DECISION ANALYSIS
Vol. 2, No. 4, December 2005, pp. 207–228
ISSN 1545-8490 | EISSN 1545-8504 | 05 | 0204 | 0207

DOI 10.1287/deca.1050.0047

Influence Diagrams for Team Decision Analysis

Apiruk Detwarasiti
The Boston Consulting Group (Thailand), 37th Floor, U-Chu-Liang Building, 968 Rama IV Road, Bangkok 10500, Thailand, detwarasiti.apiruk@bcg.com

Ross D. Shachter
Department of Management Science and Engineering, School of Engineering, Stanford University, 380 Panama Way, Stanford, California 94305-4026, shachter@stanford.edu

We consider the representation and evaluation of team decision making under uncertainty using influence diagrams. We assume that all team members agree on common beliefs and preferences, but complete sharing of information is generally impossible. As a result, the team can be represented as a single rational individual with imperfect recall, and the optimal solution with perfect recall might not be achievable, except in special cases we can recognize. An alternative solution concept is a stable solution that integrates the notion of optimality with that of equilibrium from game theory. We extend this concept from individual decisions to sets of decisions, and introduce the Strategy Improvement and its variation, Uniform Strategy Improvement, as the corresponding solution methods. We also provide a variety of simplifying transformations to the influence diagram by exploiting its graphical structure. The result is a requisite influence diagram, one that requires minimum assessment and creates additional opportunities for optimality.

Key words: team decision analysis; requisite influence diagram; imperfect recall; incomplete sharing of information; stability; uniform strategy improvement; strategic irrelevance
History: Received on May 3, 2004. Accepted by Eric Horvitz on August 26, 2005, after 1 revision.

1. Introduction

Decision analysis provides a principled framework for decision making under uncertainty for a rational individual (Howard 1966a, Raiffa 1968). Influence diagrams (Howard and Matheson 1984/2005) are compact graphical representations of these decision situations, and efficient algorithms have been developed to analyze them (Olmsted 1983, Shachter 1986, Shachter and Peot 1992, Shenoy 1992, Jensen et al. 1994, Zhang 1998). The diagrams allow us to represent the decision maker's beliefs about uncertainties, preferences for prospects, alternatives for decisions, and observations available at the time of those decisions.

When there is a single decision maker, rationality requires perfect recall, that any observations available and choices made at the time of earlier decisions are known when later decisions are made (Kuhn 1953). This leads to a globally optimal solution for all of the sequential decision situations through the solution technique of backward induction (BI) (Bellman 1957). BI is also applicable to team decision situations, those in which team members agree on probabilities for the uncertainties, preferences for the prospects, and alternatives for the decisions, as long as the observations available at the time of those decisions satisfy certain conditions, e.g., that of finite-stage Markov Decision Process (MDP) (Bellman 1957).

We seek to improve the quality of team decision making where the perfect recall condition does not hold due to the incomplete sharing of information among team members. This is usually because decisions are decentralized throughout the team and commitments are being made in parallel. It could also be that the record keeping needed for the complete sharing of information is impractical or because it is desirable to provide simple strategies to some members of the team.

There are two perspectives to approach team decision making with incomplete sharing of information, either as an individual making decisions with imperfect recall or as a game with imperfect information where every player has identical payoffs (Marschak and Radner 1972). We will take the former perspective because the latter perspective can be viewed as imperfect recall of a hypothetical decision maker who owns every decision in the game. This avoids confusion with another use of the term imperfect (perfect) information, already defined as imperfect (perfect) observations in decision analysis (Howard 1966b). It also allows the representation of team decision situations to be independent of the identity of the team members and facilitates the comparison of various information structures based on the notion of value of information gathering and sharing in team environment.

Example 1. Consider the following story adapted from a classical team decision situation about a shipyard firm from Marschak and Radner (1972). The firm has two docks (a new one and an old, less efficient one) and two markets (West and East). Each manager (West and East) is offered a price through a private negotiation for a ship to be delivered in each respective market. The price is known to each, but not to the other manager when each manager must decide whether to accept or reject the offer. A dock supervisor decides whether to move some infrastructure from the new dock to the old dock in anticipation that both docks might be used. This decision will be known to each manager prior to the negotiation. Its effect is to increase and decrease the costs of building a ship in the new and old docks, respectively. □

To represent a decision situation with imperfect recall, a standard decision tree is inadequate and must be augmented with information sets as used in game trees (Kuhn 1953). A set of decision nodes in the tree are in the same information set if the decision maker cannot distinguish which one of them applies. It follows semantically that each decision node belongs to exactly one information set, possibly a singleton set, and the available alternatives as well as the choice for every decision node in the same information set must be identical. We represent an information set by a dashed line connecting the decision nodes in the information set.

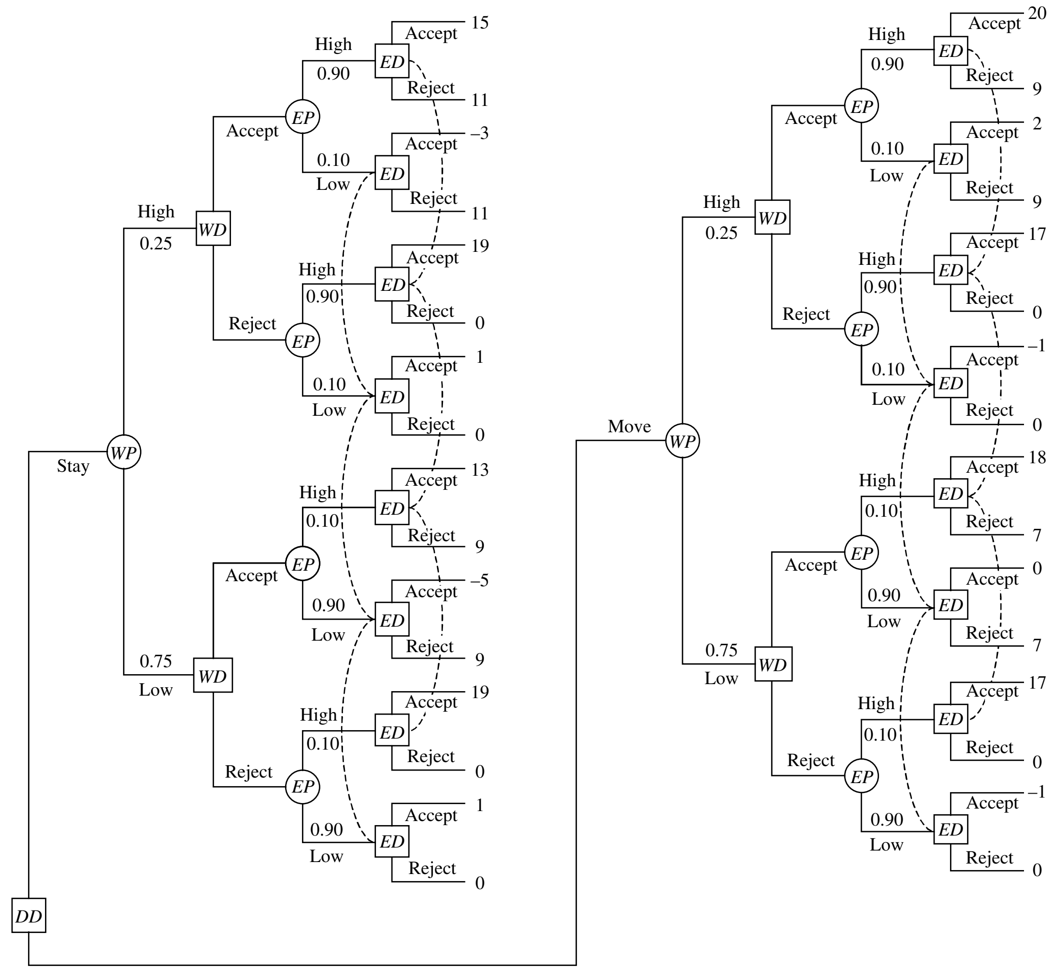
Example 2. Consider the decision tree shown in Figure 1. If interpreted from the standard decision-tree perspective (ignoring the dashed lines), the dock supervisor's decision *DD* consists of two alternatives, i.e., "Stay" or "Move." It is followed by the resolution of the West price *WP*, i.e., "High" or "Low," as well as the West manager's decision *WD*, i.e., "Accept" or "Reject." Note that the West manager knows the results of *DD* and *WP* before making her decision. The resolution of the East price *EP* comes next. It is not known to the dock supervisor and the West manager, but to the East manager. The East manager's decision *ED* would then be made with the knowledge of *DD*, *WP*, *WD*, and *EP*. This interpretation corresponding to a decision situation with perfect recall does not capture the story in Example 1 unless the information sets are added. The leftmost dashed line connecting four decision nodes of the East manager implies that she cannot distinguish among the paths leading to such nodes. In other words, she only knows for certain what all four paths share in common, i.e., $DD =$ "Stay" and $EP =$ "Low." The other dashed lines are interpreted similarly. This augmented decision tree does capture the exact story in Example 1 along with the numerical specifications. □

In addition to the fact that the augmented decision tree grows exponentially with the number of decisions and uncertainties, the tree cannot necessarily be evaluated with the BI algorithm unless it corresponds to certain types of decision situations, as noted earlier. Some possible alternative solution approaches include those that could solve games in general (McKelvey and McLennan 1996, McKelvey et al. 2004), but they are, however, not specifically developed for team decision situations with identical payoff structure. In other words, they do not take advantage of the fact that the identity of the players does not matter.

We propose to represent and evaluate the team decision situation with an influence diagram. In fact, the original influence diagram explicitly allowed the representation of imperfect recall, although perfect recall was assumed for solution (Howard and Matheson 1984/2005). Tatman and Shachter (1990) study the influence diagram with imperfect recall that corresponds to the finite-stage MDP and apply the BI algorithm to solve it. Zhang and Poole (1992) and later Zhang et al. (1994) study the influence diagram with imperfect recall in the general case and find conditions under which the BI algorithm is applicable.

Nilsson and Lauritzen (2000) also study a similar diagram under the name LIMID (LImited Memory

Figure 1 **A Decision Tree Augmented with Information Sets for Example 1**



Influence Diagram) and show more general conditions under which the diagram can be solved with the BI algorithm. They later develop an algorithm that guarantees a local optimal solution when the diagram does not satisfy their established conditions (Lauritzen and Nilsson 2001). Koller and Milch (2001, 2003) introduce MAID (Multi-Agent Influence Diagram), which extends the semantic of the influence diagram to represent and solve a game where every player has perfect recall. Their proposed solution method relies on the solution methods that solve games in general (McKelvey and McLennan 1996).

Section 2 lays the necessary foundations in team decision analysis and influence diagrams. These include formalizing the solution concepts and devising solution methods that accommodate incomplete sharing of information. Section 3 exploits the graphical structure of the influence diagram to develop more insights for decision making in §4 and simplifying the influence diagram in §5. Section 6 develops the diagram-based solution method that uses all available information to improve the decision quality. Finally, §7 demonstrates our results with Example 1, while §8 concludes the research and surveys some other related developments.

2. Foundations

This section introduces our representation for a team decision situation. Because we approach team decision making from an individual perspective, this representation does not specify, and thus is invariant to, the identities of the team members.

2.1. Terminology

A *distinction* is a feature that defines a set of mutually exclusive and collectively exhaustive possibilities, exactly one of which is resolved to be true. It is a *decision* if its resolution is under the complete control of the decision maker. Otherwise, it is an *uncertainty* and we represent the decision maker's belief about its resolution with a conditional probability distribution. A distinction is said to be *observed* if its resolution is known. An *observation* available to a decision is a set of distinctions observed at the time of making the decision. A *strategy* for a decision is a contingent plan of choices for the decision given its available observation, and we represent the decision maker's possible choices for the decision with the set of all possible strategies.

Denote by Y a set of all distinctions in a decision situation where the sets of decisions and uncertainties are denoted by $D \subseteq Y$ and $U = Y - D$, respectively. A distinction $n_1 \in Y$ is said to be irrelevant to a distinction $n_2 \in Y$ if the observation of n_1 does not change the decision maker's belief about the resolution of n_2 regardless of the strategies prescribed for every decision in D. It follows that n_1 is irrelevant to n_2 if and only if n_2 is also irrelevant to n_1. A distinction $n_1 \in Y$ is said to be irrelevant to a distinction $n_2 \in Y$, given a distinction $n_3 \in Y$, if n_1 is irrelevant to n_2, given that n_3 has already been observed.

Each $n \in Y$ is assumed to be finite, and thus we can represent its possibility set by a finite set X_n with x_n as its generic element. Similarly, we represent a possibility set of the subset of distinctions $N \subseteq Y$ by $X_N = \times_{n\in N} X_n$ with x_N as its generic element. An element in X_Y is called a *scenario*. A *prospect* is how the decision maker views the future given a scenario, and we represent the decision maker's preference among prospects with a utility function (von Neumann and Morgenstern 1947).

A decision $d \in D$ is said to have *complete sharing of information* from (*perfect recall* on) the set of decisions $F \subseteq D$ if the observation available to d includes F as well as the observation available to each decision in F. A decision always has complete sharing of information from itself by this definition. We refer to complete sharing of information and perfect recall synonymously. A set of decisions is said to have complete sharing of information if there exists an ordering of those decisions such that each decision has complete sharing of information from all earlier decisions in the set. A set of decisions is said to have incomplete sharing of information otherwise. A decision situation is said to have complete sharing of information if the set of all decisions has complete sharing of information. A decision situation is said to have incomplete sharing of information otherwise.

Finally, a *team* is a set of decision makers, called *team members*, who agree on common preferences among prospects and beliefs about uncertainties, but are, in general, responsible for making different decisions in the decision situation. As a consequence, the terminologies presented above are equally applicable to individual decision makers and teams.

2.2. Influence Diagram

A *directed graph* is defined as a set of nodes Z and a set of directed arcs A between ordered pairs of nodes such that there is at most one arc for each pair. A node $n_1 \in Z$ is a parent of a node $n_2 \in Z$ (n_2 is a child of n_1) if there is an arc directed from n_1 toward n_2. For any $n \in Z$, denote by $\text{pa}(n)$, $\text{ch}(n)$, $\text{nb}(n) = \text{pa}(n) \cup \text{ch}(n)$, and $\text{fa}(n) = \text{pa}(n) \cup n$ the sets of nodes that are parents, children, neighbors, and family of n, respectively. For any $N \subseteq Z$, denote by $\text{fa}(N) = \bigcup_{n\in N} \text{fa}(n)$ and $\text{pa}(N) = \text{fa}(N) - N$ the sets of nodes that are family and strict parents of N, respectively. A subgraph of Z induced by N is a set of nodes N and those arcs in A that are between pairs of nodes in N.

A path of length k between n_0 and n_k is a sequence $(n_i)_0^k$ of distinct nodes such that $n_i \in \text{nb}(n_{i-1})$, $i = 1, \ldots, k$. A node n_i on the path between n_0 and n_k is a head-to-head node if $n_{i-1}, n_{i+1} \in \text{pa}(n_i)$, $i \in \{1, \ldots, k-1\}$. A path is directed from n_0 to n_k if $n_i \in \text{ch}(n_{i-1})$, $i = 1, \ldots, k$. A node n_0 is an ancestor of a node n_k (n_k is a descendant of n_1) if there is a directed path from n_0 to n_k. Denote by $\text{an}(n)$ and $\text{de}(n)$ the sets of nodes that are ancestors and descendants of n, respectively. A cycle is a path between n_0 and n_k with $n_0 = n_k$. It is a directed cycle if the path comprising the cycle is directed. A *directed acyclic graph* is a directed graph with no directed cycle.

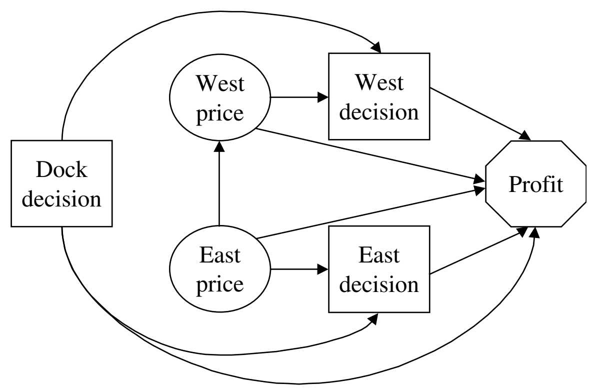
An *influence diagram* is a compact graphical representation of a decision situation. It is a directed acyclic graph with three types of nodes. A *decision node* corresponding to a decision distinction is drawn as a rectangle. An *uncertainty node* corresponding to an uncertainty distinction is drawn as an oval. Finally, a *value node* corresponding to a component of an additively separable utility function is drawn as an octagon. We refer to the distinction and its corresponding node interchangeably. Hence, D is the set of all decision nodes, U is the set of all uncertainty nodes, and V is the set of all value nodes. The set of all distinction (nonvalue) nodes is denoted by $Y = D \cup U$, and the set of all nodes is denoted by $Z = Y \cup V$. Note that an influence diagram is also called a *Bayesian network* or *belief network* if $D = \varnothing$ and $V = \varnothing$.

The arcs take on different meanings depending on the types of node they are directed toward. For any $v \in V$, the arcs directed from $\mathrm{pa}(v)$ into v are called *functional arcs*, and $\mathrm{pa}(v)$ represents the domain of some function v which is itself a component of the additively separable utility function. For any $u \in U$, the arcs directed from $\mathrm{pa}(u)$ into u are called *conditional arcs* and $\mathrm{pa}(u)$ represents the distinctions being conditioned on when assessing the belief about u. In other words, the probability distribution of u is conditioned on the resolution of $\mathrm{pa}(u)$. A special case is when $u \in U$ is a deterministic function of $\mathrm{pa}(u)$, i.e., the distribution of u conditioned on the resolution of $\mathrm{pa}(u)$ is degenerate. Such a u is also called *deterministic uncertainty*, drawn as a double-oval node, and the arcs directed from $\mathrm{pa}(u)$ into u are also called functional arcs. Finally, for any $d \in D$, the arcs directed from $\mathrm{pa}(d)$ into d are called *informational arcs* and $\mathrm{pa}(d)$ represents the observation available to d. In other words, the choice of d can be conditioned on the resolution of $\mathrm{pa}(d)$.

In addition to the acyclicity of the influence diagram, it is generally assumed that value nodes have no children (Shachter 1986). Tatman and Shachter (1990) relax this assumption by introducing the notion of supervalue nodes to represent sums and products of other value and supervalue nodes. In fact, our framework is equivalent to having a sum supervalue node, although it is not shown explicitly. Together with these restrictions, it follows semantically that every distinction node is irrelevant to its nondescendant distinction nodes given its parental nodes whenever the nodes in the influence diagram are connected, such that every arc takes on its appropriate meaning (Howard and Matheson 1984/2005).

Example 3. Consider an influence diagram representing our team decision situation from Example 1

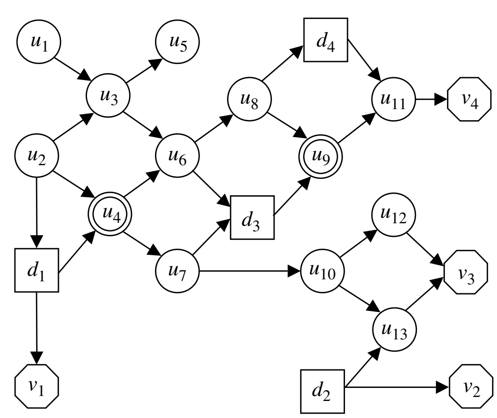
Figure 2 An Influence Diagram for Example 1



shown in Figure 2. In our terminology, the functional arcs directed to *Profit* indicate that it is a function of every distinction. The informational arcs directed to *West Decision* indicate that the West manager will only observe *West Price* and *Dock Decision* before making her choice. Similarly, the informational arcs directed to *East Decision* indicate that the East manager will only observe *East Price* and *Dock Decision* before making her choice. The only conditional arc indicates the team's belief that both prices can be relevant and their relevance is represented by the distribution of *West Price* conditioned on *East Price*. The semantic of the influence diagram also indicates, for example, that *East Price* is irrelevant to *West Decision* given *West Price*. □

Example 4. Consider a more complicated influence diagram (shown in Figure 3) modified from Jensen et al. (1994). It has four decision nodes, thirteen uncertainty nodes (including two deterministic uncertainty

Figure 3 A Complicated Influence Diagram



nodes), and four value nodes. It indicates, for example, that the team can observe u_6 and u_7 at the time of d_3. The team's belief about u_{11} is a probability distribution conditioned on d_4 and u_9, while the team's belief about u_9 is a deterministic function of d_3 and u_8. The team's preference under uncertainty is an additively separable utility function of four components, one of which is a function of u_{12} and u_{13}. We can also conclude, for example, that d_4 is irrelevant to $Y - \text{de}(d_4) = Y - \{u_{11}, v_4\}$, given $\text{pa}(d_4) = u_8$; and that u_{10} is irrelevant to $Y - \text{de}(u_{10}) = Y - \{u_{12}, u_{13}, v_3\}$, given $\text{pa}(u_{10}) = u_7$. □

2.3. Completely Specified Influence Diagram

For any $d \in D$, a strategy $\delta_d(x_{\text{fa}(d)})$ is a conditional probability distribution over X_d, representing the choice made at d, given $x_{\text{pa}(d)} \in X_{\text{pa}(d)}$, representing the observation available to d. Two notable cases are the pure strategy, which is a degenerate distribution over X_d given $x_{\text{pa}(d)} \in X_{\text{pa}(d)}$; and the uniform strategy, which is a uniform distribution over X_d given $x_{\text{pa}(d)} \in X_{\text{pa}(d)}$. We represent the set of all possible pure strategies by Δ_d^{P} and the set of all possible strategies by Δ_d. For any $F \subseteq D$, a strategy set δ_F is the set $\{\delta_d\colon d \in F\}$, and we represent the set of all possible pure-strategy sets by $\Delta_F^{\text{P}} = \bigtimes_{d\in F} \Delta_d^{\text{P}}$ and the set of all possible strategy sets by Δ_F.

For any $u \in U$, a distribution $\phi_u(x_{\text{fa}(u)})$ is a conditional probability distribution over X_u, representing the team's belief about u, given $x_{\text{pa}(u)} \in X_{\text{pa}(u)}$, representing the distinctions being conditioned on when assessing such belief. We represent the set of all possible distributions by Φ_u. For any $R \subseteq U$, a distribution set ϕ_R is the set $\{\phi_u\colon u \in R\}$, and we represent the set of all possible distribution sets by Φ_R.

For any $v \in V$, a utility $\psi_v(x_{\text{pa}(v)})$ is a function, representing a single component of the additively separable utility function, that assigns one real number to each $x_{\text{pa}(v)} \in X_{\text{pa}(v)}$. We represent the set of all possible utilities by Ψ_v. For any $W \subseteq V$, a utility set ψ_W is the set $\{\psi_v\colon v \in W\}$, and we represent the set of all possible utility sets by Ψ_W.

A *completely specified influence diagram* includes the set of all possible strategy sets Δ_D, the distribution set ϕ_U such that $\prod_{d\in D} \delta_d(x_{\text{fa}(d)}) \prod_{u\in U} \phi_u(x_{\text{fa}(u)})$ is a joint probability distribution representing the team's belief over X_Y for every $\delta_D \in \Delta_D$, and the utility set ψ_V such that $\sum_{v\in V} \psi_v(x_{\text{pa}(v)})$ is a utility function representing the team's preference over X_Y under uncertainty. These three elements (Δ_D, ϕ_U, and ψ_V) are known as the *decision basis*, a quantitative specification of a decision situation in terms of the alternatives, the information, and the preferences of the team (Howard 1984).

2.4. Algebra in an Influence Diagram

For any $N \subseteq Z$, a *probability potential* for the influence diagram is defined as

$$\mathcal{P}_N(x_{\text{fa}(N\cap Y)}) = \prod_{d\in N\cap D} \delta_d(x_{\text{fa}(d)}) \prod_{u\in N\cap U} \phi_u(x_{\text{fa}(u)}), \tag{1}$$

and for any $M \subseteq N$, a *uniformly extended probability potential* is defined as

$$\mathcal{P}_{N*M}(x_{\text{fa}(N\cap Y)}) = \mathcal{P}_{N-M}(x_{\text{fa}(\{N-M\}\cap Y)}) \prod_{d\in M\cap D} \bar{\delta}_d(x_{\text{fa}(d)}) \cdot \prod_{u\in M\cap U} \bar{\phi}_u(x_{\text{fa}(u)}), \tag{2}$$

where $\bar{\delta}_d$ and $\bar{\phi}_u$ are the uniform strategy and the uniform distribution, respectively. A *utility potential* for the influence diagram is defined as

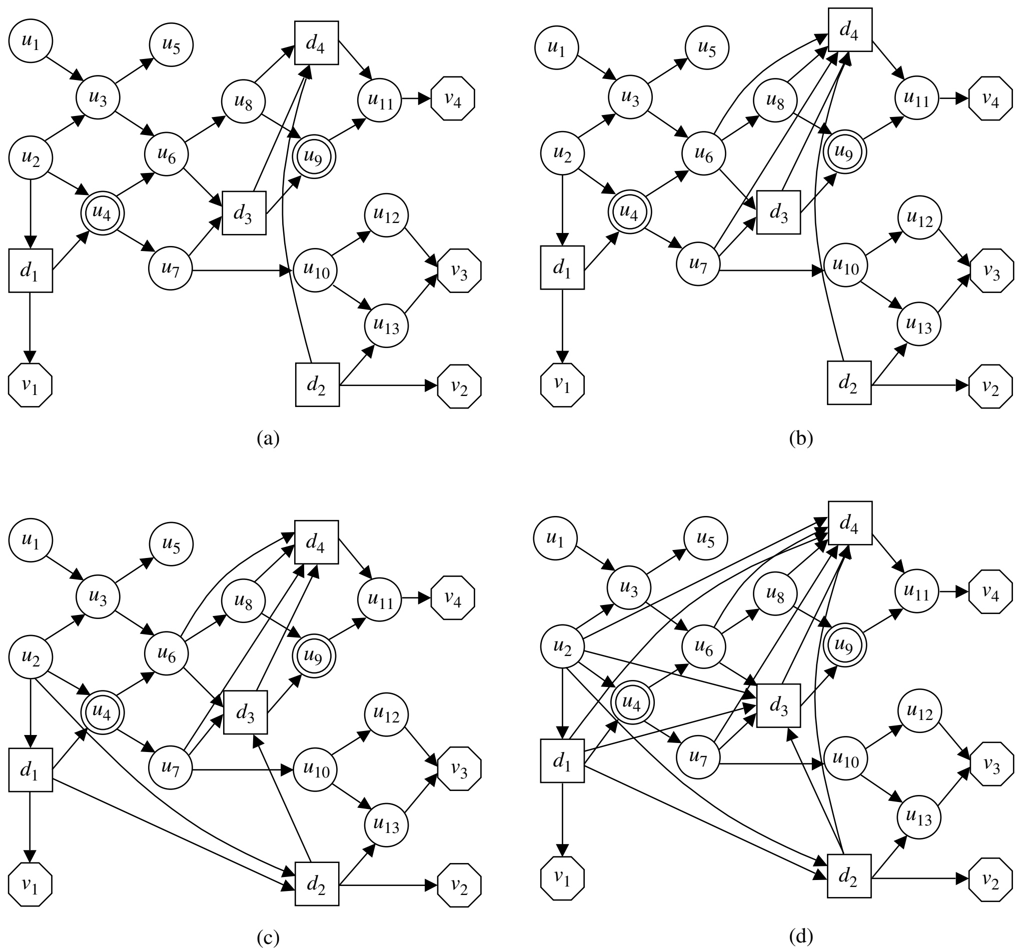
$$\mathcal{U}_N(x_{\text{pa}(N\cap V)}) = \sum_{v\in N\cap V} \psi_v(x_{\text{pa}(v)}). \tag{3}$$

A completely specified influence diagram includes Δ_D, ϕ_U, and ψ_V. ψ_V induces a utility function $\mathcal{U}_Z(x_{\text{pa}(V)})$, while a particular $\delta_D \in \Delta_D$ and ϕ_U induce a joint probability distribution $\mathcal{P}_Z(x_{\text{fa}(Y)})$. A joint probability distribution over $N \subseteq Y$ is defined as $\mathcal{P}_Z(x_N) = \sum_{x_{\text{fa}(Y)-N}} \mathcal{P}_Z(x_{\text{fa}(Y)})$. For any $N_1, N_2 \subseteq Y$ such that $\mathcal{P}_Z(x_{N_1}) > 0$, a conditional probability distribution is determined according to $\mathcal{P}_Z(x_{N_2} \mid x_{N_1}) = \mathcal{P}_Z(x_{N_1\cup N_2})/\mathcal{P}_Z(x_{N_1})$. Finally, the expected utility is written as $\text{EU}(\delta_D, \phi_U, \psi_V) = \sum_{x_{\text{fa}(Z)}} \mathcal{P}_Z(x_{\text{fa}(Y)})\mathcal{U}_Z(x_{\text{pa}(V)})$, or simply $\text{EU}(\delta_D)$, whenever it is clear in the context. In our framework, we do not need to marginalize over any value nodes.

2.5. Complete Sharing of Information in an Influence Diagram

For any $d \in D$, denote by $\text{cs}(d) = \{n \in D\colon \text{fa}(n) \subseteq \text{fa}(d)\}$ a set of decision nodes from which d has complete sharing of information. A subset $F \subseteq D$ has complete sharing of information when there exists an ordering

Figure 4 Additional Influence Diagrams



of $\{d_1, \ldots, d_k\} = F$ such that $\{d_1, \ldots, d_i\} \subseteq \text{cs}(d_i)$, $i = 1, \ldots, k$. By this definition, every subset of decisions that have complete sharing of information also has complete sharing of information, and so does every decision. An influence diagram in which D has complete sharing of information (perfect recall) is called a *decision network* (Howard and Matheson 1984/2005).

Example 5. Each decision in the influence diagram shown in Figure 3 only has complete sharing of information from itself. Note that additional observations do not imply additional complete sharing of information. The influence diagram shown in Figure 4a illustrates that d_4 only has complete sharing of information from d_2, but not d_3, because $\text{fa}(d_2) = d_2$, but not $\text{fa}(d_3) = \{d_3, u_6, u_7\}$ are observed at the time of d_4. The influence diagram shown in Figure 4b illustrates that d_4 now has complete sharing of information from d_3 because $\text{fa}(d_3)$ is now observed at the time of d_4. Although the influence diagram shown in Figure 4c illustrates that each decision has complete sharing of information from itself and the preceding decision, D does not have complete sharing of information by definition. Finally, the influence diagram shown in Figure 4d illustrates the influence diagram of a decision situation with complete sharing of information, i.e., every set of decisions has complete sharing of information. □

2.6. General Solution Concepts

According to the standard criterion in decision making under uncertainty, the team should commit to an optimal strategy set δ_D, one that maximizes the expected utility.

Definition 1. A strategy set δ_D is *optimal* when

$$\mathrm{EU}(\delta_D) \geq \mathrm{EU}(\tilde{\delta}_D) \quad \text{for } \forall\, \tilde{\delta}_D \in \Delta_D.$$

The following proposition proves the existence of an optimal strategy set that is pure.

Proposition 1. *There exists an optimal strategy set δ_D that is a pure-strategy set.*

Proof. With the utility function over the finite set $X_{\mathrm{pa}(V)}$, the expected utility of any pure-strategy set is an expected value of finitely many numbers, which is finite. Thus, there exists at least one maximal expected utility pure-strategy set $\delta_D \in \Delta_D^{\mathrm{P}}$. As any strategy set is defined as a set of conditional probability distributions, its expected utility cannot be greater than the expected utility of δ_D, a set of degenerate conditional probability distributions attaining maximum expected utility. □

Analogous to the concepts of global and local optimality in optimization, we introduce the notion of stability that precisely characterizes the scope of any local optimality.

Definition 2. A strategy set δ_F, $F \subseteq D$, is *k-stable*, $k \in \{1, \ldots, |F|\}$, with respect to strategy set δ_D, or δ_D is k-stable at F when

$$\delta_f = \arg\max_{\tilde{\delta}_f} \mathrm{EU}(\tilde{\delta}_f, \delta_{D-f})$$

for $\forall\, f \subseteq F$ such that $|f| = k$. When $k = |F|$, δ_F is also said to be *maximally stable* with respect to δ_D.

Because a maximally stable δ_F with respect to δ_D can be viewed as the optimal δ_D with fixed δ_{D-F}, its existence also follows from the proof in Proposition 1. A special case of this definition is when $F = D$ in which optimality coincides with maximal stability of δ_D. Another special case is when there are j team members responsible for distinct sets of decisions in a team decision situation. When each team member's strategy set δ_{D_i}, $i = 1, \ldots, j$, is maximally stable with respect to δ_D, the team strategy set δ_D is said to be a *Person-by-Person Satisfactory* (PBPS) team decision rule in team theory (Marschak and Radner 1972). It is also equivalent to a Nash Equilibrium (NE) obtained when we approach team decision making as a game where every player has identical payoffs (Nash 1951). There are two properties of stability worth mentioning.

Proposition 2. *If a strategy set δ_F, $F \subseteq D$, is k-stable, $k \in \{1, \ldots, |F|\}$, with respect to strategy set δ_D, δ_F is also m-stable, $m \in \{1, \ldots, k\}$, with respect to δ_D.*

Proof. Let $\delta_F = \delta_f \cup \delta_{F-f}$ for any $f \subseteq F$ such that $|f| = k$. Let $\delta_f = \delta_{f_i} \cup \delta_{f-f_i}$ for any $f_i \in f$. According to the definition of k-stability, we have

$$\begin{aligned} \mathrm{EU}(\delta_{f-f_i}, \delta_{f_i}, \delta_{D-f}) &= \max_{\tilde{\delta}_{f-f_i}, \tilde{\delta}_{f_i}} \mathrm{EU}(\tilde{\delta}_{f-f_i}, \tilde{\delta}_{f_i}, \delta_{D-f}) \\ &\geq \max_{\tilde{\delta}_{f-f_i}} \mathrm{EU}(\tilde{\delta}_{f-f_i}, \delta_{f_i}, \delta_{D-f}) \\ &= \mathrm{EU}(\delta_{f-f_i}, \delta_{f_i}, \delta_{D-f}). \end{aligned}$$

As this is true for $\forall\, f_i \in f$ and $\forall\, f \subseteq F$ by the definition of k-stability, δ_F is $(k-1)$-stable with respect to δ_D. The result follows by induction. □

Proposition 3. *If a strategy set δ_F, $F \subseteq D$, is the unique m-stable strategy set with respect to strategy set δ_D, $m \in \{1, \ldots, |F|\}$, δ_F is also maximally stable with respect to δ_D.*

Proof. There exists some δ'_F that is maximally stable that must also be m-stable by Proposition 2. If there is a unique m-stable strategy set, then $\delta_F = \delta'_F$. □

2.7. General Solution Methods

Although an optimal strategy set can always be found through the BI given any decision situation with complete sharing of information (perfect recall), it might not be applicable to a decision situation with incomplete sharing of information (imperfect recall). An approach that always guarantees an optimal δ_D is the *Exhaustive Enumeration* of all pure-strategy sets in the *enumeration set* Δ_D^{P} based on the results of Proposition 1 and the finiteness of Δ_D^{P}. Unfortunately, it is usually impractical because of the size of Δ_D^{P}, which is equal to $\prod_{d\in D} |\Delta_d^{\mathrm{P}}|$, where $|\Delta_d^{\mathrm{P}}| = |X_d|^{\prod_{e\in \mathrm{pa}(d)} |X_e|}$ such that $\prod_{e\in \mathrm{pa}(d)} |X_e| = 1$ if $\mathrm{pa}(d) = \varnothing$.

To reduce the computation of exhaustive enumeration, we can generalize the above method by enumerating the strategy sets in multiple smaller enumeration sets associated with a partition of the decisions. This results in an iterative process, called *Strategy Improvement* (SI).

Algorithm 1. *Strategy Improvement*.

Input: A completely specified influence diagram with a partition G of a subset $F \subseteq D$.

Output: A pure-strategy set δ_D that is maximally stable at g, $\forall g \in G$.

Initialization

1. Assign some initial strategy set δ_D.

Iteration

1. Assign $\delta'_F = \delta_F$.
2. For each enumeration set $g \in G$, do:
   (a) If $\delta'_g \neq \arg\max_{\tilde{\delta}_g} \mathrm{EU}(\tilde{\delta}_g, \delta_{D-g})$, update $\delta_g = \arg\max_{\tilde{\delta}_g} \mathrm{EU}(\tilde{\delta}_g, \delta_{D-g})$.
3. Repeat the iteration until $\delta'_F = \delta_F$.

The following proposition proves that the algorithm converges to a δ_D that is maximally stable at every subset of decisions that has an associated enumeration set.

Proposition 4. *A pure-strategy set δ_D obtained with the SI algorithm is maximally stable at every subset of decisions that has an associated enumeration set.*

Proof. At each iteration of the algorithm, we always update δ_g to the one that is maximally stable with respect to δ_D unless it is already so. If the algorithm stops, δ_D must be maximally stable at g, $\forall g \in G$. As the expected utility is either increasing or unchanged during the algorithm, and there is only a finite number of pure-strategy sets that will not be repeated unless they have the same expected utility, $\delta'_F = \delta_F$ eventually holds and the algorithm stops successfully. □

Exhaustive enumeration can be viewed as a special case of this algorithm when there is a single enumeration set associated with D. Another special case in which each enumeration set is associated with a distinct decision is called *Single Strategy Improvement* (SSI), and when $F = D$ *Single Policy Updating* (SPU) (Lauritzen and Nilsson 2001). By enumerating only one decision at a time, Proposition 4 only guarantees that the resulting strategy set from the SSI (or SPU) algorithm is 1-stable. However, for a decision situation with complete sharing of information (perfect recall), it is well known that the BI algorithm can find the optimal strategy set by enumerating each decision exactly once. In fact, it is even possible to have an incomplete sharing of information (imperfect recall) and yet ensures that BI will find an optimal strategy set, also by enumerating each decision exactly once.

3. Exploitation of Influence Diagram Structure

This section introduces some important semantic implications of the influence diagram that can be verified from its graphical structure and will lead to simplifying transformations.

3.1. Irrelevant Sets and Requisite Sets

We have earlier defined the notion of irrelevance with respect to a particular distribution set for the team. We now enrich this notion by defining it with respect to any distribution set representable by the graphical structure of the influence diagram.

Definition 3. For any $N_1, N_2, N_3 \subseteq Y$, N_3 is said to be an *irrelevant set* for N_1 given N_2, denoted by $N_1 \perp N_3 \mid N_2$, if $\mathcal{P}_Z(x_{N_1} \mid x_{N_2}) = \mathcal{P}_Z(x_{N_1} \mid x_{N_2 \cup N_3})$, $\forall \delta_D \in \Delta_D$, and $\forall \phi_U \in \Phi_U$.

The following notions of requisite distinctions, observations, and values build on the notion of irrelevance. For any $N \subseteq Z$, denote by $N' = Z - N$ a set of nodes that are not in N.

Definition 4. For any $N_1, N_2, N_3 \subseteq Y$, N_3 is said to be a *requisite distinction set* for N_1 given N_2 if it is a minimal set, such that $\mathcal{P}_Z(x_{N_1} \mid x_{N_2}) = \mathcal{P}_{Z*N'_3}(x_{N_1} \mid x_{N_2})$, $\forall \delta_D \in \Delta_D$, and $\forall \phi_U \in \Phi_U$.

In other words, a requisite distinction set N_3 is a minimal subset of Y such that its potential is required in order to properly compute $\mathcal{P}_Z(x_{N_1} \mid x_{N_2})$, $\forall \delta_D \in \Delta_D$, and $\forall \phi_U \in \Phi_U$. An immediate consequence of this definition is that $\mathcal{P}_Z(x_{N_1} \mid x_{N_2})$ is invariant to $\delta_{N'_3 \cap D}$ and $\phi_{N'_3 \cap U}$, and thus can also be determined with $\mathcal{P}_{Z*N'_3}(x_{N_2}) > 0$ in the case that $\mathcal{P}_Z(x_{N_2}) = 0$.

Definition 5. For any $N_1, N_2, N_3 \subseteq Y$ such that $N_3 \subseteq N_2$, N_3 is said to be a *requisite observation set* for N_1 given N_2 if it is a minimal set such that $\mathcal{P}_Z(x_{N_1} \mid x_{N_2}) = \mathcal{P}_Z(x_{N_1} \mid x_{N_3})$, $\forall \delta_D \in \Delta_D$, and $\forall \phi_U \in \Phi_U$.

In other words, a requisite observation set N_3 is a minimal subset of N_2 that is required to be observed in order to properly compute $\mathcal{P}_Z(x_{N_1} \mid x_{N_2})$, $\forall \delta_D \in \Delta_D$, and $\forall \phi_U \in \Phi_U$. It follows that N_2 is an irrelevant set for N_1 given N_3. We establish the following proposition that relates the requisite distinction set with the requisite observation set for later use.

Proposition 5. *For any $N_1, N_2, N_3 \subseteq Y$ such that $N_3 \subseteq N_2$ contains a requisite observation set for N_1 given N_2, a requisite distinction set for N_1 given N_2 is equal to a requisite distinction set for N_1 given N_3.*

Proof. Denote by N and $N' = Z - N$ a requisite distinction set for N_1 given N_2 and its complement, respectively. We have that $\mathcal{P}_Z(x_{N_1} \mid x_{N_2}) = \mathcal{P}_{Z*N'}(x_{N_1} \mid x_{N_2}) = \mathcal{P}_{Z*N'}(x_{N_1} \mid x_{N_3}) = \mathcal{P}_Z(x_{N_1} \mid x_{N_3})$, $\forall \delta_D \in \Delta_D$, and $\forall \phi_U \in \Phi_U$, by Definitions 4 and 5. This implies that N is also a requisite distinction set for N_1 given N_3. The result follows, as the proof can also be established in the other direction. □

We can further enrich the notion of requisiteness to include the value nodes. This is accomplished by defining it with respect to any strategy set, distribution set, as well as utility set the team might have as long as they are representable by the graphical structure of the influence diagram. For any $W \subseteq V$, denote by $W' = V - W$ a set of value nodes that are not in W.

Definition 6. For any $N_1, N_2 \subseteq Y$ and $W \subseteq V$, W is said to be a *requisite value set* for N_1 given N_2 if it is a minimal set such that $\mathcal{P}_Z(x_{\text{pa}(W')} \mid x_{N_1 \cup N_2}) \mathcal{U}_{W'}(x_{\text{pa}(W')}) = \mathcal{P}_Z(x_{\text{pa}(W')} \mid x_{N_2}) \mathcal{U}_{W'}(x_{\text{pa}(W')})$, $\forall \delta_D \in \Delta_D$, $\forall \phi_U \in \Phi_U$, and $\forall \psi_V \in \Psi_V$.

As ψ_V can be an arbitrary set of real numbers, the above definition can also be stated as $\mathcal{P}_Z(x_{\text{pa}(W')} \mid x_{N_1 \cup N_2}) = \mathcal{P}_Z(x_{\text{pa}(W')} \mid x_{N_2})$, $\forall \delta_D \in \Delta_D$, and $\forall \phi_U \in \Phi_U$. This is equivalent to N_1 being an irrelevant set for $\text{pa}(W')$ given N_2. In other words, a nonrequisite value set W' is a maximal subset of V such that its parental set $\text{pa}(W')$ is irrelevant to N_1 given N_2.

3.2. Bayes-Ball

Based on the semantic relationships among the nodes in the influence diagram, the concept of d-separation and its deterministic generalization, D-separation, can express the existing irrelevance among any sets of nodes (Pearl 1988, Geiger et al. 1990). For any $N_1, N_2, N_3 \subseteq Y$, an *active path* between N_1 and N_3 given N_2 is a path between $n_1 \in N_1$ and $n_3 \in N_3$ such that every head-to-head node on the path is or has a descendant in N_2, and every other node on the path is not functionally determined by N_2. N_2 is said to *D-separate* N_1 and N_3 if there is no active path between N_1 and N_3 given N_2. We establish the following proposition by Shachter (1998), which relates the notions of D-separation and irrelevance for later use.

Proposition 6. *For any $N_1, N_2, N_3 \subseteq Y$, N_2 D-separates N_1 and N_3 if and only if $N_1 \perp N_3 \mid N_2$.*

For any $N_1, N_2 \subseteq Y$, the Bayes-Ball algorithm (Shachter 1998, 1999) applies the above concepts to determine the irrelevant, requisite observation, and requisite distinction sets for N_1 given N_2 (with respect to $\mathcal{P}_Z(x_{N_1} \mid x_{N_2})$) in linear time in the size of the influence diagram, treating decisions as uncertainties. The irrelevant set is a set of nodes in Y that are not marked on the bottom, the requisite observation set is a set of nodes that are checked, and the requisite distinction set is a set of nodes that are marked on the top. See Figure 5 for illustration.

4. Implications on Decision Making

This section applies the semantic implications in the previous section to develop the insights for decision making at any individual decision in the influence diagram.

4.1. Key Perspective

According to the SI algorithm, a strategy set δ_D that is 1-stable is computed by iteratively finding a pure strategy δ_d, $\forall d \in D$, that is 1-stable with respect to δ_D. In other words, each δ_d is a degenerate conditional probability distribution that satisfies the following lemma by Nilsson and Lauritzen (1999).

Lemma 1. *A strategy δ_d, $d \in D$, is 1-stable with respect to strategy set δ_D if and only if for each $x_{\text{pa}(d)} \in X_{\text{pa}(d)}$, δ_d assigns positive mass only to $x_d \in X_d$ that satisfies*

$$x_d = \arg\max_{\tilde{x}_d} \mathcal{V}_1(\tilde{x}_d, x_{\text{pa}(d)})$$

where $\mathcal{V}_1(x_{\text{fa}(d)}) = \sum_{x_{\text{fa}(Z)-\text{fa}(d)}} \mathcal{P}_{Z*d}(x_{\text{fa}(Y)}) \mathcal{U}_Z(x_{\text{pa}(V)})$.

We will state a set of related lemmas that build on the above lemma. Each of them applies the results from the earlier section to exploit the graphical structure of the influence diagram and yields additional insights in the choice of δ_d that is 1-stable with respect to δ_D. An example based on the influence diagram shown in Figure 4b will be provided.

4.2. Requisite Values

The first insight comes from the fact that certain decisions might have no influence on some aspects of prospect valuation. In other words, some value nodes might be nonrequisite to some decision nodes. For any $d \in D$, denote by $\text{rqv}(d)$ a requisite value set for d given $\text{pa}(d)$ as defined in Definition 6.

Lemma 2. *A strategy δ_d, $d \in D$, is 1-stable with respect to strategy set δ_D if and only if for each $x_{\mathrm{pa}(d)} \in X_{\mathrm{pa}(d)}$, δ_d assigns positive mass only to $x_d \in X_d$ that satisfies*

$$x_d = \arg\max_{\tilde{x}_d} \mathcal{V}_2(\tilde{x}_d, x_{\mathrm{pa}(d)})$$

where

$$\mathcal{V}_2(x_{\mathrm{fa}(d)}) = \sum_{x_{\mathrm{fa}(Z)-\mathrm{fa}(d)}} \mathcal{P}_{Z*d}(x_{\mathrm{fa}(Y)}) \mathcal{U}_N(x_{\mathrm{pa}(N\cap V)})$$

and $N \supseteq \mathrm{rqv}(d)$.

Proof. See the appendix. □

Lemma 2 implies that the choice of δ_d that is 1-stable with respect to δ_D is invariant to the utility set $\psi_{V-\mathrm{rqv}(d)}$. In other words, the team can properly make the choice of δ_d by considering only value nodes in $\mathrm{rqv}(d)$. An immediate consequence of Lemma 2 is that every $\delta_d \in \Delta_d$ is 1-stable with respect to δ_D if $\mathrm{rqv}(d) = \varnothing$. According to Proposition 6, the requisite value set for d given $\mathrm{pa}(d)$ is equal to $\mathrm{de}(d) \cap V$.

4.3. Requisite Observations

The second insight comes from the fact that some observations might be redundant for the purpose of finding δ_d that is 1-stable with respect to δ_D. For any $d \in D$, denote by $\mathrm{rqo}(d)$ a requisite observation set for $\mathrm{pa}(\mathrm{rqv}(d))$ given $\mathrm{fa}(d)$, excluding d itself, as defined in Definition 5. Because Lemma 2 implies that every $\delta_d \in \Delta_d$ is 1-stable with respect to δ_D if $\mathrm{rqv}(d) = \varnothing$, we only need to consider the case that $\mathrm{rqv}(d) \neq \varnothing$ in the following lemma.

Lemma 3. *A strategy δ_d, $d \in D$, is 1-stable with respect to strategy set δ_D if for each $x_{\mathrm{pa}(d)} \in X_{\mathrm{pa}(d)}$, δ_d assigns positive mass only to $x_d \in X_d$ that satisfies*

$$x_d = \arg\max_{\tilde{x}_d} \mathcal{V}_3(\tilde{x}_d, x_{\mathrm{pa}(d)})$$

where

$$\mathcal{V}_3(x_{\mathrm{fa}(d)}) = \sum_{x_{\mathrm{pa}(\mathrm{rqv}(d))-\mathrm{fa}(d)}} \mathcal{P}_{Z*d}(x_{\mathrm{pa}(\mathrm{rqv}(d))} \mid x_N) \mathcal{U}_{\mathrm{rqv}(d)}(x_{\mathrm{pa}(\mathrm{rqv}(d))})$$

and

$$\mathrm{fa}(d) \supseteq N \supseteq \mathrm{rqo}(d) \cup d.$$

Proof. See the appendix. □

Lemma 3 implies that the choice of δ_d that is 1-stable with respect to δ_D does not need to condition on any nonrequisite observations in $\mathrm{pa}(d) - \mathrm{rqo}(d)$. In other words, the informational arcs from $\mathrm{pa}(d) - \mathrm{rqo}(d)$ to d can be removed without changing the team's choice of δ_d. In the case that $\mathrm{rqv}(d) = \varnothing$, we assign $\mathrm{rqo}(d) = \varnothing$ without loss of generality.

4.4. Requisite Distinctions

The third insight comes from the fact that the strategy set of some decisions, distribution set of some uncertainties, and utility set of some values might be redundant for the purpose of finding a δ_d that is 1-stable with respect to δ_D. For any $d \in D$, denote by $\mathrm{rqd}(d)$ a requisite distinction set for $\mathrm{pa}(\mathrm{rqv}(d))$ given $\mathrm{fa}(d)$ as defined in Definition 4. Note that $\mathrm{fa}(d)$ can be derived from the influence diagram before or after the removal of the nonrequisite informational arcs to d. This is because Proposition 5 implies that $\mathrm{rqd}(d)$ is invariant to the given (observed) N, $\mathrm{fa}(d) \supseteq N \supseteq \mathrm{rqo}(d) \cup d$. Again, we only need to consider the case that $\mathrm{rqv}(d) \neq \varnothing$ in the following lemma.

Lemma 4. *A strategy δ_d, $d \in D$, is 1-stable with respect to strategy set δ_D if for each $x_{\mathrm{pa}(d)} \in X_{\mathrm{pa}(d)}$, δ_d assigns positive mass only to $x_d \in X_d$ that satisfies*

$$x_d = \arg\max_{\tilde{x}_d} \mathcal{V}_4(\tilde{x}_d, x_{\mathrm{pa}(d)})$$

where

$$\mathcal{V}_4(x_{\mathrm{fa}(d)}) = \sum_{x_{\mathrm{fa}(N_1\cup N_2)-\mathrm{fa}(d)}} \mathcal{P}_{N_1*d}(x_{\mathrm{fa}(N_1\cap Y)}) \mathcal{U}_{N_2}(x_{\mathrm{pa}(N_2\cap V)}),$$

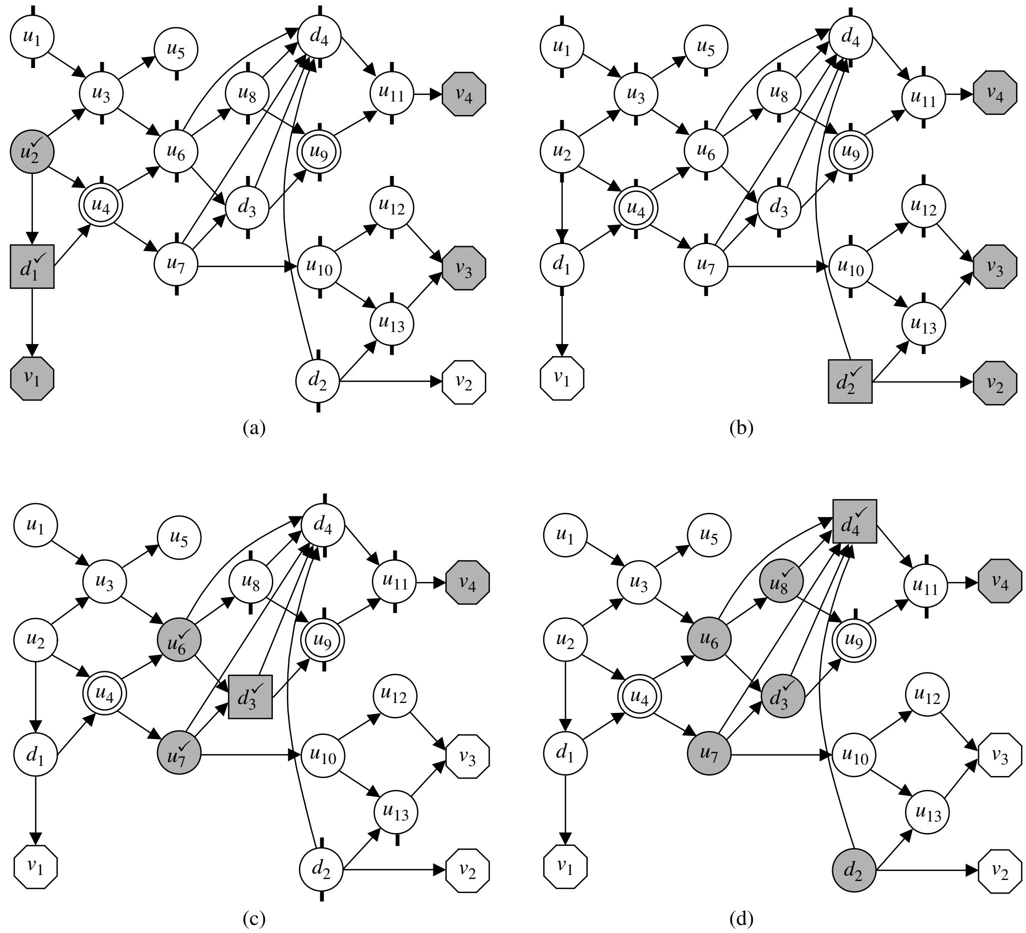
$$N_1 \supseteq \mathrm{rqd}(d) \cup d \quad \textit{and} \quad N_2 \supseteq \mathrm{rqv}(d).$$

Proof. See the appendix. □

Lemma 4 implies that the choice of δ_d that is 1-stable with respect to δ_D can be found once at least $\delta_{\mathrm{rqd}(d)\cap D}$, $\phi_{\mathrm{rqd}(d)\cap U}$, and $\psi_{\mathrm{rqv}(d)}$ are given. In the case that $\mathrm{rqv}(d) = \varnothing$, we assign $\mathrm{rqd}(d) = \varnothing$ without loss of generality.

Note that Lemma 4 in no way implies that the δ_d that is 1-stable with respect to δ_D is invariant to $\delta_{D-\mathrm{rqd}(d)}$ and $\phi_{U-\mathrm{rqd}(d)}$. A simple counterexample is with distinct $N_1, N_2 \supseteq \mathrm{rqd}(d) \cup d$ such that $\mathcal{P}_{N_1*d}(x_{\mathrm{fa}(d)}) = 0$ but $\mathcal{P}_{N_2*d}(x_{\mathrm{fa}(d)}) > 0$, for some $x_{\mathrm{pa}(d)} \in X_{\mathrm{pa}(d)}$. Lemma 4 implies that $\mathcal{P}_{N_1*d}(x_{\mathrm{fa}(d)})$ induces

Figure 5 Irrelevant, Requisite Value, Requisite Observation, and Requisite Distinction Sets in Influence Diagrams



$\mathcal{V}_4(x_{\text{fa}(d)}) = 0$ and every $x_d \in X_d$ is equally preferred. However, $\mathcal{P}_{N_2 * d}(x_{\text{fa}(d)})$ might induce $\mathcal{V}_4(x_{\text{fa}(d)}) \neq 0$ and each $x_d \in X_d$ might be valued differently.

Example 6. Consider the influence diagrams shown in Figure 5. For the purpose of decision making at d_i, $i = 1, \ldots, 4$, $\text{fa}(d_i)$ are shaded (observed) to signify the fact that d_i has complete sharing of information from itself. From the Bayes-Ball algorithm and Lemmas 2 to 4, we have that $\text{rqv}(d_i) = \text{de}(d_i) \cap V$ is the set of value nodes that are shaded, $\text{rqo}(d_i)$ is the subset of nodes in $\text{pa}(d_i)$ that are checked, and $\text{rqd}(d_i)$ is the set of nodes that are marked on the top. The influence diagram shown in Figure 5a illustrates that $\text{rqv}(d_1) = V - v_2$, $\text{rqo}(d_1) = \text{pa}(d_1) = u_2$, and $\text{rqd}(d_1) = Y - \{d_1, u_2, u_5\}$. The influence diagram shown in Figure 5b illustrates that $\text{rqv}(d_2) = V - v_1$, $\text{rqo}(d_2) = \text{pa}(d_2) = \varnothing$, and $\text{rqd}(d_2) = Y - \{d_2, u_5\}$. The influence diagram shown in Figure 5c illustrates that $\text{rqv}(d_3) = v_4$, $\text{rqo}(d_3) = \text{pa}(d_3) = \{u_6, u_7\}$, and $\text{rqd}(d_3) = \{d_2, d_4, u_8, u_9, u_{11}\}$. Finally, the influence diagram shown in Figure 5d illustrates that $\text{rqv}(d_4) = v_4$, $\text{rqo}(d_4) = \{d_3, u_8\} \neq \text{pa}(d_4)$, and $\text{rqd}(d_4) = \{u_9, u_{11}\}$. □

The notions of requisite value and requisite observation sets for decision making have been studied earlier under different names for both the influence diagram with a single value node (Shachter 1988, Fagiuoli and Zaffalon 1998) and with multiple value nodes (Shachter 1998, 1999; Zhang 1998; Nielsen and Jensen 1999). However, their definitions are order dependent, i.e., the requisite value and the requisite observation sets for a decision are derived after the graphical structure of the influence diagram has been modified to reflect the requisite value and the requisite observation sets of later decisions. Our notions are similar to those used in LIMID (Nilsson and Lauritzen 2000, Lauritzen and Nilsson 2001), which generalizes those used in Zhang and Poole (1992) as well as Zhang et al. (1994). We further generalize their notions by recognizing the distinction between an uncertainty and a deterministic uncertainty, and thus treating them differently. The notion of requisite distinction set is related to that of *barren nodes*, a set of distinction nodes that either have no children or whose children are all barren (Shachter 1986).

5. Simplification of Influence Diagram Structure

This section applies the insights developed in the previous section to simplify the graphical structure of the influence diagram while maintaining the decision quality supported by the original influence diagram.

5.1. Requisite Influence Diagram and Requisite Decision Basis

One recommended structural change prior to the evaluation is based on the result of Lemma 3, i.e., the removal of the nonrequisite informational arcs from $\text{pa}(d) - \text{rqo}(d)$ to d, $\forall\, d \in D$. At any $d \in D$, the immediate benefit is an increase in the efficiency of the solution methods once δ_d is conditioned on fewer nodes. Additional benefits are the consequences of the following lemma.

Lemma 5. *For any $d \in D$, the requisite value set $\text{rqv}(d)$, the requisite observation set $\text{rqo}(d)$, and the requisite distinction set $\text{rqd}(d)$ cannot get bigger with the removal of the nonrequisite informational arcs.*

Proof. There are two distinct cases to consider:

(1) The informational arcs are directed to d: Because $\text{pa}(d) \cap \text{de}(d) = \varnothing$, we have that $\text{rqv}(d) = \text{de}(d) \cap V$ remains unchanged. We also have that $\text{rqo}(d)$ remains unchanged by Lemma 3. Finally, we have that $\text{rqd}(d)$, which are marked on the top, remain unchanged because the Bayes-Ball algorithm does not visit the nonrequisite observations when determining $\text{rqd}(d)$ in the first place.

(2) The informational arcs are directed to $D - d$: We have that $\text{rqv}(d) = \text{de}(d) \cap V$ cannot get bigger because the removal of arcs can never increase $\text{de}(d)$. We also have that $\text{rqo}(d)$ and $\text{rqd}(d)$ cannot get bigger because the removal of arcs reduces the number of possible paths among the nodes and the Bayes-Ball algorithm marks the node only if it can be visited from its neighboring nodes.

Note that the result does not necessarily hold with the removal of arbitrary arcs, e.g., the informational arcs from requisite observation nodes. This is because such a node will be marked on the top if it were not marked in the first place according to the Bayes-Ball algorithm. □

Lemma 5 implies that by removing the nonrequisite informational arcs to d, there is a higher efficiency if the strategy δ'_d, $d' \in D - d$, also conditions on fewer nodes. The highest efficiency is achieved when the influence diagram does not contain any nonrequisite informational arcs, i.e., every $d \in D$ conditions only on $\text{rqo}(d)$.

Given the influence diagram without any nonrequisite informational arcs, there is also another recommended structural change based on the result of Lemma 4, i.e., the choice of δ_d that is 1-stable with respect to δ_D requires the following information.

- Δ_d, the set of all possible strategies at d.
- $\delta_{\text{rqd}(d)\cap D}$, the strategy set associated with the requisite distinction set $\text{rqd}(d)$.
- $\phi_{\text{rqd}(d)\cap U}$, the distribution set associated with the requisite distinction set $\text{rqd}(d)$.
- $\psi_{\text{rqv}(d)}$, the utility set associated with the requisite value set $\text{rqv}(d)$.

The rest of the information provided in the influence diagram is redundant for the purpose of decision making at d. Denote by $\text{rqn}(d) = d \cup \text{rqd}(d) \cup \text{rqv}(d)$ a *requisite node set* for d. A *requisite influence diagram* for d is a subdiagram of Z induced by $\text{fa}(\text{rqn}(d))$. It is said to be *completely specified* if it includes the above information, i.e., Δ_d, $\delta_{\text{rqd}(d)\cap D}$, $\phi_{\text{rqd}(d)\cap U}$, and $\psi_{\text{rqv}(d)}$, also called the *requisite decision basis* for d. Because neither $\delta_{\text{pa}(\text{rqn}(d))\cap D}$ nor $\phi_{\text{pa}(\text{rqn}(d))\cap U}$ is part of the requisite decision basis for d, the nodes in $\text{pa}(\text{rqn}(d))$ are shaded and have dotted borders to distinguish them from the observed nodes in the standard influence diagram.

For any $F \subseteq D$, denote by $\text{rqd}(F) = \bigcup_{d\in F} \text{rqd}(d) - F$, $\text{rqv}(F) = \bigcup_{d\in F} \text{rqv}(d)$, and $\text{rqn}(F) = F \cup \text{rqd}(F) \cup$

rqv(F) the requisite distinction, the requisite value, and the requisite node sets for F, respectively. A requisite influence diagram for F is a subdiagram of Z induced by fa(rqn(F)). It is said to be completely specified if it includes the requisite decision basis for F, i.e., Δ_F, $\delta_{\text{rqd}(F)\cap D}$, $\phi_{\text{rqd}(F)\cap U}$, and $\psi_{\text{rqv}(F)}$. Again, the nodes in pa(rqn(F)) are shaded and have dotted borders because neither $\delta_{\text{pa}(\text{rqn}(F))\cap D}$ nor $\phi_{\text{pa}(\text{rqn}(F))\cap U}$ is part of the requisite decision basis for F. The requisite influence diagram of any given influence diagram is the requisite influence diagram for D of the given influence diagram.

5.2. Influence Diagram Reduction

We now present the algorithm, called *Influence Diagram Reduction* (IDR). For any $F \subseteq D$, it determines the requisite influence diagram for F of any given influence diagram.

Algorithm 2. *Influence Diagram Reduction*.

Input: An influence diagram with a subset $F \subseteq D$.

Output: A set of requisite influence diagrams for d, $\forall d \in D$, and for F.

Iteration

1. For each decision $d \in D$, do:
   (a) Determine pa(d) and rqv(d).
   (b) Assign pa′(d) = pa(d).
   (c) If rqv(d) = ∅, assign rqo(d) = ∅ and rqd(d) = ∅; else, find rqo(d) and rqd(d) with the Bayes-Ball algorithm.
   (d) Update pa(d) = rqo(d) by removing the informational arcs from pa(d) − rqo(d) to d.
2. Repeat the iteration until pa′(d) = pa(d), $\forall d \in D$.
3. The requisite influence diagram for each $d \in D$ is the subdiagram of Z induced by fa(rqn(d)), rqn(d) = $d \cup$ rqd(d) $\cup$ rqv(d).
4. The requisite influence diagram for F is the subdiagram of Z induced by fa(rqn(F)), rqn(F) = $\bigcup_{d\in F}$ rqn(d).

The following proposition proves that the algorithm converges to a set of requisite influence diagrams for d, $\forall d \in D$, and for F.

Proposition 7. *A set of influence diagrams obtained with the IDR algorithm are requisite for d, $\forall d \in D$, and for F.*

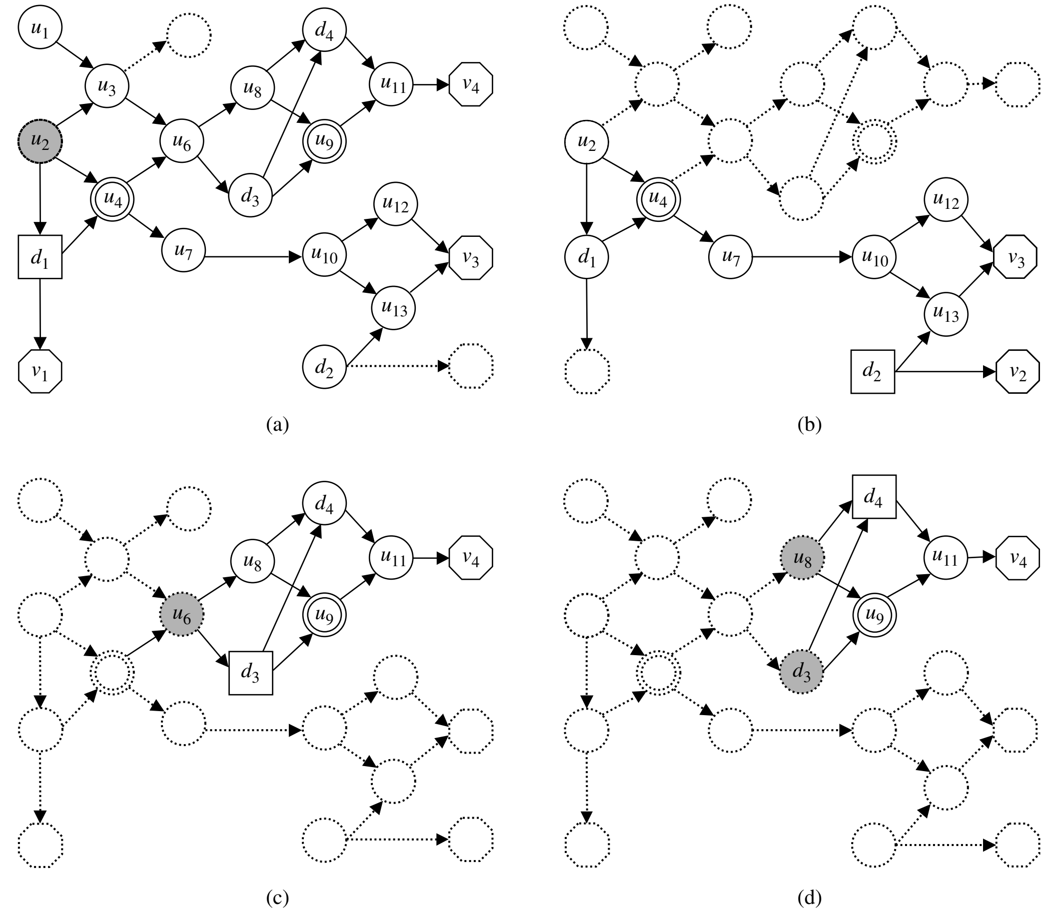
Proof. The correctness of the algorithm is based on the correctness of the Bayes-Ball algorithm and the results of Lemmas 1 to 5. The convergence follows because there is a finite number of informational arcs in the influence diagram, and at least one of them is removed in each iteration (except the last). □

A special case of this algorithm that determines the requisite influence diagram for each decision when the given influence diagram has complete sharing of information is called the *Decomposition of Influence Diagram* (Nielsen 2001). In such a case, the IDR algorithm can be completed in one cycle by applying it in reverse order of decisions having complete sharing of information. In the case of incomplete sharing of information, the IDR algorithm should be applied iteratively following the reverse order of subset of decisions having complete sharing of information.

Example 7. Consider applying the IDR algorithm to the influence diagrams shown in Figure 5, which results in the influence diagrams shown in Figures 6 and 7. For each d_i, $i = 1, \dots, 4$, the nodes in Z − fa(rqn(d_i)) are masked for the purpose of comparison between the original and the requisite influence diagrams. The requisite influence diagram for d_1 shown in Figure 6a is the subdiagram induced by fa(rqn(d_1)) = $Z - u_5 - v_2$ with pa(rqn(d_1)) = u_2 shaded. The requisite influence diagram for d_2 shown in Figure 6b is the subdiagram induced by fa(rqn(d_2)) = $\{d_1, d_2, u_2, u_4, u_7, u_{10}, u_{12}, u_{13}, v_2, v_3\}$, none of which is shaded because pa(rqn(d_2)) = ∅. The requisite influence diagram for d_3 shown in Figure 6c is the subdiagram induced by fa(rqn(d_3)) = $\{d_3, d_4, u_6, u_8, u_9, u_{11}, v_4\}$ with pa(rqn(d_3)) = u_6 shaded. Finally, the requisite influence diagram for d_4 shown in Figure 6d is the subdiagram induced by fa(rqn(d_4)) = $\{d_3, d_4, u_8, u_9, u_{11}, v_4\}$ with pa(rqn(d_4)) = $\{d_3, u_8\}$ shaded. The final step of the IDR algorithm concludes with the requisite influence diagram (for D), the subdiagram induced by fa(rqn(D)) = $Z - u_5$, none of which is shaded, as shown in Figure 7. □

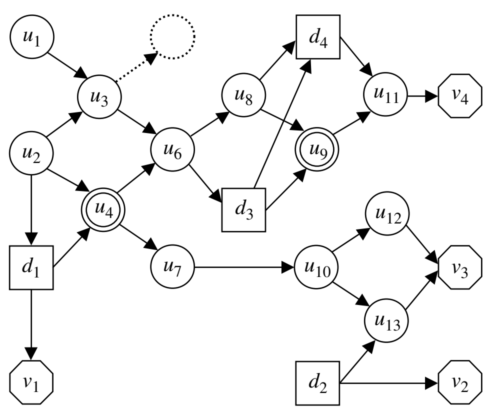
There is a distinction between decision making at D versus computing the expected utility for δ_D. Although a completely specified requisite influence diagram for D provides sufficient information to find a 1-stable δ_D, the information can be insufficient for the purpose of computing the expected utility of that δ_D. To compute the expected utility, the information must include the utility set ψ_V, the distribution set $\phi_{N\cap U}$, and the strategy set $\delta_{N\cap D}$, where N is the

Figure 6 **A Set of Requisite Influence Diagrams for** *d*



requisite distinction set for pa(V) derived from the influence diagram without nonrequisite informational arcs. Note that the nodes in $Y - N$ are exactly the barren nodes.

Figure 7 **A Requisite Influence Diagram**



6. Diagram-Based Solution Concept and Method

This section develops the diagram-based solution concept and method based on the insights exploited in the earlier sections.

6.1. Strategic Irrelevance in an Influence Diagram

A decision $d \in D$ is said to be *strategically irrelevant* to the set of decisions $F \subseteq D$ if, with respect to any strategy set δ_{D-F}, there exists a strategy δ_d that is 1-stable with respect to every possible strategy set $\delta_F \in \Delta_F$. A decision is always strategically irrelevant to itself by this definition. A set of decisions is said to have a strategic irrelevance if there exists a complete ordering of those decisions such that each decision is strategically irrelevant to all earlier decisions in the set. A set of decisions is a strategically irrelevant set if it has strategic irrelevance, and a strategically relevant set otherwise. A decision situation is said to have a strategic irrelevance if there exists a strategically irrelevant set that includes all decisions in the decision situation. A decision situation is said to have a strategic relevance otherwise.

To search for strategically irrelevant sets, we reconsider the remark at the end of Lemma 4. Although the choice of δ_d that is 1-stable with respect to δ_D can be found once at least $\delta_{\text{rqd}(d)\cap D}$, $\phi_{\text{rqd}(d)\cap U}$, and $\psi_{\text{rqv}(d)}$ are given, it does not imply that such a δ_d is invariant to $\delta_{D-\text{rqd}(d)}$ and $\phi_{U-\text{rqd}(d)}$. Fortunately, there exists a δ_d that is both 1-stable with respect to δ_D and invariant to $\delta_{D-\text{rqd}(d)}$ and $\phi_{U-\text{rqd}(d)}$. For the purpose of determining the strategically irrelevant sets, we only prove the existence of such δ_d that is invariant to $\delta_{D-\text{rqd}(d)}$. For notational convenience, denote $\text{si}(d) = D - \text{rqd}(d)$.

Theorem 1. *A strategy δ_d, $d \in D$, is 1-stable with respect to strategy set $\delta_{\text{si}(d)} \cup \delta_{D-\text{si}(d)}$, $\forall \delta_{\text{si}(d)} \in \Delta_{\text{si}(d)}$, if and only if δ_d is 1-stable with respect to any strategy set $\delta'_{\text{si}(d)} \cup \delta_{D-\text{si}(d)}$ that induces $\mathcal{P}_{Z*d}(x_{\text{fa}(d)}) > 0$ whenever $\mathcal{P}_{Z*\text{si}(d)}(x_{\text{fa}(d)}) > 0$.*

Proof. For an arbitrary $\delta_{\text{si}(d)} \cup \delta_{D-\text{si}(d)}$, there are three distinct cases for each $x_{\text{fa}(d)} \in X_{\text{fa}(d)}$.

(1) $\mathcal{P}_{Z*\text{si}(d)}(x_{\text{fa}(d)}) = 0$: This implies that $\mathcal{P}_{Z*d}(x_{\text{fa}(d)}) = 0$ when induced by either $\delta_{\text{si}(d)} \cup \delta_{D-\text{si}(d)}$ or $\delta'_{\text{si}(d)} \cup \delta_{D-\text{si}(d)}$. Lemma 1 implies that every $x_d \in X_d$ is equally preferred, and the proof for this case follows immediately.

(2) $\mathcal{P}_{Z*\text{si}(d)}(x_{\text{fa}(d)}) \neq 0$, but $\delta_{\text{si}(d)} \cup \delta_{D-\text{si}(d)}$ induces $\mathcal{P}_{Z*d}(x_{\text{fa}(d)}) = 0$: Lemma 1 implies that every $x_d \in X_d$ is equally preferred with $\delta_{\text{si}(d)} \cup \delta_{D-\text{si}(d)}$, but not necessarily with $\delta'_{\text{si}(d)} \cup \delta_{D-\text{si}(d)}$. Hence, δ_d that is 1-stable with respect to the latter is always 1-stable with respect to the former.

(3) $\mathcal{P}_{Z*\text{si}(d)}(x_{\text{fa}(d)}) \neq 0$ and $\delta_{\text{si}(d)} \cup \delta_{D-\text{si}(d)}$ induces $\mathcal{P}_{Z*d}(x_{\text{fa}(d)}) \neq 0$: $\mathcal{P}_{Z*d}(x_{\text{fa}(d)}) \neq 0$ ensures that $\mathcal{P}_{Z*d}(x_{\text{pa}(\text{rqv}(d))} \mid x_{\text{fa}(d)})$ is determinable and does not vary with either $\delta_{\text{si}(d)} \cup \delta_{D-\text{si}(d)}$ or $\delta'_{\text{si}(d)} \cup \delta_{D-\text{si}(d)}$ according to the definition of $\text{rqd}(d)$ in Lemma 4. Hence, Lemma 3 implies that δ_d that is 1-stable with respect to the former is also 1-stable with respect to the latter, and vice versa.

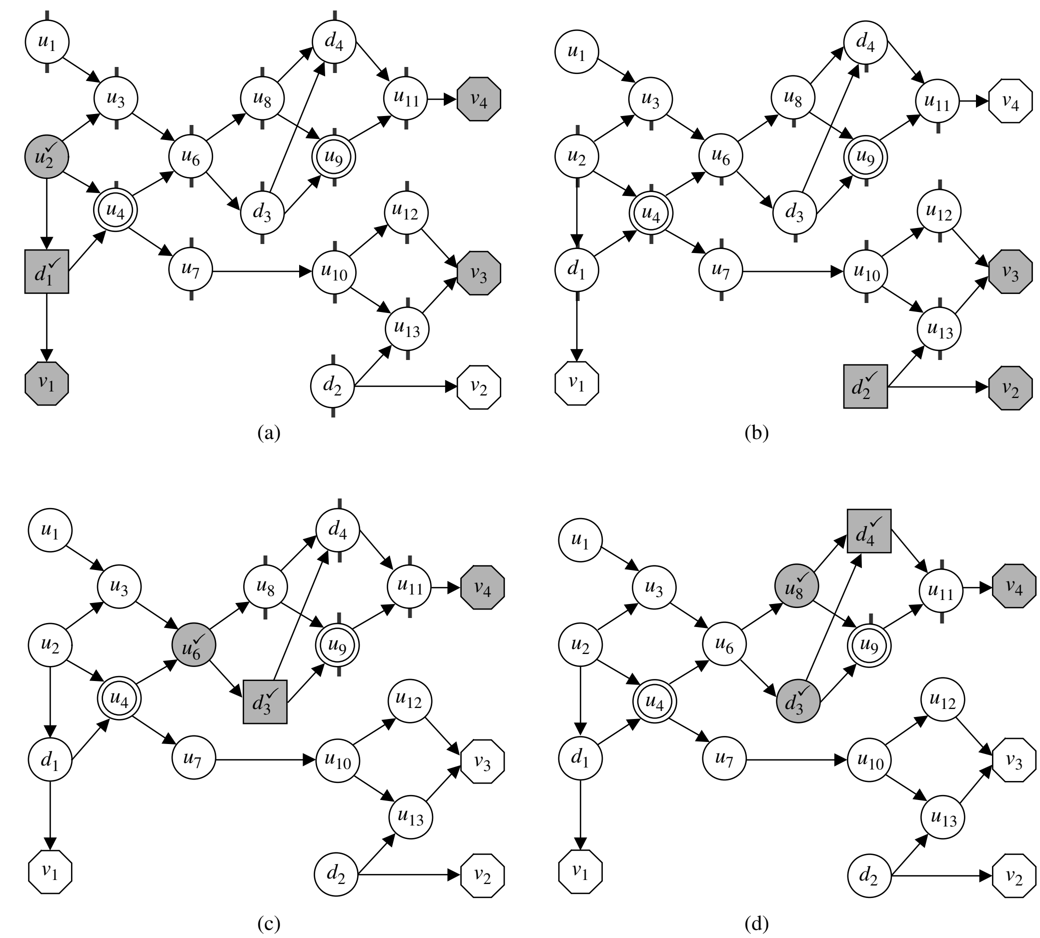
The "if" part follows because the above cases hold for any $\delta_{\text{si}(d)} \cup \delta_{D-\text{si}(d)}$, $\forall \delta_{\text{si}(d)} \in \Delta_{\text{si}(d)}$. Note that the "only if" part follows trivially because $\delta'_{\text{si}(d)} \in \Delta_{\text{si}(d)}$. □

For any $d \in D$, Theorem 1 implies that $\text{si}(d)$ is a set of decision nodes to which d is strategically irrelevant. A subset $F \subseteq D$ has strategic irrelevance when there exists an ordering of $\{d_1, d_2, \ldots, d_k\} = F$ such that $\{d_1, \ldots, d_i\} \subseteq \text{si}(d_i)$, $i = 1, \ldots, k$. Note that the sets $\text{si}(d)$, $\forall d \in D$, can and should be derived from the requisite influence diagram. This is because $\text{si}(d)$ (defined through $\text{rqd}(d)$) is never smaller in the requisite influence diagram than in the original influence diagram according to Lemma 5.

Example 8. Consider the requisite influence diagrams shown in Figure 8. For each d_i, $i = 1, \ldots, 4$, $\text{si}(d_i)$ is the set of nodes in D that are not marked on the top. The influence diagram shown in Figure 8a illustrates that d_1 is strategically irrelevant only to itself because $\text{si}(d_1) = d_1$. The influence diagram shown in Figure 8b illustrates that d_2 is strategically irrelevant to $D - d_1$ because $\text{si}(d_2) = D - d_1$. The influence diagram shown in Figure 8c illustrates that d_3 is strategically irrelevant to $D - d_4$ because $\text{si}(d_3) = D - d_4$. Finally, the influence diagram shown in Figure 8d illustrates the best situation, i.e., d_4 is strategically irrelevant to all decisions because $\text{si}(d_4) = D$. □

The notion of strategic irrelevance is similar to that defined in Koller and Milch (2001, 2003), when we approach team decision making as a game where every player has identical payoffs. The strategically irrelevant sets are, however, never smaller in our case because we work with the requisite influence diagram. Note that although the notions of *stepwise-decomposability candidate node* (Zhang and Poole 1992, Zhang et al. 1994) and *extremal decision* (Nilsson and Lauritzen 2000, Lauritzen and Nilsson 2001) have a similar motivation, they are special cases of strategic irrelevance because the formers are defined indirectly through the irrelevant set while the latter is defined through the requisite distinction set. As a result, a decision that satisfies either of the former notions always satisfies the latter notion, while the converse is not true. Consider the requisite influence diagram shown in Figure 7. If d_1 were an uncertainty, we would have that $\text{si}(d_2) = D$, and d_2 would be strategically irrelevant to $\{d_2, d_3, d_4\}$. Consequently, there exists δ_{d_2} that is 1-stable with respect to every $\delta_{d_3 \cup d_4} \in \Delta_{d_3 \cup d_4}$. Such a d_2, however, is not an extremal decision because neither $\text{rqv}(d_2) \perp \text{fa}(d_3) \mid \text{fa}(d_2)$ nor $\text{rqv}(d_2) \perp \text{fa}(d_4) \mid \text{fa}(d_2)$ holds. Note also that the former notions

Figure 8 Irrelevant, Requisite Value, Requisite Observation, and Requisite Distinction Sets in Requisite Influence Diagrams



are only defined with respect to all decisions (excluding those already converted to uncertainties), while the latter notion is defined with respect to any subset of decisions.

6.2. Strategic Stability in an Influence Diagram

The above property of the strategic irrelevance yields a new perspective in finding a δ_d that is 1-stable to δ_D. We first introduce the notion of uniform stability that satisfies the condition given in Theorem 1.

Definition 7. A strategy set δ_F, $F \subseteq D$, is *uniformly stable* with respect to strategy set δ_D when

$$\delta_d = \arg\max_{\tilde{\delta}_d} \text{EU}(\tilde{\delta}_d, \bar{\delta}_{\text{si}(d)-d}, \delta_{D-\text{si}(d)})$$

for $\forall\, d \in F$ where $\bar{\delta}_{\text{si}(d)-d}$ is a uniform strategy set.

The following corollary restates Theorem 1 with the above definition of uniform stability.

Corollary 1. *A strategy δ_d, $d \in D$, is 1-stable with respect to strategy set $\delta_{\text{si}(d)} \cup \delta_{D-\text{si}(d)}$, $\forall\, \delta_{\text{si}(d)} \in \Delta_{\text{si}(d)}$, if and only if δ_d is uniformly stable with respect to strategy set δ_D.*

Proof. Note that "δ_d is uniformly stable with respect to strategy set δ_D "is equivalent to" δ_d is 1-stable with respect to strategy set $\bar{\delta}_{\text{si}(d)} \cup \delta_{D-\text{si}(d)}$." The result follows immediately because $\bar{\delta}_{\text{si}(d)} \cup \delta_{D-\text{si}(d)}$ induces $\mathcal{P}_{Z*d}(x_{\text{fa}(d)}) = \mathcal{P}_{Z*\text{si}(d)}(x_{\text{fa}(d)})$, the condition given in Theorem 1. □

A consequence of Theorem 1 and Corollary 1 is that with respect to δ_D, if δ_d does not vary with $\delta_{\text{si}(d)}$, it always constitutes a part of maximally stable strategy set with δ_e, $\forall\, e \in \text{si}(d) - d$.

Theorem 2. *Given a strategy δ_d, $d \in D$, that is uniformly stable with respect to strategy set δ_D, $\{\delta_d, \delta_e\}$, $e \in$*

$\mathrm{si}(d) - d$, *is maximally stable with respect to* δ_D *if and only if* δ_e *is 1-stable with respect to* δ_D.

Proof. According to the definition of maximal stability and Theorem 1, we have

$$\begin{aligned}\mathrm{EU}(\delta_e, \delta_d, \delta_{D-\{e\cup d\}}) &= \max_{\tilde{\delta}_e, \tilde{\delta}_d} \mathrm{EU}(\tilde{\delta}_e, \tilde{\delta}_d, \delta_{D-\{e\cup d\}}) \\ &= \max_{\tilde{\delta}_e} \mathrm{EU}(\tilde{\delta}_e, \delta_d, \delta_{D-\{e\cup d\}}) \\ &= \mathrm{EU}(\delta_e, \delta_d, \delta_{D-\{e\cup d\}}).\end{aligned}$$

With a uniformly stable δ_d with respect to δ_D, $\{\delta_d, \delta_e\}$ is maximally stable with respect to δ_D if δ_e is 1-stable with respect to δ_D. The reverse is always true by Proposition 2. □

A corollary of Theorem 2 applies to any strategically irrelevant set.

Corollary 2. *For any* $F \subseteq D$ *that has strategic irrelevance, a strategy set* δ_F *is maximally stable with respect to strategy set* δ_D *if* δ_F *is uniformly stable with respect to* δ_D.

Proof. Without loss of generality, assume that $F = \{d_1, d_2, \ldots, d_k\}$ with the index representing an ordering of the decisions. According to the definition of maximal stability and Theorem 2, we have

$$\begin{aligned}\mathrm{EU}(\delta_F, \delta_{D-F}) &= \max_{\tilde{\delta}_F} \mathrm{EU}(\tilde{\delta}_F, \delta_{D-F}) \\ &= \max_{\tilde{\delta}_{F-d_k}, \tilde{\delta}_{d_k}} \mathrm{EU}(\tilde{\delta}_{F-d_k}, \tilde{\delta}_{d_k}, \delta_{D-F}) \\ &= \max_{\tilde{\delta}_{F-d_k}} \mathrm{EU}(\tilde{\delta}_{F-d_k}, \delta_{d_k}, \delta_{D-F}).\end{aligned}$$

Hence, maximal stability of δ_F with respect to δ_D requires maximal stability of δ_{F-d_k} and uniform stability of δ_{d_k} with respect to δ_D. By induction, it requires uniform stability of δ_F with respect to δ_D. □

We now introduce the algorithm, called *Uniform Strategy Improvement* (USI), that combines the benefits of the SSI algorithm with the uniform stability. Although it can exploit any strategically irrelevant set when present, it is applicable to any set of decisions. Each enumeration set is still associated with a single decision, yet the resulting strategy set is maximally stable at every subset of decisions (within the union of all enumeration sets) that has strategic irrelevance. We call this *Strategic Stability* (within the union of all enumeration sets) where the modifier is omitted in the case that the union of all enumeration sets is the set of decisions represented by the strategy set.

Algorithm 3. *Uniform Strategy Improvement*.

Input: A completely specified influence diagram with a subset $F \subseteq D$.

Output: A pure-strategy set δ_D that is strategically stable within F.

Initialization

1. Assign some initial strategy set δ_D.

Iteration

1. Assign $\delta'_F = \delta_F$.
2. For each decision $d \in F$, do:
   (a) If $\delta'_d \neq \arg\max_{\tilde{\delta}_d} \mathrm{EU}(\tilde{\delta}_d, \bar{\delta}_{\mathrm{si}(d)-d}, \delta_{D-\mathrm{si}(d)})$, update δ_d based on Lemma 4 with respect to $\bar{\delta}_{\mathrm{si}(d)} \cup \delta_{D-\mathrm{si}(d)}$.
3. Repeat the iteration until $\delta'_F = \delta_F$.

Note that if we replace $\bar{\delta}_{\mathrm{si}(d)}$ in the above algorithm with some $\delta'_{\mathrm{si}(d)}$ that satisfies the condition given in Theorem 1, we have another variation of the algorithm even though it is not necessarily more efficient due to the additional work required to verify such a condition. The following proposition proves that the algorithm converges to a δ_D that is strategically stable within F, the union of all enumeration sets.

Proposition 8. *A pure-strategy set* δ_D *obtained with the USI algorithm is strategically stable within the union of all enumeration sets*.

Proof. Lemma 4 ensures that at each iteration, we always update δ_d to the one that is uniformly stable with respect to δ_D unless it is already so. If the algorithm converges, δ_D must be strategically stable within F according to Corollary 2. The proof of convergence then follows from that in Proposition 4. □

Example 9. Consider the influence diagrams shown in Figure 8. Applying the above algorithm yields a δ_D that is maximally stable at $\{d_1, d_3, d_4\}$ and $\{d_2, d_3, d_4\}$. This is because $\mathrm{si}(d_4) = D$ and $\mathrm{si}(d_3) = D - d_4$, but $d_1 \notin \mathrm{si}(d_2)$ and $d_2 \notin \mathrm{si}(d_1)$ according to the derivation in Example 8. □

The USI algorithm is identical to the SSI (or SPU) algorithm if properly applied to a *soluble* influence diagram, one with an exact solution ordering of extremal decisions (Lauritzen and Nilsson 2001). Both algorithms are initialized with the uniform strategy set $\bar{\delta}_D$ and applied in reverse order of the solution ordering, although any order is allowed with the USI algorithm. The key difference between both algorithms arises when applied to a nonsoluble influence diagram. The USI algorithm always finds a strategically stable strategy set, while the SSI (or SPU) algorithm only finds a 1-stable strategy set although it might find a strategically stable one. An extreme case is when the influence diagram is nonsoluble but has strategic irrelevance, i.e., the USI algorithm guarantees a maximally stable (optimal) strategy set, while the SSI (or SPU) algorithm only guarantees a 1-stable strategy set.

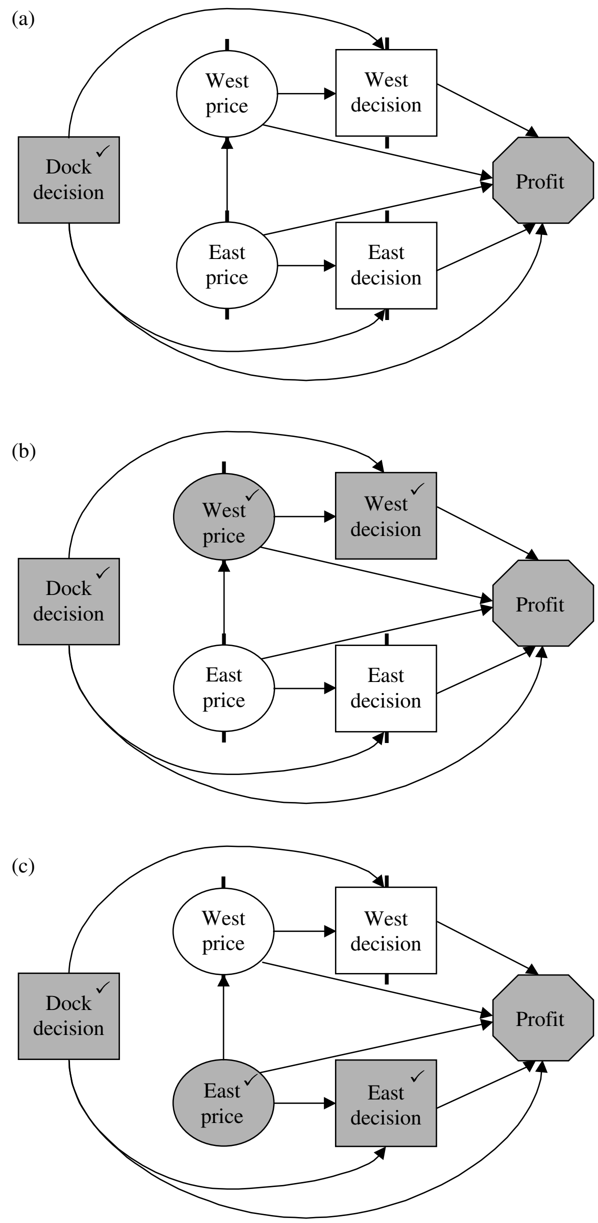
7. Solution Example

There are three decisions in Example 1. The supervisor's decision has two pure strategies, while each manager's decision has $2^{2\times 2} = 16$ pure strategies. This brings the total number of pure-strategy sets to $2 \times 16 \times 16 = 512$. Based on the numerical specifications provided in the augmented decision tree as shown in Figure 1, the exhaustive enumeration algorithm finds 16 different strategy sets to be maximally stable (optimal) with the expected utility of 11.3.

Applying the IDR algorithm to the influence diagram shown in Figure 2 results in the influence diagrams shown in Figure 9. The resulting influence diagrams illustrate that they are already requisite for their respective decisions, and no reduction is possible. They also illustrate that the supervisor's decision is only strategically irrelevant to itself, while each manager's decision is strategically irrelevant to itself and the supervisor's decision. Applying the USI algorithm guarantees that the resulting strategy set is strategically stable, i.e., maximally stable at $DD \cup WD$ and $DD \cup ED$.

Note that the influence diagram shown in Figure 2 is nonsoluble, i.e., without exact solution ordering, because there is no extremal decision. Table 1 thus compares the results obtained from applying the SSI (or SPU) algorithm and the USI algorithm following all $3! = 6$ possible solution orderings. By initializing both algorithms with the uniform strategy set, the strategically stable solutions guaranteed by the USI algorithm happen to be optimal in every possible solution ordering while the 1-stable solutions guaranteed by the SSI (or SPU) algorithm happen to be optimal in two solution orderings.

Figure 9 Irrelevant, Requisite Value, Requisite Observation, and Requisite Distinction Sets in Requisite Influence Diagrams for Example 1



8. Conclusions

We have shown how to represent and evaluate the team decision situation with an influence diagram. The solution concept of k-stability and the corresponding solution method of strategy improvement allow us to properly evaluate the team decision situation having incomplete sharing of information.

Table 1 Results of SSI (SPU) and USI Algorithms on Example 1

Ordering	SSI (SPU)	USI
$D.D. \rightarrow W.D. \rightarrow E.D.$	10.8	11.3
$D.D. \rightarrow E.D. \rightarrow W.D.$	10.8	11.3
$W.D. \rightarrow D.D. \rightarrow E.D.$	10.8	11.3
$W.D. \rightarrow E.D. \rightarrow D.D.$	10.8	11.3
$E.D. \rightarrow D.D. \rightarrow W.D.$	11.3	11.3
$E.D. \rightarrow W.D. \rightarrow D.D.$	11.3	11.3

The notions of strategic irrelevance and requisite influence diagram, resulting from exploiting and simplifying the graphical structure of the influence diagram, allow the use of all available information to improve the decision quality. These in turn enable us to find the joint strategy that is maximally stable over the largest sets of decisions possible with the USI algorithm.

The proofs of Theorem 2 and Corollary 2 imply that no reevaluation is needed for a decision if the remaining evaluations are on the set of the decisions to which it is strategically irrelevant. This insight suggests that some solution orderings can be more efficient than others, although the USI algorithm is applicable to any solution ordering. Koller and Milch (2001, 2003) suggest a partial solution ordering based on the notion of a relevance graph, which allows the BI algorithm to be performed over a subset of decisions instead of a single decision. Detwarasiti (2005) further applies this insight to organize a collection of the enumeration sets such that the SI algorithm always converges to the optimal strategy set, and also proposes an improved solution ordering for the USI algorithm that includes an efficient ordering within the subset of decisions not addressed by the relevance graph.

Because most efficient evaluation techniques for an influence diagram convert the diagram into an auxiliary graphical structure called a *junction tree* (Jensen et al. 1994, Lauritzen and Nilsson 2001), the refinement of the USI algorithm should not only improve its efficiency, but also address its applicability within the junction-tree propagation techniques (Detwarasiti 2005).

The information structure that supports the highest decision quality is the one represented by an influence diagram having complete sharing of information. This opens the possibility of quantifying the value of an information structure through the decision quality supported by it. Additional exploitation of the graphical structure of the influence diagram can determine whether an information structure represents sufficient sharing of information, one that yields the decision quality equivalent to that representing complete sharing of information but with reduced assessment and computational effort (Detwarasiti 2005).

Acknowledgments
The authors are indebted for the support by General Motors through the GM-Stanford Collaborative Research Laboratory in Work Systems, and to their colleagues. They are also grateful to Ronald Howard, the editor Eric Horvitz, and three anonymous referees for their valuable comments.

Appendix

Proof of Lemma 2. We can expand $\mathcal{V}_1(x_{\mathrm{fa}(d)})$, defined in Lemma 1 as

$$\begin{aligned}\mathcal{V}_1(x_{\mathrm{fa}(d)}) &= \sum_{x_{\mathrm{fa}(Z)-\mathrm{fa}(d)}} \mathcal{P}_{Z*d}(x_{\mathrm{fa}(Y)})\mathcal{U}_Z(x_{\mathrm{pa}(V)}) \\ &= \sum_{x_{\mathrm{fa}(Z)-\mathrm{fa}(d)}} \mathcal{P}_{Z*d}(x_{\mathrm{fa}(Y)})\{\mathcal{U}_N(x_{\mathrm{pa}(N\cap V)}) + \mathcal{U}_{Z-N}(x_{\mathrm{pa}(V-N)})\}.\end{aligned}$$

We can premarginalize the second summation over $\mathrm{fa}(Z) - \{\mathrm{pa}(V-N) \cup \mathrm{fa}(d)\}$ as

$$\begin{aligned}\mathcal{V}_1(x_{\mathrm{fa}(d)}) &= \sum_{x_{\mathrm{fa}(Z)-\mathrm{fa}(d)}} \mathcal{P}_{Z*d}(x_{\mathrm{fa}(Y)})\mathcal{U}_N(x_{\mathrm{pa}(N\cap V)}) \\ &\quad + \sum_{x_{\mathrm{pa}(V-N)-\mathrm{fa}(d)}} \sum_{x_{\mathrm{fa}(Z)-\{\mathrm{pa}(V-N)\cup \mathrm{fa}(d)\}}} \mathcal{P}_{Z*d}(x_{\mathrm{fa}(Y)})\mathcal{U}_{Z-N}(x_{\mathrm{pa}(V-N)}) \\ &= \mathcal{V}_2(x_{\mathrm{fa}(d)}) + \sum_{x_{\mathrm{pa}(V-N)-\mathrm{fa}(d)}} \mathcal{P}_{Z*d}(x_{\mathrm{pa}(V-N)\cup \mathrm{fa}(d)})\mathcal{U}_{Z-N}(x_{\mathrm{pa}(V-N)}).\end{aligned} \tag{A1}$$

Equation (A1) follows from substituting the first summation with $\mathcal{V}_2(x_{\mathrm{fa}(d)})$. In the case that $\mathcal{P}_{Z*d}(x_{\mathrm{pa}(V-N)} \mid x_{\mathrm{fa}(d)})$ can be determined, we can rewrite the summation as

$$\begin{aligned}\mathcal{V}_1(x_{\mathrm{fa}(d)}) &= \mathcal{V}_2(x_{\mathrm{fa}(d)}) + \sum_{x_{\mathrm{pa}(V-N)-\mathrm{fa}(d)}} \mathcal{P}_{Z*d}(x_{\mathrm{fa}(d)})\mathcal{P}_{Z*d}(x_{\mathrm{pa}(V-N)} \mid x_{\mathrm{fa}(d)}) \\ &\qquad \cdot \mathcal{U}_{Z-N}(x_{\mathrm{pa}(V-N)}) \qquad \text{(A2)} \\ &= \mathcal{V}_2(x_{\mathrm{fa}(d)}) + \mathcal{P}_{Z*d}(x_{\mathrm{fa}(d)}) \\ &\quad \cdot \sum_{x_{\mathrm{pa}(V-N)-\mathrm{fa}(d)}} \mathcal{P}_{Z*d}(x_{\mathrm{pa}(V-N)} \mid x_{\mathrm{pa}(d)})\mathcal{U}_{Z-N}(x_{\mathrm{pa}(V-N)}) \qquad \text{(A3)} \\ &= \mathcal{V}_2(x_{\mathrm{fa}(d)}) + \mathcal{P}_{Z*d}(x_{\mathrm{fa}(d)})\mathcal{K}_1(x_{\mathrm{fa}(d)}). \qquad \text{(A4)}\end{aligned}$$

Note that we can replace $\mathcal{P}_{Z*d}(x_{\text{pa}(V-N)} \mid x_{\text{fa}(d)})$ in Equation (A2) with $\mathcal{P}_{Z*d}(x_{\text{pa}(V-N)} \mid x_{\text{pa}(d)})$ in Equation (A3) because $\text{pa}(V-N) \perp d \mid \text{pa}(d)$ by Definitions 3 and 6. Equation (A4) follows from the fact that we can always substitute the result of marginalization with some appropriate function $\mathcal{K}_1(x_{\text{fa}(d)})$. The next step is to replace $\mathcal{K}_1(x_{\text{fa}(d)})$ with $\mathcal{K}_1(x_{\text{pa}(d)})$ because Equations (A2) and (A3) imply that $\mathcal{K}_1(x'_d, x_{\text{pa}(d)}) = \mathcal{K}_1(x''_d, x_{\text{pa}(d)})$, $x'_d, x''_d \in X_d$. Finally, we can factor $|X_d|^{-1}$ out of $\mathcal{P}_{Z*d}(x_{\text{fa}(d)})$ as

$$\mathcal{V}_1(x_{\text{fa}(d)}) = \mathcal{V}_2(x_{\text{fa}(d)}) + |X_d|^{-1}\mathcal{P}_{Z*d}(x_{\text{pa}(d)})\mathcal{K}_1(x_{\text{pa}(d)}).$$

The result follows because for each $x_{\text{pa}(d)} \in X_{\text{pa}(d)}$, $\mathcal{V}_2(x_{\text{fa}(d)})$ is simply a linear transformation of $\mathcal{V}_1(x_{\text{fa}(d)})$. □

Proof of Lemma 3. We can define

$$\mathcal{V}_2(x_{\text{fa}(d)}) = \sum_{x_{\text{fa}(Z)-\text{fa}(d)}} \mathcal{P}_{Z*d}(x_{\text{fa}(Y)})\mathcal{U}_{\text{rqv}(d)}(x_{\text{pa}(\text{rqv}(d))})$$

according to Lemma 2 without loss of generality. This allows us to premarginalize it over $\text{fa}(Z) - \{\text{pa}(\text{rqv}(d)) \cup \text{fa}(d)\}$ as

$$\begin{aligned}\mathcal{V}_2(x_{\text{fa}(d)}) &= \sum_{x_{\text{pa}(\text{rqv}(d))-\text{fa}(d)}} \sum_{x_{\text{fa}(Z)-\{\text{pa}(\text{rqv}(d))\cup\text{fa}(d)\}}} \mathcal{P}_{Z*d}(x_{\text{fa}(Y)})\mathcal{U}_{\text{rqv}(d)}(x_{\text{pa}(\text{rqv}(d))}) \\ &= \sum_{x_{\text{pa}(\text{rqv}(d))-\text{fa}(d)}} \mathcal{P}_{Z*d}(x_{\text{pa}(\text{rqv}(d))\cup\text{fa}(d)})\mathcal{U}_{\text{rqv}(d)}(x_{\text{pa}(\text{rqv}(d))}).\end{aligned}$$

In the case that $\mathcal{P}_{Z*d}(x_{\text{pa}(\text{rqv}(d))} \mid x_{\text{fa}(d)})$ can be determined, we can rewrite the summation as

$$\begin{aligned}\mathcal{V}_2(x_{\text{fa}(d)}) &= \sum_{x_{\text{pa}(\text{rqv}(d))-\text{fa}(d)}} \mathcal{P}_{Z*d}(x_{\text{fa}(d)})\mathcal{P}_{Z*d}(x_{\text{pa}(\text{rqv}(d))} \mid x_{\text{fa}(d)})\mathcal{U}_{\text{rqv}(d)}(x_{\text{pa}(\text{rqv}(d))}) \\ &= \mathcal{P}_{Z*d}(x_{\text{fa}(d)}) \sum_{x_{\text{pa}(\text{rqv}(d))-\text{fa}(d)}} \mathcal{P}_{Z*d}(x_{\text{pa}(\text{rqv}(d))} \mid x_N)\mathcal{U}_{\text{rqv}(d)}(x_{\text{pa}(\text{rqv}(d))}). \qquad \text{(A5)}\end{aligned}$$

Equation (A5) follows from the fact that $\text{fa}(d) \perp \text{pa}(\text{rqv}(d)) \mid N$ by Definitions 3 and 5. As in Lemma 2, we can factor $|X_d|^{-1}$ out of $\mathcal{P}_{Z*d}(x_{\text{fa}(d)})$, and substitute the summation with $\mathcal{V}_3(x_{\text{fa}(d)})$ as

$$\mathcal{V}_2(x_{\text{fa}(d)}) = |X_d|^{-1}\mathcal{P}_{Z*d}(x_{\text{pa}(d)})\mathcal{V}_3(x_{\text{fa}(d)}).$$

The result follows because for each $x_{\text{pa}(d)} \in X_{\text{pa}(d)}$, $\mathcal{V}_3(x_{\text{fa}(d)})$ is simply a linear transformation of $\mathcal{V}_2(x_{\text{fa}(d)})$. The implication is in one direction because $\mathcal{P}_{Z*d}(x_{\text{pa}(\text{rqv}(d))} \mid x_N)$ and $\mathcal{V}_3(x_{\text{fa}(d)})$ might be determinable when $\mathcal{P}_{Z*d}(x_{\text{fa}(d)}) = 0$, and thus $\mathcal{V}_2(x_{\text{fa}(d)}) = 0$. □

Proof of Lemma 4. Denote by $N'_1 = Z - N_1$ a set of nodes that are not in N_1. Lemma 3 implies that we can define

$$\begin{aligned}\mathcal{V}_2(x_{\text{fa}(d)}) &= \sum_{x_{\text{fa}(Z)-\text{fa}(d)}} \mathcal{P}_{Z*d}(x_{\text{fa}(Y)})\mathcal{U}_{\text{rqv}(d)}(x_{\text{pa}(\text{rqv}(d))}) \\ &= |X_d|^{-1}\mathcal{P}_{Z*d}(x_{\text{pa}(d)}) \cdot \sum_{x_{\text{pa}(\text{rqv}(d))-\text{fa}(d)}} \mathcal{P}_{Z*d}(x_{\text{pa}(\text{rqv}(d))} \mid x_N)\mathcal{U}_{\text{rqv}(d)}(x_{\text{pa}(\text{rqv}(d))}) \\ &= |X_d|^{-1}\mathcal{P}_{Z*d}(x_{\text{pa}(d)})\mathcal{V}_3(x_{\text{fa}(d)}) \qquad \text{(A6)}\end{aligned}$$

and

$$\begin{aligned}\mathcal{V}'_2(x_{\text{fa}(d)}) &= \sum_{x_{\text{fa}(Z)-\text{fa}(d)}} \mathcal{P}_{Z*\{N'_1\cup d\}}(x_{\text{fa}(Y)})\mathcal{U}_{\text{rqv}(d)}(x_{\text{pa}(\text{rqv}(d))}) \\ &= |X_d|^{-1}\mathcal{P}_{Z*\{N'_1\cup d\}}(x_{\text{pa}(d)}) \cdot \sum_{x_{\text{pa}(\text{rqv}(d))-\text{fa}(d)}} \mathcal{P}_{Z*\{N'_1\cup d\}}(x_{\text{pa}(\text{rqv}(d))} \mid x_N)\mathcal{U}_{\text{rqv}(d)}(x_{\text{pa}(\text{rqv}(d))}) \\ &= |X_d|^{-1}\mathcal{P}_{Z*\{N'_1\cup d\}}(x_{\text{pa}(d)})\mathcal{V}'_3(x_{\text{fa}(d)}). \qquad \text{(A7)}\end{aligned}$$

Definition 4 implies that

$$\mathcal{P}_{Z*d}(x_{\text{pa}(\text{rqv}(d))} \mid x_N) = \mathcal{P}_{Z*\{N'_1\cup d\}}(x_{\text{pa}(\text{rqv}(d))} \mid x_N)$$

because they are both invariant to $\delta_{N'_1\cap D}$ and $\phi_{N'_1\cap U}$. We thus have $\mathcal{V}_3(x_{\text{fa}(d)}) = \mathcal{V}'_3(x_{\text{fa}(d)})$, which implies that $\mathcal{V}_2(x_{\text{fa}(d)}) = \mathcal{K}_2(x_{\text{pa}(d)})\mathcal{V}'_2(x_{\text{fa}(d)})$, $\mathcal{K}_2(x_{\text{pa}(d)}) = \mathcal{P}_{Z*d}(x_{\text{pa}(d)})/\mathcal{P}_{Z*\{N'_1\cup d\}}(x_{\text{pa}(d)})$. Because $\mathcal{P}_{Z*\{N'_1\cup d\}}(x_{\text{pa}(d)}) = 0$ implies $\mathcal{P}_{Z*d}(x_{\text{pa}(d)}) = 0$, we have that $\mathcal{K}_2(x_{\text{pa}(d)}) = 0$ whenever $\mathcal{P}_{Z*d}(x_{\text{pa}(d)}) = 0$. This allows us to expand $\mathcal{V}_2(x_{\text{fa}(d)})$ in Equation (A6) with respect to $\mathcal{V}'_2(x_{\text{fa}(d)})$ in Equation (A7) as

$$\begin{aligned}\mathcal{V}_2(x_{\text{fa}(d)}) &= \mathcal{K}_2(x_{\text{pa}(d)}) \sum_{x_{\text{fa}(Z)-\text{fa}(d)}} \mathcal{P}_{Z*\{N'_1\cup d\}}(x_{\text{fa}(Y)})\mathcal{U}_{\text{rqv}(d)}(x_{\text{pa}(\text{rqv}(d))}) \\ &= \mathcal{K}_2(x_{\text{pa}(d)}) \sum_{x_{\text{fa}(Z)-\text{fa}(d)}} \mathcal{P}_{N_1*d}(x_{\text{fa}(N_1\cap Y)})\mathcal{P}_{N'_1*N'_1}(x_{\text{fa}(N'_1\cap Y)}) \cdot \mathcal{U}_{\text{rqv}(d)}(x_{\text{pa}(\text{rqv}(d))}).\end{aligned}$$

We can premarginalize the second probability potential over $\text{fa}(Z) - \text{fa}(N_1 \cup N_2)$ as

$$\begin{aligned}\mathcal{V}_2(x_{\text{fa}(d)}) &= \mathcal{K}_2(x_{\text{pa}(d)}) \sum_{x_{\text{fa}(N_1\cup N_2)-\text{fa}(d)}} \mathcal{P}_{N_1*d}(x_{\text{fa}(N_1\cap Y)}) \cdot \mathcal{U}_{\text{rqv}(d)}(x_{\text{pa}(\text{rqv}(d))}) \sum_{x_{\text{fa}(Z)-\text{fa}(N_1\cup N_2)}} \mathcal{P}_{N'_1*N'_1}(x_{\text{fa}(N'_1\cap Y)}) \\ &= \mathcal{K}_2(x_{\text{pa}(d)}) \sum_{x_{\text{fa}(N_1\cup N_2)-\text{fa}(d)}} \mathcal{P}_{N_1*d}(x_{\text{fa}(N_1\cap Y)}) \cdot \mathcal{U}_{\text{rqv}(d)}(x_{\text{pa}(\text{rqv}(d))})\mathcal{K}_3(x_{\text{fa}(N_1\cup N_2)\cap\text{fa}(N'_1\cap Y)}) \\ &= K \cdot \mathcal{K}_2(x_{\text{pa}(d)}) \sum_{x_{\text{fa}(N_1\cup N_2)-\text{fa}(d)}} \mathcal{P}_{N_1*d}(x_{\text{fa}(N_1\cap Y)}) \cdot \mathcal{U}_{\text{rqv}(d)}(x_{\text{pa}(\text{rqv}(d))}). \qquad \text{(A8)}\end{aligned}$$

Equation (A8) follows because $\mathcal{K}_3(x_{\text{fa}(N_1\cup N_2)\cap\text{fa}(N'_1\cap Y)})$ is, by construction, a strictly positive constant over $X_{\text{fa}(N_1\cup N_2)\cap\text{fa}(N'_1\cap Y)}$, which allows us to factor it out as some constant K. Denote by $N_2' = N_2 - \mathrm{rqv}(d)$ a set of nodes that are in N_2, but not $\mathrm{rqv}(d)$. We can rewrite the summation as

$$\mathcal{V}_2(x_{\mathrm{fa}(d)}) = K \cdot \mathcal{K}_2(x_{\mathrm{pa}(d)}) \Bigg\{ \mathcal{V}_4(x_{\mathrm{fa}(d)}) - \sum_{x_{\mathrm{fa}(N_1 \cup N_2) - \mathrm{fa}(d)}} \mathcal{P}_{N_1 * d}(x_{\mathrm{fa}(N_1 \cap Y)}) \cdot \mathcal{U}_{N_2'}(x_{\mathrm{pa}(N_2' \cap V)}) \Bigg\}$$
$$= K \cdot \mathcal{K}_2(x_{\mathrm{pa}(d)}) \{ \mathcal{V}_4(x_{\mathrm{fa}(d)}) - \mathcal{K}_4(x_{\mathrm{pa}(d)}) \}. \quad \text{(A9)}$$

Equation (A9) follows because $N_2' \cap V \subseteq V - \mathrm{rqv}(d)$, a nonrequisite value set for d given $\mathrm{pa}(d)$, and Lemma 2 implies that we can substitute the summation with some function $\mathcal{K}_4(x_{\mathrm{pa}(d)})$. The result follows because for each $x_{\mathrm{pa}(d)} \in X_{\mathrm{pa}(d)}$, $\mathcal{V}_4(x_{\mathrm{fa}(d)})$ is simply a linear transformation of $\mathcal{V}_2(x_{\mathrm{fa}(d)})$. Finally, the implication is in one direction because we might have $\mathcal{P}_{N_1 * d}(x_{\mathrm{fa}(d)}) \propto \mathcal{P}_{Z * \{N_1' \cup d\}}(x_{\mathrm{fa}(d)}) > 0$ when $\mathcal{P}_{Z * d}(x_{\mathrm{fa}(d)}) = 0$. □

References

Bellman, R. E. 1957. *Dynamic Programming*. Princeton University Press, Princeton, NJ.

Detwarasiti, A. 2005. Team decision analysis and influence diagrams. Ph.D. dissertation, Department of Management Science and Engineering, Stanford University, Stanford, CA.

Fagiuoli, E., M. Zaffalon. 1998. A note about redundancy in influence diagrams. *Internat. J. Approximate Reasoning* **19** 351–365.

Geiger, D., T. Verma, J. Pearl. 1990. Identifying independence in Bayesian networks. *Networks* **20** 507–534.

Howard, R. A. 1966a. Decision analysis: Applied decision theory. D. B. Hertz, J. Melese, eds. *Proc. 4th Internat. Conf. Oper. Res.* Wiley-Interscience, New York, 55–71.

Howard, R. A. 1966b. Information value theory. *IEEE Trans. Systems Sci. Cybernetics* **SSC-2** 22–26.

Howard, R. A. 1984. The evolution of decision analysis. R. A. Howard, J. E. Matheson, eds. *Readings on the Principles and Applications of Decision Analysis I*. Strategic Decision Group, Menlo Park, CA, 6–16.

Howard, R. A., J. E. Matheson. 1984/2005. Influence diagrams. R. A. Howard, J. E. Matheson, eds. *Readings on the Principles and Applications of Decision Analysis II*. Strategic Decision Group, Menlo Park, CA, 719–762. Reprinted, *Decision Anal.* **2**(3) 127–143.

Jensen, F., F. V. Jensen, S. L. Dittmer. 1994. From influence diagrams to junction trees. R. L. de Mantaras, D. Poole, eds. *Proc. 10th Conf. Uncertainty in Artificial Intelligence*. Morgan Kaufmann, San Francisco, CA, 367–373.

Koller, D., B. Milch. 2001. Multi-agent influence diagrams for representing and solving games. B. Nebel, ed. *Proc. 17th Internat. Joint Conf. Artificial Intelligence*. Morgan Kaufmann, San Francisco, CA, 1027–1034.

Koller, D., B. Milch. 2003. Multi-agent influence diagrams for representing and solving games. *Games Econom. Behavior* **45** 181–221.

Kuhn, H. W. 1953. Extensive games and the problem of information. H. W. Kuhn, A. W. Tucker, eds. *Contributions to the Theory of Games II*. Princeton University Press, Princeton, NJ, 193–216.

Lauritzen, S. L., D. Nilsson. 2001. Representing and solving decision problems with limited information. *Management Sci.* **47** 1235–1251.

Marschak, J., R. Radner. 1972. *Economic Theory of Teams*. Yale University Press, New Haven, CT.

McKelvey, R. D., A. McLennan. 1996. Computation of equilibria in finite games. H. M. Amman, D. A. Kendrick, J. Rust, eds. *Handbook of Computational Economics I*. Elsevier Science, Amsterdam, Netherlands, 87–142.

McKelvey, R. D., A. M. McLennan, T. L. Turocy. 2004. *GAMBIT: Software Tools for Game Theory*. http://econweb.tamu.edu/gambit.

Nash, J. F. 1951. Non-cooperative games. *Ann. Math.* **54** 286–295.

Nielsen, T. D. 2001. Decomposition of influence diagrams. S. Benferhat, P. Besnard, eds. *Proc. 6th Eur. Conf. Symbolic and Quantitative Approaches to Reasoning with Uncertainty*. Springer-Verlag, Berlin, Germany, 144–155.

Nielsen, T. D., F. V. Jensen. 1999. Well defined decision scenarios. K. B. Laskey, H. Prade, eds. *Proc. 15th Conf. Uncertainty in Artificial Intelligence*. Morgan Kaufmann, San Francisco, CA, 502–511.

Nilsson, D., S. L. Lauritzen. 2000. Evaluating influence diagrams using LIMIDs. C. Boutilier, M. Goldszmidt, eds. *Proc. 16th Conf. Uncertainty in Artificial Intelligence*. Morgan Kaufmann, San Francisco, CA, 436–445.

Olmsted, S. M. 1983. On representing and solving decision problems. Ph.D. dissertation, Department of Engineering-Economic Systems, Stanford University, Stanford, CA.

Pearl, J. 1988. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan Kaufmann, San Mateo, CA.

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

Shachter, R. D. 1986. Evaluating influence diagrams. *Oper. Res.* **34** 871–882.

Shachter, R. D. 1988. Probabilistic inference and influence diagrams. *Oper. Res.* **36** 589–604.

Shachter, R. D. 1998. Bayes-Ball: The rational pastime. G. F. Cooper, S. Moral, eds. *Proc. 14th Conf. Uncertainty in Artificial Intelligence*. Morgan Kaufmann, San Francisco, CA, 480–487.

Shachter, R. D. 1999. Efficient value of information computation. K. B. Laskey, H. Prade, eds. *Proc. 15th Conf. Uncertainty in Artificial Intelligence*. Morgan Kaufmann, San Francisco, CA, 594–601.

Shachter, R. D., M. A. Peot. 1992. Decision making using probabilistic inference methods. D. Dubois, M. P. Wellman, B. D'Ambrosio, P. Smets, eds. *Proc. 8th Conf. Uncertainty in Artificial Intelligence*. Morgan Kaufmann, San Mateo, CA, 276–283.

Shenoy, P. P. 1992. Valuation-based systems for Bayesian decision analysis. *Oper. Res.* **40** 463–484.

Tatman, J. A., R. D. Shachter. 1990. Dynamic programming and influence diagrams. *IEEE Trans. Systems, Man, Cybernetics* **20** 365–379.

von Neumann, J., O. Morgenstern. 1947. *Theory of Games and Economic Behavior*. Princeton University Press, Princeton, NJ.

Zhang, N. L. 1998. Probabilistic inference in influence diagrams. G. F. Cooper, S. Moral, eds. *Proc. 14th Conf. Uncertainty in Artificial Intelligence*. Morgan Kaufmann, San Francisco, CA, 514–522.

Zhang, N. L., D. Poole. 1992. Stepwise-decomposable influence diagrams. B. Nebel, C. Rich, W. R. Swartout, eds. *Proc. 3rd Internat. Conf. Principles Knowledge Representation and Reasoning*. Morgan Kaufmann, San Mateo, CA, 141–152.

Zhang, N. L., R. Qi, D. Poole. 1994. A computational theory of decision networks. *Internat. J. Approximate Reasoning* **11** 83–158.