ELSEVIER

Representing asymmetric decision problems using coarse valuations

Liping Liu*, Prakash P. Shenoy

Department of Management, University of Akron, Akron, OH 44325-4801, USA
Kansas University Business School, 1300 Sunnyside Ave., Room 345 Su, Lawrence, KS 66045-2003, USA

Abstract

A valuation-based system approach to knowledge representation has shown its advantages in improving computational efficiency and in allowing many decision models including belief networks. This study applies the Dempster–Shafer theory of belief functions and extends its framework to allow coarse valuations, which admit incomplete specification of probabilities and utilities and, therefore, are more flexible in representing asymmetric decision problems. It presents an algorithm for making inferences and decisions in systems using coarse valuations. It shows that a coarse valuation-based system provides a most natural and compact representation of decision problems.
© 2003 Elsevier B.V. All rights reserved.

1. Introduction

In a valuation-based system [17], we represent knowledge by functions called *valuations*. We draw inferences from such systems using two operations called combination and marginalization. Combination corresponds to the aggregation of knowledge and marginalization corresponds to the coarsening of knowledge. The valuation-based system approach to knowledge representation is expressive in the sense that it allows many decision models including the Dempster–Shafer theory of belief functions [21] and

possibility theory [5,18]. Its representation for the Bayesian decision model [17,19] is similar to influence diagrams [20].

A graphical depiction of a valuation-based system is called a *valuation network*. Like an influence diagram, a valuation network compactly represents the probabilistic dependence in a problem domain. It allows more efficient computation when implementing them in decision support systems. However, a critical hurdle in the implementation is their failure to efficiently capture asymmetric dependencies between events and choices. In order to work around this problem, systems developers have to create dummy events and acts, and degenerate probabilities and utilities. In some highly asymmetric problems, such artifacts constitute a major portion of problem inputs and cause unnecessary computation. More importantly, they are artificial constructs and hinder effec-

---

* Corresponding author. Kansas University Business School, 1300 Sunnyside Ave., Room 345 Su, Lawrence, KS 66045-2003, USA. Tel.: +1-330-972-5947; fax: +1-330-972-6588.

E-mail addresses: liu@acm.org (L. Liu), pshenoy@ku.edu (P.P. Shenoy).

tive communication between decision makers and systems developers.

This paper extends the framework of valuation-based systems to represent asymmetric decision problems using a belief function-like calculus, called *coarse valuations*. The calculus admits incomplete specifications of probabilities such as a vacuous belief for a state space and a non-additive probability for a subset of the state space. The notion of coarse valuations provides a natural and compact way of representing asymmetric decision problems. It reduces the need for artificial events and acts, and degenerate probabilities and utilities. It captures both structural asymmetry and numerical coalescence at the functional level so that the graphical representation of an asymmetric problem is as compact as that of a symmetric one.

This paper has two distinctive contributions to model representation in general and the techniques of valuation-based systems in particular. First, it proposes a new representation of domain knowledge for decision support by allowing coarse valuations, which include probabilities and belief functions as special cases. The new representation improves the flexibility and efficiency of decision-model representation. It is able to represent asymmetric decision problems as compactly as a symmetric valuation network while reducing 60–96% of dummy events and acts, and degenerate probabilities and utilities. Second, it re-defines combination and marginalization operations, re-describes Shenoy's fusion algorithm [17,19], and proposes a complete algorithm for coarse valuation-based systems.

An outline of this paper is as follows. In Section 2, we review related studies. In Section 3, we illustrate the nature of asymmetric problems and describe a valuation network representation using coarse valuations. In Section 4, we define three basic operations for coarse valuations. In Section 5, we propose a fusion algorithm for coarse valuation-based systems. In Section 6, we show the correctness of the fusion algorithm. In Section 7, we illustrate the algorithm by solving an example. Finally, in Section 8, we make some concluding remarks.

2. Related studies

Asymmetry can be viewed as an unbalanced dependence among uncertain events and acts. From this viewpoint, a problem is asymmetric if it has a decision tree representation that is unbalanced, i.e., not all scenarios contain the same sequence of acts and events. Therefore, to represent asymmetry, one has to capture the event/act dependence. Based on this idea, several methods have been proposed. For example, Call and Miller [2] describe a representation using separate decision trees and influence diagrams. Fung and Shachter [6] and Qi et al. [15] propose an adaptation of influence diagrams. Smith et al. [23] describe a representation using an influence diagram and several conditional distribution trees, which represent local event/act dependence. Covaliu and Oliver [3] use a so-called sequential decision diagram to represent the event/act dependence. Shenoy [22] proposes a new type of valuations called indicator valuations to represent the event/act dependence. Liu and Shenoy [12] decompose an asymmetric problem into several symmetric sub-problems and use a symmetric valuation network to represent each of the symmetric sub-problems. Demirer and Shenoy [4] propose a new representation called sequential valuation networks that is a hybrid of sequential decision diagrams [3] and symmetric valuation networks [17]. Nielsen and Jensen [14] describe a new representation called asymmetric influence diagrams that is a hybrid of influence diagram and sequential decision diagrams. Finally, Bielza and Shenoy [1] compare decision trees, conditional distribution trees [23], sequential decision diagrams [4], and indicator valuations [22].

3. Knowledge representation

A coarse valuation-based system represents an inference and decision-making problem using both a graphical depiction called a *coarse valuation network* and a set of normalized tables. By using set-to-point mappings, each table represents an uncertainty or payoff valuation (function). By using variable and valuation nodes as well as precedence constraints, the coarse valuation network provides a control structure for problem solving.

3.1. An example

Before we discuss our representation technique, let us describe the used car buyer's problem [7] as an

example of a highly asymmetric decision problem [15,22,23].

Joe is considering buying a 3-year-old Spartan sedan at a price of US$1000. The going rate for a similar car in the used car market is US$1100. Joe is unsure whether the car is a "lemon" or a "peach." Of the 10 major mechanical systems in the car, a peach has only 1 defective system, while a lemon has defects in 6 of the 10 systems. From historical data, 20% of Spartan cars were lemons and the other 80% were peaches. The cost of repairing one defect is US$40 and the cost of repairing six defects is US$200. For an additional US$60, Joe can buy the car from the dealer with an anti-lemon guarantee. The anti-lemon guarantee will normally pay for 50% of the repair cost, but, if the car is a lemon, it will pay the full repair cost of US$200.

The dealer gives Joe an hour to have the car examined by a mechanic. The mechanic suggests three alternative diagnostic tests—steering, transmission, fuel, and electrical—to determine the car's condition. All tests are able to detect the defects (if any) of their testing systems. The steering test costs US$9, the transmission test costs US$10, and the fuel and electrical test costs US$13. After reviewing the result of transmission test, for an additional US$4, Joe can proceed to have the differential system also tested.

A decision tree representation and solution for this problem can be found in Ref. [7]. A coarse valuation network representation is shown in Fig. 1. The optimal strategy is to do the fuel and electrical test; if both systems are non-defective, then buy with no anti-lemon guarantee, else buy with anti-lemon guarantee. The maximum expected payoff is US$32.89.

3.2. Variables

We represent a decision variable in a valuation network by a rectangular-shaped node. We use symbol Ω_D to denote the set of all alternatives available at decision D. By making a decision D, we mean that the decision-maker chooses one and only one of the elements of Ω_D. We call Ω_D the state space of D.

In the used car buyer's problem, there are three decision variables T, D, and B, respectively, representing the initial test decision (which test to conduct?), the differential test decision (whether or not to continue with the differential test?), and the final purchase decision (whether or not to buy the car or buy the car with an anti-lemon guarantee?). The state space for T has four acts: no test (t_0), do the steering test (t_1), do the fuel and electrical test (t_2), and do the transmission test (t_3). The state space for D has two acts: do the differential test (d_1) and no test (d_0). The state space for B has three alternatives: not buy the car (b_0), buy the car with anti-lemon guarantee (b_1), and buy the car without the anti-lemon guarantee (b_2).

We represent a random variable in a valuation network by an elliptical-shaped node. We use the symbol Ω_R to denote the set of all possible values for random variable R. We assume that one and only one of the elements of Ω_R is the true value of R. We call Ω_R the state space for R.
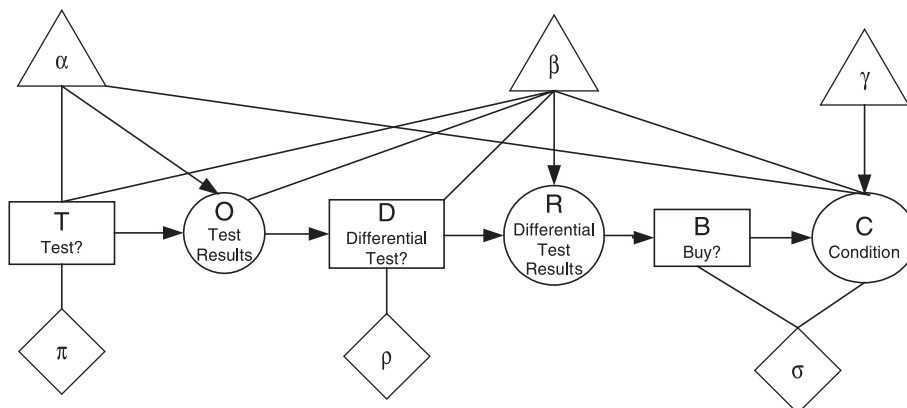

Fig. 1. A coarse valuation network for the used car buyer's problem.

In the used car buyer's problem, there are three random variables: the result of the first test O, the result of the differential test R, and the car's condition C. The state space for O has three elements: zero defect, one defect, and two defects. The state space for R has two elements: defective (0) and non-defective (1). The state space for C has two elements: peach (c_p) and lemon (c_l).

3.3. Coarse payoff valuations

For any finite set of variables X, let Ω_X denote its state space. Let x denote a state in Ω_X and $\mathbf{x}$ a subset of values in Ω_X.

A coarse payoff valuation is a set to point mapping. Suppose π is a payoff valuation for X. Then, $\pi(\mathbf{x})$ measures the payoff to the decision maker if $x \in \mathbf{x}$. The values of a payoff valuation are consequences, for example, utilities and profits.

In a valuation network, we use a diamond to represent a payoff valuation. To permit the identification of all variables in its domain, we draw undirected edges between a payoff valuation node and all the variable nodes in its domain. In the used car buyer's problem, the payoff valuations π, ρ, and σ are shown in Fig. 1, and are numerically specified in Table 1. The specifications are self-explanatory except for $\sigma(b_0, \Omega_C) = 0$, which means that, if the decision B is not to buy the car, the payoff $\sigma = 0$ regardless of the car's condition (C). In other words, given $B = b_0$, σ is independent of C. Note that σ is not totally independent of C; if $B = b_1$, then σ is dependent on C. To represent such an asymmetric event/act dependency, in a traditional valuation network or an influence diagram, $\sigma(b_0, \Omega_C) = 0$ have to be specified as $\sigma(b_0, c_p) = 0$ and $\sigma(b_0, c_l) = 0$.

The above-mentioned asymmetry is common in decision problems: the payoff of not marketing a product is independent of the market condition; the payoff of not drilling a well is independent of whether the hole has oil or not; the payoff for a 3-year bank certificate of deposit is independent of stock market conditions, macroeconomic conditions, and government controls. In general, assume π is a payoff valuation for X. If there exists a subset $\mathbf{x}$ such that for any values $x \in \mathbf{x}$, $\pi(x)$ is a constant k, then we coalesce them into one specification: $\pi(\mathbf{x}) = k$. This representation applies to the special case when $\mathbf{x}$ is a singleton or an entire state space.

For each payoff valuation π for X, we require that, for any $x \in \Omega_X$, there exists one and only one subset $\mathbf{x}$ such that $x \in \mathbf{x}$ and $\pi(\mathbf{x})$ is specified. We call this requirement *the payoff completeness*. It stipulates that there be one payoff value assigned to every point x in Ω_X.

3.4. Coarse probability valuations

In a coarse valuation network, uncertainties are represented by coarse probability valuations. Similar to a coarse payoff valuation, a coarse probability valuation is a real-valued mapping. Let H and T be two disjoint sets of variables. Suppose α is a probability valuation for H given T. Then, $\alpha(\mathbf{h}|\mathbf{t})$ measures the probability that $H \in \mathbf{h}$ given $T \in \mathbf{t}$.

Coarse probabilities have the flavor of *conditional belief functions* [13]. Mathematically, a coarse probability valuation must satisfy a set of axioms similar to that for the Dempster–Shafer theory of belief functions [16]. Semantically, the assignment of coarse probabilities follows the notion of evidential support [16]; we assign a positive probability to $(\mathbf{h}|\mathbf{t})$ if there is a piece of evidence that partially supports $(\mathbf{h}|\mathbf{t})$. We call an assertion $(\mathbf{h}|\mathbf{t})$ *focal* if $\alpha(\mathbf{h}|\mathbf{t}) > 0$. Correspondingly, we call $\mathbf{h}$ the *focal head* and $\mathbf{t}$ the *focal tail*.

The notion of coarse probabilities has three non-trivial special cases. First, if every focal head is Ω_H, then α is a *vacuous probability function* of H given T. Second, if every focal tail is Ω_T, then our knowledge about H is irrelevant to that about T. α is a Dempster–Shafer basic probability assignment function for H. Finally, if every focal head and focal tail is a singleton, then α is a conditional probability distribution function.

In a coarse valuation network, we use a triangular node to represent a coarse probability valuation. Suppose α is a conditional valuation of

Table 1
Coarse payoff valuations for the used car buyer's problem

T	π	$\{B, C\}$		σ	D	ρ
t_0	0	b_2	c_p	60	d_0	0
t_1	-9	b_2	c_l	-100	d_1	-4
t_2	-13	b_1	c_p	20		
t_3	-10	b_1	c_l	40		
		b_0	Ω_C	0		

H given T. To permit the identification of variables, we draw a directed edge pointing to each variable in H and an undirected edge to each variable in T from the α node. In the used car buyer's problem, there are three coarse valuations α, β, and γ as shown in Fig. 1 and numerically specified in Table 2, where α is a conditional probability valuation of O given $\{T, C\}$, β is a conditional probability valuation of R given $\{T, O, D, C\}$, and γ is a marginal probability valuation for C.

Note that, following the convention of Lauritzen and Spiegelhalter [10], we write the coarse probability valuations using the potential form. For example, $\beta(\Omega_T, \Omega_O, d_0, \Omega_R, \Omega_C) = 1.0$ should be interpreted as $\beta(\Omega_R | \Omega_T \times \Omega_O \times d_0 \times \Omega_C) = 1.0$. In addition, as in belief functions, a coarse probability valuation is only specified for its focal elements. Therefore, a table of coarse probability valuations has no zero values.

Much of Table 2 is self-evident. For example, $\gamma(\{c_p\}) = 0.8$ and $\gamma(\{c_l\}) = 0.2$ represent the statistic knowledge that 20% of Spartan cars are lemons. Note that a differential test is conducted only when the transmission test is initially chosen. Since a peach has one defective system, there will be a 1/9 chance that the differential system is bad if the initial transmission test does not spot any defect. On the other hand, the probability vanishes if the transmission test fails. Therefore, we have $\beta(t_3, 0, d_1, 1, c_p) = 1/9$, $\beta(t_3, 0, d_1, 0, c_p) = 8/9$, and $\beta(t_3, 1, d_1, 0, c_p) = 1.0$. The last four specifications of β in Table 2 can be interpreted in a similar manner.

The first value of β, i.e., $\beta(\Omega_T, \Omega_O, d_0, \Omega_R, \Omega_C) = 1.0$, needs some explanations. Note that if no differential system test (d_0) is conducted, there will be no differential test results. In the parlance of belief

functions, we have a vacuous belief about the differential test result (R) given non-differential system test (d_0). Therefore, the whole state space of R, Ω_R, is a focal element and takes the whole belief mass. In addition, given no differential test, $D = d_0$, our knowledge about the test results is irrelevant to that about the initial test (T), the initial test result (O), and the car's condition (C). In other words, the above vacuous belief function is independent of T, O, and C given d_0. Therefore, we have a basic probability assignment $\beta(\Omega_T, \Omega_O, d_0, \Omega_R, \Omega_C) = 1.0$.

It is interesting to note how a traditional valuation-based system and an influence diagram represent the same uncertainties. First, we need to create an artificial state "no result" (nr) for both O and R. Then, the above $\beta(\Omega_T, \Omega_O, d_0, \Omega_R, \Omega_C) = 1.0$ would have to be replaced by 96 degenerate probabilities as follows: $\beta(t, o, d_0, \text{nr}, c) = 1.0$, $\beta(t, o, d_0, 0, c) = 0$, for any $t \in \Omega_T$, $o \in \{\text{nr}, 0, 1, 2\}$, and $c \in \Omega_C$. In addition, because of the introduction of the artificial state nr, we need to specify $\beta(t, o, d_1, 0, c) = 0$ for $o \in \{\text{nr}, 0, 1, 2\}$ and $c \in \Omega_C$. Furthermore, for $t \in \{t_0, t_1, t_2\}$, there will be no differential test to follow. To make β complete, we need to add 72 more numerical specifications as $\beta(t, o, d_1, r, c) = 0$ for $t \in \{t_0, t_1, t_2\}$, $o \in \{\text{nr}, 0, 1, 2\}$, $r \in \{\text{nr}, 0, 1\}$, and $c \in \{c_p, c_l\}$. In total, we need 192 specifications for the valuation β to be completely specified. Compared with the eight specifications in Table 2, the existing approach demands a lot of extra work in representing such a piece of simple knowledge. Even worse is that many of the extra specifications are artificially created for the purpose of fitting in the model and carry neither useful data for computation nor semantic meaning to a decision maker.

The coarse probability valuation α in Table 2 represents the conditional probability of the test result O given the initial test T and the car's condition C. Of 10 major systems, a peach has only 1 defective system. Thus, if the electric system is bad, the fuel system will be not and vice versa. Thus, if the fuel and electric test t_1 is done, there are two possible test results: neither the fuel nor the electric system is bad ($O = 0$), either the fuel or the electric system is bad ($O = 1$). By doing little analysis, we have $\alpha(t_2, 0, c_p) = 0.8$ and $\alpha(t_2, 0, c_p) = 0.2$. On the other hand, a lemon can have up to six defective systems. If the fuel and electric test (t_2) is done, there are three possible

Table 2
Coarse probability valuations for the used car buyer's problem

{T, O, C}			α	{T, O, D, R, C}					β	C	γ
t_0	Ω_O	Ω_C	1	Ω_T	Ω_O	d_0	Ω_R	Ω_C	1	c_p	0.8
$\{t_1, t_3\}$	0	c_p	0.9	t_3	0	d_1	0	c_p	8/9	c_l	0.2
$\{t_1, t_3\}$	0	c_l	0.4	t_3	0	d_1	1	c_p	1/9		
$\{t_1, t_3\}$	1	c_p	0.1	t_3	1	d_1	0	c_p	1		
$\{t_1, t_3\}$	1	c_l	0.6	t_3	0	d_1	0	c_l	3/9		
t_2	0	c_p	0.8	t_3	0	d_1	1	c_l	6/9		
t_2	1	c_p	0.2	t_3	1	d_1	0	c_l	4/9		
t_2	0	c_l	0.13	t_3	1	d_1	1	c_l	5/9		
t_2	1	c_l	0.53								
t_2	2	c_l	0.33								

test results: neither the fuel nor the electric system is bad ($O = 0$), either the fuel or the electric system is bad ($O = 1$), or both are bad ($O = 2$). Thus, we have the last three specifications of α in Table 2 (hypergeometric distribution with $N = 10$, $p = 0.6$, $n = 2$).

If no initial test is done (t_0), one is ignorant about the test result and the belief is independent of the car's condition. Therefore, we have $\alpha(t_0, \Omega_O, \Omega_C) = 1.0$.

If we do the steering test (t_1) or the transmission test (t_3), the probability of finding zero defects is 0.9 given that the car is a peach (c_p). Therefore, we have $\alpha(t_1, 0, c_p) = 0.9$ and $\alpha(t_3, 0, c_p) = 0.9$. They can be coalesced into one specification $\alpha(\{t_1, t_3\} \times \{0\} \times \{c_p\}) = 0.9$, which means that our belief about the assertion $(\{0\} | \{t_1, t_3\} \times \{c_p\})$ is 0.9 according to the notion of coarse probability valuations. Similarly, we can obtain the remaining specifications of α in Table 2.

If α is specified as a probability function, we would have to first create an artificial state "no result" (nr) for O and then specify 40 numerical values; among them 26 are degenerate.

In sum, the coarse valuation network is very expressive and compact in representing asymmetric decision problems. First of all, it does not need to create any artificial states or acts such as "no result" if no test is done. Second, it eliminates all zero probabilities from further manipulations. Third, two or more numerical specifications may be coalesced into one as we did for the payoff valuation σ in Table 1 and the valuation α in Table 2. As illustrated, by representing the above α, β, and γ as coarse probabilities, the valuation network can successfully avoid most of the degenerate probabilities and all the artificial states.

To permit uncertainty reasoning and problem solving in a coarse valuation-based system, we require coarse probability valuations to satisfy three conditions: the *belief completeness condition*, *the compatibility condition*, and *the expectability condition*. Like the payoff completeness condition, the belief completeness condition stipulates that, for each random variable R, there exists at least one conditional probability function of H given T such that $R \in H$. The compatibility condition ensures no cycles of conditional probabilities such that α_1 is a conditional probability function of X given Y while α_2 is of Y given X. In addition, the compatibility condition ensures the inferability of two logical assertions. For example, to make an inference based on two rules, we naturally

need the conclusion of the first rule to be more specific than the condition of the second rule. The expectability condition enables us to multiply a coarse payoff valuation with a coarse probability valuation. Without it, we cannot apply expectation operations.

3.5. Precedence constraints

Besides variables, beliefs, and payoffs, an important ingredient in a coarse valuation network is information constraints. Some decisions have to be made before the observation of some uncertain events, and other decisions can be postponed until after some events are observed. In the used car buyer's problem, for example, the car's condition is revealed only after we purchase the car or perhaps never revealed. And the decision whether to buy the car or not may be postponed until the test result is revealed.

If a decision maker expects to be informed of the true value of random variable R before he makes a decision D, then we represent this situation by the binary relation $R \to D$ (read as R precedes D). On the other hand, if a random variable R is only revealed after a decision D is made or perhaps never revealed, then we represent this situation by the binary relation $D \to R$. It is possible that in some problems, we may have precedence constraints between two decision nodes or between two random variable nodes. For example, if random variable R_2 is only revealed after random variable R_1 is revealed, we represent this by the relation $R_1 \to R_2$.

In the used car buyer's problem, we have the precedence constraints $T \to O$, $O \to D$, $D \to R$, $R \to B$, and $B \to C$. The first test result (O) is only revealed after we make the decision to test either of steering, transmission, and fuel and electrical systems (T). The decision to do the differential test (D) is only made after observing the transmission test results (O). Finally, the decision to buy the car (B) is made after observing all test results and the car's condition is revealed after the decision of purchasing (B).

A problem can be incorrectly over-constrained permitting no solution. For example, if $D \to R$ and $R \to D$, then this will preclude a solution. Therefore, we do not permit such precedence constraints. What restrictions do we need to impose on the precedence relation $\to$? We require three conditions. First, we require that the transitive closure of $\to$, denoted by $>$,

is a partial order. We call this first condition *the partial order condition*. Second, we require that this partial order > is such that for any decision variable D and any random variable R, either $D>R$ or $R>D$. We call this second condition *the perfect recall condition*. Third, if there exists a conditional probability function of H given T, which contains a decision variable D, and D is minimum in $H \cup T$, then we require that the probability function is independent of D. We call this third condition *the semantic condition*.

Before we explain the reasons for these three conditions, we need the notions of *transitive closure* and *partial ordering*:

The transitive closure of $\rightarrow$ is defined as follows:

- $X>Y$ whenever $X \rightarrow Y$ and
- $X>Y$ whenever there exists a variable Z such that $X>Z$ and $Z>Y$.

A binary relation > is a partial order if it is irreflexive and transitive.

The reason for the partial order requirement is obvious. The reason for the perfect recall condition is as follows. Given the meaning of the precedence relation $\rightarrow$, for any decision variable D and any random variable R, either R is known when decision D has to be made, or not. This translates to either $R>D$ or $D>R$. Finally, the semantic condition is dictated by the meaning that a probability function measures the degrees of beliefs about uncertain propositions.

4. Knowledge operations

In a coarse valuation-based system, we represent knowledge by coarse valuations. We draw inferences and make decisions using two operations called combination and marginalization. In this section, we describe these two operations.

4.1. Combination

Suppose π is a payoff valuation of X and σ is a payoff valuation of Y. Their combination, denoted by $\pi \otimes \sigma$ or $\sigma \otimes \pi$, is a payoff valuation of $Z = X \cup Y$ defined as follows:

$$(\pi \otimes \sigma)(\mathbf{z}) = \pi(\mathbf{x}) + \sigma(\mathbf{y}), \qquad (1)$$

where $\mathbf{x}$ and $\mathbf{y}$ are respectively the projection of $\mathbf{z}$ to X and Y, which are defined as follows:

$$\mathbf{z}^{\downarrow X} = \{x \mid \mathbf{z} \cap (\{x\} \times \Omega_{Z-X}) \neq \emptyset\},$$

$$\mathbf{z}^{\downarrow Y} = \{y \mid \mathbf{z} \cap (\{y\} \times \Omega_{Z-Y}) \neq \emptyset\}. \qquad (2)$$

Note that Eq. (1) is well defined in the sense that, for every non-empty subset $\mathbf{z} \subset \Omega_Z$, there cannot be two or more pairs of subsets $(\mathbf{x}, \mathbf{y})$ such that $\mathbf{z}^{\downarrow X} = \mathbf{x}$ and $\mathbf{z}^{\downarrow Y} = \mathbf{y}$. Otherwise, one can verify that it violates the payoff completeness condition.

Suppose π is a payoff valuation of X and α is a probability valuation bearing on Y. Their combination, denoted by $\pi \otimes \alpha$ or $\alpha \otimes \pi$, is a payoff valuation of $Z = X \cup Y$ defined as follows:

$$(\pi \otimes \alpha)(\mathbf{z}) = \pi(\mathbf{x})\alpha(\mathbf{y}). \qquad (3)$$

where $\mathbf{z}^{\downarrow X} = \mathbf{x}$ and $\mathbf{z}^{\downarrow Y} = \mathbf{y}$. Because of the belief completeness condition and the expectability condition, Eq. (3) is well defined.

Suppose α and β are two compatible probability valuations respectively defined on X and Y. Their combination, denoted by $\alpha \otimes \beta$ or $\beta \otimes \alpha$, is a probability valuation on $Z = X \cup Y$:

$$(\alpha \otimes \beta)(\mathbf{z}) = \Sigma \{\alpha(\mathbf{x})\beta(\mathbf{y}) \mid \mathbf{z}^{\downarrow X} = \mathbf{x}, \ \mathbf{z}^{\downarrow Y} = \mathbf{y}\}. \qquad (4)$$

One can prove that Eq. (4) generalizes both Dempster's rule of combination and Bayesian rule of conditioning. If α and β are both Dempster–Shafer belief functions, Eq. (4) is the same as Dempster's rule except for normalization. On the other hand, assume that the heads of α and β are disjoint as for conditional probabilities. Because of the compatibility of α and β, there exists a unique pair of focal elements $(\mathbf{x}, \mathbf{y})$ that satisfies $\mathbf{z}^{\downarrow X} = \mathbf{x}$ and $\mathbf{z}^{\downarrow Y} = \mathbf{y}$ for $\mathbf{z} \subset \Omega_Z$. Then, Eq. (4) is reduced to:

$$(\alpha \otimes \beta)(\mathbf{z}) = \alpha(\mathbf{x})\beta(\mathbf{y}). \qquad (5)$$

The semantics of combination is justified as follows. Eq. (1) basically assumes the additive decomposition of a multi-attribute value function [8]. Eq. (3) computes expected utilities and, therefore, assumes a normative preference structure when making choices under uncertainty [11]. Eq. (4) is consistent with Dempster's rule of combination when combining marginal belief functions and the Bayesian rule of conditioning when combining conditional probability functions.

Table 3 illustrates the combination of α and β and Table 4 shows the combination of $\alpha \otimes \beta$ and γ. Since every random variable belongs to the head of one and only one conditional probability valuation, Eq. (5) is applicable. I.e., we can simply intersect focal elements and multiply their corresponding probabilities. Note that the combination of coarse probability valuations is associative and commutative. The order of combining them does not matter and $(\alpha \otimes \beta) \otimes \gamma$ can be equivalently written as $\alpha \otimes \beta \otimes \gamma$.

4.2. Marginalization

Marginalization means deleting variables from a valuation or projecting a valuation to its partial domains. Depending on the type of variable being deleted, the definition of marginalization is different.

When a decision variable D is to be eliminated from a payoff valuation, marginalization corresponds to the maximization of the payoff valuation by selecting a subset from Ω_D for D. Precisely, assume π is a payoff valuation of X and $D \in X$. The marginal of π for $Y = X - \{D\}$, denoted by $\pi^{\downarrow Y}$, is a payoff valuation for Y defined as follows:

$$\pi^{\downarrow Y}(\mathbf{y}) = \max\{\pi(\mathbf{x}) \mid \mathbf{y} \subset \mathbf{x}^{\downarrow Y}\} \qquad (6)$$

Eq. (6) corresponds to a decision tree solution process called "folding back." It assumes that the

Table 3
Computation of $\alpha \otimes \beta$

{T, O, D, R, C}					$\alpha \otimes \beta$
t_0	Ω_O	d_0	Ω_R	Ω_C	1.00
$\{t_1, t_3\}$	0	d_0	Ω_R	c_p	0.90
$\{t_1, t_3\}$	0	d_0	Ω_R	c_l	0.40
$\{t_1, t_3\}$	1	d_0	Ω_R	c_p	0.10
$\{t_1, t_3\}$	1	d_0	Ω_R	c_l	0.60
t_3	0	d_1	0	c_p	0.80
t_3	0	d_1	0	c_l	0.13
t_3	0	d_1	1	c_p	0.10
t_3	0	d_1	1	c_l	0.27
t_3	1	d_1	0	c_p	0.10
t_3	1	d_1	0	c_l	0.27
t_3	1	d_1	1	c_l	0.33
t_2	0	d_0	Ω_R	c_p	0.80
t_2	0	d_0	Ω_R	c_l	0.13
t_2	1	d_0	Ω_R	c_p	0.20
t_2	1	d_0	Ω_R	c_l	0.53
t_2	2	d_0	Ω_R	c_l	0.33

Table 4
Computation of τ, δ, and τ/δ

{T, O, D, R, C}					$\tau = (\alpha \otimes \beta) \otimes \gamma$	$\delta = \tau^{\downarrow \{T, O, D, R\}}$	τ/δ
t_0	Ω_O	d_0	Ω_R	c_p	0.800	1.000	0.800
t_0	Ω_O	d_0	Ω_R	c_l	0.200		0.200
$\{t_1, t_3\}$	0	d_0	Ω_R	c_p	0.720	0.800	0.900
$\{t_1, t_3\}$	0	d_0	Ω_R	c_l	0.080		0.100
$\{t_1, t_3\}$	1	d_0	Ω_R	c_p	0.080	0.200	0.400
$\{t_1, t_3\}$	1	d_0	Ω_R	c_l	0.120		0.600
t_3	0	d_1	0	c_p	0.640	0.666	0.961
t_3	0	d_1	0	c_l	0.026		0.039
t_3	0	d_1	1	c_p	0.080	0.134	0.600
t_3	0	d_1	1	c_l	0.054		0.400
t_3	1	d_1	0	c_p	0.080	0.134	0.600
t_3	1	d_1	0	c_l	0.054		0.400
t_3	1	d_1	1	c_l	0.066	0.066	1.000
t_2	0	d_0	Ω_R	c_p	0.640	0.666	0.961
t_2	0	d_0	Ω_R	c_l	0.026		0.039
t_2	1	d_0	Ω_R	c_p	0.160	0.266	0.600
t_2	1	d_0	Ω_R	c_l	0.106		0.400
t_2	2	d_0	Ω_R	c_l	0.066	0.066	1.000

decision-maker wants to maximize the payoff valuation. This is true if the payoff values represent utility, profits, probability of success, etc. On the other hand, if the decision-maker wishes to minimize the payoff values such as disutility, cost, and probability of failure, then we need to substitute min for max in Eq. (6).

When marginalizing a decision variable out of a payoff valuation, we need bookkeeping the subset $\mathbf{d}$, in which the maximum is achieved. Formally, we call the mapping $\psi_D: \mathbf{y} \to \mathbf{d}$ a solution for D with respect to π. Tables 5 and 6 in Section 7 show the marginalization of v to $\{T, O, D, R\}$ and the solution for B with respect to v.

Assume α is a probability valuation defined on X and is independent of D. To marginalize D out of α is to simply drop D out of X. Formally, the marginal of α for $Y = X - \{D\}$, denoted by $\alpha^{\downarrow Y}$, is a conditional probability function for Y defined as follows:

$$\alpha^{\downarrow Y}(\mathbf{y}) = \alpha(\mathbf{x}) \qquad \text{where } \mathbf{y} \subset \mathbf{x}^{\downarrow Y} \qquad (7)$$

As we will see in Section 6, we always marginalize a minimal variable out of a valuation. Therefore, according to the semantic condition, a probability valuation must be independent of D if it is to be deleted. Table 8 in Section 7 shows how $\eta = \theta^{\downarrow \{T, O\}}$ is obtained from θ in Table 7.

Table 5
The deletion of C and B

{T, O, D, R, B, C}						$(\tau/\delta)\otimes\sigma$	v	Ψ_B	$\varphi = v^{\downarrow\{T, O, D, R\}}$
t_0	Ω_O	d_0	Ω_R	b_2	c_p	48	28	b_2	28
t_0	Ω_O	d_0	Ω_R	b_2	c_1	-20			
t_0	Ω_O	d_0	Ω_R	b_1	c_p	16	24		
t_0	Ω_O	d_0	Ω_R	b_1	c_1	8			
t_0	Ω_O	d_0	Ω_R	b_0	c_p	0	0		
t_0	Ω_O	d_0	Ω_R	b_0	c_1	0			
$\{t_1, t_3\}$	0	d_0	Ω_R	b_2	c_p	54	44	b_2	44
$\{t_1, t_3\}$	0	d_0	Ω_R	b_2	c_1	-10			
$\{t_1, t_3\}$	0	d_0	Ω_R	b_1	c_p	18	22		
$\{t_1, t_3\}$	0	d_0	Ω_R	b_1	c_1	4			
$\{t_1, t_3\}$	0	d_0	Ω_R	b_0	c_p	0	0		
$\{t_1, t_3\}$	0	d_0	Ω_R	b_0	c_1	0			
$\{t_1, t_3\}$	1	d_0	Ω_R	b_2	c_p	24	-36		
$\{t_1, t_3\}$	1	d_0	Ω_R	b_2	c_1	-60			
$\{t_1, t_3\}$	1	d_0	Ω_R	b_1	c_p	8	32	b_1	32
$\{t_1, t_3\}$	1	d_0	Ω_R	b_1	c_1	24			
$\{t_1, t_3\}$	1	d_0	Ω_R	b_0	c_p	0	0		
$\{t_1, t_3\}$	1	d_0	Ω_R	b_0	c_1	0			
t_3	0	d_1	0	b_2	c_p	57.66	53.76	b_2	53.76
t_3	0	d_1	0	b_2	c_1	-3.9			
t_3	0	d_1	0	b_1	c_p	19.22	20.78		
t_3	0	d_1	0	b_1	c_1	1.56			
t_3	0	d_1	0	b_0	c_p	0	0		
t_3	0	d_1	0	b_0	c_1	0			
t_3	0	d_1	1	b_2	c_p	36	-4		
t_3	0	d_1	1	b_2	c_1	-40			
t_3	0	d_1	1	b_1	c_p	12	28	b_1	28
t_3	0	d_1	1	b_1	c_1	16			
t_3	0	d_1	1	b_0	c_p	0	0		

Table 6
The deletion of C and B (continued from Table 5)

{T, O, D, R, B, C}						$(\tau/\delta)\otimes\sigma$	v	Ψ_B	$\varphi = v^{\downarrow\{T, O, D, R\}}$
t_3	0	d_1	1	b_0	c_1	0			
t_3	1	d_1	0	b_2	c_p	36	-4		
t_3	1	d_1	0	b_2	c_1	-40			
t_3	1	d_1	0	b_1	c_p	12	28	b_1	28
t_3	1	d_1	0	b_1	c_1	16			
t_3	1	d_1	0	b_0	c_p	0	0		
t_3	1	d_1	0	b_0	c_1	0			
t_3	1	d_1	1	b_2	c_1	-100	-100		
t_3	1	d_1	1	b_1	c_1	40	40	b_1	40
t_3	1	d_1	1	b_0	c_1	0	0		
t_2	0	d_0	Ω_R	b_2	c_p	57.66	53.76	b_2	53.76
t_2	0	d_0	Ω_R	b_2	c_1	-3.9			
t_2	0	d_0	Ω_R	b_1	c_p	19.22	20.78		
t_2	0	d_0	Ω_R	b_1	c_1	1.56			
t_2	0	d_0	Ω_R	b_0	c_p	0	0		
t_2	0	d_0	Ω_R	b_0	c_1	0			
t_2	1	d_0	Ω_R	b_2	c_p	36	-4		
t_2	1	d_0	Ω_R	b_2	c_1	-40			
t_2	1	d_0	Ω_R	b_1	c_p	12	28	b_1	28
t_2	1	d_0	Ω_R	b_1	c_1	16			
t_2	1	d_0	Ω_R	b_0	c_p	0	0		
t_2	1	d_0	Ω_R	b_0	c_1	0			
t_2	2	d_0	Ω_R	b_2	c_1	-100	-100		
t_2	2	d_0	Ω_R	b_1	c_1	40	40	b_1	40
t_2	2	d_0	Ω_R	b_0	c_1	0	0		

When a chance variable is to be reduced, marginalization corresponds to the summation of a valuation over all the values of the chance variable. When the valuation is a probability valuation, the summation is simply the familiar marginalization operation in probability theory and the Dempster–Shafer theory of belief functions [9]. When the valuation is a payoff function, the summation represents a similar operation called "averaging out" in solving a decision tree.

Mathematically, assume π is a payoff valuation of X containing the chance variable R. The marginal of π for $Y = X - \{R\}$, denoted by $\pi^{\downarrow Y}$, is a payoff valuation for Y defined as follows:

$$\pi^{\downarrow Y}(\mathbf{y}) = \Sigma\{\pi(\mathbf{x}) \mid \mathbf{y} \subset \mathbf{x}^{\downarrow Y}\}. \qquad (8)$$

Assume α is a conditional probability function of H given T and H contains a chance variable R. Let

$X = H \cup T$, $K = H - \{R\}$, and $Y = K \cup T$. The marginal of α for Y, denoted by $\alpha^{\downarrow Y}$, is a conditional belief function of K given T defined as follows:

$$\alpha^{\downarrow Y}(\mathbf{y}) = \Sigma\{\alpha(\mathbf{x}) \mid \mathbf{y}^{\downarrow K} = \mathbf{x}^{\downarrow K}, \mathbf{y}^{\downarrow T} \subset \mathbf{x}^{\downarrow T}\}. \qquad (9)$$

The column δ in Table 4 shows the marginalization of $\alpha \otimes \beta \otimes \gamma$, which is a conditional probability valuation of $\{O, R, C\}$ given $\{T, D\}$, to $\{T, O, D, R\}$.

Table 7
The deletion of R

{T, O, D, R}				$\theta = \delta^{\downarrow\{T, O, D\}}$	δ/θ	$\varphi\otimes(\delta/\theta)$	ψ
t_0	Ω_O	d_0	Ω_R	1.000	1.000	28.00	28.00
$\{t_1, t_3\}$	0	d_0	Ω_R	0.800	1.000	44.00	44.00
$\{t_1, t_3\}$	1	d_0	Ω_R	0.200	1.000	32.00	32.00
t_3	0	d_1	0	0.800	0.830	44.62	49.38
t_3	0	d_1	1		0.170	4.76	
t_3	1	d_1	0	0.200	0.670	18.76	31.96
t_3	1	d_1	1		0.330	13.20	
t_2	0	d_0	Ω_R	0.666	1.000	53.76	53.76
t_2	1	d_0	Ω_R	0.266	1.000	28.00	28.00
t_2	2	d_0	Ω_R	0.066	1.000	40.00	40.00

4.3. Division

Operationally, division is opposite to combination. Let α be a probability valuation for X and $Y \subset X$. Then, we define $\alpha/\alpha^{\downarrow Y}$, called α divided by $\alpha^{\downarrow Y}$, to be a probability valuation for X as follows:

$$(\alpha/\alpha^{\downarrow Y})(\mathbf{x}) = \alpha(\mathbf{x})/\alpha^{\downarrow Y}(\mathbf{y}). \tag{10}$$

Because α is positive for all focal elements, both α and $\alpha^{\downarrow Y}$ will be positive. Thus, the division in Eq. (10) is always well defined. The last column in Table 4 shows the division of $(\alpha \otimes \beta \otimes \gamma)$ by $(\alpha \otimes \beta \otimes \gamma)^{\downarrow \{T, O, D, R\}}$.

Note that divisions are critical to the recovery of conditional probabilities when doing arc reversals in an influence diagram. However, due to the fusion algorithm [17], the division becomes less important in a valuation network. As we will see shortly, if a valuation network has only one payoff valuation as assumed by influence diagrams, the division is no longer necessary.

5. A fusion algorithm

According to Shenoy [17], solving a problem or making an inference in a valuation-based system is equivalent to marginalizing all variables out of the joint valuation, which is the result of combining all the valuations. For problems with a few variables, it is feasible that we combine all the payoff and probability valuations into one joint payoff valuation and then marginalize it by eliminating the variables. However, for a problem with a large number of variables, this global computation approach is not feasible. As an alternative, the fusion algorithm [17], based on the idea of local computation, arranges the combination and deletion process locally. A coarse valuation network, the graphical representation of a decision model, defines the computational architecture that governs in what sequence valuations are combined and variables are deleted. Specifically, Shenoy [17] specifies that the operational sequence be in accordance with the precedence constraints: It starts with a minimal variable and then proceeds to the next minimal variable until all the variables are deleted. Each time when a variable is deleted, we combine all the valuations that bear on the variable and then delete the variable out the combined valuation. Shenoy [17] calls each step of such local computation *a fusion operation*. In the rest of this section, we define fusion operations for coarse valuations.

Suppose a valuation network has n payoff valuations $\pi_1, \pi_2, \ldots, \pi_n$ and m probability valuations $\alpha_1, \alpha_2, \ldots, \alpha_m$, where π_i domain is X_i and α_j domain is Y_j for $i = 1, 2, \ldots, n$ and $j = 1, 2, \ldots, m$. The fusion operation can be described under five different cases depending on the type of the variable to be deleted.

5.1. Case 1

Suppose that D is a decision variable and $D \notin Y_j$ for $j = 1, 2, \ldots, m$. Then, the set of valuations after deleting D is:

$$(\otimes\{\pi_i \mid D \in X_i\})^{\downarrow X - \{D\}}$$
$$\cup \{\pi_i \mid D \notin X_i\} \cup \{\alpha_1, \alpha_2, \ldots, \alpha_m\}$$

where $X = \cup \{X_i \mid D \in X_i\}$. In words, after fusion, all the payoff valuations that bear on D are combined and marginalized such that D is eliminated. The other payoff valuations and all probability valuations remain unchanged.

5.2. Case 2

Suppose R is a chance variable and $R \notin X_i$ for $i = 1, 2, \ldots, n$. Then, the set of valuations after deleting R is:

$$(\otimes\{\alpha_j \mid R \in Y_j\})^{\downarrow Y - \{R\}} \cup \{\alpha_j \mid R \notin Y_j\}$$

$$\cup \{\pi_1, \pi_2, \ldots, \pi_m\}$$

where $Y = \cup \{Y_j \mid R \in Y_j\}$. In words, after fusion, those probability valuations whose domain contains R are combined and marginalized such that R is eliminated. The other probability valuations and all the payoff valuations remain unchanged.

5.3. Case 3

Suppose R is a chance variable and $R \in X_i$ for $i = 1, 2, \ldots, n$. Then, the set of valuations after deleting R is:

$$(\pi \otimes \alpha)^{\downarrow X \cup Y - \{R\}} \cup \{\alpha_j \mid R \notin Y_j\}$$

where $\pi = \otimes\{\pi_i | i = 1, 2, \ldots, n\}$, $\alpha = \otimes\{\alpha_j | R \in Y_j\}$, $X = \cup \{X_i | i = 1, 2, \ldots, n\}$ and $Y = \cup\{Y_j | R \in Y_j\}$. In words, after fusion, all payoff valuations and those probability valuations that bear on R are combined and marginalized such that R is eliminated. The other probability valuations remain unchanged.

5.4. Case 4

Suppose R is a chance variable and only a part of the payoff valuations bear on R. Then, the set of valuations after deleting R is:

$$\{\alpha^{\downarrow X - \{R\}}\} \cup \{\pi\} \cup \{\pi_i \mid R \notin X_i\} \cup \{\alpha_j \mid R \notin Y_j\}$$

where $\alpha = \otimes\{\alpha_j | R \in Y_j\}$, $\pi = [(\otimes\{\pi_i | R \in X_i\}) \otimes (\alpha / \alpha^{\downarrow X - \{R\}})]^{\downarrow X \cup Y - \{R\}}$, $X = \cup\{X_i | R \in X_i\}$, and $Y = \cup\{Y_j | R \in Y_j\}$. In words, after fusion, the payoff and probability valuations that do not bear on R remain unchanged. A new probability valuation $\alpha^{\downarrow X - \{R\}}$ and a new payoff valuation π are created.

Note that this is the only case divisions take place. If a problem has only one payoff valuation, then this case does not happen and divisions become unnecessary.

5.5. Case 5

Suppose D is a decision variable and there exists j such that $D \in Y_j$. Then, the set of valuations after deleting D is:

$$\{\pi\} \cup \{\pi_i \mid D \notin X_i\} \cup \{\alpha_j \mid D \notin Y_j\}$$

$$\cup \{\alpha_j^{\downarrow Y_j - \{D\}} \mid D \in Y_j\}$$

where $\pi = (\otimes\{\pi_i | D \in X_i\})^{\downarrow X - \{D\}}$, $X = \cup\{X_i | D \in X_i\}$. Note that, since the deletion is in accordance with the precedence constraints, D is to be deleted whenever D is minimum. Therefore, D is minimum in Y_j. The semantic condition of the precedence constraints dictates that α_j is independent of D. Thus, $\alpha_j^{\downarrow Y_j - \{D\}}$ is well defined by Eq. (7).

Note that Shenoy [17] avoids Case 5 by adding a test in Case 4 to see whether a newly generated probability valuation is vacuous. If it is, then the fusion can be simplified as:

$$\{\pi\} \cup \{\pi_i \mid R \notin X_i\} \cup \{\alpha_j \mid R \notin Y_j\}$$

where $\pi = [(\otimes\{\pi_i | R \in X_i\}) \otimes (\otimes\{\alpha_j | R \in Y_j\})]^{\downarrow X \cup Y - \{R\}}$, $X = \cup\{X_i | R \in X_i\}$, and $Y = \cup\{Y_j | R \in Y_j\}$. This approach does not always work as evidenced by θ in Table 7. Even though it works in some cases, we feel that it is not efficient to add an extra test to Case 4.

6. The correctness of the fusion algorithm

This section intends to answer two technical questions regarding the fusion algorithm. First, will the fusion algorithm produce a correct solution? Second, because a precedence order is a partial order, there may be multiple minimal variables at a certain step of problem solving. The sequence for fusion operations may be non-unique. Then, do different sequences give the same solution?

Shenoy [17] positively answers these questions in the context of conditional probabilities. His proof relies on a few critical properties of the marginalization operator on regular payoff and probability valuations. To answer the questions in the context of coarse valuations, we need to establish the equivalent properties as follows:

Property 1. *Suppose π is a coarse payoff valuation for X and R_1 and R_2 are chance variables in X. Then*

$$(\pi^{\downarrow X - \{R_1\}})^{\downarrow X - \{R_1, R_2\}} = (\pi^{\downarrow X - \{R_2\}})^{\downarrow X - \{R_1, R_2\}} \quad (11)$$

Property 2. *Suppose π is a coarse payoff valuation for X, and D_1 and D_2 are decision variables in X. Then*

$$(\pi^{\downarrow X - \{D_1\}})^{\downarrow X - \{D_1, D_2\}} = (\pi^{\downarrow X - \{D_2\}})^{\downarrow X - \{D_1, D_2\}} \quad (12)$$

Property 3. *Suppose α is a coarse probability valuation of H given T, and R_1 and R_2 are two chance variables in H. Let $X = H \cup T$. Then*

$$(\alpha^{\downarrow X - \{R_1\}})^{\downarrow X - \{R_1, R_2\}} = (\alpha^{\downarrow X - \{R_2\}})^{\downarrow X - \{R_1, R_2\}} \quad (13)$$

Property 4. *Suppose α is a coarse probability valuation of H given T, and decision variables D_1 and D_2 belong to T and are minimal in $X = H \cup T$. Then*

$$(\alpha^{\downarrow X - \{D_1\}})^{\downarrow X - \{D_1, D_2\}} = (\alpha^{\downarrow X - \{D_2\}})^{\downarrow X - \{D_1, D_2\}} \quad (14)$$

In words, these properties state that different orders of eliminating multiple chance variables or multiple decision variables do not affect the final result of the marginalization. That is, we can delete two decision or chance variables out of a coarse valuation in opposite sequences but their end results are the same. The verification of the above four properties is straightforward by using Eqs. (6)–(9). With these properties, we can prove the following property:

Property 5. *Deleting all the variables out of a coarse valuation network in any sequence in accordance with the precedence order produces a unique solution.*

To prove this property, first we need to note that the definitions of marginalization can be extended to the case of eliminating multiple variables. For example, suppose π is a payoff valuation for X and X_1 and X_2 are two variables in X. Then, $\pi^{\downarrow X - \{X_1, X_2\}}$ is well defined with respect to the precedence order. The reasoning is as follows. Because $>$ is a partial order, we have either $X_1 > X_2$ or $X_2 > X_1$ or $X_1 \sim X_2$. The marginal $\pi^{\downarrow X - \{X_1, X_2\}}$ is well defined in the first two cases because X_i is deleted after X_j whenever $X_i > X_j$ for $i, j = 1, 2$. In the case of $X_1 \sim X_2$, X_1 and X_2 are either both decision variables or both chance variables according to the perfect recall condition. Then, by Eqs. (11) and (12), the marginal $\pi^{\downarrow X - \{X_1, X_2\}}$ is the same no matter which order we take to delete X_1 and X_2. Therefore, in all the cases, $\pi^{\downarrow X - \{X_1, X_2\}}$ is well defined. By induction, it is easy to infer that the marginal $\pi^{\downarrow X - Y}$ is well defined

with respect to a precedence order for every subset $Y \subset X$.

Second, suppose a valuation network has n payoff valuations $\pi_1, \pi_2, \ldots, \pi_n$ and m probability valuations $\alpha_1, \alpha_2, \ldots, \alpha_m$, where π_i bears on X_i and α_j on Y_j for $i = 1, 2, \ldots, n$ and $j = 1, 2, \ldots, m$. Let $\pi = \otimes\{\pi_i | i = 1, 2, \ldots, n\}$, $\alpha = \otimes\{\alpha_i | i = 1, 2, \ldots, m\}$, $X = \cup\{X_i | i = 1, 2, \ldots, n\}$, $Y = \cup\{Y_j | j = 1, 2, \ldots, m\}$, and $Z = X \cup Y$. Then, according to Shenoy [17], solving a problem in a coarse valuation-based system is a process of marginalizing $\pi \otimes \alpha$ to obtain $(\pi \otimes \alpha)^{\downarrow Z - Z}$, i.e., $(\pi \otimes \alpha)^{\downarrow \varnothing}$. The above analysis implies that the marginal $(\pi \otimes \alpha)^{\downarrow \varnothing}$ is unique with respect to a precedence order, even though it may be obtained from a different sequence of deleting the variables in Z.

Property 5 proves the uniqueness. Regarding the correctness, by slightly modifying the proof to Lemma 4 in Ref. [17], we can prove the following property:

Property 6. *Applying the fusion operations to all the variables in a sequence in accordance with the precedence order will produce the marginal $(\pi \otimes \alpha)^{\downarrow \varnothing}$.*

According to Shenoy [17], problem solving in a valuation-based systems is equivalent to computing the marginal $(\pi \otimes \alpha)^{\downarrow \varnothing}$. Then according to Property 6, it is further equivalent to deleting all variables from a coarse valuation network using the fusion algorithm according to the precedence order. Therefore, Property 6 implies the correctness of the fusion algorithm.
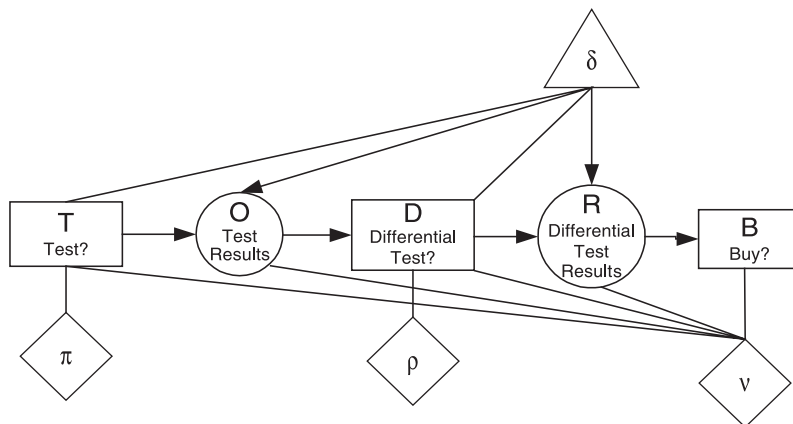

Fig. 2. The valuation network after deleting *C*.

7. An illustration

In this section, we demonstrate the fusion algorithm to the used car buyer's problem. First, as per the precedence constraints (see Fig. 1), the deletion order is CBRDOT.

7.1. Step 1

Since σ bears on the chance variable C while π and ρ do not, Case 4 of the fusion algorithm applies. After the fusion, the new probability valuation is $\delta = (\alpha \otimes \beta \otimes \gamma)^{\downarrow \{T, O, D, R\}}$, which is the conditional probability function of $\{O, R\}$ given $\{T, D\}$ (see Table 4). The new payoff valuation is $v = (\sigma \otimes (\tau/\delta))^{\downarrow \{T, O, D, B, R\}}$ (see Tables 5 and 6). The other two payoff valuations π and ρ are unchanged. The valuation network after deleting C is shown in Fig. 2.

7.2. Step 2

Since B is a decision variable and no probability valuations bear on B, Case 1 of the fusion algorithm applies. After the fusion, the probability valuation δ and the payoff valuations π and ρ remain unchanged. The payoff valuation v is marginalized such that B is eliminated and a new payoff valuation $\varphi = v^{\downarrow \{T, O, D, R\}}$ is generated (see Tables 5 and 6). Fig. 3 shows the valuation network after deleting B.

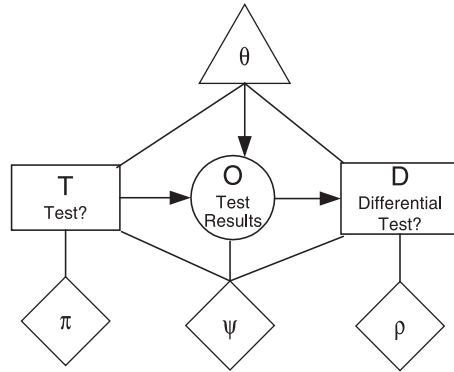While marginalizing v using Eq. (6), we track the solution to B with respect to v. For example,

Fig. 4. The valuation network after deleting R.

$\Psi_B(\{t_0\} \times \Omega_O \times \{d_0\} \times \Omega_R) = \{b_2\}$. It means that the optimal decision is to buy the car without an anti-lemon guarantee ($B = b_2$) if no test is done ($T = t_0$, $D = d_0$) and, of course, we are ignorant about the test results (Ω_O, Ω_R).

7.3. Step 3

According to Fig. 3, the payoff valuation φ bears on the chance variable R while π and ρ do not, Case 4 of the fusion algorithm applies. After the fusion, the new probability valuation is $\theta = \delta^{\downarrow \{T, O, D\}}$. The new payoff valuation is $\psi = (\varphi \otimes (\delta/\theta))^{\downarrow \{T, O, D\}}$ (see Table 7). The other two payoff valuations π and ρ are unchanged. The valuation network after deleting R is shown in Fig. 4.
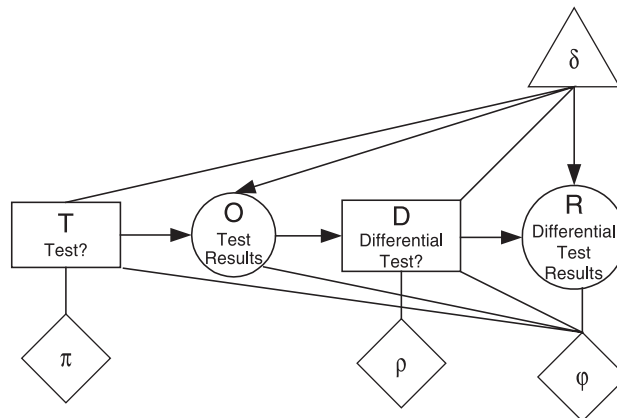

Fig. 3. The valuation network after deleting B.

Table 8
The deletion of D

$\{T, O, D\}$			η	$\psi\otimes\rho$	Ψ_D	κ
t_0	Ω_O	d_0	1.000	28.00	d_0	28.00
t_1	0	d_0	0.800	44.00	d_0	44.00
t_1	1	d_0	0.200	32.00	d_0	32.00
t_2	0	d_0	0.666	53.76	d_0	53.76
t_2	1	d_0	0.266	28.00	d_0	28.00
t_2	2	d_0	0.066	40.00	d_0	40.00
t_3	0	d_0	0.800	44.00		
t_3	0	d_1		45.38	d_1	45.38
t_3	1	d_0	0.200	32.00	d_0	32.00
t_3	1	d_1		27.96		

Table 9
The deletion of O

$\{T, O\}$		$\eta^{\downarrow T}$	$\eta/\eta^{\downarrow T}$	$\kappa\otimes(\eta/\eta^{\downarrow T})$	λ
t_0	Ω_O	1.0	1.000	28.00	28.00
t_1	0	1.0	0.800	35.20	41.60
t_1	1		0.200	6.40	
t_2	0	1.0	0.666	35.80	45.89
t_2	1		0.266	7.45	
t_2	2		0.066	2.64	
t_3	0	1.0	0.800	36.30	42.70
t_3	1		0.200	6.40	

7.4. Step 4

According to Fig. 4, the probability valuation θ bears on the decision variable D, Case 5 of the fusion algorithm applies. After the fusion, θ is reduced into $\eta = \theta^{\downarrow\{T,\ O\}}$. The payoff valuations ψ and ρ are combined and then marginalized into κ such that D is eliminated. The payoff valuation π remains unchanged. The numerical result is shown in Table 8. After the deletion of D, the valuation network is shown in Fig. 5.

7.5. Step 5

According to Fig. 5, the payoff valuation κ bears on the chance variable O while π does not. Therefore, Case 4 of the fusion algorithm applies. After the fusion, the new probability valuation is $\eta^{\downarrow T}$. The
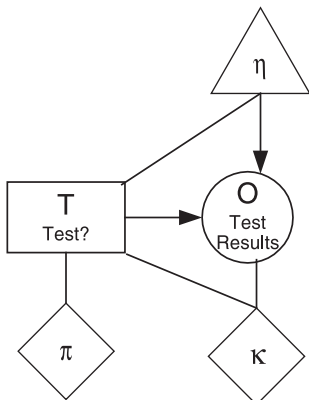
new payoff valuation is $\lambda = [\kappa\otimes(\eta/\eta^{\downarrow T})]^{\downarrow T}$ (see Table 9). The payoff valuation π remains unchanged. The valuation network after deleting O is shown in Fig. 6.

7.6. Step 6

According to Fig. 6, the probability valuation $\eta^{\downarrow T}$ bears on the decision variable T, Case 5 of the fusion algorithm applies. After the fusion, $\eta^{\downarrow T}$ is reduced into $\eta^{\downarrow\phi}$ by using Eq. (7). The payoff valuations λ and π are combined and then marginalized into $(\lambda\otimes\pi)^{\downarrow\phi}$ such that T is eliminated. The numerical result is shown in Table 10. After the deletion of T, no more variables are left and the entire valuation network is solved.

The optimal strategy for the used car buyer's problem has been registered by the three solutions Ψ_T, Ψ_D, and Ψ_B. According to Ψ_T in Table 10, t_2 (the fuel and electric test) should be chosen as an initial test. According to Ψ_D in Table 8, given $T = t_2$, the optimal choice for D is d_0 (no differential test)
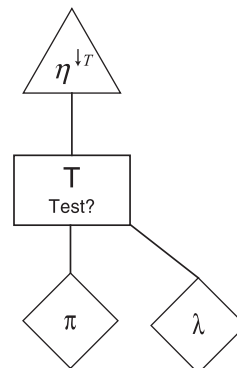

Fig. 5. The valuation network after deleting D.

Fig. 6. The valuation network after deleting O.

Table 10
The deletion of T

T	$\lambda \otimes \pi$	Ψ_T	$(\lambda \otimes \pi)^{\downarrow \phi}$
t_0	28.00		
t_1	32.60		
t_2	32.89	t_2	32.89
t_3	32.70		

regardless the initial test result. Finally, according to Ψ_B in Tables 5 and 6, given $T = t_2$, the optimal strategy is to buy the car without an anti-lemon guarantee (b_2) if the initial test t_2 does not find any errors and buy the car with an anti-lemon guarantee (b_1) if the initial test does find one or two defective systems. The optimal value for the problem is 32.89.

8. Conclusion and discussion

By applying the Dempster–Shafer theory of belief functions, this paper proposed a notion of coarse valuations for representing domain knowledge in the context of decision-making under uncertainty. By using examples, it demonstrated how we could compactly and efficiently represent payoff and probability functions. It established the conditions that coarse valuations must follow in order to permit inference reasoning in a coarse valuation-based system. It defined combination and marginalization operations that stipulate how to manipulate coarse valuations in a system. It proposed a fusion algorithm and a computational architecture for problem solving in a coarse valuation-based system. Finally, it showed the correctness of the fusion algorithm and illustrated the algorithm using a comprehensive example.

This research has made a contribution to the conceptual foundation of decision support system. The notion of coarse valuations bridges the gap between the Dempster–Shafer theory of belief functions and Bayesian theory of probability functions. Like belief functions, the coarse valuations follow the semantics of evidential support. They avoid artificial priors in the Bayesian theory by representing partial indecisiveness using focal elements and complete ignorance using vacuous valuations. In the meantime, they allow a compact model for probabilistic IF–THEN assertions. In addition, the rule of combining coarse

valuations includes as special cases both Dempster's rule of combination and the Bayesian rule of conditioning. Furthermore, the conditions that coarse valuations must follow, as a matter of fact, specify the conditions that enable belief functions to be useful not only for uncertainty reasoning in artificial intelligence but also useful for developing decision support systems.

This research has also made two contributions to the implementation of decision support systems in general and valuation-based systems in particular. First, it extends the framework of valuation-based systems to allow more flexible judgments of probabilities and payoffs in knowledge representation. An influence diagram allows only conditional probabilities; a traditional valuation-based system allows conditional as well as joint probabilities. In contrast, a coarse valuation-based system allows not only conditional and joint probabilities but also belief functions and other belief function-like calculus. Second, and most importantly, this paper made a contribution to the representation of asymmetric decision problems. Asymmetry is prevalent while symmetry is rare. To represent an asymmetric decision problem in a decision support system, both influence diagrams and valuation networks have to introduce dummy events and acts, and degenerate probabilities and utilities to make it symmetric. The symmetrization can explode the problem space by up to 90%, as illustrated by the used car buyer's problem. This is why influence diagrams still cannot replace decision trees to be implemented in decision support systems. This paper showed that coarse valuations are the most compact and natural representation of payoffs and uncertainties. They are even more expressive than a decision tree. As we have shown, a coarse valuation network can take full advantage of both conditional independence and numerical coalescence. On the other hand, because it lacks a mechanism to represent conditional independence, a

Table 11
A comparison of different representations of the used car buyer's problem

Method	# Probabilities	# Utilities
Coalesced decision trees	35	35
Symmetric influence diagrams	226	12
Coarse valuations	20	11

decision tree can exploit conditional independence only if a causal relation coincides with the chronological order. Meanwhile, a decision tree can exploit numerical coalescence only when one branch of the tree is identical to another.

Table 11 compares the representations of the used car buyer's problem using three different methods. It shows that a coalesced decision tree (Fig. 5 on p. 702 in Ref. [7]) needs 35 probabilities and 35 utilities to represent the problem. Of course, the probabilities in the decision tree representation are obtained after preprocessing from 14 probabilities. A symmetric influence diagram representation (or a symmetric valuation network representation) requires specifying 226 probabilities and 12 utilities. On the other hand, a coarse valuation network needs only 20 probabilities and 11 utilities, as shown in Tables 1 and 2.

The larger the problem space, the more calculation it requires. A comparison of computational efficiencies is important. However, because the three methods use different algorithms to solve problems, any comparison is naturally tied to their implementations. Here, we only give an approximate magnitude to show the differences. For example, to delete the variable C in the used car buyer's problem, no matter which method we employ, we have to multiply the probability valuations α, β, and γ. To do so, the decision tree method requires 64 multiplications. A symmetric influence diagram requires 12,288 multiplications. The coarse valuation method requires 160 multiplications. Of course, both the decision tree method and the influence diagram method require additional divisions and multiplications to make Bayesian revisions complete. Also, if no special data structures are designed to avoid them, the coarse valuation method requires additional 160 set intersections.

This paper has dealt with theoretical issues involved in the implementation of a decision support system for decision-making under uncertainty. However, it does not describe any complete implementation. Future research and development along this line of inquiries is definitely needed in order to practically validate the efficiency of our proposed knowledge representation technique and the fusion algorithm. Theoretically, this paper applied coarse valuations to valuation-based systems. We expect that they are also applicable to influence diagrams. However, this has not been done to date.

References

[1] C. Bielza, P.P. Shenoy, A comparison of graphical techniques for asymmetric decision problems, Management Science 45 (11) (1999) 1552–1569.

[2] H.J. Call, W.A. Miller, A comparison of approaches and implementations for automating decision analysis, Reliability Engineering and System Safety 30 (1990) 115–1990.

[3] Z. Covaliu, R.M. Oliver, Formulation and solution of decision problems using decision diagrams, Management Science 41 (12) (1995) 1860–1881.

[4] R. Demirer, P.P. Shenoy, Sequential valuation networks and asymmetric decision problems, Working Paper No. 286, University of Kansas, School of Business, Lawrence, KS, 2000.

[5] D. Dubois, H. Prade, Inference in possibilistic hypergraphs, Proceedings of the Third International Conference on Information Processing and Management of Uncertainty in Knowledge-based Systems (IPMU-90), Paris, France, 1990, pp. 228–230.

[6] R.M. Fung, R.D. Shachter, Contingent influence diagrams, Working Paper, Department of Engineering-Economic Systems, Stanford University, Stanford, CA, 1990.

[7] R.A. Howard, The used car buyer, in: R.A. Howard, J.E. Matheson (Eds.), The Principles and Applications of Decision Analysis, vol. 2 (1962), Strategic Decisions Group, Menlo Park, CA, 1984, pp. 689–718.

[8] R.L. Keeney, H. Raiffa, Decisions with Multiple Objectives, Wiley, New York, NY, 1976.

[9] A. Kong, Multivariate belief functions and graphical models, PhD Dissertation, Department of Statistics, Harvard University, Cambridge, MA, 1986.

[10] S.L. Lauritzen, D.J. Spiegelhalter, Local computations with probabilities on graphical structures their application to expert systems (with discussion), Journal of Royal Statistical Society 50 (1988) 157–224 (series B).

[11] L. Liu, P.P. Shenoy, A theory of coarse utility, Journal of Risk and Uncertainty 11 (1995) 17–49.

[12] L. Liu, P.P. Shenoy, A decomposition method for asymmetric decision problems, Proceedings of the 1995 Decision Sciences Institute Annual Meeting, Boston, MA, vol. 2, 1995, pp. 589–591.

[13] L. Liu, P.P. Shenoy, Conditional belief functions, Proceedings of 1998 Decision Sciences Institute Annual Meeting, Las Vegas, NV, vol. 2, 1998, pp. 589–591.

[14] T.D. Nielsen, F.V. Jensen, Representing and solving asymmetric decision problems, in: C. Boutilier, M. Goldszmidt (Eds.), Uncertainty in Artificial Intelligence: Proceedings of the Sixteenth Conference, Morgan Kaufmann, San Francisco, CA, 2000, pp. 416–425.

[15] R. Qi, L. Zhang, D. Poole, Solving asymmetric decision problems with influence diagrams, in: R.L. Mantaras, D. Poole (Eds.), Uncertainty in Artificial Intelligence: Proceedings of the Tenth Conference, Morgan Kaufmann, San Francisco, CA, 1994, pp. 491–497.

[16] G. Shafer, A Mathematical Theory of Evidence, Princeton Univ. Press, Princeton, NJ, 1976.

[17] P.P. Shenoy, Valuation-based systems for Bayesian decision analysis, Operations Research 40 (3) (1992) 463–484.

[18] P.P. Shenoy, Using possibility theory in expert systems, Fuzzy Sets and Systems 52 (2) (1992) 129–142.

[19] P.P. Shenoy, A new method for representing and solving Bayesian decision problems, in: D.J. Hand (Ed.), Artificial Intelligence and Statistics, vol. III, Chapman and Hall, London, 1993, pp. 119–138.

[20] P.P. Shenoy, A comparison of graphical techniques for decision analysis, European Journal of Operational Research 78 (1) (1994) 1–21.

[21] P.P. Shenoy, Using Dempster–Shafer's belief-function theory in expert systems, in: R.R. Yager, M. Fedrizzi, J. Kacprzyk (Eds.), Advances in the Dempster–Shafer Theory of Evidence, Wiley, New York, NY, 1994, pp. 395–414.

[22] P.P. Shenoy, Valuation network representation and solution of asymmetric decision problems, European Journal of Operational Research 121 (3) (2000) 579–608.

[23] J.E. Smith, S. Holtzman, J.E. Matheson, Structuring conditional relationships in influence diagrams, Operations Research 41 (2) (1993) 280–297.

Liping Liu is an Associate Professor of Information Systems at University of Akron. He received his Master in Systems Engineering from Huazhong University of Science and Technology, China, in 1991, and his PhD in Business from University of Kansas, USA, in 1995. His research interests have been in the areas of Electronic Business, Object-Oriented Systems Design, and Uncertainty Reasoning in Artificial Intelligence. His articles have appeared in *International Journal of Approximate Reasoning*, *Journal of Risk and Uncertainty*, *European Journal of Operational Research*, and others. He has served as a Guest Editor for *International Journal of Intelligent Systems*. He has strong practical and teaching interests in Client/Server and Web-Based Systems Design and Development using advanced DBMS, CASE, and RAD tools. He won two teaching awards. His recent consulting experience includes designing a patient record management system, a payroll system, a course management system, and an e-travel agent, and administering Oracle databases for companies like Trailmobile Parts and Service, Southern Illinois Healthcare, and Beck Bus Transportation.

Prakash P. Shenoy is the Ronald G. Harper Distinguished Professor of Artificial Intelligence in Business, University of Kansas at Lawrence. He received an MS and a PhD in Operations Research from Cornell University in 1975 and 1977, respectively. His research interests are in the areas of uncertain reasoning and decision analysis. He is the inventor of valuation-based systems, an abstract framework for knowledge representation and inference that includes Bayesian probabilities, Dempster–Shafer belief functions, and other domains. His articles have appeared in journals such as *Operations Research*, *Management Science*, *Artificial Intelligence*, and *International Journal of Approximate Reasoning*. He serves as the North-American Editor of *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, and as an Associate Editor of *Operations Research*, *Management Science*, and *International Journal of Approximate Reasoning*, and as an ad-hoc referee for over 30 journals and conferences in Artificial Intelligence and Management Science/Operations Research.