Market-Based Belief Aggregation and Group Decision Making

Dissertation Proposal

David M. Pennock
University of Michigan
Artificial Intelligence Laboratory
1101 Beal Avenue
Ann Arbor, MI 48109-2110 USA
dpennock@umich.edu

Committee Members:
Edmund H. Durfee
Daniel Koditschek
Stephen M. Pollock

Advisor: Michael P. Wellman

1 Overview

Research toward my dissertation will investigate the problem of belief aggregation, and its application to group decision making. *Belief aggregation* is a process of combining or summarizing the probabilistic beliefs of a collection of individual agents in order to form a single consensus group belief. The problem has been studied from the perspective of a variety of disciplines, though I focus on formal, mathematical aggregation models since these are most tenable for computational agents. My efforts stem mostly from work in the statistics, economics, and artificial intelligence (AI) communities. Researchers have proposed many belief aggregation methods, although on the question of which is best the general consensus is that there is no consensus. My approach takes inspiration from the power of economic markets to effectively summarize the opinions of a large collection of agents into a single number, namely price. In this model, belief aggregation is facilitated through a securities market in the relevant uncertain events: each security pays off \$1 if its associated uncertain event occurs, nothing otherwise. Participating agents buy and sell these "contingent claims" according their own individually rational beliefs and preferences, causing prices to rise and fall in step with supply and demand. The resulting price of a security represents the group's aggregate assessment of the probability of its corresponding uncertain event. The introduction of a securities market has several advantages over the purely mathematical pooling functions common in the statistics literature. It provides monetary incentives for agents to participate, gather relevant evidence, and bet according to their true beliefs. It also provides a precise and well-studied protocol for distributed agents to interact without directly revealing their internal beliefs. My dissertation work will extend previous results [16, 17] through both theoretical and empirical studies, preferring analysis over experiment, but realizing that in many realistic situations explicit mathematical proofs will not be possible.

It has been argued that pure belief aggregation has little practical value outside the more general context of *group decision making* [24], in which *both* beliefs and utilities (preferences) of participating agents are combined to allow inference of "optimal" group decisions. Regardless of the validity of this argument, the prospect of making good group decisions is of interest in many situations. While all of my work to date has focused on handling asymmetric uncertainty for belief aggregation, future work will address treatment of heterogeneous preferences and generalizations of our market model to the group decision problem. I also plan a "proof of concept" demonstration of the practical applicability of the model.

The next section provides a conceptual foundation for the rest of the proposal, describing my "worldview" or modeling conventions, and motivating research on market-based aggregation. Section 3 reviews the relevant background economic theory. Section 4 reports on the status of my research thus far, avoiding most mathematics. More specific expositions and detailed derivations appear in Appendices A and B, which are papers appearing at the Conference of Uncertainty in AI in '97 and '96, respectively. The more relevant is Appendix A, a self-contained, informative account of our past work and research methodology. Finally, Section 5 maps out my proposed future theoretical, empirical, and practical investigations of market-based belief aggregation and group decision making.

2 Worldview and Motivations

This mostly non-technical section serves two functions. First, it introduces my picture of the world and how it relates to my research goals and directions. Subsection 2.1 delineates the assumptions and modeling conventions that I subscribe to: how the "world" or environment is represented, and how agent "citizens" view and interact within this world. Subsection 2.2 provides a clear statement of the research problem I am addressing. Second, Subsection 2.3 addresses my underlying motives for pursuing research on aggregate assessments, and outlines some basic justifications supporting the pursuit of a market-oriented avenue toward solutions. Why, when, and how are belief aggregation and group decision making useful endeavors? What added value does an economic approach to these problems have over more traditional methods? I do not pretend to have complete or foolproof answers to these questions, but attempt to formulate reasonable preliminary responses.

2.1 Multiple Distributed Agents in an Uncertain World

Uncertainty is ubiquitous, ... almost surely. In general, agents cannot be certain about the current state of their environment, the effects of their actions, or the full influence of outside forces on the world, even if we assume that at some lowest level (particle physics?) the world is purely deterministic. There are two equally endemic reasons for this unfortunate reality [18]. The first is *laziness*, or the inability to list the exponential (infinite?) number of deterministic rules needed to fully describe all of the ways the world may unfold—i.e., finding and storing all of the rules and all of the exceptions to all of the rules and all of the addendums to all of the exceptions is simply intractable. The second reason is *ignorance*: we may not have a complete and correct theory delineating every rule that governs the world, and even if we did we may not know all of the relevant factors that trigger the various rules. Thus any realistic "mental model" of the world should include a method to handle uncertainty.

The most widely accepted representation language for uncertainty is standard probability theory, and I join this particular bandwagon confidently and without much hesitation. Some advantages of this theory are its mathematical rigor (making it an excellent choice for computer implementation), its solid and storied history, and its convincing axiomatic justifications [18, 12]. I also choose a *subjective* interpretation of probabilities over a *frequentist* one, although this decision does not affect much of my analysis. There are significant populations of researchers that either do not adhere to a subjective interpretation or dismiss the religion of probabilities altogether. Philosophical discussions are beyond the scope of this proposal; see Section 1.4 and Chapter 2 of [12] for an introduction to some of the issues.

I also accept standard probabilistic decision theory as the correct foundation for a single agent to make optimal, or at least rational, *decisions* in an uncertain environment. In addition to probabilistic beliefs, agents have a utility or preference function ranking the desirability of the potential states of the world. Each agent arrives at an optimal decision (chooses the best action from those available) by "simply" maximizing expected utility.

A common worldview in artificial intelligence (AI) is that of a single "intelligent" or rational agent operating in and acting upon its "dumb" or relatively passive environment. Certainly from the perspective of one agent, outside actors can be categorized under the umbrella term "environment", but it may often be beneficial to ascribe "agenthood" to fellow intelligent decision makers to facilitate inferences about their behavior. The generic situation in almost any domain includes multiple agents interacting either cooperatively or competitively (or both). This realization, along with general trends toward distributed computing across networks (especially the Internet), has fostered rapid growth in research in the subfield of *multi-agent systems* within the larger AI community.

Distributed systems are a special case of multi-agent systems, where

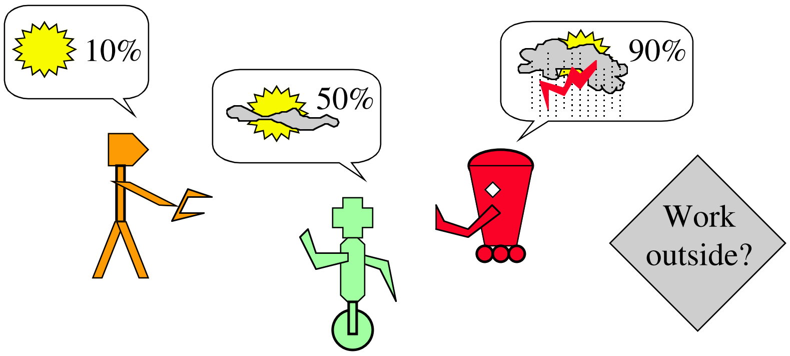

Figure 1: Cartoon depiction of "multiple distributed agents in an uncertain world". Agents are separate individual entities, each with its own subjective probabilistic beliefs (say, concerning the weather) and preferences, and may face a group decision (e.g., "should we work outside today?").

agents are conceptually separated and/or shielded from one another. In principle, optimization problems involving multiple agents can always be solved in a *centralized* manner, on a single (perhaps massively parallel) computer with full knowledge of all relevant inputs; for example, any algorithm designed to be distributed can be simulated on a centralized machine. Motivating a distributed design, then, requires some exogenous constraint(s). Agents may be physically separated (across the Internet?) with high communication costs. Agents (and their human designers or owners) may fear breeches of security or trust, and thus refuse to reveal their detailed inner workings to a central location. Distribution may also facilitate design, by breaking an algorithm up into logically distinct parts. My proposed mechanism for belief aggregation is inherently distributed, and thus its application should be most advantageous when one of these external conditions is an important factor.

And so we arrive at my worldview mantra: "multiple distributed agents in an uncertain world", which is pictured in abstract cartoon form in Figure 1. Beyond this expansive catch phrase, the specific model of the world I employ in this research is of course more restrictive, and I will attempt to list the major assumptions now. First, the set of all possible states of nature, denoted as Ω, is discrete and finite. Statements about the world are limited in expressiveness to probabilities over propositional events. All agents agree on the existence of some core subset of relevant states $\Omega' \subseteq \Omega$, and all group assessments concern these states only (non-shared states may still affect a particular agent's actions). For simplicity, in the remainder of this proposal I will assume that $\Omega' = \Omega$. Participating agents are assumed cooperative in the sense that they agree to respect the derived consensus probability distribution and abide by group decisions. At the same time, whenever possible, agents are allowed to act independently or with *individual rationality*, meaning that each uses its own probabilistic beliefs and utility to optimize personal decisions. I feel this respect for individuality is important for two reasons (besides to espouse idealized political principles). First, in moving to multi-agent environments, we would prefer *not* to throw away all of the positive progress made in single-agent AI studies. Second, this stress on individual rationality may entice agents to join the group dynamics without fear of an inevitable socialist sacrifice of "one for the many".

Having outlined the basic premises and conventions adopted for my "world model", I am now in a position to more formally state the research questions that I am pursuing within this domain.

2.2 What is Belief Aggregation? What is Group Decision Making?

In the beginning there was standard single-agent probabilistic decision theory—and it was good. In this conceptualization, agents have probabilistic beliefs and utility functions across the states of the world Ω, and choose actions in order to maximize expected utility. This theory offers a principled, mathematically sound, and well-accepted framework for an *individual* agent to make optimal or rational decisions under uncertainty. Given several reasonable axioms, this method can be shown to be the *only* way to arrive at optimal decision policies [18, 9]. And so all was right and good with the world called Ω.

And then there were *groups*—collections of agents that wanted to pool their knowledge and opinions, and come to mutual decisions. Suddenly solving problems on Ω becomes a little messier. The analogous task of making *optimal* group decisions proves elusive simply to define, let alone solve unambiguously. In keeping with the principle of maintaining agent independence, groups are assumed composed of rational individuals, each with its own beliefs and preferences. But given inconsistent preferences, which possible outcome is "best" for the group? Given disagreeing beliefs, what can be considered the "true" group assessment of the probability of various out-

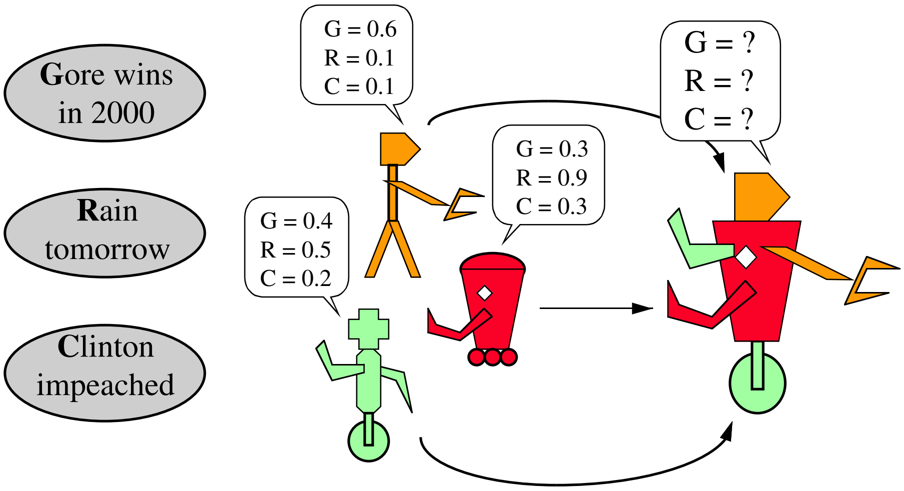

Figure 2: Abstract representation of a belief aggregation function: it inputs individual probability distributions across a set of (possibly related) events—for example concerning politics and weather—and outputs a group consensus distribution. A concrete definition of *optimal* aggregation remains elusive.

come states? Which group members are more confident, more trusted, more accurate, more important? How does one measure such quantities, and what effect should they have on aggregate assessments? What properties of the aggregate probability distribution are most essential?

Postponing these nagging questions regarding *optimal* group belief assessments and decisions (which are largely unanswered and will likely remain so for the foreseeable future) until the next subsection, it *is* possible at least to precisely state the form of inputs and outputs we desire. The operation of *belief aggregation* takes as input the joint probability distributions across Ω of n agents and returns a single "summary" probability distribution. This weakly constrained process is pictured symbolically in Figure 2. Much of the previous work in this field involves characterizing and/or axiomatizing the properties of the output probability distribution for various proposed pooling operators [3, 5, 6, 13, 14, 21, 23]. The take-home message from these studies is that each method has different desirable properties, but (provably) no method has them all. *Group decision making* does not have a completely standard form, though a common one takes as input the same n probability distributions, the agents' n utility or preference functions, and a set of possible group actions, and outputs a single action [8]. Motivated mostly by a desire for simplicity and elegance, many researchers assume that the group decision process generates separately a summary probability distribution and a summary utility function, and thus can make use of ordinary single-agent utility maximization to find the action of choice. In this case, belief aggregation becomes a sub procedure within the larger task of group decision making.

2.3 Why Do We Care?

As promised, we return to the question of *optimality* in multi-agent aggregation, though notably not to an answer. No single solution has emerged in decades of study, and given the well-characterized barriers (i.e., impossibility theorems [8, 6, 19] and definitional controversies [5]) it seems clear that no hands-down winner is likely to prevail. If even basic definitions of optimality for group probabilities and decisions seem unattainable, why should we embark on designing and studying such models at all? What benefits can we expect (or hope for) given the well-known limitations? Finally, in what circumstances is a market-based approach to the problem justified and advantageous? As unanimously lamented by researchers everywhere (and marking an unofficial theme of this proposal), questions come easier than answers. I am just beginning to get a handle on some of the relevant issues; this section traces some of the motivations and reasoning behind my chosen investigatory path.

At a most basic level, there exist many common situations in which aggregate beliefs and collective decisions seem to be both "natural" and beneficial for groups of agents. As an example, consider a group of doctors with disparate subjective beliefs and heterogeneous preferences evaluating the probabilities of diseases and weighing the merits of various treatments. Or, from a more AI-centric perspective, consider culling the advice of these doctors into a probabilistic expert system [13]. A formal, grounded method for combining their opinions and utilities seems almost requisite, if the resulting diagnoses and remedies are to be trusted. Similar examples abound: a company board assessing uncertain business conditions and choosing appropriate strategies; a network of computers gauging potential loads and traffic and allocating memory, processor time, or other resources.

Given current trends toward decentralization in computation and "agent-oriented" programming, rationale for aggregation mechanisms may also arise out of practical convenience. Probability and utility combination may simply be well suited for realizing competent cooperative behavior in an increasingly distributed world. Software agents representing distinct interests and possessing individual knowledge and information-gathering capabilities will form their own beliefs and utilities. No overarching authority will be technically or computationally able to gather all of the relevant information in order to centrally orchestrate their "optimal" group activity. For that matter, no would-be coordinator will obtain permission to access all agents' beliefs and preferences or enforce group behavior. In an agent-centric world, a mechanism for pooling disparate opinions from distinct entities (especially one that respects individuality and operates with distributed protocols) may prove of great practical value.

As discussed above, many researchers agree that the prospect of optimality for group assessments and decisions is doubtful. This sobering realization is partial impetus for my particular research methodology. Since perfect *top-down* design (i.e., posit optimality and derive the implied aggregation function) seems futile, I suggest a mostly *bottom-up* approach: start from desirable, reasonable and/or convenient agent characteristics and mechanisms for interaction, and attempt to formalize the nature of the resulting aggregate assessments. Central to this philosophy is the desire to retain, as much as possible, each agent's individual rationality. I prefer *not* to assume that agents are fully cooperative and trusting, or that agents agree to personal sacrifice to achieve some measure of group optimality—conditions often required of a top-down designed mechanism. Instead I propose creation of a system that provides direct incentives for participation and truthfulness (the magic incentive is, of course, money), and allows well-understood notions of individual rationality to drive group dynamics. Prioritized emphasis on agent independence may also provide for a more practical system from an engineering standpoint: the natural separation into autonomous individuals should aid design and implementation in an actual distributed computing environment.

Classical economic theory provides a solid, formal, and relevant foundation on which to build such a distributed, agent-oriented aggregation system. Notions of distinct, interacting, rational individuals are pervasive in the field of economics. Trading protocols are well documented, are in current widespread use, and transfer smoothly to computer implementation. The elicitation of individual subjective probabilities through a lottery valuation is commonly understood [10]. Market mechanisms are often applied in uncertain domains, and their potential function as aggregators of belief is generally recognized. For example, the price of a stock represents the "market evaluation" of the expected present value of future dividends, and odds in a horse race summarize the bettors' beliefs about the winning horse's identity. Some of the effects of asymmetric uncertainty on security prices have been examined [20]. Our motivating philosophy is that, while decision theory enables an *individual* agent to choose optimal actions under uncertainty, economic or market-based systems provide the necessary extensions to form a general tool for distributed *group* decision making [22].

3 Economics Background

This section covers the relevant economic concepts appealed to in our models. As computer scientists, we feel most comfortable utilizing basic or classical theories, as we are more confident in our understanding of these principles than of more recent, more controversial, or less clearly delineated areas of the field.

Uncertainty often enters into economics in the form of *contingent goods*: coupons that pay off in tangible goods (grain, hogs, widgets, money, whatever) only contingent upon the outcome of some uncertain event. A *security* is a specific type of contingent good that pays off in monetary units (in our models, all securities pay off $1). A rational participant in the market will *demand* quantities of goods to buy (positive demand) or sell (negative) that maximizes its utility.

A market is in *competitive equilibrium* at prices for which aggregate or total demand is zero; that is, supply (negative demand or orders to sell) and positive demand (orders to buy) balance across all participating agents. At equilibrium prices, the market's allocation of goods is *pareto optimal*: no other allocation can increase one agent's utility without a corresponding decrease for another agent.

The details of particular trading and price adjustment protocols designed to facilitate market *convergence* to equilibrium are not necessary for theoretical equilibrium analysis. However experimental investigation and real applications do require specific mechanisms. Two software tools developed in our research group can be employed in this regard: the WALRAS market-oriented programming environment and the Michigan Internet AuctionBot. Each provides fairly general facilities to set up and run distributed computational markets, usually implementing an iterative, incremental, "greedy" price adjustment policy. Given a generic collection of agents and goods, existence of an equilibrium is not guaranteed; furthermore, given that a (unique) equilibrium *does* exist, greedy price adjustment is not guaranteed to converge to (it) one. However there are precisely characterized and well-documented conditions under which existence, uniqueness, and/or convergence *are* guaranteed.

4 The Products of Previous Progress: Possibly Profound / Probably Promising / Purely Preliminary

My initial inquiries into the proposed dissertation topic have focused primarily on markets for belief aggregation, postponing investigation of utility combination and the general group decision problem until a later date (see Section 5). Research thus far has yielded several intriguing results, in my opinion displaying enough promise to warrant further in-depth (i.e., dissertation-level) study. This section is organized in roughly chronological order of research developments. Subsection 4.1 briefly describes our first stab at market-based belief aggregation, which established an existence proof that any joint probability distribution can be represented in a specialized economy, called a *MarketBayes* economy. The more recent results summarized in Subsection 4.2 stem from what I believe is a more direct, intuitive approach. That subsection sets up the basis for a *securities market* for belief aggregation, highlights some analysis of the model, and begins to assess its usefulness for the task at hand.

4.1 MarketBayes

A *Bayesian network* is a widely used, compact way to represent arbitrary joint probability distributions. In initial efforts toward handling belief aggregation, we demonstrated that the same probabilistic information can be represented in a general equilibrium economic model. Specifically, we define a precise mapping from any Bayesian network with binary nodes to an economy (called a *MarketBayes* economy) of securities, consumers, and producers. There is a direct (provable) correspondence between the unique equilibrium prices of securities in this economy and the original probabilities from the Bayesian network. This subsection only outlines the basic concepts—more details and discussions appear in Appendix B.

A Bayesian network defines the necessary conditional dependencies among a set of uncertain events or propositions. In the translation to a MarketBayes economy, some (possibly compound) events become securities, each paying off \$1 if its corresponding event occurs, nothing otherwise.

In a Bayesian network, links between event nodes encode conditional probabilities. For example, a single link from node A to B is accompanied by the information $\Pr(B|A) = k$ where k is some probability. The same equation can be rewritten as: $\Pr(AB) = k \Pr(A)$. In a MarketBayes economy, the *consumers* effectively implement equations of this form. AB and A are securities, and the consumer's preference for AB is k times that of A. If the ratio of the prices $\Pr(AB)/\Pr(A)$ diverges from k, the consumer will buy or sell according to its preference, driving the ratio toward k. Note that in the MarketBayes framework, consumers technically have preference directly over securities, *not* actual beliefs in the probability of events.

In a Bayesian network, the laws of probability are inherent in the inference mechanism. In a MarketBayes economy, *producers* encode probabilistic identities. For example the relation $\Pr(A) = \Pr(AB) + \Pr(A\bar{B})$ is enforced by a producer that has the technology to "transform" one A into one AB and one $A\bar{B}$, and vice versa. If the price $\Pr(A)$ diverges from the price $\Pr(AB) + \Pr(A\bar{B})$, the producer will purchase the cheaper security (bundle), and sell the equivalent one in the higher priced market—thus driving the two prices together. This type of producer is an *arbitrageur* since it capitalizes on inconsistencies between related prices.

We have found that consumers and producers of the forms described above are sufficient to encode any Bayesian network with binary nodes. So far, the MarketBayes system has been analyzed only in the case of one consumer per conditional event. In order to address belief aggregation, the economy would include multiple consumers per conditional, each with varying preferences, confidence or risk aversion, and endowments or resources. Thus currently the main contribution of this work is an existence argument that any joint probability distribution can be encoded in a computational market economy.

4.2 A Securities Market for Belief Aggregation

This subsection presents more recent theoretical work that *does* specifically address pooling the beliefs of multiple agents. This model begins from more fundamental postulates than MarketBayes: agents have beliefs over events and utility for money, instead of direct preferences for securities. I describe the results at a course-grained level, highlighting important points. More detailed information, discussion, and formal analysis appears in Appendix A, a self-contained conference paper that should be valuable for understanding and assessing our research. The subsection is organized into four parts. The first lays the groundwork: it introduces the basic market setup (agents and securities) to facilitate belief aggregation, and defines relevant notation. The second part delineates any assumptions made regarding agent behavior in order to arrive at closed-form solutions, and attempts to motivate them. The third part lists some of the implied characteristics of a single utility-maximizing agent's demand when participating in certain simple markets. The last part covers properties of equilibrium prices (consensus beliefs) that arise from group members' collective interactions in the market.

4.2.1 Market Setup

I adopt relatively common and straightforward agent and market characteristics. An influential philosophy of design is to start with fundamental principles, emphasize simplicity throughout, explicitly catalog and attempt to motivate assumptions, and then derive the implied results. This research strategy is embraced partially out of a general desire for elegance and an attraction to Okham's razor. Such formal clarity should also make it easier for us, and other interested researchers, to both understand and judge the outcome.

Our notation conforms to the worldview conventions introduced in Subsection 2.1. The set of all possible world states is $\Omega = \{\omega_1, \omega_2, \ldots\}$, where the ω are disjoint atomic states or events. Events of interest are denoted with capital letters near the beginning of the alphabet $(A, B, C, \ldots)$, each a (possibly degenerate) set of atomic events (e.g., $A = \{\omega_4\}, B = \{\omega_2, \omega_9, \omega_{11}\}, C = \emptyset$). Each agent has a subjective joint probability distribution $\Pr(\cdot)$ over all states. Its prior probability for the event A is then just $\Pr(A) = \sum_{\omega \in A} \Pr(\omega)$. Each agent also has a utility $u(y)$ for y dollars. Do not confuse this utility for money with utility or preference for the possible states of nature Ω. We have *not* defined the latter here, since we are not yet delving into issues of utility combination or group decisions. When there are multiple agents, we denote the probability and utility functions of agent i by $\Pr_i$ and u_i, respectively.

The market consists of one security for every event of interest. The security "\$1 if A" is worth one dollar if A occurs (if the current world state

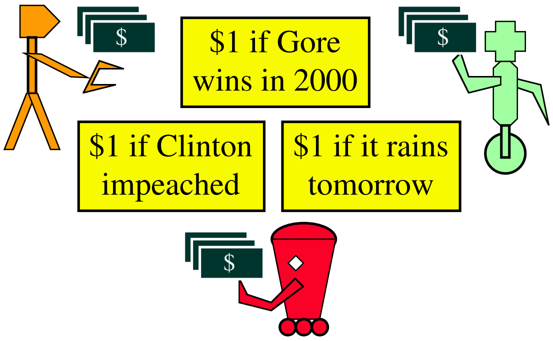

Figure 3: Basic setup of a market for belief aggregation. Securities pay off contingent on events of interest. Each agent purchases or sells securities in order to maximize its own expected utility.

$\omega_c \in A$), zero otherwise. If the state of the world (and of the event A) is uncertain, an agent will value this security according to its $\Pr(A|\cdot)$ (prior or conditional on observed evidence) and its utility for money $u(y)$. Specifically, an agent's demand for securities is chosen to maximize its expected utility. This generic setup is pictured in Figure 3.

The equilibrium prices of securities are interpreted as the consensus probabilities for the corresponding events. There is a certain "leap of faith" initially needed to justify this interpretation of prices, though it seems intuitive and others have made the same connection (see the discussion in "Related Work" Section 5 of Appendix A about Idea Futures, the Iowa Electronic Markets, and Hanson's work [7]). An important part of my research is to find more formal justifications by examining the properties of equilibrium prices implied by various instantiations of agents and securities. Another facet of my research is to characterize (theoretically and/or empirically) the necessary and sufficient conditions for existence of, uniqueness of, and convergence to equilibrium prices.

4.2.2 Assumptions

All of my formal results assume that agents are *risk averse* for money; that is, $u(y)$ is an increasing concave function. An agent is risk averse if and only if it always prefers the expected payoff of a lottery *for certain* to the (risky) lottery itself. A *risk neutral* agent will tend to invest all of its resources in the one most attractively priced security. Risk aversion places natural bounds on investment and encourages an agent to diversify. Demand functions for risk neutral agents will in general *not* be continuous or differentiable; demand for risk-averse agents should be more readily amenable to analysis.[1] The majority of my derivations (so far) require a stronger assumption that agents have *constant* risk aversion, or that $u(y) = -e^{-cy}$ where c is its risk aversion coefficient. I chose this functional form mainly for its simplicity: it seemed a natural starting point, and allows for tractable closed-form derivations.

Agents are assumed *competitive*: they maximize utility at current or given prices, ignoring any effect that their demands may have on prices. This is a common assumption in economic theory, and generally holds when the agent's resources are small compared to the market as a whole. Agents also have *fixed beliefs*, uninfluenced by the potentially valuable information that prices reveal about other agent's beliefs. This assumption certainly aids mathematical analysis, but is not otherwise easily justified. More realistic market models for belief aggregation may need to consider agents that "learn" from prices. On the other hand, since we consider only properties of aggregate prices *in equilibrium*, beliefs may well be considered "fixed" only *after* Bayesian updating from market price information. Finally, each agent's utility for dollars $u(y)$ does not depend on the state of the world—this is also an oft invoked assumption, though admittedly adopted here primarily for analytic convenience.

4.2.3 Individual Demand

The following list outlines several formal results concerning individual agent demand. Details and proofs appear in Appendix A.

- A generic risk averse agent's demand for a single security is qualitatively "correct": positive (zero, negative) when the price is less than (equal to, greater than) its probability for the corresponding event. If the agent is *sufficiently* risk averse, and prices are consistent with the logical structure of events, then its optimal demand is bounded. If prices are inconsistent (for example, two equivalent securities have different prices, or the sum of the prices of a set of disjoint events is greater than

[1] Human activity can often be classified as risk averse although, as might be expected, a variety of other (often seemingly irrational) behaviors have been observed [11]. In this work I am not interested in modeling human behavior, only building competent machines.

one) then agents have an opportunity to increase utility without bound through *arbitrage*.

- For an agent with constant risk aversion, we can derive an equation for its demand for a single security as a function of price, probability and risk coefficient. As an agent's risk aversion approaches zero (approximating risk neutrality) it is willing to buy or sell increasing amounts of the security. In a sense, a smaller risk aversion indicates greater confidence in beliefs—the agent is willing to wager more to take advantage of a discrepancy between price and probability.

- Also for constant risk aversion, we can formulate an agent's expected utility for *two* securities, and derive two (numerically solvable) coupled equations that implicitly define the agent's demand for each security. Appendix A proves several propositions about the qualitative correlations between demand for one security and demand for the other.

4.2.4 Equilibrium Prices ⇔ Consensus Beliefs

The previous subsection dealt with single agent demand. We now consider some theoretical properties of equilibrium prices that emerge from *group* dynamics.

We can derive a formula for the competitive equilibrium price in a market consisting of agents (with constant risk aversion) bidding on a single security. Appendix A enumerates several desirable or otherwise interesting properties of this price when interpreted as the group's aggregate probability. Three significant attributes deserve special attention:

- The functional form the equilibrium price equation is a normalized geometric mean. In the context of the belief-aggregation literature, this is a normalized version of the *logarithmic opinion pool* (LogOP) for a single event [4, 5, 21, 23]. This establishes a direct connection from our market model, based on individually rational distributed interacting agents, to a well-known *centralized* pooling mechanism. It also provides a decision-theoretic interpretation for the notoriously slippery concept of "expert weights". In the usual interpretation, the weighting exponents in the geometric mean encode some sort of degree of expertise, confidence, or reliability, and are almost always chosen in an ad

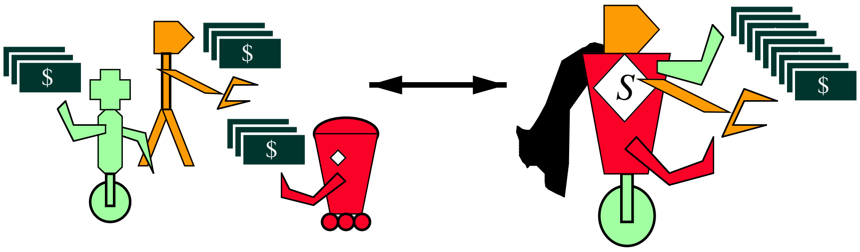

Figure 4: The aggregate demand of a group of agents is the same as that of a single representative "super-agent" with belief equal to the group's equilibrium price, and risk aversion less than that of any individual member agent.

hoc manner [3, 5, 6, 14]. In our model, the *derived* weights are a normalized inverse measure of risk aversion. In general, the unnormalized LogOP does *not* generate a probability ($\Pr(A) + \Pr(\bar{A})$ may be < 1), while the normalized version does.

- The equilibrium price of a security in a *two* security market, where the events are *disjoint*, also obeys the normalized LogOP form. I tentatively conjecture that the same correspondence holds for an arbitrary number of disjoint events. A proof of this would be significant since any joint probability distribution can be (and often is) represented across disjoint events.

- Suppose that a group of N agents with beliefs $\Pr_1(A), \ldots, \Pr_N(A)$ and constant risk aversion coefficients $c_1, \ldots, c_N$ effect an equilibrium price $p*$ in a single-security market. Then their aggregate (total) demand is equal to the demand of a single representative "super-agent" with belief $\Pr(A) = p*$ and risk aversion c such that $1/c = (1/c_1 + \ldots + 1/c_N)$. To an outside observer, this super-agent behaves as a rational "individual" in exactly the same sense as expected of a single agent—its demand can be rationalized as maximization of an expected group utility for money. Moreover, for any collection of agents (assuming finite positive c_i), the super-agent risk aversion is *strictly less than* that of any participating individual. Thus the group as a whole is willing to take on more risk, acting in some sense with more "confidence" than any member alone. Continuing with our surreal series of symbolic cartoon abstractions, "Super-Agent Man" is pictured in Figure 4.

5 Proposals and Predictions:
Plans / Potential Paths and Pitfalls / Pace of Progress

This section catalogs my future goals and directions, suggesting a road map for progress toward the eventual dissertation. Subsection 5.1 discusses theoretical and experimental extensions to the current belief aggregation framework. Subsection 5.2 explores possible generalizations our system to handle utility combination and thus group decisions. Subsection 5.3 arranges these various goals in chronological order of expected completion date.

5.1 Belief Aggregation

The most concrete avenue for further study continues where Section 4 leaves off, generalizing theoretical analysis of our market model for belief aggregation. An obvious extension would be to consider more complex sets of securities. One specific near term focal point is to verify (or falsify) the conjecture (mentioned in Subsection 4.2.4) that, for constant risk aversion, our mechanism corresponds to the normalized LogOP for an arbitrary number of disjoint events. An affirmative answer would establish the (surface) equivalence of the two mechanisms for agents with any joint probability distribution (since it can always be factored into disjoint events). Another direction would be to look at broader classes of risk-averse utility functions (non-constant), and even possibly risk-prone and risk-neutral forms. This seems a natural way to understand and motivate generalizations of the normalized LogOP. I will also pursue formal characterizations concerning the existence and uniqueness of price equilibria, and the convergence to these equilibria via distributed bidding protocols and classical economic price adjustment.

I plan concurrent empirical investigations in richer market domains where theoretical analysis becomes intractable. Such economies may simply involve non-constant or non-risk-averse agents, or multiple correlated securities. They may also allow non-competitive agents and/or "learning" agents that update beliefs from observed prices; each extension will entail tests of existence, convergence, and properties of aggregate prices. As mentioned earlier, our research group's WALRAS market-oriented programming environment and Michigan Internet AuctionBot are both suitable computational market testbeds for these experiments.

There are several so-called "impossibility results" associated with belief aggregation mechanisms, essentially proving that no one pooling function can possess all potentially desirable attributes [6, 19]. Stated another way, given certain reasonable axioms, *no* reasonable pooling function is possible. Formalizing exactly which properties our mechanism has (and thus implying which it cannot have) is a necessary step to placing my work in its proper conceptual place within the larger body of belief aggregation results.

I also need to better assess how my research fits into the context of the finance literature. Economists are familiar with the concepts of asymmetric uncertainty and heterogeneous preferences, and the interpretation of price as an aggregate probability. Varian has studied the effect of the dispersion of agents' subjective probabilities on security prices [20]. There is a body of work concerning *price revelation*, where agents update their beliefs according to observed price information. A subset of this literature looks at *myopic* agents that can learn from prices, but do not consider the effect of their (or others') actions past the current time horizon. Up to this point, I have not tapped into most of these resources; obviously I have considerable homework left to do.

5.2 Group Decision Making

Thus far the bulk of my efforts have centered on combining the disparate beliefs of a collection of agents. This task does have a large, mostly theoretical legacy and limited direct applicability. However, the next step in building a more practical multi-agent uncertain reasoner is to add the capability of coming to group decisions. This more general problem must take into account agents' utilities or preferences, in addition to beliefs. The collective decision does not necessarily require, as an intermediate step, the formation of aggregate beliefs and utilities. If, on the other hand, we *can* somehow arrive at both a group probability distribution and a group utility, then good old reliable single-agent theories allow us to infer rational decisions.

For the degenerate case where all agents have equivalent utility functions, the "optimal" group utility is almost unarguably that of the individuals. In this situation we already have at least a well-defined decision making mechanism, though any evaluation of the *quality* of such decisions remains to be seen.

For the generic case of heterogeneous preferences, I envision some sort of distributed market-based approach to utility combination, with the same advantage of monetary incentives for participation and truthfulness, and with a continued emphasis on individual rationality. Any details of a final realization of this goal remain admittedly very hazy. One possible starting point is the group's representative "super-agent", described in Subsection 4.2.4 and in more detail in Appendix A. This conceptual group-agent *does* have *both* an aggregate belief and an aggregate utility; however its consensus utility is for *money*, not for the relevant states of nature. Perhaps a missing link can be grafted on that connects the two types of utility in a way meaningful for decision making (or perhaps not).

Although the exact blueprint of a group decision maker is currently difficult to imagine, some tools for evaluating the effectiveness of such a system are already available. In general, assessments would be based on *game-theoretic* notions: one natural test would be for pareto optimality [8], another for incentive compatibility.

The study of group decision making, and heterogeneous preferences in general, is also fraught with its own history of impossibility results [8, 19, 1, 2]. Coming to a full understanding of these limitations should constrain the search for, and thus hopefully aid in the discovery of, any workable mechanism.

Some of my underlying motives for studying market-based aggregation (of beliefs *and* of utilities) were introduced in Subsection 2.3, where I admitted that current justifications are less than airtight. Concurrent to advancing toward the above stated end goals, I would like to shore up my beginning motives and thus refine the foundational reasons for my proposed research path.

Finally, I intend to build a "proof of concept" application of the proposed market-based group decision maker. For this "real-world" test, the AuctionBot will likely be the vehicle for implementation; it is an actual distributed (Internet-based) programming environment, whereas WALRAS is only a conceptually distributed algorithm simulated on a single machine. A convenient experimental domain may be supplied from some wing of our research group, or from one of my other research interests, which include satisfiability testing [15], complexity[2], and machine learning. I do not expect this to be a large, fully practical test, only a "toy" version of a real problem.

[2] I am pursuing a graduate certificate in the new Program for the Study of Complex Systems (PSCS) at the University of Michigan.

5.3 Timetable

This subsection arranges my projected goals in expected chronological order of completion, indexed by ETA[3]. Each item is titled with a unifying conceptual theme, though boundaries will certainly be crossed. The list is intended only as a rough guide for my own use in focusing and assessing progress, as well as for readers to judge the legitimacy of achieving my scheduled objectives.

September 1, 1997. A strengthened understanding of belief aggregation. Prove (disprove) correspondence between LogOP and securities market with disjoint events. Explore some theoretical properties of a standard decreasingly risk-averse agent bidding on a single security. Catalog properties of our initial belief aggregation mechanism and how they fit in with "impossibility" theorems. Conduct search of finance literature. Write computer implementation to handle multiple securities. Conduct preliminary empirical examination of existence, uniqueness, convergence, and equilibrium price properties for two and more securities. Brainstorm ideas for an example application domain.

January 1, 1998. Joining belief and utility aggregation for group decision making. Continue literature search: finance, group decision making, multi-agent and distributed AI, and of course belief aggregation. Formulate possibilities for utility combination and place in context of group decision making impossibility results. Define a group decision mechanism (possibly appealing to a utility aggregation "oracle"): how to marry utility and belief combination for group decisions. List relevant and/or potential game-theoretic tests for evaluating group decisions. Continue theoretical belief aggregation work: pair risk aversion forms with derived pooling functions, analyze multiple security case in more detail, try proofs of existence, uniqueness, and convergence. If working in WALRAS probably move experiments to AuctionBot. Continue empirical work: various risk forms and combinations of securities. Attempt to refine answers to "why" questions from Subsection 2.3. Narrow down choice of application domain and suggest some possible example decision problems. Think about how the original MarketBayes work may fit into the now "bigger" picture.

[3]Estimated Time of Altering course.

May 1, 1998. Agents that learn. Understanding, analyzing, and assessing group decision making. Continue literature search (see previous two items). Empirical work on "price revelation": add some "learning" agents that update beliefs according to observed prices and test as thoroughly as possible. Theoretical work: finalize a precise method for market-based utility combination, and examine some implied properties for simple decision making problems. Add utility aggregation and decision making capability to computer implementation. Decide which tests (game-theoretic or otherwise) of decision quality are most appropriate, or at least seem most valuable. Continue to refine motives from Subsection 2.3. Finalize application domain and define example decision problem.

September 1, 1998. Experiments and practical application. Continue literature search (see previous three items). Conduct simple experiments of decision maker, and evaluate mechanism (equilibrium convergence) and decision quality (using previously identified tests). Define a precise and rigorous empirical methodology for exploring the application domain and assessing results. Begin "proof of concept" experiments.

December 15, 1998. Write dissertation. Continue literature search (see previous four items)—it never ends! Run carefully planned experiments, carefully log results, carefully analyze outcomes, carefully argue that "concept is now proven". Organize all previous results; try as best as possible to demarcate an overall conceptual "whole". Continue to fill in any obvious theoretical and/or empirical gaps. Collect together all results, flaws, unfinished work, future ideas and begin to write, continue to write, keep on writing, write often, write well, write some more, and then revise often, revise well, and revise some more. Complete and defend dissertation. Drink beer regardless of outcome.

Acknowledgments

I am grateful for frank and helpful comments, and pointers to relevant literature, from Robin Hanson, Mike West, Jeffrey MacKie-Mason, and students in the UM Decision Machines Group. Special thanks to Michael Wellman for extensive editing comments as well as contributions to the reported motivations, goals, and results.

References

[1] Kenneth J. Arrow. *Social Choice and Individual Values.* Yale University Press, second edition, 1963.

[2] Kenneth J. Arrow. Values and collective decision-making. In Peter Laslett and W. G. Runciman, editors, *Philosophy, Politics and Society (third series)*, pages 215–232. Basil Blackwell Oxford, 1967.

[3] Jon Atli Benediktsson and Philip H. Swain. Consensus theoretic classification methods. *IEEE Transactions on Systems, Man, and Cybernetics*, 22(4):688–704, 1992.

[4] Norman C. Dalkey. Toward a theory of group estimation. In H. A. Linstone and M. Turoff, editors, *The Delphi Method: Techniques and Applications*, pages 236–261. Addison-Wesley, Reading, MA, 1975.

[5] Simon French. Group consensus probability distributions: a critical survey. *Bayesian Statistics*, 2:183–202, 1985.

[6] Christian Genest and James V. Zidek. Combining probability distributions: A critique and an annotated bibliography. *Statistical Science*, 1(1):114–148, 1986.

[7] Robin D. Hanson. Could gambling save science? Encouraging an honest consensus. *Social Epistemology*, 9(1):3–33, 1995.

[8] Aanund Hylland and Richard Zeckhauser. The impossibility of Bayesian group decision making with separate aggregation of beliefs and values. *Econometrica*, 47(6):1321–1336, 1979.

[9] Savage L. J. *The Foundations of Statistics.* Wiley, New York, 1954.

[10] Edward E. Leamer. Bid-ask spreads for subjective probabilities. In P. Goel and A. Zellner, editors, *Bayesian inference and decision techniques*, pages 217–232. Elsevier Science, North-Holland–Amsterdam, 1986.

[11] Mark J. Machina. Choice under uncertainty: Problems solved and unsolved. *Journal of Economic Perspectives*, 1(1):121–54, Summer 1987.

[12] Richard E. Neapolitan. *Probabilistic Reasoning in Expert Systems: Theory and Algorithms*. John Wiley and Sons, New York, 1990.

[13] Keung-Chi Ng and Bruce Abramson. Consensus diagnosis—A simulation study. *IEEE Transactions on Systems, Man, and Cybernetics*, 22(5):916–928, 1992.

[14] Keung-Chi Ng and Bruce Abramson. Probabilistic multi-knowledge-base systems. *Applied Intelligence*, 4(2):219–236, 1994.

[15] David M. Pennock and Quentin F. Stout. Exploiting a theory of phase transitions in three-satisfiability problems. In *Proceedings of the 13th National Conference on Artificial Intelligence*, pages 253–258, Portland, OR, USA, August 1996. AAAI Press.

[16] David M. Pennock and Michael P. Wellman. Toward a market model for Bayesian inference. In *Proceedings of the 12th Conference on Uncertainty in Artificial Intelligence*, pages 405–413, Portland, OR, USA, August 1996.

[17] David M. Pennock and Michael P. Wellman. Representing aggregate belief through the competitive equilibrium of a securities market. In *Proceedings of the 13th Conference on Uncertainty in Artificial Intelligence*, Providence, RI, USA, August 1997. To Appear.

[18] Stuart J. Russell and Peter Norvig. *Artificial Intelligence: A Modern Approach*. Prentice-Hall, New Jersey, 1995.

[19] Donald G. Saari. Chaotic exploration of aggregation paradoxes. *SIAM Review*, 37(1):37–52, March 1995.

[20] Hal R. Varian. Differences of opinion in financial markets. In Courtenay C. Stone, editor, *Financial Risk: Theory, Evidence and Implications*, pages 3–37. Kluwer Academic, Norwell, MA, 1989.

[21] S. Weerahandi and J. V. Zidek. Multi-Bayesian statistical decision theory. *Journal of the Royal Statistical Society. Series A. General*, 144(1):85–93, 1981.

[22] Michael P. Wellman. The economic approach to artificial intelligence. *ACM Computing Surveys*, 27(3):360–362, 1995.

[23] Mike West. Bayesian aggregation. *Journal of the Royal Statistical Society. Series A. General*, 147(4):600–607, 1984.

[24] James V. Zidek. Group decision analysis and its application to combining opinions. *Journal of Statistical Planning and Inference*, 20(3):307–325, 1988.

Appendix A

David M. Pennock and Michael P. Wellman. "Representing Aggregate Belief through the Competitive Equilibrium of a Securities Market", *Proceedings of the 13th Conference on Uncertainty in Artificial Intelligence (UAI-97)*, Providence, RI, USA, 1997. To Appear.

Appendix B

David M. Pennock and Michael P. Wellman. "Toward a Market Model for Bayesian Inference", *Proceedings of the 12th Conference on Uncertainty in Artificial Intelligence (UAI-96)*, pp. 405-413, Portland, OR, USA, August 1996.