Graphical Models for Groups: Belief Aggregation and Risk Sharing

David M. Pennock
Yahoo! Research Labs
74 N. Pasadena Ave., 3rd Floor
Pasadena, CA 91103 USA
pennockd@yahoo-inc.com

Michael P. Wellman
University of Michigan AI Laboratory
1101 Beal Avenue
Ann Arbor, MI 48109-2110 USA
wellman@umich.edu

June 14, 2004

Abstract

We investigate the use of graphical models in two fundamental problems of group coordination: (1) reaching a consensus on beliefs, and (2) allocating risk. On the negative side, we prove that under mild assumptions, even if every member of a group agrees on a graphical topology, no method of combining their beliefs can maintain that structure. Even weaker conditions rule out local aggregation within the conditional probability tables of the graphical models. We show that the linear opinion pool (LinOP) and the logarithmic opinion pool (LogOP) are both NP-hard to compute, even for queries easy to compute for every individual. In terms of risk sharing, we show that securities markets structured like graphical models are generally no more tractable than complete securities markets, the unattainable gold standard for optimal risk allocation. On the positive side, we show that if probabilities are combined with LogOP, then commonly-held Markov independencies are maintained. We give procedures for constructing a graphical model of LogOP-aggregated beliefs, and from this computing any LogOP query with time complexity comparable to that of exact Bayesian inference. We give conditions under which optimal risk sharing can be obtained more efficiently by structuring a securities market like a graphical model. One sufficient condition is for all agents' risk-neutral independencies to agree with the independencies encoded in the securities market. A second sufficient condition is agreement on Markov independencies among agents all with constant absolute risk aversion.

1 Introduction

Mathematical models of group coordination are well studied across a number of disciplines, including statistics, decision theory, economics, political science, and artificial intelligence. Common to nearly all formal treatments is a hobbling and unsettling number of paradoxes and impossibility theorems. We focus on two common forms of group coordination: belief aggregation and risk allocation. Belief aggregation is commonly framed as the task of extracting a summarized report of the opinions of a panel of experts, where "opinions" are in the form of subjective probability assessments. More formally, given N agents, indexed $i = 1, \ldots, N$, each with a subjective probability distribution $\Pr_i$ over a state space Ω, an *opinion pool* f is a function that aggregates the agents' beliefs into a single belief, denoted $\Pr_0$:

$$\Pr_0 \equiv f(\Pr_1, \Pr_2, \ldots, \Pr_N). \tag{1}$$

The two most common opinion pool functions are (1) the linear opinion pool (LinOP), or weighted algebraic average, and (2) the logarithmic opinion pool (LogOP), or weighted geometric average.

Risk allocation is another important group activity, perhaps best exemplified by the insurance industry, where risk averse individuals give up a small amount of their *expected* future wealth in return for a decrease in the *variance* of their future wealth. Since the insurance company is less risk averse than the individual, the exchange is mutually beneficial. More generically, agents can reallocate their risks arbitrarily by exchanging *securities*, or state-contingent payoffs: for example, the sender of a package might purchase the state-contingent payoff "$100 if the package is lost" from a courier service. An allocation of risk is said to be *Pareto optimal* if all other possible allocations are worse for at least one of the agents (ignoring ties).

We examine the extent that graphical models can facilitate belief aggregation and risk allocation. Graphical models utilize a language for encoding independencies among random variables. The great power of graphical models is their ability to efficiently encode joint probability distributions when there is sufficient structural independence among variables. How can graphical models help in belief aggregation or risk allocation? The most direct way is by making whatever aggregation or allocation is required more computationally efficient, in terms of both time and space complexity. In this regard, we derive both negative and positive results. For belief aggregation we show that the computational savings enjoyed by representing each individual's probability distribution graphically in general does not carry over to the aggregate distribution, regardless of the aggregation function used. Moreover, for the LinOP and LogOP, computing the aggregate cannot be (trivially) split up into local computations on graph modules: in fact, in both cases the computation is NP-hard. On the positive side, we show that, although the LogOP does not maintain all independencies, it does maintain *Markov* independencies, which correspond directly to a type of graphical model called a Markov network (and the independencies in a moralized and triangulated Bayesian network "join tree"). We give a corresponding algorithm for computing the LogOP that is comparable in time complexity to exact Bayesian inference: exponential in the largest clique size rather than the graph size, in some cases providing an exponential speedup.

The picture for risk sharing is similarly mixed. Securities markets can be structured in analogy to graphical models, leading to a potentially exponential savings in the number of markets required to support a Pareto optimal equilibrium. However, because independencies are not preserved in aggregation, in general the securities market graph must be fully connected, and is thus no less intractable than a *complete* securities market, the unattainable gold standard for optimal risk allocation. On the other hand, we provide some (fairly strong) conditions under which optimal risk sharing can be obtained more efficiently. One sufficient condition is for all agents' risk-neutral independencies to agree with the independencies encoded in the securities market. A second sufficient conditions is agreement on Markov independencies among agents all with constant absolute risk aversion.

The second way that graphical models might help in a group coordination setting is more subtle. Most impossibility theorems arise when the space of possible inputs is completely generic (i.e., the *universal domain*). Although graphical models can describe a rich set of independence relations, they cannot describe *all* possible types of independencies. So it is conceivable that some impossibility theorems might be circumvented when inputs must take a form representable by a graphical model. This would not be a terribly severe restriction, since the types of independencies representable by graphical models are generally regarded as very diverse, natural, and reasonable. Unfortunately, although we show that restricting inputs to graphical representations does circumvent at least one impossibility theorem, it does so in a fairly trivial way, and we further find that most impossibility theorems continue to hold. Moreover, new impossibility theorems arise that confound the conventional wisdom of how people expect graphical models to combine.

This paper is organized as follows. Section 2 presents all the requisite background material and introduces our notation. Section 3 describes the use of graphical models for belief aggregation. Section 4 explores the value of graphical models for structuring securities markets and facilitating risk sharing. We conclude in Section 5.

2 Background and Notation

We consider a group of N agents, indexed $i = 1, 2, \ldots, N$, each with a subjective probability distribution $\Pr_i$ over states of the world and a utility function u_i for money. Denote the set of all possible states of the world as $\Omega = \{\omega_1, \omega_2, \ldots\}$. The ω are mutually exclusive and exhaustive.

State is often more concisely and naturally characterized as the set of outcomes of *events*. Denote the set of modeled events as $Z = \{A_1, A_2, \ldots, A_M\}$. Underlying M arbitrary events is a state space Ω of size $|\Omega| = 2^M$, consisting of all possible combinations of event outcomes. Conversely, any set of states can be factored into a set of $M = \lceil \lg |\Omega| \rceil$ events. Without further assumption, the two representations are equivalent in both expressivity and size, although the event factorization may be more natural. In most of what follows, the events $\{A_j\}$ are the focus of attention, with Ω the implied joint outcome space. We refer to the $\{A_j\}$ as the *primary events*, so as to distinguish them from the other $2^{2^M} - M$ possible sets of states, each of which is also an event.

2.1 Decision Making Under Uncertainty

In general, an agent's utility is defined over the cross product of available actions and possible states. We assume here that utility arises from an underlying utility for *money*. If agent i's utility for μ dollars is $u(\mu)$, then its utility U for a particular action a is its expected utility for money,

$$U_i(a) = E_i\left[u_i\left(\Upsilon_i^{\langle\omega\rangle}\right)\right] = \sum_{\omega\in\Omega} \Pr_i(\omega) u_i\left(\Upsilon_i^{\langle\omega\rangle}\right), \tag{2}$$

where $\Upsilon_i^{\langle\omega\rangle}$ is agent i's wealth in dollars when action a is taken in state ω (the dependence of $\Upsilon_i^{\langle\omega\rangle}$ on a is implicit). Agent i's *decisions* are made by maximizing expected utility, or choosing the action a that maximizes (2).

We assume throughout that utility increases monotonically with wealth. *Local risk aversion* at μ, denoted $r_i(\mu)$, is defined as $r_i(\mu) \equiv -u_i''(\mu)/u_i'(\mu)$. Agent i is *risk-averse* if $r_i(\mu) > 0$ for all μ, or, equivalently, if u_i is everywhere concave. Under this condition, the agent always prefers a guaranteed payment equal to the expected value of a lottery rather than the lottery itself, thus exhibiting an "aversion" to gambling. The agent is *risk-neutral* if $r_i(\mu) = 0$ for all μ, or u_i is linear; in this case, maximizing (2) coincides with maximizing expected payoff.

2.2 Risk-Neutral Probability

Notice that an outside observer, privy only to agent i's chosen actions, cannot uniquely discern either the agent's belief or its utility: the two quantities are inextricably linked (Kadane & Winkler, 1988). Any one of a continuous family of belief–utility pairs offers an equally valid rationalization for the agent's actions. That is, for any function $f(\omega)$, subjective probabilities proportional to $\Pr_i(\omega)f(\omega)$ matched with utilities $u_i\left(\Upsilon_i^{\langle\omega\rangle}\right)/f(\omega)$ result in strategically equivalent utilities for actions $U_i(a)$.

Risk-neutral probabilities are defined as

$$\Pr_i^{\mathrm{RN}}(\omega) \propto \Pr_i(\omega) u_i'\left(\Upsilon_i^{\langle\omega\rangle}\right), \tag{3}$$

where u_i' is the derivative of utility (Nau, 1995). Agent i's observable behavior, manifested as actions, is indistinguishable from that of a hypothetical agent with transformed probabilities $\Pr_i^{\mathrm{RN}}(\omega)$ and reciprocally transformed utility $u_i^{\mathrm{RN}}(\mu) \equiv u_i(\mu)/u_i'\left(\Upsilon_i^{\langle\omega\rangle}\right)$. It turns out that the observer *can* uniquely assess agent i's risk-neutral probabilities. In fact, all standard elicitation procedures designed to reveal agent i's beliefs based on monetary incentives (de Finetti, 1974; Winkler & Murphy, 1968)—for example, querying the prices at which the agent would buy or sell various lottery tickets—essentially reveal $\mathrm{Pr}_i^{\mathrm{RN}}$, and *not* Pr_i (Kadane & Winkler, 1988). The agent's *observable* beliefs are in effect its risk neutral probabilities, not its true probabilities.

2.3 Opinion Pools

A variety of authors have proposed or advocated a corresponding variety of aggregation functions of the form (1); Genest and Zidek (1986) and French (1985) provide comprehensive surveys. Three approaches are generally distinguished. The first assumes a single Bayesian decision maker h (real or fictitious, within or outside the group), called the *supra Bayesian*, with a joint distribution over all events and all participants' beliefs. The supra Bayesian updates its beliefs via Bayes's rule, given the "evidence" of everyone else's beliefs. The resulting posterior is taken to be the consensus belief.

The second approach is to apply a prespecified function, usually some form of weighted average, that maps any set of probability distributions to a singleton. Note that some such functions can be interpreted as the updating procedure of a supra Bayesian. Each pooling function is usually justified axiomatically, by assuming a "reasonable" set of properties of the aggregate distribution. The two most common and well-studied aggregation functions are the linear and logarithmic opinion pools (LinOP, LogOP). The LinOP is a weighted arithmetic mean of the members' probabilities,

$$\mathrm{Pr}_0(\omega) = \sum_{i=1}^{N} \alpha_i \mathrm{Pr}_i(\omega), \tag{4}$$

and the LogOP is a normalized, weighted geometric mean,

$$\mathrm{Pr}_0(\omega) \propto \prod_{i=1}^{N} [\mathrm{Pr}_i(\omega)]^{\alpha_i}, \tag{5}$$

where the α_i are called *expert weights*, usually nonnegative numbers that sum to one. The LinOP and LogOP can actually be characterized as two instances of a parameterized family of weighted aggregation functions (Cooke, 1991). A third approach to pooling opinions is based on maximum entropy inference. The consensus is the unique probability distribution that maximizes

Shannon entropy, chosen from among the distributions that are consistent with all available information, including the experts' beliefs, their past performance, and/or dependencies among experts (Levy & Delic, 1994; Myung, Ramamoorti, & Bailey, 1996).

The debate over which aggregation method is best continues to rage (Benediktsson & Swain, 1992; Cooke, 1991; Jacobs, 1995; Ng & Abramson, 1992; Winkler, 1986). Several authors (most emphatically Lindley (1985, 1988)) argue that the supra Bayesian approach is superior, as it is grounded in standard normative Bayesian theory (Clemen & Winkler, 1993; Morris, 1974, 1977; Rosenblueth & Ordaz, 1992; West & Crosse, 1992; Winkler, 1981).

Attempts to justify more symmetric opinion pools often proceed by posing axioms on the combination function, and arguing that they represent desirable properties (Dalkey, 1975; Genest, 1984c, 1984b, 1984a; Genest & Zidek, 1986; Genest, McConway, & Schervish, 1986; Genest & Wagner, 1987; Wagner, 1984). Many of these properties seem reasonable, but disagreement persists on which are essential.[1] Researchers have proved that certain pooling formulae are implied by certain sets of properties. We begin with two seemingly incontrovertible assumptions.

Property 1 (Unanimity (UNAM)) *If* $\Pr_h(\omega) = \Pr_i(\omega)$ *for all agents* h *and* i*, and for all states* $\omega \in \Omega$*, then* $\Pr_0(\omega) = \Pr_1(\omega)$*.*

Property 2 (Nondictatorship (ND)) *There is no single agent* i *such that* $\Pr_0(\omega) = \Pr_i(\omega)$ *for all* $\omega \in \Omega$*, and regardless of the agents' beliefs.*

UNAM states that if everyone's assessments are in complete agreement, then the consensus agrees as well. ND simply ensures that what is inherently a multiagent problem is not reduced to the single-agent case.

Property 3 (Marginalization property (MP)) *Let* E *be an* arbitrary *event, that is, any subset of* Ω*. Then,*

$$f(\Pr_1, \Pr_2, \ldots, \Pr_n)(E) = \\ f(\Pr_1(E), \Pr_2(E), \ldots, \Pr_n(E)).$$

[1] For example, Lindley (1985) regards the so-called *marginalization property* as an "adhockery" while Cooke (1991) characterizes any consensus function that does *not* respect it as "downright queer".

Property 4 (External Bayesianity (EB)) *Let E and F be arbitrary events. Then,*

$$f(\mathrm{Pr}_1, \mathrm{Pr}_2, \ldots, \mathrm{Pr}_n)(E|F) = \\ f(\mathrm{Pr}_1|F, \mathrm{Pr}_2|F, \ldots, \mathrm{Pr}_n|F)(E).$$

MP and EB require consistency for probabilistic operations performed before and after pooling. MP states that we obtain the same probability for an event E whether we pool the opinions first, and then compute $\mathrm{Pr}_0(E) = \sum_{\omega \in E} \mathrm{Pr}_0(\omega)$, or if we first compute $\mathrm{Pr}_i(E) = \sum_{\omega \in E} \mathrm{Pr}_i(\omega)$ for each agent i, and then pool their opinions only over E. Similarly, EB holds that we obtain the same $\mathrm{Pr}_0(E|F)$ whether we combine opinions first and condition on F second, or condition on F first and combine opinions second. It has been shown that any f satisfying both MP and UNAM is a LinOP (Genest, 1984c), and any satisfying EB and UNAM is a LogOP (Genest, 1984a). Genest (1984b) also shows that f cannot simultaneously satisfy MP, EB, UNAM, and ND.

Property 5 (Proportional dependence on states (PDS))

$$\mathrm{Pr}_0(\omega) \propto f(\mathrm{Pr}_1(\omega), \mathrm{Pr}_2(\omega), \ldots, \mathrm{Pr}_n(\omega)).$$

PDS is sometimes called *independence of irrelevant states*, or termed a *likelihood principle*. It assures that the consensus likelihood ratio between two states does not depend on the agents' assessments of any other "irrelevant" state. The LinOP, LogOP, and most other proposed opinion pools satisfy PDS.

Property 6 (Independence preservation property (IPP)) *Let E and F be* arbitrary *events. If* $\mathrm{Pr}_i(E|F) = \mathrm{Pr}_i(E)$ *for all agents i, then* $\mathrm{Pr}_0(E|F) = \mathrm{Pr}_0(E)$.

IPP requires that *all* unanimously held independencies are preserved in the consensus. Advocates of IPP reason that identifying the independencies in a model is central to understanding the underlying phenomena, and that complete agreement on this dimension should be embraced. On the other hand, Genest and Wagner (1987) make a compelling case *against* the use of IPP by proving that *no* aggregation function whatsoever can satisfy it along with PDS and ND, when $|\Omega| \geq 5$.

Another point of contention is how best to determine the expert weights. In most cases they are chosen in an ad hoc manner to encode some measure of confidence, reliability, or importance (Benediktsson & Swain, 1992; French, 1985; Winkler, 1968). Some more formal methods to derive weights have been proposed, by making assumptions concerning the form of, or interdependence among, participants' beliefs (Cooke, 1991; Degroot & Mortera, 1991; Jacobs, 1995; Morris, 1977), or through iterative self-weighting procedures (Degroot, 1974). Little work, in any of the traditional categories of opinion pools, explicitly addresses truth incentives for reporting either self-assessed weights or the probabilities themselves.

2.4 Securities Markets for Allocating Risk

Under uncertainty, risk-averse agents will desire to *hedge* or *insure* against their risks by distributing wealth across states. For example, insuring the delivery of a package effectively transfers wealth from the *package-received* state to the *package-lost* state. The *Arrow-Debreu securities market* is the fundamental theoretical framework in economics and finance for resource allocation under uncertainty (Arrow, 1964; Dreze, 1987; Mas-Colell, Whinston, & Green, 1995). A *security*, denominated in money or other exchangeable good, pays off variously contingent upon the realization of an uncertain state. Let $\langle A \rangle$ denote a security that pays off one dollar if and only if the event A occurs. If the price of this security is $p^{\langle A \rangle}$ per unit, then agent i's decision to purchase $x_i^{\langle A \rangle}$ units is equivalent to accepting a lottery with payoff $(1 - p^{\langle A \rangle})x_i^{\langle A \rangle}$ if A occurs, and $-p^{\langle A \rangle}x_i^{\langle A \rangle}$ otherwise. Positive $x_i^{\langle A \rangle}$ indicates a quantity to buy, and negative $x_i^{\langle A \rangle}$ a quantity to sell.

In a market of S such securities, let $\mathbf{p} = \langle p^{\langle 1 \rangle}, p^{\langle 2 \rangle}, \ldots, p^{\langle S \rangle} \rangle$ denote the securities' prices, and $\mathbf{x}_i = \langle x_i^{\langle 1 \rangle}, x_i^{\langle 2 \rangle}, \ldots, x_i^{\langle S \rangle} \rangle$ denote the quantities of the securities held by agent i. Agent i's utility for securities is its expected utility for money (2), where the agent's choice of actions is how much to buy or sell of each security.

Agents trade securities with each other prior to revelation of the world state. In an economy of N agents, each continually maximizing (2), prices adjust until all buy orders match with sell orders for all securities. A market is in *competitive equilibrium* at prices $\mathbf{p}$ if and only if

$$\sum_{i=1}^{N} \mathbf{x}_i(\mathbf{p}) = \mathbf{0}, \tag{6}$$

where $\mathbf{x}_i(\mathbf{p})$ is agent i's optimal demand vector at prices $\mathbf{p}$.

A securities market is termed *complete* if it contains at least $|\Omega| - 1$ *linearly independent* securities. Such a market guarantees, under classical assumptions, that equilibrium entails a *Pareto optimal*, or *efficient*, allocation of risk.

A *conditional security* $\langle A_1 | A_2 \rangle$ pays off *contingent* on A_1 and *conditional* on A_2. That is, if A_2 occurs, then it pays out exactly as $\langle A_1 \rangle$; on the other hand, if $\bar{A}_2$ occurs, then the bet is called off and any price paid for the security is refunded (de Finetti, 1974). The canonical complete market consists of one security paying out in each state of nature. In general, though, any set of securities (possibly including conditionals) with a payoff-by-state matrix of rank $|\Omega| - 1$ is complete.

When one unit of each security pays out one dollar, the equilibrium prices in a securities market form a coherent probability distribution. For example, $p^{\langle A_1 \rangle} = p^{\langle A_1 A_2 \rangle} + p^{\langle A_1 \bar{A}_2 \rangle}$, or $p^{\langle A_1 A_2 \rangle} = p^{\langle A_1 | A_2 \rangle} p^{\langle A_2 \rangle}$. In fact, the equilibrium prices coincide with the agents' risk-neutral probabilities (3) for the available securities, which must be in complete agreement (Dreze, 1987; Nau & McCardle, 1991). Derived formally in Section 4.1, we simply sketch the intuition here. Since a risk-neutral agent buys $\langle A_j \rangle$ if $p^{\langle A_j \rangle} < \Pr_i(A_j)$ (it simply maximizes expected payoff), then *any* agent buys $\langle A_j \rangle$ if $p^{\langle A_j \rangle} < \Pr_i^{\mathrm{RN}}(A_j)$. Similarly, the agent sells if $p^{\langle A_j \rangle} > \Pr_i^{\mathrm{RN}}(A_j)$. If two agents h and i have differing risk neutral probabilities—that is, $\Pr_h^{\mathrm{RN}}(A_j) \neq \Pr_i^{\mathrm{RN}}(A_j)$—then there is an intermediate price at which they are both willing to trade. It follows that, at equilibrium, when by definition opportunities for exchange have been exhausted, all agents' risk neutral probabilities agree across available securities. Furthermore, since offers to buy and sell must match, the equilibrium prices equal these consensus probabilities.

There are two, largely inseparable, reasons for agents to trade in securities: to insure against risk ("hedge") and to profit from perceived mispricings ("speculate"). The more averse to risk, the more the former consideration dominates an agent's decision making. On the other hand, risk-neutrality—the limit of diminishing risk aversion—is synonymous with pure speculation.

These two behaviors are aligned with the two central roles of securities markets in the theory of economics under uncertainty. The first, as mentioned, is to support the reallocation of risk. The second is to *aggregate* and *disseminate* information. Agents that disagree on the likelihood of states may seek to exchange securities at prices that yield, according to each's subjective viewpoint, an increase in expected returns. Moreover, each agent is privy, albeit implicitly, to the evidence gathered by other agents (perhaps at great cost) via fluctuations in price.

2.5 Graphical Models

A joint probability distribution can often be represented more compactly as a graphical model (Darroch, Lauritzen, & Speed, 1980). Conciseness is achieved by exploiting conditional independence among the primary events. Let $\mathrm{CI}[A_j, W, X]$ be shorthand for $\Pr(A_j|WX) = \Pr(A_j|W)$, indicating that A_j is conditionally independent of the set of events X, given another set W. Consider the event $A_k \in Z$, with predecessors $\mathbf{pred}(A_j) \equiv \{A_1, A_2, \ldots, A_{k-1}\}$. Suppose that, given the outcomes of a subset $\mathbf{pa}(A_k) \subseteq \mathbf{pred}(A_k)$ of its predecessors—called A_k's *parents*—the event A_k is conditionally independent of all other preceding events, or $\mathrm{CI}[A_k, \mathbf{pa}(A_k), \mathbf{pred}(A_k) - \mathbf{pa}(A_k)]$. This structure can be depicted graphically as a *directed acyclic graph* (DAG): each event is a node in the graph, and there is a directed edge from node A_j to node A_k if and only if A_j is a parent of A_k. We also refer to A_k as the *child* of A_j. A DAG has no directed cycles and thus defines a partial order over its vertices. We assume without loss of generality that the event indices are consistent with this partial ordering; in other words, if A_j is a predecessor of A_k then $j < k$. We can write the joint probability distribution in a (usually) more compact form:

$$\Pr(A_1 A_2 \cdots A_M) = \prod_{k=1}^{M} \Pr(A_k|\mathbf{pa}(A_k)).$$

For each event A_k, we record a *conditional probability table* (CPT), which contains probabilities $\Pr(A_k|\mathbf{pa}(A_k))$ for all possible combinations of outcomes of events in $\mathbf{pa}(A_k)$. Thus, it is possible to implicitly represent the full joint with $O\left(M \cdot 2^{\max\{q(k)\}}\right)$ probabilities, instead of $2^M - 1$, where $q(k) = |\mathbf{pa}(A_k)|$ is the number of parents of A_k.

A *Markov network* (MN) is another graphical language for modeling conditional independence and for implicitly describing joint distributions (Whittaker, 1990; Darroch et al., 1980). Events are again associated with nodes in a graph, and edges encode probabilistic dependencies. The underlying structure of a MN is an *undirected* graph. Given the outcomes of its direct neighbors, an event A_j is conditionally independent of *every* other event in the network, not just preceding events. The neighbors of an event form a *Markov blanket* around it, "shielding" it from direct influence from the rest of the events (Pearl, 1988).

A *Markov independence* is a special type of conditional independence (Darroch et al., 1980; Pearl, 1988; Whittaker, 1990). The node A_j and the set of nodes $X \subseteq Z - A_j$ are Markov independent, given another set $W \subseteq Z - X - A_j$, if $\text{CI}[A_j, W, X]$ and $A_j \cup W \cup X = Z$. Recall that Z is the set of all modeled events.

The Markov blanket of a node in a BN consists of its direct parents, its direct children, and its children's direct parents (Pearl, 1988). Therefore a BN can be converted into a MN by *moralizing* the network, or fully connecting ("marrying") each node's parents, and dropping edge directionality (Lauritzen & Spiegelhalter, 1988; Neapolitan, 1990). A MN can be converted into a BN by *filling in* or *triangulating* (Kloks, 1994) the graph, and adding directionality according to the fill-in ordering (Jensen, 1996; Lauritzen & Spiegelhalter, 1988; Neapolitan, 1990; Pearl, 1988). Both transformations are sound with respect to independence, but neither is complete.

A DAG is an *independency map*, or an *I-map*, of a probability distribution Pr if every independency implicit in the graph holds within Pr (Pearl, 1988). Note that a complete graph is a trivial I-map of any distribution over Ω.

A DAG is *decomposable* if there is an edge between every two nodes that share a common child (Chyu, 1991; Darroch et al., 1980; Pearl, 1988; Shachter, Andersen, & Poh, 1991). Trees are a subset of decomposable DAGs, since every node has at most one parent. Complete graphs are also decomposable since *every* two nodes are connected. Any BN can be made decomposable by reorienting some edges and introducing new edges where needed (Chyu, 1991; Shachter et al., 1991). A two step procedure of moralization plus fill-in (triangulation) will render a BN decomposable. Finding the smallest decomposable representation (finding the optimal fill-in ordering) is NP-hard, and even the smallest decomposable representation can be exponentially larger than the original BN. Still, the decomposable representation can be exponentially more compact than the full joint distribution. The independencies encoded in a decomposable BN are all Markov independencies (Pearl, 1988).

3 Graphical Models for Belief Aggregation

In this section, we address the problem of representing aggregate beliefs concisely. The implications of results in this section for securities markets will be examined in Section 4.

We presume that each agents' beliefs are given as a graphical model, and that the combined beliefs are to be represented as well with a graphical model. Two intuitively reasonable assumptions in this context, made *a priori* by other authors, are (1) if all agents agree on a single topology, then that structure should be maintained, and (2) probability aggregation can be isolated within each conditional probability table (CPT). Section 3.2 demonstrates that each of these properties leads to an impossibility theorem when combined with other reasonable, oft-invoked assumptions. On a more positive note, Section 3.3 shows that the logarithmic opinion pool (LogOP) maintains all agreed-upon Markov independencies, and describes procedures for constructing consensus Markov networks (MNs) and consensus Bayesian networks (BNs) that are consistent with the LogOP. Section 3.4 presents an algorithm that can, in some cases, compute the LogOP exponentially faster than the brute force approach. That section also characterizes the computation complexity of the linear opinion pool (LinOP) and the LogOP. Section 3.1 defines the properties necessary to state the various impossibility and possibility results appearing later in the section.

3.1 Property Definitions

Recall the independence preservation property (IPP), defined in Section 2.3. For an aggregation function f to satisfy IPP, any independencies that are agreed-upon by all agents must be maintained within the consensus distribution. Genest and Wagner (1987) prove that *no* aggregation function can simultaneously satisfy IPP, proportional dependence on states (PDS) (Property 5), and nondictatorship (ND) (Property 2). But one might argue that IPP is overly strong. It requires preservation of, for example, a unanimous independence between the events $E = A_3\bar{A}_7$ and $F = \bar{A}_2 A_4 \vee A_7$. This kind of independence seems of little descriptive value to a modeler, and indeed cannot be represented with a BN. One may be willing to forgo preserving *all* independencies, being content to preserve independencies among the *primary* events, $A_1, A_2, \ldots, A_M$. With this in mind, we define a weaker independence property.

Property 7 (Event independence preservation property (EIPP)) *If* $\mathrm{Pr}_i(A_j|A_k) = \mathrm{Pr}_i(A_j)$ *for all agents* i*, then* $\mathrm{Pr}_0(A_j|A_k) = \mathrm{Pr}_0(A_j)$.

Note that, when $|\Omega| = 4$, the conditions IPP and EIPP are essentially equivalent. In this situation, the only way for two events to be independent is if each consists of exactly two atomic states, and if they overlap at exactly one state (Genest & Wagner, 1987).

In Section 3.2, we see that substituting EIPP for IPP does admit a possibility that is consistent with both PDS and ND, though not a very satisfactory one. In search of a nontrivial possibility, we define two even weaker independence conditions.

Property 8 (Markov event independence preservation property (MEIPP)) *If* $\mathrm{Pr}_i(A_j|WA_k) = \mathrm{Pr}_i(A_j|W)$ *for all agents* i *and for all* $W \subseteq Z$ *(including* $W = \emptyset$*), then* $\mathrm{Pr}_0(A_j|A_k) = \mathrm{Pr}_0(A_j)$.

Property 9 (Non-Markov event independence preservation property (NMEIPP)) *If* $\mathrm{Pr}_i(A_j|A_k) = \mathrm{Pr}_i(A_j)$ *for all agents* i*, and* $\mathrm{Pr}_h(A_j|WA_k) \neq \mathrm{Pr}_h(A_j|W)$*, for some agent* h *and some* $W \subseteq Z$*, then* $\mathrm{Pr}_0(A_j|A_k) = \mathrm{Pr}_0(A_j)$.

These two properties are purposely constructed so that EIPP $\Leftrightarrow$ (MEIPP $\wedge$ NMEIPP). We see in Section 3.2 that the source of the impossibility lies entirely within the latter. Finally, we define a stronger version of the MEIPP.

Property 10 (Markov independence preservation property (MIPP)) *Let* $W, X \subseteq Z - A_j$ *be disjoint sets of events such that* $A_j \cup W \cup X = Z$*. If* $\mathrm{Pr}_i(A_j|WX) = \mathrm{Pr}_i(A_j|W)$ *for all agents* i*, then* $\mathrm{Pr}_0(A_j|WX) = \mathrm{Pr}_0(A_j|W)$.

The relative strengths of these various independence conditions can be summarized as follows:

$$\text{IPP} \Rightarrow \text{EIPP} \Leftrightarrow (\text{MEIPP} \wedge \text{NMEIPP})$$
$$\text{MIPP} \Rightarrow \text{MEIPP}$$

Finally, we define a property that captures what seems to be a natural assumption within the context of graphical models, advocated independently by other authors (Matzkevich & Abramson, 1992). We say that an aggregator satisfies the *family aggregation* (FA) property if it operates locally, within each conditional probability table (CPT) of the consensus structure.

Property 11 (Family aggregation (FA))

$$\begin{aligned}&\mathrm{Pr}_0(A_j|\mathbf{pa}(A_j)) = \\ &\quad f(\mathrm{Pr}_1(A_j|\mathbf{pa}(A_j)), \ldots, \mathrm{Pr}_N(A_j|\mathbf{pa}(A_j))).\end{aligned}$$

In Sections 3.2.1 and 3.2.2, we consider the implications of the properties EIPP and FA, respectively.

3.2 Combining Bayesian Networks: Examples and Impossibility

3.2.1 Event Independence Preservation

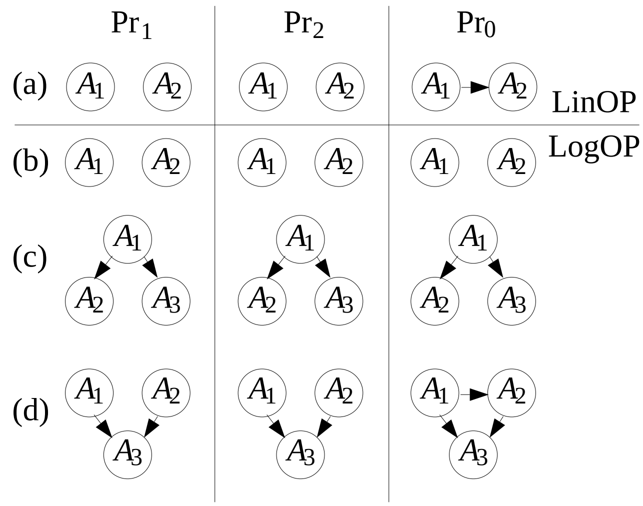
We begin with an example the build the underlying intuition.

Example 1 (EIPP and the LinOP)

Suppose that two agents agree that two primary events, A_1 and A_2, are independent, as pictured in Figure 1(a), but disagree on the associated marginal probabilities.

For concreteness, let the first agent hold beliefs $\mathrm{Pr}_1(A_1) = \mathrm{Pr}_1(A_2) = 0.5$, and the second $\mathrm{Pr}_2(A_1) = 0.8$ and $\mathrm{Pr}_2(A_2) = 0.6$. Thus,

$$\begin{array}{ll}\mathrm{Pr}_1(A_1A_2) = 0.25 & \mathrm{Pr}_2(A_1A_2) = 0.48 \\ \mathrm{Pr}_1(A_1\bar{A}_2) = 0.25 & \mathrm{Pr}_2(A_1\bar{A}_2) = 0.32 \\ \mathrm{Pr}_1(\bar{A}_1A_2) = 0.25 & \mathrm{Pr}_2(\bar{A}_1A_2) = 0.12 \\ \mathrm{Pr}_1(\bar{A}_1\bar{A}_2) = 0.25 & \mathrm{Pr}_2(\bar{A}_1\bar{A}_2) = 0.08.\end{array}$$



1

Figure 1: Independence preservation behavior of (a) LinOP and (b)–(d) LogOP. If two agents' beliefs Pr_1 and Pr_2 have the dependency structures shown, then the consensus Pr_0 will in general have the dependency structure depicted in column three.

Now if we apply the LinOP (4) with, say, equal weights of $w_1 = w_2 = 0.5$, we get:

$$\begin{aligned}
\mathrm{Pr}_0(A_1A_2) &= 0.365 \\
\mathrm{Pr}_0(A_1\bar{A}_2) &= 0.41 \\
\mathrm{Pr}_0(\bar{A}_1A_2) &= 0.185 \\
\mathrm{Pr}_0(\bar{A}_1\bar{A}_2) &= 0.165.
\end{aligned}$$

In particular, $\mathrm{Pr}_0(A_1)\,\mathrm{Pr}_0(A_2) \neq \mathrm{Pr}_0(A_1A_2)$, and so the two events are *not* independent in the consensus.[2] Even though the precondition of the EIPP is met, the postcondition is not: a BN representation of the derived consensus would have to include an edge between A_1 and A_2. □

Example 2 (EIPP and the LogOP)

[2]As early as Yule (1903) it was recognized that averaging two distributions may mask a commonly held independence.

Suppose that two agents' beliefs over two primary events are as described in Example 1. If we apply the LogOP with equal weights, we get:

$$\begin{aligned}
\Pr_0(A_1A_2) &= 0.367007 \\
\Pr_0(A_1\bar{A}_2) &= 0.29966 \\
\Pr_0(\bar{A}_1A_2) &= 0.183503 \\
\Pr_0(\bar{A}_1\bar{A}_2) &= 0.14983.
\end{aligned}$$

In this case, $\Pr_0(A_1)\Pr_0(A_2) = \Pr_0(A_1A_2)$, and the two events remain independent, as shown in Figure 1(b). This is not a numerical coincidence; in fact, independence between only two events is always maintained by the LogOP (Genest & Wagner, 1987). Now suppose that among three primary events, both agents agree that A_3 is independent of A_2 given A_1. That is, both agents agree that dependencies conform to a tree structure, with A_1 the parent of both A_2 and A_3, as depicted in Figure 1(c). Then once again, the LogOP will maintain this structure. One might conjecture that the LogOP maintains all BN structures, but this is not the case. For example, suppose that, among three primary events, the two agents agree that A_1 and A_2 are mutually independent, and that A_3 depends on both A_1 and A_2. That is, both agents agree on the polytree structure in Figure 1(d). In this case, when we compute the consensus with the LogOP, A_1 and A_2 will in general become mutually dependent, the EIPP is not satisfied, and a consensus BN will require an arc between the two nodes. □

Having seen that both the LinOP and the LogOP violate the EIPP, we seek a more general characterization of the class of functions that do obey it. We begin by showing that Lemma 3.2 in (Genest & Wagner, 1987), originally proved with respect to the IPP, is also applicable under the weaker EIPP.

Lemma 1 (Adapted from (Genest & Wagner, 1987)) *If f obeys EIPP and PDS, then there exist constants $\alpha_1, \alpha_2, \ldots, \alpha_N$, and c such that*

$$\Pr_0(\omega_j) = \sum_{i=1}^{N} \alpha_i \Pr_i(\omega_j) + c. \tag{7}$$

Proof. Consider three events A_1, A_2, and A_3, with agents' beliefs described as follows:

$$\Pr_i(A_1A_2A_3) = \Pr_i(A_1A_2\bar{A}_3) \quad = \quad \frac{(1-z_i)^2}{4(1+z_i)}$$

$$\begin{aligned}
\Pr_i(A_1\bar{A}_2A_3) = \Pr_i(A_1\bar{A}_2\bar{A}_3) &= \frac{1-z_i}{4} \\
\Pr_i(\bar{A}_1\bar{A}_2A_3) &= x_i \\
\Pr_i(\bar{A}_1\bar{A}_2\bar{A}_3) &= y_i, \qquad (8)
\end{aligned}$$

where $z_i = x_i + y_i$ for all i. In this case, all agents agree that A_1 and A_2 are independent and, as long as $z_i < 1$, these equations describe a legal probability distribution. Since f obeys PDS, there must be some function g such that,

$$\Pr_0(\bar{A}_1\bar{A}_2A_3) = \frac{g(x_1, x_2, \ldots, x_N)}{\sum_{k=1}^{8} g(\Pr_1(\omega_k), \ldots, \Pr_N(\omega_k))}$$

and similarly for $\Pr_0(\bar{A}_1\bar{A}_2\bar{A}_3)$. Now imagine a second situation exactly as in (8), except with $\Pr_i(\bar{A}_1\bar{A}_2A_3) = x'_i$ and $\Pr_i(\bar{A}_1\bar{A}_2\bar{A}_3) = y'_i$. Genest and Wagner show that, as long as $x_i + y_i = x'_i + y'_i < 1$, then

$$\begin{aligned}
& g(x_1, x_2, \ldots, x_N) + g(y_1, y_2, \ldots, y_N) \\
&= g(x'_1, x'_2, \ldots, x'_N) + g(y'_1, y'_2, \ldots, y'_N). \qquad (9)
\end{aligned}$$

From here, they show that *since x_i and y_i can be chosen arbitrarily* (as long as their sum is less than one), then f must have the form specified. □

Genest and Wagner go on to show, without further assumption, that f must be a dictatorship. However, that proof does *not* carry through under the weaker condition EIPP. This can be seen via a simple counterexample. Let f always ignore the agents' opinions, and simply assign a uniform distribution over all $\omega \in \Omega$. In this case, the consensus distribution holds that *all* primary events A_j are independent, and thus any agreed upon independencies are trivially maintained. One might wonder whether EIPP admits any other, more appealing, aggregation functions. The following proposition essentially establishes that it does not.

Proposition 2 *No aggregation function f can simultaneously satisfy EIPP, PDS, UNAM, and ND.*

Proof. With the addition of UNAM, it is clear that c must be zero in (7), and thus f must have the form of a standard LinOP (4). From Example 1, we know that the LinOP does not maintain independence even between just two events. The fact that the LinOP cannot satisfy both IPP and ND is proved formally by several authors (Genest, 1984c; Lehrer & Wagner, 1983;

Wagner, 1984). Their proofs are applicable to EIPP as well, since they hold even when $|\Omega| = 4$, in which case EIPP and IPP coincide. $\Box$

A careful examination of the proof of Lemma 1 also suggests one more possibility when the full generality of IPP is relaxed. Suppose that all agents agree that *all three* events, A_1, A_2, and A_3, are completely independent. Then it can be shown that $\Pr_i(\bar{A}_1 A_2 A_3) = z_i/(1+z_i) + y_i$ and, furthermore, that $x_i = y_i$ for all i. In this case, (9) holds only vacuously, since $x_i' = x_i$ and $y_i' = y_i$. Moreover, since x_i and y_i are no longer arbitrary, the proof does not go through. Thus, under this fully independent condition, the conclusion of Lemma 1 is no longer valid.

This insight leads us to characterize the inherent impossibility more sharply, by dividing EIPP into two, weaker conditions, NMEIPP and MEIPP, and showing that the former retains the impossibility while the latter does not.

Corollary 3 *No aggregation function f can simultaneously satisfy NMEIPP, PDS, UNAM, and ND.*

Proof. The proof of Lemma 1 still follows under NMEIPP, and thus so does the proof of Proposition 2. $\Box$

Section 3.3 demonstrates that in fact, MEIPP is perfectly consistent with PDS, UNAM, and ND in a nontrivial way. Indeed, the stronger MIPP is consistent as well.

3.2.2 Family Aggregation

Example 3 (Family aggregation)

Consider two agents, each with a BN consisting of two primary events, with A_1 the parent of A_2 and with beliefs as follows:

$$\begin{aligned} \Pr_1(A_1) &= 0.2 \qquad & \Pr_2(A_1) &= 0.8 \\ \Pr_1(A_2|A_1) &= 0.4 \qquad & \Pr_2(A_2|A_1) &= 0.8 \\ \Pr_1(A_2|\bar{A}_1) &= 0.6 \qquad & \Pr_2(A_2|\bar{A}_1) &= 0.3 \end{aligned}$$

We compute each consensus CPT as an average of the corresponding individual CPTs. That is, $\Pr_0(A_1) = (.2+.8)/2 = .5$, $\Pr_0(A_2|A_1) = (.4+.8)/2 = .6$, etc. This results in the following consensus joint distribution:

$$\begin{aligned}
\Pr_0(A_1 A_2) &= 0.3 \\
\Pr_0(A_1 \bar{A}_2) &= 0.2 \\
\Pr_0(\bar{A}_1 A_2) &= 0.225 \\
\Pr_0(\bar{A}_1 \bar{A}_2) &= 0.275.
\end{aligned}$$

Next suppose that both agents reverse their edge between the two events, such that A_2 is the parent of A_1, but that their joint distributions remain unchanged. Now the agents' CPTs are:

$$\begin{aligned}
\Pr_1(A_2) &= 0.56 & \Pr_2(A_2) &= 0.7 \\
\Pr_1(A_1|A_2) &= 0.142857 & \Pr_2(A_1|A_2) &= 0.914286 \\
\Pr_1(A_1|\bar{A}_2) &= 0.272727 & \Pr_2(A_1|\bar{A}_2) &= 0.533333
\end{aligned}$$

and if we average locally within each CPT, we get a different consensus distribution:

$$\begin{aligned}
\Pr_0(A_1 A_2) &= 0.333 \\
\Pr_0(A_1 \bar{A}_2) &= 0.149121 \\
\Pr_0(\bar{A}_1 A_2) &= 0.297 \\
\Pr_0(\bar{A}_1 \bar{A}_2) &= 0.220878.
\end{aligned}$$

Thus averaging only within each family of the BN violates the form of the opinion pool itself (1), which insists that the consensus joint distribution depend only on the underlying joint distributions of the agents involved. □

We now show that this inconsistency is not confined solely to the averaging aggregator.

Proposition 4 *No aggregation function f can simultaneously satisfy FA, UNAM, and ND.*

Proof. Let the first event in the consensus BN be A_{j_1} the second A_{j_2}, ..., and the last A_{j_M}. The FA property requires both of the following:

$$\begin{aligned}
&\Pr_0(A_{j_1}) \\
&\quad = \quad f(\Pr_1(A_{j_1}), \Pr_2(A_{j_1}), \ldots, \Pr_N(A_{j_1})) \qquad (10) \\
&\Pr_0(A_{j_M}|Z - A_{j_M}) \\
&\quad = \quad f(\Pr_1(A_{j_M}|Z - A_{j_M}), \ldots, \Pr_N(A_{j_M}|Z - A_{j_M})). \qquad (11)
\end{aligned}$$

By the definition of an opinion pool (1), the consensus belief depends only on the agents' underlying joint distributions, and not on the particular ordering of events in each BN. Thus, we must arrive at the same consensus distribution as long as $\{j_1, j_2, \ldots, j_M\}$ is some permutation of $\{1, 2, \ldots, M\}$. Consider two permutations, one where $j_1 = 1$ and one where $j_M = 1$. Then (10) and (11) become:

$$\begin{aligned} &\Pr_0(A_1) \\ &\quad = \quad f(\Pr_1(A_1), \Pr_2(A_1), \ldots, \Pr_N(A_1)) \qquad (12) \\ &\Pr_0(A_1|Z - A_1) \\ &\quad = \quad f(\Pr_1(A_1|Z - A_1), \ldots, \Pr_N(A_1|Z - A_1)). \qquad (13) \end{aligned}$$

Dalkey (1975) proves that no function can simultaneously satisfy (12), (13), UNAM, and ND. Alternatively, the two equations essentially require that f satisfy both MP and EB, defined in Section 2.3, which Genest (1984b) shows are incompatible with UNAM and ND. □

3.3 The LogOP and Consensus Markov Networks

The results in Section 3.2 suggest that insisting upon general event independence preservation has rather severe consequences. In this section, we see that preserving *Markov* independencies is in fact compatible with PDS, UNAM, and ND. Let A_j be a primary event, and $W \subseteq Z - A_j$ and $X = Z - W - A_j$ be sets of events. Then A_j is Markov independent of X given W if $\Pr(A_j|WX) = \Pr(A_j|W)$.

Proposition 5 *The LogOP satisfies MIPP.*

Proof. Since the LogOP is defined in terms of atomic states ω, we make use of the following two identities:

$$\Pr_0(A|WX) \equiv \frac{\Pr_0(AWX)}{\Pr_0(AWX)+\Pr_0(\bar{A}WX)}$$

$$\Pr_0(A|W) \equiv \frac{\sum_X \Pr_0(AWX)}{\sum_X \Pr_0(AWX)+\sum_X \Pr_0(\bar{A}WX)}$$

where $\sum_X$ represents a sum over all possible combinations of outcomes of events in the set X. Then we have that,

$$
\begin{aligned}
&\Pr_0(A|WX) = \frac{\prod_{i=1}^{N}[\Pr_i(AWX)]^{\alpha_i}}{\prod_{i=1}^{N}[\Pr_i(AWX)]^{\alpha_i}+\prod_{i=1}^{N}[\Pr_i(\bar{A}WX)]^{\alpha_i}}\\
&= \frac{\prod\left[\frac{\Pr_i(AW)\Pr_i(WX)}{\Pr_i(W)}\right]^{\alpha_i}}{\prod\left[\frac{\Pr_i(AW)\Pr_i(WX)}{\Pr_i(W)}\right]^{\alpha_i}+\prod\left[\frac{\Pr_i(\bar{A}W)\Pr_i(WX)}{\Pr_i(W)}\right]^{\alpha_i}}\\
&= \frac{\prod[\Pr_i(AW)]^{\alpha_i}}{\prod[\Pr_i(AW)]^{\alpha_i}+\prod[\Pr_i(\bar{A}W)]^{\alpha_i}}\\
&= \frac{\prod[\Pr_i(AW)]^{\alpha_i}}{\prod[\Pr_i(AW)]^{\alpha_i}+\prod[\Pr_i(\bar{A}W)]^{\alpha_i}}\cdot\frac{\sum_X\prod[\Pr_i(WX)]^{\alpha_i}}{\sum_X\prod[\Pr_i(WX)]^{\alpha_i}}\\
&= \frac{\sum_X\prod[\Pr_i(AW)\Pr_i(WX)]^{\alpha_i}}{\sum_X\prod[\Pr_i(AW)\Pr_i(WX)]^{\alpha_i}+\sum_X\prod[\Pr_i(\bar{A}W)\Pr_i(WX)]^{\alpha_i}}\\
&= \frac{\sum_X\prod\left[\frac{\Pr_i(AW)\Pr_i(WX)}{\Pr_i(W)}\right]^{\alpha_i}}{\sum_X\prod\left[\frac{\Pr_i(AW)\Pr_i(WX)}{\Pr_i(W)}\right]^{\alpha_i}+\sum_X\prod\left[\frac{\Pr_i(\bar{A}W)\Pr_i(WX)}{\Pr_i(W)}\right]^{\alpha_i}}\\
&= \frac{\sum_X\prod[\Pr_i(AWX)]^{\alpha_i}}{\sum_X\prod[\Pr_i(AWX)]^{\alpha_i}+\sum_X\prod[\Pr_i(\bar{A}WX)]^{\alpha_i}}\\
&= \frac{\sum_X\Pr_0(AWX)}{\sum_X\Pr_0(AWX)+\sum_X\Pr_0(\bar{A}WX)}\\
&= \Pr_0(A|W)
\end{aligned}
$$

$\square$

Suppose that each agent's belief is given as a MN, and we wish to generate a consensus MN structure that can encode the results of the LogOP. As discussed in Section 2.5, graph connectivity in a MN represents probabilistic dependence, and the neighborhood relation represents direct influence. For each node A_j, the set of its neighbors plays the role of W in Proposition 5, and all other nodes constitute the set X. The proposition ensures that, if all agents agree on a common MN structure, then the consensus distribution derived by the LogOP will respect the same structure. When agents are not in complete agreement on the structure, then the consensus can be represented as a MN defined by the union of all the individual MNs. In other words, there is an edge between A_j and A_k in the consensus MN if and only if there is an edge between those two nodes in at least one of the agents' MNs.

Pearl (1988) gives axiomatic descriptions of both MNs and BNs. Only the former includes an axiom called *strong union*, which states that if $\Pr(A_j|A_k) = \Pr(A_j)$, then $\Pr(A_j|WA_k) = \Pr(A_j|W)$ for all $W \subseteq Z$. Notice that, if the precondition of the EIPP is met, and strong union holds for all agents, then the precondition of the MEIPP must also hold. This axiom is the key distinction that allows common MN structures to be maintained in the LogOP consensus, whereas common BN structures in general are not.

Given a collection of BNs, generating a consensus BN structure that is consistent with the LogOP is also relatively straightforward. We first convert each BN into a MN by moralizing the graphs, or fully connecting each node's parents and dropping edge directionality (Lauritzen & Spiegelhalter, 1988; Neapolitan, 1990). Next, we compute the union of the individual MNs, and finally we convert the resulting consensus MN back into a BN by filling in or triangulating the network, reintroducing directionality according to the fill-in order[3] (Jensen, 1996; Lauritzen & Spiegelhalter, 1988; Neapolitan, 1990; Pearl, 1988).

We have outlined how to derive consensus MN or BN *structures*; what of computing the associated probabilities? In Section 3.4, we give an algorithm for computing the probabilities in a consensus BN that is polynomial in the size of its CPTs. Note that, even when all agents agree on a BN structure, the size of the final representation may grow exponentially during fill-in, and computing the union of the intermediate MNs when agents disagree will only exacerbate this problem. Nevertheless, even a decomposable representation can be exponentially smaller than the full joint distribution, and the most popular algorithms for exact Bayesian inference do operate on decomposable models in practice.

3.4 Computing LogOP and LinOP

Since the LinOP (4) and LogOP (5) are defined over atomic states, computing, for example, the consensus marginal probability of a single event involves in the worst case a summation over 2^{M-1} terms. Moreover, even computing the LogOP consensus for a single state requires a normalization factor that is itself a sum over all 2^M states. In this section, we see that if each agent's belief is represented as a BN, the LinOP and LogOP consensus for any probabilistic query can be computed more efficiently. In particular, for the LogOP, we can compute the CPTs of a consensus BN with time complexity $O\left(NM^2 \cdot 2^{\max\{q(j)\}}\right)$, where $q(j)$ is the number of parents of A_j in the consensus structure.

[3] We do not claim that these consensus structures are minimal, or even that LogOP is the preferred aggregation method. My goal is more to guide a modeler's decision process by delineating what representations are consistent under what circumstances.

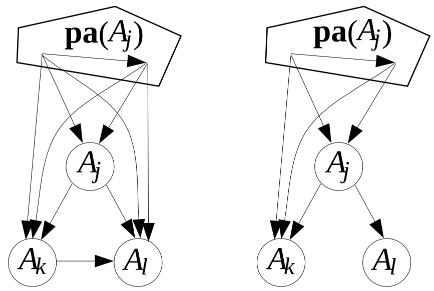


Figure 2: Two potential sections of a decomposable BN. A_j's children can be either in the same clique or in separate cliques.

We focus first on the task of generating a LogOP-consistent consensus BN. We compute its structure as described in Section 3.3. Consider computing the CPT at A_j, that is, $\Pr_0(A_j|\mathbf{pa}(A_j))$ for all combinations of outcomes of events in $\mathbf{pa}(A_j)$. From Proposition 4, we know that simply combining each agent's assessment of this conditional probability will not succeed in general. However, we *can* compute the *last* CPT, $\Pr_0(A_M|\mathbf{pa}(A_M))$, in terms of only the $\Pr_i(A_M|\mathbf{pa}(A_M))$, by computing the LogOP over the single event A_M:

$$\Pr_0(A_M|\mathbf{pa}(A_M)) = \frac{\prod_{i=1}^{N}\left[\Pr_i(A_M|\mathbf{pa}(A_M))\right]^{\alpha_i}}{\prod\left[\Pr_i(A_M|\mathbf{pa}(A_M))\right]^{\alpha_i} + \prod\left[\Pr_i(\bar{A_M}|\mathbf{pa}(A_M))\right]^{\alpha_i}}. \tag{14}$$

Because the LogOP satisfies EB, if we condition on *all* other events $Z - A_M$ in the network, then the LogOP over just A_M will return the same result as if we had computed the LogOP over all events, and then conditioned on $Z - A_M$. Equation 14 also reflects the fact that $\Pr_0(A_M|\mathbf{pa}(A_M)) = \Pr_0(A_M|Z - A_M)$ and $\Pr_i(A_M|\mathbf{pa}(A_M)) = \Pr_i(A_M|Z - A_M)$, by the semantics of the BNs.

We can compute the remainder of the CPTs in reverse index order. Assume that the CPTs $\Pr_0(A_k|\mathbf{pa}(A_k))$ have been calculated for all $k > j$, and that next we need to calculate $\Pr_0(A_j|\mathbf{pa}(A_j))$. To simplify the discussion, let A_j have exactly two children, A_k and A_l, with $j < k < l$; the analysis generalizes easily to more children (or one child). Since the BN is decomposable, its topology is a tree of cliques (Chyu, 1991; Pearl, 1988; Shachter et al., 1991), and A_k and A_l can either be in the same clique or in separate cliques, as depicted in Figure 2. Note that decomposability also ensures that A_j's neighbors, $A_l \cup A_k \cup \mathbf{pa}(A_j)$, constitute its Markov blanket. We can query each of the agent's BNs for the probabilities $\Pr_i(A_j|A_l \cup A_k \cup \mathbf{pa}(A_j))$

using a standard BN inference algorithm. From these, we can compute the corresponding consensus probability as a LogOP only over A_j, as before:

$$\begin{aligned}
&\mathrm{Pr}_0(A_j|A_l \cup A_k \cup \mathbf{pa}(A_j)) \\
&\quad \propto \prod_{i=1}^{N} \left[\mathrm{Pr}_i(A_j|A_l \cup A_k \cup \mathbf{pa}(A_j))\right]^{\alpha_i}.
\end{aligned} \tag{15}$$

We now need only eliminate the conditioning on A_l and A_k. By Bayes's rule, we have that

$$\begin{aligned}
&\frac{\mathrm{Pr}_0(A_j|A_l \cup A_k \cup \mathbf{pa}(A_j))}{\mathrm{Pr}_0(\bar{A}_j|A_l \cup A_k \cup \mathbf{pa}(A_j))} \\
&\quad = \frac{\mathrm{Pr}_0(A_l \cup A_k|A_j \cup \mathbf{pa}(A_j))}{\mathrm{Pr}_0(A_l \cup A_k|\bar{A}_j \cup \mathbf{pa}(A_j))} \cdot \frac{\mathrm{Pr}_0(A_j|\mathbf{pa}(A_j))}{\mathrm{Pr}_0(\bar{A}_j|\mathbf{pa}(A_j))} \\
&\quad = \frac{\mathrm{Pr}_0(A_l|A_k \cup A_j \cup \mathbf{pa}(A_j))}{\mathrm{Pr}_0(A_l|A_k \cup \bar{A}_j \cup \mathbf{pa}(A_j))} \cdot \frac{\mathrm{Pr}_0(A_k|A_j \cup \mathbf{pa}(A_j))}{\mathrm{Pr}_0(A_k|\bar{A}_j \cup \mathbf{pa}(A_j))} \\
&\qquad \cdot \frac{\mathrm{Pr}_0(A_j|\mathbf{pa}(A_j))}{\mathrm{Pr}_0(\bar{A}_j|\mathbf{pa}(A_j))}.
\end{aligned}$$

Because the BN is decomposable, and regardless of whether A_k and A_l are in the same or different cliques, $\mathrm{Pr}_0(A_l|A_k \cup A_j \cup \mathbf{pa}(A_j)) = \mathrm{Pr}_0(A_l|\mathbf{pa}(A_l))$ and $\mathrm{Pr}_0(A_k|A_j \cup \mathbf{pa}(A_j)) = \mathrm{Pr}_0(A_k|\mathbf{pa}(A_k))$, both of which have already been computed. Therefore we can calculate the CPT at A_j as follows:

$$\begin{aligned}
&\frac{\mathrm{Pr}_0(A_j|\mathbf{pa}(A_j))}{\mathrm{Pr}_0(\bar{A}_j|\mathbf{pa}(A_j))} \\
&\quad = \frac{\mathrm{Pr}_0(A_j|A_l \cup A_k \cup \mathbf{pa}(A_j))}{\mathrm{Pr}_0(\bar{A}_j|A_l \cup A_k \cup \mathbf{pa}(A_j))} \cdot \frac{\mathrm{Pr}_0(A_l|\tilde{\mathbf{pa}}(A_l))}{\mathrm{Pr}_0(A_l|\mathbf{pa}(A_l))} \\
&\qquad \cdot \frac{\mathrm{Pr}_0(A_k|\tilde{\mathbf{pa}}(A_k))}{\mathrm{Pr}_0(A_k|\mathbf{pa}(A_k))},
\end{aligned} \tag{16}$$

where $\tilde{\mathbf{pa}}(A_k)$ and $\tilde{\mathbf{pa}}(A_l)$ contain $\bar{A}_j$, and $\mathbf{pa}(A_k)$ and $\mathbf{pa}(A_l)$ contain A_j. Once we compute the likelihood ratio on the LHS of (16), the desired probabilities are uniquely determined, since $\mathrm{Pr}_0(A_j|\mathbf{pa}(A_j)) + \mathrm{Pr}_0(\bar{A}_j|\mathbf{pa}(A_j)) = 1$. The psuedocode for the full algorithm is given in Figure 3.

A consensus BN consistent with the LinOP would in general be fully connected, and thus not an object of particular value. However, if all agents' beliefs are given as BNs, we can retain their separation and still compute

LOGOP-CONSENSUS-BN$(\mathrm{Pr}_1, \mathrm{Pr}_2, \ldots, \mathrm{Pr}_N)$
INPUT: N Bayesian networks: $\mathrm{Pr}_1, \mathrm{Pr}_2, \ldots, \mathrm{Pr}_N$
OUTPUT: LogOP-consistent consensus BN: Pr_0

1. Structure of Pr_0 = TRIANGULATE $[\cup_{i=1}^N$ MORALIZE $[\mathrm{Pr}_i]]$
2. $\mathrm{Pr}_0(A_M|\mathbf{pa}(A_M)) \propto \prod_{i=1}^N [\mathrm{Pr}_i(A_M|\mathbf{pa}(A_M))]^{\alpha_i}$
3. **for** $j = M-1$ **downto** 1
4. $\quad\mathrm{Pr}_0(A_j|A_l \cup A_k \cup \mathbf{pa}(A_j)) \propto \prod_{i=1}^N [\mathrm{Pr}_i(A_j|A_l \cup A_k \cup \mathbf{pa}(A_j))]^{\alpha_i}$
5. $\quad\frac{\mathrm{Pr}_0(A_j|\mathbf{pa}(A_j))}{\mathrm{Pr}_0(A_j|\mathbf{pa}(A_j))} = \frac{\mathrm{Pr}_0(A_j|A_l \cup A_k \cup \mathbf{pa}(A_j))}{\mathrm{Pr}_0(A_j|A_l \cup A_k \cup \mathbf{pa}(A_j))} \cdot \frac{\mathrm{Pr}_0(A_l|\tilde{\mathbf{pa}}(A_l))}{\mathrm{Pr}_0(A_l|\mathbf{pa}(A_l))} \cdot \frac{\mathrm{Pr}_0(A_k|\tilde{\mathbf{pa}}(A_k))}{\mathrm{Pr}_0(A_k|\mathbf{pa}(A_k))}$

Figure 3: Algorithm for computing the CPTs of a LogOP-consistent consensus BN.

LinOP queries more efficiently. We exploit the fact that the LinOP obeys MP, and thus that the LinOP of any compound, marginal event can be computed as a LinOP over only that event. For example,

$$\mathrm{Pr}_0(A_2\bar{A}_5A_9) = \sum_{i=1}^N \alpha_i \mathrm{Pr}_i(A_2\bar{A}_5A_9),$$

where the terms on the RHS are calculated using a standard algorithm for Bayesian inference. Any conditional probability can be computed as the division of two compound, marginal probabilities.

Finally, we characterize the computational complexity of LinOP when all input models are BNs. Clearly, computing an arbitrary query $\mathrm{Pr}_0(E|F)$ is NP-hard. Proposition 6 establishes that, even when all topologies agree, and even when only computing the LinOP of a CPT entry, the problem remains intractable.

Proposition 6 *Let all input BNs have identical topologies. Then computing* $\mathrm{Pr}_0(A_j|\mathbf{pa}(A_j))$ *consistent with LinOP is NP-hard.*

Proof. Suppose that $N = 2$. Let Pr_1 be an arbitrary BN and let Pr_2 have an identical topology, but encode a uniform distribution—that is, $\mathrm{Pr}_2(\omega) =$

$1/2^M$. We have shown that, if $\Pr_0(A_M|\mathbf{pa}(A_M))$ were computable in polynomial time, then $\Pr_1(A_M)$ could be inferred in polynomial time. Computing the later query is NP-hard (Cooper, 1990), and so the former must be as well. □

3.5 Related Work

Faria and Smith (1996) examine a group decision making situation where agents agree on a common decomposable BN structure and have identical preferences. They define a weaker form of EB, called *conditional external Bayesianity* (CEB), which requires EB to hold only for CPT entries, and only when evidence updates are based on *cutting* likelihood functions—those which can be factored according to the model structure. They show that a generalized LogOP, called a conditional LogOP, is the only pooling function that satisfies both CEB and UNAM. The conditional modified LogOP preserves the agreed-upon structure and allows expert weights to vary across families in the structure. The authors also present an associated procedure for iteratively revising weights that reflects the relative alignment of the experts' predictions with actual observed outcomes.

Ng and Abramson (1994) describe an architecture called the *probabilistic multi-knowledge-base system*, which consists of a collection of BNs, each encoding the knowledge of a single expert. The BNs are kept separate and probabilities are combined *at run time* with a variable-weight variant of the LinOP. The authors address a variety of engineering issues, including the elicitation and propagation of expert confidence information, and build a working prototype to diagnose pathologies of the lymph system. Xiang (1996) describes conditions under which *multiply sectioned Bayesian networks*, originally developed for single agent reasoning, can represent the combined beliefs of multiple agents. The main assumption is that, whenever two agents' BNs contain some of the same events, they must agree on the joint distribution over these common events. Bonduelle (1987) prescribes both normative and behavioral techniques for a decision maker (DM) to identify and reconcile differences of opinion among experts. When those opinions are expressed as graphical models, he suggests that the DM first choose a consensus topology, and then calculate aggregate probabilities. Jacobs (1995) compares the LinOP and supra Bayesian approaches as methods for combining the multiple feature analyzers found in real and artificial neural systems.

Matzkevich and Abramson (1992) give an algorithm for explicitly combining two BN DAGs into a single DAG, or *fusing* the two topologies. The algorithm transfers one arc at a time from the second DAG to the first, possibly reversing the arc in order to remain consistent with the current partial ordering. Reversing arcs may add new arcs to the second DAG (Shachter, 1988), which would in turn need to be transferred. In a second paper, the same authors show (1993) that the task of minimizing the number of arcs in their combined DAG is NP-hard, as are several other related tasks. They argue that, intuitively, the consensus model should capture independencies agreed upon by at least $c \leq n$ of the agents; in particular, when $c = n$ and the orderings are mutually consistent, the consensus DAG should be a union of the individual DAGs. In both of these papers, and in Bonduelle's work, it is essentially assumed that the EIPP, or a stronger version thereof, should hold.

Though Matzkevich and Abramson make no commitment on how to combine probabilities, they do give an example (1992) where the LinOP is applied *locally*, or separately within each CPT, thus satisfying the FA property. Although such a constraint on aggregation may seem natural, we saw in Section 3.2 that it actually has very severe implications.

4 Graphical Models for Risk Sharing

We turn now to a second common group coordination problem: risk allocation. Securities markets allow agents to transfer risk among themselves, and *complete* securities markets, defined in Section 2, support a Pareto-optimal allocation of risk. Unfortunately, a complete market required a number of securities exponential in the number of primary events, and so is for all practical purposes impossible to achieve in general.

In this section, we explore the extent to which graphical models can help reduce the number of securities required. Though we find some strict conditions under which graphical structures can yield exponential savings, our findings are mainly negative: even unanimously agreed upon independencies may disappear once the group begins interacting.

We begin by noting the strong correspondence between equilibrium prices in a securities market and aggregate beliefs. The connection allows us to track the consequences of the negative and positive results of Section 3 as they apply to risk sharing via securities markets.

4.1 Equilibrium as Consensus

The standard formulation of competitive equilibrium (6) is as a fixed point where each agent's demand is optimal at current prices, and each security's price balances aggregate demand. In this section, we examine an alternative characterization of equilibrium, recognized first by Drèze (1987). Agent i's first-order condition for $x_i^{\langle j \rangle}$ is:

$$\frac{\partial U_i(\mathbf{x})}{\partial x_i^{\langle j \rangle}} = \sum_{\omega \in \Omega} \Pr_i(\omega) \frac{\partial u_i\left(\Upsilon_i^{\langle \omega \rangle}\right)}{\partial x_i^{\langle j \rangle}} = 0,$$

where $\Upsilon_i^{\langle \omega \rangle} = \sum_k \left(1_{\omega \in A_k} - p^{\langle k \rangle}\right) x_i^{\langle k \rangle}$ is its payoff in state ω, and $1_{\omega \in A_k}$ is the indicator function that equals one if $\omega \in A_k$, and zero otherwise. Applying the chain rule

$$\sum_{\omega \in \Omega} \Pr_i(\omega) \left(1_{\omega \in A_j} - p^{\langle j \rangle}\right) u_i'\left(\Upsilon_i^{\langle \omega \rangle}\right) = 0$$

$$\sum_{\omega \in A_j} \Pr_i(\omega) u_i'\left(\Upsilon_i^{\langle \omega \rangle}\right)$$

$$-p^{\langle j \rangle} \sum_{\omega \in \Omega} \Pr_i(\omega) u_i'\left(\Upsilon_i^{\langle \omega \rangle}\right) = 0,$$

and solving for $p^{\langle j \rangle}$, we find that:

$$p^{\langle j \rangle} = \frac{\sum_{\omega \in A_j} \Pr_i(\omega) u_i'\left(\Upsilon_i^{\langle \omega \rangle}\right)}{\sum_{\omega \in \Omega} \Pr_i(\omega) u_i'\left(\Upsilon_i^{\langle \omega \rangle}\right)} = \Pr_i^{\mathrm{RN}}(A_j). \tag{17}$$

In words, equilibrium can also be considered a fixed point where exchanges among agents induce a *consensus* on risk-neutral probabilities across available securities, and where the security prices themselves match these agreed-upon values.

4.2 Complete Markets, Complete Consensus, and Pareto Optimality

As described in Section 2.4, a securities market is *complete* when $S = |\Omega| - 1$ and all securities are *linearly independent*. In such a market, equilibrium allocations of risk are Pareto optimal: any gamble, contingent on *any* event

$E \subseteq \Omega$, that is an acceptable purchase for one agent is *not* an acceptable sale for any other (Arrow, 1964).

A probability distribution over Ω has dimensionality $|\Omega| - 1$ (normalized likelihoods for the $|\Omega|$ states). Prices of securities in a complete market constitute $|\Omega| - 1$ linearly independent equations for these $|\Omega| - 1$ unknowns, and thus define unique probabilities for all states $\omega \in \Omega$, also called the *state prices* (Huang & Litzenberger, 1988; Varian, 1987). Denote these probabilities as $\Pr_0(\omega)$, and let $\Pr_0(E) = \sum_{\omega \in E} \Pr_0(\omega)$ be the price-probability of any event E, perhaps not directly corresponding to an available security.

The agents' risk-neutral distributions also have dimensionality $|\Omega| - 1$, subject to the S constraints defined by (17). If the market is complete, it follows that $\Pr_i^{\mathrm{RN}}$ is uniquely determined, and equals $\Pr_0$ for all i. That is, a complete market induces a compete consensus on risk-neutral probabilities. This suggests an intuitive explanation of why equilibrium allocations are Pareto optimal. All agents behave *as if* they are risk-neutral (payoff-maximizing) with identical beliefs. In such a situation, there are simply no differences of risk-preference or opinion on which to trade.

If $S < |\Omega| - 1$, then the consensus on risk-neutral probabilities is generally incomplete. Whenever $\Pr_h^{\mathrm{RN}}(\omega) \neq \Pr_i^{\mathrm{RN}}(\omega)$ for any ω, there exists an acceptable exchange between agents h and i, though perhaps not supported by the S available securities. An equilibrium allocation in an incomplete market is not necessarily Pareto optimal.[4] But it *can* be, depending on the particular belief structures of the agents. Call a market *operationally complete* if its competitive equilibrium $(\mathbf{x}, \mathbf{p})$ is Pareto optimal (with respect to the agents involved), even if the market contains less than $|\Omega| - 1$ securities. As a degenerate example, an empty market is operationally complete for an economy of completely identical agents. Although such a market does not support all *conceivable* trades, it does support all *acceptable* trades among the given agents.

4.3 Structured Markets: An Analogy to Bayesian Networks

A complete securities market contains $|\Omega| - 1$ securities, essentially one for each $\omega \in \Omega$. In attempting to represent probability distributions over Ω,

[4]Allocations are always efficient with respect to *available* securities, but not necessarily with respect to all states.

researchers in uncertain reasoning are faced with an analogous combinatorial explosion. The typical solution is to work with the factored event space, rather than the state space, and to exploit any independencies among events using graphical models.

Continuing the analogy, securities markets can be structured according to the directed acyclic graph D of any BN. Simply introduce one conditional security $\langle A_j|\mathbf{pa}(A_j)\rangle$ for every conditional probability $\Pr(A_j|\mathbf{pa}(A_j))$ in the network. For each event A_j with $q(j) = |\mathbf{pa}(A_j)|$ parents, this adds $2^{q(j)}$ securities, one for each possible combination of outcomes of events in $\mathbf{pa}(A_j)$. Call such a market *D-structured.* Imagine for the moment that D is fully connected (that is, no independencies are represented). Then a D-structured market contains $\sum_{j=1}^{M} 2^{j-1} = 2^M - 1 = |\Omega| - 1$ linearly independent securities, and is thus complete.

The benefit of a BN representation, and likewise a structured market, obtains when D is less than fully connected, and thus the market contains less than $|\Omega| - 1$ securities. What can be said in this case? Certainly, depending on the beliefs and utilities of the agents, inefficient allocations are possible. Nonetheless, under circumstances explored below, the smaller market may suffice for operational completeness.

4.4 Compact Markets I

4.4.1 Consensus On Risk-Neutral Independencies

Call a D-structured market a *risk-neutral independency market*, or an *RNI-market*, if, in equilibrium, D is an I-map of $\Pr_i^{\mathrm{RN}}$ for all agents i. That is, all agents' risk-neutral distributions agree with the independencies encoded in the market's structure. Paralleling our notation for true conditional independence, let $\mathrm{CI}_i^{\mathrm{RN}}[A_j, W, X]$ denote the risk-neutral conditional independence $\Pr_i^{\mathrm{RN}}(A_j|WX) = \Pr_i^{\mathrm{RN}}(A_j|W)$.

Proposition 7 *At equilibrium in an RNI-market,* $\Pr_h^{\mathrm{RN}}(\omega) = \Pr_i^{\mathrm{RN}}(\omega)$ *for all agents h, i and all states $\omega \in \Omega$.*

Proof. The market contains $\sum_{j=1}^{M} 2^{q(j)}$ securities, imposing an equal number of constraints on every agent's risk-neutral distribution via (17). For each event, I-mapness further imposes $2^{q(j)}(2^{j-1-q(j)} - 1)$ conditional independence constraints of the form $\mathrm{CI}_i^{\mathrm{RN}}[A_j, \mathbf{pa}(A_j), \mathbf{pred}(A_j) - \mathbf{pa}(A_j)]$, for all combinations of outcomes of events in $\mathbf{pa}(A_j)$ and all but one combination of outcomes of events in $\mathbf{pred}(A_j) - \mathbf{pa}(A_j)$ (the remaining one is implied by the others). Then every agent's risk-neutral distribution is subject to

$$\sum_{j=1}^{M} 2^{q(j)} + 2^{q(j)}(2^{j-1-q(j)} - 1)$$
$$= \sum_{j=1}^{M} 2^{j-1} = 2^M - 1 = |\Omega| - 1$$

identical, linearly independent constraints. Therefore $\Pr_h^{\mathrm{RN}} = \Pr_i^{\mathrm{RN}}$ for all h, i. $\square$

In an RNI-market, define the *state prices* $\Pr_0(\omega) = \Pr_i^{\mathrm{RN}}(\omega)$ as the unique probabilities over Ω that are consistent with the prices of available securities and the independencies of D. The following corollary establishes that equilibrium prices for any of the $|\Omega| - 1 - S$ "missing" securities are also derivable from $\Pr_0$.

Corollary 8 *Let $\langle p^{\langle 1\rangle}, \ldots, p^{\langle S\rangle}\rangle$ be the equilibrium prices in an RNI-market. Introduce a new security $\langle E\rangle$. Then $\langle p^{\langle 1\rangle}, \ldots, p^{\langle S\rangle}, \Pr_0(E)\rangle$ are equilibrium prices in the expanded market.*

Proof. Before the extra security is introduced, all agents' risk-neutral probabilities $\Pr_i^{\mathrm{RN}}(E)$ already equal $\Pr_0(E)$, without buying or selling any quantity of the security. It follows that, with the additional security, the equilibrium condition (17) is satisfied with $x_i^{\langle E\rangle} = 0$ for all i, $p^{\langle E\rangle} = \Pr_0(E)$, and all other prices unchanged. $\square$

The number of securities in an RNI-market, $O\left(M \cdot 2^{\max\{q(j)\}}\right)$, can be exponentially smaller than the $2^M - 1$ required for traditional completeness. The following corollary shows that the more compact market supports allocations that are equally efficient.

Corollary 9 *Every RNI-market is operationally complete. That is, the equilibrium allocations $\mathbf{x}$ and state prices $\Pr_0$ in an RNI-market constitute an equilibrium in a (truly) complete market composed of the same agents.*

Proof. By repeated application of Corollary 8, we can add the $|\Omega| - 1 - S$ securities necessary to complete the market.[5] For each new security, a price

[5] A natural set to add are the $\sum_{j=1}^{M} 2^{q(j)}(2^{j-1-q(j)} - 1)$ securities of the form $\langle A_j|\mathbf{pred}(A_j)\rangle$, for all events A_j, all combinations of outcomes of $\mathbf{pa}(A_j)$, and all but one combination of outcomes of $\mathbf{pred}(A_j) - \mathbf{pa}(A_j)$.

consistent with $\Pr_0$, coupled with zero demand from all agents, satisfies (17). All complete markets, regardless of structure, support the same equilibrium allocations and state prices (Huang & Litzenberger, 1988; Mas-Colell et al., 1995; Varian, 1987). □

Proposition 7 and its corollaries are equilibrium results only. We sketch here one possible procedure for *reaching* agreement on the market structure.[6] Begin with securities in only the M events: $\langle A_1 \rangle, \ldots, \langle A_M \rangle$. If any agent's demand for $\langle A_k | A_j \rangle$ (for any $j < k$) at price $p^{\langle A_k \rangle}$ is nonzero, then it creates a new market in $\langle A_k | A_j \rangle$. If, at some future time, the agent has zero demand for its new security, then it may retract the security. An additional condition for equilibrium is that no agent desires to create or withdraw any markets. Then, in equilibrium, it should be the case that all agents' risk-neutral independencies agree with the market structure, and that the market is operationally complete. We might want to add a transaction cost for opening new markets, so that equilibrium only ensures that risks are hedged up to a threshold cost.

4.4.2 Computational Complexity Of Arbitrage

Imagine that, after equilibrium is reached in an RNI-market, a redundant security is introduced, say $\langle A_M \rangle$. The equilibrium price of $\langle A_M \rangle$ is already determined (Corollary 8): it must equal $\Pr_0(A_M) = \Pr_i^{\mathrm{RN}}(A_M)$. Furthermore, if the current price does *not* equal $\Pr_0(A_M)$, then the market is not in equilibrium, and arbitrage is possible. For example, if $p^{\langle A_M \rangle} < \Pr_0(A_M)$, then an outside observer O could purchase it at the going price and sell it to any of the agents at price p^* such that $p^{\langle A_M \rangle} < p^* < \Pr_i^{\mathrm{RN}}(A_M) = \Pr_0(A_M)$. Although O does not have direct access to $\Pr_0(A_M)$, it is uniquely computable given the other prices and the independence structure of D.

If O can find an arbitrage opportunity by correctly pricing the redundant security, then O can perform Bayesian inference, which is #P-complete (Cooper, 1990).

[6] This procedure is similar to Geiger's (1990) protocol for eliciting independence structures from experts.

4.5 Compact Markets II: Consensus on True Independencies

Equilibrium agreement on risk-neutral independencies may seem a somewhat strange condition, especially considering that the $\Pr_i^{\mathrm{RN}}$ are changing as transactions occur. Some authors argue that, since agents appear to act according to $\Pr_i^{\mathrm{RN}}$ and standard elicitation techniques reveal $\Pr_i^{\mathrm{RN}}$, risk-neutral probabilities are in fact no less "real" than true probabilities (Kadane & Winkler, 1988; Nau & McCardle, 1991). However, while it seems reasonable that agents would have true independencies in common (Pearl, 1993; Smith, 1990), it is harder to justify why their risk-neutral independencies would coincide. This section develops a theory of compact markets based on consensus on true independencies. If, despite any quantitative differences between $\Pr_i$ and $\Pr_i^{\mathrm{RN}}$, an agent's true independencies were always manifest as risk-neutral independencies, then results concerning RNI-markets would carry over unchanged. Section 4.5.1 demonstrates that this is indeed the case for a subclass of agents and a subset of independencies. Section 4.5.2 discusses how known limitations of belief aggregation procedures restrict the possibility of obtaining compact markets under more general circumstances.

4.5.1 Consensus on Markov Independencies

A commonly assumed risk-averse utility form is *exponential utility*: $u_i(\mu) = -e^{-c_i\mu}$. This utility form is synonymous with *constant absolute risk aversion* (CARA), where c_i is agent i's coefficient of risk aversion, or $1/c_i$ its risk tolerance. As the agent's wealth increases, its marginal utility for unit dollars decreases (since it is risk-averse), but the *amount* of its aversion to risk remains constant at c_i.

In this section, we show that, in economies composed of agents with CARA, markets structured according to agreed upon (true) *Markov* independencies are operationally complete. Define an *independency market*, or an *I-market*, as a D-structured market such that D is an I-map of $\Pr_i$ for all agents i (i.e., all agents' true distributions agree with the independencies in D). An I-market is *decomposable* if D is decomposable—every node's parents are fully connected.

Let $Z = \{A_1, \ldots, A_M\}$ be the set of all events, $A_j \in Z$ a particular event, and $W \subseteq Z - A_j$ and $X = Z - W - A_j$ subsets of events. We are interested in whether agent i's Markov independencies $\mathrm{CI}_i[A_j, W, X]$ are reflected as a risk-neutral independencies $\mathrm{CI}_i^{\mathrm{RN}}[A_j, W, X]$, and are thus observable. For brevity, we drop the subscript i when only one agent is under consideration.

Proposition 10

$$\mathrm{CI}[A_j, W, X] \& \left(\frac{u'\left(\Upsilon^{\langle \bar{A}_j W X \rangle}\right)}{u'\left(\Upsilon^{\langle A_j W X \rangle}\right)} = \frac{u'\left(\Upsilon^{\langle \bar{A}_j W \tilde{X} \rangle}\right)}{u'\left(\Upsilon^{\langle A_j W \tilde{X} \rangle}\right)} \right)$$
$$\Rightarrow \mathrm{CI}^{\mathrm{RN}}[A_j, W, X], \tag{18}$$

where the second precondition must hold for all possible joint outcomes of the events in W, and all pairs $(X, \tilde{X})$ of different joint outcomes of events in X.

Proof.

$$\frac{u'\left(\Upsilon^{\langle \bar{A}_j W X \rangle}\right)}{u'\left(\Upsilon^{\langle A_j W X \rangle}\right)} = \frac{u'\left(\Upsilon^{\langle \bar{A}_j W \tilde{X} \rangle}\right)}{u'\left(\Upsilon^{\langle A_j W \tilde{X} \rangle}\right)}$$

$$\Pr(A_j W) + \Pr(\bar{A}_j W)\frac{u'\left(\Upsilon^{\langle \bar{A}_j W X \rangle}\right)}{u'\left(\Upsilon^{\langle A_j W X \rangle}\right)}$$
$$= \Pr(A_j W) + \Pr(\bar{A}_j W)\frac{u'\left(\Upsilon^{\langle \bar{A}_j W \tilde{X} \rangle}\right)}{u'\left(\Upsilon^{\langle A_j W \tilde{X} \rangle}\right)}$$

$$\frac{\frac{\Pr(A_j W)\Pr(WX)}{\Pr(W)} u'\left(\Upsilon^{\langle A_j W X \rangle}\right)}{\frac{\Pr(A_j W)\Pr(WX)}{\Pr(W)} u'\left(\Upsilon^{\langle A_j W X \rangle}\right) + \frac{\Pr(\bar{A}_j W)\Pr(WX)}{\Pr(W)} u'\left(\Upsilon^{\langle \bar{A}_j W X \rangle}\right)}$$
$$= \frac{\frac{\Pr(A_j W)\Pr(W\tilde{X})}{\Pr(W)} u'\left(\Upsilon^{\langle A_j W \tilde{X} \rangle}\right)}{\frac{\Pr(A_j W)\Pr(W\tilde{X})}{\Pr(W)} u'\left(\Upsilon^{\langle A_j W \tilde{X} \rangle}\right) + \frac{\Pr(\bar{A}_j W)\Pr(W\tilde{X})}{\Pr(W)} u'\left(\Upsilon^{\langle \bar{A}_j W \tilde{X} \rangle}\right)}$$

$$\frac{\Pr(A_j W X) u'\left(\Upsilon^{\langle A_j W X \rangle}\right)}{\Pr(A_j W X) u'\left(\Upsilon^{\langle A_j W X \rangle}\right) + \Pr(\bar{A}_j W X) u'\left(\Upsilon^{\langle \bar{A}_j W X \rangle}\right)}$$
$$= \frac{\Pr(A_j W \tilde{X}) u'\left(\Upsilon^{\langle A_j W \tilde{X} \rangle}\right)}{\Pr(A_j W \tilde{X}) u'\left(\Upsilon^{\langle A_j W \tilde{X} \rangle}\right) + \Pr(\bar{A}_j W \tilde{X}) u'\left(\Upsilon^{\langle \bar{A}_j W \tilde{X} \rangle}\right)}$$

$$\frac{\Pr^{\mathrm{RN}}(A_j W X)}{\Pr^{\mathrm{RN}}(A_j W X) + \Pr^{\mathrm{RN}}(\bar{A}_j W X)} = \frac{\Pr^{\mathrm{RN}}(A_j W \tilde{X})}{\Pr^{\mathrm{RN}}(A_j W \tilde{X}) + \Pr^{\mathrm{RN}}(\bar{A}_j W \tilde{X})}$$
$$\Pr^{\mathrm{RN}}(A_j | W X) = \Pr^{\mathrm{RN}}(A_j | W \tilde{X})$$

□

The second precondition in (18) says that the ratio of marginal utility in states where A_j does not occur to marginal utility in states where A_j does occur cannot depend of the outcomes of events in X. This is true (and indeed $\mathrm{Pr}^{\mathrm{RN}} = \mathrm{Pr}$) if the agent's marginal utility u' is constant across states. This holds if the agent is risk neutral, and holds approximately if utility is state-independent and $\Upsilon^{\langle\omega_j\rangle} \approx \Upsilon^{\langle\omega_k\rangle}$. But this approximation is not realistic for an agent engaged in trading securities, since a central role of the market is precisely to enable the transfer of wealth across states.

Let $\Upsilon^{\langle A_j W\rangle}$ be the agent's payoff from all securities that depend only the outcomes of events in $A_j \cup W$. Examples are $\langle A_j\rangle$, $\langle A_j W\rangle$, and $\langle A_j|W\rangle$, which return the same dollar amount regardless of the realizations of events in $X = Z - W - A_j$. Similarly, let $\Upsilon^{\langle WX\rangle}$ be the payoff from securities that do not depend on A_j.

Suppose that the agent exhibits CARA, and that its payoffs are separable according to $\Upsilon^{\langle A_j WX\rangle} = \Upsilon^{\langle A_j W\rangle} + \Upsilon^{\langle WX\rangle} - \Upsilon^{\langle W\rangle}$. Separability essentially means that any of the agent's securities (or prior stakes) whose payoff depends on A_j cannot also depend on events in X. In this case,

$$
\begin{aligned}
&\frac{u'\left(\Upsilon^{\langle \bar{A_j} WX\rangle}\right)}{u'\left(\Upsilon^{\langle A_j WX\rangle}\right)} = \frac{u'\left(\Upsilon^{\langle \bar{A_j} W\rangle} + \Upsilon^{\langle WX\rangle} - \Upsilon^{\langle W\rangle}\right)}{u'\left(\Upsilon^{\langle A_j W\rangle} + \Upsilon^{\langle WX\rangle} - \Upsilon^{\langle W\rangle}\right)} \\
&= \frac{ce^{-c\Upsilon^{\langle \bar{A_j} W\rangle}} e^{-c\Upsilon^{\langle WX\rangle}} e^{c\Upsilon^{\langle W\rangle}}}{ce^{-c\Upsilon^{\langle A_j W\rangle}} e^{-c\Upsilon^{\langle WX\rangle}} e^{c\Upsilon^{\langle W\rangle}}} \\
&= \frac{ce^{-c\Upsilon^{\langle \bar{A_j} W\rangle}} e^{-c\Upsilon^{\langle W\tilde{X}\rangle}} e^{c\Upsilon^{\langle W\rangle}}}{ce^{-c\Upsilon^{\langle A_j W\rangle}} e^{-c\Upsilon^{\langle W\tilde{X}\rangle}} e^{c\Upsilon^{\langle W\rangle}}} \\
&= \frac{u'\left(\Upsilon^{\langle \bar{A_j} W\rangle} + \Upsilon^{\langle W\tilde{X}\rangle} - \Upsilon^{\langle W\rangle}\right)}{u'\left(\Upsilon^{\langle A_j W\rangle} + \Upsilon^{\langle W\tilde{X}\rangle} - \Upsilon^{\langle W\rangle}\right)} = \frac{u'\left(\Upsilon^{\langle \bar{A_j} W\tilde{X}\rangle}\right)}{u'\left(\Upsilon^{\langle A_j W\tilde{X}\rangle}\right)}
\end{aligned}
$$

Thus the constraint on utility in (18) is satisfied, and any Markov independencies are observable.

We are now in a position to derive the main result of this section.

Proposition 11 *When all agents have CARA, every decomposable I-market is an RNI-market.*

Proof. Let W_j be the set of direct parents and direct children of event A_j, and X_j all other events. From decomposability and I-mapness, we can infer that

1. $\mathrm{CI}_i[A_j, W_j, X_j]$ for all agents i and events j,
2. none of the securities $\langle A_j|\mathbf{pa}(A_j)\rangle$ that are contingent on A_j depend on X_j, and

3. none of the securities $\langle A_k | \mathbf{pa}(A_k) \rangle$ such that $A_j \in \mathbf{pa}(A_k)$ that are conditional on A_j depend on X_j.

Items 2 and 3 ensure separability of payoffs from the available securities (we assume that any prior stakes are also separable). Then, invoking Proposition 10, $\mathrm{CI}_i^{\mathrm{RN}}[A_j, W_j, X_j]$ for all agents i and events j. As a result, D is an I-map of every $\Pr_i^{\mathrm{RN}}$ regardless of allocations or prices, including those at equilibrium. $\Box$

Proposition 7 and Corollaries 8 and 9 are immediately applicable. In particular, for agents with CARA, every decomposable I-market is operationally complete.

4.5.2 Inherent Limitations

One might wonder whether compact I-markets are possible for larger classes of agents or independencies. It can be shown via counterexample that, even when all agents have CARA, a market conforming to agreed-upon (possibly non-Markov) independencies will not always be operationally complete. Moreover, when all agents have logarithmic utility for money (another commonly assumed utility form), even a market conforming to agreed-upon Markov independencies will not always be operationally complete.

Although we do not have a formal statement of impossibility, results from statistical belief aggregation suggest that agreement on true independencies will *not* be sufficient in general to yield compact and operationally complete markets. The state prices $\Pr_0$ in a securities market are a function of all the agents' beliefs (and their utilities), and as such essentially constitute a measure of aggregate belief. Many researchers have studied belief aggregation functions (Genest & Zidek, 1986), and several impossibility theorems severely restrict the class of functions that preserve unanimously held independencies (Genest & Wagner, 1987), even when restricted to independencies among the primary events (Pennock & Wellman, 1999). The aggregation "function" of a securities market is of course subject to the same limitations. We suspect that, for many configurations of agents, markets structured according to unanimously-held true independencies will not yield provably optimal allocations of risk. Nevertheless, it may well be the case that structured markets can yield *approximately* optimal allocations over a wider range of agent populations.

The examples in Section 3 where the LinOP (Example 1) and LogOP

(Example 2) fail to preserve unanimous independencies have direct, negative implications for the possibility of compact I-markets for larger classes of agents or independencies.

Let two agents have beliefs over two events, as prescribed in Example 1 and pictured in Figure 1(a). Both agents agree that the events are independent. Suppose additionally that they both have GLU with equal wealth parameters. In a complete market (of $2^2-1=3$ linearly independent securities), the unique state prices (and the unique consensus risk-neutral probabilities) are the same as derived by the LinOP in the example, and do *not* reflect the independence. Thus an I-market consisting of only the two securities $\langle A_1\rangle$ and $\langle A_2\rangle$ is not operationally complete.

In Example 2 (Figure 1(d)), two agents agree that, among three primary events, A_1 and A_2 are independent, and A_3 depends on both. If both agents have CARA, then the unique state prices in a complete market equal the LogOP consensus probabilities, and do not preserve the independence between A_1 and A_2. Thus an I-market mirroring the agreed-upon polytree structure of Figure 1(d) (containing the six securities $\langle A_1\rangle$, $\langle A_2\rangle$, $\langle A_3|A_1A_2\rangle$, $\langle A_3|A_1\bar{A}_2\rangle$, $\langle A_3|\bar{A}_1A_2\rangle$, and $\langle A_3|\bar{A}_1\bar{A}_2\rangle$) is not operationally complete.

5 Conclusion

Positive progress in research in group coordination, though certainly not lacking, is circumscribed by controversy and impossibilities. For example, a proliferation of results in the 1980s, exemplified by groundbreaking contributions by Genest (1984a, 1984b, 1984c) and his coauthors (Genest et al., 1986; Genest & Wagner, 1987), do much to demarcate the impassable boundaries in the context of belief aggregation. Contributions in this paper as well fall on both sides of the impossibility fence. Section 5.1 catalogues those results that entail new limitations, Section 5.2 those that uncover new possibilities.

5.1 For the Pessimist...

A subset of results in this paper further confine and confound the search for reasonable aggregation procedures, by extending the impossibility theorems to new domains, and by raising new concerns.

A series of theorems (Lehrer & Wagner, 1983; Wagner, 1984) culminating in that of Genest and Wagner (1987) show that very weak and reasonable constraints on an aggregation function are enough to rule out independence preservation (i.e., the retention of all agreed-upon independencies within the aggregate distribution). But these theorems apply to functions that preserve all possible independencies between any events—even those not representable in a graphical model. A potential loophole remained that some reasonable function might preserve the independencies among primary events in a graphical model. Indeed, in Section 3, we see that the same conditions sufficient to rule out general independence preservation are *not* sufficient to rule out this weaker form. However, we show that, with the additional (uncontroversial) assumption of unanimity, the impossibility returns.

This result resurfaces in the study of structured securities markets in Section 4. The intuitive inclination to structure the market according to agreed-upon independencies proved fatally flawed. Prices in a securities market are essentially the output of an aggregation function—the one defined by market equilibrium—and are thus subject to all the general limitative theorems.

We derive a second impossibility theorem in Section 3 regarding the combination of BNs. A natural policy—that other authors have advocated or assumed—is to confine the aggregation locally, within each conditional probability table of the BN. We prove that any such local aggregation function necessarily fails to satisfy either unanimity or nondictatorship, two seemingly incontrovertible assumptions.

Other results demonstrate that desirable operations, while not impossible, are instead (worst-case) intractable. Someone interested in computing the LinOP of several probability distributions, each represented as a BN, would not want to construct a consensus BN, as it would in general be fully connected. This suggests keeping the individual BNs separate, and computing the LinOP of any desired query at runtime. In Proposition 6, we prove that performing this computation is NP-hard, even if answering the same query is easy within each individual BN. Similarly, we prove that LogOP is NP-hard as well.

Even the positive results in Section 3 describing BN representations of the logarithmic opinion pool (LogOP), are shaded by potential computational barriers. The consensus network structure must be made decomposable, a process that can increase the size of the representation exponentially. Similar computational concerns arise in Section 4 when the structure of the securities market is required to be decomposable. we also show in Section 4 that properly pricing securities and finding arbitrage opportunities within in a compact market is NP-hard.

5.2 For the Optimist...

On the other hand, some results in this paper can be characterized as possibility results. Each identifies a weakening of an impossibility theorem that exposes a (hopefully nontrivial) solution.

One possibility result arises by weakening Genest and Wagner's (1987) axiom for preserving independence, to require only the preservation of Markov independencies. Section 3 demonstrates that the LogOP does in fact maintain all agreed-upon Markov independencies. This suggests that, if the preservation of independence structure is important—and many authors argue that it is (Laddaga, 1977; Raiffa, 1968)—then the LogOP may be the most viable option. Markov independencies play an important role in the theory of graphical models, and are precisely the type representable in Markov networks (MNs) and decomposable BNs. We describe procedures for constructing MN and BN structures consistent with the LogOP consensus. We also delineate an algorithm for computing all of the conditional probability tables in a LogOP-consensus BN. This structured representation is potentially exponentially smaller than the standard representation.

The preservation of Markov independencies has a direct corollary in Section 4's investigation of structured securities markets. For a certain class of agents, true Markov independencies are always observable as risk-neutral independencies. Thus, if all agents are of this type, all beliefs agree with the independencies encoded in the market structure, and this structure is decomposable, then the market is operationally complete.

Acknowledgments

We thank to Didier Dubois, C. Lee Giles, Robin Hanson, Eric Horvitz, Jeffrey MacKie Mason, Robert Nau, Charles Plott, Stephen Pollock, Ross Shachter, James Smith, Mike West, Fredrik Ygge, the members of the Decision Machine Research Group at the University of Michigan, and the anonymous reviewers of conference versions of this work. Parts of this investigation were conducted while the first author was at Microsoft Research and at NEC Laboratories America. This work was partially supported by AFOSR Grant F49620-97-0175.

References

Arrow, K. J. (1964). The role of securities in the optimal allocation of risk-bearing. *Review of Economic Studies*, *31*(2), 91–96.

Benediktsson, J. A., & Swain, P. H. (1992). Consensus theoretic classification methods. *IEEE Transactions on Systems, Man, and Cybernetics*, *22*(4), 688–704.

Bonduelle, Y. (1987). *Aggregating expert opinions by resolving sources of disagreement*. Ph.D. thesis, Stanford University.

Chyu, C. C. (1991). Decomposable probabilistic influence diagrams. *Probability in the Engineering and Informational Sciences*, *5*, 229–243.

Clemen, R. T., & Winkler, R. L. (1993). Aggregation of point estimates: A flexible modeling approach. *Management Science*, *39*(4), 501–515.

Cooke, R. M. (1991). *Experts in Uncertainty: Opinion and Subjective Probability in Science*. Environmental Ethics and Science Policy Series. Oxford University Press, New York.

Cooper, G. (1990). The computational complexity of probabilistic inference using Bayes belief networks. *Artificial Intelligence*, *42*, 393–405.

Dalkey, N. C. (1975). Toward a theory of group estimation. In Linstone, H. A., & Turoff, M. (Eds.), *The Delphi Method: Techniques and Applications*, pp. 236–261. Addison-Wesley, Reading, MA.

Darroch, J. N., Lauritzen, S. L., & Speed, T. P. (1980). Markov fields and log-linear interaction models for contingency tables. *Annals of Statistics*, *8*(3), 522–539.

de Finetti, B. (1974). *Theory of Probability: A Critical Introductory Treatment*, Vol. 1. Wiley, New York.

Degroot, M. H. (1974). Reaching a consensus. *Journal of the American Statistical Association*, *69*(345), 118–121.

Degroot, M. H., & Mortera, J. (1991). Optimal linear opinion pools. *Management Science*, *37*(5), 546–558.

Dreze, J. H. (1987). Market allocation under uncertainty. In *Essays on Economic Decisions under Uncertainty*, pp. 119–143. Cambridge University Press.

Faria, A. E., & Smith, J. Q. (1996). Conditional external bayesianity in decomposable influence diagrams. In *Bayesian Statistics 5*, pp. 551–560.

French, S. (1985). Group consensus probability distributions: A critical survey. *Bayesian Statistics*, *2*, 183–202.

Geiger, D. (1990). *Graphoids: A Qualitative Framework for Probabilistic Inference*. Ph.D. thesis, UCLA.

Genest, C. (1984a). A characterization theorem for externally Bayesian groups. *Annals of Statistics*, *12*(3), 1100–1105.

Genest, C. (1984b). A conflict between two axioms for combining subjective distributions. *Journal of the Royal Statistical Society*, *46*(3), 403–405.

Genest, C. (1984c). Pooling operators with the marginalization property. *Canadian Journal of Statistics*, *12*(2), 153–163.

Genest, C., McConway, K. J., & Schervish, M. J. (1986). Characterization of externally Bayesian pooling operators. *Annals of Statistics*, *14*(2), 487–501.

Genest, C., & Wagner, C. G. (1987). Further evidence against independence preservation in expert judgement synthesis. *Aequationes Mathematicae*, *32*(1), 74–86.

Genest, C., & Zidek, J. V. (1986). Combining probability distributions: A critique and an annotated bibliography. *Statistical Science*, *1*(1), 114–148.

Huang, C.-F., & Litzenberger, R. H. (1988). *Foundations for Financial Economics*. Elsevier, Amsterdam.

Jacobs, R. A. (1995). Methods for combining experts' probability assessments. *Neural Computation*, *7*(5), 867–888.

Jensen, F. V. (1996). *An Introduction to Bayesian Networks*. Springer, New York.

Kadane, J. B., & Winkler, R. L. (1988). Separating probability elicitation from utilities. *Journal of the American Statistical Association*, *83*(402), 357–363.

Kloks, T. (1994). *Treewidth: Computations and Approximations*. Springer-Verlag, Berlin.

Laddaga, R. (1977). Lehrer and the consensus proposal. *Synthese*, *36*, 473–477.

Lauritzen, S. L., & Spiegelhalter, D. J. (1988). Local computations with probabilities on graphical structures and their application to expert systems. *Journal of the Royal Statistical Society, Series B*, *50*, 157–224.

Lehrer, K., & Wagner, C. G. (1983). Probability amalgamation and the independence issue: A reply to Laddaga. *Synthese*, *55*(3), 339–346.

Levy, W. B., & Delic, H. (1994). Maximum entropy aggregation of individual opinions. *IEEE Transactions on Systems, Man, and Cybernetics*, *24*(4), 606–613.

Lindley, D. V. (1985). Reconciliation of discrete probability distributions. In *Bayesian Statistics 2*, pp. 375–390 Amsterdam. North-Holland.

Lindley, D. V. (1988). The use of probability statements. In *Accelerated Life Testing and Experts' Opinions in Reliability*, pp. 25–57 Amsterdam. Elsevier.

Mas-Colell, A., Whinston, M. D., & Green, J. R. (1995). *Microeconomic Theory*. Oxford University Press, New York.

Matzkevich, I., & Abramson, B. (1992). The topological fusion of Bayes nets. In *Eighth Conference on Uncertainty in Artificial Intelligence*, pp. 152–158.

Matzkevich, I., & Abramson, B. (1993). Some complexity considerations in the combination of belief networks. In *Ninth Conference on Uncertainty in Artificial Intelligence*, pp. 152–158.

Morris, P. A. (1974). Decision analysis expert use. *Management Science*, *20*(9), 1233–1241.

Morris, P. A. (1977). Combining expert judgments: A Bayesian approach. *Management Science*, *23*(7), 679–693.

Myung, I. J., Ramamoorti, S., & Bailey, A. D. (1996). Maximum entropy aggregation of expert predictions. *Management Science*, *42*(10), 1420–1436.

Nau, R. F. (1995). Coherent decision analysis with inseparable probabilities and utilities. *Journal of Risk and Uncertainty*, *10*(1), 71–91.

Nau, R. F., & McCardle, K. F. (1991). Arbitrage, rationality, and equilibrium. *Theory and Decision*, *31*, 199–240.

Neapolitan, R. E. (1990). *Probabilistic Reasoning in Expert Systems: Theory and Algorithms*. Wiley, New York.

Ng, K.-C., & Abramson, B. (1992). Consensus diagnosis—A simulation study. *IEEE Transactions on Systems, Man, and Cybernetics*, *22*(5), 916–928.

Ng, K.-C., & Abramson, B. (1994). Probabilistic multi-knowledge-base systems. *Applied Intelligence*, *4*(2), 219–236.

Pearl, J. (1988). *Probabilistic Reasoning in Intelligent Systems*. Morgan Kaufmann.

Pearl, J. (1993). From Bayesian networks to causal networks. In *Adaptive Computing and Information Processing 1*, pp. 165–194.

Pennock, D. M., & Wellman, M. P. (1999). Graphical representations of consensus belief. In *Fifteenth Conference on Uncertainty in Artificial Intelligence (UAI–99)*, pp. 531–540.

Raiffa, H. (1968). *Decision Analysis: Introductory Lectures on Choices under Uncertainty*. Addison-Wesley, Reading, MA.

Rosenblueth, E., & Ordaz, M. (1992). Combination of expert opinions. *Journal of Scientific and Industrial Research*, *51*, 572–580.

Shachter, R. D. (1988). Probabilistic inference and influence diagrams. *Operations Research*, *36*(4), 589–604.

Shachter, R. D., Andersen, S. K., & Poh, K. L. (1991). Directed reduction algorithms and decomposable graphs. In *Uncertainty in Artificial Intelligence*, Vol. 6, pp. 197–208. North Holland, Amsterdam.

Smith, J. (1990). Statistical principles on graphs. In Oliver, R., & Smith, J. (Eds.), *Influence Diagrams, Belief Nets and Decision Analysis*, pp. 89–120. Wiley, New York.

Varian, H. R. (1987). The arbitrage principle in financial economics. *J. Economic Perspectives*, *1*(2), 55–72.

Wagner, C. (1984). Aggregating subjective probabilities: Some limitative theorems. *Notre Dame Journal of Formal Logic*, *25*(3), 233–240.

West, M., & Crosse, J. (1992). Modeling probabilistic agent opinion. *Journal of the Royal Statistical Society. Series B. Methodological*, *54*(1), 285–299.

Whittaker, J. (1990). *Graphical Models in Applied Multivariate Statistics.* Wiley, Chichester.

Winkler, R. L. (1968). The consensus of subjective probability distributions. *Management Science*, *15*(2), B61–B75.

Winkler, R. L. (1981). Combining probability distributions from dependent sources. *Management Science*, *27*(4), 479–488.

Winkler, R. L. (1986). Expert resolution. *Management Science*, *32*(3), 298–303. Introduces commentary by Lindley, Schervish, Clemen, French, and Morris.

Winkler, R. L., & Murphy, A. H. (1968). Good probability assessors. *J. Applied Meteorology*, *7*, 751–758.

Xiang, Y. (1996). A probabilistic framework for cooperative multi-agent distributed interpretation and optimization of communication. *Artificial Intelligence*, *87*(1–2), 295–342.

Yule, G. U. (1903). Notes on the theory of association of attributes in statistics. *Biometrika*, *2*, 121–134.