Variable Elimination for Interval-Valued Influence Diagrams

Rafael Cabañas[1(✉)], Alessandro Antonucci[2], Andrés Cano[1], and Manuel Gómez-Olmedo[1]

[1] Department of Computer Science and Artificial Intelligence CITIC, University of Granada, Granada, Spain
{rcabanas,acu,mgomez}@decsai.ugr.es

[2] Istituto Dalle Molle di Studi sull'Intelligenza Artificiale (IDSIA), Galleria 2, Manno-Lugano, Switzerland
alessandro@idsia.ch

Abstract. Influence diagrams are probabilistic graphical models used to represent and solve decision problems under uncertainty. Sharp numerical values are required to quantify probabilities and utilities. Yet, real models are based on data streams provided by partially reliable sensors or experts. We propose an interval-valued quantification of these parameters to gain realism in the modelling and to analyse the sensitivity of the inferences with respect to perturbations of the sharp values. An extension of the classical influence diagrams formalism to support interval-valued potentials is provided. Moreover, a variable elimination algorithm especially designed for these models is developed and evaluated in terms of complexity and empirical performances.

Keywords: Influence diagrams · Bayesian networks · Credal networks · Sequential decision making · Imprecise probability

1 Introduction

Influence diagrams are probabilistic graphical models able to cope with decision problems with uncertainty. The parameters of an influence diagram are conditional probabilities for single variables given some other variables, or utilities depending on given sets of variables. The quantification of these parameters is based on a statistical processing of data or on the elicitation of expert knowledge.

Exactly as Bayesian networks, influence diagrams require sharp estimates of their parameters. Yet, when coping with expert knowledge, sharp values can be unfit to express judgements (e.g., which is the number modelling the probability for an option *more probable* than its negation?). This issue appears also when coping with scarce or missing data (e.g., probabilities conditional on rare events).

For reasons of this kind, in the last two decades, various extensions of Bayesian networks to support more general probabilistic statements have been proposed. These models have been developed in the field of possibility theory [2],

S. Destercke and T. Denoeux (Eds.): ECSQARU 2015, LNAI 9161, pp. 541–551, 2015.
DOI: 10.1007/978-3-319-20807-7_49

evidence theory [17], and imprecise probability [4]. The latter models, called *credal networks*, offer a direct sensitivity analysis interpretation: a credal network is a collection of Bayesian networks, all over the same variables and with the same graph, whose parameters are consistent with constraints (e.g., intervals) modelling a limited ability in the assessment of sharp estimates. Similarly, various extensions of decision trees have also been proposed [9,11]. The situation is different for influence diagrams. The early attempts of Fertig and Breese [7] first, and Zaffalon [6] after, to extend these models to non-sharp quantification are to the best of our knowledge the only works in this direction.[1] This is unfortunate as the above considerations about the difficulty of assessing sharp estimates for probabilities are even more compelling for utilities, which are supposed to model intrinsically qualitative objects such as preferences.

In this paper we extend to the interval-valued case the formalism of influence diagrams by keeping the same sensitivity-analysis interpretation of credal networks: a generalized influence diagram is a collection of classical influence diagrams consistent with the interval constraints. When coping with interval-valued utilities we might have overlaps between the different expectations. In these cases we adopt a conservative approach which rejects decisions leading to certainly dominated options and keeps all the other ones. A first example of *variable elimination* to compute inferences in these generalized models (arc reversal was considered in [7]) is also proposed together with some preliminary tests.

2 Basics

Let us first define the basic notation. We use upper-case letters for variables and lower-case for states. Given a variable X, x is an element of the domain of X, which we denote as Ω_X. Given a set of n variables $\boldsymbol{X} := (X_1, \ldots, X_n)$, and a multi-valued index $J \subseteq [1, n]$, X_J is the joint variable including any X_i such that $i \in J$. Thus, $\Omega_{X_J} = \times_{i \in J} \Omega_{X_i}$, where $\times$ is the Cartesian product. Given a second index I, notation $x_I \sim x_J$ is used to express *consistency*, i.e., to denote the fact that the two states have the same values on $X_{I \cap J}$. *Chance* variables are those whose actual value might be unknown, *decision* variables are those whose actual value can be set by the decision maker. A *potential* over X_J is a map $\psi : \Omega_{X_J} \to \mathbb{R}$. *Probability potentials* (PP, also called conditional probability tables) are special potentials. Given two disjoint set of variables X_I and X_J, $\phi(X_I, X_J)$ is a PP over X_I given X_J if and only if it is a nonnegative potential such that $\sum_{x_I \in \Omega_{X_I}} \phi(x_I, x_J) = 1$ for each $x_J \in \Omega_{X_J}$.

Influence Diagrams. Influence diagrams (IDs) [8] are a class of graphical models designed to formalize sequential decision problems with uncertainty. The uncertainty is represented by PPs, while the user preferences are represented by generic potentials called here *utility potentials* (UPs).

[1] Sensitivity analysis does not require the specification of more general class of models, being only focused on the results of the inferences. Thus, it should be regarded as a different topic, which, as a matter of fact, received more attention (e.g., [14]).

An ID over a set of variables $(\boldsymbol{X}, \boldsymbol{D})$ is made of a qualitative and a quantitative part. The qualitative part is a directed acyclic graph (DAG) $\mathcal{G}$ with three types of nodes. *Chance* nodes are depicted as circles and are in one-to-one correspondence with the chance variables, i.e., the variables in $\boldsymbol{X}$. Decision nodes are depicted as squares and associated to decision variables $\boldsymbol{D}$. Utility nodes are depicted as a diamonds and should be barren. For chance and decision variables we use the terms node and variable interchangeably. Utility nodes are not associated to variables. Still, these nodes are jointly denoted as $\boldsymbol{U}$. The immediate predecessors of a node Y according to the $\mathcal{G}$ are called *parents* and denoted as Π_Y. From the quantitative point of view, for each chance node, a PP over the corresponding variable and its parents is defined, while, for each utility node, a UP potential over the parents should be assessed. The formal definition of ID is the following.

Definition 1. *An influence diagram is a tuple $\langle \mathcal{G}, \boldsymbol{X}, \boldsymbol{D}, \boldsymbol{U}, \Phi, \Psi \rangle$, where $\mathcal{G}$ is a DAG over $\boldsymbol{X} \cup \boldsymbol{D} \cup \boldsymbol{U}$, while $\Phi = \{\phi(X, \Pi_X)\}_{X \in \boldsymbol{X}}$ and $\Psi = \{\psi(\Pi_U)\}_{U \in \boldsymbol{U}}$ are collections of, respectively, PPs and UPs.*

To model sequential decision problems with IDs, some additional information should be provided. A complete order of decision variables (e.g., $D_1 \prec \ldots \prec D_n$), and a partial order for $\boldsymbol{X} \cup \boldsymbol{D}$ consistent with the partial one for $\boldsymbol{D}$ should be formulated. Accordingly the partial order has form $\mathcal{I}_0 \prec D_1 \prec \mathcal{I}_1 \prec \cdots \prec D_n \prec \mathcal{I}_n$, with $\cup_{j=0}^{n} \mathcal{I}_j = \boldsymbol{X}$ and $\mathcal{I}_i \cap \mathcal{I}_j = \emptyset$ for each i, j. This reflects a temporal interpretation: the chance variables in $\mathcal{I}_i$ are observed before decision D_{i+1} is taken, and the ordering over $\boldsymbol{D}$ reflects the order in which the different decisions are taken. An ID of this kind is called *regular. Non-forgetting assumption* is usually required as well: previous decisions and observations are known at each decision. Here we only consider regular IDs with the non-forgetting assumption. A classical example is here below.

Example 1 (The oil wildcatter [15,16]). An oil wildcatter must decide whether to drill or not. He is uncertain whether the amount of oil (O) in the place is empty (e), wet (w) or soaking (s). The wildcatter can make seismic tests (S) that will give a closed reflection pattern (c) indicating much oil, an open pattern (o) indicating for some oil, or a diffuse pattern (d) denoting almost no hope for oil. These two are chance variables, while the decision variables are T [to test (t) or not (nt)] and D [to drill (d) or not (nd)]. The utility nodes P and C describe the profit possibly obtained from the presence of oil and the cost of the tests. The DAG of an ID modelling this problem is in Fig. 1. Decision T precedes decision D, while the partial order is complete being $T \prec \{S\} \prec D \prec \{O\}$.

ID Evaluation. A *policy* for a decision variable D is a map $\delta_D : \Omega_{\Pi_D} \to \Omega_D$, i.e., a rule to assign D on the basis of the values of the parents. A *strategy* Δ is a collection of policies, one for each $D \in \boldsymbol{D}$. From a joint observation $\boldsymbol{x}$ of the chance variables and a strategy Δ, we deduce a joint specification of the decision variables, and hence of the values of the utilities, which are assumed additive. Accordingly, we can compute the expected utility of a strategy as follows:

$\phi(O)$

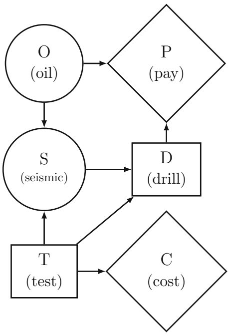
O: e	w	s
.5	.3	.2

$\phi(S|O, T = t)$

S	O: e	w	s
c	.1	.3	.5
o	.3	.4	.4
d	.6	.3	.1

$\phi(S|O, T = nt)$

S	O: e	w	s
c	.33	.33	.33
o	.33	.33	.33
d	.33	.33	.33



$\psi(O, D)$

T	O: e	w	s
d	-70	50	200
nd	0	0	0

$\psi(T)$

T: t	nt
-10	0

Fig. 1. The oil wildcatter's decision problem as an ID

$$\text{EU}(\Delta) := \sum_{\boldsymbol{x}} \prod_{X \in \boldsymbol{X}} \phi(x, \pi_X) \sum_{U \in \boldsymbol{U}} \psi(\pi_U), \tag{1}$$

with x, π_X, $\pi_U \sim \boldsymbol{x}$, and the values of π_X and π_U also consistent with the values of $\boldsymbol{D}$ obtained from $\boldsymbol{x}$ and Δ. A typical task with IDs is to detect the strategy Δ^* maximizing the expected utility. This should respect the partial order, i.e.,

$$EU(\Delta^*) = \sum_{\mathcal{I}_0} \max_{D_1} \cdots \max_{D_n} \sum_{\mathcal{I}_n} \prod_{X \in \boldsymbol{X}} \phi(X|\Pi_X) \sum_{U \in \boldsymbol{U}} \psi(\Pi_U) \tag{2}$$

The components of this optimal strategy, are called *optimal policies*, and they can be regarded as the intermediate steps of the above maximization. The optimal policy $\delta^*_{D_i}$ associated to D_i is therefore:

$$\delta^*_{D_i}(\Pi_{D_i}) = \arg\max_{D_i} \sum_{\mathcal{I}_i} \max_{D_{i+1}} \cdots \max_{D_n} \sum_{\mathcal{I}_n} \prod_{X \in \boldsymbol{X}} \phi(X|\Pi_X) \sum_{U \in \boldsymbol{U}} \psi(\Pi_U). \tag{3}$$

Example 2. In the oil wildcutter's problem (Ex. 1), when seismic tests have been done and returned a closed reflection pattern the right thing to do is to drill, i.e. $\delta^*_D(S = c, T = t) = d$ *as* $EU_D(S = c, T = t) = 77.5$.

Variable Elimination. *Variable elimination* (VE) is a typical approach to inference in graphical model. VE algorithms for IDs [10,18] are commonly used to solve Eq. (2). Unlike VE for Bayesian networks, in regular IDs the elimination order is not arbitrary: it should be the inverse of an order consistent with the partial order associated to the ID [12]. Furthermore, while chance variables are removed by sum, decision variables are instead eliminated by maximization. To describe VE, we first show how to remove a single variable Y in Algorithm 1. The whole algorithm consists in iterating the procedure over $\boldsymbol{X} \cup \boldsymbol{D}$. When the last variable is eliminated, the algorithm returns a potential with no arguments (i.e., a constant) with the value in Eq. (2).

The operator dom returns the variables in the argument of a potential. Sums in line 4 and maxima in line 6 of Algorithm 1 are two different forms (the first for

Algorithm 1. Removing a single (chance or decision) variable Y

1: $(\Phi_Y, \Psi_Y) \leftarrow (\{\phi \in \Phi | Y \in \mathrm{dom}(\phi)\}, \{\psi \in \Psi | Y \in \mathrm{dom}(\psi)\})$ ▷ Select
2: $(\phi_Y, \psi_Y) \leftarrow (\otimes_{\phi \in \Phi_Y} \phi, \otimes_{\psi \in \Psi_Y} \psi)$ ▷ Combine
3: **if** $Y \in \boldsymbol{X}$ **then**
4: $\quad (\phi'_Y, \psi'_Y) \leftarrow (\sum_Y \phi_Y, \frac{\sum_Y \phi_Y \otimes \psi_Y}{\sum_Y \phi_Y})$ ▷ Remove by sum (chance vars)
5: **else**
6: $\quad (\phi'_Y, \psi'_Y) \leftarrow (\phi_{Y=y}, \max_Y \psi_Y)$ ▷ Remove by max (decision vars)
7: $\quad \delta^*_Y \leftarrow \arg\max_Y \psi_Y$ ▷ Optimal policy (as a byproduct)
8: **end if**
9: $(\Phi, \Psi) \leftarrow (\Phi \backslash \Phi_Y \cup \{\phi'_Y\}, \Psi \backslash \Psi_Y \cup \{\psi'_Y\})$ ▷ Update
10: **return** (Φ, Ψ)

decision, the second for chance variables) of *marginalization*, removing the variable from the argument of the potential. The division (line 4) is defined element-wise. For the PP in line 6, Y can be eliminated by instantiating an arbitrary value [12]. When eliminating a decision variable, the maximization of the UP also gives the corresponding optimal policy (line 7).

The operator $\otimes$ in Algorithm 1 combines pairs of potentials as explained here below. It is easy to check that these definitions are well-posed and that the operator is associative and commutative.

(i) given two UPs, say $\psi(X_I)$ and $\psi'(X_J)$, their combination $\psi \otimes \psi'$ is a UP over $X_{I \cup J}$ obtained by element-wise sums, i.e., $(\psi \otimes \psi')(x_{I \cup J}) := \psi(x_I) + \psi'(x_J)$ for each $x_{I \cup J} \in \Omega_{X_{I \cup J}}$, with $x_I, x_J \sim x_{I \cup J}$;
(ii) given a PP $\phi(X_I, X_J)$ and a UP $\psi(X_K)$, their combination $\phi \otimes \psi$ is a UP over $X_L := X_{I \cup J \cup K}$ defined by element-wise products, i.e., $(\phi \otimes \psi)(x_{I \cup J \cup K}) := \phi(x_I, x_J) \cdot \psi(x_K)$, for each $x_{I \cup J \cup K} \in \Omega_{X_{I \cup J \cup K}}$, with $x_I, x_J, x_K \sim x_{I \cup J \cup K}$;
(iii) finally, given two PPs, say $\phi(X_I, X_J)$ and $\phi'(X_K, X_L)$, their combination $\phi \otimes \phi'$ is a PP over $X_{I \cup K}$ given $X_{(J \cup L) \backslash (I \cup K)}$ defined by element-wise products, i.e., $(\phi \otimes \phi')(x_{I \cup K}, x_{(J \cup L) \backslash (I \cup K)}) := \phi(x_I, x_J) \cdot \phi(x_K, x_L)$ for each $x_{I \cup K} \in \Omega_{X_{I \cup K}}$ and $x_{(J \cup L) \backslash (I \cup K)} \in \Omega_{X_{(J \cup L) \backslash (I \cup K)}}$, with $x_I, x_J, x_K, x_L \sim x_{I \cup K}, x_{(J \cup L) \backslash (I \cup K)}$.

3 Interval-Valued Potentials

The main goal of this paper is to extend IDs to support interval-valued specifications. To do that, we first formalize the basic notion of interval-valued potential.

Definition 2. *An* interval-valued utility potential *(IUP) over X_I is a pair of UPs over X_I. We use the compact notation $\underline{\overline{\psi}}(X_I)$ for a IUP over X_I, $\underline{\psi}$ and $\overline{\psi}$ are the two UPs involved in the specification and are called, respectively, the lower and upper* bounds *of the IUP. The* extension *$\underline{\overline{\psi}}^*(X_I)$ of this IUP is the set of UPs consistent with the bounds, i.e.,*

$$\underline{\overline{\psi}}^*(X_I) := \left\{\psi : \Omega_{X_I} \to \mathbb{R} \,\middle|\, \underline{\psi}(x_I) \leq \psi(x_I) \leq \overline{\psi}(x_I), \forall x_I \in \Omega_{X_I}\right\}. \tag{4}$$

The extension of a IUP $\overline{\underline{\psi}}$ is non-empty if and only if $\underline{\psi}(x_I) \leq \overline{\psi}(x_I)\ \forall x_I \in \Omega_{X_I}$. We similarly define an *interval-valued probability potential* (IPP) over X_I given X_J as a pair of (not necessarily normalized) PPs over X_I given X_J. We denote such a IPP as $\overline{\underline{\phi}}(X_I, X_J)$, where $\underline{\phi}(X_I, X_J)$ and $\overline{\phi}(X_I, X_J)$ are the two (unnormalized) bounds. The extension is also defined in terms of consistency.

$$\overline{\underline{\phi}}^*(X) := \left\{ \phi : \Omega_{X_I} \times \Omega_{X_J} \to \mathbb{R}_0^+ \;\middle|\; \begin{array}{l} \sum_{x_I} \phi(x_I, x_J) = 1, \\ \underline{\phi}(x_I, x_J) \leq \phi(x_I, x_J) \leq \overline{\phi}(x_I, x_J), \\ \forall (x_I, x_J) \in \Omega_{X_I} \times \Omega_{X_J} \end{array} \right\}. \quad (5)$$

Condition $\underline{\phi}(x_I, x_J) \leq \overline{\phi}(x_I, x_J)$ for each x_I, x_J, together with $\sum_{x_I} \underline{\phi}(x_I, x_J) \leq 1 \leq \sum_{x_I} \overline{\phi}(x_I, x_J)$, for each $x_J \in \Omega_{X_J}$ is necessary and sufficient for the extension of the IPP to be non-empty. The additional condition $\overline{\phi}(x_i') + \sum_{x_i \neq x_i'} \underline{\phi}(x_i) \leq 1$ and the analogous expression for the lower instead of the upper bounds is called *reachability* [5]. The meaning is that for each $p \in [\underline{\phi}(x_I, x_J), \overline{\phi}(x_I, x_J)]$, there is at least a PP $\phi \in \overline{\underline{\phi}}^*$ s.t. $\phi(x_I, x_J) = p$. Note also that an IPP with non-empty extension can be always reduced to a reachable one by shrinking its bounds. Both the extensions of a IUP and a IPP are therefore convex sets of, respectively, UPs and PPs. Vice versa, while any convex set of UPs can be regarded as the extension of a IUP, the same does not hold for IPPs.

Example 3. Figure 2 reports an interval-valued specification of the five potential associated to the ID in Fig. 1. It is a trivial exercise to check that: the IPPs have non-empty extensions and are reachable, and the UPs and PPs of the original model are included in the extensions of their interval-valued counterparts.

$\phi(O)$

O: e	O: w	O: s
[.4875, .5125]	[.2925, .3175]	[.195, .22]

$\psi(O, D)$

		O: e	O: w	O: s
T	d	$[-75, -65]$	[45, 55]	[195, 205]
	nd	$[-5, 5]$	$[-5, 5]$	$[-5, 5]$

$\psi(T)$

T: t	T: nt
$[-15, -5]$	$[-5, 5]$

$\phi(S \mid O, T = t)$

		O: e	O: w	O: s
S	c	[.0975, .1225]	[.2925, .3175]	[.4875, .5125]
	o	[.2925, .3175]	[.39, .415]	[.39, .415]
	d	[.585, .61]	[.2925, .3175]	[.0975, .1225]

$\phi(S \mid O, T = nt)$

		O: e	O: w	O: s
S	c	[.325, .35]	[.325, .35]	[.325, .35]
	o	[.325, .35]	[.325, .35]	[.325, .35]
	d	[.325, .35]	[.325, .35]	[.325, .35]

Fig. 2. A set of IUPs and IPPs for the oil wildcatter's decision problem

Combining Interval-Valued Potentials. The combination operation over potentials in Algorithm 1 can be extended to interval-valued potentials as follows:

(i) given two IUPs, say $\overline{\underline{\psi}}(X_I)$ and $\overline{\underline{\psi}}'(X_J)$, their combination $\overline{\underline{\psi}} \otimes \overline{\underline{\psi}}'$ is a IUP over $X_{I \cup J}$ s.t. $(\underline{\psi} \otimes \underline{\psi}')(x_{I \cup J}) := \underline{\psi}(x_I) + \underline{\psi}'(x_J)$ for each $x_{I \cup J} \in \Omega_{X_{I \cup J}}$ with $x_I, x_J \sim x_{I \cup J}$; and similarly for the upper bounds;

(ii) given a IPP $\overline{\underline{\phi}}(X_I, X_J)$ and a IUP $\overline{\underline{\psi}}(X_K)$, their combination $\overline{\underline{\phi}} \otimes \overline{\underline{\psi}}$ is a IUP over $X_{I \cup J \cup K}$ s.t. $(\underline{\psi} \otimes \underline{\psi})(x_{I \cup J \cup K}) := \underline{\phi}(x_I, x_J) \cdot \underline{\psi}(x_K)$ for each $x_{I \cup J \cup K} \in \Omega_{X_{I \cup J \cup K}}$, with $x_I, x_J, x_K \sim x_{I \cup J \cup K}$; if $\underline{\psi}(x_K) < 0$ the lower bound of the combination is obtained by multiplying the lower bound of the IUP for the upper bound of the IPP (and vice versa for the upper bound);

(iii) given two IPPs, say $\underline{\overline{\psi}}(X_I, X_J)$ and $\underline{\overline{\psi}}'(X_K, X_L)$, their combination $\underline{\overline{\psi}} \otimes \underline{\overline{\psi}}'$ is a IPP over $X_{I\cup K}$ given $X_{(J\cup L)\setminus(I\cup K)}$ s.t. $(\underline{\psi} \otimes \underline{\psi}')(x_{I\cup K}, x_{(J\cup L)\setminus(I\cup K)}) := \underline{\psi}(x_I, x_J) \cdot \underline{\psi}(x_K, x_L)$, for each $x_{I\cup K} \in \Omega_{X_{I\cup K}}$ and $x_{(J\cup L)\setminus(I\cup K)} \in \Omega_{X_{(J\cup L)\setminus(I\cup K)}}$, with $x_I, x_J, x_K, x_L \sim x_{I\cup K}, x_{(J\cup L)\setminus(I\cup K)}$. If $\underline{\overline{\psi}} \otimes \underline{\overline{\psi}}'$ is not reachable, the transformation to make it reachable is performed.

The following result, whose proof is left to the reader provides a sensitivity-analysis justification for the proposed generalization of the combination operator.

Proposition 1. *Given potentials (no matter whether IUPs or IPPs) $\underline{\overline{\psi}}$ and $\underline{\overline{\phi}}$ the extension of their combination is s.t.* $(\underline{\overline{\psi}} \otimes \underline{\overline{\phi}})^* = \left\{ \psi \otimes \phi \,\middle|\, \psi \in \underline{\overline{\psi}}^*, \phi \in \underline{\overline{\phi}}^* \right\}$.

4 Interval Influence Diagrams

IDs can be extended to the interval framework by simply replacing the PPs and UPs in Definition 1 with an equal number of IPPs and IUPs defined on the same domains. We call a model of this kind an *interval-valued influence diagram* (IID). As an example, the ID in Example 1 with the interval-valued quantification in Fig. 2 is an IID over the graph in Fig. 1. Before applying to IIDs the VE scheme in Algoritm 1, the different operations over the potentials involved in the algorithm should be extended to intervals. In the previous section we discussed how to do that with the combination. Here we generalize the operations in lines 4 and 6.

Eliminating Chance Variables. We start from line 4. To sum out a variable from a IPP, we sum out the variable from the two bounds and, if needed, we make the result reachable. Concerning the sum and division in the second term, given an IPP $\underline{\overline{\phi}}(X_I \cup Y, X_J)$ and a IUP $\underline{\overline{\psi}}(X_K \cup Y)$, we set the result as a IUP $\underline{\overline{\hat{\psi}}}$ over $X_{I\cup J\cup K}$ s.t.

$$\underline{\overline{\hat{\psi}}}^*(X_{I\cup J\cup K}) := \left\{ \hat{\psi}(X_{I\cup J\cup K}) \,\middle|\, \begin{array}{l} \hat{\psi}(x_{I\cup J\cup K}) = \frac{\sum_y \phi(x_I, y, x_J) \cdot \psi(y, x_K)}{\sum_y \phi(x_I, y, x_J)} \\ \forall x_{I\cup J\cup K}, \forall \phi \in \underline{\overline{\phi}}^*, \forall \psi \in \underline{\overline{\psi}}^* \end{array} \right\}. \tag{6}$$

This allows to obtain a result as in Proposition 1. Note that the argument of the sum in the numerator can be regarded as an element of $(\underline{\overline{\phi}} \otimes \underline{\overline{\psi}})^*$. By computing the bound of the extension in Eq. (6), we eventually obtain the required IUP. Because of Eqs. (4) and (5), the extensions $\underline{\overline{\psi}}^*$ and $\underline{\overline{\phi}}^*$ are convex regions defined by linear constraints. Furthermore, the objective function to optimize is linear-fractional. So the task reduces to a linear program by the Charnes-Cooper transformation. Equivalently, the task can be regarded as a combinatorial optimization by considering only the extreme points of the feasible region.[2] Faster, but approximate, approaches can be also considered. To obtain the lower bound, we can take the lower bound of the numerator and the upper bound of the denominator in Eq. (6) (and vice versa for the upper bound). This induces an outer

[2] The solution of a linear program is an extreme point of the feasible region.

approximation. A heuristic alternative consists in consider as PP specifications a lower bound for a value of Y and the upper bounds for the other values (or vice versa).

Eliminating Decision Variables. Here we discuss the operations in lines 6 and 7. The arg max operation is intrinsically related to the fact that a UP has sharp values. The problem of deciding the "maximal" options in the interval case might be arguable. The most conservative approach is the following.

Definition 3. *Let* $\overline{\psi}$ *be a IUP over* $Y \cup \boldsymbol{X}_I$. *An element* $y \in \Omega_Y$ *is* interval-maximal *given* $x_I \in \Omega_{\boldsymbol{X}_I}$ *if there is no* $y' \in \Omega_Y$ *s.t.* $\underline{\psi}(y', x_I) > \overline{\psi}(y, x_I)$.

Let D be a decision to be eliminated from $\overline{\psi}(D, \boldsymbol{X}_I)$. To detect the optimal policy $\delta^*_D(\boldsymbol{X}_I)$ we compute the interval-maximal states of D given each $x_I \in \Omega_{\boldsymbol{X}_I}$. This corresponds to a so-called *credal* policy allowing for indecision between two or more possible options. Finally the maximization of the IUP is done as usual by acting separately on the two bounds.

Example 4. In the oil wildcutter's problem (Example 1), if the combinatorial approach is used, $\delta^*_D(S = c, T = t) = \{d\}$ *as* $EU_D(S = c, T = t) \in [46.24, 115.78]$.

Complexity Analysis. VE in IDs takes time exponential in the maximum clique size (i.e., the arity of the biggest potential generated during the evaluation). The same holds for IIDs, apart from a possible bottleneck during the elimination of the chance variables as in Eq. (6). This is the case when the combinatorial optimization is adopted: the number of extreme points to be evaluated is exponential in the number of states. This is not the case with the outer approximation and the heuristic, as well as the linear programming which is polynomial in the number of constraints, and hence in the number of states.

5 Empirical Validation

For a preliminary validation of the VE algorithm we consider six IDs [1,3,13,15]. Table 1 details the number of nodes of each type for these models. These IDs are transformed in IIDs by a perturbation of the original parameters. The approaches in Sect. 4) are compared. Figure 3 shows the computation times. As expected, the outer approximation and the heuristic method roughly take the double of the time required by the precise evaluation. The exact approach with the extreme

Table 1. Number of chance, decision and utility nodes for the benchmark IIDs

	NHL	Jaundice	Appendicitis	Comp. Assym	Oil	Thinkbox
$\lvert\boldsymbol{X}\rvert$	17	21	4	3	2	5
$\lvert\boldsymbol{D}\rvert$	3	2	1	5	2	2
$\lvert\boldsymbol{U}\rvert$	1	1	1	2	1	4

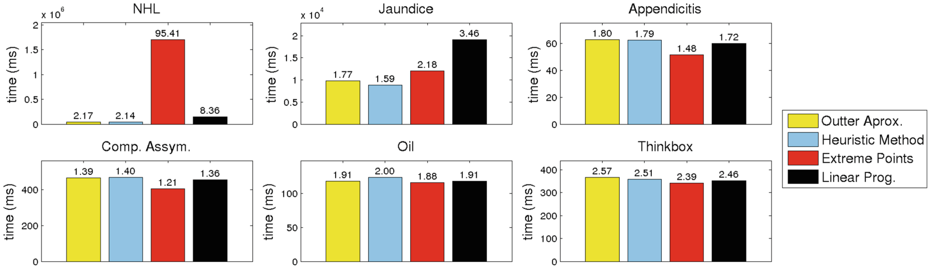


Fig. 3. Evaluation times for the IIDs in Table 1 and relative duration w.r.t. the evaluation of the corresponding IDs

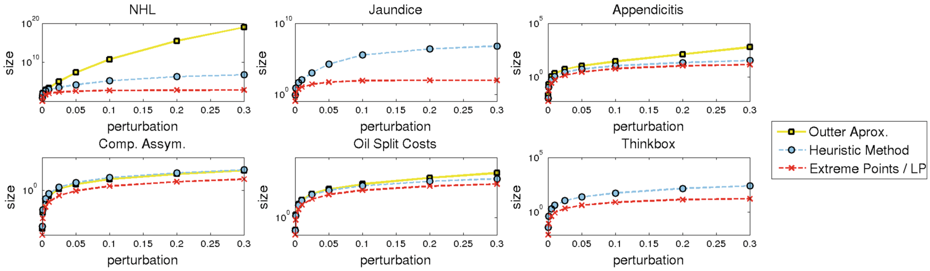


Fig. 4. Size of $EU(\Delta^*)$ for different sizes of the IPPs

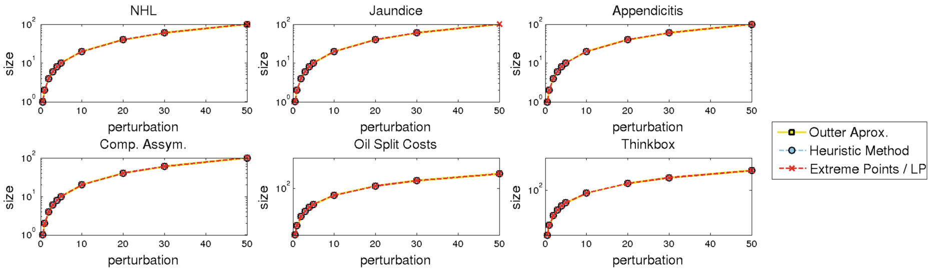


Fig. 5. Size of $EU(\Delta^*)$ for different sizes for the IUPs

points might be very slow if there are chance variables with many states, this being the case of NHL, which has a chance variable with 12 states. The exact approach based on linear programming is slightly slower than the heuristic and the outer approximation.

A sensitivity analysis has also been done to evaluate the effect of the size of the intervals of the initial potentials affects the informativeness of the solutions. Figure 4 shows the size of the (interval-valued) expected utility of the optimal policy $EU(\Delta^*)$ as a function of the sizes of the intervals of the potentials. The results based on the linear programming and on the enumeration of the extreme points are exact and therefore coincide. As expected the heuristic is more precise than the outer approximation, which in some cases returns infinite values because of a division by zero. If sharp values are assumed for the IPPs and the intervals are only in the IUPs, the results are those in Fig. 5. The four methods, which differs only in the treatment of the IPPs, produce the same results. Comparing the scales of Figs. 4 and 5, it can be observed that the imprecision in the IPPs has a stronger effect that in the IUPs.

6 Conclusions and Future Work

We have generalized the formalism of influence diagrams to the interval framework by allowing both probabilities and utilities to take interval values. A variable elimination algorithm has been also proposed and preliminary tested. In the experimental part, four different methods for eliminating chance variables have been compared, showing that the best results are obtained if the linear programming approach is considered. As a future work intend to develop approximate evaluation algorithms for these models.

Acknowledgments. This research was supported by the Spanish Ministry of Economy and Competitiveness under project TIN2013-46638-C3-2-P, the European Regional Development Fund (FEDER), the FPI scholarship program (BES-2011-050604) and the short stay in foreign institutions scholarship EEBB-I-14-08102. The authors have also been partially supported by "Junta de Andalucía" under projects TIC-06016 and P08-TIC-03717.

References

1. Hugin Expert network repository. http://www.hugin.com/technology/samples
2. Benferhat, S., Smaoui, S.: Hybrid possibilistic networks. Int. J. Approximate Reasoning **44**(3), 224–243 (2007)
3. Bielza, C., Gómez, M., Insua, S.R., del Pozo, J.A.F., Barreno, P.G., Caballero, S., Luna, M.S.: IctNEO system for jaundice management. Revista de la Real Academia de Ciencias Exactas, Físicas y Naturales **92**(4), 307–315 (1998)
4. Cozman, F.G.: Credal networks. Artif. Intell. **120**, 199–233 (2000)
5. de Campos, L.M., Huete, J.F., Moral, S.: Probability intervals: a tool for uncertain reasoning. Int. J. Uncertainty Fuzziness Knowl. Based Syst. **2**(02), 167–196 (1994)
6. Fagiuoli, E., Zaffalon, M.: Decisions under uncertainty with credal influence diagrams. Technical report, pp. 51–98, IDSIA (1998). (unpublished)
7. Fertig, K.W., Breese, J.S.: Probability intervals over influence diagrams. IEEE Trans. Pattern Anal. Mach. Intell. **15**(3), 280–286 (1993)
8. Howard, R.A., Matheson, J.E.: Influence diagram retrospective. Decision Anal. **2**(3), 144–147 (2005)
9. Huntley, N., Troffaes, M.C.M.: Normal form backward induction for decision trees with coherent lower previsions. Ann. Oper. Res. **195**(1), 111–134 (2012)
10. Jensen, F.V., Nielsen, T.D.: Bayesian Networks and Decision Graphs. Springer Verlag, New York (2007)
11. Kikuti, D., Cozman, F.G., de Campos, C.P.: Partially ordered preferences in decision trees: computing strategies with imprecision in probabilities. In: IJCAI Workshop on Advances in Preference Handling, pp. 118–123 (2005)

12. Kjaerulff, U.: Triangulation of graphs - algorithms giving small total state space. Research report R-90-09, Department of Mathematics and Computer Science, Aalborg University, Denmark (1990)
13. Lucas, P.J.F., Taal, B.: Computer-based decision support in the management of primary gastric non-hodgkin lymphoma. In: UU-CS, vol. 33 (1998)
14. Nielsen, T.D., Jensen, F.V.: Sensitivity analysis in influence diagrams. IEEE Trans. Syst. Man Cybern. Part A Syst. Hum. **33**(2), 223–234 (2003)
15. Raiffa, H.: Decision Analysis: Introductory Lectures on Choices Under Uncertainty. Addison-Wesley, Boston (1968)
16. Shenoy, P.P.: Valuation-based systems for Bayesian decision analysis. Oper. Res. **40**(3), 463–484 (1992)
17. Xu, H., Smets, P.: Reasoning in evidential networks with conditional belief functions. Int. J. Approximate Reasoning **14**(2–3), 155–185 (1996)
18. Zhang, N.L., Poole, D.: Exploiting causal independence in Bayesian network inference. J. Artif. Intell. Res. **5**, 301–328 (1996)