G shares a rank order R over a set of discrete valued options O then R can define a collective C.

Definition 3. *Rank Order Collective: If* $\mathrm{R_j}$ *is a rank order over a set of options* O*, and* $\mathrm{C_j}$ *is a subset of a population* P *s.t.* $\forall \mathrm{p} \in \mathrm{C_j}$, $\mathrm{R_p} = \mathrm{R_j}$ *and* $\forall \mathrm{q} \notin \mathrm{C_j}$, $\mathrm{R_q} \neq \mathrm{R_j}$*, then* $\mathrm{C_j}$ *is a collective defined by the property* $\mathrm{R_j}$ *and is called a* rank order collective*.*

Bayesian outcomes will now be mapped to the rank order concept. The terms *variable* and *node* will be used interchangeably in this dissertation when referring to Bayesian networks. Each variable X_i in a Bayesian network has a number of possible values ($\{x_{i1}, x_{i2}, ..., x_{ir}\}$), where r is the arity of X_i. The posterior probability of a variable in a Bayesian network being a specific value (x_{ij}) is derived through inference. Given the posterior probabilities of the values of a variable, there will be an order from *most likely* to *least likely* for these possible values. For example, given a binary variable X, with a probability distribution $P(X = T) = 0.25, P(X = F) = 0.75$, the order of values is FT. This order is analogous to the rank order in Definition 1. In the case of a Bayesian decision network, the result of inference is a set of *expected* utilities for the possible decision options [30, 59]. The decision options can be ranked by order of *highest expected utility* to *lowest expected utility*, or *best* to *worst* option. Given a Bayesian network, the rank order can be determined for an arbitrary variable or decision in the network.

Definition 4. *Bayesian Rank Order: A Bayesian rank order* R^* *with respect to a variable* X *in a Bayesian network is a rank order over the posterior probabilities* $(\mathrm{P}(\mathrm{X}=\mathrm{x_1}),\mathrm{P}(\mathrm{X}=\mathrm{x_2}), ..., \mathrm{P}(\mathrm{X}=\mathrm{x_r}))$ *of the values of* X*. In a Bayesian decision network,* R^* *is the rank order of the expected utilities* $(\mathrm{U}(\mathrm{o_1}), \mathrm{U}(\mathrm{o_2}), ..., \mathrm{U}(\mathrm{o_r}))$ *of the options of* D*, where* $\mathrm{U}(\mathrm{o_i})$ *is the expected utility of decision option* $\mathrm{o_i}$*.*

Before a collective is defined in terms of the Bayesian rank order, it must be shown that an aggregate of $m > 1$ equivalent Bayesian rank orders R over the probabilities of a variable X, or the utilities of a decision D, results in the same rank order, R. For example, given two sets of probability estimates for a variable that are in the same Bayesian rank order $X : P_1(X) = [0.25, 0.75]$, $P_2(X) = [0.3, 0.7]$, will the mean of the two probability estimates result in the same Bayesian rank order? This can be illustrated by showing that the mean of $m > 1$ arbitrary sets of n ordered values in $\mathbb{R}$ results in an ordered set of values.

Proposition 1. *The sum of* $\mathrm{m} > 1$ *sets of* n *real valued (*$\mathrm{r} \in \mathbb{R}$*), ordered numbers will result in an ordered set.*

Proof 1. *This is proven by contradiction. Imagine that there are* $m = 2$ *sets of* n *cups (*C_1 *and* C_2*) that are ordered from left to right such that a cup contains at least as much liquid as the cup to its left. Now consider that the liquid in each of* C_2*'s cups is added in order to each of* C_1*'s cups. So the liquid in the* i^{th} *cup in* C_2 *(*c_{2i}*) is added to the* i^{th} *cup in* C_1 *(*c_{1i}*). The amount of liquid in the cups in* C_1 *will remain in order. If this were not the case, then cup* c_{2j} *would have had to have more in it than cup* c_{2i} *(where* i *is farther left than* j*). However, this would mean that* C_2 *was not in order, which would be a contradiction. If* $m > 2$*, this applies for combining all sets of cups.*

The arithmetic mean of the values added to each cup is found by dividing the amounts of liquid in each cup c_{1i} by the same amount (m). Therefore finding the mean will not effect the rank order. The geometric mean $(\sqrt[n]{c_{11} * c_{12} * ... * c_{1k}})$ is not additive but is multiplicative. However, the same result can be shown with a slight change. When finding the product of two values ≥ 0, the result of $x * y_1$ will always be less than or equal to the result of $x * y_2$ where $y_1 \leq y_2$. When aggregating values < 0, simply add the absolute value of the minimum value provided to all values. The cup demonstration is equivalent to the utility or probability values for a collective. Each set of cups represents the probabilities of the values of a variable $P(X = x)$. Finding the mean of the distributions for a collective is equivalent to finding the mean of the probabilities of each value of X. For example, given the probability distributions $X : P_1(X) = [0.25, 0.75]$, $P_2(X) = [0.3, 0.7]$, the arithmetic mean of $P_1(X)$ and $P_2(X)$ is $[0.275, 0.725]$, resulting in the same rank order as the original values.

The definition of *Bayesian rank order* is based on the relative likelihood of a probabilistic variable's possible values or the relative expected utility of a set of decision options. The *collective belief* of a collective is the aggregate of the probabilities or utilities that define the *Bayesian rank order* for the collective.

Definition 5. *Collective Belief: If* R^* *is a Bayesian rank order over the posterior probabilities of a variable* X, *(or the expected utilities of a decision D), and* C *is a rank order collective s.t.* $\forall \mathrm{c_i} \in \mathrm{C}$, *the Bayesian rank order of* $\mathrm{c_i}$ *is* R^*, *then the* collective belief ϕ *of* C *is the aggregate of the* k *probability distributions (or expected utilities) supplied by the members of* C.

Definition 6. *Rational Social Choice: A collective belief* ϕ *is a rational social choice if it conforms to the properties for rational belief aggregation defined in Figure 3.1 and Section 3.1.3. This is shown in Section 4.3.4.*

The following definitions define the *set* of rank order collectives for a population P.

Definition 7. *Partition: A partition* T *of a population* P *is a set of collectives such that all individuals in* P *are in exactly one collective* $\mathrm{C_j}$. *A partition will be either weak or strong as defined below.*

Definition 8. *Strong Partition: A strong partition* $\mathrm{T_s}$ *of a population* P *is a partition such that each individual* $\mathrm{i} \in \mathrm{P}$ *can be assigned to one and only one collective* $\mathrm{C_j}$. *In other words, there is a unique partition* $\mathrm{T_s}$ *for the population* P.

Definition 9. *Weak Partition: A weak partition* $\mathrm{T_w}$ *of a population* P *is a partition such that each individual* $\mathrm{i} \in \mathrm{P}$ *could fit into multiple collectives, but is assigned to one collective* $\mathrm{C_j}$.

Definition 10. *Collective Belief Model: A collective belief model* $\mathrm{CBM} = (\mathrm{T}, \Phi)$ *is composed of a partition* T *containing* k *collectives* $\mathrm{C_j} \in \mathrm{T}$*, and each collective's collective belief* $\phi_{\mathrm{j}} \in \Phi$ *for a given set of options* O*.*

4.3.1 Rank Order Relations

The rank order collectives are formed from the pairwise preference order relations defined in Section 3.1. In particular, the *society* relations are relevant to collectives, which are generalizations about a society, population, or in this case, a collective (C):

1. xPy: C prefers x to y. xPy is equivalent to $x < y$.
2. xIy: C is indifferent to x or y. xIy is equivalent to $x \leq y$ or $y \leq x$.
3. xRy: C prefers x to y or is indifferent to them (is P or I). xRy is equivalent to $x \leq y$.

A preference order among more than two options can be made up of any combination of the R, P and I relations. The *rank order* for each rank order collective will be a preference order formed in this manner. The following is a list of characteristics of specific relations.

- By the definition of collective, it is not possible for a collective to be both indifferent to x and y and strictly prefer x over y, therefore between any two options x and y, relations I and P are mutually exclusive.
- I and P are *exhaustive*, meaning that a preference order made from a combination of I and P relations can represent any relationship between two or more options.
- Given the previous characteristic of the I and P relations, all individuals in a population can provide a rank order over a set of options O. Therefore a partition of an arbitrary population can be discovered using rank orders formed from the above relations that will include *all* individuals in the population.
- A strong partition can be formed from rank orders defined by the P and I relations since the relations are mutually exclusive and exhaustive. Given a set of options O, an

individual i will only have one rank order R_i over O formed from the P and I relations. Thus that individual can be placed in one and only one collective defined by the P and I relations.

- If the R relation is used to define a pairwise relationship in a population, then only a *weak* partition can be created. For example, if x and y represent two possible options, and a collective C believes that both are equally likely ($P(x) = 0.5$ and $P(y) = 0.5$) then the relations xRy and yRx could both hold. In other words, the collective rank order for a collective C could be xRy or yRx, and therefore a partition made using the R relation would not be unique.
- The number of collectives in a partition is dependent on the size of the set of options (O) and the set of relations allowed (Ψ). If $r = |O|$ and $\psi = |\Psi|$, then there will be $r! * \psi^{r-1}$ possible collectives, since there are $r!$ permutations of r options and ψ^{r-1} combinations of relations between the r options in each rank order.
- If Ψ is the set of allowed relations and O is the set of options, then $\Gamma(O, \Psi)$ is the set of $r! * \psi^{r-1}$ possible rank order relations.

4.3.2 Collective Choice Function

Finally, a social choice function that discovers a partition for a given population is defined as follows.

Definition 11. *Collective Choice Function: Given a population* P *and a set of* r *options* O*, over which each individual* i *in* P *provides a rank order* $\mathrm{R_i}$*:*

i Select Ψ*, the set of rank order relations allowed. If only* P *and/or* I *are selected, then a strong partition will be formed. If* R *is allowed, then a weak partition will be formed.*

ii Separate a population P *into* $\mathrm{m} \leq \mathrm{r}! * \psi^{\mathrm{r}-1}$ *groups, each group representing a unique ordering* $R_j \in \Gamma(O, \Psi)$*, such that each individual* i *in group* $\mathrm{G_j}$ *provides a rank order* $\mathrm{R_i} = \mathrm{R_j}$ *over the* r *options in* O*.*

iii Each group $\mathrm{G_j}$ *becomes a collective* $\mathrm{C_j}$ *defined by its collective rank order* $\mathrm{R_j}$*.*

iv In a Bayesian environment, use an opinion pool function (e.g. arithmetic or geometric mean) to compute each collective C_j's collective belief.

4.3.3 Rational Social Choice Properties

The following properties defined by Kenneth Arrow [1, 2] were introduced in Section 3.1.2 and are reiterated here for reference.

1. *Universal Domain (UDP)*: All preference orders over a set of options O are allowed
2. *Completeness (CP)*: Social choice function returns an order that includes all relevant alternatives
3. *Transitivity (TP)*: if a society S prefers A to B and B to C then S prefers A to C (also replace "prefers" to "is indifferent to")
4. *Pareto optimality (should be at least weekly Pareto optimal)*:
   (a) Weak Pareto principle (WP). For all x and y, if xP_iy for all i, then xPy:
   (b) Strong Pareto principle (SP). For all x and y, if xR_iy for all i, and xP_iy for some i, then xPy:
5. *Independence of irrelevant alternatives: (IIP)* If the society has a preference order $xPyPz$, then xPy and yPz also hold. Also, a change in preference from wPz to zPw does not affect the preference order $xPyPz$.
6. *Non-dictatorship and non-imposition (NDIP).* There is no dictator. Individual i is a dictator if, $\forall x$ and $\forall y$, $xP_iy \rightarrow xPy$. Non-imposition means that no order has been pre-determined for any individual.

Bayesian social choice theorists Hylland and Zeckhauser added the following properties for Bayesian social choice functions [31].

1. *Unanimity (UP):* If each individual i has the same probability estimate P_i, then the aggregate P_o should also be equivalent.

2. *Bayesian group decision (BGP):* The aggregate is a function of only the probability estimates (P) and utilities (U) supplied by the group of interest.

4.3.4 Enabling a Rational Social Choice Using Collectives

This section shows that the *collective* choice function upholds the RSC properties by creating a partition T, such that each property holds for each collective in T. The collective choice function creates a partition and model based on a snapshot of a situation, composed of the options O, a population P, and each member of P's beliefs about O. No claims are made about the rationality of human behavior over time or in different situations, only about the the mathematical properties of a snapshot of the beliefs provided at a single point in time. The properties are

Proposition 2. *A rational social choice can be formed from a rank order collective: The properties defined in Section 3.1 hold for a rank order collective as defined in Definition 3.*

Proof 2. *Each property is discussed separately below:*

1. *Universal Domain (UDP): If all preference (rank) orders were not allowed, then there would be some rank order $R \in \Gamma(O, Psi)$ that would not be allowed to form a collective. However, each of the $r! * \psi^{r-1}$ orders in $\Gamma(O, \Psi)$ could be submitted by the members of a population P. By definition of the collective choice function (Def. 11), any unique rank order submitted by an individual will form a collective. Thus all rank orders over O can be represented by a collective within a partition and no rank order is disallowed.*

2. *Completeness (CP): If completeness did not hold, then some option $o \in O$ would not be included in a cluster's rank order. As defined in Def. 11, each collective C_j is defined by its ordering R_j over all options $o \in O$, thus all o are included in R_j.*

3. *Transitivity (TP): Since there is only one rank order R_j per collective over a set O of r options, and since a rank order is partial order (by Def. 1), then transitivity holds for each collective by the properties of a partial order.*

4. *Weak Pareto principle (WP): The weak Pareto principle is upheld by the definition of rank order collective. For all individuals $i \in C_j$, if the rank order for $i = R_i$ and the rank order for $C_j = R_j$ then $R_i = R_j$. Thus, for all options x and y and all $i \in C_j$, if xP_iy then xP_jy for the collective C_j. Seidenfeld, et al. [56] also state this finding when there is consensus among the preference order of expected utilities.*

5. *Independence of Irrelevant Alternatives (IIP): If IIP did not hold, then the consideration of a new option $o' \notin O$ would change the rank order between the options $o \in O$, or the removal of an option $o'' \in O$ would change the rank order for the options $O' = O - o''$. A partition T as defined in Def. 7 is based on a model formed from a given set of options O and a population P's beliefs about those options. If a new set of options O' is given that is either missing an option that was in O or has an additional option not in O, then a new model will be formed, from which a new partition T' will be discovered. T' is not guaranteed to be the same as T, however T' will also not affect the original partition T nor the rank orders that represent the collectives in T. In other words, a partition and its collectives is dependent only on the set of options and beliefs about the options used to form it. Thus, any rank orders from T will be maintained, regardless of the rank orders in T'.*

6. *Non-dictatorship (NDIP): By the definition of rank order collective (Def. 3), if an individual i is in collective C_j then $xP_jy \rightarrow xP_iy$. However, if i supplied a different rank order such that his belief $R_i \neq R_j$, then i will no longer be in C_j, and would instead be in a C_k such that $R_i = R_k$. Thus, $xP_jy \rightarrow xP_iy$, does not imply that the complement $xP_iy \rightarrow xP_jy$ applies.*

7. *Unanimity (UP): This property states that the aggregate (or opinion pool) function used to combine probabilistic beliefs must result in $P(X) = p$ if all i in C_j provide $P(X) = p$. This property holds for both the arithmetic and geometric means.*

8. *Bayesian group decision (BGP): This property states that the aggregate of a group is a function of only the beliefs and utilities supplied by the individuals in a group. By the definition of collective belief (Def. 5), the collective belief ϕ_j of a collective C_j is an aggregate of only the beliefs and utilities supplied by the members of C_j.*

4.4 Game Theoretic Concepts

A preference relation, D, for dis-prefers is added that is the complement of P: $xPy = yDx$. Given alternative options x, y and individuals i, j: xD_iy: Person or group i dis-prefers x to y $(= yP_ix)$.

4.4.1 Game theoretic definitions

Definition 12. *Weak Pareto optimal solution: A solution or strategy* s *is a weak Pareto optimal solution if there is no other strategy* $\mathrm{s'}$ *such that all game players* $\mathrm{a} \in \mathrm{A}$ *prefer* $\mathrm{s'}$ *to* s *when only strict preference is considered. Mathematically,* s *is weakly Pareto optimal if and only if there exists no* $\mathrm{s'}$*, s.t.* $\forall \mathrm{a} \in \mathrm{A}, \mathrm{u_a}(\mathrm{s'}) > \mathrm{u_a}(\mathrm{s})$ *[56, 37].*

Definition 13. *Strong Pareto optimal solution: A solution or strategy* s *is a strong Pareto optimal solution if there is no other strategy* $\mathrm{s'}$ *such that all game players* A *are either indifferent to* $\mathrm{s'}$ *and* s*, or prefer* $\mathrm{s'}$ *to* s*. Mathematically,* s *is strongly Pareto optimal if and only if there exists no* $\mathrm{s'}, \mathrm{s.t.} \forall \mathrm{a} \in \mathrm{A}, \mathrm{u_a}(\mathrm{s'}) > \mathrm{u_a}(\mathrm{s})$*, and there exists some* $\mathrm{j} \in \mathrm{A}$ *s.t.* $\mathrm{u_j}(\mathrm{s}) > \mathrm{u_j}(\mathrm{s'})$*. [56, 37].*

Definition 14. *Normal form game: From [37], a normal-form game is tuple* $(\mathrm{N}, \mathrm{D^*}, \mathrm{u})$*, where:*

- *N is a finite set of n players;*
- $\mathrm{D^*} = \mathrm{D_1} \times \mathrm{D_2} \times ... \times \mathrm{D_n}$*, such that* $\mathrm{D_i}$ *is the finite set of actions (or decisions) that player* i *can take on;*
- $\mathrm{u} = (\mathrm{u_1}, \mathrm{u_2}, ...\mathrm{u_n})$*, where* u *is a real-valued utility function for player* i

A normal form game is typically represented by an $n-$dimensional matrix, where each dimension is the vector of actions D_i for each agent i. Each cell represents the utility for each of the players, where each agent takes on the actions specified by the cell.

Definition 15. *Strategy: A strategy* $\mathrm{s_i}$ *is the action that a player* i *will take given the actions that the other players take (*$\mathrm{s_{-i}}$*) [37].*

Definition 16. *Strategy profile: A strategy profile* $\sigma = (\mathrm{s}_1, \mathrm{s}_2, ...\mathrm{s}_\mathrm{n})$ *is the set of actions that each player will take given the actions that the other players takes [37].*

Definition 17. *Nash equilibrium: A strategy profile* σ *is a Nash equilibrium if, for all players* i*,* $\mathrm{s}_\mathrm{i} \in \sigma$ *is the best action (maximizes* u_i*) given the actions of the other players* ($\mathrm{s}_{-\mathrm{i}}$*) [37].*

Definition 18. *Minimax and maximin strategies: A minimax strategy for players* $-\mathrm{i}$ *against a player* i *is one such that players* $-\mathrm{i}$ *act to minimize the maximum payoff of* i*. A maximin strategy for a player* i *is the complement of the minimax strategy, it is the strategy* s_i *for player* i *that maximizes* i'*s minimum payoff in the event that players* $-\mathrm{i}$ *play a minimax strategy [37].*

4.4.2 Multi-agent influence diagrams

Definition 19. *A Multi-agent Influence Diagram (MAID), defined by Koller and Milch [35] is an influence diagram (Bayesian decision network) that represents the beliefs, actions of preferences of multiple-agents.*

MAIDs can be used to represent games in which each agent is a player. In MAIDs, the following notation is used:

i. D: A decision, represented by a decision node in a MAID

ii. A: The set of agents, $a \in A$, represented by a MAID

iii. D_i: The set of decisions that agent i will take on

iv. $Pa(D)$: The set of variable and decision nodes that affect a decision D

v. σ: A strategy profile for all agents in a game, indicating the decision option that each agent will take given the decisions that the other agents will take.

4.4.3 Graph Theoretic Concepts

These definitions will be used in chapter 7

Definition 20. *Node width: The width of a node* $\mathrm{w(d)}$ *in a graph is the number of nodes preceding it in a tree decomposition of a graph [14].*

Definition 21. *Width of an ordering: The width of an ordering is the maximum node width given an ordering* d *of the graph in a tree decomposition [14].*

Definition 22. *Treewidth or induced width: Treewidth or induced width,* $\mathrm{w} * (\mathrm{d})$ *or* w*, is the width of the ordering in an optimal tree decomposition of a graph, where an optimal decomposition minimizes treewidth [7, 14].*

Definition 23. *Poly-tree: A poly-tree is a graph in which the unlabeled graph (directed edges converted to undirected edges) contains at most one path between any two nodes* i *and* j *[50].*

Chapter 5

Collective Belief Aggregation

This chapter provides an overview of the collective belief aggregation approach to Bayesian belief aggregation and social decision-making. Two significant limitations of existing belief and preference aggregation approaches are that they (1) can form consensus models that under-represent divergent beliefs and (2) they may result in an irrational *social choice* solution in the presence of divergence.

This chapter will demonstrate two aggregation approaches that address these limitations. Section 5.1 introduces an example that will demonstrate the potential of utilizing this approach for policy-making. Section 5.2 revisits the irrationality results described in Chapter 3 with a simple but realistic decision network. I first show that averaging contradicting beliefs can cause a social choice solution to be selected over a more preferred solution. I then show that a single individual can sway the social choice solution to his benefit, against the preferences of all other individuals. Section 5.3 will highlight the differences in behavior of prior and posterior aggregation. Section 5.4 presents an approach that partitions a population into clusters before aggregation. The consensus clusters increase the representation of diverging beliefs. Section 5.5 introduces a more refined aggregation approach that partitions a population into collectives whose consensuses will uphold the rational social choice properties defined in Figure 3.1 and Section 3.1.3. By forming collectives whose members agree on the relative likelihood or desirability of an outcome, the two situations described in Section 5.2 can be avoided. Finally, Section compares prior and posterior collective belief aggregation and shows that only posterior aggregation can guarantee collectives that uphold the RSC properties.

5.1 A Motivating Example

Policy-making and politics are fields that exemplify the challenges and rewards of combining diverse and often conflicting beliefs. Diversity can result in harmonious policy decisions that consider everyone's point of view, or it can cause polarization and increase feelings of detachment from a community. My goal is to enable policy-making that represents all significant beliefs, and enables groups to objectively cooperate to achieve a goal without attempting to force consensus. To this end, the examples in this dissertation address various political and policy-making situations.

I begin with a situation that can be generalized to other policy decisions. Suppose there is a logging interest that would like to clear a forest to sell the lumber or to sell the land for development. The public may be opposed to the logging, due to the effects it has on the environment, or they may support the loggers because it will bring jobs and commerce into the area. The public could also be ambivalent, in which case they do not support or oppose the loggers, or they see equivalent benefits of both situations. If the government were to get involved in the decision of whether to log or not, policy could be developed that is either pro-logging interest or pro-environment.

The decision network shown in Figure 5.1 represents a group decision about whether a vote should be held to introduce new policy. If no vote occurs then no policy will be enacted. Each individual in the group has a utility for each policy decision. Each individual also has a belief in which policy would win if a vote is held. In Figure 5.1, the oval represents the conditional probability of $Policy$ given the $Vote$ decision, represented by $P(Policy|Vote)$. Policy will be one of $[E, L, N]$, where E= environmental, L= logging, N = none. The utility of each policy option will be in the range $[-2, 2]$. Inference on the network will determine the expected utility of the decision options— to vote or not to vote.

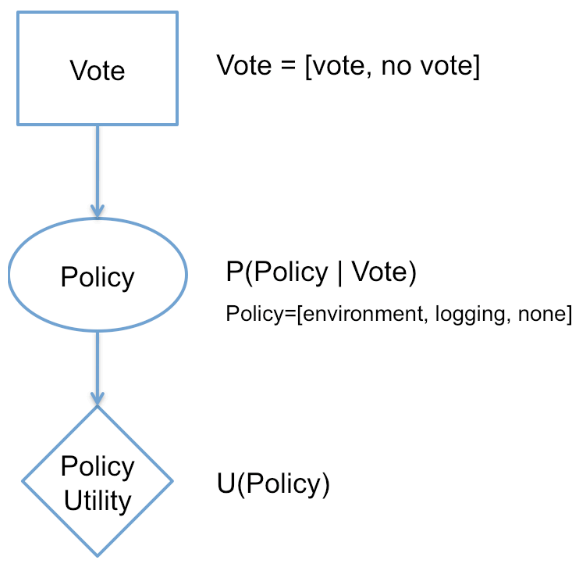


Figure 5.1: A decision network representing a decision to put a new policy to vote.

5.2 Irrational Social Choice Results Revisited

This section revisits the impossibility theorems discussed in Sections 3.2.2 and 3.3.2 using the network in Figure 5.1.

5.2.1 A Non-Pareto Optimal Solution

If the individuals in a group making the vote decision have opposing beliefs and utilities, it is possible that the aggregate of their beliefs will result in a non-Pareto optimal solution (as defined in Figure 3.1). In other words, all individuals prefer one vote option, but the social choice is the other option. This is demonstrated in the following example. Suppose there are two individuals i_1 and i_2. The individuals' conditional probability tables for $P(Policy|Vote)$ are shown in Figure 5.2. In this case, if there is a vote the policy will be either E or L. If no vote is held, then there will be no policy, therefore $P(Policy = N|Vote = NoVote) = 1.0$. The individuals' utilities for each policy decision are shown in Table 5.1.

The results of applying an arithmetic mean to find the consensus probability distribution, $P_0(Policy|Vote)$, is shown in Table 5.2. The consensus on policy utilities, $U_0(Policy)$, is shown in Table (Y). The geometric mean will result in similar values.

$P_1(Policy\|Vote)$		Vote	
		$Vote$	$NoVote$
Policy	E	0.75	0.0
	L	0.25	0.0
	N	0.0	1.0

$P_2(Policy\|Vote)$		Vote	
		$Vote$	$NoVote$
Policy	E	0.25	0.0
	L	0.75	0.0
	N	0.0	1.0

Figure 5.2: Conditional probabilities $P_i(Policy|Vote)$ for two individuals i_1 and i_2.

$U_1(Policy)$		
E	L	N
1.0	-1.2	0.0

$U_2(Policy)$		
E	L	N
-1.2	1.0	0.0

Table 5.1: Utilities $U(Policy)$ for policy options for two individuals i_1 and i_2.

$P_0(Policy\|Vote)$		Vote	
		$Vote$	$NoVote$
Policy	E	0.5	0.0
	L	0.5	0.0
	N	0.0	1.0

Table 5.2: Consensus conditional probability table for $P_0(Policy|Vote)$ computed using the arithmetic mean.

$U_0(Policy)$		
E	L	N
-0.1	-0.1	0.0

Table 5.3: Consensus utilities $U_0(Policy)$ for the policy options computed using the arithmetic mean.

The expected utility of each decision option is found using the following formula:

$$EU(Vote) = P(Policy|Vote)U(Policy) \tag{5.1}$$

Table 5.4 compares the results of applying the formula to each individual's beliefs and utilities with the results of applying the formula to the consensus beliefs and utilities. The best option for each individual and the consensus is shown in bold. We can see that the consensus favors the opposite decision option as *both* individuals. In other words, the consensus option is not Pareto optimal. While in this situation the group was composed of only two individuals, the same outcome will arise if all members of a group prefer one option, but their consensus results in the same or similar values as Tables 5.2 and 5.3.

	$Vote$	$NoVote$
$EU_1(Vote)$	**0.45**	0.0
$EU_2(Vote)$	**0.45**	0.0
$EU_0(Vote)$	-0.1	**0.0**

Table 5.4: Expected utilities of each individual and their consensus. The options with the highest expected utility are shown in bold. Note that the consensus results in the option that neither individuals preferred.

5.2.2 Dictatorship

Section 3.3.2 demonstrated how a single individual can skew the consensus solution in his favor using quantitative beliefs and utilities. According to Arrow's axioms for preference aggregation, this is considered a dictatorship [1]. The more general phenomenon that is occurring is that the mean of a set of quantitative values can be skewed by a small number of highly divergent values. In addition to the dictatorship situation, this can cause the consensus to "lose" the representation of some of the values. In other words, as a set of values becomes more divergent, the set's consensus will become less similar to the original values.

The next example demonstrates a dictatorship situation using the decision network in Figure 5.1. Suppose that the table on the left of Figure 5.3 contains the conditional probabilities for a group g of three individuals who all happen to have the same beliefs. The table on the right of Figure 5.3 contains the probabilities for an individual d who waits to supply his values until the others supply theirs. Since he can see their values, he can compute what he needs to provide in order to skew the vote decision in his direction. The consensus of the group composed of $g \cup d$ is shown in Table 5.5. The utilities $U(Policy)$ are identical for all individuals and are shown in Table 5.6.

$P_g(Policy \mid Vote)$		Vote	
		$Vote$	$NoVote$
	E	0.4	0.0
Policy	L	0.6	0.0
	N	0.0	1.0

$P_d(Policy \mid Vote)$		Vote	
		$Vote$	$NoVote$
	E	0.9	0.0
Policy	L	0.1	0.0
	N	0.0	1.0

Figure 5.3: The table on the left contains conditional probabilities for a group g of three individuals with identical beliefs. The table on the right shows the conditional probabilities for a single individual d.

$P_0(Policy \mid Vote)$		Vote	
		$Vote$	$NoVote$
	E	0.525	0.0
Policy	L	0.475	0.0
	N	0.0	1.0

Table 5.5: The consensus conditional probabilities of the group composed of $g \cup d$.

$U_0(Policy)$		
E	L	N
1.0	-1.0	0.0

Table 5.6: Consensus utilities $U_0(Policy)$ for the policy options computed using the arithmetic mean.

Table 5.7 shows the expected utility for the group g, the dictator d and their combined consensus computed using equation 5.1. Again the best option for the group or individual is shown in bold. We see that in this situation, the dictator is able to flip the preference of the other individuals by a slim margin.

	$Vote$	$NoVote$
$EU_g(Vote)$	-0.2	**0.0**
$EU_d(Vote)$	**0.8**	0.0
$EU_0(Vote)$	**0.05**	0.0

Table 5.7: Expected utilities of a group $EU_g(Vote)$ of three individuals, an individual $EU_d(Vote)$ and their consensus $EU_g(Vote)$. Individual d is able to sway the decision in his favor by a slim margin.

5.3 Prior Versus Posterior Aggregation

Section 3.3.1 introduced some distinctions between prior and posterior aggregation. To review, prior aggregation occurs when the beliefs for each variable or utility in a network are aggregated before inference is performed on the network. In other words a consensus *network* is formed. Posterior aggregation occurs when inference is run on each individual's network separately, and the results of inference are then aggregated. A "belief" can represent either an *a priori* probability estimate for a variable in the Bayesian network, or a conditional probability estimate for a child given the probability of its parents. It can also refer to an inferred belief, or the posterior probability of an outcome.

Prior and posterior aggregation do not always result in the same consensus values or even the same solutions. Both examples in Section 5.2 used prior aggregation. If posterior aggregation is used on the example in Section 5.2.1 the consensus solution is the Pareto optimal option (the average of the individual's expected utility for $Vote$ is 0.45). However the results of the dictatorship example in Section 5.2.2 are the same with posterior aggregation. In fact, the dictatorship example could occur with a range of group sizes and expected utilities for the vote decision. If we assume that the expected utility for $Vote = NoVote$ is always 0 for all individuals, any sets of values for which the Eq. 5.2 is met will cause a single individual to override the beliefs and preferences of the other individuals. If there are x individuals in a group g who lean towards $Vote = NoVote$ and one "dictator" who leans the other direction, then the following equation holds.

$$\begin{aligned}\frac{1}{x+1}\left(\sum_{i=1}^{x} EU_i(Vote = Vote) + EU_d(Vote = Vote)\right) > 0 \\ \sum_{i=1}^{x} EU_i(Vote = Vote) + EU_d(Vote = Vote) > 0 \\ \sum_{i=1}^{x} EU_i(Vote = Vote) > -EU_d(Vote = Vote)\end{aligned} \tag{5.2}$$

While using posterior aggregation may avoid some irrational social choice results that are specific to prior aggregation, it will not eliminate all issues. The next sections will introduce an approach that addresses the challenges described thus far in this chapter.

5.4 Clusters of Consensus

The first approach to partitioning a group of individuals uses a clustering algorithm to detect groups of individuals with similar beliefs. The following example will show that clustering and then aggregating the beliefs within each cluster forms a set of consensus models that are more representative of the original population's beliefs than a single consensus. The k-means clustering algorithm is used to discover the clusters. The k-means clustering algorithm is simple, commonly used, and does well on linearly separated data. Given an input parameter k, the algorithm works as follows:

1. Select k random points as cluster centers
2. Assign each point to its closest center using Euclidean distance metrics
3. Compute the center of each cluster and re-assign the center
4. Repeat until clusters have stabilized

The input to the clustering algorithm is a matrix containing a row for each individual in a population. In this case, the population contained 10,000 individuals with randomly generated beliefs. The row is composed of the individual's beliefs used to form the network in Figure 5.1. Figures 5.4-5.6 show histograms of expected utilities for the $Vote$ option, comparing the original population distribution with a single mean ($k = 1$) and cluster means ($k = 2, 4, 8$). The distribution of the original data shows the percent of individuals that fit into each bin. Each bar that represents a k value shows the percent of individuals that fit in each bin if each individual were to provide the mean of their cluster. Each consecutive graph shows increased divergence in the original expected utilities, measured by standard deviation.

The *Kullback-Leibler divergence measure* measures the difference between two probability distributions P and Q [36]. A low divergence means that the distributions have a high degree of overlap or similarity. Thus, it can be used to measure how well a consensus represents a population's beliefs. In this case, the K-L divergence of the cluster consensus distribution from the distribution of the population's expected utilities is compared with the K-L divergence of the single consensus (mean of all expected utilities) from the population's distribution. The K-L divergence measure is defined formally in Section 4.2. Equation 4.1 was used in which X contains one variable— the expected utility of $Vote$. The K-L divergence of each consensus distribution from the original distribution is shown in the bottom right of each graph in Figures 5.4-5.6.

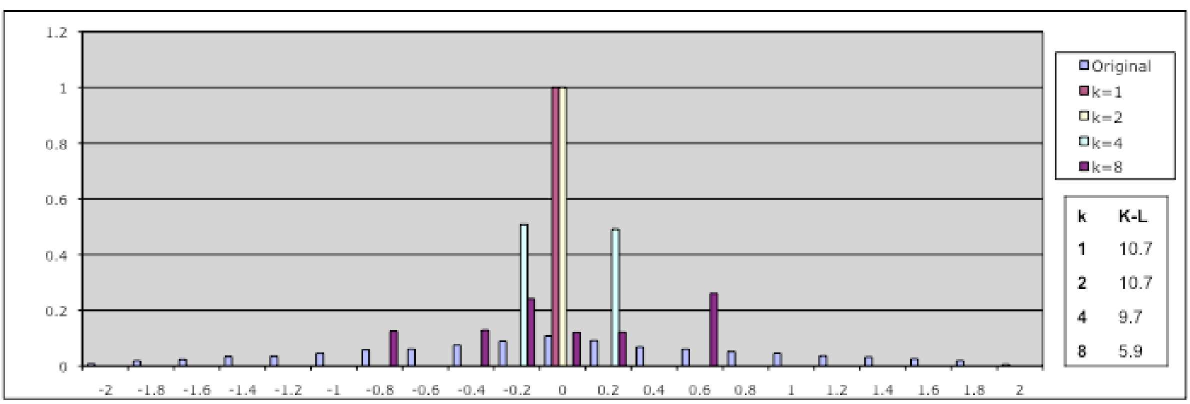


Figure 5.4: A histogram of expected utilities comparing the original population distribution with a single mean (k=1) and cluster means (k=2,4,8). The expected utilities in this graph had a standard deviation of 0.88. The x axis shows the range of expected utility: [-2.0, 2.0].

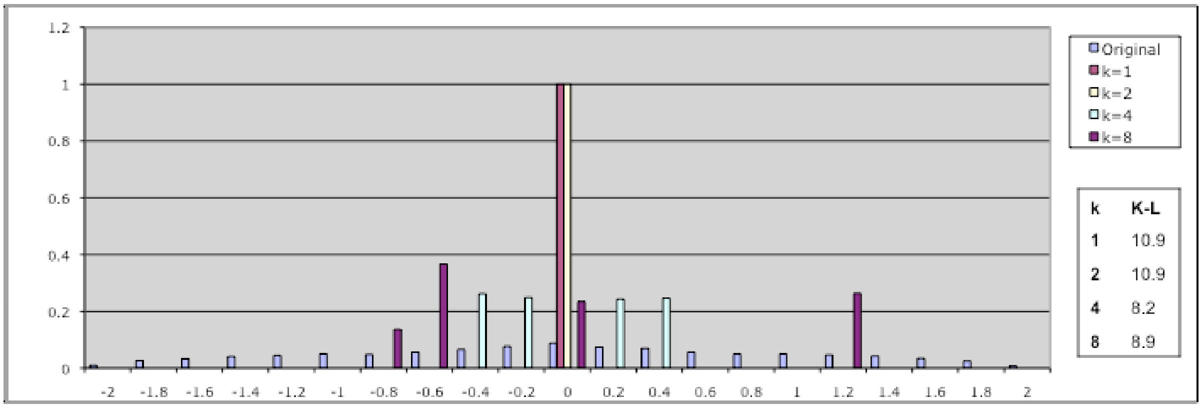


Figure 5.5: A histogram of expected utilities comparing the original population distribution with a single mean (k=1) and cluster means (k=2,4,8). The expected utilities in this graph had a standard deviation of 0.97. The x axis shows the range of expected utility: [-2.0, 2.0].

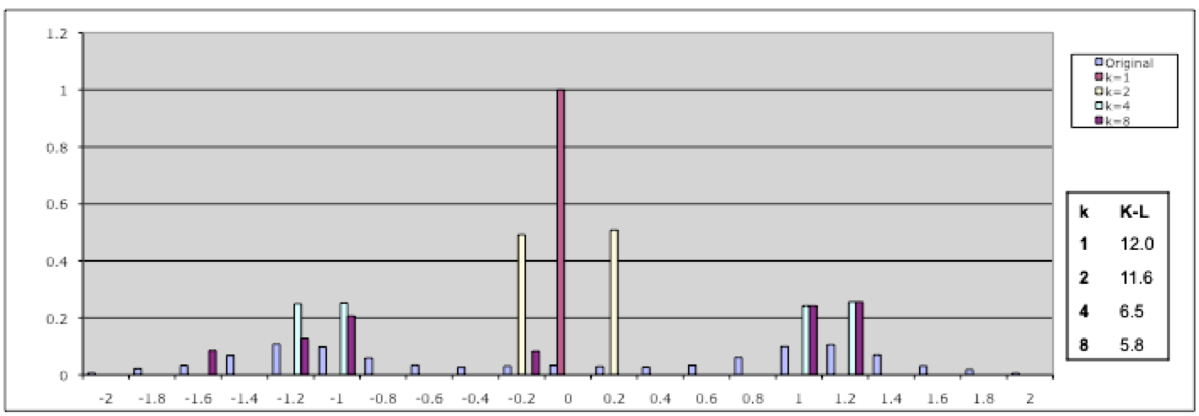


Figure 5.6: A histogram of expected utilities comparing the original population distribution with a single mean (k=1) and cluster means (k=2,4,8). The expected utilities in this graph had a standard deviation of 1.11. The x axis shows the range of expected utility: [-2.0, 2.0].

All graphs show that clustering reduces the K-L divergence of the consensus distribution. Figure 5.6 illustrates that clustering works particularly well for modeling divergent beliefs. In two of the three graphs a larger number of clusters decreased the K-L divergence. However, K-L divergence of $k = 8$ was higher than $k = 4$ in Fig. 5.5. This indicates that in some cases, no additional representational gain can be made by continuing to increase k. K-L divergence could be in this manner to determine the optimal number of clusters.

5.5 Collective Belief Models

The previous section demonstrated a first approach at partitioning a population before aggregation. While the clusters of consensus increased the representation of the original population, there is no guarantee that individuals placed in clusters would have selected the consensus solution of their cluster. In other words, while the mean of a cluster could prefer the $Vote$ option, an individual in the cluster could actually prefer the $NoVote$ option.

This section introduces the *collective belief aggregation* approach to partition a population into groups that agree on the relative likelihood or desirability of an outcome, determined by the *Bayesian rank order* of the posterior probabilities or expected utilities. The Bayesian rank order is formally defined in Section 4.3. In summary, in a decision network it is the preferred order of the decision options based on the decreasing order of the options' expected utilities. For example, if the expected utility for the $Vote$ option is -1.2 and the expected utility for the $NoVote$ option is 0.0, then the decreasing order of their expected utilities is $[0.0, -1.2]$ and the Bayesian rank order (or just rank order) is $NoVote, Vote$. In a rank order, the most preferred option is first, followed by the less preferred options and ending with the least preferred option. Using the preference relations defined in Section 3.1.1, the relation would be $NoVote\mathbf{P}Vote$, meaning $NoVote$ is preferred over $Vote$. A Bayesian rank order could also include indifference if the expected utilities of two options are equivalent.

The new partitioning approach forms *collectives* from individuals who have the same Bayesian rank order of the decision options. Collectives are also formally defined in Section 4.3. In the previous examples all individuals with the rank order $NoVote\mathbf{P}Vote$ would be placed in one collective and all the individuals with the rank order $Vote\mathbf{P}NoVote$ would be placed in another collective. If all individuals have a strict preference, then in the vote decision there would only be two collectives. However, the number of collectives is dependent on the number of decision options. If there are d decision options, then there are $O(d!)$ possible collectives. The number of actual collectives in a population is the number of unique rank orderings that the individuals in the population provide.

Since all members of a collective provide the *same* rank ordering, the consensus (mean) of the collective will also have the same rank ordering. The proof for this is demonstrated by

Proposition 1 in Section 4.3. The consensus of each collective, called the *collective belief*, is the mean of all members' expected utilities. The benefit of a collective maintaining the same rank ordering as its members is that *no individual member can prefer a different solution than the collective's consensus solution*. This fact means that the aggregate of a collective will uphold the rational social choice properties defined in Section 4.3, Definition 6 (demonstrated in Proposition 2).

The collective belief aggregation approach partitions a population into collectives and finds the aggregate of each collective. The result is that there will be multiple consensus solutions if there is any disagreement on the rank order of the decision options. If all the consensus solutions happen to be the same, then there will be one solution that suits the whole population. Otherwise there will be competing solutions that will need to be resolved through other means (such as through the game theoretic analysis discussed later in this chapter). In summary, a single consensus approach may result in an irrational social choice when there is a stalemate or significant divergence in belief. In these situations, the collective belief aggregation approach will result in a *set* of rational social choice solutions.

5.5.1 Revisiting Non-Pareto Optimal Solutions

This section revisits the irrational social choice result demonstrated in Section 5.2.1 to show that the collective belief aggregation approach returns a Pareto optimal solution when the single consensus aggregation approach returns a non-Pareto optimal solution. The example in Section 5.2.1 used prior aggregation, therefore this example will also use a form of prior aggregation (called incremental aggregation) that is discussed in detail in Section 7. In summary, collectives can be formed for the prior distribution of any variable in a decision network. During inference multiple variables are combined through propagation. A new collective will be formed from the intersection of each variable's collectives. For example, if there are two collectives for $U(Vote)$, and two collectives for $P(Policy|Vote)$, then a new set of collectives for $EU(Vote) = P(Policy|Vote)U(Vote)$ will be formed from the intersections of the original variables' collectives.

Revisiting the example in Section 5.2.1 using incremental collective belief aggregation, each individual will be placed in her own collective for each variable. This will also result in each individual being placed in her own collective for the expected utility. The preferred solution for each collective is then the $Vote$ option, which is the Pareto optimal solution. It is important to note that while the example in Section 5.2.1 results in a Pareto optimal solution using incremental collective belief aggregation, this is not the case in general. Only posterior collective aggregation can guarantee a rational social choice, as discussed in Section 5.5.3.

5.5.2 Revisiting Dictatorship

This section revisits the dictatorship example in Section 5.2.2. Regardless of whether prior or posterior aggregation is used, the collective belief aggregation approach will form separate collectives for the group of individuals who prefer the $NoVote$ option and the individual who prefers the $Vote$ option. In this case, each collective's solution has equal representation and the would-be dictator can no longer "flip" the result in his favor. The more general result of this observation is that all unique preference orders will be represented in the output of the collective belief aggregation approach. Therefore the quantitative aspects of averaging divergent values cannot change the structure of the collectives, although it can change the consensus values. However, since each collective maintains the rank order of its members, the relative preferences between options are always maintained.

5.5.3 Prior Versus Posterior Collective Belief Aggregation

As with traditional belief aggregation, collective belief aggregation can be preformed such that the collectives are discovered prior to inference or the collectives are discovered based on the posterior results of inference. Prior collective belief aggregation uses clustering to make an initial guess at how the collectives will form based on the population's expected utilities. However, prior aggregation does not have the same RSC guarantees that posterior aggregation has. In other words, it is possible that an individual will be placed into a collective whose posterior rank order is different than the consensus rank order. The distinctions between prior and posterior collective belief aggregation are discussed in more detail in Section 7.1.

Figures 5.7- 5.8 compare prior collective belief aggregation with posterior collective belief aggregation using similar belief data as the examples in Section 5.4. The bars for Prior and Posterior show the size and mean of each collective. For example, the Posterior bar on the left of Fig 5.7 represents the collective whose rank order was $NoVote\mathbf{P}Vote$. The collective contained just under half the population. The figures show that the collectives are a better representation of the original beliefs when the expected utilities are more divergent. The percent value in the bottom right of each figure shows the percent of individuals that prior collective aggregation placed in their correct collective. In other words, they have the same rank order of the vote options as their collective. The accuracy of the prior aggregation approach varies. In Fig. 5.8 prior aggregation was almost the same as posterior. This is probably due to having the majority of the population in two small but relatively tight peaks.

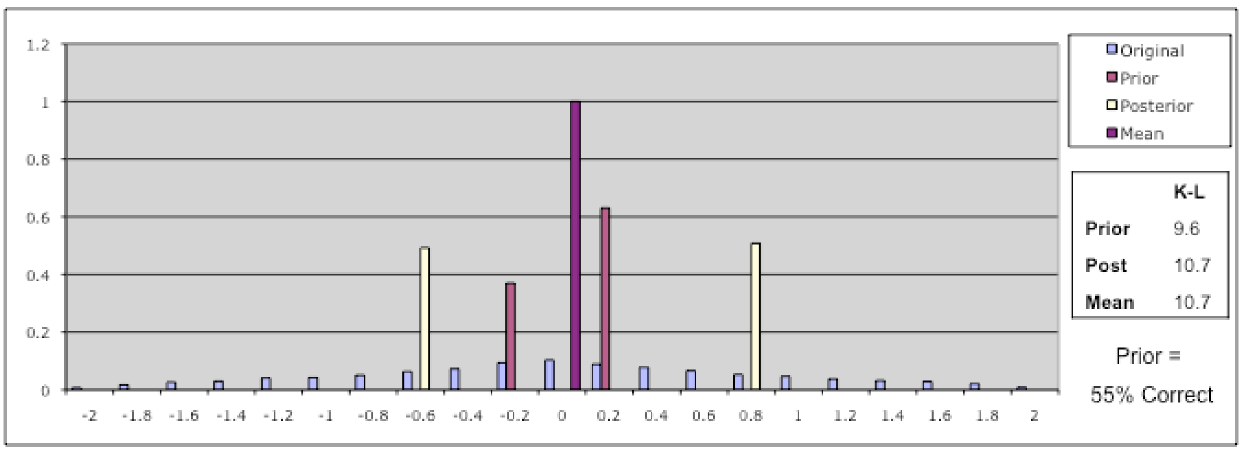


Figure 5.7: A histogram of expected utilities comparing the original population distribution with prior and posterior collective belief aggregation. The expected utilities in this graph had a standard deviation of 0.88. The x axis shows the range of expected utility: [-2.0, 2.0].

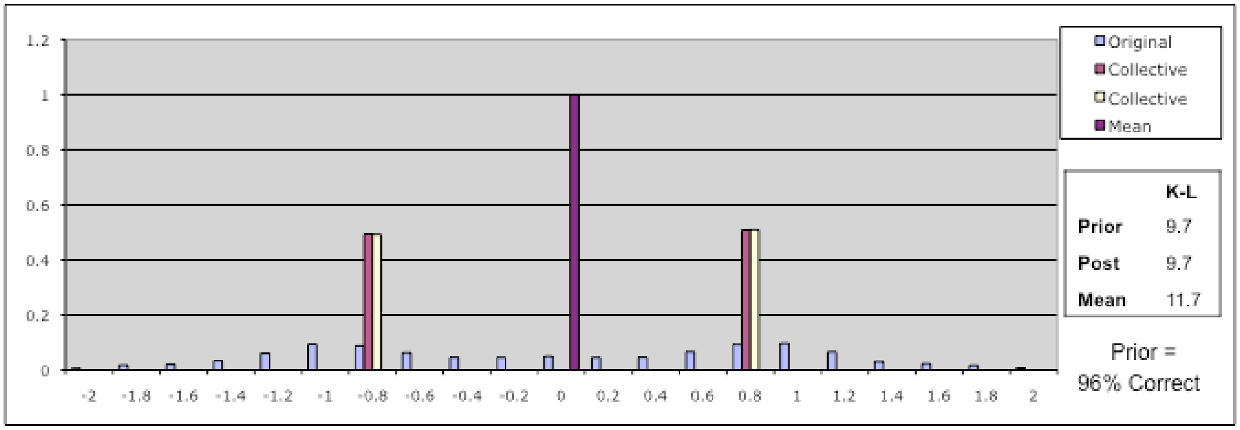


Figure 5.8: A histogram of expected utilities comparing the original population distribution with prior and posterior collective belief aggregation. The expected utilities in this graph had a standard deviation of 0.97. The x axis shows the range of expected utility: [-2.0, 2.0].

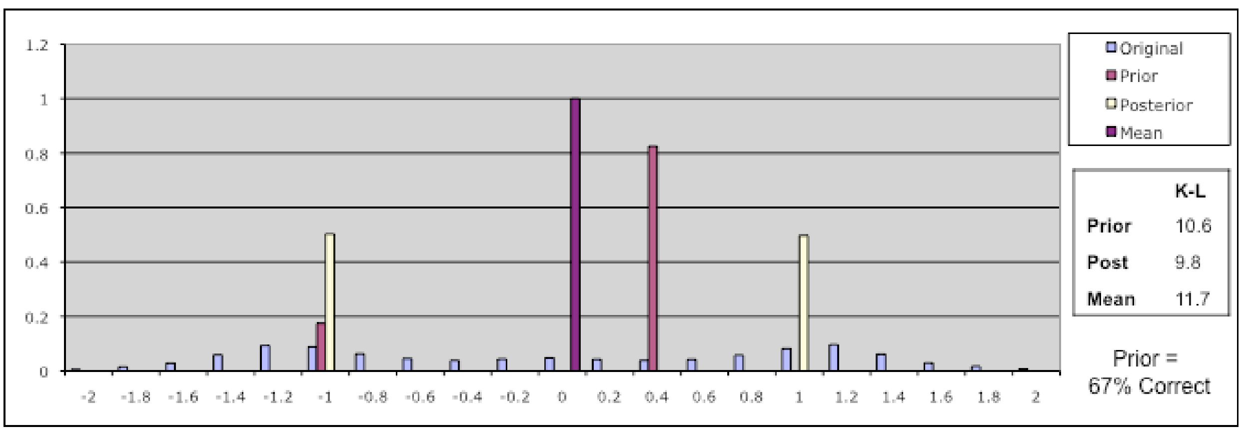


Figure 5.9: A histogram of expected utilities comparing the original population distribution with prior and posterior collective belief aggregation. The expected utilities in this graph had a standard deviation of 1.05. The x axis shows the range of expected utility: [-2.0, 2.0].

5.6 Summary

The examples in this chapter demonstrate two different approaches to partitioning a population prior to belief aggregation. Clustering greatly increased the representation of the original beliefs. It may be more appropriate if the data has many different "peaks" and if one is not concerned about upholding rational social choice principles or accurately placing individuals in collectives that matched their preferences. Posterior collective aggregation is more appropriate when one needs to guarantee that individuals are not placed in the wrong collective according to their preferences. This would be particularly important if a policy decision were to be made based on the members of a collective. In this case the members would want to be certain that they were placed in the collective that represents their preferences. The next chapter demonstrates how one can form a "super-agent" from each collective that will represent the collective in a decision-making game.

This chapter successfully demonstrated the first two objectives described in Section 1.3. The examples illustrate situations in which the individuals in a population had diverging beliefs, in some situations resulting in a solution that was in complete opposition to the preferences a subgroup of the population. Section B shows a situation in which the diverging opinions occur using beliefs elicited from humans. The chapter then showed that partitioning the population into clusters or collectives before aggregation resulted in more representative consensus models. The collective belief aggregation approach forms collectives whose aggregate will uphold rational social choice principles defined by social choice theorists. This is verified in Section 4.3.

Chapter 6

Applying Game Theoretic Analysis

The previous chapter introduced the collective belief aggregation approach and discussed how collectives can be formed from a population that uphold rational social choice properties. This chapter will demonstrate how to form decision-making games in which each player is a “super-agent” derived from a collective. Game theory has its own definitions for rational behavior, defined by its many solutions and equilibria [37]. Applying game theoretic analysis to the set of collectives will enable decision-makers to discover the rational solutions for a population in the context of competitive game theory. The this will allow the natural competition between individuals with divergent objectives to emerge. In this manner, game theoretic analysis, including finding Pareto optimal and Nash equilibrium solutions, can be applied within a large population.

This chapter first shows how the Pareto optimal solutions for a population can be found from the set of collective beliefs. I then expand the decision network in Figure 5.1 to demonstrate a more complex decision situation from which several collectives will emerge. The normal form games that represent the relative utilities and actions of the different collectives are then discussed. Finally, I demonstrate how multi-agent influence diagrams (MAIDs) [35] can be used to find Nash equilibrium solutions and the maximin strategy [37] in strategic situations. When a MAID represents the decisions of a large population, the agents may be super-agents that emerge from collectives.

6.1 Extracting the Pareto Optimal Solutions

The goal of social choice theory is to find one Pareto optimal solution for a population. In reality, there may be more than one solution that meets the Pareto condition. After partitioning a population into collectives, a logical next step in the social decision-making process might be to find the *set* of Pareto optimal solutions for the population. In the games described in this section, each "player" will be a super-agent whose belief is the collective belief of a collective. Each player's preferences are represented by its collective's shared rank order of the options. Discovery of the Pareto solutions involves eliminating the solutions that are preferred by no one (or no collective). In other words, there is always another solution that everyone prefers, therefore a rational social choice function will not select a non Pareto solution. Mathematically, a Pareto optimal solution is one in which no players can do better (have a higher utility) without another player doing worse [37].

An algorithm was developed to extract the Pareto solutions from a strong partition T_s, given the rank orders from the resulting collectives. A strong partition, as defined in Def. 8 in Section 4.3, is a partition that is derived from collectives defined using rank order relations containing only preference or indifference. The algorithm finds the set of options that uphold the strong Pareto condition by first finding those that do not. The algorithm is described in Appendix A.

6.2 An Expanded Decision Network

In the examples in Chapter 5 only two collectives were formed from the decision options; those who preferred the $Vote$ option and those who preferred the $NoVote$ options. More collectives will arise when multiple decision options are given. In Figure 6.1 the decision network in Fig. 5.1 has been expanded to include more decision options, variables and utilities. Suppose that the groups using the previous network decided to go ahead and put the policy to a vote. Each individual now must decide whether to vote for the pro-environment policy, pro-logging policy, or not to vote. The expected utility of each option is dependent on two variables and three utilities. The utility *Effort* represents the amount of "effort" an individual associates with voting. *Environment* and *Jobs* are the utilities assigned to the quality of environment and effect on jobs, consecutively, if the loggers decide to log. The logger's decision is represented by a variable node and is dependent on the policy that is enacted by the government. This policy is dependent on the votes of the community.

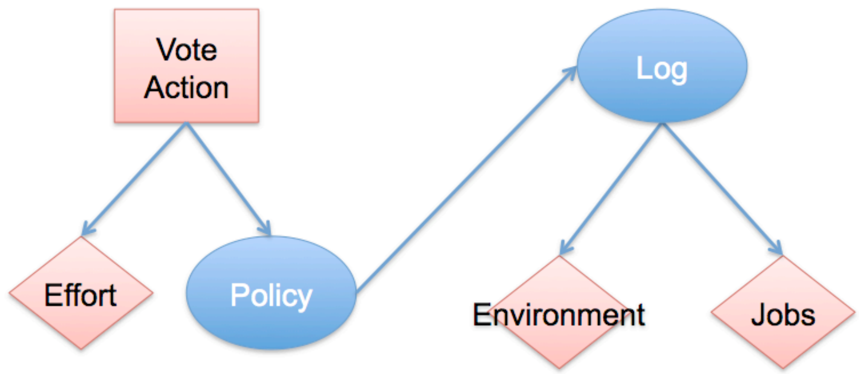


Figure 6.1: An expanded decision network representing a decision to vote for a specific policy based on the utility of logging and the utility of "making an effort". The utility of logging is a sum of the utilities for the environment and jobs given logging.

In this example the $VoteAction$ decision variable has three different options: E = vote pro-environment, L = vote pro-logging and N = no vote. Since there are three different options there will $3! = 6$ possible strict preference orderings. Therefore six possible collectives can emerge from the population. A simulation was created that computes the expected utility of each individual and discovers the collectives in a population given randomly generated beliefs and utilities. Table 6.1 shows the collectives that emerge from a population in which approximately half the population leans towards not voting and the other half is split evenly towards voting pro-environment and pro-logging. The left column shows the rank order representing the collective. The middle three columns show the collective belief, or average expected utility, for each collective. The right column shows the percent of the population that the collective contains. All but one of the possible collectives (ELN) is seen in this simulation. Adjusting the ratio of vote tendencies in the simulation will change the relative size of the collectives and their collective beliefs.

This logging policy example used simulated beliefs and preferences. A similar example demonstrating collectives using beliefs elicited from humans is shown in Appendix B.

Collective Rank Order	Collective Belief E	L	N	Percent of Population
NEL	-0.11	-0.16	1.8	0.256
NLE	1.6	3.9	5.0	0.247
LNE	3.2	7.9	5.9	0.247
ENL	-2.4	-5.9	-4.4	0.242
LEN	5.2	5.0	7.8	0.008

Table 6.1: The collectives that emerge from a population based on the expected utility of the decision variable $VoteAction$.

6.3 Normal Form Games

This section demonstrates how a normal form game (Section 4.4, Def. 14) can be formed from the collectives in Table 6.1. A normal form game compares the utility for each player given the action of all other players. It can be used to find the Pareto optimal solutions and Nash equilibrium solutions in strategic situations. Figure 6.2 shows the normal form matrices for two pairs of collectives. A full normal form matrix would contain k dimensions, one dimension for each collective. Figure 6.2 shows only two of six dimensions, using two different pairs of collectives. The matrix on the left of figure compares ENL to LNE, the two collectives in direct opposition to each other. Their utility indicates that they both prefer opposing actions. In this case, each action is a Pareto solution. The normal form matrix on the right of the figure compares the collectives ENL and NEL. If the public were composed of only these two collectives, then the pro-logging vote action would not be Pareto optimal. However, since all collectives must be considered, then each option is a Pareto solution.

ENL/ LNE	E	L	N
E	-2.4, 3.2	-2.4, 7.9	-2.4, 5.9
L	-5.9, 3.2	-5.9, 7.9	-5.9, 5.9
N	-4.4, 3.2	-4.4, 7.9	-4.4, 5.9

ENL/ NEL	E	L	N
E	-2.4, -0.11	-2.4, -0.16	-2.4, 1.8
L	-5.9, -0.11	-5.9, -0.16	-5.9, 1.8
N	-4.4, -0.11	-4.4, -0.16	-4.4, 1.8

Figure 6.2: Two partial normal form matrices showing the utility of each action for two pairs of collectives. The left matrix compares opposing collectives ENL (rows) and LNE (columns). The matrix on the right shows two less opposing collectives, ENL and NEL.

6.4 Finding Nash Equilibria with Multi-agent Influence Diagrams

This section demonstrates how a concept developed by Koller and Milch in [35] forms the basis for more extensive game theoretic analysis with collectives. The authors introduced *multi-agent influence diagrams* (MAIDs) that represent strategic situations between multiple agents. MAIDs are defined formally in Section 4.4.2. In the MAIDs, each agent a represented in the scenario will have a *decision rule* for each of its decisions $D \in D_a$ that is a probability distribution over the options in D given $Pa(D)$, where $Pa(D)$ is the set of decision and variable nodes that affect D.

A strategy profile σ is an assignment for all agents of decision rules to all decisions in a MAID. A strategy profile σ is a Nash equilibrium solution, if for all agents $a \in A$, D_a is optimal given the strategy profile. In other words each agent has selected the option for each decision that optimizes their expected utility given the strategy of the other agents. The MAID in [35] assumes that an agent represents a single entity. This chapter demonstrates that an agent can be a super-agent representing a collective as defined in Section 4.3.

6.4.1 Model

This section extends the network in Figure 6.1 to form a MAID that represents a strategic situation between a community and a logging interest. In addition to the agents representing each emergent collective from the previous example, the players in the game will include a logger agent. As in the MAIDs in [35], the dashed lines represent decisions or variable nodes that affect a decision. These are the parents of a decision D, or $Pa(D)$. In the expanded MAID the variable node *Log* is a decision node, because the logging interest will make the decision to *log* or to *not log*, based on the which policy wins the vote. The loggers also have the option to lobby the policy-makers, which may increase the likelihood of pro-logging policy. However, lobbying is also associated with a cost. In addition, logging is associated with a profit for the logging interest, depending on the policy. A pro-logging policy allows for a positive profit, whereas a pro-environment policy results in a loss (or negative profit) for the logging interest.

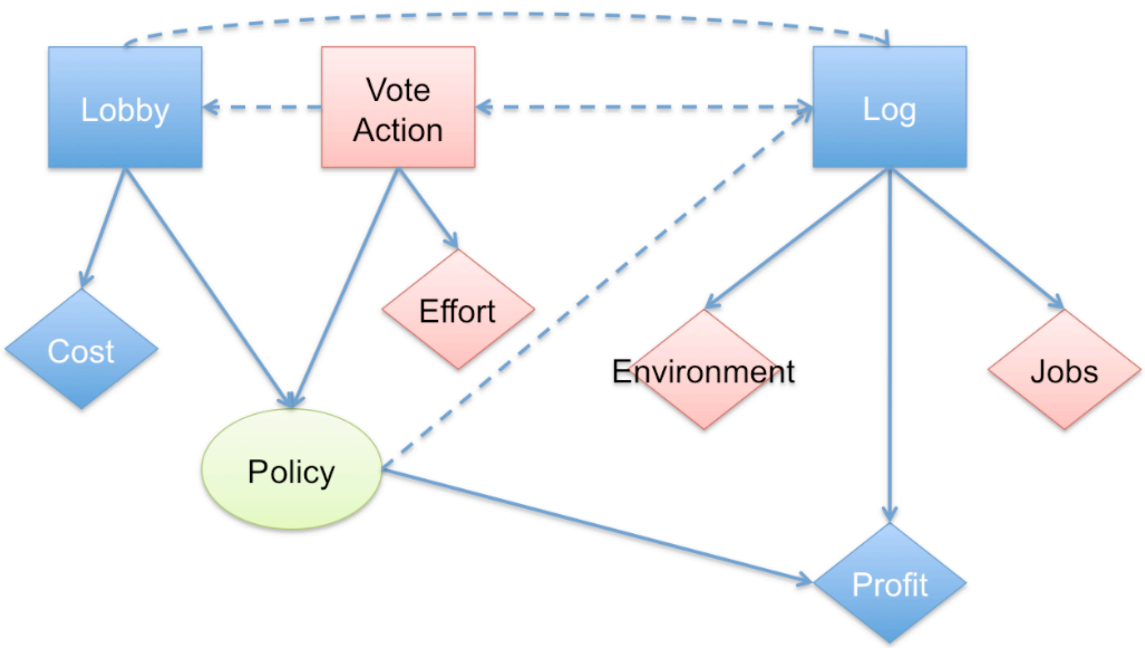


Figure 6.3: A MAID that extends Fig. 6.1 with the decisions and utilities of the logging interest. The dashed edges represent the values that an agent considers when he makes a decision.

6.4.2 Game theoretic analysis

This section demonstrates how to find the Nash equilibrium solutions and minimax/maximin strategies using the MAID in the previous section. Figure 6.4 shows the partial normal form for two agents, the ENL collective and the logger. As in the simpler MAID, a full normal form matrix would contain a dimension for each agent, however only two dimensions are shown for simplicity. The matrix in fig. 6.4.a shows the utility of each agent given both agent's decisions. The ENL agent's highest expected utility is to vote for the environmental policy because the vote will decrease the likelihood of pro-logging policy and increase the agent's utility for the environment. In this example the logging interest also decides to lobby, therefore the utility of the *don't log* action is negative.

The gray cell in the matrix shows the Nash equilibrium solution for this scenario. Given the logger's strategy of logging, the best strategy of the ENL collective is to vote pro-environment. Likewise, given the combined strategies (or strategy profile) of all the collectives, the logger's best action in this situation is to log. The logger's utility for logging is derived from the expected action of each of the collectives. Specifically, the probability distribution of vote action, $P(VoteAction = d)$ is derived from the collective's preferred actions using the following formula:

$$\forall d \in D : P(D = d) = \sum_{j \in C_d} \frac{|C_j|}{N}$$

Where $D = \mathit{VoteAction}$, $d \in \{E, L, N\}$, and C_d is the set of collectives that prefer action d. The logger then computes the likelihood of pro-logging policy given the probability distribution of vote action. In this particular situation in which only about 25% of the population votes pro-environment, policy is likely to lean in the logger's favor and the loggers will make a profit. The normal form matrices in Figure 6.2 that compare the collective agents will not change with the additional consideration of the logger's utility and actions.

I now discuss the maximin strategy, defined as the strategy for an agent a that maximizes the payoff given that the other agents' aims are to minimize a's payoff [37]. In other words, it is the best choice for a given his worst case scenario. The table in Figure 6.4.b shows the logger's utility of logging given hypothetical situations in which the public leans 100% towards one of the actions. The worst case scenario for the loggers is the case in which all a.

ENL	Loggers log	Loggers don't log
E	-2.4, 415	0,-100
L	-5.9, 415	0,-100
N	-4.4, 415	0,-100

b.

Vote Action	Loggers log	Loggers Don't log
E	-129	-100
L	451	-100
N	885	-100

Figure 6.4: (a.) A normal form for the MAID comparing the logging interest to the ENL collective, whose preferred action is to vote pro-environment. The Nash equilibrium solution is the shaded cell in the matrix. (b.) Table showing the logger's utility of logging given situations in which the public leans 100% towards one of the actions.

collectives prefer to vote pro-environment. Since the E action is the strategy that minimizes the logger's utility, the maximin strategy for the logging interest is to *not log*, since pro-environment policy will result in a loss if they log. The minimax strategy is the complement of the maximin strategy. One could consider voting pro-environment to be a minimax strategy for the public because it would minimize the payoff for the logger if the public were to "gang up" on the logging interest [37].

The previous example results in a Nash equilibrium solution in which the logging interest logs. If the input parameters are changed, the Nash equilibrium solution for the logging interest may also change. Figure 6.5 shows the best strategy for the logging interest if the vote ratio of the community is changed. The amount of potential loss is also varied, where potential loss is the percent of potential profit that the logger's experience if policy is pro-environment. The figures show the logger's decision to *log* or to *not log*, given the $loss$ in rows, and $ratio$ in columns and the logger's decision to $lobby$. A white cell containing 0 indicates that the loggers best option given the input parameters is always to *not log*. A dark gray cell containing 1 indicates that the loggers best option is always to *log*. A light gray cell containing 1/0 means that the best option was to *log* when $lobby = true$ and to *not log* when $lobby = false$. Therefore the light gray options show the effect of lobbying.

loss/ratio	0.99 0.00 0.01	0.90 0.05 0.05	0.8 0.1 0.1	0.7 0.1 0.2	0.6 0.2 0.2	0.5 0.2 0.3	0.4 0.3 0.3	0.25 0.25 0.50	0.2 0.5 0.3	0.2 0.6 0.2	0.1 0.7 0.2
1.0	0	0	0	0	0	0	1/0	1	1	1	1
.9	0	0	0	0	0	1/0	1	1	1	1	1
.8	0	0	0	0	0	1/0	1	1	1	1	1
.7	0	0	0	0	1/0	1/0	1	1	1	1	1
.6	0	0	0	1/0	1/0	1	1	1	1	1	1
.5	0	0	1/0	1/0	1/0	1	1	1	1	1	1
.4	0	0	1/0	1/0	1	1	1	1	1	1	1
.3	1/0	1/0	1	1	1	1	1	1	1	1	1
.2	1/0	1	1	1	1	1	1	1	1	1	1
.1	1	1	1	1	1	1	1	1	1	1	1
0	1	1	1	1	1	1	1	1	1	1	1

Figure 6.5: A table showing the logger's Nash equilibrium solution to *log* (1) or to *not log* (0), varying three options: the loss for loggers (rows represent the possible loss as the percent of possible profit), the ratio of vote actions (columns represent the ratio of the community that leans towards vote actions [E, L, N]), and the decision to lobby. The light gray cells containing (1/0) show the effect of lobbying. Not logging is preferred when the logger does not lobby, while logging is preferred when the logger decides to lobby.

6.5 Sequential Games and Imperfect Information

Section 6.3 described the approach to form multi-agent influence diagrams from collectives to enable game theoretic analysis. The logging simulation was essentially a single action game, in which the strategy of each player was known or assumed before each agent analyzed his decision options. In real world situations, the strategy of the other players is not always immediately apparent [37]. In many cases players will not admit their strategy to the other players, thus strategy must be inferred over multiple moves. Alternatively, players may switch strategies depending on the actions taken by the other players over multiple plays. In these situations it may take time for a game to settle into a state in which a Nash equilibrium solution can be determined. As defined, the MAIDs represent only the known strategies with a dashed line from the known decision for agent a_i, to the corresponding decision for agent a_j. A possible extension to the MAID would be to represent decisions as a probabilistic variables until they are known with a high degree of certainty.

Appendix C describes an an election polling simulation that demonstrates game theoretic analysis in a sequential game. In a sequential game the moves of each player depend on the moves of the other players in the previous time step [37]. In addition, the strategy of each individual may change or may not be immediately evident to all other players. In the described situation the Nash equilibrium solution is found only after the game had stabilized to the extent that individuals stopped changing their strategies. I also discussed an interesting phenomenon in which the average expected utility of the population predicted an outcome that was different than the outcome indicated by the individuals' expected. This result may be relevant the price of anarchy, defined as the difference between individuals acting to maximize the social optimum and individuals acting to maximize their own utility [55].

6.5.1 Summary

This chapter demonstrated how to apply game theoretic analysis using the collectives generated with the collective choice function. Section 6.1 introduced an algorithm to extract the Pareto optimal solutions from the collective beliefs of each collective. It demonstrated that the algorithm will return a set of Pareto optimal solutions in situations that fail using a single consensus approach. Section 6.3 discussed normal form games and showed the normal form matrices for pairs of collectives. The remainder of the chapter expanded the concept of multi-agent influence diagrams to form agents in a game from a set of collectives. The game-playing agents with potentially competitive goals emerged from the collectives that were derived from the MAIDs. The MAIDs were used to find Nash equilibrium solutions and the minimax and maximin strategies.

This chapter demonstrated a general approach to applying game theoretic analysis using collectives and therefore successfully achieved the objective described in Section 1.3.3. The `findParetoSet` algorithm will find the Pareto solutions given a set of collective beliefs in situations that can be described with a Bayesian network. The extended MAIDs can be used to find Nash equilibrium solutions in strategic situations between agents and super-agents that can be described with a MAID. The logging simulation highlights the applicability of my approach in policy and decision- making on a large scale. By considering the beliefs, preferences and behavior of a population, we can see the possible effects of public action on policy, as well as the effects of public action and policy on the actions of corporations and other organizations.

Chapter 7

Algorithms for Collective Belief Aggregation

This dissertation has discussed an approach to social decision-making using collective belief models and game theoretic analysis. I have described how this approach achieves the first three objectives set forth in Section 1.3. I now present an algorithmic approach to discovering collectives, detailing the collective choice function discussed at a high level in Section 4.3. The algorithms discussed are combined aggregation and inference algorithms, differing only in the order of aggregation and inference. Specifically, I discuss and compare three different algorithmic approaches to inference: *prior aggregation, posterior aggregation*, and my *incremental aggregation* approach. Posterior aggregation will serve as the exact, brute force algorithm according to the collective choice function. Posterior aggregation performs inference, based on a query, on each individuals' Bayesian network prior to performing aggregation on their results. Prior aggregation attempts to aggregate the individuals to form consensus networks, on which inference is then performed. Incremental is a hybrid between prior and posterior aggregation. This chapter discusses the algorithms, then compare their complexity. Finally runtime and accuracy of the approaches are compared using simulated models and data. This chapter address the final objective, defined in Section 1.3.4.

7.1 Incremental Collective Discovery

This section discusses an algorithm for discovering a partition of an elicited population including the collectives and their collective beliefs. The intuition for the alternative aggregation approaches is first given. A formal description of the incremental algorithm is then provided. Since any Bayesian decision network can be converted into a Bayesian network [42, 35], the inference algorithm for a Bayesian network is discussed. The collective choice function defined in Section 4.3 utilized the *inferred* belief in a queried variable, X_i, also known as its posterior probability $P'(X_i)$. Thus, the inference algorithm introduced in this dissertation will be for the posterior belief assessment of a variable $P'(X_i)$ given X_i's network. The network could contain prior distributions for X_i's ancestors and observations about X_i's descendents.

Since I am interested in the inferred collectives, the exact, brute force algorithm is simply to 1) query each individual's network, 2) run an inference algorithm on each network and return the result of the query, and 3) run a *binning* algorithm to place each individual into a collective based on their inferred beliefs. I will call this approach the *posterior* aggregation algorithm (after posterior compromise [40]) because it does aggregation *after* inference. While this algorithm will result in collectives that accurately represent the Bayesian rank orders $P(X_i)$, no benefits in using aggregation to reduce runtime will be seen.

Another algorithmic approach would be to form the consensus networks *before* running inference on them. In other words one forms the collectives based on the individuals' a priori beliefs and conditional distributions. I call this approach the *prior* aggregation algorithm (after prior compromise [40]). This algorithm could result in reduced runtime over the posterior algorithm because the computationally intensive inference will occur on a smaller number (k) of consensus networks than the m individual networks in the posterior algorithm. However, since the collectives are formed based on the *inferred* beliefs, the prior aggregation approach may not guarantee that the collectives will represent the inferred Bayesian rank orders for the population accurately. Essentially, the prior aggregation algorithm would be making a *guess* about the posterior probability $P(X_i)$ based on simply observing the a priori probability distributions. If this could be done accurately, then the inference algorithm would not be needed!

The prior aggregation algorithm can be compared to a classification algorithm. A classification algorithm attempts to determine the class of an object based on the object's *features* [64]. Similarly, the prior aggregation algorithm attempts to classify the Bayesian rank order and therefore the collective (class) that an individual (object) belongs to based on his prior distributions (features). Classification algorithms are measured by their accuracy. If an algorithm is able to correctly classify each object correctly, it is be considered 100% accurate. Likewise, If an aggregation algorithm is able to correctly determine each individual's collective, as does the posterior aggregation algorithm, then that algorithm is 100% accurate. Therefore, an algorithm's ability to determine the posterior Bayesian rank order will be referred to as its *accuracy*.

An approach that balances runtime and accuracy would provide the most flexibility. In other words, it should have a lower runtime than the posterior aggregation algorithm, but a higher accuracy than the prior aggregation algorithm. I introduce a combination of the two approaches, that I call the *incremental* aggregation approach. The incremental aggregation algorithm uses the classification *guesses* of the prior aggregation algorithm to "bootstrap" the algorithm, but finds the posterior collectives for each variable X_j during inference. The result is a set of collectives for each variable in X_i's network. The algorithm cannot guarantee that the collectives reflect the true rank order of each individual for each variable. However, it is expected that the algorithm's accuracy will be higher than the prior aggregation approach.

The incremental aggregation algorithm is integrated with a Bayesian inference algorithm (to be discussed later). It returns the collectives and collective beliefs for the posterior probability of a given query $P(X_q)$, where X_q is the variable of interest. The algorithm works intuitively as follows:

Bootstrap step:
The algorithm finds the set of collectives and collective beliefs for the prior distributions of all variables in the network for which individuals have supplied prior probability distributions, $P_C(X_i)$, and the conditional probability distributions, $P_C(X_i|Pa_i)$ for all variables with parents.

Inference step:
For the purpose of this introduction, assume that the posterior probability of a node has been

computed or is available before it is passed to its children. The algorithm computes the collective posterior probability $P_C(X_i)$ of each node using the collective beliefs for the posterior distributions of X_i's parents, $P_C(Pa_i)$, and the collective beliefs for the prior distributions of $P_C(X_i|Pa_i)$. Thus, $P_C(X_i)$, the posterior probability is computed as follows:

$$P_C(X_i) = P_C(X_i|Pa_i)P_C(Pa_i) \tag{7.1}$$

The potential benefits of the incremental algorithm are derived by reducing the number of values being propagated for each node. Instead of propagating each individual's probability distributions throughout the network, the algorithm propagates the collectives' distributions. This results in propagating $k \leq m$ values instead of m values, where m is the population size and k is the number of collectives. First, the bootstrap step jump-starts aggregation so the algorithm begins with k_i collectives and their means for the prior or conditional probability of each node X_i. The algorithm then updates the collectives for each node X_i during inference, after computing X_i's posterior probability, $P'(X_i)$. The mean posterior probabilities of these collectives are then propagated to X_i's neighbors instead of each individual's posterior probabilities.

An existing Bayesian inference algorithm was extended to develop the incremental aggregation algorithm. The well-known variable elimination algorithm, developed by Zhang and Poole [68] and Dechter [15], was chosen because it returns an *exact* posterior probability distribution and because it is simpler than the junction tree algorithm by Jensen [32] that requires a Bayesian network to be converted into an intermediate structure [13].

7.2 Algorithms

I will now summarize the original variable elimination (also known as bucket elimination) algorithm for belief assessment of a query Q. The algorithm below summarizes Kevin Murphy's implementation of the algorithm from his Matlab toolkit [43]. Further discussion of the algorithm can be found in [68, 14].

Procedure `variableElim`:

- Inputs:
  - PTs: A PT for all variables $X_i \in X$, including observations
  - d: an ordering over those variables
  - Q: a query
- Outputs: $P'(Q)$, the posterior probability distribution of Q

```
variableElim(PTs, Q, d)
    for each node X_i
        initialize X_i's bucket with its PT and its family's PTs
        compute the product of the PTs assigned to each bucket (a JPT)
    end
    place Q at the beginning of d
    for each node X_i in reverse order of d
        marginalize last bucket containing X_i to sum it out
        place result in the previous bucket containing X_i
        find product of PTs in bucket
    end
    the first bucket will contain P'(Q)
end
```

The next algorithm, `binPT(PTs)`, discovers the collectives for a population based on the PTs provided for a variable. The population is either composed of a set of individuals, each with his own PT; or a set of collectives, each of which has a collective PT. The algorithm "bins" each individual into a new collective based on the symbolic order of the values in each PT. If the input is a set of collectives and their PTs, then each individual in a collective is binned based on his collective's PT. The `findPTOrder(PT)` procedure is a helper procedure that returns the rank order of a single PT, formed by appending the order of the conditional probabilities for each combination of parent values. Symbolic orders are discussed in depth in Section 7.3. Procedure `computeCollectiveBelief(PTs,IDs)` finds the mean of `PTs` using the function in eq. 7.2 if `IDs` is a set of individuals, or the weighted mean function in eq. 7.3 if `IDs` is a set of collectives, each containing the IDs of its members.

Procedure `binPT(PTs, IDs)`:

- Inputs:
  - `PTs`– A set of PTs for a variable X
  - IDs– The IDs associated with each PT; could be a single ID or a set of IDs for each PT
- Outputs: A set of collectives for X's PT

```
binPT(PTs)
   /* each individual or collective has a cpt in PTs */
   for all cpt in PTs
       order=findPTOrder(cpt)
       place owner of cpt in a collective defined by order
   end
   find mean PT for each collective
end
```

Procedure `findPTOrder(PT)`:

- Inputs: `PT`— A single PT
- Outputs: The symbolic order of the values in the PT

```
findPTOrder(PT)
   /* append the order of vals for each parent combination */
   order = empty
   allParentCombos = enumerate all combinations of parent values
   for each combo of values in allParentCombos
```

```
            o=find order of node's values given node's parents in PT
            append o and combo to order
       end
   end
```

Procedure `computeCollectiveBelief(PTs,IDs)`:

- Inputs: `PTs`— A set of PTs; `IDs` the IDs for a set of individuals, or a set of collectives
- Outputs: The mean of the PTs. If PTs belong to collectives, then result is a weighted mean of the PTs

```
computeCollectiveBelief(PTs,IDs)
    if IDs contains a set of separate individuals
        collective belief is computed as in e.q. 7.2
    else IDs contains a set of collectives
        new collective belief is computed as in e.q. 7.3
    end
```

$$\frac{1}{|IDs|}\sum_i PT_i \tag{7.2}$$

$$\sum_j PT_j \times \frac{|PT_j|}{|P|} \tag{7.3}$$

I now discuss the extension to the variable elimination algorithm, `variableElimAgg(PTs, Q, d)`, which also performs incremental aggregation. Aggregation is performed when probability distributions are combined to form a joint distribution or the posterior distribution of a node. This will occur when the PTs in a bucket are multiplied to find the new PT for the bucket. This is the propagation step in inference, in which one finds the product of two probability distributions. The helper procedure, `productOfIntersectingBins(...)` performs the aggregation, but first it must combine two sets of collectives, representing the collectives for two different distribution tables. When the two sets of collectives are combined, the product of only the overlapping collectives is found. The process is illustrated in Figure 7.1. The A and B nodes represent two sets of collectives, which are the two sets of PTs being combined in a bucket. Collective A_i's collective belief will be in PT_{ai} and collective B_j's will be in PT_{bj}. The first step in `productOfIntersectingBins(...)` is to find the cartesian product of all collectives in A, with all collectives in B. For all A_i and all B_j, a temporary collective $T_{ij} = A_i \cap B_j$ is found, represented by the small ovals attached to the edges in Fig. 7.1. If T_{ij} is non-empty, then PT_{ij} will be the product of PT_{ai} and PT_{bj}, the collective belief of A_i and B_j, respectively. If T is the set of non-empty intersections, then $|T| \leq m$. In other words there can be no more than m non-empty intersections– one for each individual in P.

The overall result of `productOfIntersectingBins(...)` is that it finds the product of the groups of individuals that agree on the preference order between two PTs. A side effect is that the runtime of inference can potentially be reduced in the cases when the number of non-empty intersections is less than m, compared to exact inference which does propagation on all m networks. After finding T, the algorithm re-bins T into a new set of collectives by calling `binPTs`(PT_T,T), where PT_T contains the products of the intersections. In the bucket elimination algorithm, PT_T could contain conditional probability tables or joint probability tables, depending on the PTs that are combined.

Procedure `variableElimAgg(PTs, Q, d)`:

- Inputs:
  - PTs– A PT for variable all variables $X_i \in X$, for all m individuals
  - d– an ordering over all variables X_i
  - Q- a query
- Outputs: A set of approximate collectives based on Q and the posterior probability distribution of Q for each collective

```
variableElimAgg(PTs, Q, d)
    for each node X_i
        call binPT(PTs) to find collectives for X_i
```

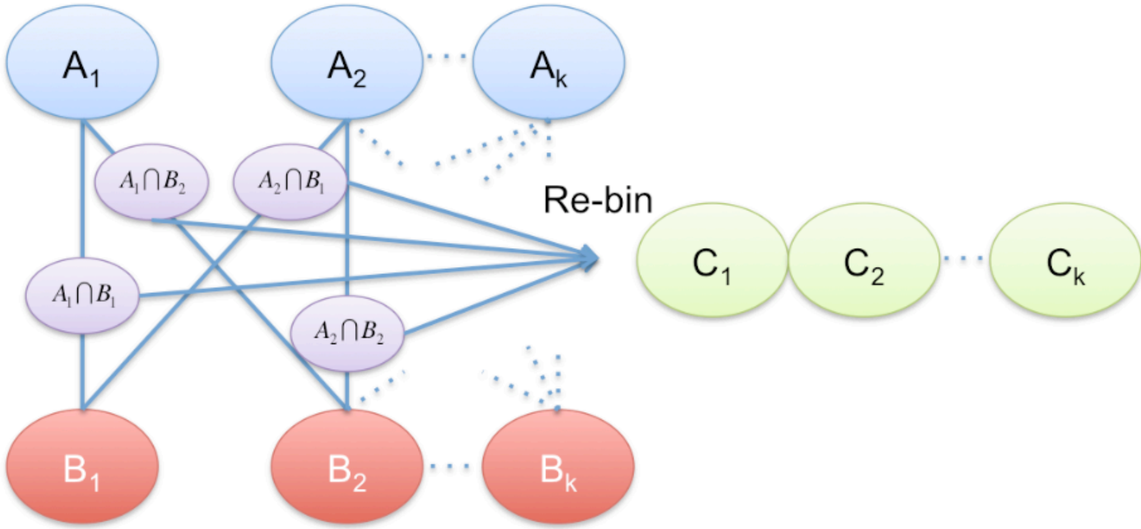


Figure 7.1: Illustration of the process of combining two sets of collectives during propagation.

        find PT_{ij}: the collective belief for each collective C_j
    end
    for each node X_i
        initialize X_i's bucket with each collective's PT and the PTs ↩
            of X_i's family
        call productOfIntersectingBins($Bins_a$, $Bins_b$, PTs_a, PTs_b) ↩
            to compute the products of the PTs assigned to each bucket
      place result in bucket
    end
    place Q at the beginning of d
    for each node X_i in reverse order of d
      for each collective in last bucket b containing X_i
            marginalize the bucket to sum X_i out
            place result in the previous bucket b' containing X_i
      end
      call productOfIntersectingBins($Bins_a$, $Bins_b$, PTs_a, PTs_b) ↩
            to compute the new product of the PTs in the bucket b'

```
      place result in bucket b'
   end
   the first bucket will contain P'(Q) for each collective
end
```

Procedure `productOfIntersectingBins`($Bins_a$, $Bins_b$, PTs_a, PTs_b)

- Inputs:
  - $Bins_a$ and $Bins_b$: The collectives for the two PTs to be multiplied;
  - PTs_a, PTs_b: The PTs for each collective
- Outputs: New collectives and PTs for the product of the two sets of PTs

```
productOfIntersectingBins( Bins_a, Bins_b, PTs_a, PTs_b)
  for all bins Bin_i in Bins_a and Bin_j in Bins_b
      T_ij = the intersection of Bin_i and Bin_j
      if intersection is non-empty
          PT_ij = the product of PT_i and PT_j
          place PT_ij in Prods
      end
  end
  call binPTs(Prods, T) to bin the results into new collectives
end
```

As with many algorithms that attempt to approximate the results of an exact, brute force algorithm, the incremental aggregation algorithm balances efficiency with accuracy. An unfortunate side effect of the algorithm is that it may inaccurately predict the final collective for an individual, since it propagates the means of collectives instead of each individual's PTs. This phenomena is an effect of computing a product of probability distributions, which can result in a different preference order than the products of an individual's PTs. The following example illustrates this situation. Suppose we have two individuals A and B with the following probability tables for $P(Y|X)$ and $P(X)$:

A

$P(Y\|X)$		X = F	X = T
Y	F	.49	.99
	T	.51	.01

$P(X)$	
F	.75
T	.25

B

$P(Y\|X)$		X = F	X = T
Y	F	.01	.51
	T	.99	.49

$P(X)$	
F	.75
T	.25

A and B both have the same preference orderings for both tables: $P(Y|X) = FTTF$ and $P(X) = FT$, however, the product of these tables does not yield the same preference order for $P(Y)$:

$$P_A(Y = T) = .51 * .25 + .01 * .75 = .135 \rightarrow FT \tag{7.4}$$

$$P_B(Y = T) = .99 * .25 + .49 * .75 = .615 \rightarrow TF \tag{7.5}$$

Clearly combining the mean of the two individual's probability tables indicates that one individual's preference order will not be equivalent to the mean's preference order. This phenomenon was proven in general in [56]. They showed that it is not possible to find a rational social choice solution using prior aggregation. Only posterior aggregation, used in the exact aggregation algorithm, can guarantee accurate collectives, in which each individual in a collective has the same preference order for the posterior belief in a query as the collective to which she belongs.

While the prior aggregation can result in incorrect estimation of collectives, the incremental algorithm uses a combination of prior and posterior aggregation. Prior aggregation is used in the initialization step to aggregate each PT and posterior aggregation is used after combining PTs in buckets. It is expected that the incremental aggregation algorithm will correctly predict the appropriate collective (preference order) for an individual in a population more often than a full prior aggregation approach. A full prior aggregation approach is one in which the individuals are aggregated to find consensus networks prior to inference. Other belief aggregation techniques use this approach to form a single consensus network [52, 40]. Prior aggregation can also be used to form multiple consensus networks. In this case individuals are first clustered based on their CPTs and then a consensus network is formed for each cluster. This approach was used in 5.4. In Section 7.3 I will compare the accuracy and runtime of three aggregation approaches: prior, incremental and posterior.

7.3 Complexity

Bayesian inference is equivalent to SAT in the worst case. The variable elimination algorithms I have described are used to assess belief in a Bayesian network to compute the posterior probability of a queried variable $P(X_q)$. There are approximations for belief assessment, such as belief propagation, that can achieve improved space and runtime [50, 16]. The complexity of Bayesian network inference is highly dependent on the structure of the network, in particular the degree of the nodes in the network. Researchers show that Bayesian inference is exponential in the *induced width* (see Def. 22) [13]. Therefore runtime of the bucket elimination algorithm is highly dependent on the order in which nodes are processed. An optimal ordering would be one in which a node's parents always precede it in the ordering [14].

The time and space complexity for finding the posterior probability of a queried variable $P(X_q)$ using bucket elimination is $O(n \cdot exp(w*(d))$, where $w*(d)$ is the induced width of an optimal ordering d and n is the number of variables in the network [14]. I will refer to $w*(d)$ as w in further discussions. Thus, the complexity of inference on a single network is $O(n \cdot exp(w))$. In other words, $O(exp(w))$ computations are required to compute the posterior probability of each node in the network.

In the case of a poly-tree, the induced width is equivalent to $|F|$, the maximum family size (node and its parents). In order to compare the complexity of the exact aggregation algorithm with the incremental aggregation algorithm I will assume a poly-tree because family size is a key factor in the complexity of the incremental algorithm.

The complexity of inference on a poly-tree is now analyzed in more detail. At each node X_i, one must consider two important components:

- $|Pa_i|$: the number of incoming nodes, or parents
- r_i: the arity of the each of the variables, including X_i and Pa_i, r will be the maximum arity considered

The size of the conditional probability table (CPT) for the probability of a variable X_i given its parents Pa_i, representing the values for $P(X_i|Pa_i)$ will be $O(r^{|Pa_i|+1})$, which is equivalent to $O(r^{|F_i|})$. This is also equivalent to the time to compute the posterior probability $P(X_i) = P(X_i|Pa_i)P(a_i)$. Assuming a poly-tree, this results in the runtime of inference as described above: $O(n \cdot exp(|F|))$. I have simply specified $exp()$ to be r. Thus, for purposes of comparison, I describe the runtime of inference on a poly-tree as $O(n \cdot r^{|F|})$.

7.3.1 Complexity of exact algorithm

The exact combined aggregation and inference algorithm runs inference on m networks, one for each individual in a population, and then "bins" the individuals into collectives based on the rank order of the individual's belief. Assuming each individual has an identical poly-tree, the runtime of the exact algorithm is $O(mn \cdot r^{|F|})$ plus the runtime of binning. Binning and finding the means of the collectives will be $O(m)$, and therefore will not affect the asymptotic runtime.

7.3.2 Complexity of incremental algorithm

My objective is to reduce the asymptotic runtime of approximate inference. I now analyze the complexity of the incremental algorithm to see under which conditions this is possible. The algorithm is split into two steps:

1. *Initialization*: Finds the set of collectives for the prior or conditional probability table for each variable X_i in a network. Uses the procedures `binPT(PTs, IDs)`, `findPTOrder(PT)` and `computeCollectiveBelief(PTs,IDs)`
2. *Inference*: Computes the posterior probability of each variable X_i in the network using the collectives beliefs discovered in the initialization step. In the process the approximate collectives for each bucket are incrementally updated. The inference algorithm returns the collectives and collective beliefs for a query Q. Inference uses the `variableElimAgg(PTs, Q, d)` and `productOfIntersectingBins(`$Bins_a$, $Bins_b$, PTs_a, PTs_b`)` procedures as well as the procedures used in the initialization step.

I first analyze the number of possible collectives for the value $P(X_i)$ (or $P'(X_i)$). If X_i has r possible values then there are $r!$ possible orderings on those values, each of which is a potential collective. For example, if X_i is binary, then there are $2! = 2$ possible collectives representing the orderings TF and FT.

I now analyze the number of possible collectives for the CPT of a variable X_i if X_i has parents. In this case, each variable in X_i's family will have up to r possible values. The possibility that there is a different ordering for each combination of parent values must be considered. For instance, if a variable X_i has one parent X_j, then the following orderings (followed by their symbolic representation) are possible:

1. $X_i = TF|X_j = T$ and $X_i = TF|X_j = F$ $(TFTF)$
2. $X_i = TF|X_j = T$ and $X_i = FT|X_j = F$ $(TFFT)$
3. $X_i = FT|X_j = T$ and $X_i = TF|X_j = F$ $(FTTF)$
4. $X_i = FT|X_j = T$ and $X_i = FT|X_j = F$ $(FTFT)$

Unfortunately, the number of possible collectives grows super-exponentially with the number of parents. The following table shows only four of the 16 possible collectives if X_i gains a second parent, X_k:

X_i	X_j	X_k	Ordering
TF	T	T	TFTFTFTF
TF	T	F	
TF	F	T	
TF	F	F	
TF	T	T	TFTFTFFT
TF	T	F	
TF	F	T	
FT	F	F	
TF	T	T	TFTFFTTF
TF	T	F	
FT	F	T	
TF	F	F	
TF	T	T	TFTFFTFT
TF	T	F	
FT	F	T	
FT	F	F	
...			

In general, the number of collectives, given r and $|Pa|$ is $O(r!^{r^{|Pa|}})$, since there are $r^{|Pa|}$ combinations of parent values and $r!$ possible orderings of $P(X_i)$ for each combination. Fortunately, variables are often binary or at least r will be small, bounding the factorial growth of $r!$ to a constant. The number of collectives is also bound by the population size m. In the worst case, each individual will be in his own collective. Therefore, the upper bound on the collectives is $O(min(m, r!^{r^{|Pa|}}))$.

I now analyze each step of the incremental algorithm to determine the asymptotic behavior. First the initialization step is analyzed. In the `binPT(PTs, IDs)` procedure in Section 7.2, each individual is placed in a collective for $P(X_i|Pa_i)$. Each individual j will have a PT for each node X_i. While there are $|C_i|$ possible collectives for each X_i, this number will only be reached if each possible ordering is seen and $m \geq |C_i|$. The `findPTOrder(PT)` procedure determines the ordering of a single PT and the collective in which its owner belongs. The ordering is derived by appending the order of $P(X_i)$ given each combination of its parent's values. The number of combinations of parent values for X_i is $r^{|Pa_i|}$. For each combination *combo*, the r possible values of X_i are sorted (in line 13), and then appended to the whole ordering (line 14). The sort will require $O(r \cdot log(r))$ operations, but since r is likely to be small, this value is a constant. Thus, the runtime of the `findPTOrder(PT)` procedure is $O(r^{|Pa|})$. The `binPT(PTs, IDs)` procedure will run `findPTOrder(PT)` on each individual's PT. If there are m individuals and n PTs in a network, then the runtime for the initialization step is:

$$O(mn \cdot r^{|Pa|}) \tag{7.6}$$

Finding the mean of each collective takes an additional $O(m)$ time and is dominated by the runtime of binning. The runtime of the initialization step is only slightly lower than the runtime of the exact inference algorithm because $|Pa| = |F| - 1$, however, recall that the estimate for the exact inference algorithm is only considering poly-trees. In general graphs, the runtime is $O(mn \cdot r^w)$, where $|F| \leq w$. The runtime of the incremental algorithm's initialization is for a general, possibly multiply-connected graph. Therfore, the runtime of binning is in general asymptotically lower than the runtime of exact inference.

I now analyze the inference step. The incremental algorithm attempts to improve efficiency by propagating the collective beliefs throughout the networks instead of each individual's value. The collective beliefs were discovered in the initialization step, and they now serve to bootstrap the algorithm. Like the original bucket elimination algorithm, complexity of `variableElimAgg(PTs, Q, d)` is dominated by the time to process each bucket. In the original algorithm this is bounded by $O(exp(w))$, where w is the induced width of the optimal ordering. In Murphy's implementation of the algorithm the parents of node X_i will be processed before X_i [43]. Therefore the order is near optimal.

The runtime of the `variableElimAgg()` algorithm is also affected by the number of possible collectives for each probability distribution, which was previously shown to be $O(min(m, r!^{r^{|Pa|}}))$. The `productOfIntersectingBins(...)` procedure finds the intersection of two sets of collectives, thus if we let $c = r!^{r^{|Pa|}}$, then there are $O(min(m, c^2))$ possible collectives for each bucket. The algorithm then bins the results of the intersections into new collectives, at a runtime of $O(r^{|Pa|})$ per bucket. Since $r^{|Pa|} \leq exp(w)$, then this means each bucket requires $O(2 \cdot min(m, c^2) \cdot exp(w))$ operations. The factor of two is added because the process occurs twice, between lines 6-12 and 14-22. Thus, the runtime of the `variableElimAgg(PTs, Q, d)` is:

$$O(n \cdot min(m, c^2) \cdot exp(w)) \tag{7.7}$$

This runtime dominates the initialization step when $min(m, c^2) = m$ and $exp(w) \geq r^{|Pa|}$, which is likely in any general, multiply-connected network. The asymptotic complexity is equivalent to the exact inference algorithm, whose runtime is $O(mn \cdot exp(w))$, when $min(m, c^2) = m$. If $c^2 < m$ then the incremental algorithm will be faster. Thus, we can see that the incremental algorithm is highly dependent on the number of possible collectives, which is based on the size of the probability tables, and is a factor of the number of parents each node has. Note that the upper bound on the number of possible collectives is not always reached in practice. Reduced runtime would occur in situations in which there is a high degree of agreement across multiple probability tables. In other words, a reduction in the number of non-empty intersections would reduce the number of actual collectives.

Reduced complexity could also be derived from re-structuring the network to reduce the maximum number of parents [13] or by using combining rules such as the noisy-OR gate, in which each parent-child relationship is considered independent of the others [50, 18, 68]. This would eliminate the exponential growth on the size of the conditional and joint probability tables. In fact, since the belief networks are derived from beliefs elicited from people, combining rules would also be a more natural way for individuals to provide their beliefs on conditional probabilities. For instance, it seems more natural for a human to describe the effect of one factor in the outcome of a situation than the effect of all factors combined. Extending noisy-OR gate inference is a natural future extension of the work. However, I have described the complexity using general Bayesian inference for completeness.

7.4 Experiments

This section discusses several experiments that measure the runtime and accuracy of the variable elimination algorithm. Three aggregation approaches are compared:

- *Posterior aggregation:* runs inference on each individual's belief network based on a query Q and bins the posterior probability, $P'(Q)$ into collectives.
- *Incremental aggregation:* runs inference using the incremental algorithm described in Section 7.2
- *Prior aggregation:* uses the Weka k-means clusterer [64] to cluster individuals based on their prior and conditional probability tables into k clusters. Consensus networks are then formed from the clusters and inference is run on them to find $P'(Q)$. The clusters that formed the consensus networks are then binned into collectives based on their rank orders.

The section focuses on comparing posterior aggregation with the incremental aggregation algorithm. The accuracy of prior aggregation is poor enough to not make it a contender for practical use, at least without significant modification.

7.4.1 Metrics

The experiments in this section test three metrics: *runtime*, *accuracy* and *representation*. Runtime is as it sounds, simply measuring the time (in seconds) to complete a combined aggregation and inference task. In most of the experiments, a single task is composed of computing the collectives and collective beliefs for a query, given a Bayesian network structure and a population of individuals with randomly generated parameters. Accuracy specifically refers to correctly predicting the preference order of an individual's posterior belief, which in turn indicates the collective to which an individual belongs. The *actual* preference order is the preference order derived from posterior aggregation. Thus, posterior aggregation has 100% accuracy. To measure accuracy of incremental and prior aggregation, I compare the results of running aggregation and inference on the same network using posterior aggregation. If $\hat{C}_r$ is the collective to which an individual r has been found to belong using posterior aggregation and C_r is the collective to which the same individual has been found to belong using a different aggregation technique T, then the accuracy of the technique, α_T

given the set of collectives C derived from T is measured as:

$$c_r = \begin{cases} 1 & \text{if } C_r = \hat{C}_r \\ 0 & \text{otherwise} \end{cases}$$

$$\alpha_T(C) = \frac{1}{|P|} \sum_{r=1}^{|P|} c_r \tag{7.8}$$

Finally, *representation* is a measure of how well the aggregation technique represents the population's beliefs. As discussed in Section 5.4, Kulback-Leibler (K-L) divergence is a measure of how divergent a distribution Q is from a distribution P. Given a "target" or "actual" distribution P and an estimate Q, the K-L divergence of Q from P over n data points is:

$$KL(P||Q) = \sum_{i=1}^{n} P(i) log \frac{P(i)}{Q(i)} \tag{7.9}$$

If the K-L divergence of a distribution Q from P is near zero, then the distributions are very similar. In this case Q is considered a "good" representation of P. My hypothesis is that the collective beliefs are more representative of a population than a single consensus model. The following measure uses K-L divergence to measure the divergence of the collective's belief distribution from the distribution of all individuals' beliefs. This is compared to the K-L divergence of the mean of all beliefs in a population. The distributions are computed from a histogram of n bins over the values in $[0, 1]$. Inference on each individual r's network results in a posterior probability $P'_r(X_q = T)$ for a queried binary variable X_q, the value of which will fit into one of the n bins. $P_{Iq}(i)$ is the probability that any individual's network resulted in a posterior probability $P'_r(X_q) = T$ that fits in bin i. For each collective C_j in C, $P'_{0j}(X_q = T)$ is the collective belief derived from aggregation and inference. $P_{Cq}(i)$ is the probability that an individual's belief is in bin i if each individual in collective C_j provides the belief $P'_{0j}(X_q)$. Given the set C of k collectives, I compute the K-L divergence between P_I, the distribution of the individuals and P_C, the distribution of the set of collectives as follows:

$$KL(P_I||P_C) = \sum_{i=1}^{n} P_{Iq}(i) log \frac{P_{Iq}(i)}{P_{Cq}(i)} \tag{7.10}$$

Equation 7.10 is also used to find the K-L divergence of the single consensus model (mean of all individual's posterior beliefs) from the individual's posterior beliefs, in which case there is only one value for $P_0'(X_q = T)$. Finally, K-L divergence is used for a second measure of accuracy. This measure indicates how similar the collective beliefs derived from an approximate aggregation algorithm are from the exact algorithm. I compute the divergence of the collective beliefs $P_C'(X_q)$; derived from the prior and incremental aggregation algorithms, from the collective beliefs $P_{\hat{C}}'(X_q)$; derived from the posterior aggregation algorithm using the following formula:

$$KL(P_{\hat{C}}'(X_q)||P_C'(X_q)) = \sum_{x_q} P_{\hat{C}}'(X_q = x_q) log \frac{P_{\hat{C}}'(X_q = x_q)}{P_C'(X_q = x_q)} \tag{7.11}$$

7.4.2 Models and data

The Bayesian networks structure for the experiments are randomly generated DAGs with the following parameters:

- *Number of nodes:* number of nodes in the network
- *Max parents:* maximum *indegree* for a node
- *Max children:* maximum *outdegree* for a node

The parameters for the networks were randomly generated using the Matlab Bayes Net Toolkit [43]. The probability distributions for each node in the network were sampled using a Gaussian mixture model, shown in Figure 7.2, centered around the means $[0.0, 0.25, 0.5, 0.75, 1.0]$ and adjusted to fall in $[0, 1]$. Since I am simulating elicited beliefs from humans, I used the Gaussian mixture model instead of a uniform distribution, since humans may be more likely to provide probability estimates near a common set of probabilities. Tools previously developed for belief elicitation used English terms such as "very unlikely," "unlikely," "fifty-fifty," "likely," "very likely," and converted the terms into these numeric values [27, 26, 25].

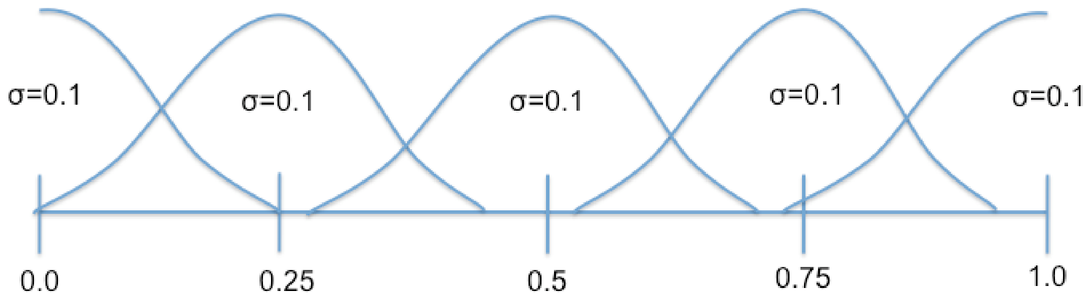


Figure 7.2: Gaussian mixture model used to generate probability tables for randomly generated networks; $\sigma = 0.1, \mu = \{0.0, 0.25, 0.5, 0.75, 1.0\}$.

7.4.3 Runtime

I now discuss experiments comparing the runtime of posterior aggregation with incremental aggregation, varying input parameters for Bayesian networks and population size. Since many of the experiments show the results of only one run with a given set of input parameters, I first show that runtime varies minimally when the same network structure and input parameters are used. The bottom table in Figure 7.3 shows runtime over five runs, the average of all runs, and the standard deviation. The top table shows the input parameters used for the experiments. Many of the following figures will contain this table, showing the input parameters used. *Node depth of query* refers to the maximum path length to a root node from the queried node. This value has an impact on the runtime of the aggregation algorithm. The columns in Figure 7.3 refer to the following runtimes:

- *binSpeed*: the runtime of the initialization step of the incremental algorithm. In this step each individual's probability table for each node is "binned" into a collective for the node. This is the runtime of the `binPT(PTs, IDs)` algorithm in Section 7.2.
- *infSpeed*: the runtime of the inference step of the incremental algorithm. This is the `variableElimAgg(PTs, Q, d)` described in Section 7.2.
- *incSpeed*: the total runtime for the incremental algorithm = *binSpeed* + *infSpeed*.
- *postSpeed*: the total runtime of the posterior aggregation algorithm, or `variableElim(PTs, Q, d)` + `binPT(PTs, IDs)`.

Nodes	Max parents	Node depth of query	Population
20	3	4	1000

Run	binSpeed	infSpeed	incSpeed	postSpeed
1	39.7664	222.63	262.3964	104.9661
2	37.8935	221.7353	259.6288	105.0301
3	37.8985	221.7721	259.6706	104.9917
4	37.9738	221.4475	259.4213	104.9834
5	37.8875	222.4427	260.3302	104.9551
Average	38.28394	222.00552	260.28946	104.98528
Stdev	0.829464667	0.504973189	1.226174644	0.028862987

Figure 7.3: The bottom table shows the standard deviation of runtime over five runs using the same network structure. The input parameters are shown in the top table.

The runtime values in Figure 7.3 show that the incremental algorithm is slower than the aggregation algorithm in some cases. In particular, the runtime is highly affected by the maximum number of parents, as discussed in Section 7.3. The table in Figure 7.4 shows the runtime when varying the maximum number of parents (and children). The runtime of the incremental algorithm is lower than the posterior aggregation algorithm when there is a maximum of two parents and children per node, but with the given input parameters, overtakes the runtime of the posterior algorithm with just three parents.

Nodes	Max parents	Node depth of query	Population
10	vary	3	1000

parents	binSpeed	infSpeed	incSpeed	postSpeed
2	14.6236	20.2877	34.9113	45.3685
3	18.7208	67.0646	85.7854	50.3992
4	23.9171	273.2118	297.1289	51.39
5	37.0291	541.1687	578.1978	52.1371

Figure 7.4: Comparison of runtimes when varying the maximum number of parents.

Nodes	Max parents	Node depth of query	Population
20	2	4	vary

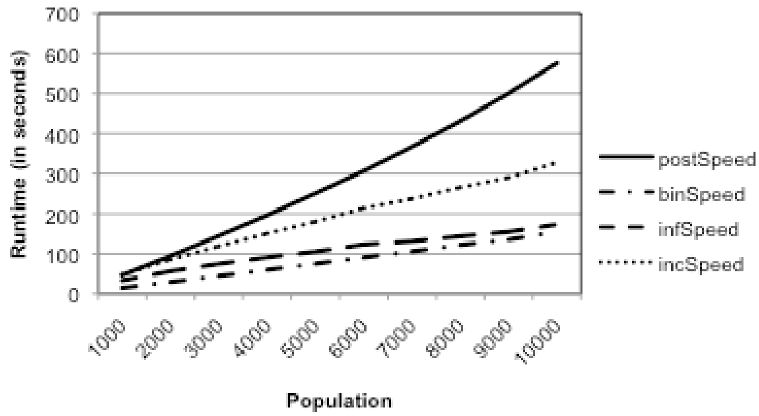


Figure 7.5: Runtime of algorithms when varying the population size, using a network with a maximum of two parents. Runtime was measured in population size increments of 1000.

Nodes	Max parents	Node depth of query	Population
20	3	4	vary

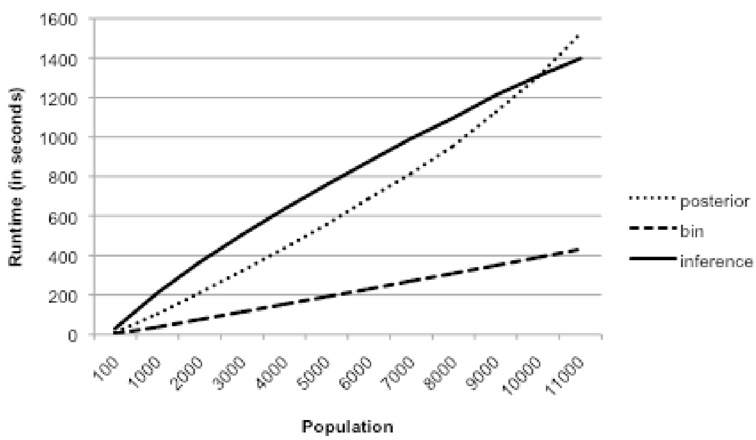


Figure 7.6: Runtime of algorithms when varying the population size, using a network with a maximum of three parents.

Nodes	Max parents	Node depth of query	Population
vary	2	varies	1000

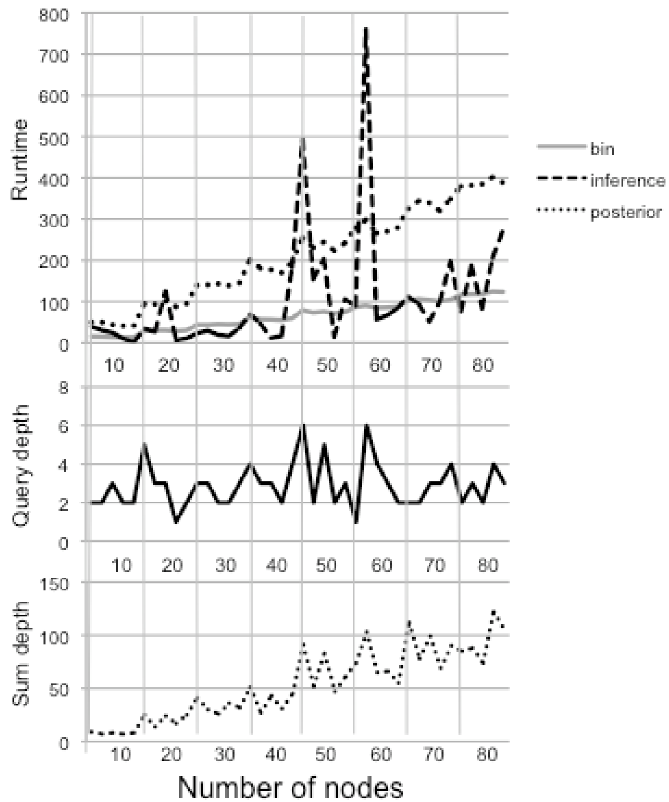


Figure 7.7: Runtime of algorithms when varying the number of nodes in a network as well as the depth of the queried node.

The graph in Figure 7.5 shows the results of varying the population m, from 1000-11000 in increments of 1000. The graph shows that the runtime of the incremental algorithm is lower and has a lower growth rate than the posterior algorithm when a maximum of two parents are allowed. The graph in Figure 7.6 shows that the initial runtime of the posterior algorithm is lower, but has a higher growth-rate than the incremental algorithm. The runtime of the posterior algorithm overtakes the inference step of the incremental algorithm at a population of around 10,000. Eventually the runtime of the posterior algorithm will overtake the total incremental runtime, given the growth-rates indicated by the graph. The slower growth rate is likely due to the cap on the number of possible collectives having been reached. While the population may be higher, the number of collectives reaches a limit based on the number of parents each node has.

The graphs in Figure 7.7 compare the runtimes over different networks when varying the number of nodes in the network. The randomly generated networks also resulted in varying topologies. The depth of the queried node is shown in the middle graph, and the sum of the node depths of each node in the network is shown in the bottom graph. The networks with a higher sum depth are more "stringy" than "bushy." Five experiments were run for each network size, between 10 and 80 nodes (shown in the horizontal axis). It is evident that the node depth has a strong effect on the runtime of the incremental aggregation algorithm. For most runs, the runtime of incremental inference (dashed line) is far lower than the runtime of the posterior algorithm (dotted line). However, when the query depth or sum depth spikes, much higher runtimes are seen. This is specific to the behavior of the variable elimination algorithm. In deeper networks, the number of probability tables added to each bucket is larger. While this has a minimal effect on the original variable elimination algorithm, it increases the size of the joint probability tables created during inference, and therefore increases the number of collectives.

Aside from the expected increase in runtime of the incremental aggregation algorithm with higher numbers of parents, the algorithm is also susceptible to the general topology of the network. Deeper networks are more problematic because the size of the probability tables. The number of possible collectives increases exponentially. I conclude that a different inference algorithm, such as belief propagation [50], may be more appropriate as an aggregation and collective discovery algorithm for general networks. This is because the size of the probability tables is only dependent on each node's immediate family, not its entire ancestry. Re-structuring the networks, as in [9] would also be a possible solution in the general case. Again, potential benefits to using the noisy OR-gate [50, 18, 68] or similar method to reduce the size of the probability tables can be seen.

7.4.4 Accuracy

I now compare the accuracy of the incremental aggregation algorithm with the exact posterior aggregation algorithm. Again, accuracy is the classification accuracy of the approximation algorithm, in which the approximation algorithm attempts to predict the collective to which each individual in a population belongs, based on a query. Equation 7.8 was used to measure accuracy in the following experiments. Figure 7.8 shows the accuracy of the incremental aggregation algorithm for a 20-node network with a maximum of three parents per node, and a population size of 1000. Select nodes were queried, and the accuracy of the query is shown next to the node. An accuracy of 0.81 or 81% means that 810 of the 1000 individuals were placed in their appropriate collective according to their posterior belief. Again, it is evident that node depth has a strong effect on the accuracy of the network. Nodes with a higher depth require propagation of collective beliefs along a longer path, decreasing accuracy at each node. It can be seen that the number of parents affects accuracy. For instance, the node labeled 18 has a higher accuracy than its single parent, who has three parents.

I was curious about the characteristics of the correctly versus incorrectly classified individuals. My theory was that individuals whose belief estimates hovered around 0.5 (often coined "on the fence") would be harder to classify because a small variance in belief has the potential to "flip" them from one preference order to another. Conversely, people whose beliefs tended to be more extreme, or "opinionated," would be easier to classify. To test this theory, two measurements were made of each individual's set of probability distributions:

1. *Distance from 0.5*: The average distance from 0.5 of each individual's probability estimate, over all n variables: $\frac{1}{n}\sum_{i=1}^{n}|P(X_i|Pa_i) - 0.5|$
2. *Entropy*: Entropy is a measure of uncertainty of a variable [61]. Since binary values were used, values towards 0.5 would be considered more "random" and have a higher entropy.

Figure 7.9 compares these measurements for incorrectly and correctly classified individuals using a two-node network. In this experiment, 92% of the individuals were correctly classified. It can be seen from the top graphs in the figure that incorrectly classified individuals always had a distance of ≤ 0.35 from 0.5 and an entropy of ≥ 1.25. Conversely, all individuals with the complement of these values were correctly classified. The histograms in the bottom of figure 7.9 show the distribution of correctly classified individuals for 11 bins over the range of distances. The line across the histogram approximates the rate of correctly

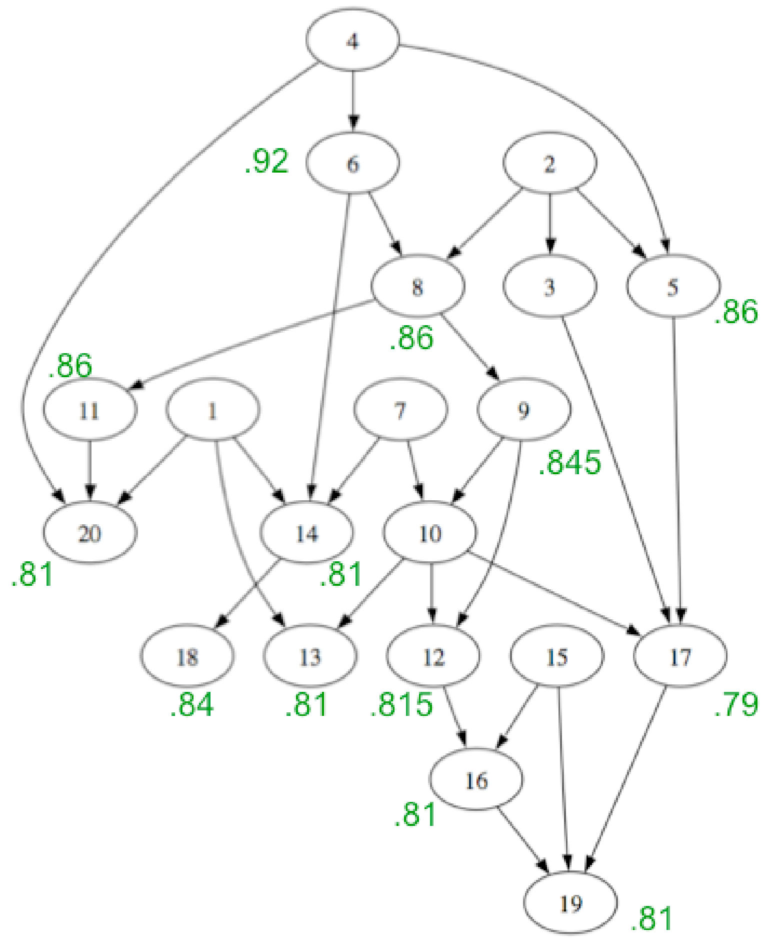


Figure 7.8: A 20 node network with up to three parents per node. The accuracy of the algorithm querying select nodes is shown next to the nodes.

classified individuals. Any of the bins that have values higher than the bar contain individuals that were classified correctly more often than average. Any bins with values below the bar contain individuals that were classified incorrectly more frequently than the average. Figure 7.9 shows that the values closer to 0.5 were classified incorrectly more frequently. The same conclusions can be made for the histogram over the entropy values. Figure 7.10 shows the same measurements for a four node network, whose accuracy was 82%. While the separation of always-correctly classified individuals was not as distinct in the larger network, more extreme probability estimates tend to be classified correctly more frequently.

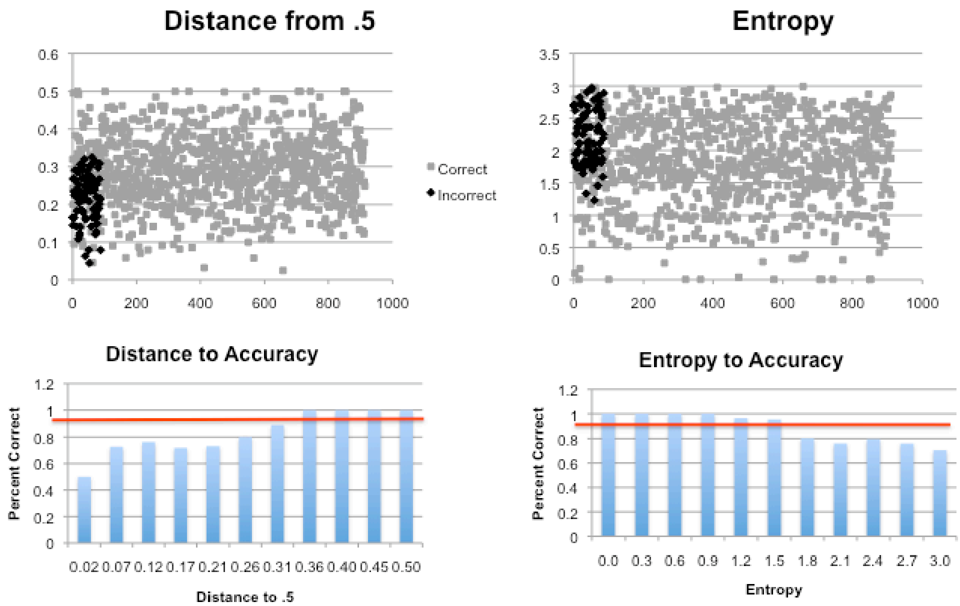


Figure 7.9: Measurements of the characteristics of individual's beliefs who were correctly classified versus incorrectly classified. A two-node network was used for these graphs.

These results demonstrate the validity of my hypothesis. I conclude that a potential aggregation approach that leverages these characteristics would be able to separate a population into groups based on the characteristics of their beliefs. Then one could apply incremental aggregation to the high distance groups, and posterior aggregation to those who are more difficult to classify.

	High Dist	Med Dist	Low Dist	total
size of group	225	453	322	1000
number correct	203	373	252	828
% correct	0.90	0.82	0.78	0.83

7.4.5 Prior Aggregation

I now compare the runtime and accuracy of the prior aggregation algorithm to the incremental and posterior aggregation algorithms. Figure 7.11 shows the results of varying k in the k-means clusterer. The posterior and incremental measurements are shown as horizontal

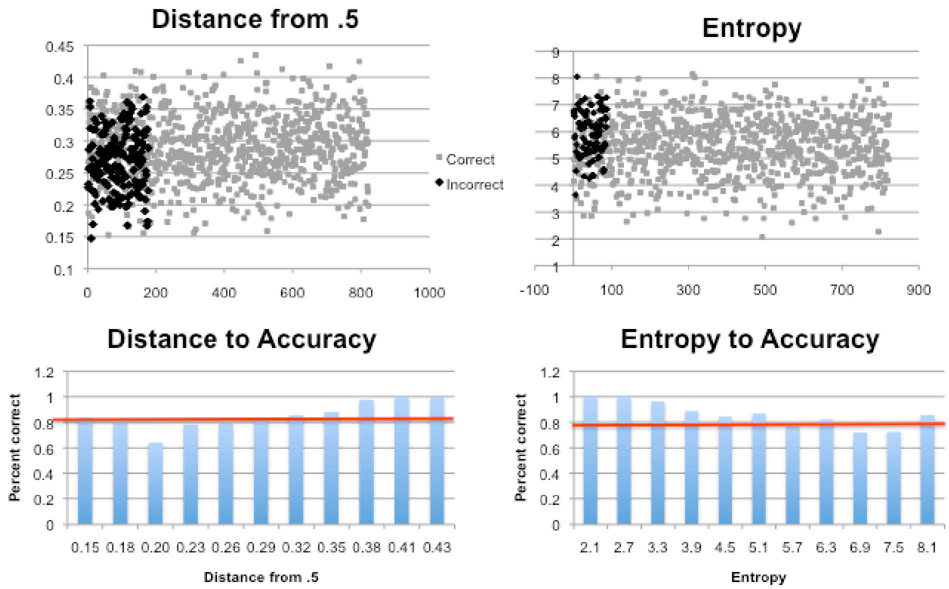


Figure 7.10: Measurements of the characteristics of individual's beliefs who were correctly classified versus incorrectly classified. A four-node network was used for these graphs.

lines in all graphs. The bottom two graphs only show lines for the incremental and prior aggregation algorithms because they are being measured against the posterior algorithm. The bottom graph shows the K-L divergence of the posterior probabilities for each collective, using equation 7.11. This is a second measure of accuracy and shows how similar the posterior probabilities from the estimated collectives are from the actual collectives. It can be seen that with lower values of k, the runtime of the prior algorithm is much lower than the other algorithms. However, the accuracy of prior aggregation is also much lower than incremental aggregation. Runtime and quality increase with the value of k, however, higher values of k can no longer be considered prior aggregation. For example, given a population of 1000, $k = 512$ is roughly equivalent to "aggregating" pairs of individuals and then running posterior aggregation on the set of pairs. In addition, the clustering algorithm has a restrictively high space requirement and hit memory limitations sooner than the incremental and posterior aggregations. The clustering algorithm considered all values in each individual's probability tables, therefore the dimensionality of the feature space was very large.

Nodes	Max parents	Node depth of query	Population
10	2	3	1000

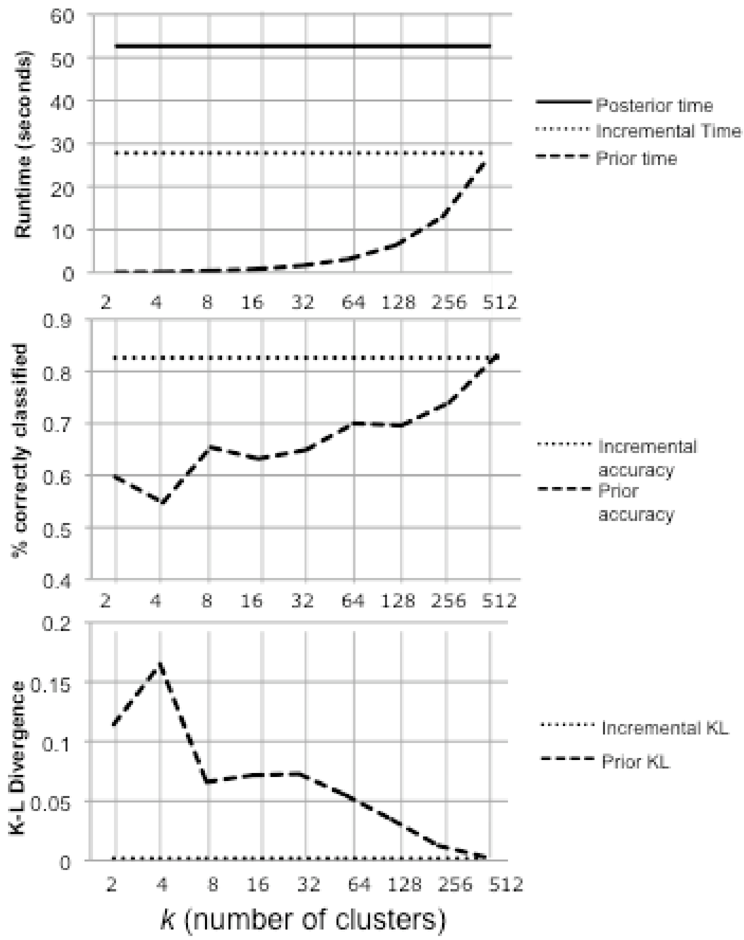
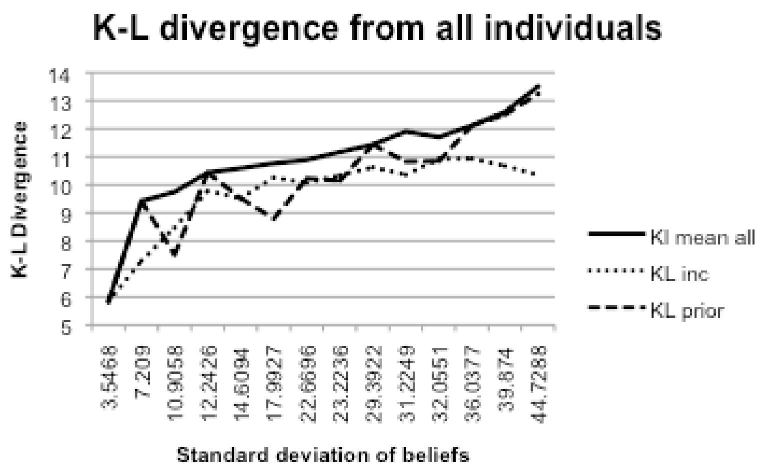


Figure 7.11: Runtime, classification accuracy, and K-L divergence accuracy of prior, incremental and posterior algorithms. The number of clusters k was varied and is shown in the horizontal axis.

7.4.6 Representation

The final experiment demonstrates that the collective belief aggregation approach is more "representative" of a population than a single consensus model. Figure 7.4.6 compares the K-L divergence, measured using equation 7.10, of the incremental aggregation algorithm (KL inc), the prior aggregation algorithm (KL prior), and the mean of all individuals (KL mean all), from the posterior probability of all individuals considered separately. The standard deviation of posterior beliefs varied from 3.5 to 44.7. In the experiments a high deviation indicates that the population is more polarized. The graph in Figure 7.4.6 shows that the K-L divergence of the incremental algorithm (in the dotted line) is always lower than or equal to the K-L divergence of the mean of all individuals (solid line). The K-L divergence of the prior aggregation approach is often lower, and never higher than the mean of all individuals. This graph demonstrates that the described approach results in a belief model that is a better representation of a population than a single consensus model. Figure 7.4.6 shows histograms of this experiment over the posterior probabilities in $[0, 1]$. The graphs are ordered from highest standard deviation of the posterior probabilities (highly polarized) to lowest (strong consensus). It can be seen from these graphs that the collective belief aggregation approach is more effective for representing polarized beliefs than the single consensus model. They also show that the approach is effective even in situations in which there is high consensus.

Nodes	Max parents	Node depth of query	Population
10	2	3	1000



7.5 On Prior versus Posterior Aggregation

This section compared prior aggregation with incremental and posterior aggregation, showing that prior aggregation results in very poor accuracy on queries. Nearly all previous aggregation methods discussed in this dissertation have used prior aggregation. The benefit is that a single model can theoretically represent a group or population of individuals. However, my research leads me to question the necessity of attempting prior aggregation over posterior aggregation. We see that posterior aggregation is more effective when one is interested in aggregating the beliefs of a group of individuals to determine the outcome of a situation. This is logical if one considers that different factors may play a role in the outcome of a situation. Shared belief in a final outcome does not mean that individuals agreed on all the factors that contributed to the outcome.

If one is still interested in forming consensus models from the posterior beliefs, they could work backwards from the queried variable, highlighting the similarities and differences between individuals that result in a shared belief in the outcome. Figure 7.5 shows an example of the collectives that emerge from varying beliefs on a relationship $X \rightarrow Y$. Each node represents a collective for a belief and is labeled with the preference order of the node's collective. The relative size of the nodes indicates the relative size of the collectives. The visualization was created by beginning at the collectives for the posterior belief in $P'(Y)$, represented by the nodes at the bottom, and moving towards the roots of the graph to show the collectives for the intermediate nodes. The middle nodes represent the collectives for the CPT of $P(Y|X)$. The top nodes represent the collectives for $P(X)$. This visualization of the collective belief model reveals interesting information about the agreement and disagreement between individuals. Divergence on beliefs that cannot be explained by the given model could expose hidden nodes indicating factors that should be included in the model.

Distributing the aggregation process may eliminate the need to find an efficient algorithm for aggregation. For example, I envision that the collective discovery approach will be used to aggregate the beliefs of a population elicited in a distributed environment. In this case, inference on each individual's network could be run at each node in the computer network, and the posterior results could be sent to another node (or nodes) to be aggregated.

7.6 Summary

This chapter described the incremental aggregation algorithm that approximates a posterior aggregation approach. Posterior aggregation means that collectives are formed after inference is run on each individual's network. Posterior aggregation results in a set of collectives that conform to the rational social choice properties defined in Section 4.3. The incremental algorithm was an attempt to create a more efficient algorithm. The variable elimination algorithm [68, 15] was extended because it is an exact inference algorithm and is relatively simple. Analysis showed that the incremental algorithm is highly susceptible to the structure of the network, in particular the number of parents each node has and the depth of the network. I conclude that a different inference algorithm, such as belief propagation [50] or utilizing combining rules such as the noisy OR-gate [18] may result in improved runtime. Alternatively, the networks could be restructured to reduce the number of parents each node has.

I discussed the loss in classification accuracy that is a result of incremental aggregation using some prior aggregation. I showed that even if two groups share a preference order over two sets of prior beliefs, finding the product of those beliefs can result in a different preference orders for the two groups. I analyzed the characteristics of individuals who were incorrectly classified— meaning that the collective they were estimated to belong to was not the same collective discovered using posterior aggregation. Analysis showed that individuals whose beliefs tended to be close to 50/50, indicating a lack of strong opinion, were misclassified more often than more opinionated individuals. This indicates that analyzing these individuals separately using posterior aggregation may improve overall accuracy and runtime.

Finally, I discussed the possibility that posterior aggregation may be sufficient and preferable in many situations, particularly if there is conflict in the beliefs and preferences. Distributed computing could replace the need for an efficient algorithm to run inference and aggregation on all individuals at once. This chapter successfully achieved the fourth and final objective of this dissertation, discussed in Section 1.3.4.

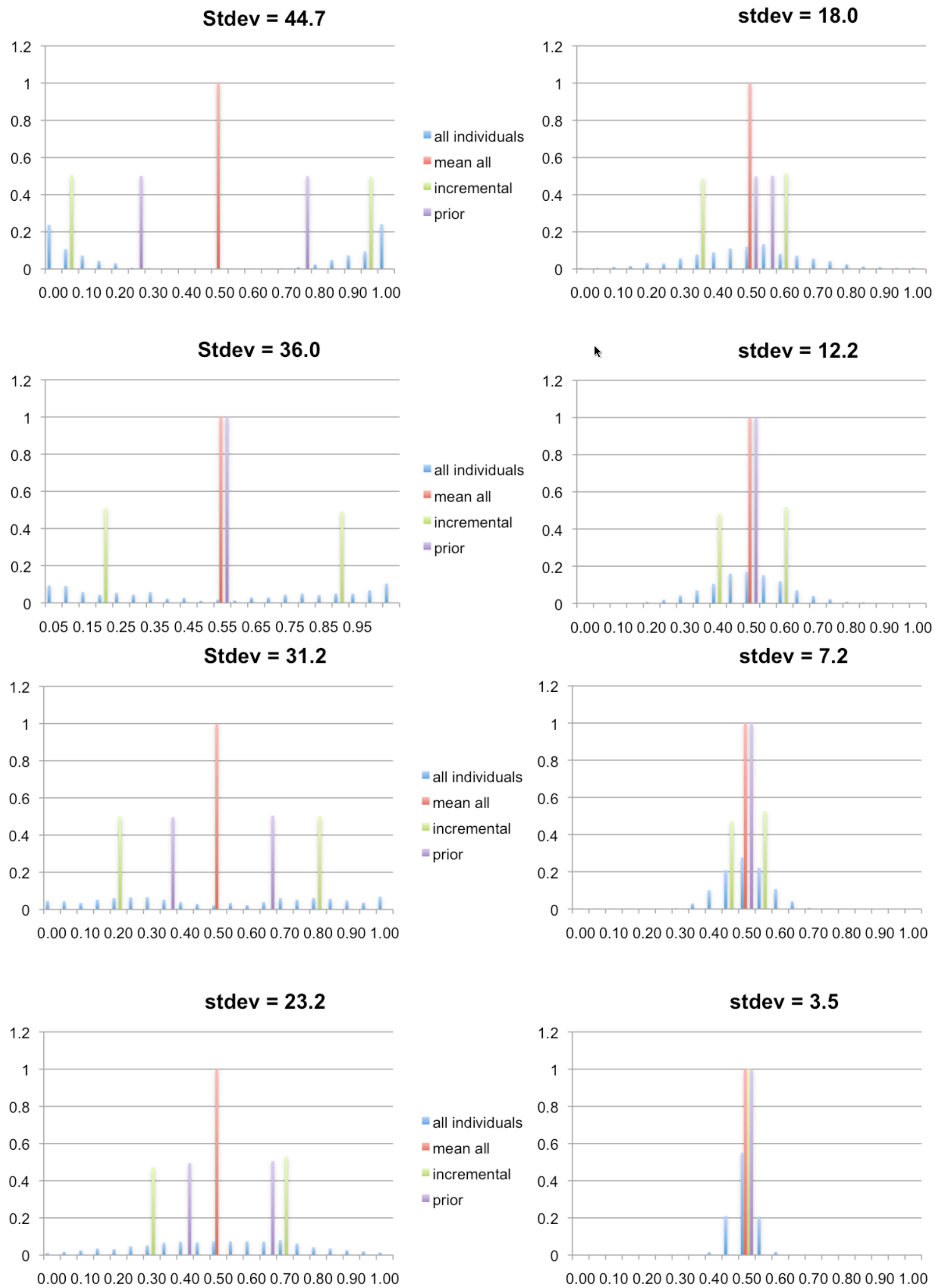


Figure 7.12: Histogram of the distribution of posterior probabilities, comparing the distribution over $[0, 1]$ for all individuals considered separately (all individuals), the collectives derived from incremental aggregation (Incremental), the collectives derived from prior aggregation (prior) and the single consensus model (mean all). The graphs are ordered vertically first and then horizontally, from high variance in the posterior probabilities, to low variance.

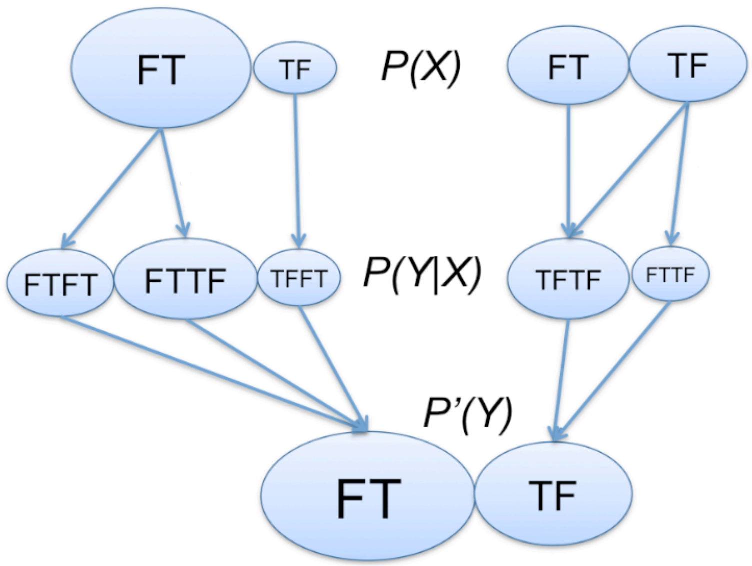


Figure 7.13: A visualization of the collectives for a small network representing $X \rightarrow Y$. The figure shows the posterior collectives for $P'(Y)$ and the collectives that formed the intermediate nodes.

Chapter 8

Summary and Future Work

8.1 Summary

This dissertation describes a new approach to aggregating the beliefs and preferences of many individuals to form *collective belief models* that capture the diversity of a population and enable game theoretic analysis for decision-making in a large population. Super-agents are formed from collective belief models that represent the preferences of all collective members in decision-making games. By allowing the competitive nature of a situation to emerge instead of "averaging away" the differences, the collective belief aggregation approach enables decision-making that conforms to mathematical principles of rationality defined by social choice and game theorists. Belief aggregation approaches that average the beliefs and preferences of all individuals may return inconsistent or irrational social choice solutions when faced with conflicting opinions. In contrast, the described approach elegantly handles these situations by returning a *set* of solutions that the collectives can evaluate using game theoretic analysis and negotiation. The following steps summarize the approach to social decision-making described in this dissertation.

1. Elicit beliefs and preferences from a population of individuals and form Bayesian belief or decision networks from these beliefs.

2. Extract the symbolic preference order used in social choice theory from the inferred

Bayesian beliefs.

3. Form *collectives* from the groups who share a preference order and compute their *collective belief*, which is the aggregate of the beliefs of the individuals in the collectives. It was shown that if a group of individuals share a preference order, their aggregate will uphold principles of rationality defined by social choice theorists.

4. Form a decision-making *game* in which each of the "players" is derived from one of the collectives. Game theoretic analysis can then be applied to find Pareto optimal solutions, Nash equilibria and other solutions and strategies.

The four objectives set forth in Section 1.3 were successfully achieved. Section 5.4 demonstrated that the high-level approach to separating a population of individuals based on their beliefs and preferences before performing aggregation. This caused opposing objectives to emerge that were not represented by the consensus model. Kullback-Leibler divergence was used to show that the described approach forms a more representative model than the single consensus approach.

Section 4.3 defined the building blocks of a new type of consensus model called *collective belief models*. *Rank order collectives* were defined to be groups of individuals that share the same preference order over a set of symbolic or Bayesian options. A *collective choice function* was then defined that will form a consensus belief model from a population of individuals. It was shown that the aggregate, *collective belief* of each collective upholds the rational social choice principles defined by social choice theorists.

Section 6 described decision-making games formed from "super-agents" that represent each collective that emerges from a population. I introduced an algorithm to find the Pareto optimal solutions given a set of collectives and their collective beliefs. I demonstrated that the set of Pareto optimal solutions can be found using collectives even in situations in which the single consensus approach returned results that were not Pareto optimal solutions. Multi-agent influence diagrams were extended with the collective choice function to show the Nash equilibrium solutions and minimax strategies.

Section 7 discussed the differences between prior and posterior aggregation and introduced a hybrid approach called *incremental aggregation*. It was shown that the incremental aggregation algorithm's runtime can be lower than the brute force (posterior aggregation) algorithm given certain constraints on the structure of the networks.

8.2 Future Work

The research described in this dissertation forms the foundation for continued research in the area of social decision-making. With continued research into game theory and negotiation techniques I hope to invigorate research into decision- and policy- making techniques that incorporate the beliefs and preferences of stakeholders.

The described research originates from computer science, artificial intelligence and probabilistic reasoning, but crosses disciplinary boundaries into the social sciences— particularly sociology, political science and economics. Continued research relevant to computer scientists includes utilizing collective belief models to combine evidence from multiple sources. This is relevant to sensor networks, robotics, multi-agents and any automated techniques that combine heterogeneous knowledge sources. In addition, investigating the nature of divergence could have implications for causal inference and learning the structure of large networks. For example, hidden nodes may be revealed by observing that a single node has varying effects on another.

Continued research relevant to the social sciences includes expanding collective belief models to incorporate the factors that contribute to human beliefs and preferences. In particular social scientists may be interested in the variations in backgrounds that describe divergent beliefs and the commonalities that form a common ground between individuals. Economists may be interested in utilizing computational models to analyze the competitive and cooperative behavior of societies. Political scientists could use collective belief models to more fully understand the preferences of voters and to analyze the behavior of voters over time, including factors that cause individuals to switch preferences. Policy-makers could use my approach to analyze the effects of policy changes on populations and organizations. Organizations could use the techniques described to make decisions given the direct input of their members. Corporations could expand awareness of their customers and competitors. Finally, greater understanding of human behavior and a more structured approach to social decision-making could improve global relations and help resolve conflicts.

The following list enumerates some potential research directions:

- Combining belief aggregation, structure learning and elicitation techniques may help capture the factors that contribute to beliefs and preferences and form more complete and robust models. Expanding the models may reveal that apparent irrationality in human behavior can be explained by not fully representing all the factors in human decision-making. In addition, stabilizing the belief models so that they capture the temporal effects of public opinion such as group-think, hype and panic may improve understanding and decision-making.

- In the approach described in this dissertation preference (rank) orders are deterministic. Investigating probabilistic preference orders that would form a distribution over all possible preference orders may help capture the inconsistencies in human belief over time and in different situations. In addition, probabilistic preference orders would help differentiate individuals who lean towards indifference from those with strong opinions.

- A hybrid approach using clusters of consensus and collectives may be appropriate in situations where a set of rational social choice solutions is desired as well as a good representation of the population's beliefs. One could first discover the collectives and then find any significant clusters within each collective. The drawback would be an increased number of collectives and a less concise representation for game theoretic analysis.

- Further investigation into the different outcomes that arise from the average expected utility and individual utilities, discussed in Sections 3.3.3 and 6.5 may be relevant to the game theoretic concept called the *price of anarchy*. Putting a "price" on selfish behavior seems to imply that the social optimum is more desirable than selfish behavior. However, in this dissertation I have demonstrated the dangers of relying only on the social optimum. Perhaps collective belief models offer a compromise between altruism and selfishness.

- While the approaches described in this dissertation utilized probabilistic beliefs and preferences, the collective belief aggregation approach could also be used for deterministic preferences. Further research into voting and social choice theories might reveal additional areas where the approach can be applied.

- The decision networks and examples in this dissertation did not take into consideration the effect of an individual's community on their own beliefs and desires. For example, a family member might adjust his preferences in order to please another family member. The richness of models could be greatly increased by connecting the networks of individuals to those in their communities. Belief propagation could occur across the individuals' networks, and the results of inference could be aggregated after convergence. In this manner, one could model the naturally occurring consensus that occurs in social groups before discovering collectives.

- Investigation into negotiation techniques could enable collectives to come to an agreement and make decisions when conflict arises. By observing the similarities and differences between collectives that emerge from multiple issues, one may gain a deeper understanding of conflict and consensus in societies. These issues may indicate the decisions that require the most attention, and where negotiation techniques could have the most effect. In addition, the size of the collectives could play a significant role in negotiation techniques.

Appendix A

An Algorithm to Find Pareto Optimal Solutions

The procedure `findParetoSet`(T, Φ, O) finds the set of options in O that meet the strong Pareto condition by first finding the solutions that do not. An option o_a does **not** meet the condition if there exists an option o_b such that someone prefers o_b to o_a and no one prefers o_a to o_b. The following requirements, defined by [56], are for options that do **not** meet the strong Pareto condition. They are used to guide the algorithm.

Definition 24. *Strong Pareto Condition: An option o_a does not meet the strong Pareto condition if and only if there is another option o_b such that:*

1. *Everyone either prefers o_b to o_a or is indifferent to them ($\forall_i$, $o_b P_i o_a$ or $o_b I_i o_a$), and*
2. *Someone strictly prefers o_b to o_a ($\exists_i$ s.t. $o_b P_i o_a$)*

The `findParetoSet`(T, Φ, O) algorithm, emulates this definition by finding the set of options that do not meet the strong Pareto condition. The complement of this set is the set of strong Pareto optimal solutions. The pairwise relation D was added to mean *dislikes*, the opposite of P (prefers), as defined in Section 4.4. R contains the preference relations between a candidate option o_a and another option o_b for each collective. The test in the line marked with (1) in the `findParetoSet`(T, Φ, O) algorithm will determine if the option o_a does not meet the strong Pareto condition.

Procedure `findParetoSet`(T, Φ, O):

- Inputs:
  - T: a partition containing a set of collectives
  - Φ: a set containing each collective's collective belief
  - O: the set of options
- Outputs: The set of Pareto optimal solutions for the partition

```
/* relations = the possible pairwise relations:
P = strict preference, D= dis-preference, I= indifference*/
relations={P, D, I}

S_np = findParetoSet(T, Φ, O)
    /* S_np is the set of non- Pareto options*/
    S_np = {}
    for each option o_a in O
        for each option o_b in O (a ≠ b)
           /* R will store each collective's preference
           relation for o_b to o_a*/
           R = {}
           for each collective C_j in T
              /*place C_j's preference between o_b and o_a in R*/
              R(j)= preference(φ_j, o_b, o_a)
           end
           /*Check the conditions for o_a not to be Pareto
           (P ∈ R) means someone prefers o_b to o_a
           (D ∉ R) means no one prefers o_a to o_b*/
           if (P ∈ R) and (D ∉ R)
              S_np = S_np ∪ o_a; break
           end
```

```
      end
    end
    /*The options that meet the Pareto condition are
    those in O that are not in S_np*/
    S_sp = O - S_np
end
```

Procedure `preference(`ϕ, o_1, o_2) is a helper function for findParetoSet(T, Φ, O):

- Inputs:
  - ϕ: a collective belief
  - o_1, o_2: two options to be compared in ϕ
- Outputs: The relation indicating preference for o_1 compared to o_2

```
/*returns collective C's preference between options o1 and o2*/
function r = preference(φ, o1, o2)
    if o1 < o2 in φ /* o1 P o2 */
      r = P
    else if o2 < o1 in φ /* o2 P o1 */
      r = D
    else
      r = I
    end
end
```

The `findParetoSet()` procedure creates an $r \times r \times k$ matrix, for r options and k collectives. The cell (a, b, j) indicates collective $C_j's$ preference relation for o_a and o_b. If C_j prefers o_a to o_b, then $o_a P_j o_b$ and cell $(a, b, j) = P$. Otherwise, if collective C_j prefers o_b to o_a, then cell $(a, b, j) = D$. If C_j is indifferent to the options, then cell $(a, b, j) = I$. The algorithm determines if o_a is non-Pareto by comparing o_a to each other option o_b. If the column $(b, a, *)$ does not contain D and it contains at least one P, then o_b is always preferred or indifferent to o_a, meaning that o_a is not Pareto. The set of Pareto solutions is the complement of the set of non-Pareto solutions. Since the strong Pareto solution is a superset of the weak Pareto solution, the algorithm can be used to find either one. Note that a Pareto optimal solution cannot be extracted from a weak partition T_w using this approach, because one cannot distinguish preference and indifference.

Appendix B

A Decision Model for Stem Cell Research

This appendix describes an experiment applying the clustering and collective belief aggregation approaches to actual human beliefs. The beliefs and utilities were elicited from people using Mechanical Turk, a website for hiring people to do simple online tasks for very low cost. The issue of stem cell research was selected because of the issue's polarizing effect on individuals with diverse backgrounds and motivations. The hypothetical decision in this experiment is whether the (US) government should fund embryonic stem cell research, adult stem cell research, both, or neither based on opinions collected from an online survey.

B.1 Decision Model and Data

The stem cell dataset is composed of opinions elicited from 293 people using Mechanical Turk and Survey Monkey (surveymonkey.com). Fig. B.1 shows the questions and their options that composed the survey. Individuals were asked to select one of the options for each question. The answers provided by the individuals were used to form the Bayesian decision network in Fig. B.2. The Key in Fig. B.1 indicates the corresponding node in the decision network.

Key	Question	Options
GF	Should the government fund stem cell research?	a. Yes, the government should fund embryonic and adult research b. The government should only fund embryonic stem cell research c. The government should only fund adult stem cell research d. The government should not fund any stem cell research
PE	What is the potential for embryonic stem cell research to cure diseases and provide new methods to test cures?	a. The potential is very high (90%) b. There is some potential (75%) c. It is not clear yet what the potential is (50%) d. It is unlikely to provide cures (25%) e. It is very unlikely to provide cures (10%)
PA	What is the potential for adult stem cell research to cure diseases and provide new methods to test cures?	a. The potential is very high (90%) b. There is some potential (75%) c. It is not clear yet what the potential is (50%) d. It is unlikely to provide cures (25%) e. It is very unlikely to provide cures (10%)
GI	Would government funding of stem cell research improve its potential?	a. Yes, government funding is important to advance research (99%) b. No, private funding is sufficient to advance research (1%)
EI	Are you concerned about the ethical issues of using embryonic stem cells?	a. I am very concerned. The potential does not outweigh the ethical issues. (-20) b. I am a little bit concerned. I think there is potential, but some issues need to be addressed. (-10) c. I am not very concerned. I think the potential outweighs any ethical issues. (0)
MR	How important is advancing medical research and curing diseases in general to you?	a. Very important. (+15) b. Somewhat important. (+10) c. Not very important. (0)
KS	Do you know someone who could potentially benefit from stem cell research?	a. Yes, myself or a loved one (+15) b. Yes, a friend or associate (+10) c. Not personally(0)

Figure B.1: Elicited beliefs for stem cell research model.

Fig. B.3 shows the probability distribution for each question (also referred to as belief) given all individuals surveyed. Each cluster of bars represents one belief and its options. For instance, the first belief labeled with **GF** indicates the preferred options for the question "Should the government fund stem cell research?" The first bar to the left is the probability for decision option a in Fig. B.1. The second bar corresponds to option b and so on. There is one bar for each option provided in a question. For consistency, the distribution for each question was over five options. However, not all questions had five options. In this case, the non-options are represented by cells containing the probability 0.0001.

The first observation is that there is significant divergence across most of the decision options. In particular, the question labeled, **EI**, representing ethical issues with stem cell research, has a relatively normal distribution across all of its options. On the other hand,

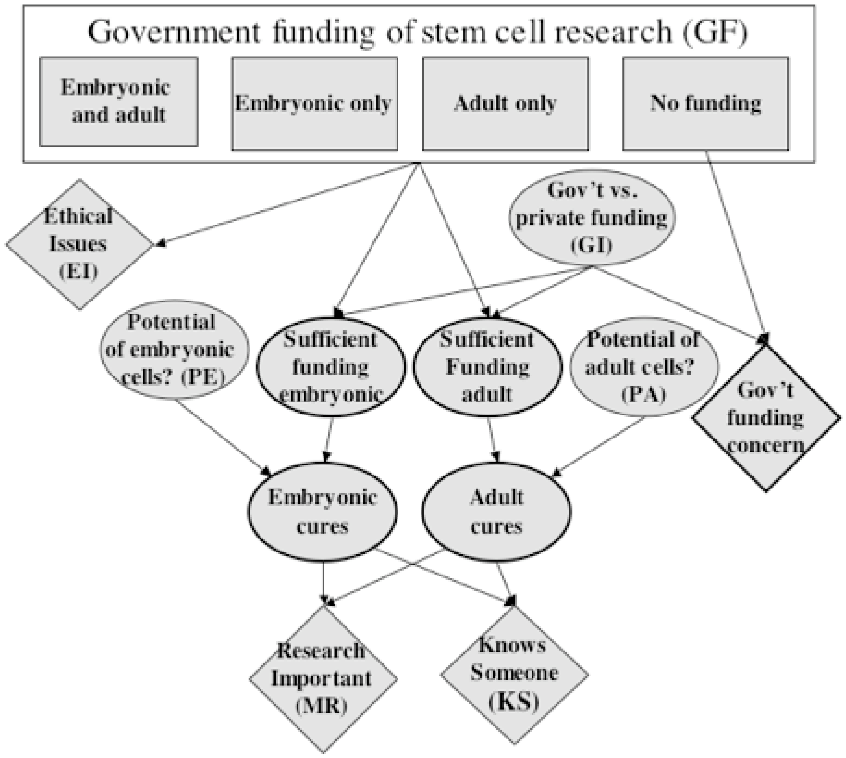
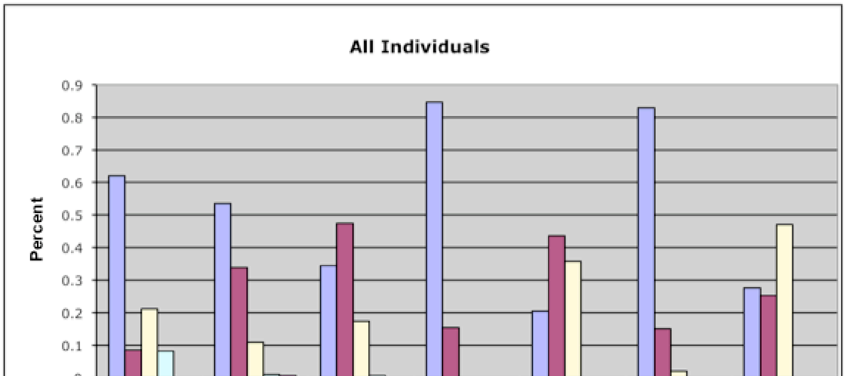


Figure B.2: An influence diagram for the stem cell research issue.

questions such as the one labeled **GI**, representing the importance of government research investment, strongly favors one option over the other. This indicates that there is more consensus on **GI** than **EI**.

This experiment first clustered the individuals based on the options they selected for the questions using the k-means clustering algorithm. Figures B.4- B.7 show the probability distribution for each cluster. Each cluster has been labeled with its defining characteristic.



	GF	PE	PA	GI	EI	MR	KS
Option 1	0.621	0.5358	0.3446	0.8461	0.2047	0.8291	0.2764
Option 2	0.0853	0.3379	0.4743	0.1535	0.4367	0.1501	0.2525
Option 3	0.2116	0.1092	0.174	0.0001	0.3583	0.0205	0.4708
Option 4	0.0819	0.0102	0.0068	0.0001	0.0001	0.0001	0.0001
Option 5	0.0001	0.0068	0.0001	0.0001	0.0001	0.0001	0.0001

Figure B.3: The probability distribution for each belief given all individuals surveyed. The beliefs are shown in the x axis, with each color indicating a different option. The abbreviations under each cluster of bars correspond to the key in Fig. B.1

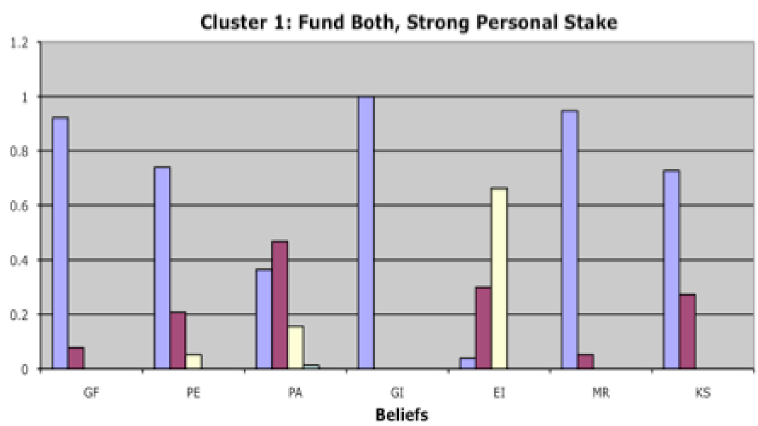


Figure B.4: The probability distribution for the first cluster, indicating strong support for embryonic research and a high personal stake in the outcome.

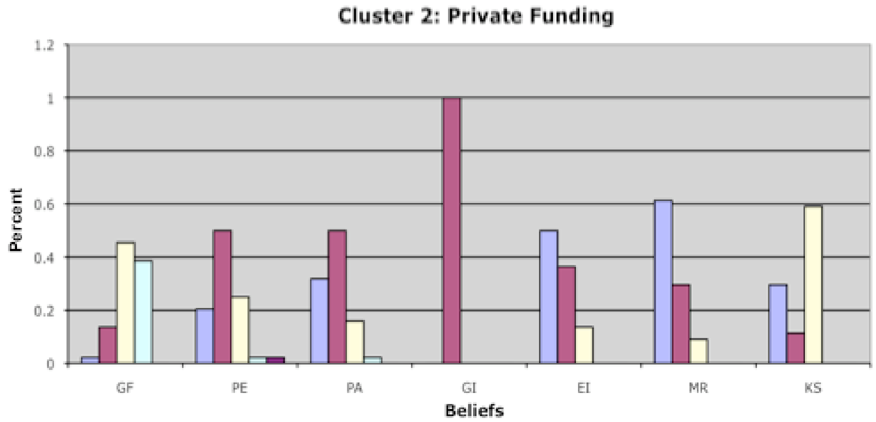


Figure B.5: The probability distribution for the second cluster, which overwhelmingly supports private funding for research.

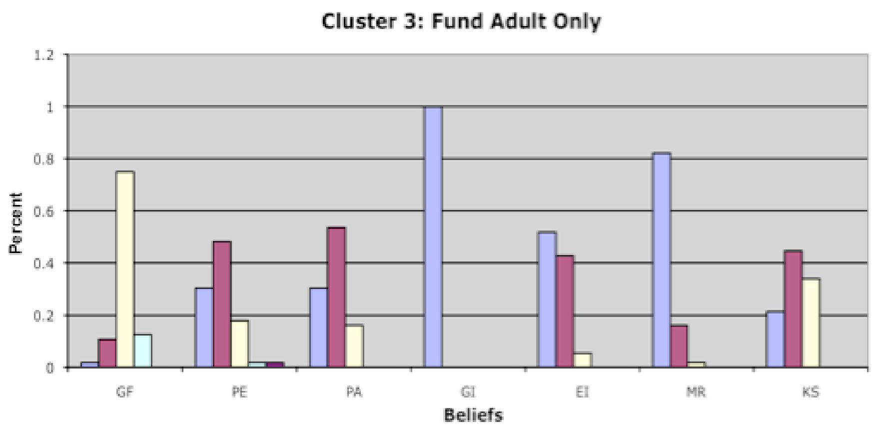


Figure B.6: The probability distribution for the third cluster, preferring that the government only fund adult stem cell research.

B.2 Inference

The next step in the experiment was to run inference on each of the individuals' networks within a cluster to find the expected utility of each decision option for each individual. The final step was to aggregate the expected utilities in a cluster to determine the cluster consensus. Before inference, each option selected by an individual in the survey was converted into a numeric value (shown next to each option in Fig. B.2). Appropriate utility values were determined from experimentation. The inference algorithm was run on each individ-

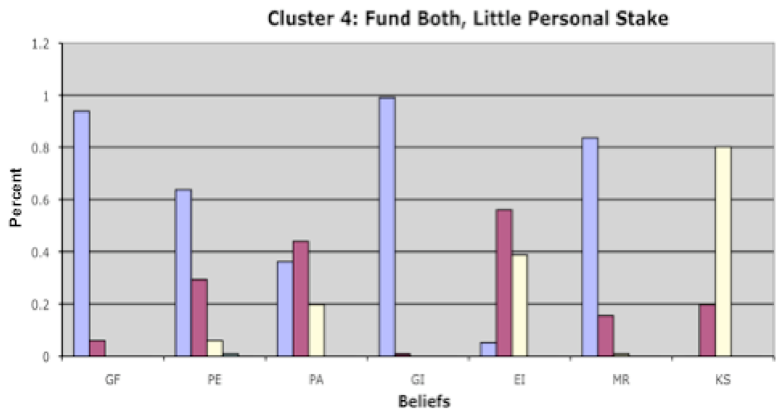


Figure B.7: The probability distribution for the fourth cluster, favoring funding for both types of research, but with a lower personal stake than in cluster 1.

ual's supplied beliefs given a set of candidate utility values. If the decision option that the individual had selected matched the highest utility option, this was considered a *correct* prediction. A number of candidate utility values were tested until the highest number of correct predictions was reached.

Once each individual's selected options were converted to their numerical counterparts, the expected utility of each decision was computed through inference, as described in [59]. Finally, the consensus belief (aggregate) for each cluster was computed as the arithmetic mean of the expected utilities for each individual in the cluster.

B.3 Comparing single consensus versus consensus cluster results

The average expected utilities for all individuals is shown in Table B.1. According to these values, the social choice is to "fund both embryonic and adult stem cell research." While the single consensus approach determines the optimal social choice option given the averages of the beliefs, it does not represent a deep understanding of the underlying opinions. In fact, there is a significant subset of the population (8%) who believe that the government should not fund any research according to Fig. B.3. While a majority of the surveyed population

Decision option	**Utility**
Fund embryonic and adult stem cell research	*16.3*
Fund embryonic stem cell research only	6.1
Fund adult stem cell research only	13.8
Do not fund any research	3.8

Table B.1: Average expected utility for each decision option.

may believe that the government should fund both embryonic and adult stem cell research, policy-makers may also be interested in who is likely to oppose such measures.

Decision	**Cluster**				**Pareto Optimality**
	Cluster 1	**Cluster 2**	**Cluster 3**	**Cluster 4**	
Fund Both	38.64	-19.42	14.81	16.74	Optimal
Fund Embryo only	18.82	-9.29	-0.15	5.75	Not optimal
Fund adult only	19.83	3.35	14.97	11.17	Optimal
Fund neither	0.01	19.45	0.01	0.17	Optimal
Characteristics					
Size	77	44	56	116	
Percent	26%	15%	19%	40%	
Optimal solution	Fund both	Fund neither	Adult only	Fund both	
Worst solution	Fund neither	Fund both	Embryo only	Fund neither	

Figure B.8: Overall utility results using belief clusters and some characteristics of the clusters.

Fig. B.8 shows each cluster's expected utility for each decision option. I make several observations about these results:

- The optimal solution for the whole population is in fact the worst case scenario for a significant minority (15%) of the population.
- Cluster 2 is in complete opposition to the social choice.

These observations successfully demonstrate two expected situations, listed in objective 1.3.1. The clustering approach also enables one to visualize the Pareto optimal solutions.

The single consensus approach, represented by the average expected utility in Table B.1 is unable to distinguish these solutions. Nor does the social choice function distinguish the opposition that occurs in this relatively simple experiment.

B.4 Measuring the *representation* of a population

Finally, the K-L divergence measure was utilized to determine whether the consensus clusters more accurately represent the beliefs of a population than the single consensus approach. Figures B.9-B.12 show the K-L divergence of the expected utility for each decision option. These graphs represent histogram over the range of expected utilities from lowest (left) to highest (right). Each cluster of bars represents the number of individuals whose expected utility fit into that bin. The K-L divergence of the single consensus approach (mean) was compared to the K-L divergence of the cluster consensuses using four and eight clusters. As in the example described in Section 4.2, K-L divergence is a measure of how much the distribution of the consensus diverges from the distribution of all individuals' expected utility.

In each figure, the bars labeled **All** are the distribution of all individuals over the range of expected utility. The bar labeled **mean** is the single consensus, or average expected utility. The bars labeled **8 clusters** and **4 clusters** are the distributions of the cluster consensuses. K-L divergence of each group from **All** is shown in the bottom right of each figure. In most cases, a greater number of clusters typically reduces the K-L divergence. However, in Figure B.10, the K-L divergence of four clusters was slightly lower than eight clusters. This may be an indicator that additional clustering fails to discover more significant divergence. Also, Figure B.12 shows that the mean actually has a lower K-L divergence than four clusters and is quite close to that of eight clusters. In this case there is a fairly strong consensus on the expected utility, indicating that clustering may not be required to achieve a good representation.

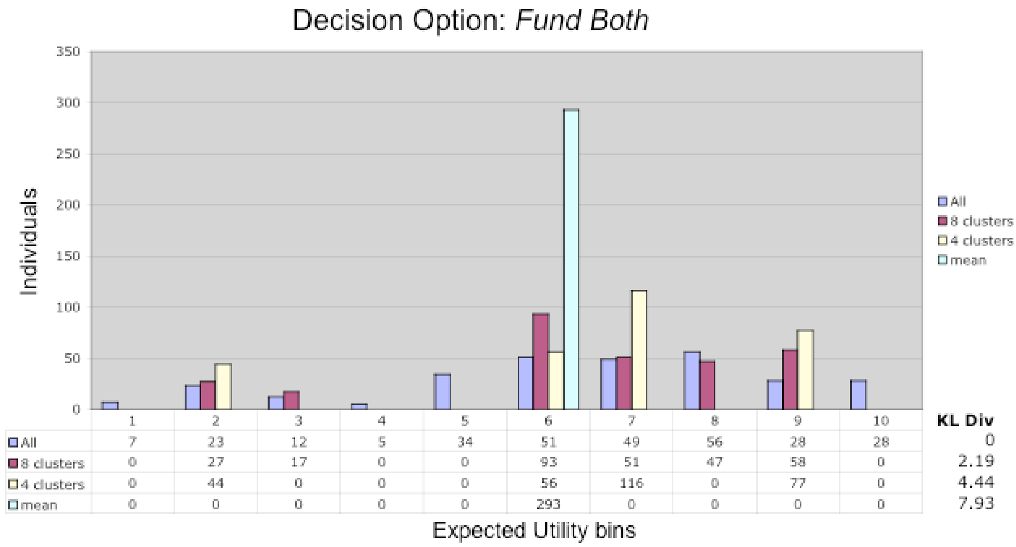


Figure B.9: Histogram of the expected utilities for the decision option *Fund Both* over the range of expected utilities. K-L divergence of 8, 4 and 1 cluster from the distribution of all individuals is shown in the bottom right.

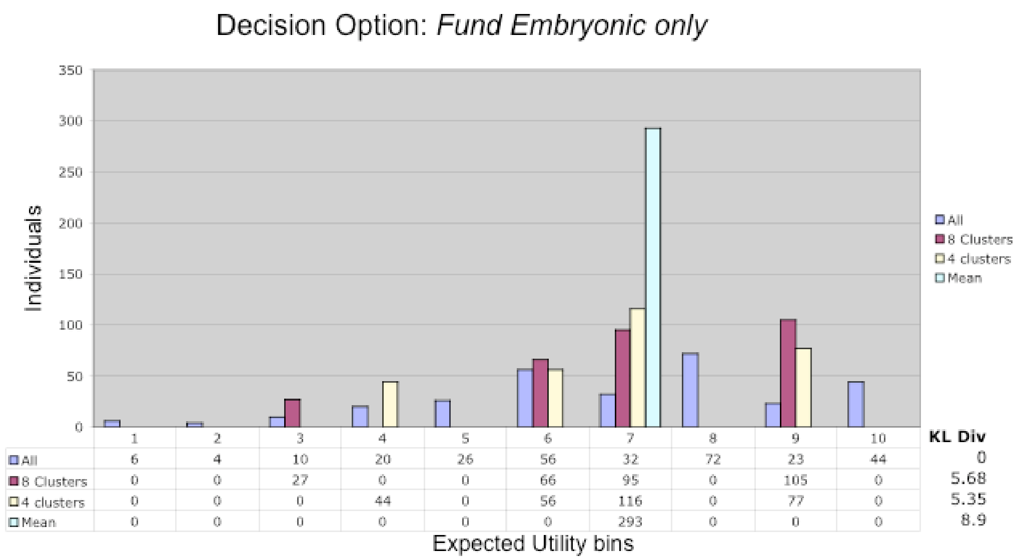
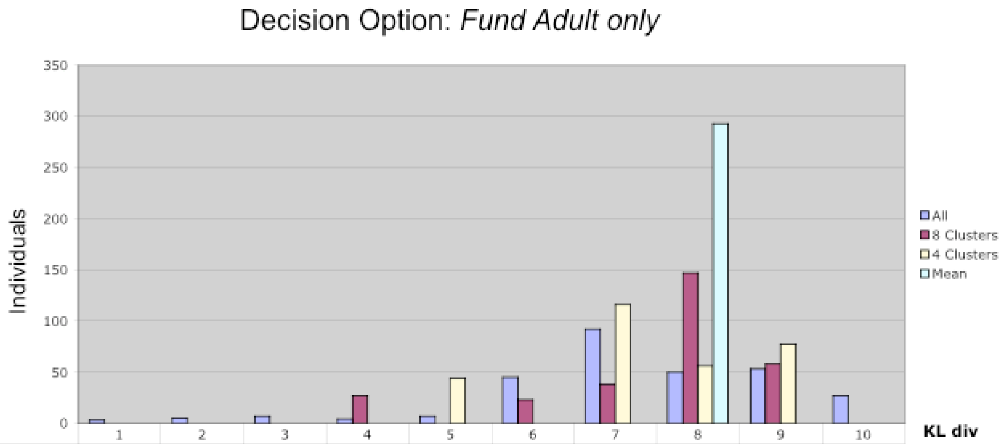


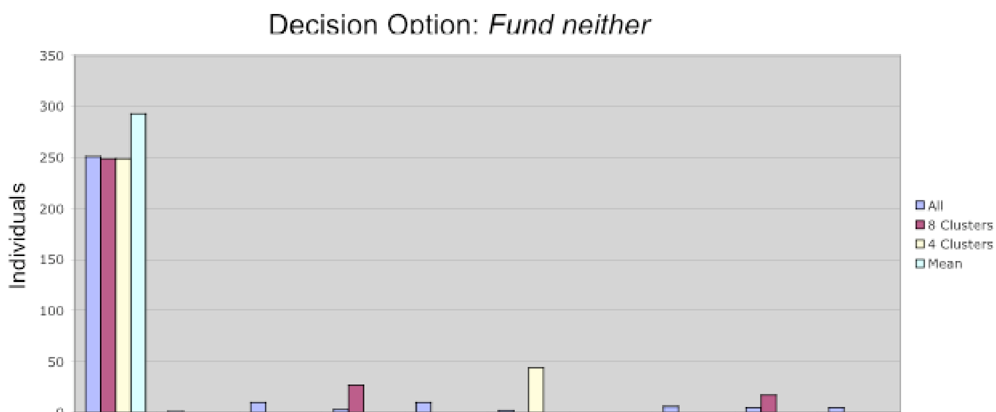
Figure B.10: Histogram of the expected utilities for the decision option *Fund Embryonic Only* over the range of expected utilities. K-L divergence of 8, 4 and 1 cluster from the distribution of all individuals is shown in the bottom right.



	1	2	3	4	5	6	7	8	9	10	KL div
All	3	5	7	4	7	45	92	50	53	27	0
8 Clusters	0	0	0	27	0	23	38	147	58	0	1.7
4 Clusters	0	0	0	0	44	0	116	56	77	0	2.72
Mean	0	0	0	0	0	0	0	293	0	0	8.37

Expected Utility bins

Figure B.11: Histogram of the expected utilities for the decision option *Fund Adult Only* over the range of expected utilities. K-L divergence of 8, 4 and 1 cluster from the distribution of all individuals is shown in the bottom right.



	1	2	3	4	5	6	7	8	9	10	KL Div
All	251	1	10	3	10	2	0	6	5	5	0
8 Clusters	249	0	0	27	0	0	0	0	17	0	0.87
4 Clusters	249	0	0	0	0	44	0	0	0	0	1.05
Mean	293	0	0	0	0	0	0	0	0	0	0.93

Expected Utility bins

Figure B.12: Histogram of the expected utilities for the decision option *Fund Neither* over the range of expected utilities. K-L divergence of 8, 4 and 1 cluster from the distribution of all individuals is shown in the bottom right.

B.5 Stem Cell Collectives

Finally, the posterior collective belief aggregation algorithm was run on the expected utilities for the decision options. The rank order for each collective uses the following symbols to represent each decision option:

Symbol	Decision Option
B	Fund both embryonic and adult
E	Fund embryonic only
A	Fund adult only
N	Fund neither

Table B.2 shows that ten out of the 24 possible collectives emerged from the population and their collective beliefs. The majority of the population fit into the top five collectives, with a few stragglers in their own collectives. Each collective is guaranteed to accurately reflect the relative preferences of its members. For instance, the table shows that all decision options are Pareto optimal, since there is no option that everyone prefers over another. In contrast, the clustering approach inaccurately inferred that everyone preferred A over E (see Figure B.8). In a real situation, one could elicit beliefs from a larger population and use the collectives to apply game theoretic analysis and negotiation techniques to help form policy.

Collective Rank Order	**Collective Belief** B	E	A	N	**Collective Size**
$BAEN$	28.2	13.4	18.6	3.8	144
$ABNE$	10.3	-1.2	14.2	2.8	45
$NAEB$	7.4	7.3	20.7	26.3	45
$BEAN$	31.5	19.9	15.2	3.6	37
$ANBE$	-1.6	-8.9	9.0	1.7	15
$NABE$	-18.9	-24.9	-7.2	1.8	2
$BNAE$	8.1	-3.6	-0.17	3.1	2
$BANE$	16.3	-6.5	13.5	5.7	1
$BENA$	23.2	8.2	4.2	4.3	1
$NBAE$	-1.5	-7.5	-7.5	1.5	1

Table B.2: The collectives that emerge from a population based on the expected utility of the stem cell decision options.

Appendix C

An Election Polling Simulation

The experiment in this section demonstrates an election polling simulation, in which the strategy of an individual is to support the candidate she intends to vote for. In this simple simulation, an individual may change her strategy depending on the expected utility of the candidates. The expected utility will be dependent in part on an individual's belief in the candidate's likelihood of winning, which depends on how many people plan to vote for each candidate. In this simulation, the Nash equiliqrium solution is determined only after the game has "stabilized." Stability is reached only when individuals cease to change strategy for a number of moves.

In the election polling simulation there are three candidates, two sharing a majority of the vote within a few percentage points of each other, while a third has a small minority. The simulation shows that the collective choice function is a better predictor of the voting behavior of a population than a social choice function that finds a single consensus. In fact, the social choice function *incorrectly* predicts the outcome of the election in a contrived (but believable) situation. In particular, the social choice, or average expected utility, does not predict the same outcome as would the individuals voting for the candidate with the highest personal expected utility. This phenomenon was discovered in this simulation, and was discussed in general in Section 3.3.3. The cause of this situation may be related to the game theoretic concept, the *price of anarchy.* The price of anarchy is the difference in the utility if everyone acts for the social good, versus if everyone acts for their own selfish good

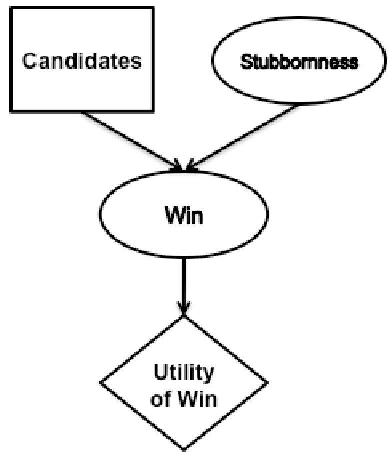


Figure C.1: A Bayesian decision network for an election simulation in which the rectangle represents the decision options (candidates), the oval labeled *Win* represents the belief that each of these candidates will win and the diamond represents the conditional utility of the winner. *Win* is also dependent on *Stubbornness*, which is the probability that an individual will be resistant to changing beliefs.

[49]. Mathematically, it is the ratio of the social optimum to the Nash equilibrium solution [55].

Given a simulated population, the election polling simulation computes each individual's expected utility for each candidate using a simple decision network, shown in Figure C.1. Each individual will provide a *conditional utility* for each candidate that represents the utility of a candidate *given* that the candidate wins. It is possible that an individual may prefer one candidate, but believe that the likelihood of him winning is low that the expected utility of the individual's second favorite candidate may actually be higher. An individual's belief that each candidate will win is also dependent on her "stubbornness." The more stubborn an individual is, the more likely she is to believe that her preferred candidate will win, despite receiving evidence to the contrary. Finally, each individual will declare her vote based on her highest expected utility.

C.1 Initialization

Notation:

P:	a population of voters
O:	a set of candidates
$P(W)$:	the likelihood of each candidate winning
$U(O\|W)$:	a conditional utility for each candidate given the candidate wins
$P(S)$:	the likelihood of a given individual being stubborn

The simulation is initialized with a population P of individuals and a set O of candidates. Each simulated individual will provide: $U(O|W)$, $P(W)$, and $P(S)$. $P(S) = 1.0$ means that the individual is completely resistant to changing her belief. The simulation then repeats a "polling" process until *convergence* occurs. Convergence occurs if no individual has switched votes for a specified number of repetitions. Convergence is intended to detect stability in the simulation and is not guaranteed to occur.

Simulation Parameters:

- *weights*: contains a weight for each candidate to indicate the initial lean of the population towards that candidate
- N, the population size
- C, the number of repetitions to test for convergence. The simulation must meet the requirement for convergence for C steps in order to be considered "converged."

Each of the N individuals is randomly assigned a preferred candidate based on the *weights* parameter. Each individual will set the conditional utilities for each candidate based on her preference. The candidates are on a range from "left" to "right," such that an individual that prefers the leftmost or rightmost candidate will give second preference to the middle candidate (not the opposing candidate). The stubbornness value for each individual is assigned randomly between 0.5 and 1.0. Each individual initializes her *Win* beliefs based on the candidate's utility and her stubbornness. Every individual will consider her preferred candidate more likely, but a more stubborn individual will overwhelmingly favor her preferred candidate. Each individual will then compute her expected utility for each candidate $o \in O$ using the following formula:

$$U(o) = U(o|w) * P(w) \tag{C.1}$$

The individual will then set her initial "vote" to the candidate with the highest expected utility. The initial *vote count* is the count of votes to each candidate for the whole population.

C.2 Social choice functions

This simulation compares a *single consensus* social choice function and the *collective* choice function. The single consensus choice function is an average of all individuals' expected utility for each candidate:

$$\forall o \in O, U_0(o) = \frac{1}{n}\sum_{i=1}^{n} U_i(o) \tag{C.2}$$

Where O is the set of candidates and n is the population size. The collective choice function will find the expected utility for each candidate, *for each collective* C_j:

$$\forall C_j, \forall c \in C, U_j(o) = \frac{1}{m_j}\sum_{i=1}^{m_j} U_i(o) \tag{C.3}$$

Where m_j is the number of individuals in collective C_j.

C.3 Results

At each time-step of the simulation, each individual will update her beliefs based on the *vote count*, which is provided to all individuals, and her stubbornness. She will then update her expected utility and select the candidate with the highest expected utility as her "vote." After all individuals have updated their belief, a new *vote count* will be determined. The simulation was run several times with different initialization parameters to observe the final results after convergence.

Simulation parameters:

weights: $\{0.05, 0.46, 0.49\}$
N: 10,000
C: 100

In this run, the leftmost candidate begins with a small portion of the votes, while the rightmost candidate has the highest proportion of the votes by a small margin. The simulation was repeated with these parameters several times and always had results similar to the following. Convergence, where $C = 100$ typically occurs after about 250 repetitions, meaning individuals stopped switching at around 150 repetitions. Table C.1 shows the final vote count and the single consensus social choice, determined by the average expected utility of each candidate.

	Left	Center	Right
Final vote count	314	4692	**4994**
Expected utility	-0.13	**1.52**	1.27

Table C.1: Final vote count and the single consensus social choice of each candidate.

Table C.2 shows each collective's average expected utility. Each collective is defined by its rank order, shown in the left column. The second column shows the collective's size and the remaining columns show the collective belief, or average expected utility for each candidate for that collective, followed by the variance in parentheses. Note that the average expected utilities for each candidate over the whole population P can be derived from these results using the formula, with $U_j(o)$ derived as in eq. C.3:

$$U(o) = \sum_{j=1}^{k} U_j(o) * \frac{|C_j|}{|P|} \tag{C.4}$$

Where k is the number of collectives.

After the simulation has converged, the results of this simulation imply the Nash equilibrium solution for this election simulation. Each individual– and therefore each collective– has decided their strategy with the awareness of the *vote count*, which represents the strategy of the rest of the population. During the simulation, individuals may change their strategy

RO	**Size**	**Left**	**Center**	**Right**
CLR	3601	-0.12 (0.030)	4.13 (0.36)	-2.86 (1.94)
RLC	2901	-0.23 (0.006)	-2.30 (1.40)	4.51 (0.12)
RCL	2093	-0.13 (0.015)	1.50 (0.94)	4.51 (0.09)
CRL	1091	-0.08 (0.030)	4.17 (0.24)	1.46 (1.02)
LCR	313	0.32 (0.120)	-2.22 (1.70)	-3.54 (1.35)
LRC	1	5.38 (0.000)	-1.11 (0.00)	-0.91 (0.00)

Table C.2: Average expected utility (collective belief) for each candidate for each collective followed by the variance in parentheses. The collectives are defined by their rank order (**RO**).

to maximize their expected utility given the vote count. However, once individuals have stopped switching votes, they have settled on a strategy that maximizes their utility given the strategies of the rest of the population. Each collective in Table C.2 reflects the strategy of its members through its expected utility.

C.4 Inconsistency in Average Expected Utility

An interesting discovery in this simulation is that the candidate with the highest average expected utility is *not* always the candidate with the highest vote count, even though each individual is voting according to her highest expected utility. Figure C.2 illustrates the range of start weights that results in inconsistent results. The y axis shows the start weights for *Center* and *Right*. *Left* was always 0.05. In the legend, **U** is the highest expected utility and **O** is the highest vote count. The x axis indicates the number of runs (5 total) of the simulation using the same input parameters. In the graph in fig. C.2, the middle, light colored, start weights always resulted in inconsistent behavior, while the top and bottom, darker colored, weights were accurately predicted. Interestingly, the most inconsistent results occur when the weight ratio is $Right = Left + Center$.

This result may be related to the *price of anarchy*. Mathematically, the price of anarchy is the ratio of the social optimum to the Nash equilibrium solution. If each individual votes according to their highest expected utility given the situation, then the results shown in

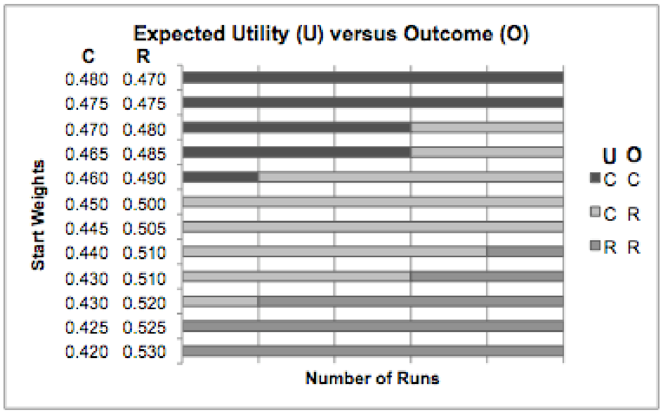


Figure C.2: Comparison of highest expected utility and vote count. The light gray bars indicate runs for which the option with the highest average expected utility was different than the vote outcome.

Table C.1 are the *social choice* of a Nash equilibrium solution. The open question in this case is what defines the social optimum and is it relevant in this situation? This dissertation demonstrates the risks in assuming a social optimum by way of average expected utility. My approach takes on a middle ground between social optimum and selfishness that may result in new implications for comparing selfish and altruistic behavior. I defer further discussion of this topic to future investigation.

While the average expected utilities over P incorrectly predicts the outcome, the average expected utility of each collective's preferred candidate results in collective beliefs that accurately reflect the outcome of the election (*Right* is highest *and Right* wins), shown in table C.3. In the context of the price of anarchy, this may be due to the fact that the collectives are a representation of the selfish goals of each individual, albeit possibly shared by a community.

Left	.0313*.32+.0001*5.38 = **0.011**
Center	.3601*4.13+.1091*4.17 = **1.94**
Right	.2901*4.51+.2093*4.51 = **2.25**

Table C.3: The average collective belief correctly reflects the outcome of an election if each individual votes according to their highest expected utility.

C.5 Predicting Voter Behavior

The rank order of the individuals that switch votes is a useful predictor of voter behavior. In the run described, 180 individuals switched votes during the simulation because their candidate with the highest expected utility changed. All of the individuals that switched moved from the collective represented by the rank order LCR, to the collective represented by the rank order CLR. The fact that all individuals who switched were from the same collective is evidence that the collective choice function is a better predictor of voting behavior than the single consensus social choice function. This social choice function does not distinguish these potential "vote switchers" from the rest of the population. As we might expect, the individuals that switched collectives and votes had expected utilities that indicated that they were nearly indifferent to their first and second preferred options. As in a real election, "swing votes" prove to be particularly important. The election simulation demonstrates how my approach allows the competition between opposing collectives to emerge, resulting in a more predictable situation than the single consensus social choice function allows.

References

[1] Kenneth J. Arrow. A difficulty in the concept of social welfare. *Journal of Political Economy*, 58, 1950.

[2] Kenneth J. Arrow. *Social Choice and Individual Values, Second edition*. John Wiley & Sons, 2nd edition, September 1963.

[3] Robert Bell, Yehunda Koren, and Chris Volinsky. The bellkor 2008 solution to the netflix prize. document generated for the 'netflix prize' competition, 2008.

[4] Jesus Mario Bilbao. *Cooperative games on combinatorial structures*. Klewer Academic, 2000.

[5] Duncan Black. On the rationale of group decision-making. *The Journal of Political Economy*, 56(1):23–34, 1948.

[6] Simon Blackburn. Rational animal? *Behavioral and Brain Sciences*, 4(3), 1981.

[7] Hans L. Bodlaender. A tourist guide through treewidth. *Acta Cybernetica*, 11:1–23, 1993.

[8] Daren C. Brabham. Crowdsourcing as a model for problem solving: An introduction and cases. *Convergence: The International Journal of Research into New Media Technologies*, 14(1):75–90, 2008.

[9] Jack S. Breese and Eric Horvitz. Ideal reformulation of belief networks. In *UAI '90: Proceedings of the Sixth Annual Conference on Uncertainty in Artificial Intelligence*, pages 129–144, New York, NY, USA, 1991. Elsevier Science Inc.

[10] A. Chinchuluun, P.M. Pardalos, A. Migdalas, and L. Pitsoulis, editors. *Pareto Optimality, Game Theory and Equilibria*. Springer, 2008.

[11] L. J. Cohen. Can human irrationality be experimentally demonstrated? *Behavioral and Brain Sciences*, 4:317–70, 1981.

[12] James S. Coleman. *The mathematics of collective action / James Coleman*. Heinemann, London :, 1973.

[13] Adnan Darwiche. A differential approach to inference in bayesian networks. In *Journal of the ACM*, pages 123–132, 2000.

[14] Rina Dechter. Bucket elimination: A unifying framework for probabilistic inference. In *Uncertainty in Artificial Intelligence UAI-96*, pages 211–219, 1996.

[15] Rina Dechter. Bucket elimination: A unifying framework for reasoning. *Artificial Intelligence*, 113(1-2):41–85, 1999.

[16] Rina Dechter and Irina Rish. A scheme for approximating probabilistic inference. In *Proceedings of Uncertainty in AI (UAI)*, pages 132–141, 1997.

[17] Arthur P. Dempster. Upper and lower probabilities induced by a multivalued mapping. In *Classic Works of the Dempster-Shafer Theory of Belief Functions*, pages 57–72. 2008.

[18] Francisco Javier Diez and F. J. D'iez. Parameter adjustment in bayes networks. the generalized noisy or-gate. In *In Proceedings of the 9th Conference on Uncertainty in Artificial Intelligence*, pages 99–105. Morgan Kaufmann, 1993.

[19] Allan M. Feldman and Roberto Serrano. Arrowś impossibility theorem: Two simple single-profile versions. Working Papers 2007-07, Instituto Madrileo de Estudios Avanzados (IMDEA) Ciencias Sociales, 2007.

[20] Maarten Franssen. Arrow's theorem, multi-criteria decision problems and multi-attribute preferences in engineering design. *Research in Engineering Design*, 16(1-2):42–56, 2005.

[21] Gerald F. Gaus. The rational, the reasonable and justification. *Journal of Political Philosophy*, 3(3):234–258, 1995.

[22] Christian Genest and James V. Zidek. Combining probability distributions: A critique and an annotated bibliography. *Statistical Science*, 1(1):114–135, 1986.

[23] Itzhak Gilboa, Dov Samet, and David Schmeidler. Utilitarian aggregation of beliefs and tastes. *Journal of Political Economy*, 112(4):932–938, August 2004.

[24] David Goldberg, David Nichols, Brian M. Oki, and Douglas Terry. Using collaborative filtering to weave an information tapestry. *Commun. ACM*, 35(12):61–70, 1992.

[25] Kshanti Greene, TJ Goan, and Eugene R. Creswick. A collaborative eye to the future. In *Proceedings of SPIE 2008*, 2008.

[26] Kshanti Greene, Joseph Kniss, George Luger, and Carl Stern. Satisficing the masses: Applying game theory to large-scale, democratic decision problems. In *Proceedings of the IEEE Conference on Social Computing (SocialCom-2009), Social Intelligence in Applied Gaming Workshop*, 2009.

[27] Kshanti Greene and George Luger. Aggreeing to disagree: Leveraging consensus and divergence in bayesian belief aggregation. In *Papers from the AAAI Spring Symposium, Technosocial Predictive Analytics Workshop*, pages 49–54, 2009.

[28] D. L. Hall and J. Llinas. An introduction to multisensor data fusion. *Proceedings of the IEEE*, 85(1):6–23, August 2002.

[29] Harry Halpin, Valentin Robu, and Hana Shepherd. The complex dynamics of collaborative tagging. In *WWW '07: Proceedings of the 16th international conference on World Wide Web*, pages 211–220, New York, NY, USA, 2007. ACM.

[30] R.A. Howard and J.E. Matheson, editors. *Readings on the Principles and Applications of Decision Analysis*, volume 2, pages 6–16. Strategic Decisions Group, 1984.

[31] Aanund Hylland and Richard J Zeckhauser. The impossibility of bayesian group decision making with separate aggregation of beliefs and values. *Econometrica*, 47(6):1321–36, November 1979.

[32] Finn Jensen. *An Introduction to Bayesian Networks*. Springer-Verlag, 1996.

[33] Rudolph Emil Kalman. A new approach to linear filtering and prediction problems. *Transactions of the ASME–Journal of Basic Engineering*, 82(Series D):35–45, 1960.

[34] Don Koks and Subhash Challa. An introduction to bayesian and dempster-shafer data fusion. Technical report, 2003.

[35] Daphne Koller and Brian Milch. Multi-agent influence diagrams for representing and solving games. In *IJCAI*, pages 1027–1036, 2001.

[36] S. Kullback and R. A. Leibler. On information and sufficiency. *The Annals of Mathematical Statistics*, 22(1):79–86, 1951.

[37] Kevin Leyton-Brown and Yoav Shoham. *Essentials of Game Theory: A Concise, Multidisciplinary Introduction*. Morgan and Claypool Publishers, 2008.

[38] Marc Maes and Michael Faber. Issues in utility modeling and rational decision making. In *Reliability and Optimization of Structural Systems*, pages 95–104. Taylor & Francis Group, 2004.

[39] Marc A. Maes and Michael H. Faber. Bayesian framework for managing preferences in decision-making. *Reliability Engineering & System Safety*, 91(5):556 – 569, 2006.

[40] Izhar Matzkevich and Bruce Abramson. The topological fusion of bayes nets. In *Proceedings of the Eighth Conference on Uncertainty in Artificial Intelligence*, 1992.

[41] Pedrito II Maynard-Reid and Urszula Chajewska. Aggregating learned probabilistic beliefs. In *Proceedings of the 17th Conference in Uncertainty in Artificial Intelligence*, 2001.

[42] Brian Milch. Probabilistic models for agents beliefs and decisions. In *In Proc. 16th UAI*, pages 389–396. Morgan Kaufmann, 2000.

[43] Kevin Murphy. Bayes net toolbox for matlab, 1997.

[44] Kevin Murphy. A brief introduction to graphical models and bayesian networks, 1998.

[45] Clotilde Napp and Elyes Jouini. Aggregation of heterogeneous beliefs. *Journal of Mathematical Economics*, 42(6):752–770, 2006.

[46] Klaus Nehring. The (im)possibility of a paretian rational. Economics Working Papers 0068, Institute for Advanced Study, School of Social Science, 2005.

[47] Luis Ortiz. Algorithms for graphical games. In *DIMACS Workshop on Bounded Rationality*, 2005.

[48] Marco Ottaviani and Peter N. Sorensen. Aggregation of information and beliefs in prediction markets. Technical report, University of Copenhagen, 2006.

[49] Christos H. Papadimitriou. Algorithms, games, and the internet. In *ICALP '01: Proceedings of the 28th International Colloquium on Automata, Languages and Programming,*, pages 1–3, London, UK, 2001. Springer-Verlag.

[50] Judea Pearl. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan Kaufman, 1988.

[51] David M. Pennock and Michael P. Wellman. Representing aggregate belief through the competitive equilibrium of a securities market. In *Proceedings of the Thirteenth Conference on Uncertainty in Artificial Intelligence*, 1997.

[52] David M. Pennock and Michael P. Wellman. Graphical representations of consensus belief. In *Proceedings of the Fifteenth Conference on Uncertainty in Artificial Intelligence*, 1999.

[53] David M. Pennock and Michael P. Wellman. Graphical models for groups: belief aggregation and risk sharing. *Decision Analysis*, 2(3):148–164, 2005.

[54] Howard Raiffa. *Decision Analysis*. Addison-Wesley, 1968.

[55] Tim Roughgarden. The price of anarchy is independent of the network topology. In *Journal of Computer and System Sciences*, pages 428–437, 2002.

[56] Teddy Seidenfeld, Joseph Kadane, and Mark Schervish. On the shared preferences of two bayesian decision makers. *The Journal of Philosophy*, 86(5):225–244, 1989.

[57] Amartya K Sen. Personal utilities and public judgements: Or what's wrong with welfare economics? *Economic Journal*, 89(355):537–58, September 1979.

[58] Kari Sentz and Scott Ferson. Combination of evidence in dempster-shafer theory. Technical report, Sandia National Labs, 2002.

[59] R. D. Shachter. Evaluating influence diagrams. *Operations Research*, 34(6):871–882, 1986.

[60] G. Shafer. *A mathematical theory of evidence*. Princeton university press, 1976.

[61] C. E. Shannon. Prediction and entropy of printed english. *Bell Systems Technical Journal*, 30:50–64, 1951.

[62] A Steinberg and C Bowman. Rethinking the jdl data fusion levels. In *The Dual Node Network (DNN) Data Fusion & Resource Management (DF&RM) Architecture*, 2004.

[63] Thayer Watkins. Arrow's impossibility theorem for aggregating individual preferences into social preferences, 2009.

[64] Ian Witten and Eibe Frank. *Data Mining: Practical Machine Learning Tools and Techniques*. Morgan Kaufmann, 2005.

[65] Xiaofeng Wu, Peter J. F. Lucas, Susan Kerr, and Roelf Dijkhuizen. Learning bayesian-network topologies in realistic medical domains. In *ISMDA '01: Proceedings of the Second International Symposium on Medical Data Analysis*, pages 302–308, London, UK, 2001. Springer-Verlag.

[66] Eliezer Yudkowsky. Cognitive biases potentially affecting judgment of global risks. In Nick Bostrom and Milan Cirkovic, editors, *Global Catastrophic Risks*. Oxford University Press, 2007.

[67] Lotfi A. Zadeh. Is there a need for fuzzy logic? *Inf. Sci.*, 178(13):2751–2779, 2008.

[68] Nevin Lianwen Zhang and David Poole. Exploiting causal independence in bayesian network inference. *Journal of Artificial Intelligence Research*, 5:301–328, 1996.