

# CREDO: a Military Decision-Support System based on Credal Networks

Alessandro Antonucci

David Huber

Marco Zaffalon

IDSIA (Switzerland)

alessandro@idsia.ch

{david,zaffalon}@idsia.ch

Philippe Luginbühl

Armasuisse (Switzerland)

philippe.luginbuehl@armasuisse.ch

Ian Chapman

Defence Research &

Development (Canada)

ian.chapman@forces.gc.ca

Richard Ladouceur

ADGA (Canada)

rladouceur@adga.ca

**Abstract**—A software tool especially designed for military domains to create and query decision-support systems is presented. Credal networks, which are Bayesian networks whose parameters have the freedom to vary in convex sets, are used to model the relations among the system variables. A novel elicitation procedure of these sets, which allows the military experts to report their knowledge by purely qualitative judgements, is proposed. Two high-level fusion procedures to cope with multiple experts in this framework are also derived. All these features are supported by the software and demonstrated in an application to space security tested during the last NATO multinational experiment.

## I. INTRODUCTION

Human decision making can become problematic when coping with complex systems. Even if the interactions among the factors are well understood, the impact of a decision and/or an observation might be hardly predictable. This is the main motivation for the development of *decision-support systems* (DSSs), intended as software tools providing a rational and measurable backbone to decision problems. Nowadays DSSs are popular in fields like business and industrial processes. This is also the case of military decision making, where several DSSs have been already developed [1]. The possibility of showing the output of a DSS as an objective basis for a particular choice is particularly appreciated in hierarchical organizations such as armies. When modeling military expert knowledge by a DSS, it is important to note that this kind of knowledge is mostly qualitative and hard to encode by deterministic rules. This suggests the opportunity of a probabilistic modeling. Yet, eliciting probabilistic statements from qualitative expert judgements might involve a high level of arbitrariness (e.g., which number denotes the probability of an event which is more probable than its negation?). A more reliable elicitation can be obtained by converting expert qualitative judgements in (linear) constraints (e.g., intervals) for the related probabilities. The corresponding model of uncertainty is the set of consistent probability distributions, i.e., a *credal set* (Sect. II). Credal sets are used to generalize Bayesian networks to a more expressive class of probabilistic graphical models called *credal networks* (Sect. III).

In this paper we use credal networks to implement DSSs. First we describe how multiple expert opinions (modeled by credal sets) can be easily merged in the framework of credal networks (Sect. IV). The discussion is indeed specialized to military domains by presenting a procedure to elicit credal sets

from qualitative judgements (Sect. V). These ideas have been implemented in a Java software tool, called CREDO, as part of a project granted by Armasuisse (the Swiss federal body for the procurement of technologically complex systems and equipment, security-related technologies and quality management). Some implementation details about the architecture and the interface are provided (Sect. VI). Finally, as a demonstrative example, we present a DSS for space security based on a credal network, which have been developed and tested with CREDO during the last NATO multinational experiment (Sect. VII).

## II. FROM PRECISE TO IMPRECISE PROBABILITIES

Given a variable  $X$  taking its possible values in a finite set  $\Omega_X$ , a condition of (subjective) uncertainty about the actual state of  $X$  can be modeled in different ways. Bayesian probability theory adopts a probability mass function  $P(X)$ , i.e., a normalized nonnegative function over  $\Omega_X$ . For each  $x \in \Omega_X$ ,  $P(x)$  is the “probability that  $X = x$ ”. This number, modeling the strength of the beliefs in favour of the event  $X = x$ , has a clear behavioural interpretation based on de Finetti’s betting scheme [2]. Although very effective in most of the cases, classical probability can hardly model scenarios with conflicting information and partial or complete ignorance. This is the motivation behind various generalizations of the Bayesian theory. A popular example is *evidence theory* [3], where  $P(X)$  is replaced by a (normalized and nonnegative) belief function which assigns the probability masses over the power set of  $\Omega_X$ . When assigned to non-singletons, no information about how the mass distributes over the states is given and masses can be also assigned to overlapping sets. Thus, going back to the probabilistic interpretation, the probability mass function associated to a belief function is not univocally specified, being in general characterized only by lower and upper bounds. Walley’s theory of *imprecise probability* [4] offers a de-Finetti-like behavioural interpretation of these generalized uncertainty models, which can be equivalently regarded as sets of distributions consistent with the bounds. This is the basis for an even more general uncertainty theory, which natively copes with *credal sets* (i.e., sets of mass functions). We use notation  $K(X)$  to denote a credal set over  $X$  and put it on a subscript when computing inferences over it. As an example:

$$\underline{E}_{K(X)}[f(X)] := \min_{P(X) \in K(X)} \sum_{x \in \Omega_X} P(x) \cdot f(x) \quad (1)$$

is the lower expectation of a function  $f(X)$ . As a consequence of (1), a credal set and its convex hull return the same inferences. Accordingly we only consider convex credal sets and denote by  $\text{ext}[K(X)]$  the extreme points of  $K(X)$ , whose number is assumed to be finite.

### III. CREDAL NETWORKS

Given a joint variable  $\mathbf{X} := (X_0, X_1, \dots, X_n)$ , a compact specification of a credal set  $K(\mathbf{X})$  can be achieved by a *credal network* [5], which is defined by: (i) an acyclic directed graph  $\mathcal{G}$  whose nodes are in one-to-one correspondence with the variables in  $\mathbf{X}$ ; (ii) a collection of (conditional) credal sets  $\{K(X_i|\text{pa}(X_i))\}_{\text{pa}(X_i) \in \Omega_{\text{Pa}(X_i)}}$  over  $X_i$ , one for each state of the joint variable  $\text{Pa}(X_i)$  (denoting the parents of  $X_i$  in  $\mathcal{G}$ ), for each  $X_i \in \mathbf{X}$ . The above definition coincides with that of a Bayesian network in the special case where the conditional credal sets are “precise”, i.e., they contain a single mass function. Like for Bayesian networks, graph  $\mathcal{G}$  is a visual model of conditional independence relations based on the Markov condition: “every variable is conditionally independent of its non-descendant non-parents given its parents”. Yet, the notion of stochastic independence should be extended to cope with credal sets. Given  $K(X_i, X_j)$ ,  $X_i$  and  $X_j$  are independent if stochastically independent for each  $P(X_i, X_j) \in \text{ext}[K(X_i, X_j)]$ . This definition easily extends to the case of conditional independence. Under these assumptions, a joint credal set  $K(\mathbf{X})$  is obtained by combining the conditional credal sets in (ii) as follows:

$$K(\mathbf{X}) = \text{CH} \left\{ P(\mathbf{X}) \left| \begin{array}{l} P(\mathbf{x}) = \prod_{i=0}^n P(X_i|\text{pa}(X_i)) \\ \forall \mathbf{x} \in \Omega_{\mathbf{X}} \text{ with } (x_i, \text{pa}(X_i)) \sim \mathbf{x} \\ \forall P(X_i|\text{pa}_i) \in K(X_i|\text{pa}(X_i)) \\ \forall i = 0, \dots, n \forall \text{pa}(X_i) \in \Omega_{\text{Pa}(X_i)} \end{array} \right. \right\}, \quad (2)$$

where CH denotes the convex hull and  $\sim$  consistency between states. The credal set in (2) is an imprecise-probabilistic model of the knowledge about the whole joint set of variables. Specific inference algorithms have been developed to extract information from models of this kind according to (1).

### IV. AVERAGING EXPERT OPINIONS

In this section we show how classical approaches to the fusion of probabilistic expert knowledge (e.g., see [6]) can be embedded within the framework of credal networks, and thus extended to support knowledge modeled by credal sets. Let us start from the Bayesian case. Given variable  $X$ ,  $P(X_a)$  and  $P(X_b)$ , with  $\Omega_{X_a} = \Omega_{X_b} = \Omega_X$ , are modeling the knowledge of two experts about  $X$ . Arithmetic and geometric averages are typically considered for an unbiased fusion of these two pieces of probabilistic information:

$$P_+(X = x) \propto P(X_a = x) + P(X_b = x), \quad (3)$$

$$P_\times(X = x) \propto P(X_a = x) \cdot P(X_b = x). \quad (4)$$

Consider a Bayesian network with nodes  $X_a$  and  $X_b$  and no arcs. Both these averages can be easily embedded in the network by simple transformations.

*a) Arithmetic average:* Add to the model the variable  $X$  and an auxiliary binary variable  $D$ , with  $\Omega_D = \{a, b\}$ . Let  $D$  be a parentless node and  $X$  a child of  $X_a$ ,  $X_b$  and  $D$

(see Fig. 1.a). Set a uniform prior over  $D$  and the following conditional model for  $X$ :

$$P(X = x|x_a, x_b, D) = \begin{cases} \delta(x_a, x) & \text{if } D = a \\ \delta(x_b, x) & \text{if } D = b \end{cases}, \quad (5)$$

where  $\delta(x', x'') = 1$  if  $x' = x''$  and zero otherwise. Conditional probabilities in (5) say that  $X = X_a$  if  $D = a$  and  $X = X_b$  if  $D = b$ . In other words  $D$  decides which expert is right (assuming that one and only one of them can be right) and a uniform prior over  $D$  expresses a lack of preference between the two experts. Thus, it is not surprising to check that, marginalizing out the other variables, the Bayesian network states that  $P(X = x)$  reproduces the arithmetic mean in (3). Similarly, a non-uniform  $P(D)$  corresponds to a weighted average between  $P(X_a)$  and  $P(X_b)$ .

*b) Geometric average:* Add to the (original) model an auxiliary Boolean variable  $D'$ , which is a child of  $X_a$  and  $X_b$  (see Fig. 1.b). Set the following conditional model for  $D'$ :

$$P(D' = \text{true}|x_a, x_b) = \delta(x_a, x_b). \quad (6)$$

This corresponds to the logical constraint:  $D'$  is true if and only if  $X_a = X_b$ . By marginalizing out  $X_b$  and conditioning on  $D$  it is possible to verify that  $P(X_a = x|D = \text{true})$  (or, equivalently,  $P(X_b = x|D = \text{true})$ ) reproduces the geometric average in (4).

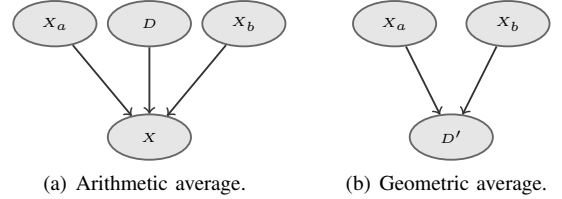


Fig. 1. Averaging expert opinions with directed graphical models.

Both the procedures easily generalize to the case of many experts. The derivation is indeed identical if the variables associated to the expert opinions are not disconnected nodes of the Bayesian network. Furthermore, as credal networks are a generalization of Bayesian networks, the two average procedures can be used also with credal networks, thus providing an average for credal sets too. For the arithmetic average, when coping with credal networks the quantification of  $D$  can be also based on a credal set  $K(D)$ . This corresponds to a *credal* averaging of different models analogous to what has been done for classification in, e.g., [7]. As an example, if  $K(D)$  is the vacuous credal set (i.e., the whole probability simplex) the average is the (convexification of the) union of the credal sets.

$P(X_a = x)$	$P(X_b = x)$	$P(D = a)$	$P_+(x)$	$P_\times(x)$
.20	.40	.50	.30	.14
.20	.80	.50	.50	.50
[.10, .30]	[.30, .50]	.50	[.20, .40]	[.05, .30]
[.10, .30]	[.70, .90]	.50	[.40, .60]	[.21, .79]
[.10, .30]	[.70, .90]	[0, 1]	[.10, .90]	[.21, .79]

TABLE I. GEOMETRIC AND ARITHMETIC AVERAGING OF PRECISE AND IMPRECISE ASSESSMENTS OVER A BOOLEAN VARIABLE  $X$ .

Table I reports some examples of this generalized fusion technique with two experts and  $\Omega_X = \{x, \neg x\}$ . In this case

assessing the probability for a single state implicitly define the whole probability mass function, i.e., if  $P(x) \in [l, u]$  then  $P(\neg x) = [1 - u, 1 - l]$ . Notably, in the imprecise framework, the effect of a dissensus among the experts affects the level of imprecision (i.e., size of the intervals) in the averaged outputs.

## V. DSSS WITH CREDAL NETWORKS

According to the discussion in Section III, the specification of a DSS based on credal networks requires the three following sequential steps:

- (i) identification of the system variables  $\mathbf{X}$ ;
- (ii) specification of an acyclic directed graph  $\mathcal{G}$  modeling conditional independence relations in  $\mathbf{X}$ ;
- (iii) assessment of the conditional credal sets (for each variable given every possible value of the parents).

Regarding (i), an issue is, for each variable, the identification of the possible states, which should correspond to exhaustive and exclusive options. As credal networks are defined here over categorical variables, numerical variables should be discretized. In doing that it is important to set a bound to the maximum number of states to avoid an exponential growth in the number of extremes of the credal sets (see Footnote 2).

Regarding (ii), there are no significant differences from the case of Bayesian networks (see e.g., [8] for some insights). Here we just note that, although the orientation of the arcs does not necessarily express causal relations, connecting the variables with causal arcs certainly produces a proper graph identification.

Step (iii) is generally the most demanding and critical as it requires a (quantitative) assessment of the credal sets from the experts, whose knowledge is, especially in military domains, often qualitative. We therefore present a simple elicitation procedure for credal sets which does not require explicit quantitative assessments by the experts.

### Credal sets elicitation by qualitative judgements

In our assumptions experts express their knowledge about the system by qualitative judgements formulated in natural language. A *dictionary* to convert these judgements in quantitative assessments is therefore needed. Each judgement is intended to evaluate how likely or unlikely, according to the expert, a particular event is. The dictionary can be therefore reduced to a number of *keywords* (adjectives, in practice) expressing different strengths for beliefs.<sup>1</sup> This is also done with Bayesian networks, where keywords should be associated to sharp probabilistic values. Yet, by adopting credal sets, intervals (i.e., lower and upper bounds of probabilities) can replace these sharp values: this provides a more robust and (potentially) reliable model of expert knowledge.

An example is in Fig. 2. Besides adjectives like “very likely” or “fifty-fifty”, note the keyword “-” used to express the lack of an assessment (i.e., complete ignorance about that specific event) and corresponding to no constraint (i.e.,  $0 \leq P(x) \leq 1$ ). Note also that, if a keyword is regarded as a logical negation of another one, we take the complements

of the bounds. E.g., in the dictionary in Fig. 2, the keyword “likely” means  $.6 \leq P(x) \leq .85$  while “ $x$  is unlikely” is intended as “ $\neg x$  is likely” and hence  $.15 \leq P(x) \leq .4$ . Keywords like “fifty-fifty” are intended to model the fact that an event and its negation have approximately the same probability. In the Bayesian case, this would correspond to a sharp symmetric assessment (e.g.,  $[\cdot.5, \cdot.5]$ ), replaced here by a symmetric interval (e.g.,  $[\cdot.35, \cdot.65]$ ).

In summary the above procedure translates qualitative keyword-based judgements into a set of probability intervals  $\{\underline{P}(x), \bar{P}(x)\}_{x \in \Omega_X}$ , with zero/one bounds if no assessments are provided. This defines the consistent credal set:<sup>2</sup>

$$K(X) = \left\{ P(X) \mid \begin{array}{l} \underline{P}(x) \leq P(x) \leq \bar{P}(x), \forall x \in \Omega_X, \\ \sum_{x \in \Omega_X} P(x) = 1 \end{array} \right\}. \quad (7)$$

An example of a consistent credal set is in Fig. 3.

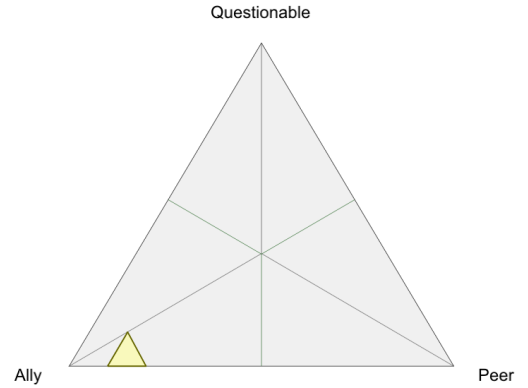


Fig. 3. Credal sets visualization in the CREDO software. A ternary variable called *Type of Partner* whose possible values are *Ally*, *Peer*, and *Questionable* is used in this figure and in the following to geometrically show credal sets. Judgements “very unlikely” for the first state and “very likely” for the second is considered. The tridimensional coordinates of the points are the probabilities of the three states, and the big triangle is the probability simplex.

Assessments over different states of a single variable (given a same configuration of the parents) are not independent. E.g., two states of  $X$  cannot be both “likely” (if the keyword is intended as before) as this would imply a sum of the corresponding probabilities (i.e., the probability that  $X$  is in one of these states) greater than one. As noted in [10], to avoid this kind of problems when coping with probability intervals, the following constraints should be satisfied for each  $x' \in \Omega_X$ :

$$\sum_{x \in \Omega_X \setminus \{x'\}} \underline{P}(x) + \bar{P}(x') \leq 1, \quad (8)$$

$$\sum_{x \in \Omega_X \setminus \{x'\}} \bar{P}(x) + \underline{P}(x') \geq 1. \quad (9)$$

At the software level, the specification of the judgements is a sequential process. To prevent inconsistent assessments, if the user wants to add a new judgement, the system evaluates the above constraints for each keyword and only those leading to valid assessments are visualized (see e.g., Fig. 4).

<sup>1</sup>In our experience with this kind of knowledge engineering, the number of keywords needed by the experts is rarely greater than ten.

<sup>2</sup>The extremes of a credal set of this kind can be computed by a formula which iterates over the permutations of  $\Omega_X$  [9]. Thus,  $|\text{ext}[K(X)]| \leq |\Omega_X|!$ .

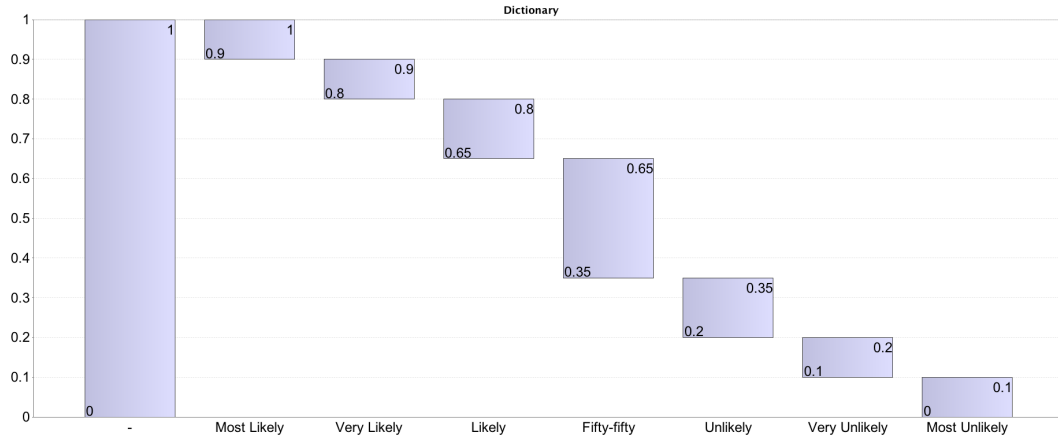


Fig. 2. Histogram view of a dictionary as specified in the CREDO software. The keywords are, from left to right, “-” (i.e., no judgement), “most likely”, “very likely”, “fifty-fifty”, “unlikely”, “very unlikely” and “most unlikely”.

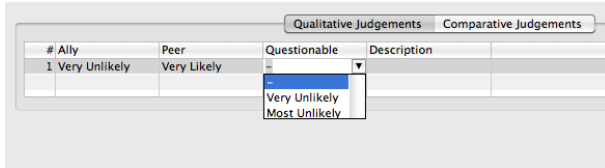


Fig. 4. Preventing inconsistent judgements in the CREDO software. When judgements “very unlikely” and “very likely” have been already specified for the first two states, the third state has only three possible keywords (“-”, “very unlikely” and “most unlikely”), as the other ones would violate (8) or (9).

### Comparative judgements

In the previous paragraph we forced the expert to express his/her beliefs about the probability of a particular event in “absolute” terms by means of a keyword-based judgement translated into a (linear) constraint on a single probability. Here we also consider “comparative” judgements which can be translated into constraints (again linear) over pairs of probabilities. The keywords are in this case models of the relative strength of the beliefs associated to a state when compared to another one. Consider for instance the comparative keyword “much more probable”, which can be used to express that the probability of a state is significantly greater than that of another one. In the dictionary the keyword should be associated to the numerical values involved in the constraint specification. Also in this case we use a positive interval, say  $[\underline{\gamma}, \bar{\gamma}]$  such that, if  $x'$  and  $x''$  are the two states involved in the comparison, the constraint is:

$$\underline{\gamma} \leq \frac{P(x')}{P(x'')} \leq \bar{\gamma}. \quad (10)$$

In our case study (see Section VII), the adjective “much more probable” is associated to  $[\underline{\gamma}, \bar{\gamma}] = [2, 5]$ , this meaning that  $x'$  is at least two and at most five times more probable than  $x''$  (see Fig. 5). Single inequalities can be obtained by setting  $\underline{\gamma} = 0$  or  $\bar{\gamma} = +\infty$ . Both comparative and absolute judgements can be combined in the specification of a credal set. As already explained for absolute judgements, once the user wants to add a new comparative judgements, the software only visualizes the

comparative keywords which do not produce inconsistencies.<sup>3</sup>

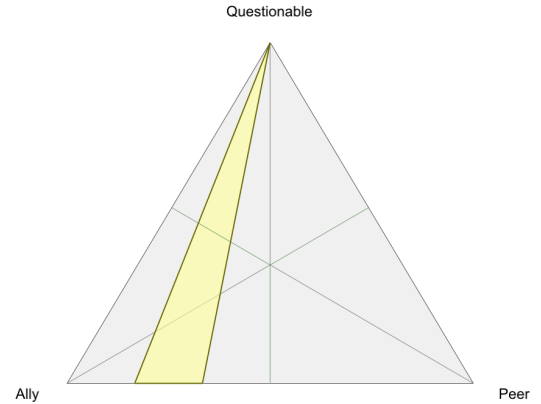


Fig. 5. The credal set associated to a comparative judgement of “much more probable” for the relation between the first and the second state of a ternary variable. See the caption of Fig. 3 for details about the geometric visualization.

### Querying a credal-network-based DSS

Once all the required conditional credal sets of a credal network have been assessed, the model specification is complete. The result is a mathematical, autonomous, model of the expert knowledge about a particular domain. Such a model can be automatically queried by appropriate inference algorithm. In DSS, one is typically interested in the updated beliefs about a variable of interest, say  $X_0$ , given some evidential information gathered for some other system variables, say  $\mathbf{x}' \in \Omega_{\mathbf{X}'}$ , with  $\mathbf{X}' \subset \mathbf{X} \setminus \{X_0\}$ . In the imprecise-probabilistic framework this corresponds to the posterior credal set  $K(X_0|\mathbf{x}')$  corresponding to the marginalization and conditioning of (2). Two examples of posterior credal sets are in Fig. 6. Unlike for credal sets specified in the modeling phase, it might be not possible to express the posterior credal sets in terms of

<sup>3</sup>Conditions for the non-emptiness of the consistent credal set like (8) and (9) cannot be derived for comparative judgements. A linear system to identify possible intersections between the hyper-planes associated to the comparative judgements and the edges of the original credal set should be solved instead.

qualitative keyword-based judgements. We therefore provide a numerical characterization in terms of probability intervals like in the example in Fig. 7.

Posterior credal sets are the typical support we can extract from a DSS based on credal networks. Compared to models based on sharp probabilistic assessment (e.g., Bayesian networks), the support we receive is twofold as it includes: (i) indication about the most likely option for the variable of interest; (ii) indication about the level of indeterminacy (also called imprecision) of the output. Consider for instance the two credal sets in Figs. 6 and 7: both these outputs point to the state “Questionable”, yet in the case of (a) the information is much more precise than in (b). In situations where the output is very precise, the user can decide that the most probable option clearly reflect the actual value of the variable of interest. On the other side, a very imprecise output reflect that, for that particular scenario, the available evidence is not sufficient to clearly identify the decision to take. We point the reader to [11] for a deeper discussion about decision making based on credal sets and to [12] for a focus on decisions in credal networks.

### Explaining outputs

Consider the posterior credal set  $K(X_0|\mathbf{x}')$  associated to an updating task as introduced in the previous paragraph. Assuming that the variable of interest  $X_0$  has no children in  $\mathcal{G}$ , the posterior probabilities, for each  $x_0 \in \Omega_{X_0}$ , can be expressed as:

$$P(x_0|\mathbf{x}') = \sum_{\text{pa}(X_0) \in \Omega_{\text{Pa}(X_0)}} P(x_0|\text{pa}(X_0)) \cdot \left[ \prod_{\tilde{x} \in \text{Pa}(X_0)} P(\tilde{x}|\mathbf{x}') \right], \quad (11)$$

where the fact that each parameter varies in a credal set is left implicit for the sake of readability. According to (11), the posterior probability is a weighted average of the conditional probabilities for the variable of interest given the different possible values of the parents. The weights are the products of the posterior probabilities for each parent of the variable of interest. This remark can be exploited to identify, with respect to a particular updating task, which judgements associated to  $X_0$  contributed more to the observed output (see Fig. 8).

## VI. THE CREDO SOFTWARE TOOL

Here we provide some detail about the CREDO software tool for the development of DSSs based on credal networks. First, let us note that all the above mentioned features, namely, the fusion of expert opinions, the elicitation of credal sets by qualitative judgements and the explanation of the output, are supported by the software.

Regarding implementation, CREDO is a Java application built on top of the *Eclipse Rich Client Platform*.<sup>4</sup> This framework offers a plugin-based environment for a modular development. We extensively used this flexibility to implement a modular and reusable set of features, declaratively assembled into the final application. To compute inferences we currently use the inference algorithm described in [13]. An extension point to allow new algorithms to be added to the system was also implemented. See Fig. 9 for a simplified overview.

#	Partnership Context	Sharing Provisions	Caveats	Acceptable	Not Acceptable	Description
1	Favourable	Altruist-Broad	Strongly Limited	Likely	-	
2	Favourable	Altruist-Broad	Mildly Limited	Very Likely	-	
3	Favourable	Altruist-Broad	No Limitations	Most Likely	-	
4	Favourable	Limited	Strongly Limited	Unlikely	-	
5	Favourable	Limited	Mildly Limited	Fifty-fifty	-	
6	Favourable	Limited	No Limitations	Likely	-	
7	Favourable	Commercial	Strongly Limited	Very Unlikely	-	
8	Favourable	Commercial	Mildly Limited	Fifty-fifty	-	
9	Favourable	Commercial	No Limitations	Likely	-	
10	Neutral	Altruist-Broad	Strongly Limited	Likely	-	
11	Neutral	Altruist-Broad	Mildly Limited	Very Likely	-	
12	Neutral	Altruist-Broad	No Limitations	Most Likely	-	
13	Neutral	Limited	Strongly Limited	Unlikely	-	
14	Neutral	Limited	Mildly Limited	Fifty-fifty	-	
15	Neutral	Limited	No Limitations	Likely	-	
16	Neutral	Commercial	Strongly Limited	Very Unlikely	-	
17	Neutral	Commercial	Mildly Limited	Fifty-fifty	-	
18	Neutral	Commercial	No Limitations	Likely	-	
19	Unfavourable	Altruist-Broad	Strongly Limited	Very Unlikely	-	
20	Unfavourable	Altruist-Broad	Mildly Limited	Fifty-fifty	-	
21	Unfavourable	Altruist-Broad	No Limitations	Likely	-	
22	Unfavourable	Limited	Strongly Limited	Very Unlikely	-	
23	Unfavourable	Limited	Mildly Limited	Unlikely	-	
24	Unfavourable	Limited	No Limitations	Fifty-fifty	-	
25	Unfavourable	Commercial	Strongly Limited	Most Unlikely	-	
26	Unfavourable	Commercial	Mildly Limited	Very Unlikely	-	
27	Unfavourable	Commercial	No Limitations	Very Unlikely	-	

Explanation threshold: 0.6

Fig. 8. Ranking the importance of judgements over the variable of interest in the CREDO software. Judgements such that the lower bound of the corresponding weight in (11) exceed a given threshold are highlighted.

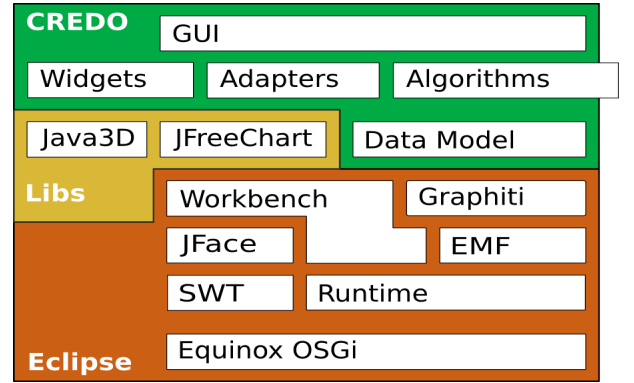


Fig. 9. The different sets of plugins used in the implementation of the CREDO software.

Regarding the data model, we adopted the *Eclipse Modeling Framework*,<sup>5</sup> which offers a number of features to work with structured data models. In particular, we used XMI serialization, the transactions and the observer patterns.

The core classes of the data model are in the UML diagram in Fig. 10. Each credal network (called model in the diagram) is composed by a collection of variables. Each variable contains a number (at least two) states. The model is also composed by a collection of credal sets (credal set interface) associated to both a variable and a (possibly empty) collection of states by the conditioning relation. The concrete classes implementing the credal set interface are in the bottom part of the diagram. Four different specifications are supported: (i) the feasible region of a collection of linear inequalities (H-specification); (ii) the enumeration of the extremes (V-specification); (iii) the probability intervals (interval specification); and (iv) the qualitative judgements (qualitative specification). The latter is based on a combination of single-state and comparative judgements.

<sup>4</sup><http://www.eclipse.org/home/categories/rcp.php>.

<sup>5</sup><http://www.eclipse.org/modeling/emf>.

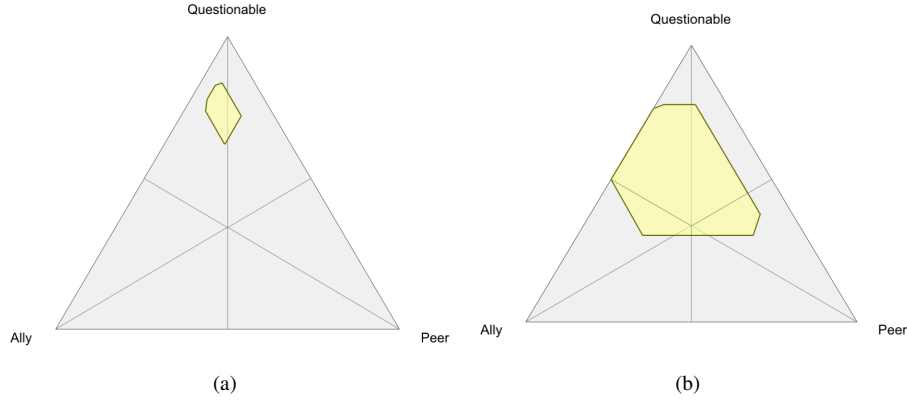


Fig. 6. Posterior credal sets over a ternary variable.

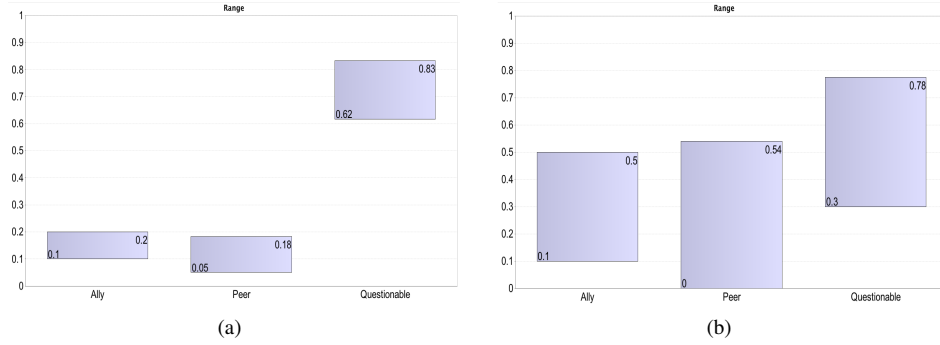


Fig. 7. Probability intervals induced by the posterior credal sets in Fig. 6.

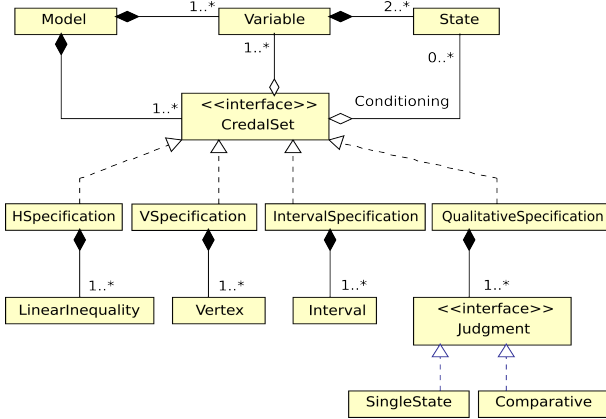


Fig. 10. UML class diagram for credal networks in the CREDO software.

## VII. A DSS FOR SPACE SECURITY WITH CREDO

The *Multinational Experiment Series* is an ongoing unclassified multinational and interagency concept development and experimentation initiative. The series has run since 2001 and conducted by a core group of nations: NATO has been an active participant since 2004. The Multinational Experiment 7 is concerned with protecting our access to the global commons. These are parts of the Earth to which all nations have legal access and contain resources, which have economic or social value when extracted or because of their location. Space is an important domain within the global commons. The vulnerabilities of space systems and nations' dependency on

those systems provide a compelling argument in favour of developing a strategy for increased resiliency to protect space capabilities.

Space deterrence and space defence seek to protect space assets. Nonetheless, there remains a possibility that these systems will fail. Space mitigation strategy manages that risk; considering the probability and impact of a failure, it provides a cost-effective plan of action to maintain critical capabilities to mitigate the potential impact of failure. Inspired by collective response, the *Collaborative Space Mitigation Concept* (CSMC) uses partnership agreements and interoperability to present a space mitigation strategy framework that leverages excess space capacity to increase space systems resiliency efficiently. Space capability developers require a mitigation plan that balances the cost of minimizing the impact against the probability of loss or degradation of space assets. The CSMC is intended to help space capability developers to develop space mitigation strategies to minimize the potential impact of space deterrence and space defence failure. CSMC implementation depends on nation willingness to share space capability. An important issue is therefore to assess the *political acceptability* of bilateral cooperation arrangements in a particular mitigation strategy.

As part of a cooperation between the Swiss and the Canadian Armed Forces, the variables and the graphical structure of a credal network implementing a DSS for the evaluation of the political acceptability has been implemented. The main model variables, together with the list of possible state and some information are reported here below, while Fig. 11 depicts the acyclic directed graph modeling the conditional independence



relations among the model variables.

- *Political acceptability* has possible states *acceptable* and *unacceptable*: this is the variable of interest.
- *Space pillar* has possible states *SATCOM* (command, control, communication and computer systems that are dependent on satellites communication capabilities) , *ISR* (synchronized and integrated planning and operation of all collection capabilities) and *SSA* (ability to obtain information and knowledge about the space beyond the Earth atmosphere).
- *Type of partner* has possible states *ally*, *peer* and *questionable*.
- *Partner capability* has possible states *most advanced*, *average* and *new to space*.
- *Access sharing* has possible states *C2 payload and raw data* (ability to directly manage the beam), *raw data only* and *no direct access*.
- *Compensation* has possible states *in-kind*, *small*, *medium* and *large compensation*.
- *Purpose limitation* has possible states *peaceful and non-economic*, *peaceful only* and *no limitations*.
- *Geographical limitation* has possible states *peace-keeping exclusion*, *partner exclusion* and *no limitations*.

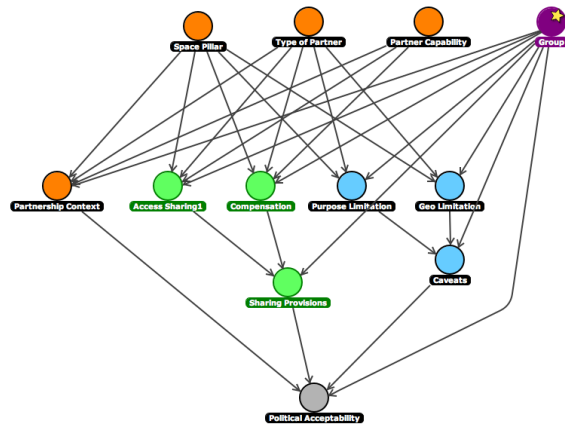


Fig. 11. A DSS to assess the political acceptability of a CSMC.

During the final meeting of the experiment a group of six subject matter experts was formed. This group was divided into two groups of three individuals. Each group developed its own conclusions for 27 different scenarios corresponding to all the possible combinations of variables *space pillar*, *type of partner* and *partner capability*.

Two subsequent approaches were considered: an informal group discussion and a formal individual approach based on the use of CREDO. One group developed the operational acceptability and constraints acceptability first through the informal discussion and then repeated the exercise through the use of CREDO, while the other group proceeded in the reverse order.

At the beginning of the experiment, the participants defined their own set of keywords and relative interpretation (see Fig. 2). A shared dictionary among the experts with seven judgments proved to be an acceptable trade-off and has been applied throughout the experiment as a subjective scale for the quantification of the network. The complete freedom in creating the dictionary and the possibility of using intervals were greatly appreciated by the participants. Those two advantages were clearly a motivation in favor of an increase in the degree of acceptance of the model and the CREDO support tool. With the new network, both groups were able to assess 27 scenarios completely, during three CREDO sessions, each session being a half day or four hours.

During the sessions with CREDO the groups were therefore required to assess the conditional credal sets for the variable in the network in Fig. 11. Differences have been identified in the quantification of the different variables by the two groups. As an example Fig. 12 reports a comparison of the assessments reported for the variable *political acceptability* by the two groups.

#	Partnership Conte	Sharing Provisions	Caveats	Acceptable	Acceptable
1	Favourable	Altruist-Broad	Strongly Limited	Very Likely	Likely
2	Favourable	Altruist-Broad	Mildly Limited	Most Likely	Very Likely
3	Favourable	Altruist-Broad	No Limitations	1	Most Likely
4	Favourable	Limited	Strongly Limited	Likely	Unlikely
5	Favourable	Limited	Mildly Limited	Likely	Fifty-fifty
6	Favourable	Limited	No Limitations	Very Likely	Likely
7	Favourable	Commercial	Strongly Limited	Fifty-fifty	Very Unlikely
8	Favourable	Commercial	Mildly Limited	Likely	Fifty-fifty
9	Favourable	Commercial	No Limitations	Likely	Likely
10	Neutral	Altruist-Broad	Strongly Limited	Likely	Likely
11	Neutral	Altruist-Broad	Mildly Limited	Very Likely	Very Likely
12	Neutral	Altruist-Broad	No Limitations	Most Likely	Most Likely
13	Neutral	Limited	Strongly Limited	Likely	Unlikely
14	Neutral	Limited	Mildly Limited	Likely	Fifty-fifty
15	Neutral	Limited	No Limitations	Very Likely	Likely
16	Neutral	Commercial	Strongly Limited	Fifty-fifty	Very Unlikely
17	Neutral	Commercial	Mildly Limited	Likely	Fifty-fifty
18	Neutral	Commercial	No Limitations	Very Likely	Likely
19	Unfavourable	Altruist-Broad	Strongly Limited	Unlikely	Very Unlikely
20	Unfavourable	Altruist-Broad	Mildly Limited	Fifty-fifty	Fifty-fifty
21	Unfavourable	Altruist-Broad	No Limitations	Likely	Likely
22	Unfavourable	Limited	Strongly Limited	Most Unlikely	Very Unlikely
23	Unfavourable	Limited	Mildly Limited	Unlikely	Unlikely
24	Unfavourable	Limited	No Limitations	Fifty-fifty	Very Unlikely
25	Unfavourable	Commercial	Strongly Limited	Most Unlikely	Most Unlikely
26	Unfavourable	Commercial	Mildly Limited	Unlikely	Very Unlikely
27	Unfavourable	Commercial	No Limitations	Fifty-fifty	Very Unlikely

Fig. 12. Expert judgements about *political acceptability* given the possible values of the parent variables as reported by the two groups (last two columns).

After quantification, the model was queried to evaluate the posterior probabilities of the political acceptability for each scenario (and each group). Results are in Fig. 13. In its aim to represent the subject of political acceptability, the network proved to be an acceptable model, despite the fact that the experts were not involved in the construction of the network.

In order to increase the acceptability of a network as a representation of a complex situation, the experiment confirmed the fact that it is always difficult for an expert to trust or accept a model that has been created by others. Despite its simplicity, the so-obtained DSS facilitated valuable discussions and the results generated by both groups proved to be very similar, being also consistent with the assessment provided by the experts during the informal discussions.

Feedback from the participants revealed that CREDO performed well. The CREDO approach was likened to the U.S. military decision-making procedure, and the suggestion was made that CREDO could be applied to deliberate planning and decision-making processes, provided that the decision model is properly designed. Another comment praised the visualization of the results, as the decision and uncertainty can be easily

communicated to a commander or executive who has asked his staff to provide a recommendation. The benefit of displaying the uncertainty level as a range is that a commander can direct his staff to do more research into key portions of the decision network in order to increase the reliability of the result.

## VIII. CONCLUSIONS

A software to implement decision-support systems by means of a class of directed imprecise-probabilistic graphical models called credal networks have been presented and tested on a real-world application. As a future work we intend to develop novel and better modeling features for this tool in order to continue to improve the decision support.

## REFERENCES

- [1] P. Davis, J. Kulick, and M. Egner, *Implications of Modern Decision Science for Military Decision-Support Systems*. Santa Monica: RAND, 2005.
- [2] B. de Finetti, *Theory of Probability: A Critical Introductory Treatment*. Chichester: John Wiley & Sons, 1974–1975.
- [3] G. Shafer, *A Mathematical Theory of Evidence*. Princeton University Press, 1976.
- [4] P. Walley, *Statistical Reasoning with Imprecise Probabilities*. Chapman and Hall, 1991.
- [5] F. G. Cozman, “Credal networks,” *Artificial Intelligence*, vol. 120, pp. 199–233, 2000.
- [6] C. Genest and J. Zidek, “Combining probability distributions: A critique and an annotated bibliography,” *Statistical Science*, vol. 1, pp. 114–148, 1986.
- [7] G. Corani and A. Antonucci, “Credal Ensembles of Classifiers,” *Computational Statistics and Data Analysis*, 2012.
- [8] D. Koller and N. Friedman, *Probabilistic Graphical Models*. MIT press, 2009.
- [9] A. Chateauneuf and J.-Y. Jaffray, “Some characterizations of lower probabilities and other monotone capacities through the use of möbius inversion,” *Mathematical Social Sciences*, vol. 17, no. 3, pp. 263–283, 1989.
- [10] 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*, vol. 2, pp. 167–196, 1994.
- [11] M. Troffaes, “Decision making under uncertainty using imprecise probabilities,” *International Journal of Approximate Reasoning*, vol. 45, no. 1, pp. 17–29, 2007.
- [12] A. Antonucci and C. de Campos, “Decision making by credal nets,” in *Proceedings of the International Conference on Intelligent Human-machine Systems and Cybernetics (IHMSC 2011)*, vol. 1. Hangzhou (China): IEEE, 2011, pp. 201–204.
- [13] A. Antonucci, C. de Campos, D. Huber, and M. Zaffalon, “Approximating credal network inferences by linear programming,” in *Proceedings of the 12th European Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty (ECSQARU 2013)*, 2013.

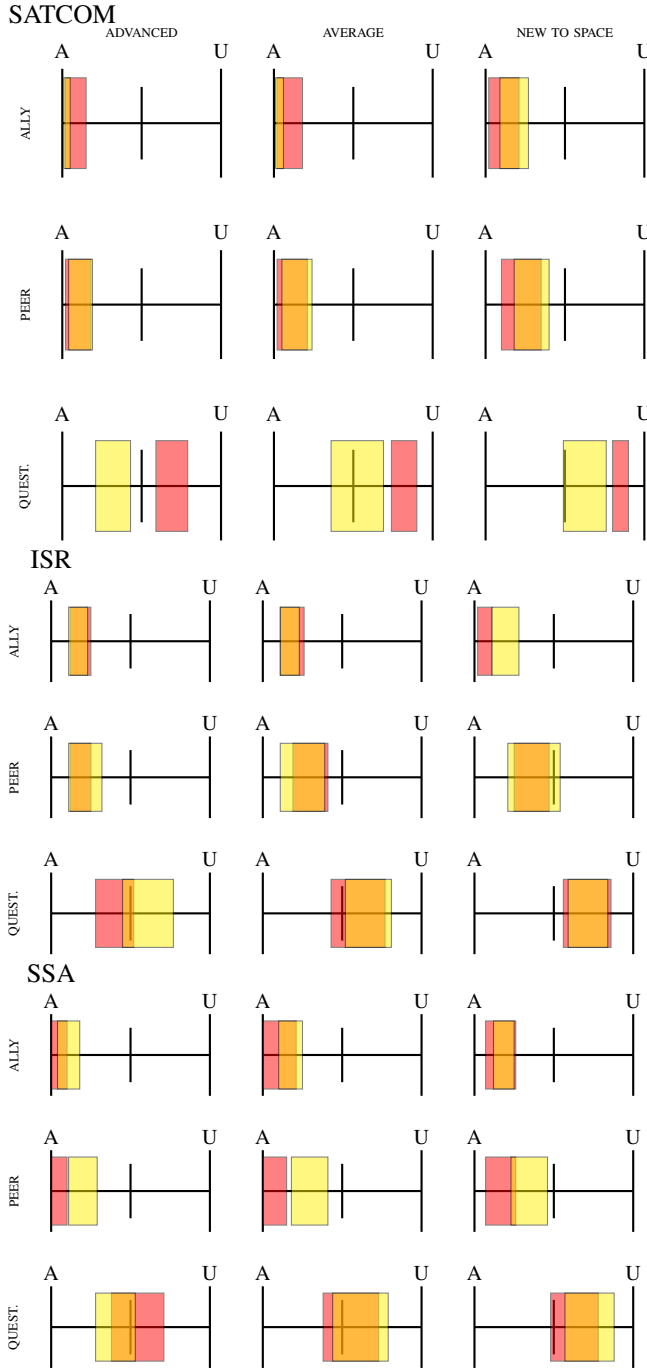


Fig. 13. Inferences for the political acceptability. Labels describe the values of the observed variables. The preference between acceptable (A) and unacceptable (U) is visualized as probability interval for the second option. Yellow and red histograms denote the inferences for the two groups. Orange areas corresponds to partial overlaps between the two intervals.