Representing diversity in communities of Bayesian decision-makers

Kshanti A. Greene, Joe M. Kniss and George F. Luger
Department of Computer Science
University of New Mexico
kshanti, jmk, luger@cs.unm.edu

***Abstract*—High-quality information has emerged from the contributions of many using the wiki paradigm. A logical next step is to use the wisdom of the crowd philosophy to solve complex problems and produce informed policy. We introduce a new approach to aggregating the beliefs and preferences of many individuals to form models that can be used in *social* policy and decision-making. Traditional *social choice functions* used to aggregate beliefs and preferences attempt to find a *single* solution for the whole population, but may produce an irrational social choice when a stalemate between opposing objectives occurs. Our approach, called *collective belief aggregation*, partitions a population into *collectives* that share a preference order over the expected utilities of decision options or the posterior likelihoods of a probabilistic variable. It can be shown that if a group of individuals share a preference order over the options, their aggregate will uphold principles of rational aggregation defined by social choice theorists. Super-agents can then be formed for each collective that accurately represent the preferences of their collective. These super-agents can be used to represent the collectives in decision analysis and decision-making tasks. We demonstrate the potential of using collective belief aggregation to incorporate the objectives of stakeholders in policy-making using preferences elicited from people about healthcare policy.**

I. Introduction

We introduce a new approach to aggregating the beliefs and preferences of many individuals to form models that maintain the population's diversity. These models can be used to form policy and make decisions at societal scales. Computational models for social decision-making that attempt to incorporate the "wisdom of the crowd" philosophy have the potential to change the decision-making paradigm in a society. There is promise in incorporating the experiences and objectives of individuals to form models for collective policy-making that will enable them to have a direct influence in the social, economic, and political decisions that affect them. We have seen that high-quality knowledge can emerge from large numbers of contributors using the wiki paradigm. Our goal is to enable informed and appropriate policies to emerge from the wisdom of the crowd.

Several theoretical limitations must be addressed before an organization or society can take actions that are based on aggregated beliefs and preferences of its members. In particular, the spirit and objectives of the a community's contributions should be preserved in a community solution. However, techniques such as voting and averaging will not always produce rational results. In particular, social choice and Bayesian theorists state that it is not possible to combine, or *aggregate* arbitrary beliefs or preferences to form a single rational solution [1], [2], [3]. For example, the Nobel-laureate economist Kenneth Arrow showed that attempting to find a "social preference" order given a ranked order of options from multiple individuals can result in a situation that does not conform to principles of rational aggregation [1]. These findings were extended for Bayesian decision-making when aggregating probabilistic beliefs and utilities [3], [2].

Our goal is to form an accurate representation of a population's diverse beliefs and preferences from which people can create new paradigms for rational decision-making. In [4] we introduced our approach for clustering individuals based on their beliefs and preferences prior to aggregation. In this paper, we partition a population into *collectives* based on each individual's expected utility of a set of decision options. In particular, we show that if a group of individuals share a common preference order over the decision alternatives, that their aggregate will uphold the rationality properties defined by social choice theorists. Super-agents that accurately represent each collective can then be used to initialize decision analysis, policy-making and decision-making techniques.

In this article we will demonstrate some techniques for decision analysis and decision-making using super-agents. For example, we apply our approach to form super-agents that negotiate healthcare policy options "on behalf" of their collective members. The decision-making techniques presented in this paper are not meant to be exhaustive. Our intention is to present an approach that forms an accurate representation of a diverse population and to spark new discourse in policy formation and decision-making that leverages competing objectives instead of attempting to average it away.

The article is organized as follows. Section II summarizes the foundational theories of this research, including Bayesian, social choice and game theories. Section III discusses rational social choice in more depth and shows two irrational social choice results that occur in Bayesian aggregation. Section IV introduces our collective belief aggregation approach that forms a representation of a diverse population. Section V demonstrates that super-agents can be used to represent collectives' preferences in decision analysis, game theoretic analysis and negotiation techniques.

II. Background

A. Bayesian Networks

This research is based on a framework that is used to model reasoning in the presence of uncertainty. A Bayesian network is a form of graphical model that integrates the concepts of graph theory and probabilistic reasoning [5], [6], [7]. 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 are the parents of X [5].

Bayesian networks can be extended to address decision problems using *influence diagrams* [8], 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.

In the work described in this article, Bayesian decision networks are used to represent decisions and the factors that influence decisions. For example, instead of asking deterministic questions for complex issues such as "do you support healthcare reform?" we ask individuals to consider which factors would increase or decrease their support for healthcare reform.

B. Social Decision-making

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 beliefs and preferences of a group of individuals [1], [9], [3], [2]. 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 [3], [2]. 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 that are relevant to the research described in this article are discussed in more detail in Section III.

Game theory is the mathematical study of interacting agents, each with the self-interested goal of improving their own situation [10]. The value of a situation is determined by each player's utility or expected utility. Game theory describes many types of solutions and strategies that are considered rational behavior in competitive and strategic environments. For instance, 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 [10].

III. Rational Social Choice

This article addresses a precise definition of rationality, introduced by social choice theorists for the purpose of combining beliefs and preferences [1], [9], [3], [2]. Our 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. We do not attempt to address the issue of whether humans behave rationally. In particular, no judgment is made about the correctness of the individuals' beliefs. Nor are any assumptions made that the individuals or group would act consistently according to their beliefs and preferences.

This section discusses the principles of rationality that have been defined by social choice theorists and illustrates the failure of Bayesian aggregation methods with a simple decision network. The remainder of the article presents an approach that enables belief and preference aggregation for decision-making that upholds these properties for rational aggregation.

A. Preference Relations

The notation in Table I is used to describe the *pairwise* preference relationships that indicate 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. For example, 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 $vPcRs$.

TABLE I
Pairwise preference relations

x, y:	alternative options
i, j:	individuals
P:	a preference relation representing *strict preference*
I:	a preference relation representing *indifference*
xPy:	a group prefers x to y
xIy:	a group is indifferent to x and y
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)

B. Arrow's Axioms

The properties in Figure 1 are a subset of the properties for rational aggregation introduced by economist Kenneth Arrow [1], and later adapted by Arrow and other researchers [9], [2]. These social choice theorists determined that these properties must hold for belief and preference aggregation to be considered *rational*. The full set of properties is discussed in [11]. The properties in Fig. 1 can be shown to be broken in a Bayesian environment and are relevant to the research described in this article.

Given a population P a social choice solution should have the following properties:

1) *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:
2) *Non-dictatorship and non-imposition (NDIP)*. There is no dictator. Individual i is a dictator if, $\forall x$ and $\forall y$, $xP_iy \to xPy$. Non-imposition means that no order has been pre-determined for any individual.

Fig. 1. Two of the properties for rational preference aggregation defined by Kenneth Arrow [1], [9]

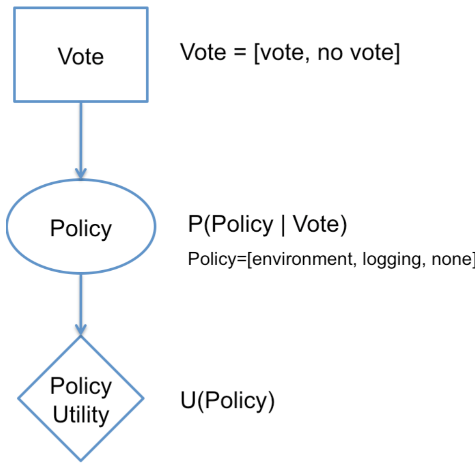

Fig. 2. A decision network representing a decision to put a new policy to vote.

C. 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. Our 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 article address various political and policy-making situations.

We 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. 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 2 represents a group decision about whether a vote should be held to introduce a 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 2, the rectangle represents the decision to put the policy to vote or not. The oval represents the conditional probability of $Policy$ given the $VoteAction$ decision, represented by $P(Policy|VoteAction)$. Policy will be one of $[E, L, N]$, where E= environmental, L= logging, N = none. The utility of each policy option, $U(Policy)$ is represented by the diamond. The range of utility is $[-2, 2]$, with a negative utility indicating an objection to the policy option. Inference on the network will determine the expected utility of the decision options— to vote or not to vote. Equation 1 computes the *expected utility* of each decision option.

$$EU(VoteAction) = P(Policy|VoteAction)U(Policy) \tag{1}$$

D. A Non-Pareto Optimal Solution

We adapt a demonstration by Hylland and Zeckhauser [3] to the decision network described in the previous section. 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. In other words, there is some option *other than* the social choice that is preferred by all individuals. Suppose there are two individuals i_1 and i_2. The individuals' conditional probability tables (CPTs) for $P_i(Policy|VoteAction)$ are shown in Figure 3, indicating their belief in the likelihood of each policy given that the policy is put to a vote. If there is a vote the policy will be either E or L. If no vote is held, then there will be no policy. The individuals' utilities for each policy decision are shown in Table II.

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

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

Fig. 3. $P_i(Policy|VoteAction)$ for two individuals i_1 and i_2.

TABLE II
UTILITIES $U_i(Policy)$ OF THE POLICY OPTIONS FOR TWO INDIVIDUALS.

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

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

The results of applying an arithmetic mean to find the consensus CPT, $P_0(Policy|VoteAction)$, are shown in Table III. The consensus on policy utilities, $U_0(Policy)$, is shown in Table IV. The geometric mean will result in similar values. Table V compares the results of applying equation 1 to each individual's beliefs and utilities with the results of applying the equation to the consensus beliefs and utilities. The best option for each individual and the social choice ($EU_o(VoteAction)$) is shown in bold. We can see that the consensus favors the opposite decision option that *both* individuals favor. In other words, the consensus option is not Pareto optimal.

TABLE III
CONSENSUS CONDITIONAL PROBABILITY TABLE FOR $P_0(Policy|VoteAction)$ COMPUTED USING THE ARITHMETIC MEAN.

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

TABLE IV
CONSENSUS UTILITIES $U_0(Policy)$ FOR THE POLICY OPTIONS COMPUTED USING THE ARITHMETIC MEAN.

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

TABLE V
EXPECTED UTILITIES OF EACH INDIVIDUAL AND THEIR CONSENSUS. THE OPTIONS WITH THE HIGHEST EXPECTED UTILITY ARE SHOWN IN BOLD.

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

E. Dictatorship

The next example demonstrates a situation that breaks the NDIP (non-dictatorship) property in Figure 1 using the decision network in Figure 2. Suppose that the table on the top of Figure 4 contains the conditional probabilities for a group g of three individuals who all happen to have the same beliefs. The table on the bottom of Figure 4 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 VI. The utilities $U(Policy)$ are identical for all individuals and are shown in Table VII.

Table VIII shows the expected utility for the group g, the dictator d and their combined consensus computed using equation 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. This example demonstrates 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]. A more general phenomenon is occurring that causes the mean of a set of quantitative values to be skewed by a small number of highly divergent values. In addition to the dictatorship situation, this causes the consensus to "lose" the representation of the population's beliefs and preferences. In other words, as a set of values becomes more divergent, the set's mean will become less similar to the original beliefs.

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

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

Fig. 4. The table on the top contains conditional probabilities for a group g of three individuals with identical beliefs. The table on the bottom shows the conditional probabilities for a single individual d.

TABLE VI
THE CONSENSUS CONDITIONAL PROBABILITIES OF THE GROUP COMPOSED OF $g \cup d$.

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

TABLE VII
CONSENSUS UTILITIES $U_0(Policy)$ FOR THE POLICY OPTIONS COMPUTED USING THE ARITHMETIC MEAN.

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

TABLE VIII
EXPECTED UTILITIES OF A GROUP $EU_g(VoteAction)$ OF THREE INDIVIDUALS, AN INDIVIDUAL $EU_d(VoteAction)$ AND THEIR CONSENSUS $EU_g(VoteAction)$.

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

F. Rational Social Choice Definition

We define a *rational social choice* (RSC) as a social choice solution for a group of individuals that upholds the properties for rational aggregation defined by Arrow [1], [9] and extended for a Bayesian environment by Hylland and Zeckhauser [3]. Two of these properties were shown in Figure 1. The remainder of the properties are discussed in [12].

IV. Collective Belief Aggregation

Two significant limitations of existing belief and preference aggregation approaches are that (1) they can form consensus models that under-represent divergent objectives (2) they may result in an irrational *social choice* solution in the presence of divergence, as shown in Section III. In [4] we introduced an approach that clusters individuals based on the similarity of their beliefs and preferences and then finds an aggregate for each cluster. The clustering approach improved the *representation* of the population's beliefs and preferences, measured by the Kullback-Liebler divergence measure.

We now present an aggregation approach that addresses the second limitation of belief and preference aggregation approaches. The *collective belief aggregation* (CBA) approach partitions a population into subgroups that agree on the relative desirability of a set of decision options. The relative desirability is determined by the *Bayesian rank order* of the expected utilities for the decision options computed using a Bayesian decision network. Inference on a decision network computes the expected utility of a set of decision options. The Bayesian rank order is the preference order of the decision options based on the decreasing order of the options' expected utilities. For example, if an individual's 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 the expected utilities is $[0.0, -1.2]$ and the individual's Bayesian rank order (or just rank order) is $[NoVote, Vote]$. Using the preference relations defined in Section III-A, 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.

Our partitioning approach forms *collectives* from individuals who have the same Bayesian rank order of the decision options. We formally define a collective as a subset of a population such that a specific property holds for all members of the subset [12]. In this case the property is the Bayesian rank order. 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 the vote action decision would result in only two collectives. In general, 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 in [12]. 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 a rational social choice can be derived from the collective belief (also shown in [12]).

The CBA approach can be applied to any discrete probability distribution. In a traditional (non decision) Bayesian network, the rank order represents the relative likelihood of the possible outcomes, ordered from most to least likely.

Aggregation approaches that attempt to find a single consensus solution may result in an irrational social choice when there is a stalemate or significant divergence in belief. In contrast, collective belief aggregation finds a *set* of rational social choice solutions that may need to be resolved through other means, such as through the game theoretic analysis in Section V-B and negotiation techniques in Section V-C).

A. Posterior versus Prior Collective Belief Aggregation

Collective belief aggregation can be performed such that the collectives are discovered based on the posterior results of inference or the collectives are discovered prior to inference. The accuracy of a CBA algorithm is measured by the percent of individuals that are placed in the correct collective according to their Bayesian rank order over the decision options. In order to guarantee rational social choice for each collective, a CBA algorithm must guarantee 100% accuracy. Otherwise, it is possible that an individual will be placed into a collective whose consensus rank order is different than the individual's rank order.

Given a population of individuals, each having a network containing a decision D with d options, the exact, brute force CBA algorithm to discover and aggregate the collectives for D works as follows:

1) Run an inference algorithm on each network to compute the expected utility for each of the d decision options.
2) Place each individual into a collective based on the individual's expected utility for the decision options.
3) Compute the average expected utility, or collective belief, for each decision option and each collective.
4) Order the d options in decreasing order of their expected utilities to determine each collective's Bayesian rank order.

This approach is called the *posterior* CBA algorithm (after posterior compromise [13]) because it forms the collectives and does aggregation *after* inference. While this algorithm will result in 100% accuracy, the number of networks that inference must be run on is the population size.

Prior CBA uses a partitioning algorithm, such as clustering [14], to make an initial guess at how the collectives will form based on the population's prior beliefs (in the form of prior and conditional probabilities) and utilities. A consensus network is then formed from the mean of each cluster's prior beliefs, as in the examples in Sections III-D and III-E. This approach would allow one to reduce the number of networks that inference needs to be performed on. However 100% accuracy and an RSC cannot be guaranteed for the collectives.

Using the network in Fig. 2, a simulation was created in which the beliefs and utilities of the individuals in a population were initialized using a Gaussian mixture model for $P(Policy|VoteAction)$ and a uniform distribution for $U(Policy)$ in the range $[-2, 2]$. The mixture model had $\mu = [0.0, 0.25, 0.5, 0.75, 1.0]$ and $\sigma = [0.1, 0.2, 0.2, 0.2, 0.1]$. The prior and posterior CBA algorithms were applied to a population of 10,000. The accuracy of the prior CBA algorithm is shown in Figure 5. The different groups of points represent different mixture weights for $P_i(Policy|VoteAction)$ over a number of runs of the simulation. The diamond group's mixture weights resulted in a stronger consensus centered around an expected utility of 0.0. The triangle group's beliefs were more polarized, resulting in two peaks of expected utility near -1.0 and 1.0.

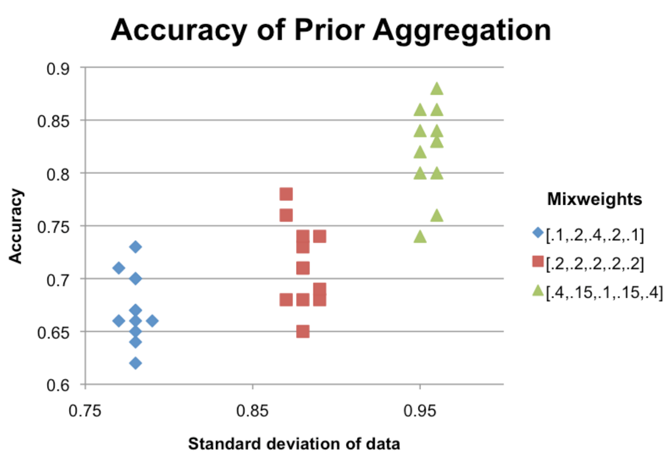

Fig. 5. The accuracy of prior aggregation using three mixture models.

We can see from Fig. 5 that the accuracy of the prior aggregation algorithm increased as the standard deviation increased. Since the network contained only two variables, the high deviation of one variable resulted in clusters that fairly accurately predicted the rank order of the expected utilities. However, as the network size increases, the accuracy of the prior aggregation algorithm may decrease because more variables (dimensions) are involved in clustering. The prior and posterior CBA algorithms and a hybrid algorithm called *incremental aggregation* are discussed in depth in [11].

B. Revisiting Non-Pareto Optimal Solutions

We now revisit the example in Section III-D using collective belief aggregation. Posterior CBA will result in both individuals being placed in the same collective, with an average expected utility of 0.45 for the vote option, which is also the Pareto optimal solution. Using prior aggregation, if a partitioning algorithm first separates the two individuals into their own collectives, each collective's expected utility will be 0.45 for the vote option, again resulting in a Pareto solution. Again, only posterior collective aggregation can guarantee a rational social choice, as discussed in Section IV-A.

C. Revisiting Dictatorship

We first note that the dictatorship situation in the example in Section III-E would also occur if posterior aggregation were used without forming collectives. In fact, the irrational social choice will result any time the individuals' expected utility of the $Vote$ option ($EU_i(Vote)$) causes the following inequality to hold, where x is the number of individuals in the g group.

$$\sum_{i=1}^{x} EU_i(Vote) > -EU_d(Vote)$$

The posterior 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. Since each collective maintains the rank order of its members, the relative preferences between options are always maintained.

V. Social Decision-making with Super-agents

The previous section introduced the collective belief aggregation approach and discussed how collectives can emerge from a population whose aggregate will uphold rational social choice properties. Since the relative preference order of all members of a collective is the same, some generalizations can be made about each collective. These generalizations are the collective belief— which is the mean of the expected utilities for the decision options, and the Bayesian rank order— which is the deterministic preference order over the decision options. A "super-agent" is formed for each collective that takes on the collective belief and Bayesian rank order. These super-agents can then accurately represent the relative preferences of their collective members in decision-making. This section will demonstrate some techniques for analyzing and making decisions using the super-agents including game theoretic analysis and negotiation.

A. 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. Discovery of the Pareto optimal 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.

An algorithm was developed to extract the Pareto solutions from the super agent's Bayesian rank orders. The algorithm, described in [11], finds the set of options that uphold the strong Pareto condition by first finding those that do not. The Pareto optimal solutions are the minimally acceptable social choice solutions for a population. Other decision-making techniques and analyses can be used to reduce this set further.

B. Game Theoretic Analysis

A concept developed by Koller and Milch [15] forms the basis for game theoretic analysis using super-agents. The authors introduced *multi-agent influence diagrams* (MAIDs) that represent strategic situations between multiple agents. The MAIDs in [15] assume that an agent represents a single entity. We allow an agent to be a super-agent that represents a collective. Figure 6 contains a MAID that represents a strategic situation between a community and a logging interest. In this case, multiple super-agents that have different Bayesian rank orders over the $VoteAction$ decision may emerge, representing different factions in the community. Another agent represents the logging interest and its decision to log and lobby depending on which policy is more likely.

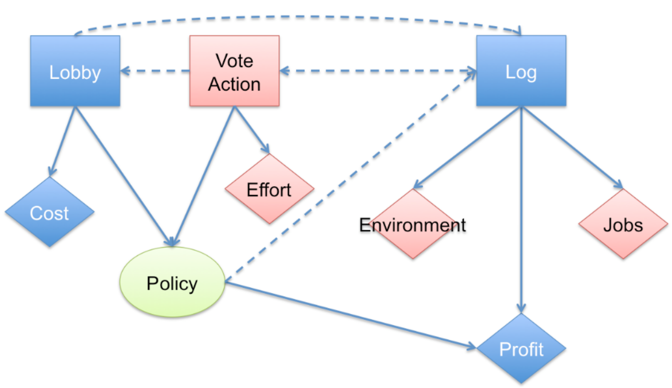

Fig. 6. A MAID that represents the community's beliefs and actions ($VoteAction$, $Effort$, $Environment$ and $Jobs$ nodes) and a logging interest's beliefs and actions ($Lobby$, Log, $Cost$ and $Profit$ nodes) The dashed edges represent the values that an agent considers when he makes a decision.

As in the MAIDs in [15], the dashed lines in Fig. 6 represent decisions or variable nodes that affect a decision. These are the parents of a decision D, or $Pa(D)$. The Nash equilibrium solutions for the super-agents and the logging interest agent are the decision options that maximize the expected utility of each agent given the parents, $Pa(D)$, of each decision [15]. Simulations that find the Nash equilibrium solutions as well as the maximin and minimax strategies for this MAID are discussed in [11]. Using MAIDs representing the beliefs and actions of super-agents, we are able to efficiently apply game theoretic analysis in large populations.

C. Negotiating Super-agents

The next demonstration uses preferences elicited from people using Amazon.com's Mechanical Turk (mturk.com), which is an online source for low-cost human labor. A few days before the 2010 healthcare bill was voted on in the US Congress, we asked Mechanical Turk workers to state whether they supported the bill as is, did not support the bill, or would support it with changes. They were then given a number of options (suggested by other Mechanical turk workers) and asked if adding each option would decrease, increase or have no effect on their support for the healthcare bill. 200 people provided their opinion.

The goal of the following experiment is to find the set of options that will maximize the number of people that are likely to support the bill. Each super-agent represents the interest of its collective. Negotiation between super-agents occurs in the following manner. Suppose there are two super-agents, i and j, and two options, a and b. Super-agent i's collective supports a but super-agent j's collective does not. However super-agent j suggests adding b, which would gain the support of its collective. If adding b will not lose the support of i's collective, then the two super-agents have "negotiated" to include both options.

In the experiment, a super-agent represents each collective containing individuals that have the same opinion about a set of changes. We use a Bayesian network to represent the likelihood that an individual will support the healthcare bill given a set B of b possible options to add to the bill. Each option in B will be a parent of $Support$. The conditional probability table representing $P(Support|B)$ will then indicate the likelihood that $Support$ is true or false given that each option in B could be true or false.

The posterior probability $P'(Support)$ when all options are false is the original support likelihood each individual provided. In a real-world situation we would ask people to quantify how much an option increases or decreases their support. For this demonstration, a small value (0.15) was then added or subtracted from the original likelihood for each option in B that is true, if the change increased or decreased their likelihood of support, respectively.

Collectives were then formed from individuals that had the same rank ordering over the support values for all combinations of options in B. In other words, the rank order of $P(Support|b_i)$ was the same for each $b \in B$. This means that each super-agent can accurately represent all members of its collective for negotiation on all options. The rank ordering possibilities were TF– indicating a likelihood of support, FT– indicating a likelihood of no support, and $\{FT\}$– indicating that the collective was undecided, in other words its likelihood for $P(Support = true|B)$ was near 0.5. Given the set B, the subset of options set to $true$ that maximizes the likelihood of support for the bill is the optimal set of options.

Table IX shows four different options for which the individuals provided their opinions. Given these options, 31 out of $3^4 = 81$ possible collectives emerged from the population, ranging in size from one to 79. The largest collective represented the individuals who supported the bill given any combination of the four options. Another collective containing 59 individuals would not support the bill no matter which options were added. A number of collectives, with 1-7 members each, were initially undecided or did not support the bill, but would support the bill given different subsets of the four options. Without any options added, 39.5% of the surveyed population supported the bill, 32% did not and 28.5% were undecided.

Table X shows each combination of options that were set to true, and the total number of individuals who would likely support the bill given those options. The combination that maximized likelihood of support– to 62.5%– was adding mental health coverage and legislation to regulate the cost of healthcare. This combination corresponds to the conditional probability $P(Support|M = true, C = true, T = false, A = false)$.

TABLE IX
FOUR OPTIONS TO THE HEALTHCARE BILL.

Symbol	Option Description
M	Add mental health coverage
C	Add legislation to regulate the cost of healthcare
A	Limit abortion coverage
T	Tax premium insurance plans

TABLE X
TOTAL SUPPORT FOR THE BILL (OUT OF 200) GIVEN THE SELECTED OPTIONS. THE BEST SET OF OPTIONS IS M,C (IN BOLD).

Options	Support
All false	79
A	86
A,T	88
C,A	94
M,A	95
T	95
M,A,T	101
C,A,T	103
M,T	109
C,T	109
M,C,A	116
M	117
C	118
M,C,A,T	118
M,C,T	122
M,C	**125**

We can show that negotiation produces an optimal combination of options when a simple majority approach does not. Table XI shows three options followed by the total number of individuals whose support is gained or lost by introducing the option when considered independently. A negative number (for example for option S) means that a majority of the population does not support the option. Attempting to increase support by adding options that have a simple majority would result in options C and E. However, negotiation agents that consider all combinations of options results in a total support of 123 for $[C, E]$ and a total gain of 125 for $[C, E, S]$. In this case the S option gained support from a collective without losing that collective's support on the other two options. The surveyed population's diversity meant that a simple majority vote for each option would not result in an optimal combination of policies.

TABLE XI
THREE OPTIONS TO THE HEALTHCARE BILL AND THE NUMBER OF INDIVIDUALS WHOSE SUPPORT IS GAINED (POSITIVE) OR LOST (NEGATIVE) IF THE OPTION IS INTRODUCED.

Symbol	Option Description	Gain/Loss
C	Add legislation to regulate the cost of healthcare	125
E	Expand medicare benefits	77
S	Introduce a single-payer system	-2

VI. SUMMARY

Instead of "averaging away" any conflict, our aggregation approach allows competing objectives to emerge by partitioning a population into collectives based on their expected preference order over the decision options. Using posterior aggregation, the aggregate of each collective will result in a rational social choice for that collective. Super-agents can then accurately represent each collective in policy and decision-making endeavors. Future research directions include investigating additional decision-making techniques using the super-agents and elicitation techniques that form richer decision models.

REFERENCES

[1] K. J. Arrow, "A difficulty in the concept of social welfare," *Journal of Political Economy*, vol. 58, 1950. [Online]. Available: http://econpapers.repec.org/RePEc:ucp:jpolec:v:58:y:1950:p:328

[2] T. Seidenfeld, J. Kadane, and M. Schervish, "On the shared preferences of two bayesian decision makers," *The Journal of Philosophy*, vol. 86, no. 5, pp. 225–244, 1989.

[3] A. Hylland and R. J. Zeckhauser, "The impossibility of bayesian group decision making with separate aggregation of beliefs and values," *Econometrica*, vol. 47, no. 6, pp. 1321–36, November 1979.

[4] K. Greene, J. Kniss, G. Luger, and C. 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.

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

[6] F. Jensen, *An Introduction to Bayesian Networks*. Springer-Verlag, 1996.

[7] K. Murphy, "A brief introduction to graphical models and bayesian networks," 1998. [Online]. Available: http://www.cs.ubc.ca/ murphyk/Bayes/bnintro.html

[8] R. D. Shachter, "Evaluating influence diagrams," *Operations Research*, vol. 34, no. 6, pp. 871–882, 1986.

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

[10] K. Leyton-Brown and Y. Shoham, *Essentials of Game Theory: A Concise, Multidisciplinary Introduction*. Morgan and Claypool Publishers, 2008.

[11] K. Greene, "Collective belief models for representing consensus and divergence in communities of bayesian decision-makers," Ph.D. dissertation, University of New Mexico Department of Computer Science, 2010.

[12] K. A. Greene, J. M. Kniss, and G. F. Luger, "Enabling rational democratic decision-making with collective belief models and game theoretic analysis," UNM Department of Computer Science, Tech. Rep., 2010.

[13] I. Matzkevich and B. Abramson, "The topological fusion of bayes nets," in *Proceedings of the Eighth Conference on Uncertainty in Artificial Intelligence*, 1992.

[14] I. Witten and E. Frank, *Data Mining: Practical Machine Learning Tools and Techniques*. Morgan Kaufmann, 2005.

[15] D. Koller and B. Milch, "Multi-agent influence diagrams for representing and solving games," in *IJCAI*, 2001, pp. 1027–1036.