Credal Valuation Network for Ongoing Threat Assessment

Branko Ristic
School of Engineering
RMIT University
Melbourne, Australia
branko.ristic@rmit.edu.au

Alessio Benavoli
Trinity College Dublin
The University of Dublin
Dublin, Ireland
alessio.benavoli@tcd.ie

***Abstract*—The paper develops a valuation network for sequential assessment of threat under epistemic uncertainty based on theoretical foundations and semantics of imprecise probability theory. The valuations are expressed as credal sets defined by coherent probability intervals on singletons. The combination rule is the generalized Bayes rule introduced by Walley. The model of a single-target threat is based on the classical "capability-intent" paradigm in an air surveillance context. Numerical results illustrate the performance of developed credal valuation network (with imprecise probabilities) against the valuation network with precise probabilistic models.**

I. Introduction

Classical information theory [1] deals with communication and storage of information. In the modern world, with the volumes of data growing at unprecedented levels, the focus of information theory has shifted to machine intelligence, and in particular, to automated systems for reasoning and decision making. A reasoning system must be able to combine the information coming from different sources and the available knowledge-base, in order to extract those parts that are relevant to specific questions. Knowledge and data appear in many different forms, such as statistical models, physical measurements, human declarations, images, video clips, contextual information, domain knowledge. If the knowledge-base and data are certain, we use logic for reasoning. Typically, however, data is affected by some level of uncertainty [2]. Therefore, the first requirement for reliable machine reasoning is a mathematical formulation of a trustworthy quantitative model of uncertain information. The second requirement is that the combination and extraction of desired information must obey certain rules of "coherence" [3]. The rules are natural: 1. The result of combining two pieces of information must be information; 2. The result of extracting a piece of information must be information; 3. The combination process must not add information that is not contained in the pieces of information being combined; 4. The result of the combination of two or more pieces of information must be independent of the order in which they are combined; 5. Extracting the information of interest after combining two pieces of information must be equivalent to extracting the information of interest from each piece of information separately followed by the combination of the resulting information. A system that satisfies the aforementioned rules is referred to as an algebraic system. The algebra of information [3] is an algebraic system made up of information. A particular type of information algebra is called the valuation algebra, where by *valuation* we refer to a quantified representation of an uncertain piece of information. The rules (axioms) of valuation algebra are satisfied in practically all frameworks of uncertainty modeling, e.g. probability theory [3], possibility theory [4], Dempster-Shafer theory [3] and imprecise probability theory [5], leading to the development and application of the corresponding valuation networks [4], [6]–[9].

This paper develops a credal valuation network (CVN) [5] for sequential assessment of threat under epistemic uncertainty based on theoretical foundations and semantics of imprecise probability theory. This machine reasoning system is developed to support decision making by human operators. The valuations in CVN are expressed as credal sets defined by coherent probability intervals on singletons, while the combination rule is the generalized Bayes rule introduced by Walley [10]. The model of threat for a single target is proposed based on the classical "capability-intent" paradigm. The developed CVN for ongoing threat assessment is demonstrated using a temporal sequence of inputs from surveillance sensors. Numerical results are compared against the corresponding valuation network which is using the precise values of probabilities (rather in intervals). We point out that imprecise probabilities have have been used earlier in a somewhat similar context of supporting military decisions [11], by application of *credal networks*, that is, Bayesian networks whose parameters have the freedom to vary in convex sets. Note that credal networks represent a special case of a CVNs [5, Appendix A].

The paper is organised as follows. Sec. II presents a brief review of credal valuation networks. Sec. III describes the proposed valuation network for a single-target threat assessment, consisting of 11 binary variables and 12 valuations which codify the (local) relationships among the subsets of variables. Numerical results obtained using a temporal sequence of incoming information are presented in Sec. V, followed by conclusions of the study in Sec. VI.

II. A BRIEF REVIEW OF CREDAL VALUATION NETWORKS

A. Valuation networks (VNs)

Valuation network (a.k.a. valuation based system) was introduced by Shenoy [12] as a general framework for knowledge representation and reasoning under uncertainty in expert systems. Practical problems are modeled in this framework by a network of interrelated entities, called variables. Let $\mathbf{V}$ be the set of all variables in the network. Each variable can take values in a discrete-state space, called the *frame*. The frame of variable $X \in \mathbf{V}$ is denoted Θ_X.

The (uncertain) relationships between variables are represented by the functions called valuations. Let the set of all valuations in a network be denoted by Φ. A valuation $\varphi \in \Phi$ specifies the relationship between a subset of variables, referred to as its domain $d(\varphi) \subseteq \mathbf{V}$. Mapping $d : \Phi \rightarrow 2^{\mathbf{V}}$ is referred to as the *labeling operation*. In order to explain how the specification of a valuation $\varphi \in \Phi$ is expressed mathematically, let us introduce the notion of a frame (sample space) of a set of variables (domain) $\mathbf{D} = d(\varphi)$. This frame, denoted $\Theta_{\mathbf{D}}$, represents a set of possible configurations of $\mathbf{D}$. Suppose the frame of variable $X \in \mathbf{D}$ is Θ_X. Then, the frame of $\mathbf{D}$ is given by $\Theta_{\mathbf{D}} \stackrel{\triangle}{=} \times\{\Theta_X : X \in \mathbf{D}\}$, where $\times$ denotes the Cartesian product. A valuation φ specifies the relationship between the variables in $\mathbf{D} = d(\varphi)$ by assigning beliefs, expressed as numerical values, to the elements of the frame of $\Theta_{\mathbf{D}}$.

A graphical representation of a valuation network is a hypergraph, where variables are nodes, while valuations are edges that join any number of nodes. This graphical representation codifies the domain knowledge and input (measured) data for automated reasoning about the problem.

There are two basic operations with valuations.

- **Combination** $\otimes$**.** If $\varphi_1, \varphi_2 \in \Phi$ are two valuations, then the combined valuation $\varphi_1 \otimes \varphi_2$ represents the aggregated knowledge from φ_1 and φ_2.
- **Marginalization** $\downarrow$**.** If $\varphi \in \Phi$ and $\mathbf{C} \subseteq d(\varphi)$, then the marginalized valuation $\varphi^{\downarrow \mathbf{C}}$ represents the knowledge obtained by focusing φ from $d(\varphi)$ to $\mathbf{C}$.

Given a finite set of valuations $\Phi = \{\varphi_1, \ldots, \varphi_r\}$, inference refers to marginalization (focusing) of all available knowledge, expressed by the joint valuation $\otimes\Phi = \varphi_1 \otimes \cdots \otimes \varphi_r$, to a subset of variables $\mathbf{D}^o \subseteq \mathbf{V}$, called *decision variables*.

A straightforward approach to inference would be to compute the joint valuation $\otimes\Phi$ first, followed by its marginalisation to $\mathbf{D}^o$. Unfortunately, this would be cumbersome in practice, even for a small scale valuation network. The computational load grows because the domain size increases with each combination, whereas the complexity grows exponentially with the domain size. By imposing certain axioms for *combination*, *marginalization* and *labeling* operations [13]–[15], however, it is possible to compute the marginal $(\otimes\Phi)^{\downarrow \mathbf{D}^o}$ on local domains, without the need to explicitly compute the joint valuation. The list of axioms was given in Introduction [3]. The concept of local computations is carried out by the *fusion algorithm*, which eliminates sequentially all variables $X \in \mathbf{V} \setminus \mathbf{D}_o$ which are of no interest to the inference problem [7], [8], [14].

The fusion algorithm is applied over a structure called the binary joint tree (BJT), where all combinations are carried on pairs of valuations, that is on a binary basis (two-by-two). Finally, marginals are computed by means of a message-passing scheme among the nodes of the BJT. Full details of software implementation of a generic valuation network can be found in [7], [8], [14].

The concept of valuation algebra is very general with some of the best known instantaniations listed next: (1) In the context of probability theory, valuations are expressed by probability mass functions (PMFs), where the combination rule is a point-wise multiplication (with normalisation carried out by summation) [3]; (2) In the context of possibility theory, valuations are possibility functions, while the combination operator is a point-wise multiplication with normalisation by the maximum operation [4]; (3) Valuations are belief functions in the context of evidence theory, with Dempster's rule as the combination operation [16]. In this paper we will adopt the theoretical foundations and semantics of imprecise probability theory and represent valuations as a special class of credal sets, defined by probability intervals on singletons [5], [17]–[19]. The corresponding valuation network is referred to as the credal valuation network.

B. Credal valuation networks

In this paper, we assume that the possibility space (frame) of all the variables is discrete. A credal set is then a closed convex set of PMFs. Consider a valuation with domain $d(\varphi) = \mathbf{D} \subseteq \mathbf{V}$, whose frame is $\Theta_{\mathbf{D}}$. The totally uninformative credal set on $\Theta_{\mathbf{D}}$, referred to as the *vacuous* credal set, contains all possible PMFs on $\Theta_{\mathbf{D}}$ and is defined as:

$$\mathcal{P}^{\mathbf{D}} = \Big\{ p : p(x_i) \geq 0, i = 1, \ldots, |\Theta_{\mathbf{D}}|, \text{ and } \sum_{i=1}^{|\Theta_{\mathbf{D}}|} p(x_i) = 1 \Big\}. \tag{1}$$

Any other (more informative) credal set over $\Theta_{\mathbf{D}}$ can be defined by imposing additional constraints to $\mathcal{P}^{\mathbf{D}}$. The most informative credal set is the one that contains a single (precise) PMF. The case where all valuations in the network are precise is treated as the valuation algebra of PMFs, mentioned earlier [3].

We consider valuations expressed as a credal set defined by probability intervals on singletons, i.e.

$$K^{\mathbf{D}} = \big\{ p \in \mathcal{P}^{\mathbf{D}} : \underline{p}_i \leq p(x_i) \leq \overline{p}_i, i = 1, \ldots, |\Theta_{\mathbf{D}}| \big\}. \tag{2}$$

The choice of probability intervals on singletons in (2) is not arbitrary. First, in order to avoid that $K^{\mathbf{D}}$ defined by (2) is empty, the following condition must be satisfied [18]:

$$\sum_{i=1}^{|\Theta_{\mathbf{D}}|} \underline{p}_i \leq 1 \leq \sum_{i=1}^{|\Theta_{\mathbf{D}}|} \overline{p}_i. \tag{3}$$

Furthermore, probability intervals should also satisfy the two conditions of *reachability* [18]. If the credal set is defined

with probability intervals $[\underline{p}_i, \overline{p}_i]$, for $i = 1, \ldots, |\Theta_{\mathbf{D}}|$, then the reachability conditions are:

$$\sum_{j \neq i} \underline{p}_j + \overline{p}_i \leq 1, \text{ and } \sum_{j \neq i} \overline{p}_j + \underline{p}_i \geq 1, \tag{4}$$

for $i = 1, \ldots, |\Theta_{\mathbf{D}}|$. According to Walley [10, Sec.2.7], probability intervals which satisfy (3) and (4) are *coherent*. We will only consider credal sets defined by (2), with probability intervals that satisfy (3) and (4). Credal sets represent an epistemic generalisation of valuations specified by PMFs in probabilistic valuation networks.

Combination operator. Suppose two beliefs from independent sources are expressed on domain $\mathbf{D}$ as credal sets $K_1^{\mathbf{D}} \in \Phi_{\mathbf{D}}$ and $K_2^{\mathbf{D}} \in \Phi_{\mathbf{D}}$. The credal set of the combined (fused) belief on $\mathbf{D}$, i.e.

$$K_{12}^{\mathbf{D}} = K_1^{\mathbf{D}} \otimes K_2^{\mathbf{D}}, \tag{5}$$

can be expressed in the from (2):

$$K_{12}^{\mathbf{D}} = \{p \in \mathcal{P}^{\mathbf{D}} : \underline{p}_i \leq p(x_i) \leq \overline{p}_i, \text{ for } i = 1, \ldots, |\Theta_{\mathbf{D}}|\}, \tag{6}$$

where the lower probability of configuration $x_i \in \Theta_{\mathbf{D}}$ is defined as [5]:

$$\underline{p}_i = \quad = \min_{\substack{p_1 \in K_1^{\mathbf{D}};\, p_2 \in K_2^{\mathbf{D}} \\ \text{s.t. } \sum_{x_j \in \Theta_{\mathbf{D}}} p_1(x_j) p_2(x_j) > 0}} \frac{p_1(x_i) p_2(x_i)}{\sum_{x_j \in \Theta_{\mathbf{D}}} p_1(x_j) p_2(x_j)}. \tag{7}$$

Eq. (7) minimises the normalised point-wise multiplication of PMFs p_1 and p_2 over all $p_1 \in K_1^{\mathbf{D}}$ and $p_2 \in K_2^{\mathbf{D}}$, such that p_1 and p_2 are not in total conflict. It is easy to verify that if credal sets K_1 and K_2 are singletons (i.e. two PMFs), then (7) reduce to the combination rule as a point-wise multiplication with normalisation, as in the probabilistic valuation networks [3].

The upper probability of configuration $x_i \in \Theta_{\mathbf{D}}$, i.e. $\overline{p}_i$, is defined similarly, with minimisation in (7) replaced by maximisation. Further theoretical details as well as an explanation for practical computations, can be found in [5].

III. Credal Valuation Network for Threat Assessment

According to the JDL data fusion panel [20], threat assessment is a higher-level fusion function which projects the current situation (on the battlefield) to the future and in combination with the adversary doctrine and objectives, predicts the risks and consequences. Threat models are typically based on the classic "opportunity-capability-intent" paradigm [21]. Various threat models have been proposed in the literature, either in the form of a Bayesian network [22]–[25] or an evidential network [8], [26].

A. A valuation-based graphical model of threat

We consider a threat model whose graphical representation is shown in Fig. 1. This is a single-target model in the context of air surveillance, resembling the model in [8]. There are 11 variables and 12 valuations. The set of decision variables is a singleton $\mathbf{D}^o = \{\mathrm{T}\}$. A complete list of variables and their frames are given in Table I. Note that all variables, except P, are binary, that is taking values as either *false* (0) or *true* (1).

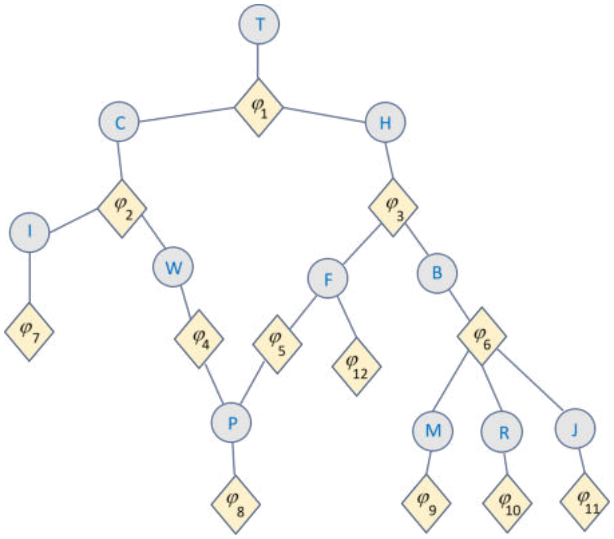

Fig. 1. Valuation network representing the threat model

TABLE I
Variables of the valuation network in Fig. 1

Variable	Name	Frame
T	Threat	$\{0, 1\}$
C	Capability	$\{0, 1\}$
H	Hostile intent	$\{0, 1\}$
I	Imminence	$\{0, 1\}$
W	Weapons	$\{0, 1\}$
P	Platform	$\{0, 1, 2, 3, 4, 5\}$
F	Foe	$\{0, 1\}$
B	Behaviour	$\{0, 1\}$
M	Attacking maneuver	$\{0, 1\}$
R	Illuminating radar	$\{0, 1\}$
J	Deceptive jamming	$\{0, 1\}$

Valuations can be categorised into those which codify the domain-knowledge and those which describe observations (inputs). Domain-knowledge valuations are φ_1, φ_2, φ_3, φ_4, φ_5, φ_6. Their specification is given in Table II. Threat T is specified by conjunction of capability C and hostile intent H. Note that threat was defined in [8] as a superposition of C and H. Superposition is not used here because it increases the frame size of variable T and consequently the computational load. The rationale for using superposition was that it gives a scale of the levels of threat. However, we can express the level of threat even when T is binary by the level of belief mass assigned to *true* (1) or *false* (0). Similarly, capability is conjunction of imminence I and target weapons W. Hostile intent H is specified by disjunction of threatening behaviour B and the fact that the target is hostile (or foe F). Similarly, behaviour B is a disjunction of the observation that the target has performed an attacking manoeuvre M, turned-on its illuminating radar (to direct its weapon) and applied deceptive

jamming to hide its position. Valuations φ_4 and φ_5 are defined by conditional probability tables (CPTs). The CPT for φ_4 contains probabilities $Pr\{\mathrm{W}=1|\mathrm{P}=i\}$, for $i=0,1,\ldots,5$. Due to normalisation, $Pr\{\mathrm{W}=0|\mathrm{P}=i\}=1-Pr\{\mathrm{W}=1|\mathrm{P}=i\}$, for $i=0,1,\ldots,5$. Similarly, the CPT for φ_5 contains probabilities $Pr\{\mathrm{F}=1|\mathrm{P}=i\}$, for $i=0,1,\ldots,5$.

TABLE II
Domain-knowledge valuations in Fig. 1

Valuation	Specfication	Confidence
φ_1	T = C $\wedge$ H	α_1
φ_2	C = I $\wedge$ W	α_2
φ_3	H = F $\vee$ B	α_3
φ_4	CPT	
φ_5	CPT	
φ_6	B = M $\vee$ R $\vee$ J	α_6

Fig. 2 shows the BJT constructed for the valuation network in Fig. 1, using the following elimination sequence: M, R, J, P, B, I, F, W, H, C. The nodes in the BJT are labelled by integer numbers from 1 to 11. The leaves of the tree (the nodes labelled from 1 to 12) represent the original valuations $\varphi_1,\ldots,\varphi_{12}$. The remaining nodes in the BJT represent the intermediate steps of the fusion algorithm; as such they specify the order in which the valuations must be pairwise combined in order to calculate the valuation for the variable T. The vertical labels next to the nodes of the BJT denote the domains (the subsets of variables) of the nodes.

B. Credal set representations

Valuations are expressed as credal sets. This section briefly explains a practical computer implementation of such a representation.

Consider first a simple input valuation, such as φ_7. This valuation expresses our belief about imminence of an attack, using for example the measurements of target range and range-rate. Suppose the confidence in an imminent attack is α_7. Table III shows the computer representation of such a (simple) valuation, for two cases: (1) when α_7 is a precise value (probability); (2) when α_7 is specified by a probability interval, i.e. $\alpha_7\in[\underline{\alpha}_7,\overline{\alpha}_7]$.

TABLE III
Computer representation of a simple valuation, such as φ_7

variable I	prob.	prob. interval
0	$1-\alpha_7$	$[1-\overline{\alpha}_7, 1-\underline{\alpha}_7]$
1	α_7	$[\underline{\alpha}_7,\overline{\alpha}_7]$

Valuations φ_1 and φ_2 are specified by the logical AND operation. Consider for example φ_1, defined in Table II. Table IV lists its configurations and assigned probability masses (precise and interval valued). There are four configurations (number 1, 3, 5 and 8) which satisfy the logical AND operation T = C $\wedge$ H. The probability mass α_1 is equally distributed among them, and thus they are allocated probability mass $\alpha_1/4$. The remaining $(1-\alpha_1)$ is equally distributed across the other four configurations (i.e. 2, 4, 6 and 7). If confidence α_1 is specified by a probability interval, i.e. $\alpha_1\in[\underline{\alpha}_1,\overline{\alpha}_2]$, then the corresponding probability mass intervals assigned to the configurations of $\Theta_{T,C,H}$ are given in the sixth column of Table IV.

TABLE IV
Computer representation of AND valuation φ_1

Config.	T	C	H	prob.	prob. interval
1	0	0	0	$\alpha_1/4$	$[\underline{\alpha_1}/4, \overline{\alpha_1}/4]$
2	1	0	0	$(1-\alpha_1)/4$	$[1-\overline{\alpha_1}/4, 1-\underline{\alpha_1}/4]$
3	0	1	0	$\alpha_1/4$	$[\underline{\alpha_1}/4, \overline{\alpha_1}/4]$
4	1	1	0	$(1-\alpha_1)/4$	$[1-\overline{\alpha_1}/4, 1-\underline{\alpha_1}/4]$
5	0	0	1	$\alpha_1/4$	$[\underline{\alpha_1}/4, \overline{\alpha_1}/4]$
6	1	0	1	$(1-\alpha_1)/4$	$[1-\overline{\alpha_1}/4, 1-\underline{\alpha_1}/4]$
7	0	1	1	$(1-\alpha_1)/4$	$[1-\overline{\alpha_1}/4, 1-\underline{\alpha_1}/4]$
8	1	1	1	$\alpha_1/4$	$[\underline{\alpha_1}/4, \overline{\alpha_1}/4]$

Similarly, we represent valuations φ_3 and φ_6 by the logical OR operation. Details are omitted, but can be worked out easily from the explanation of the logical AND. Finally, valuations φ_4 and φ_5 consist of 12 configurations of pairs (P,W) and (P,F), respectively, The probability masses, as precise and interval values, for φ_4 and φ_5 will be given in the next section.

IV. Eliciting and estimating probability intervals

In the previous sections, we explored probabilistic inference in credal valuation networks focusing on valuations expressed as probability intervals. Now, we will delve into eliciting probability intervals from experts and learning them from data.

Eliciting credal sets as probability intervals from human experts is straightforward by using pairwise judgments. For example, consider the binary variable 'imminence of an attack' which can be True (T) or False (F). Stating that "probability of imminence being true is greater or equal than probability of being false" directly translates into $p_T\geq p_F=(1-p_T)$ leading to the probability interval $p_T\in[0.5,1]$.

When human experts are not available, one can learn probability intervals (credal sets) from data. In this case, probability intervals can arise from two main sources: (1) a lack of knowledge about the probabilistic model generating the data and (2) uncertainty in the parameters of the probabilistic model we use for learning from data. This area has been extensively studied in robust statistics [10], [27]–[31]. In this paper the focus is on estimating probability intervals for discrete binary variables. Consider again variable 'imminence' $\mathrm{I}\in\{0,1\}$. Our goal is to learn $p(\mathrm{I}=1)=\theta$ and $p(\mathrm{I}=0)=1-\theta$ from observations of the variable I. For instance, let the sequence of observations be $Z=\{1,1,1,0,0\}$. This sequence could be, for example, the output of a signal processing unit which computes in real-time the measured target range-rate divided by its range, and compares it with a certain threshold. By denoting with n the number of observed ones in Z (i.e. $n=3$ in our case) and by N the length of the sequence Z, the likelihood is:

$$P(Z|\theta)=\theta^n(1-\theta)^{N-n},$$

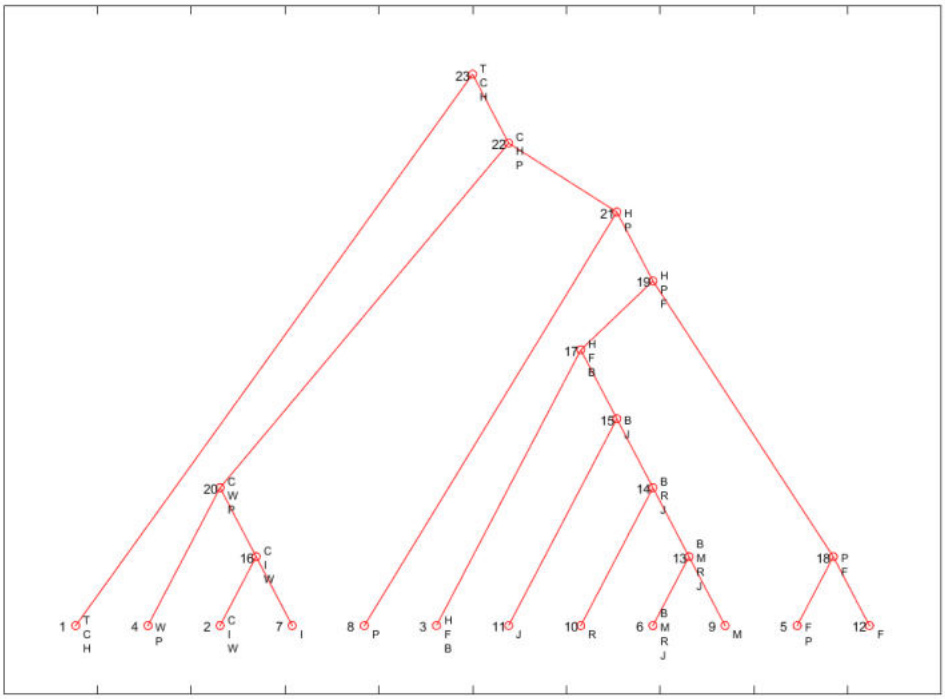

Fig. 2. BJT constructed for the valuation network in Fig. 1

assuming independence in the sequence of observations. In order to estimate the posterior distribution of θ (via Bayes' rule), we assume as prior the standard Beta distribution

$$p(\theta) = \frac{\Gamma(s)}{\Gamma(st)\Gamma(s(1-t))}\theta^{st-1}(1-\theta)^{s(1-t)-1},$$

where $s > 0$ and $t \in (0,1)$ are its parameters and $\Gamma(\cdot)$ is the Gamma function. The parameter t is the mean of θ, $E[\theta] = t$, and s is called sample-size parameter. This is the so-called mean and sample size parametrisation of the Beta distribution.

In case of lack of prior information, an issue in Bayesian analysis is how to choose these parameters to reflect this condition of prior ignorance. To address this issue, Walley in [10] proposed the so-called *Imprecise Beta model* (IBM),[1] which considers as prior the set of all possible Beta distributions:

$$p(\theta) \in \left\{ \frac{\Gamma(s)}{\Gamma(st)\Gamma(s(1-t))}\theta^{st-1}(1-\theta)^{s(1-t)-1} : t \in (0,1) \right\},$$

where s is usually selected in order to guarantee some desirable (asymptotic) properties [10], [33]. The posterior is still an IBM with

$$p(\theta|Z) \in \left\{ \frac{\Gamma(N+s)}{\Gamma(n+st)\Gamma(s(1-t)+N-n)}\theta^{n+st-1}(1-\theta)^{N-n+s(1-t)-1} : t \in (0,1) \right\}.$$

[1]The Imprecise Dirichlet model can be applied in the case of a variable with more than two possible outcomes [10], [32].

The posterior mean of θ is a probability interval [10]:

$$E[\theta|Z] \in \left[\frac{n}{N+s}, \frac{n+s}{N+s} \right]. \tag{8}$$

Wen $s = 1$, this interval agrees with those derived in [34], [35] using a different method. Going back to our initial example with $Z = \{1,1,1,0,0\}$, and adopting $s = 1$, we have that

$$E[\theta|Z] \in \left[\frac{1}{2}, \frac{4}{6} \right],$$

and $E[1-\theta|Z] \in [\frac{2}{6}, \frac{1}{2}]$. Parameter s controls the extent of the interval.

V. Ongoing threat assessment results

The developed credal and probabilistic valuation networks received the sequence of (uncertain) input information presented in Table V, with the value of $\delta = 0.001$. In the absence of any input information, $\varphi_7, \ldots, \varphi_{11}$ are set to the uniform distribution. Furthermore, confidence for φ_1 and φ_2 was set to 1 (precise value) and $\alpha_1 \in [1-\Delta, 1]$ with $\Delta = 0.005$ (interval value). Confidence for φ_3 and φ_6 was set to 0.95 (precise value) and $[0.95-\Delta, 0.95+\Delta]$ (interval value). Finally, the PMF for CPT representing φ_4 was set to $p_4 = [0.01, 0.99, 0, 1, 0, 1, 1, 0, 0.8, 0.2, 0.99, 0.01]/6$ in the probabilistic VN. For the corresponding credal VN, the

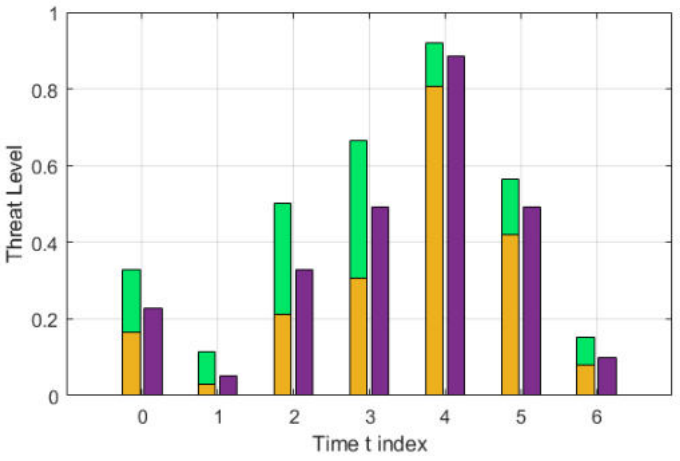

Fig. 3. A bar plot of threat level over time: (a) the probabilistic VN with precise PMFs (pink colour); (b) credal VN using probability intervals (yellow bars represent lower probabilities, green bars represent upper probabilities.).

lower and upper probability limits are specified as $[p_4 - \delta, p_4 + \delta]$. The PMF for CPT representing φ_5 was set to $p_5 = [0, 1, 1, 0, 1, 0, 0.05, 0.95, 0.05, 0.95, 1, 0]/6$ in the probabilistic VN. For the corresponding credal VN, the lower and upper probability limits are specified as $[p_5 - \delta, p_5 + \delta]$.

Elicitation of φ_7 was explained in Sec. IV. The rate of input sequence of obseravations $Z = \{z_1, \ldots, z_L\}$ is much higher than the rate of updates of the valuation network. Hence, we consider the input sequence in a sliding window of length 9, moving by 5 steps ahead. Thus at t_1 we use subsequence $\{z_1, \ldots, z_5, \ldots, z_9\}$ in computation of the probability interval (8), while at t_2 we use $\{z_6, \ldots, z_{10}, \ldots, z_{14}\}$, etc. The actual input sequence of observations is: $Z = \{0, 0, 0, 0, 0, 0, 0, 0, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 0, 0, 0, 0, 0, 0, 0, 0, 1\}$. Thus we obtain the values for φ_7 in Table V. Both probability (4th column) and probability interval (5th column) of φ_7 were calculated using (8); the only difference is that for probability we use $s = 0$, while probability interval, $s = 0.2$. The choice of a value for s can be based on asymptotic properties of the resulting estimator as discussed in [33].

TABLE V
Sequence of input data for threat assessment

time	valuation	info	prob.	prob. interval
t_1	φ_7	I=1	0.11	$[0.11, 0.13]$
t_2	φ_9	M=1	0.99	$[0.99 - \delta, 0.99 + \delta]$
	φ_7	I=1	0.67	$[0.65, 0.67]$
t_3	φ_7	I=1	1.00	$[0.98, 1.00]$
t_4	φ_8	P=1	0.9	$[0.9 - \delta, 0.9 + \delta]$
t_5	φ_7	I=1	0.56	$[0.54, 0.57]$
t_6	φ_7	I=1	0.11	$[0.11, 0.13]$

Fig. 3 shows the results produced by the two valuation networks: the credal VN for probability intervals and the probabilistic VN for precise PMFs. The ordinate in the graph of Fig. 3 represents the probability that T= 1, and can be interpreted as the level of threat. The yellow and green bars in Fig. 3 represent the output of the credal VN, indicating the lower probability (yellow) and the probability interval (green) of threat. This interval is a consequence of epistemic uncertainty in modeling domain knowledge and the input data. The output of the probabilistic VN is shown as a precise probability with pink bars. We observe the following. The threat probability intervals always contain the true probability values. Initially, at time $t_0 = 0$ (before any data is received), the threat level is low, but non-zero, because of the "unknown" factor. At t_1, when according to Table V, the imminence is low, although we know nothing about the hostile intent, the probability of threat falls bellow 0.1. However, at t_2, t_3 and t_4, it grows because the input information indicates that the target has a hostile intent and that it carries (lethal) weapons. Then at t_5 and t_6, as a result of a decrease in the probability of imminence, the probability of threat reduces to lower levels.

VI. Summary

The paper developed and demonstrated a credal valuation network for ongoing assessment of threat under epistemic uncertainty. The threat model was developed for a single target using the classical capability-intent paradigm in the context of air surveillance. The model is fairly general and can be easily adapted to other similar applications. Numerical results indicate reliable performance in the presence of epistemic uncertainty, characterising both the knowledge-base and input observations. Further work will develop a threat model for more realistic situations with multiple appearing/disappearing targets.

References

[1] C. E. Shannon, "A mathematical theory of communication," *Bell system technical journal*, vol. 27, no. 3, pp. 379–423, 1948.

[2] B. Ristic, C. Gilliam, M. Byrne, and A. Benavoli, "A tutorial on uncertainty modeling for machine reasoning," *Information Fusion*, vol. 55, pp. 30–44, 2020.

[3] J. Kohlas, *Information algebras: Generic structures for inference*, Springer-Verlag, London, 2003.

[4] P. P. Shenoy, "Using possibility theory in expert systems," *Fuzzy Sets and Systems*, vol. 52, no. 2, pp. 129 – 142, 1992.

[5] B. Ristic, A. Benavoli, and S. Arulampalam, "Credal valuation networks for machine reasoning under uncertainty," *IEEE Transactions on Artificial Intelligence*, vol. 5, no. 1, pp. 51–60, 2024.

[6] R. G. Almond, *Graphical Belief Modeling*, Chapman and Hall, 1995.

[7] R. Haenni, "Ordered valuation algebras:A generic framework for approximate inference," *Int. Journal of Approximate Reasoning*, vol. 37, pp. 1–41, 2004.

[8] A. Benavoli, B. Ristic, A. Farina, M. Oxenham, and L. Chisci, "An application of evidential networks to threat assessment," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 45, no. 2, pp. 620–639, April 2009.

[9] P. Kowalski, M. Zocholl, and A.-L. Jousselme, "Explaining the impact of source behaviour in evidential reasoning," *Information Fusion*, vol. 81, pp. 41–58, 2022.

[10] P. Walley, *Statistical reasoning with imprecise probabilities*, Chapman and Hall, 1991.

[11] A. Antonucci, D. Huber, M. Zaffalon, P. Luginbühl, I. Chapman, and R. Ladouceur, "CREDO: A military decision-support system based on credal networks," in *Proc. of the 16th Intern. Conf. on Information Fusion*. IEEE, 2013, pp. 1942–1949.

[12] P. P. Shenoy, "A valuation-based language for expert systems," *Int. J. Approx. Reason.*, vol. 3, no. 5, pp. 383 – 411, 1989.

[13] P. P. Shenoy and G. Shafer, "Axioms for probability and belief-function propagation," in *Readings in uncertain reasoning*, J. Pearl G. Shafer, Ed., pp. 575–610. San Mateo, CA: Morgan Kaufmann, 1990.

[14] P. P. Shenoy, "Valuation based systems: A framework for managing uncertainty in expert systems," in *Fuzzy Logic and the Management of Uncertainty*, L. A. Zadeh and J. Kacprzyk, Eds., chapter 4, pp. 83–104. Wiley, New York, 1992.

[15] R. Haenni, "Ordered valuation algebras: a generic framework for approximate inference," *Int. Journal of Approximate Reasoning*, vol. 37, pp. 1–41, 2004.

[16] A. Benavoli and B. Ristic, "Evidential networks for decision support in surveillance systems," in *Integrated tracking, classification, and sensor management*, M. Mallick, V. Krishnamurthy, and B.-N Vo, Eds., chapter 17, pp. 661–704. Wiley, 2013.

[17] T. Augustin, F. Coolen, G. de Cooman, and M. Troffaes, Eds., *Introduction to imprecise probabilities*, Wiley, 2014.

[18] 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, no. 02, pp. 167–196, 1994.

[19] A. Antonucci, C. P. De Campos, and M. Zaffalon, "Probabilistic graphical models," in *Introduction to imprecise probabilities*, T. Augustin *et al*, Ed. John Wiley & Sons, 2014.

[20] E. Waltz and J. Llinas, *Multisensor data fusion*, vol. 685, Artech house, Boston, 1990.

[21] J. D. Singer, "Threat-perception and the armament-tension dilemma," *Journal of Conflict Resolution*, vol. 2, no. 1, pp. 90–105, 1958.

[22] N. Okello and G. Thoms, "Threat assessment using Bayesian networks," in *Proceedings of the 6th International Conference on Information fusion*, 2003, pp. 1102–1109.

[23] S. Kumar and B. K. Tripathi, "Modelling of threat evaluation for dynamic targets using Bayesian network approach," *Procedia technology*, vol. 24, pp. 1268–1275, 2016.

[24] Z.-H. Fan, B.-H. Shi, J.-Y. Chen, and T.-L. Duan, "A novel dynamic Bayesian network based threat assessment algorithm," in *2017 4th International Conference on Systems and Informatics (ICSAI)*, 2017, pp. 611–615.

[25] H. Yao, H. Wang, Y. Li, Y. Wang, and C. Han, "Research on unmanned underwater vehicle threat assessment," *IEEE Access*, vol. 7, pp. 11387–11396, 2019.

[26] L. Hammond, "An evidential network approach applied to threat evaluation in above water warfare," Tech. Rep. DST Group-TR-3349, Defence Science and Technology Group, 2017.

[27] J.O. Berger, E. Moreno, L.R. Pericchi, et al., "An overview of robust Bayesian analysis," *Test*, vol. 3, no. 1, pp. 5–124, 1994.

[28] T. Augustin and R. Hable, "On the impact of robust statistics on imprecise probability models: a review," *Structural Safety*, vol. 32, no. 6, pp. 358–365, 2010.

[29] I. Montes, E. Miranda, and S. Destercke, "Unifying neighbourhood and distortion models: Part I – new results on old models," *Intern. Journal of General Systems*, vol. 49, no. 6, pp. 602–635, 2020.

[30] S. Destercke and D. Dubois, "The role of generalised p-boxes in imprecise probability models," in *Proc. of the 6th Int. Symp. on Imprecise Probability: Theories and Applications*, 2009, pp. 179–188.

[31] I. Montes, E. Miranda, and S. Destercke, "Unifying neighbourhood and distortion models: Part II – new models and synthesis," *Intern. Journal of General Systems*, vol. 49, no. 6, pp. 636–674, 2020.

[32] J.-M. Bernard, "An introduction to the imprecise Dirichlet model for multinomial data," *Intern. Journal of Approximate Reasoning*, vol. 39, no. 2-3, pp. 123–150, 2005.

[33] C. P. de Campos and A. Benavoli, "Inference with multinomial data: Why to weaken the prior strength," in *Twenty-Second International Joint Conference on Artificial Intelligence*, 2011.

[34] Arthur P Dempster, "New methods for reasoning towards posterior distributions based on sample data," *The Annals of Mathematical Statistics*, pp. 355–374, 1966.

[35] P Smets, "Belief induced by the knowledge of the probabilities," *Technical Report 94-4. I.*, 1994.