Credal Networks for Military Identification Problems [⋆]

[a]*IDSIA, Galleria 2, CH-6928 Manno (Lugano), Switzerland*

[b]*Armasuisse (W+T), Feuerwerkerstrasse 39, CH-3600 Thun, Switzerland*

Alessandro Antonucci [a], Ralph Brühlmann [b], Alberto Piatti [a,∗], Marco Zaffalon [a]

Abstract

Credal networks are imprecise probabilistic graphical models generalizing Bayesian networks to convex sets of probability mass functions. This makes credal networks particularly suited to model expert knowledge under very general conditions, including states of qualitative and incomplete knowledge. In this paper, we present a credal network for risk evaluation in case of intrusion of civil aircrafts into a restricted flight area. The different factors relevant for this evaluation, together with an independence structure over them, are initially identified. These factors are observed by sensors, whose reliabilities can be affected by variable external factors, and even by the behaviour of the intruder. A model of these observation processes, and the necessary fusion scheme for the information returned by the sensors measuring the same factor, are both completely embedded into the structure of the credal network. A pool of experts, facilitated in their task by specific techniques to convert qualitative judgements into imprecise probabilistic assessments, has made possible the quantification of the network. We show the capabilities of the proposed model by means of some preliminary tests referred to simulated scenarios. Overall, we can regard this application as a useful tool to support military experts in their decision, but also as a quite general imprecise-probability paradigm for information fusion.

Keywords: Credal Networks, Information Fusion, Sensor Management, Tracking Systems.

⋆ This research was supported by Armasuisse, and partially by the Swiss NSF grants n. 200020-116674/1 and n. 200020-121785/1.

∗ Corresponding author.

Email addresses: `alessandro@idsia.ch` (Alessandro Antonucci), `ralph.bruehlmann@ar.admin.ch` (Ralph Brühlmann), `alberto.piatti@idsia.ch` (Alberto Piatti), `zaffalon@idsia.ch` (Marco Zaffalon).

1 Introduction

In the recent times, the establishment of a restricted or prohibited flight area around important potential targets surveyed by the Armed Forces has become usual practice, also in neutral states like Switzerland, because of the potential danger of terror threats coming from the sky. A prohibited flight area is an airspace of definite dimensions within which the flight of aircrafts is prohibited. A restricted flight area is an airspace of definite dimensions within which the flight of aircrafts is restricted in accordance with certain specified conditions [17].

Once a restricted flight area is established for the protection of a single strategic object, all the aircrafts flying in this region without the required permissions are considered *intruders*. The restricted flight area can be imagined as divided in two concentric regions: an external area, devoted to the identification of the intruder, where the intruder is observed by many sensors of the civil and military Air Traffic Control and by the interceptors, and an internal area, which is a small region containing the object to protect and the military units, where fire is eventually released if the intruder is presumed to have bad aims.

Clearly, not all the intruders have the same intentions: there are intruders with bad aims, called *renegades*, intruders with provocative aims, erroneous intruders, and even aircrafts that are incurring an emergency situation. Since only renegades represent a danger for the protected object, the recognition of the intruder's aim plays a crucial role in the following decision. This is the identification problem we address in this paper.

The problem is complex for many reasons: (i) the risk evaluation usually relies on qualitative expert judgements; (ii) it requires the fusion of information coming from different sensors, and this information can be incomplete or partially contradictory; (iii) different sensors can have different levels of reliability, and the reliability of each sensor can be affected by exogenous factors, as geographical and meteorological conditions, and also by the behaviour of the intruder. A short review of the problem and some details about these difficulties are reported in Sect. 2.

In this paper, we propose *credal networks* [7] (Sect. 3) as a mathematical paradigm for the modeling of military identification problems. Credal networks are imprecise probabilistic graphical models representing expert knowledge by means of sets of probability mass functions associated to the nodes of a directed acyclic graph. These models are particularly suited for modeling and doing inference with qualitative, incomplete, and also conflicting information. All these features appear particularly important for the military problem under consideration.

More specifically, we have developed a credal network that evaluates the level of risk associated to an intrusion. This is achieved by a number of sequential steps: determination of the factors relevant for the risk evaluation and identification of a dependency structure between them (Sect. 4.1); quantification of this qualitative structure by imprecise probabilistic assessments (Sect. 5.1); determination of a qualitative model of the observation process associated to each sensor, together with the necessary *fusion scheme* of the information collected by the different sensors (Sect. 4.2); quantification of this model by probability intervals (Sect. 5.2). An analysis of the main features of our imprecise-probability approach to information fusion is indeed reported in Sect. 6.

The credal network is finally employed to evaluate the level of risk, which is simply the probability of the risk factor conditional on the information collected by the sensors in the considered scenario. A description of the procedure used to update the network, together with the results of some simulations, is reported in Sect. 7.

Summarizing, we can regard this model as a practical tool to support military experts in their decisions for this particular problem. [1] But, at the same time, our credal network can be regarded as a prototypical modeling framework for general identification problems requiring information fusion.

2 Military Aspects

This section focuses on the main military aspects of the identification problem addressed by this paper. Let us first report the four possible values of the RISK FACTOR [2] by which we model the possible intentions of the intruder.

(i) `Renegade`. The intruder intends to use his aircraft as a weapon to damage the strategic target defended by the restricted flight area. [3]
(ii) `Agent provocateur`. The aim is to provoke or demonstrate. The intruder knows exactly what he is doing and does not want to die, therefore he is expected to react positively to radio communication at a certain moment.
(iii) `Erroneous`. The intruder is entering the restricted flight area because of an error in the flight path due to bad preparation of the flight or insufficient training level.
(iv) `Damaged`. This is an intruder without bad aims that is incurring an emergency situation due to a technical problem. The pilot does not necessarily

[1] The support we provide is represented by the probabilistic information about the actual level of risk associated to an intrusion. Decisions about possible interventions can be based on this information, but are still taken by military experts. A model of such decision process, to be embedded into the network structure, could be explored (e.g., by considering the ideas in [1] and their development in [10]), but is beyond the scope of this paper.
[2] The following typographical convention is used: the variables considered in our probabilistic model are written in SMALL CAPITALS and their possible states in `typewriter`.
[3] There are also some subcategories of terrorists (e.g., poison sprayers), which will be considered only in a future work.

knows what he is doing because of a possible situation of panic. A damaged intruder can react negatively to radio communications, as their instruments could be switched off because of electrical failures. A proper identification of damaged intruders is very important because they can be easily confused with renegades.

In order to decide which one among these four categories reflects the real aim of the intruder an appropriate identification architecture should be set up. Fig. 1 displays the structure typically employed in Switzerland. When a restricted flight area is set up for the protection of an important object, the *Air Defence Direction Center* (ADDC) is in charge of the identification of possible intruders. The ADDC collects the information provided by three main sources: (i) the sensors of the civil *Air Traffic Control* (ATC), (ii) the sensors of the military ATC, (iii) the interceptors of the Swiss Air Force devoted to Air Police missions. Once this evidential information has been collected, the ADDC performs the identification of the aim of the intruder.

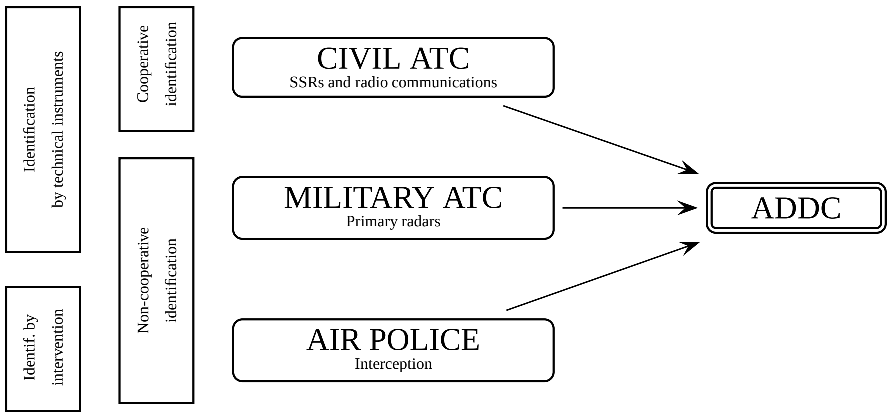

Figure 1. The identification architecture.

The civil ATC sensors are based on a collaborative communication between the ATC and the intruder. In fact, the detection of the intruder by the ATC is possible only if the intruder is equipped and operates with a *transponder*. Transponders are electronic devices that, if interrogated by the civil ATC radar, emit a signal enabling a three-dimensional localization. Radars based on this principle are called *Secondary Surveillance Radars* (SSRs). Transponders emit also an identification code. We consider the identification code *Mode 3/A*, which, in certain cases, does not allow the exact identification of the intruder to be realized (e.g., all the aircrafts flying according to the visual flight rules emit the same code). [4] It should be also pointed out that, if the transponder is switched off, the intruder remains invisible to the civil ATC because the SSR is unable to detect it.

[4] The most informative identification code *Mode S* is not considered here, because it has not yet been implemented extensively in practice.

Overall, we summarize the information relevant for the identification gathered by the civil ATC in terms of two distinct factors: the TRANSPONDER MODE 3/A, indicating if and possibly what type of identification code has been detected by the SSR; and the ATC REACTION, describing the response of the intruder to the instructions that civil ATCs report to aircrafts flying in the direction of the restricted area in order to deviate them from their current flight route.

Unlike civil ATC sensors, sensors managed by the military ATC and military Air Police units are based on a non-collaborative observation of the intruder. The main military ATC sensors are radar stations detecting the echoes reflected by the intruder of radio pulses emitted by the radar. These radars provide a continuous three-dimensional localization of the intruder. The other military sensors, which are particularly suited for the identification of intruders flying at relatively low height, are the pointing devices of anti-air firing units (two-dimensional and tracking radars, cameras) and the *Ground Observer Corps* (GOCs), which are military units equipped with optical instruments to observe the intruder from the ground.

The information gathered by these sensors which is relevant for the identification of the intentions of the intruder can be summarized by the following factors: AIRCRAFT HEIGHT, HEIGHT CHANGES, ABSOLUTE SPEED, FLIGHT PATH, AIRCRAFT TYPE, and also REACTION TO ADDC, which is the analogous of REACTION TO ATC, but referred to the case of detection by the military ATC.

Finally, regarding the information gathered by the interceptors of the Air Force, which is reported to the ADDC, the possible identification missions of the interceptors are divided into three categories according to the International Civil Aviation Organization: *surveillance*, *identification* and *intervention*. In the first type of mission, the interceptor does not establish a visual contact with the intruder but observes its behaviour using sensors; the interceptor is therefore considered as a sensor observing the same factors as the other sensors of the civil and military ATC. In the second and in the third type of missions, the interceptor establishes a visual contact with the intruder with the intention of observing it (identification), or giving it instructions in order to deviate the aircraft from the current flight route, or also to land it (intervention). The reaction of the intruder to interception is very informative about its intentions. We model this reaction to the latter two types of mission by the factor REACTION TO INTERCEPTION.

The intruder is assumed to be observed during a sufficiently long time window called observation period. All the factors we have defined in order to describe the behaviour of the intruder during this observation period are discrete variables, whose (mutually exclusive and exhaustive) possible values are defined with respect to the dynamic component of the identification process, in order to eliminate the dependency of the model on local issues. To explain how these aspects are taken into account, we detail the definitions of the factors FLIGHT PATH and HEIGHT CHANGES. The first factor describes the route followed by the intruder during the observation period from an intentional point of view and not from a physical or geographical perspective. Accordingly, their possible values are defined as follows.

(i) `Suspicious route`. The intruder follows a suspicious flight route in the direction of the protected objects.
(ii) `Provocative route`. The intruder flights in the restricted area without approaching significantly the protected objects in an apparently planned way.
(iii) `Positive reaction route`. The intruder corrects its flying route according to the instructions of the ATC or of the interceptors or spontaneously.
(iv) `Chaotic route`. The intruder follows an apparently chaotic flight path.

The definition of these possible states is independent of the specific geographical situation. In practice, we assume that a route observed on the radar or on other sensors by the ADDC is interpreted from an intentional point. Similarly, the factor HEIGHT CHANGES describes the behaviour of the intruder with respect to its height, by the following possible values.

(i) `Climb`. The intruder is climbing, i.e., increasing its height.
(ii) `Descent`. The intruder is descending. i.e., decreasing its height.
(iii) `Stationary`. The intruder maintains roughly the same height.
(iv) `Unstable`. The intruder climbs and descends in an alternate way.

These values reflect an observation of the dynamic behaviour of the intruder during the observation period.

Another important issue of our model is the distinction between the number of sensors available in the identification architecture and the evaluation of their efficiency, for a characterization of the quality of the observations. Note, for instance, that the quality of the observation of the FLIGHT PATH provided by a low number of GOCs would be scarce, exactly as the observation provided by a high number of GOCs, working under bad meteorological conditions.

By this example, we intend to point out that a proper description of the identification architecture can be obtained by distinguishing between the presence and the reliability of each sensor. The presence depends on the specific identification architecture, on the technical limits of the sensors, and also on the behaviour of the intruder itself, being in particular affected by the AIRCRAFT HEIGHT (e.g., some sensors can observe the intruder only if it is flying at low heights). The reliability depends on the meteorological and geographical conditions, on specific technical limits of the sensors (e.g., radars have low quality in the identification of the AIRCRAFT TYPE, independently of its presence) and on the AIRCRAFT HEIGHT. All these aspects are implicitly considered during the specification of the presence and the reliability of the different sensors. This model of the identification architecture is detailed in Sect. 4 and Sect. 5.

3 Mathematical Aspects

In this section, we briefly recall the definitions of *credal set* and *credal network* [7], which are the mathematical objects we use to model expert knowledge and fuse different kinds of information in a single coherent framework.

3.1 Credal Sets

Given a variable X, we denote by Ω_X the possibility space of X, with x a generic element of Ω_X. Denote by $P(X)$ a mass function for X and by $P(x)$ the probability of x.

We denote by $K(X)$ a closed convex set of probability mass functions over X. $K(X)$ is said to be a *credal set* over X. For any $x \in \Omega_X$, the lower probability for x according to the credal set $K(X)$ is $\underline{P}(x) = \min_{P(X)\in K(X)} P(x)$. Similar definitions can be provided for upper probabilities, and more generally lower and upper expectations. A set of mass functions, its convex hull, and its set of *vertices* (i.e., extreme points) produce the same lower and upper expectations and probabilities. Accordingly, a credal set can be defined by an explicit enumeration of its vertices.

A set of *probability intervals* over Ω_X, say $\mathbb{I}_X = \{\mathbb{I}_x : \mathbb{I}_x = [l_x, u_x], 0 \leq l_x \leq u_x \leq 1, x \in \Omega_X\}$ induces the specification of a credal set $K(X) = \{P(X) : P(x) \in \mathbb{I}_x, x \in \Omega_X, \sum_{x\in\Omega_X} P(x) = 1\}$. $\mathbb{I}_X$ is said to *avoid sure loss* if the corresponding credal set is not empty, and to be *reachable* (or *coherent*) if $u_{x'} + \sum_{x\in\Omega_X, x\neq x'} l_x \leq 1 \leq l_{x'} + \sum_{x\in\Omega_X, x\neq x'} u_x$, for all $x' \in \Omega_X$. $\mathbb{I}_X$ is reachable if and only if the intervals are tight, i.e., for each lower or upper bound in $\mathbb{I}_X$ there is a mass function in the credal set at which the bound is attained [19]. A non-reachable set of probability intervals avoiding sure loss can be always refined in order to become reachable [5]. The vertices of a credal set defined by a reachable set of probability intervals can be efficiently computed using standard *reverse search enumeration* techniques [3,5].

Here, we focus on credal sets defined through reachable probability intervals, as they appear as a very natural way to capture the kind of human expertize we want to reproduce by our model (see Sect. 5). Nevertheless, apart from the case of binary variables, it is possible to consider credal sets that cannot be obtained by probability intervals.

Dealing with credal sets instead of single probability mass functions needs also a more general concept of independence. The most commonly adopted concept is *strong independence*. Two generic variables X and Y are strongly independent when every vertex of the underlying credal set $K(X, Y)$ satisfies standard stochastic independence of X and Y. Finally, regarding conditioning with credal sets, we can compute the posterior credal set $K(X|Y = y)$ as the union, obtained by element-wise application of Bayes' rule, of all the posterior mass functions $P(X|Y=y)$ (the lower probability of the conditioning event is assumed positive). [5]

3.2 Credal Networks

Let $\mathbf{X}$ be a vector of variables and assume a one-to-one correspondence between the elements of $\mathbf{X}$ and the nodes of a *directed acyclic graph* $\mathcal{G}$. Accordingly, in the following we will use *node* and *variable* interchangeably. For each $X_i \in \mathbf{X}$, Π_i denotes the set of the *parents* of X, i.e., the variables corresponding to the immediate predecessors of X according to $\mathcal{G}$.

The specification of a *credal network* over $\mathbf{X}$, given the graph $\mathcal{G}$, consists in the assessment of a conditional credal set $K(X_i|\pi_i)$ for each possible value $\pi_i \in \Omega_{\Pi_i}$ of the parents of X_i, for each variable $X_i \in \mathbf{X}$. [6]

The graph $\mathcal{G}$ is assumed to code strong independencies among the variables in $\mathbf{X}$ by the so-called strong *Markov condition*: every variable is strongly independent of its nondescendant non-parents given its parents. Accordingly, it is therefore possible to regard a credal network as a specification of a credal set $K(\mathbf{X})$ over the joint variable $\mathbf{X}$, with $K(\mathbf{X})$ the convex hull of the set of joint mass functions $P(\mathbf{X}) = P(X_1, ..., X_n)$ over the n variables of the net, that factorize according to $P(x_1, \ldots, x_n) = \prod_{i=1}^{n} P(x_i|\pi_i)$. Here π_i is the assignment to the parents of X_i consistent with $(x_1, \ldots, x_n)$; and the conditional mass functions $P(X_i|\pi_i)$ are chosen in all the possible ways from the extreme points of the respective credal sets. $K(\mathbf{X})$ is called the *strong extension* of the credal network. Observe that the vertices of $K(\mathbf{X})$ are joint mass functions $P(\mathbf{X})$. Each of them can be identified with a *Bayesian network* [13], which is a precise probabilistic graphical model. In other words, a credal network is equivalent to a set of Bayesian networks.

3.3 Computing with Credal Networks

Credal networks can be naturally regarded as expert systems. We query a credal network to gather probabilistic information about a variable given evidence about some other variables. This task is called *updating* and consists in the computa-

[5] If only the upper probability is positive, conditioning can be still obtained by *regular extension* [20, App. J]. The stronger condition assumed here is required by the inference algorithms we adopt, and it is always satisfied in our tests (see Section 7).
[6] This definition assumes the credal sets in the network to be *separately specified*, i.e., selecting a mass function from a credal set does not influence the possible choices in others. Other possible specifications can be considered. We point the reader to [1] for a general unified graphical language that allows for these specifications.

tion, with respect to the network strong extension $K(\mathbf{X})$, of $\underline{P}(X_q|X_E = x_E)$ and $\overline{P}(X_q|X_E = x_e)$, where X_E is the vector of variables of the network in a known state x_E (the evidence), and X_q is the node we query. Credal network updating is an NP-hard task (also for polytrees) [9], for which a number of exact and approximate algorithms have been proposed (e.g., [8] for an overview).

4 Qualitative Assessment of the Network

We are now in the condition to describe the credal network developed for our application. According to the discussion in the previous section, this task first requires the qualitative identification of the conditional dependencies between the variables involved in the model, which can be coded by a corresponding directed graph.

As detailed in Sect. 2, the variables we consider in our approach are: (i) the RISK FACTOR, (ii) the nine variables used to assess the intention of the intruder, (iii) the variables representing the observations returned by the sensors, and (iv) for each observation two additional variables representing the level of PRESENCE and RELIABILITY of the relative sensor. In the following, we refer to the variables in the categories (i) and (ii) as *core variables*.

4.1 Risk Evaluation

Fig. 2 depicts the conditional dependencies between the core variables according to the military and technical considerations of the Expert. [7] The specification of this part of the network has required a considerable amount of military and technical expertise that, due to confidentiality reasons, cannot be explained in more detail here.

4.2 Observation and Fusion Mechanism

In this paper, we follow the general definition of *latent* and *manifest variables* given by [18]: a *latent variable* is a variable whose realizations are unobservable (hidden), while a *manifest variable* is a variable whose realizations can be directly observed. According to [4], there may be different interpretations of latent variables. In our model, we consider a latent variable as an unobservable variable that exists independent of the observation. The *core variables* in Fig. 2 are regarded as latent variables

[7] In this paper we briefly call *Expert* a pool of military experts, we have consulted during the development of the model.

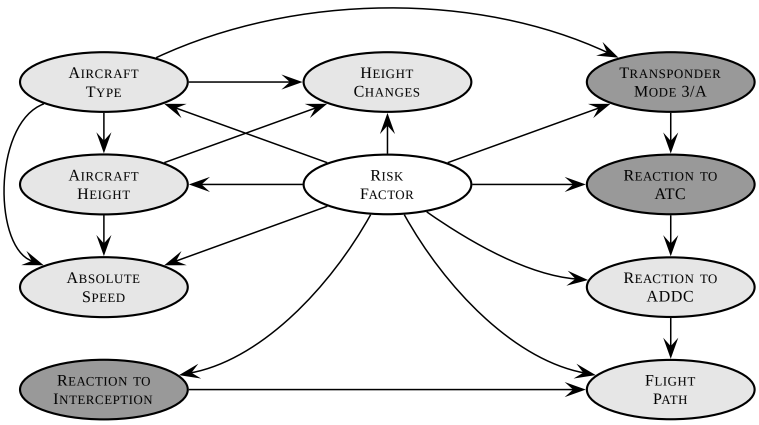

Figure 2. The core of the network. Dark gray nodes are observed by single sensors, while light gray nodes are observed by sets of sensors for which an information fusion scheme (see Sect. 4.2) is required.

that, to be determined, usually require the fusion of information coming from different sensors, with different levels of reliability. The observations of the different sensors are considered manifest variables.[8] Nevertheless, in the case of the identification code emitted by the intruder (TRANSPONDER MODE 3/A), the REACTION TO INTERCEPTION observed by the pilot, and the REACTION TO ATC observed by the controllers through SSR, the observation process is immediate; thus we simply identify the latent with the corresponding manifest variable. Clearly, if the RISK FACTOR was the only latent variable, the network in Fig. 2 would be the complete network needed to model the risk evaluation. But, because we are dealing with latent variables observed by many sensors, a model of the observation process and a fusion mechanism have to be added to the current structure.

Observation Mechanism We begin by considering observations by single sensors, and then we explain the fusion scheme for several sensors. Consider the following example: suppose that an intruder is flying at low height and is observed by ground-based observation units in order to evaluate its FLIGHT PATH. For this evaluation, the intruder should be observed by many units. If our identification architecture is characterized by too a low number of observation units, it is probable that the observation would be incomplete or even absent, although the meteorological and geographical conditions are optimal. In this case, the poor quality of the observation is due to the scarce presence of the sensor. Suppose now that the architecture is characterized by a very large number of observation units but the weather is characterized by a complete cloud cover with low clouds, then the quality of the observation is very poor although the presence of units is optimal. In this case the poor quality of the observation is due to the scarce reliability of the sensor under this meteorological condition.

[8] The manifest variables we consider are therefore referred to the observations of corresponding latent variables. Thus, if X is a latent variable, the possibility space Ω_O of the corresponding manifest variable O takes values in the set Ω_X augmented by the supplementary possible value `missing` (we denote this value by '$*$').

This example motivates our choice to define two different factors in order to model the quality of an observation by a sensor. Fig. 3 illustrates, in general, how the evidence provided by a sensor about a latent variable is assessed. The manifest variable depends on the relative latent variable, on the PRESENCE of the sensor (with possible values `present`, `partially present` and `absent`), and its RELIABILITY (with possible values `high`, `medium` and `scarce`).

According to the military principles outlined in Sect. 2, the RELIABILITY of a sensor can be affected by the meteorological and geographical situation and also by the AIRCRAFT HEIGHT, while, regarding the PRESENCE, only the AIRCRAFT HEIGHT and the identification architecture affect the quality of the observations. The influence of the latent variable AIRCRAFT HEIGHT is related to the technical limits of the sensors: there are sensors that are specific of the low and very low heights, like tracking radars and cameras; other sensors, like the primary surveillance radars, are always present at high and very high heights, but are not always present at low and very low heights.

Meteorological and geographical conditions do not affect the PRESENCE of a sensor, but only its RELIABILITY. It is worthy to point out that these exogenous factors are always observed and we will not display them explicitly as network variables, being considered by the Expert during his quantification of the RELIABILITY. [9]

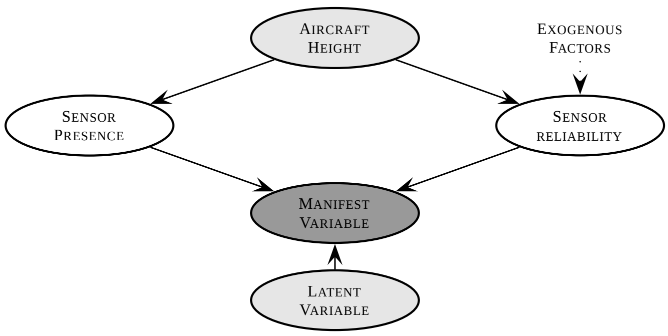

Figure 3. Observation by a single sensor. The *latent variable* is the variable to be observed by the sensor, while the *manifest variable* is the value returned by the sensor itself.

Sensors Fusion At this point we can explain how the information collected by the different observations of a single latent variable returned by different sensors can be fused together. Consider, for example, the determination of the latent variable AIRCRAFT TYPE. This variable can be observed by four types of sensors: TV cameras, IR cameras, ground-based observation units and air-based interceptors.

[9] As noted in Section 5.2, the Expert is not required to quantify these nodes for each possible configuration of the exogenous factors, but only for the specific conditions observed at the moment of the quantification. For this reason, these factors are not included among the variables of the model.

For each sensor, we model the observation using a structure like the network in Fig. 3: there is a node representing the PRESENCE of the sensor and a node representing the RELIABILITY, while the variable AIRCRAFT HEIGHT influences all these nodes. Accordingly, for each combination of PRESENCE and RELIABILITY, a different model of the relation between the manifest and the latent variable (i.e., a model of the sensor performances) should be specified. [10] Overall, we obtain a structure like in Fig. 4, which permits the fusion of the evidence about the latent variables coming from the different sensors, taking into account the reliability of the different observations and without the need of any external specification of explicit fusion procedures. This is obtained by simply assuming the conditional independence among the different sensors given the value of the ideal variable. Sect. 6 reports a note on the main features of this approach, which has been inspired by similar techniques adopted for Bayesian networks [11].

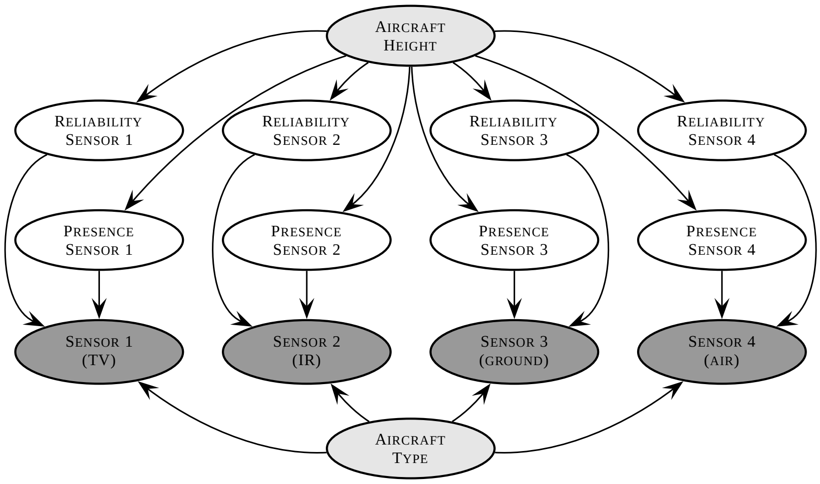

Figure 4. The determination of the latent variable *type of aircraft* by four sensors.

Whole Network The procedure considered in the previous paragraph for the node AIRCRAFT TYPE is applied to every latent variable requiring information fusion from many sensors. This practically means that we add a subnetwork similar to the one reported in Fig. 4 to each light gray node of the network core in Fig. 2. The resulting directed graph, which is still acyclic, is shown in Fig. 5. A more compact specification could be obtained by extending the formalism of *object-oriented Bayesian networks* [12] to credal networks. Accordingly, we can regard the boxed subnetworks in Fig. 5, modeling the observations of the ideal factors, as different instances of a given class, for which appropriate specifications of the attributes (possible values, number of sensors, etc.) have been provided.

[10] Here the RELIABILITY is intended as a global descriptor of the sensor performances. The quality of a particular observation can be clearly affected by the value of the ideal variable, but this is modeled in the relation between the latent and the ideal variable.

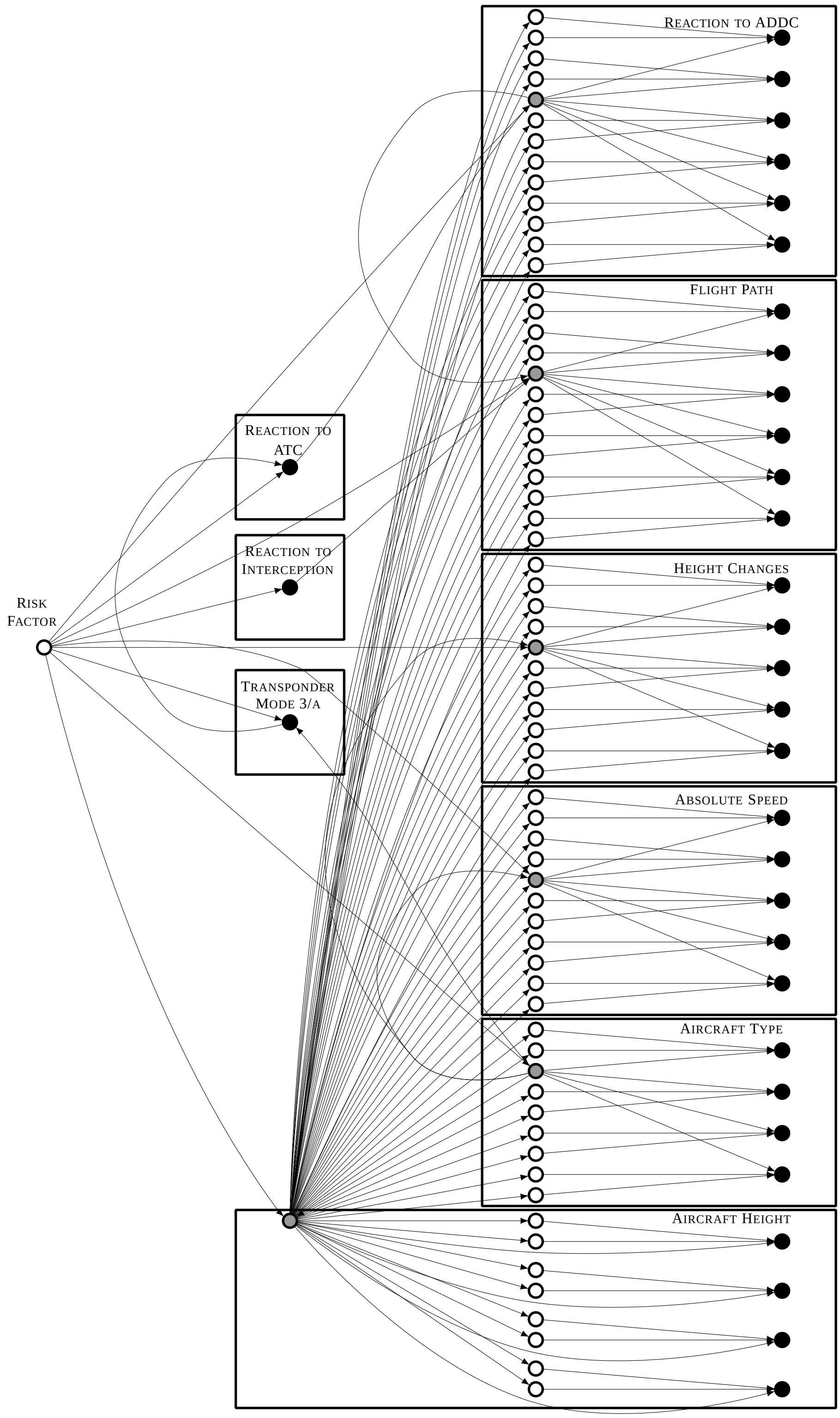

Figure 5. The complete structure of the network. Black nodes denote manifest variables observed by the sensors, latent variables corresponding to the unobserved ideal factors are gray, while presences, reliabilities, and the risk factor are white. Boxes highlight the different subnetworks modeling the observation process for the ideal factors.

5 Quantitative Assessment of the Network

As outlined in Sect. 3, the specification of a credal network over the variables associated to the directed acyclic graph in Fig. 5 requires the specification of a conditional credal set for each variable and each possible configuration of its parents. Specific procedures for the quantification of these credal sets based on Expert's qualitative judgements have been developed for the core variables (Sect. 5.1) and for the nodes modeling the observation process (Sect. 5.2).

5.1 Quantification of the Network Core

Because of the scarcity of historical cases, the quantification of the conditional credal sets for the core variables in Fig. 2 is mainly based upon military and technical considerations. The Expert provided a number of qualitative judgements like *"erroneous intruders are light aircrafts with good chance"* and *"erroneous intruders are business jets with little chance"*, later translated into the following specifications for the bounds of the probability intervals:

$$P(\texttt{light aircraft}|\texttt{erroneous}) \geq .65,$$

$$P(\texttt{business jet}|\texttt{erroneous}) \leq .20.$$

This kind of elicitation has been obtained following Walley's guidelines for the translation of natural language judgements [21, p. 48]. Clearly, there is a degree of arbitrariness in choosing single numbers for the bounds of the probability intervals, but much less than in similar approaches based on precise probabilities.

In some situations, the Expert was also able to identify logical constraint among the variables. As an example, the fact that *"balloons cannot maintain high levels of height"* represents a constraint between the possible values of the variables AIRCRAFT TYPE and AIRCRAFT HEIGHT, that can be embedded into the structure of the network by means of the following zero probability assessment:

$$P(\texttt{very high}|\texttt{balloon}) = 0.$$

Overall, the conditional credal sets corresponding to elicited probability intervals have been computed according to the procedure outlined in Sect. 3.1 and a well-defined credal network over the graphical structure in Fig. 2 has been concluded.

To complete the quantification of the credal network over the whole graphical structure in Fig. 5, we should discuss, for each sensor, the quantification of the variables modeling the observation process.

We begin by explaining how PRESENCE and RELIABILITY are specified. Consider the network in Fig. 3. The Expert should quantify, for each of the four possible values of AIRCRAFT HEIGHT, two credal sets, one for the PRESENCE and one for the RELIABILITY. For the first, he takes into consideration only the structure of the identification architecture; while for the second, also the actual meteorological and geographical situation should be considered.

In principle, this quantification task would require that the Expert answer questions like, *"what is the probability (interval) that the ground-based observers have scarce (or medium, or high) reliability in observing an aircraft flying at low height, if the meteorological condition is characterized by dense low clouds and we are in the plateau?"*. Clearly, it can be extremely difficult and time-consuming to answer dozens of questions of this kind in a coherent and realistic way. For this reason, we simply ask the Expert to provide *characteristic levels* of PRESENCE and RELIABILITY. That can be obtained by questions like the following, *"what is the reliability level that you expect from ground-based observations of an aircraft flying at low height, if the meteorological condition is characterized by dense low clouds and we are in the plateau?"*. The latter question is much simpler, because one is required to specify something more qualitative than probabilities. Together with the characteristic levels, the Expert also indicates whether or not he is uncertain about these values. Finally, for each combination of expected levels and relative uncertainty, a fixed credal set is defined together with the Expert. That substantially simplifies the quantification task, while maintaining a large flexibility in the specification of the model. As an example, assuming that Expert's expected level of reliability is high with no uncertainty, a degenerate (precise) specification is adopted, i.e.,

$$P(\text{RELIABILITY} = \texttt{high} | \text{AIRCRAFT HEIGHT} = \texttt{low}) = 1;$$

while, in case of uncertainty about such expected level, an interval $[.9, 1]$ is considered instead, and a corresponding non-zero probability for the medium level of reliability should be assumed. Analogous procedures have been employed for the quantification of the PRESENCE.

Regarding the observations, a conditional credal set for each possible value of the corresponding latent variable and each possible level of RELIABILITY and PRESENCE should be assessed.

Let X be a latent variable denoting an ideal factor and O the manifest variable corresponding to the observation of X as returned by a given sensor. For each possible joint value of RELIABILITY and PRESENCE, say (r, p), we should assess lower and upper bounds for $P(O = o|X = x, r, p)$, for each $x \in \Omega_X$ and $o \in \Omega_O = \Omega_X \cup \{*\}$, and then compute the corresponding credal sets.

This quantification step can be simplified by defining a symmetric non-transitive relation of *similarity* among the elements of Ω_X. The similarities between the possible values of a latent variable according to a specific sensor can be naturally represented by an undirected graph as in the example of Fig. 6. In general, given a latent variable X, we ask the Expert to determine, for each possible outcome $x \in \Omega_X$, the outcomes of X that are similar to x and those that are not similar to x.

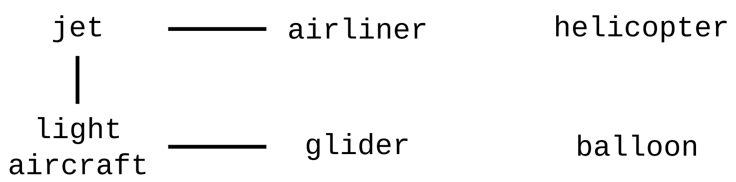

Figure 6. An undirected graph depicting similarity relations about the possible values of the variable AIRCRAFT TYPE according to the observation of a TV camera. Edges connect similar states. The sensor can mix up a `light aircraft` with a `glider` or a `jet`, but not with a `balloon` or a `helicopter` or an `airliner`.

Having defined, for each latent variable and each corresponding sensor, the similarities between its possible outcomes, we can then divide the possible observations in four categories: (i) observing the actual value of X; (ii) confounding the real value of X with a similar one; (iii) confounding the actual value of X with a value that is not similar; (iv) the observation is `missing`. The idea is to quantify, instead of a probability interval for $P(O = o|X = x, p, r)$ for each $x \in \Omega_X$ and each $o \in \Omega_O$, only four probability intervals, corresponding to the four categories of observations described above. As an example, Tab. 1 reports an interval-valued quantification of the conditional probability table $P(O|X, p, r)$ for the ideal variable AIRCRAFT TYPE, for a combination (p, r) of the values of PRESENCE and RELIABILITY that models a good (although not perfect) observation process.

Let us finally explain how the four probability intervals are quantified in our network for each combination of RELIABILITY and PRESENCE and each sensor. The probability interval assigned to the case where the observation is `missing` depends uniquely on the PRESENCE. In particular, the value `absent`, makes the probability of having a `missing` observation equal to one and therefore the probability assigned to all the other cases are equal to zero. It follows that we have only seven combinations of RELIABILITY and PRESENCE to quantify. To this extent, we use constraints based on the concept of *interval dominance* to characterize the different combinations. [11] In order of accuracy of the observation, the combinations are the following:

[11] Given a credal set $K(X)$ over a variable X, and two possible values $x, x' \in \Omega_X$, we say that the x *dominates* x' if $P(X = x') < P(X = x)$ for each $P \in K(X)$. It is easy to show that interval dominance, i.e., $\overline{P}(X = x') < \underline{P}(X = x)$, is a sufficient condition for

	light air.	glider	jet	airliner	balloon	helicopter
light aircraft	$[.600, .700]$	$[.250, .350]$	$[.125, .175]$	0	0	0
glider	$[.125, .175]$	$[.600, .700]$	0	0	0	0
jet	$[.125, .175]$	0	$[.600, .700]$	$[.250, .350]$	0	0
airliner	0	0	$[.125, .175]$	$[.600, .700]$	0	0
balloon	0	0	0	0	$[.900, .950]$	0
helicopter	0	0	0	0	0	$[.900, .950]$
missing	$[.050, .100]$	$[.050, .100]$	$[.050, .100]$	$[.050, .100]$	$[.050, .100]$	$[.050, .100]$

Table 1
A good quality observation of the AIRCRAFT TYPE based on the graph in Fig. 6. A small probability for `missing` has been assumed for each value of the ideal variable. This determines also the probability of observing the actual variable for values that are not similar to any other state. In the other cases, the right observation has been described by a lower intervals, while the probabilities of confounding the actual value with a similar one are obtained by symmetry and reachability properties.

(i) `High`, `present`. The correct observation dominates (clearly) the similar observations. The probability for non-similar observations is zero and is therefore dominated by all the other categories.
(ii) `High`, `partially present`. The correct observation dominates the similar observations and dominates (clearly) the non-similar observations. The similar observations dominates the non-similar observations.
(iii) `Medium`, `present`. The correct observation dominates the similar observations and dominates the non-similar observations. The similar observations dominates the non-similar observations.
(iv) `Medium`, `partially present`. The correct observation does not dominate the similar observations but dominates the non-similar observations.
(v) `Low`, `present`. No dominance at all.
(vi) `Low`, `partially present`. No dominance at all, but more overlapping among the intervals than in (v).
(vii) `Absent` (no matter what the reliability is). The probability of a `missing` observation is equal to one, this value dominates all the other values.

As an example, note that the intervals specified in Tab. 1 correspond to the first combination. Specifications of probability intervals for the other combinations have been obtained by considerations analogous to the ones described in the caption of Tab. 1.

dominance.

6 Information Fusion by Imprecise Probabilities

The procedure described in Sect. 4.2 and 5.2 in order to merge the observations gathered by different sensors can be regarded as a possible imprecise-probability approach to the general *information fusion* problem. In this section, we take a short detour from the military aspects to illustrate some key features of such an approach by simple examples.

Let us first formulate the general problem. Given a latent variable X, and the manifest variables $O_1, \ldots, O_n$ corresponding to the observations of X returned by n sensors, we want to update our beliefs about X, given the values $o_1, \ldots, o_n$ returned by the sensors.

The most common way to solve this problem is to assess a (precise) probabilistic model over these variables, from which the conditional probability mass function $P(X|o_1, \ldots, o_n)$ can be computed. That may be suited to model situations of relative *consensus* among the different sensors. The precise models tend to assign higher probabilities to the values of X returned by the majority of the sensors, which may be a satisfactory mathematical description of these scenarios.

The problem is more complex in case of *disagreement* among the different sensors. In these situations, precise models assign similar posterior probabilities to the different values of X. But a uniform posterior probability mass function seems to model a condition of *indifference* (i.e., we trust the different observed values with the same probability), while sensors disagreement reflects instead a condition of *ignorance* (i.e., we do not know which is the most likely value among the observed ones).

Imprecise-probability models are more suited for these situations. Posterior ignorance about X can be represented by the impossibility of a precise specification of the conditional mass function $P(X|o_1, \ldots, o_n)$. The more disagreement we observe among the sensors, the wider we expect the posterior intervals to be, for the different values of X.

The case where the bounds of a conditional probability strictly contain those of the corresponding unconditional probability, and that happens for all the conditional events of a partition, is known in literature as *dilation* [16], and is relatively common with coherent imprecise probabilities.

The following small example, despite its simplicity, is sufficient to outline how these particular features are obtained by our approach.

Example 1 *Consider a credal network over a latent variable X, and two manifest variables O_1 and O_2 denoting the observations of X returned by two identical sensors. Assume to be given the strong independencies coded by the graph in Fig. 7.*

Let all the variables be Boolean. Assume $P(X)$ uniform and both $P(O_i = 1|X = 1)$ and $P(O_i = 0|X = 0)$ to take values in the interval $[1-\gamma, 1-\epsilon]$, for each i=1,2, where the two parameters $0 < \epsilon < \gamma < .5$ model a (small) error in the observation process. Since the network in Fig. 7 can be regarded as a naive credal classifier, *where the latent variable X plays the role of the class node and the observations correspond to the class attributes, we can exploit the algorithm presented in [22, Sect. 3.1] to compute the following posterior intervals:*

$$P(X = 1|O_1 = 1, O_2 = 1) \in \left[\frac{1}{1 + (\frac{\gamma}{1-\gamma})^2}, \frac{1}{1 + (\frac{\epsilon}{1-\epsilon})^2}\right] \simeq [.941, .988],$$

$$P(X = 1|O_1 = 1, O_2 = 0) \in \left[\frac{1}{1 + \frac{\gamma(1-\epsilon)}{\epsilon(1-\gamma)}}, \frac{1}{1 + \frac{\epsilon(1-\gamma)}{\gamma(1-\epsilon)}}\right] \simeq [.308, .692],$$

where the numerical values have been computed for $\epsilon = .1$, $\gamma = .2$. Note that for these specific values, as well as in the general case, the first interval is strictly greater than $.5$, a value that represents the middle point of the second interval. Thus, we can conclude that: (i) consensus between the sensors increases the posterior probability for X; (ii) disagreement increases our ignorance about X (the probability dilates). [12]

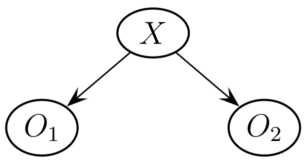

Figure 7. The credal network for Example 1.

These calculi can be easily generalized to the case of n sensors. As an example, Fig. 8 depicts the posterior intervals for the observation of a Boolean latent variable by nine identical sensors for different levels of consensus among them. Note that our approach based on credal networks reproduces the cautious behaviour we want to model, while a Bayesian network would abruptly change its estimates with the number of correct observations passing from four to five. Unlike the imprecise case, precise conditional probabilities might therefore produce unreliable extreme values in the posterior beliefs because of high sensitivity to small changes in the error rate.

It should be also pointed out that the only assumption required by this approach is the conditional independence between the manifest variables (observations) given

[12] The value $\epsilon = 0$ has been excluded, as it models a situation where both the sensors can be perfectly reliable. Clearly, such a scenario is not compatible with a disagreement between the observations.

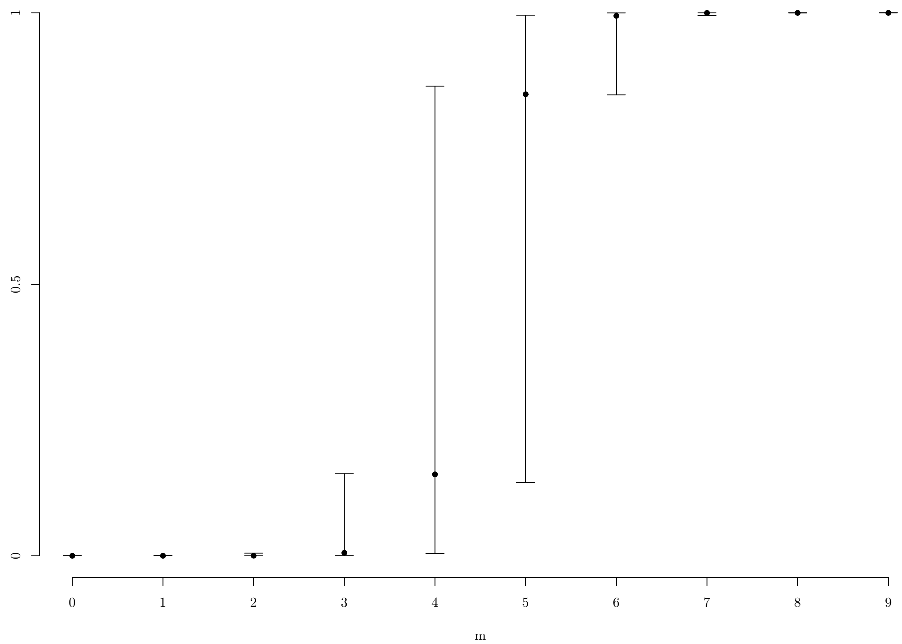

Figure 8. Posterior intervals for $P(X = 1|O_1 = 1, \ldots, O_m = 1, O_{m+1} = 0, \ldots O_9 = 0)$ as a function of m, assuming both $P(O_j = 1|X = 1)$ and $P(O_j = 0|X = 0) \in [.8, .9]$, for each $j = 1, \ldots, 9$, and a uniform $P(X)$. Black dots denote precise posterior probabilities computed assuming the precise value $.85$ for the conditional probabilities.

the latent variable (actual value of the quantity to be measured). A condition which seems to be verified in many concrete cases, as for example that of the military problem we address in this paper.

In fact, assuming fixed levels of AIRCRAFT HEIGHT, RELIABILITY and PRESENCE, Fig. 4 reproduces the same structure of the prototypical example in Fig. 7, with four sensors instead of two. The same holds for any subnetwork modeling the relations between a latent variable and the relative manifest variables.

7 Algorithmic Issues and Simulations

The discussion in Sect. 4 and Sect. 5 led us to the specification of a credal network, associated to the graph in Fig. 5, defined over the whole set of considered variables, i.e., core variables, observations collected by the different sensors, reliability and presence levels.

At this point, we can evaluate the risk associated to an intrusion, by simply updating the probabilities for the four possible values of the risk factor, conditional on the values of the observations returned by the sensors.

The size of our credal network prevents an exact computation of the posterior probability intervals. [13] Approximate procedures should be therefore employed, unless we do not want to restrict our analysis to the core of the network by assuming perfectly reliable observations for all the ideal factors.

The high computational complexity of the updating problem on the whole network should be regarded as the mathematical counterpart of the difficulties experienced by the military experts during the identification of the intruder. On the other side, when all the factors are observed in a perfectly reliable way, the goal of the intruder can be easily detected, exactly as the mathematical task of updating the core of the network is equivalent to computing class probabilities in a *naive credal classifier*, a task efficiently solved by the algorithm in [22].

Thus, we have first performed extensive simulations on the core of the network. Expert has considered several combinations of values for the ideal factors and checked whether or not the set of undominated classes returned by the core of the network was including his personal evaluation of the goal of an intruder for that scenario. Every time a mismatch between the human and the artificial expert was detected, the quantification of the probability for the network was updated. Remarkably, at the end of this validation task, we have obtained a network core able to simulate the Expert's evaluation in almost every considered scenario.

Then, as a test for the whole network, we have considered a simulated restricted flight area for the protection of a single object in the Swiss Alps, surveyed by an identification architecture characterized by absence of interceptors and relatively good coverage of all the other sensors. We assumed as meteorological conditions discontinuous low clouds and daylight. The simulated scenario reproduces a situation where a provocateur is flying very low with a helicopter and without emitting any identification code. The corresponding evidences are reported in Tab. 2.

In order to compute the posterior intervals we have considered an approximate algorithm called *generalized loopy 2U* [2], whose performances in terms of accuracy and scalability seems to be quite good. The posterior intervals were computed in few seconds on a 2.8 GHz Pentium 4 machine.

For this simulation we have assumed uniform prior beliefs about the four classes of risk. [14] Fig. 9.a depicts the posterior probability intervals for this simulated scenario. The upper probability for the outcome `renegade` is zero, and we can therefore exclude a terrorist attack. Similarly, the lower probability for the outcomes `agent provocateur` and `damaged` are strictly greater than the upper proba-

[13] The existing algorithms for exact updating of credal networks (e.g., [6,15]) are typically too slow for models with dense topologies and more than 50 nodes.

[14] Any credal set can be used to model decision maker's prior beliefs about the risk factor. Nevertheless, as noted in [14], a *vacuous* prior (i.e., a credal set equal to the whole probability simplex) would make vacuous also the posterior inferences.

VARIABLE	SENSOR 1 (SSR)	SENSOR 2 (3D)	SENSOR 3 (2D)	SENSOR 4 (TV)	SENSOR 5 (GROUND)	SENSOR 6 (TRACK)
AIRCRAFT HEIGHT	very low	very low	-	-	low	very low / low
TYPE OF AIRCRAFT	-	-	-	helicopter	helicopter	-
FLIGHT PATH	U-path	U-path	U-path	U-path	missing	U-path
HEIGHT CHANGES	descent	descent	descent	-	missing	descent
ABSOLUTE SPEED	slow	slow	slow	-	slow	slow
REACTION TO ADDC	positive	positive	positive	positive	positive	positive

Table 2
Sensors observations for the simulations in Fig. 9. The AIRCRAFT HEIGHT according to the tracking radar (SENSOR 6) is `very low` in (a) and (b), and `low` in (c).

bility for the state `erroneous`, and we can reject as less credible also this latter value because of interval dominance.

The ambiguity between `agent provocateur` and `damaged` is due, in this case, to the bad observation of the AIRCRAFT HEIGHT. In fact, a damaged helicopter is expected to land as soon as possible. While, in the modeled scenario, a provocateur is not expected to land. With a bad observation of the height, we are unable to understand if the helicopter has landed or not and therefore the ambiguity between the two risk categories is reasonable.

Indecision between `agent provocateur` and `damaged` disappears if we assume higher expected levels of RELIABILITY and PRESENCE for the sensors devoted to the observation of the AIRCRAFT HEIGHT. The results in Fig. 9.b state that the intruder is an `agent provocateur`, as we have assumed in the design of this simulation.

In Fig. 9.c we still consider a high quality observation, but more disagreement between the sensors (see Tab. 2). This produces, also in this case, indecision between two classes. Remarkably, as we expect from our model of the information fusion in case of disagreement, the intervals we observe seem to reproduce the union of the intervals computed on the network core assuming respectively `very low` (Fig. 9.d) and `low` (Fig. 9.e) AIRCRAFT HEIGHT.

Remarkably, these results have been recognized by the Expert as reasonable estimates for the considered scenarios. Yet, an extensive validation process, consisting in analyses of this kind by different military experts on many other scenarios, should be regarded as a necessary future work.

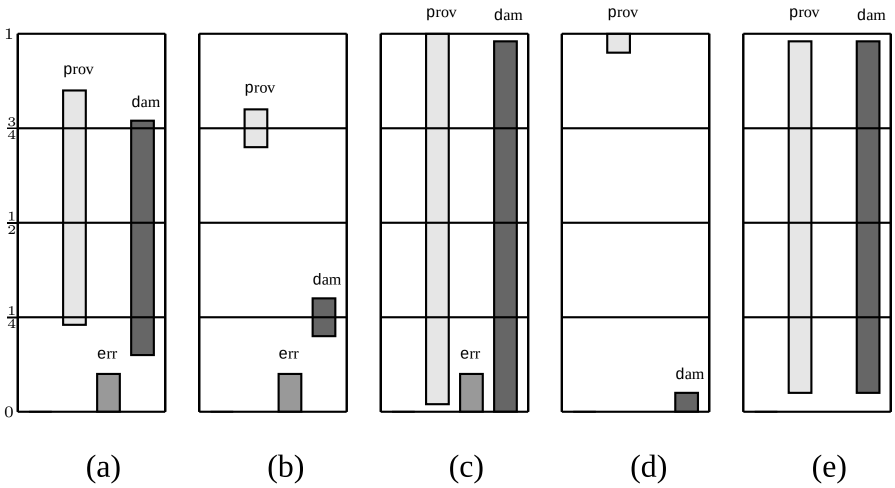

Figure 9. Posterior probability intervals for the risk factor, corresponding to the evidences reported in Tab. 2. The histogram bounds denote lower and upper probabilities. The quality of the observation of the AIRCRAFT HEIGHT is assumed to be higher in (b) than in (a). The histograms in (c) refers to a situation of increased disagreement between the sensors observing the AIRCRAFT HEIGHT. Finally, (d) and (e) report respectively the exact posterior intervals in a situation where the observation of the factors is assumed to be perfectly reliable and corresponds to the value returned by the majority of the sensors.

8 Conclusions and Future Work

A model for determining the risk of intrusion of a civil aircraft into restricted flight areas has been presented. The model embeds in a single coherent mathematical framework human expertise expressed by imprecise-probability assessments, and a structure reproducing complex observation processes and corresponding information fusion schemes.

The risk evaluation corresponds to the updating of the probabilities for the risk factor conditional on the observations of the sensors and the estimated levels of presence and reliability. Preliminary tests considered for a simulated scenario are consistent with the judgements of an expert domain for the same situation.

As future work we intend to test the model for other historical cases and simulated scenarios. The approximate updating procedure considered in the present work, as well as other algorithmic approaches will be considered, in order to determine the most suited for this specific problem.

In any case, it seems already possible to offer a practical support to the military experts in their evaluations. They can use the network to decide the risk level corresponding to a real scenario, but it is also possible to simulate situations and verify the effectiveness of the different sensors in order to design an optimal identification architecture.

Finally, we regard our approach to the fusion of the information collected by the different sensors as a sound and flexible approach to this kind of problems, able to work also in situations of contrasting observations between the sensors.

Acknowledgments

We are very grateful to the experts of Armasuisse and the Swiss Air Force for their help and support and for the interesting discussions.

A Python/C++ software implementation of the generalized loopy 2U algorithm (http://www.idsia.ch/∼yi/gl2u.html) developed by Sun Yi has been used to update the credal networks. The software tool *lrs* (http://cgm.cs.mcgill.ca/∼avis/C/lrs.htm) has been used to compute the vertices of the conditional credal sets corresponding to the probability intervals provided by the Expert. The authors of these public software tools are gratefully acknowledged.

References

[1] A. Antonucci and M. Zaffalon. Decision-theoretic specification of credal networks: A unified language for uncertain modeling with sets of Bayesian networks. *International Journal of Approximate Reasoning*, 49(2):345–361, 2008.

[2] A. Antonucci, M. Zaffalon, Y. Sun, and C.P. de Campos. Generalized loopy 2U: A new algorithm for approximate inference credal networks. In M. Jaeger and T.D. Nielsen, editors, *Proceedings of the Fourth European Workshop on Probabilistic Graphical Models*, pages 17–24, Hirtshals (Denmark), 2008.

[3] D. Avis and K. Fukuda. A pivoting algorithm for convex hulls and vertex enumeration of arrangements and polyhedra. *Discrete and Computational Geometry*, 8(3):295–313, 1992.

[4] D. Boorsbom, G.J. Mellenbergh, and J. van Heerden. The theoretical status of latent variables. *Psychological Review*, 110(2):203–219, 2002.

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

[6] A. Cano, M. Gómez, S. Moral, and J. Abellán. Hill-climbing and branch-and-bound algorithms for exact and approximate inference in credal networks. *International Journal of Approximate Reasoning*, 44(3):261–280, 2007.

[7] F.G. Cozman. Credal networks. *Artificial Intelligence*, 120:199–233, 2000.

[8] F.G Cozman. Graphical models for imprecise probabilities. *International Journal of Approximate Reasoning*, 39(2-3):167–184, 2005.

[9] C.P. de Campos and F.G. Cozman. The inferential complexity of Bayesian and credal networks. In *Proceedings of the 19th International Joint Conference on Artificial Intelligence*, pages 1313–1318, Edinburgh, 2005.

[10] C.P. de Campos and Q. Ji. Strategy selection in influence diagrams using imprecise probabilities. In *Proceedings of the 24th Conference on Uncertainty in Artificial Intelligence*, pages 121–128. AUAI Press, 2008.

[11] E. Demircioglu and L. Osadciw. A Bayesian network sensors manager for heterogeneous radar suites. In *IEEE Radar Conference*, Verona, NY, 2006.

[12] D. Koller and A. Pfeffer. Object-oriented Bayesian networks. In *Proceedings of the 13th Conference on Uncertainty in Artificial Intelligence*, pages 302–313, 1997.

[13] J. Pearl. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan Kaufmann, San Mateo, 1988.

[14] A. Piatti, M. Zaffalon, F. Trojani, and M. Hutter. Learning about a categorical latent variable under prior near-ignorance. In G. de Cooman, J. Vejnarová, and M. Zaffalon, editors, *Proceedings of the 5th International Symposium on Imprecise Probability: Theories and Applications*, pages 357–364, Prague, 2007. Action M.

[15] J.C. Rocha and Cozman F.G. Inference in credal networks: branch-and-bound methods and the A/R+ algorithm. *International Journal of Approximate Reasoning*, 39:279–296, 2005.

[16] T. Seidenfeld and L. Wasserman. Dilation for sets of probabilities. *Annals of Statistics*, 21(3):1139–1154, 1993.

[17] AIP Services. *Aeronautical Information Publication Switzerland*. Skyguide, 2007.

[18] A. Skrondal and S. Rabe-Hasketh. *Generalized latent variable modeling: multilevel, longitudinal, and structural equation models*. Chapman and Hall/CRC, Boca Raton, 2004.

[19] B. Tessem. Interval probability propagation. *International Journal of Approximate Reasoning*, 7(3):95–120, 1992.

[20] P. Walley. *Statistical Reasoning with Imprecise Probabilities*. Chapman and Hall, New York, 1991.

[21] P. Walley. Measures of uncertainty in expert systems. *Artificial Intelligence*, 83(1):1–58, 1996.

[22] M. Zaffalon. The naive credal classifier. *Journal of Statistical Planning and Inference*, 105(1):5–21, 2002.