A Linear Programming Based Approach for Evaluating Interval-valued Influence Diagrams

Rafael Cabañas
Department of Mathematics
University of Almería,
Almería, Spain
rcabanas@ual.es

Andrés Cano, Manuel Gómez-Olmedo
Department of Computer Science and
Artificial Intelligence, University of Granada,
Granada, Spain
{acu,mgomez}@decsai.ugr.es

Alessandro Antonucci
Istituto Dalle Molle di Studi
sull'Intelligenza Artificiale (IDSIA)
Lugano, Switzerland
alessandro@idsia.ch

***Abstract*—This paper states the key ideas of a generalized version of variable elimination for evaluating interval-valued influence diagrams. This extension, which is based on linear programming, does not increase the computational complexity and avoids unnecessarily large outer approximations.**

***Index Terms*—Influence diagrams, credal networks, imprecision, probability intervals, probabilistic graphical models**

I. Introduction

Influence diagrams (IDs) [5] are probabilistic graphical models used to solve decision-making problems under uncertainty. Sharp numerical values are required to quantify their parameters (i.e., potentials). This might be an issue with real models, whose parameters are typically obtained from expert judgements or partially reliable data. We consider an interval-valued quantification of the parameters to gain realism. Even though, inference in such models could be done by replacing the operations over sharp potentials with the analogous ones for interval-valued potentials, this might produce unnecessarily large outer approximations. To avoid that, we propose a sophistication of *variable elimination (VE)* based on linear programming. The content of this paper is discussed more in detail in a previous work [1].

II. Background

A discrete ID is defined over a set of chance variables $\mathbf{X}$ and a set of decisions $\mathbf{D}$. The qualitative part is an acyclic directed graph $\mathcal{G}$ with a node for each chance and decision variable. IDs contain a third type of node, namely utility nodes, representing the user preferences. These nodes are jointly denoted as $\mathbf{U}$. The quantitative part is made of a set of *probability potentials* (PPs) that represents the uncertainty, and a set of *utility potentials* (UPs) that represents the user preferences. A PP over two disjoint sets of variables X_I and X_J, denoted as $\phi(X_I|X_J)$, is a map $\phi : \Omega_{X_{I\cup J}} \rightarrow [0,1]$ such that $\sum_{x_I \in \Omega_{X_I}} \phi(x_I|x_J) = 1$ for each $x_J \in \Omega_{X_J}$. Similarly, a UP over X_K, denoted as $\psi(X_K)$, is a map $\psi : \Omega_{X_K} \rightarrow \mathbb{R}$. For each chance node, a PP over the corresponding variable and its parents is defined, while, for each utility node, a UP over the parents should be assessed.

The set of all PPs specifies a multiplicative factorization of the joint probability of $\mathbf{X}$ given $\mathbf{D}$. Thus, an ID is a compact representation of a joint expected utility $EU(\mathbf{X},\mathbf{D}) := \prod_{X\in\mathbf{X}} \phi(X|\Pi_X) \sum_{U\in\mathbf{U}} \psi(\Pi_U)$, where Π_Y is the set of parents of a given node Y. A *policy* for a decision variable D_i is a mapping $\delta_{D_i} : \Omega_{\Pi_{D_i}} \rightarrow \Omega_{D_i}$ associating a state of D_i (i.e., a decision) to its past observations and decisions. Evaluating IDs (i.e., making inference) consists in the identification of the set of optimal strategies, which maximizes the expected value of the sum of the UPs. The *maximum expected utility (MEU)* is the expected value of the utilities when the decision maker takes the optimal policies.

VE is an algorithm for evaluating IDs [5] that eliminates all the variables one by one. In the version for IDs, chance variables are removed by sum while decision are instead eliminated by maximization. In order to remove a variable Y, all the potentials containing such variable in their domains are selected and combined[1], giving as a result a PP and a UP denoted as ϕ_Y and ψ_Y. In case of a chance variable, the new potentials replacing those containing Y are computed with Eq. (1). In case of a decision, Eq. (2) is used instead[2].

$$(\phi'_Y, \psi'_Y) \leftarrow (\sum_Y \phi_Y, \frac{\sum_Y \phi_Y \cdot \psi_Y}{\sum_Y \phi_Y}) \tag{1}$$

$$(\phi'_Y, \psi'_Y) \leftarrow (\phi_Y^{R(Y=y)}, \max_Y \psi_Y) \tag{2}$$

III. Interval-valued IDs

A. Definitions

Interval-valued IDs (IIDs) are a generalization of IDs containing imprecise parameters. For the utilities we base on the interval utilities proposed by Fertig and Brese [4], and we will call them *interval-valued utility potentials* (IUPs) and will be denoted by $\overline{\underline{\psi}}$. For the probabilities, we use the notion of probability interval proposed by de Campos et al. [3], and we will use the term *interval-valued probability potential* (IPP) and will be denoted by $\overline{\underline{\phi}}$. An example of both is given below.

$$\overline{\underline{\psi}}(Y) = \begin{array}{cc} \begin{bmatrix} [-10,-5] \\ [-5,5] \end{bmatrix} & \begin{array}{c} y_1 \\ y_2 \end{array} \end{array} \quad \overline{\underline{\phi}}(X) = \begin{array}{cc} \begin{bmatrix} [.475,.525] \\ [.285,.335] \\ [.190,.240] \end{bmatrix} & \begin{array}{c} x_1 \\ x_2 \\ x_3 \end{array} \end{array} \tag{3}$$

[1]PPs are combined using *multiplication* whereas *addition* is used for UPs.
[2]$\phi^{R(\bullet)}$ denotes the restriction operation.

The difference w.r.t. the precise potentials is that, instead of single value, associated to each configuration there is an interval. Note that an interval-valued potential represents a bounded and infinite set of precise ones (its extension). We will assume that all the IUPs and IPPs have a non-empty extension and satisfy the reachability condition [3]. IIDs offer a direct sensitivity analysis interpretation. An IID can be regarded as a collection of so-called consistent IDs, all with the same graph and set of variables, with PPs and UPs taking their values from the extensions of the IPPs and IUPs of the IID.

B. Evaluation by linear programming

IID evaluation is intended as the calculation of the interval spanned by the MEU values of the consistent IDs. The evaluation could be done by replacing the operations over sharp potentials with the analogous ones for interval-valued potentials [1], e.g., the multiplication of two IPP will be obtained by separately multiplying the lower and the upper bounds. This might produce unnecessarily large outer approximations. Note that the division Eq. (1) could lead to intervals with $-\infty$ or $+\infty$ in their bounds. Instead, we propose to use linear programming for the computation of the potentials resulting at each elimination step. Yet, the combination for interval-valued potentials will be still required for obtaining these programs.

First, let us consider the removal of a chance variable Y from a set of probabilities (i.e., left-hand side in Eq. (1)), then the equivalent operation with IPPs is defined as follows.

Definition 1 (eliminating chance variables from IPPs): Consider the elimination of the chance variable Y during VE. Let $\overline{\underline{\phi}}(X_I|X_J,Y)$ denote the IPP obtained by combining all the IPPs with Y on the right-hand side, and $\overline{\underline{\phi}}(Y,X_K|X_L)$ the only IPP with Y on the left. The elimination of Y from the combination of these two IPPs generates an IPP $\overline{\underline{\phi}}(X_K,X_I|X_L,X_J)$. For each $x_{I\cup K}\in\Omega_{X_{I\cup K}}$ and $x_{L\cup J}\in\Omega_{X_{L\cup J}}$, an outer approximation of the lower bound $\underline{\phi}(x_{K\cup I}|x_{L\cup J})$ is the solution of the following task:

$$
\begin{aligned}
\text{minimize} \quad & \sum_{y\in\Omega_Y}\phi(x_I|x_J,y)\cdot\phi(y,x_K|x_L)\,,\\
\text{subject to} \quad & \underline{\phi}(x_I|x_J,y)\le\phi(x_I|x_J,y)\le\overline{\phi}(x_I|x_J,y)\,,\\
& \underline{\phi}(y,x_K|x_L)\le\phi(y,x_K|x_L)\le\overline{\phi}(y,x_K|x_L),\forall y\in\Omega_Y\,.
\end{aligned}
$$

In case of computing the lower bound, we will replace $\phi(x_I|x_J,y)$ with the lower bound $\underline{\phi}(x_I|x_J,y)$. This reduces the task to a linear program over the optimization variables $\{\phi(y,x_K|x_L)\}_{y\in\Omega_Y}$. Analogously, an outer approximation of the upper bound $\overline{\phi}(x_{K\cup I}|x_{L\cup J})$ can be calculated by maximizing the previous objective function instead and replacing $\phi(x_I|x_J,y)$ with $\overline{\phi}(x_I|x_J,y)$. Now consider the right-hand side in Eq. (1), then the equivalent operation with IUPs is defined as follows.

Definition 2 (eliminating chance variables from IUPs): Let $\overline{\underline{\phi}}(X_I|X_J,Y)$ be the IPP obtained by combining all the IPPs with Y on the right-hand side, $\overline{\underline{\phi}}(Y,X_K|X_L)$ be the only IPP with Y on the left-hand side and $\overline{\underline{\psi}}(Y,X_M)$ the combination of all the IUPs with Y in the argument. The elimination of a chance variable Y from the combination of these potentials produces a new IPP $\overline{\underline{\psi}}(X_I,X_J,X_K,X_L,X_M)$. For each $x_{I\cup J\cup K\cup L\cup M}\in\Omega_{X_{I\cup J\cup K\cup L\cup M}}$, an outer approximation of the lower bound $\underline{\psi}(x_{I\cup J\cup K\cup L\cup M})$ is the solution of the task

$$
\begin{aligned}
\text{minimize} \quad & \frac{\sum_{y\in\Omega_Y}\phi(x_I|x_J,y)\cdot\phi(y,x_K|x_L)\cdot\psi(y,x_M)}{\sum_{y\in\Omega_Y}\phi(x_I|x_J,y)\cdot\phi(y,x_K|x_L)}\,,\\
\text{subject to} \quad & \underline{\phi}(x_I|x_J,y)\le\phi(x_I|x_J,y)\le\overline{\phi}(x_I|x_J,y)\,,\\
& \underline{\phi}(y,x_K|x_L)\le\phi(y,x_K|x_L)\le\overline{\phi}(y,x_K|x_L)\,,\\
& \underline{\psi}(y,x_M)\le\psi(y,x_M)\le\overline{\psi}(y,x_M),\forall y\in\Omega_Y\,.
\end{aligned}
$$

The task has a linearly constrained cubic-fractional objective function. When calculating the lower bound, the minimization with respect to the optimization variables associated to an IUP can be trivially achieved by setting $\psi(y,x_M)=\underline{\psi}(y,x_M)$. We can also regard the product $\phi(y,x_K|x_L)\cdot\underline{\phi}(x_I|x_J,y)$ as a single optimization variable. In this way the task becomes a linear-fractional program which can be reduced to a linear program using the classical Charnes-Cooper transformation [2]. Analogously, a similar procedure can be done for the upper bound $\overline{\psi}(x_{I\cup J\cup K\cup L\cup M})$.

The elimination of a decision (i.e., Eq. (2)) is done using the analogous operations for interval-valued potentials without introducing any imprecision. To obtain the optimal policy, we adopt a conservative approach, called *interval dominance* in the imprecise-probability jargon, which rejects all the decisions leading to certainly sub-optimal strategies.

IV. Conclusions

In this paper we have proposed an extension of the classical VE for evaluating IDs whose parameters are interval-valued potentials. This extension is achieved by local optimization tasks, reduced to linear programs. An empirical analysis of this method can be found in a previous work [1]. As a future work we intend to extend this formalism to more general imprecise frameworks, e.g., credal sets represented by extreme points.

Acknowledgment

Authors have been jointly supported by the Spanish Ministry of Science, Innovation and Universities and by FEDER under the projects TIN2015-74368-JIN, TIN2016-77902-C3-2-P and TIN2016-77902-C3-3-P.

References

[1] R. Cabañas, A. Antonucci, A. Cano, and M Gómez-Olmedo. Evaluating interval-valued influence diagrams. *International Journal of Approximate Reasoning*, 80:393–411, 2017.

[2] A. Charnes and W. W. Cooper. Programming with linear fractional functionals. *Naval Research Logistics Quarterly*, 9(3-4):181–186, 1962.

[3] L. M. de Campos, J. F. Huete, and S. Moral. Probability intervals: a tool for uncertain reasoning. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 2(02):167–196, 1994.

[4] K. W. Fertig and J. S. Breese. Probability intervals over influence diagrams. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15(3):280–286, 1993.

[5] F. V. Jensen and T. D. Nielsen. *Bayesian networks and Decision Graphs*. Springer Verlag, 2007.