Semantic Criteria for the Assessment of Uncertainty Handling Fusion Models

Anne-Laure Jousselme
NATO STO Centre for Maritime Research and Experimentation
Viale San Bartolomeo, La Spezia, Italy
Anne-Laure.Jousselme@cmre.nato.int

***Abstract*—This paper proposes an illustration of the Uncertainty Representation and Reasoning Evaluation Framework (URREF) for the comparison of two classical fusion schemes. We revisit the classical works comparing Bayes' rule and Dempster's rule for fusion, and identify the criteria that have been used for both semantic theoretical and algorithmic implementation comparisons. The discussion is illustrated on a practical maritime use case previously experimented as a game with experts during a Table Top eXercise (TTX). The maritime use case described is simple enough to clearly highlight the differences and similarities between the different approaches, but complex enough to reflect real practical situations where sources of information of various nature are involved, the information they provide reflect different type of imperfection, several features are measured or observed, concurrent events happen. We highlight how the elements of the URREF ontology of** UNCERTAINTYNATURE, UNCERTAINTYDERIVATION, UNCERTAINTYTYPE **and** SOURCE **Quality influence the assessment of** UNCERTAINTYMODELS **through** REPRESENTATION **and** REASONING **evaluation criteria. We propose a list of additional elements which could be considered in the URREF ontology: Type of Problem (revision, prediction, fusion), Information Type (generic vs singular), Uncertainty Supports (variable, link, uncertainty statement) and Measurement Scale. We illustrate the different semantics of the two rules and how they may use different information.**

Keywords: URREF; Semantics; Bayes' rule; Dempster's rule; Evaluation criteria.

I. INTRODUCTION

Since the emergence of alternative frameworks to probabilistic reasoning (see for instance [1] for a survey) the debate about which approach should be followed for uncertainty handling is still on-going. Handling uncertainty in fusion problems is indeed a major issue to most of the algorithm designers as it generates many questions such as what "uncertainty" means, where it comes from, on what it bears, how to interpret the associated numerical values or measures, how to distinguish between its different nature, etc. The characterisation of uncertainty concepts and how they relate to information quality provides partial answers (*e.g.* [2], [3], [4]). It is expected that a deep understanding of the different Uncertainty Representation and Reasoning Techniques (URRTs), their underlying mathematical framework and associated hypotheses will help identify the evaluation criteria which would further guide the fusion system's designer in making informed choices about the most suitable technique to the problem at hand. It is expected as well that this would provide clearer explanations of the algorithm to the user for an improved synergy between the human and the machine [5].

The assessment of URRTs and their underlying fusion systems has been the focus of the Evaluation of Techniques for Uncertainty Representation (ETUR) Working Group since 2011. The main contribution so far is an ontology which relates evaluation criteria specifically focused on the uncertainty handling, with other information quality aspects such as the nature of uncertainty (aleatory vs epistemic), the derivation of uncertainty (objective vs epistemic), the type of uncertainty (imprecision vs uncertainty) [6].

The comparison of the probabilistic and evidential uncertainty representation models together with their associated classical rule for fusion, say Bayes' and Dempster's respectively, has been the topic of several papers. It appears that rather than competitors the two approaches are dedicated to different problems and different types of information. For instance, in Shafer's view, Dempster's rule is specifically dedicated to combine uncertain and imprecise singular information, such as testimonies. While as stated by Diaconis *et al.* [7], Bayes' rule is not applicable in case of probable knowledge, unanticipated knowledge and introspective knowledge. Although Bayes' rule is an updating rule, it has been widely used for fusion purposes as well (*e.g.* [8], [9]).

The assessment criteria used for comparing the two approaches differ depending on the perspective adopted by the authors. On the one hand, in [10], [11], [12] for instance, the criteria are the meaning or interpretation of the measures (frequentist vs subjectivist), how it impacts the meaning of conditioning operator, and the type of problem to be solved (revision, prediction, fusion). These authors also consider the expressiveness of the framework such as its ability to distinguish between ignorance and equal chance, between epistemic and aleatory uncertainty, between betting rates and degrees of confidence. The theoretical semantic comparison also considers hypotheses such as the unicity of probability measure, or the compatibility relations between frames (due to refinement) [13].

On the other hand, from a practical perspective when actually implementing the algorithms, other studies discuss criteria such as the classification rates (true positive, false negative), the conclusiveness in term of the output probability of the favored hypothesis (or cautiousness in decision), the computational complexity, the sensitivity to perturbations or robustness, the availability of prior information (*e.g.* [8], [9],

[14], [15]).

The work presented in this paper is an illustration of use of the URREF on a practical use case from the maritime domain, to compare Bayes' and Dempster's rule with a focus on semantic assessment criteria. In Section II, the maritime use case is presented and a formalisation of the problem is proposed, as a common basis for the representation and reasoning schemes to be compared. The main URREF ontology elements and associated criteria are reviewed in Section III and some extension is proposed. The characterisation of the information available in our use case is performed along different URREF dimensions and the new proposed ones. The two URRTs of Bayes' and Dempster's rules for fusion are compared in Section IV along the representation and reasoning criteria. Finally, Section V concludes on future works.

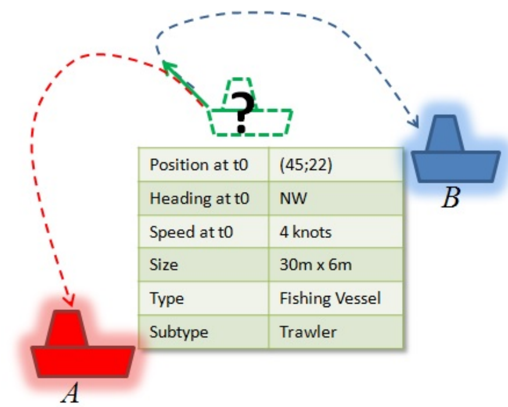
II. A MARITIME USE CASE

In order to gain a better understanding of how the military operators take decisions facing with uncertainty and conflicting information, a *Risk Game* has been designed and played at the Centre for Maritime Research and Experimentation (CMRE) as part of a Table Top eXercise (TTX) in November 2014 [16]. A pool of 32 experts in large majority maritime officers participated in this experience. The goal for the players was to find a missing vessel among two possible options, based on a series of pieces of information available through several sources. The information was provided by abstract cards that the player would sequentially query, analyse, combine with the previous ones obtained, to be able to then rate step by step his/her belief regarding the location of the vessel. The sequential processing aimed solely at decomposing the reasoning process, while actually the pieces of information were all available at the same instant t, so the order of combination would not matter. Once the player was confident enough in his/her assessment, he/she had to decide to send or not to send a patrol aircraft for further checking and confirm his/her belief, a step which concluded the game.

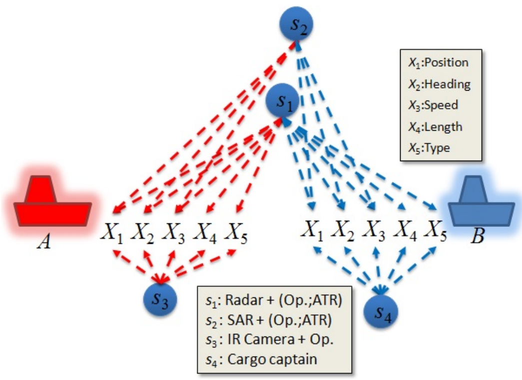
The main purpose of the Risk Game was to study the impact of information quality on the belief assessment and decision making. The information quality was thus made varying along the three dimensions of *certainty*, *precision* and *trueness* and randomly selected by a dice roll. This fusion task was selected as an example of a routine task for military operators, which benefits the design of fusion algorithms to support decision making. We will consider this use case as an example of application of the URREF for the assessment of a fusion algorithm which would automatically combine the different pieces of information and come up with a belief assessment of the two hypotheses under consideration.

A. The scenario

After being informed of the loss of the AIS contact with a particular fishing vessel one hour ago (at time 0), the Watch Officer (WO) now (at time t) needs to recover the track and locate the vessel. The locations of two unidentified tracks, called Vessel A and Vessel B, are provided as the only two possible locations for the missing vessel (see Figure 1). The Watch Officer has to match the known features of the missing vessel, as reported by its last AIS contact, with the ones of the two unidentified tracks, as reported by the on-site sources. Hence, its name, MMSI, IMO, type, length, width, etc are known with a very high confidence to the Watch Officer.[1]



(a) Last AIS information sent by the missing vessel and possible current locations, A or B.



(b) Four sources are providing information about five attributes of the two tracks.

Fig. 1. A missing vessel scenario.

The sources of information available to the Watch Officer combine a variety of sensors both cooperative (*e.g.* Automatic Identification System (AIS)) and non-cooperative (*e.g.* radar, camera), whose measurement is processed either by automatic algorithms (*e.g.* tracker, Automatic Target Recognition (ATR) algorithm)) or human analysts (*e.g.* camera analyst, cargo captain). The radar covers the whole area, the Infra-Red (IR) camera covers only the area around Track A, a cargo is in the vicinity of Track B but too far from Track A for visual identification, and a Synthetic Aperture Radar (SAR) imagery covering the whole area has been taken 30 minutes ago. Sources are imperfect and provide information which can be *uncertain* (the source itself is uncertain about its estimation or statement), *imprecise* (the source provides several possible values for the attribute estimated) and *false* (the value provided by the source does not correspond to the true value). Consequently, when combining the different pieces of information, the Watch Officer may face conflicting information.

[1]This scenario is inspired by the vehicle-borne Improvised Explosive Device (IED) scenario of the Uncertainty Forum at Fusion 2010.

B. Problem formalisation

The design of a fusion algorithm highly depends on the designer's previous experience and attitudes. There are obviously many modeling solutions to the same problem, as highlighted during the exercise of the *Uncertainty Forum* held during the Information Fusion conference in Edinburgh in 2010. As summarised in [17], three different experts came to three different solutions to the same problem with the same mathematical uncertainty representation of belief functions. It appears that these experts in fusion algorithm design made personal choices about (1) the selection of the variables and the definition of their domain (universe of discourse, frame of discernment), (2) selection of the sources and associated pieces of information to be combined, (3) modeling of sources' reliability, (4) definitions of belief functions, (5) the combination rule and (6) the decision rule. These different abstraction processes are referred in [18] as the *isolation abstraction*, the *datum abstraction*, the *data generation abstraction*. We will focus in the following on elements (3), (4) and (5) and we will fix the elements (1) and (2) by providing a basic formalisation and modeling of the problem. Also, the decision rule will not be considered in the analysis.

- $\mathcal{S}$ is the finite set of available sources of information. A source of information is understood in its most general sense and covers the cases of a sensor, an algorithm, a human, a database, a social medium, etc. Moreover, we distinguish between the information container (*e.g.* a radar signal) and the information provider (*e.g.* the tracking algorithm):
  $$\begin{aligned}\mathcal{S} = \{&\text{AIS}; (\text{Radar;Tracker/ATR}); (\text{Radar;Operator C});\\ &(\text{SAR;ATR}); (\text{SAR;Analyst B}); (\text{IR camera;Analyst A});\\ &(\text{Cargo; Captain})\} = \{s_0; s_{11}; s_{12}; s_{21}; s_{22}; s_3; s_4\}\end{aligned}$$
- $\mathcal{O}$ denotes the set of objects of interest about which the information is gathered, $\mathcal{O} = \{V_A, V_B\}$, where V_A and V_B denote respectively the physical objects Vessels A and B.
- $\mathcal{A}$ is a set of object attributes relevant to our problem. Here, we consider for each vessel:
  $$\mathcal{A} = \{\text{LOCATION}; \text{HEADING}; \text{SPEED}; \text{TYPE}; \text{LENGTH}\}$$
  denoted by the respective subscripts L, H, S, T and l. The attributes are further distinguished by vessel so LA is the attribute Length of Vessel A, for instance.
- X_a is the variable associated to the attribute a of $\mathcal{A}$ for an object of $\mathcal{O}$, resulting in our case in $5 \times 2 = 10$ variables.
- $\mathcal{X}_a$ is the domain of variable X_a associated with attribute a and contains the set of possible values for X_a. For instance:
  $$\mathcal{X}_{TA} = \mathcal{X}_{TB} = \{FV; \neg FV\}$$
  are the domains of the variables X_{TA} and X_{TB}, the "Type of Vessel A" and "Type of Vessel B" respectively, and FV means "Fishing Vessel" while $\neg FV$ means "Not a Fishing Vessel.
- The observation (or measurement) space $\mathcal{X}$ is the Cartesian product of the domains of all the observation variables:
  $$\mathcal{X} = \mathcal{X}^{(A)} \times \mathcal{X}^{(B)}$$
  where $\mathcal{X}^{(A)} = \mathcal{X}_{LA} \times \mathcal{X}_{HA} \times \mathcal{X}_{SA} \times \mathcal{X}_{TA} \times \mathcal{X}_{lA}$ is the observation space for Vessel V_A only, and $\mathcal{X}^{(B)}$ is the observation space for Vessel B.
- The corresponding "ground truth" space is denoted by $\mathcal{X}^*$ and contains the real values for both vessels V_A and V_B. To each value in $\mathcal{X}^*$ correspond possibly several measurements in $\mathcal{X}$, a single one being "true".
- The decision space describes the set of possible answers to our question of interest, here the location of the missing vessel among two possible ones:
  $$\Theta = \{A; B\}$$
  where A is the hypothesis "The missing vessel is V_A" and B is the hypothesis "The missing vessel is V_B".
- $\phi(t, s, o, a)$ (or simply ϕ) denotes the information provided at time t by source s about the value of attribute a of object o. For instance, $\phi(t, s_{11}, V_A, \text{SPEED}) = [2.8; 3.2]$ knots means that the source s_{11} (in this case the radar and its associated tracker) reports at time t that the speed of V_A is between 2.8 and 3.2 knots. ϕ can be expressed in natural language, formalised in a specific mathematical language such as probabilities or the like, be a logical statement, a single numerical value, a score vector, etc.
- Φ_0 is the set of pieces of information available at "time 0". We use the convention of "time 0" to denote the background or prior knowledge (or information). This background information is called *generic information* [1]. Although it is barely described, it may be convenient to assign to generic knowledge the source from which it originates.
- Φ_t is the set of pieces of information about the situation at t. This factual information about the current state of the world is called *singular information* [1]. In the specific case of precise, certain and numerical measurements, Φ_t is a vector of $\mathcal{X}$. Examples of generic and singular information are provided in Table II.

Let us denote by $\eta^Y_{t,s}$ a formal uncertainty representation defined over a domain Y, originating from source s at time t. Y is any subset of $\mathcal{X} \times \mathcal{X}^* \times \Theta$. We would like to design a fusion algorithm F which computes:

$$\eta^{\Theta}_{t,F}(.|\Phi_0, \Phi_t) \tag{1}$$

where $\eta^{\Theta}_{t,F}$ is the uncertainty or belief output by the fusion algorithm to be designed about the location of the missing vessel, and $\{\Phi_0; \Phi_t\}$ is the information available. Note that the vertical bar $|$ is not necessary the conditioning operator.

C. Basic logical reasoning

Let us consider only the variables X^*_{TA} and X^*_{TB} corresponding to the true type of the two vessels, and let us assume that their respective domain is an *exhaustive* and *exclusive* set of hypotheses. The basic underlying reasoning of the algorithm to be designed should satisfy the following rules:

- If Vessel A **is not** a fishing vessel, then the missing vessel is not A and thus is B.
- If Vessel A **is** a fishing vessel, then the missing vessel is either A or B. We have no information about B.

Similar rules hold for Vessel B. These logical rules can be seen as part of the generic information (see information ϕ_b in Table II) and written:

$$\begin{cases}(X^*_{TA} = \neg FV) \to B \\ (X^*_{TA} = FV) \to A \vee B\end{cases} \quad \begin{cases}(X^*_{TB} = \neg FV) \to A \\ (X^*_{TB} = FV) \to A \vee B\end{cases} \tag{2}$$

They define two multivalued (one-to-many) mappings from $\mathcal{X}^*$ to Θ as:

$$\begin{cases}\Gamma^*_{TA}(\neg FV) = \neg A = B \\ \Gamma^*_{TA}(FV) = A \vee B = \Theta\end{cases} \quad \begin{cases}\Gamma^*_{TB}(\neg FV) = \neg B = A \\ \Gamma^*_{TB}(FV) = A \vee B = \Theta\end{cases} \tag{3}$$

III. Assessment criteria of uncertainty representation and reasoning techniques

The task of the fusion algorithm designer is to model the problem into a formal mathematical framework which "properly" handles the uncertainty and provides and "adequate" output $\eta^{\Theta}_{t,F}(.|\Phi)$. The purpose of the URREF ontology is to further define what may be understood by "properly" and "adequate" and detail the corresponding criteria.

A. Evaluation subjects

An EvaluationSubject is any item which may have an impact on the system's output, that can be made varying, and thus be compared and evaluated according to a series of corresponding criteria [19]. A fusion algorithm may be assessed as a "black-box", considering only its outputs. That's the purpose of the DataOutputCriteria, such as precision and accuracy. A detailed description and characterisation of uncertainty handling in a fusion system highlights additional elements or components subject to evaluation. These components are the *uncertainty supports* which tell "what is uncertain" (see Section III-C), the *modeling of uncertainty* over the uncertainty supports (assessed by the RepresentationCriteria) and the *uncertainty calculus* (assessed by the ReasoningCriteria). In Section IV, we will evaluate and compare Bayes' rule and Dempster's rule (calculus) and their underlying mathematical representation of uncertainty (UncertaintyModel).

B. Uncertainty model

An UncertaintyModel is mathematical framework for representing and reasoning with uncertainty. We rely on the clear overview of uncertainty representation frameworks proposed by Dubois and Prade in [1]. It is not exhaustive but the main frameworks of probabilities, possibilities and belief functions are presented and compared on their respective expressiveness power and underlying semantics of uncertainty representation and calculus. The authors clearly stress how the selection of a suitable framework and its fusion or inference rule highly depends on (1) the type of problem to solve and (2) the characterisation of information (and associated uncertainty) at our disposal.

C. Uncertainty supports

We define the *uncertainty supports* [20], as items for which some uncertainty need to be represented (other said, what we are uncertain about) and distinguish between

- *individual variables* from either the measurement space $\mathcal{X}$ or the decision space Θ. For instance, the type of Vessel A, the location of the missing vessel (Θ) are uncertainty support in the form of individual variables.
- *links between variables* either from $\mathcal{X}$ or between $\mathcal{X}$ and Θ. For instance, the couple (length; type) is an uncertainty support of type "link between variables" within the observation space. The couple (type; location of missing vessel) is an uncertainty support between the observation and decision space.
- *uncertainty expression* as expressed over the above supports. For instance, the joint distribution of length and types of vessels is itself the support of some uncertainty since its estimation may not reflect the real distribution (due to a lack of data for instance). This is a second-order uncertainty.

D. Type of problem

The type of problem to solve is an element that should drive the uncertainty modeling and handling. At least the problems of *prediction*, *revision* and *fusion* should be distinguished [1], [21]:

- Revision: Assessing the new belief of an agent considering a new piece of information, as a singular information.
- Prediction: Assessing the new belief about some property of the world based on some model of the world and some new observation on the current state.
- Fusion: Merging in parallel pieces of information provided by several sources to deduce the most useful plausible information they contain [1].

The problem we address in this paper is a *fusion* problem.

E. Source quality

The source providing the information has a high impact on the result of the reasoning. Indeed, information from the IR camera ATR processing may be handled differently than information from the testimony of a cargo captain. We consider here the only source quality dimension of *reliability* which the ability of the source to provide *correct* (*i.e.* conform to truth) output. The reliability of a source is generic information which can be derived beforehand upon several trials of the source where its answers are compared to some ground truth. The result is for instance captured by confusion matrices as displayed in Table I for the ATR algorithm and the Cargo Captain, reflecting their ability to recognise fishing vessels.

TABLE I
Confusion matrices of the ATR algorithm observing location A (ϕ_c) and the Cargo Captain observing location B (ϕ_d).

$p_0(X\|X^*)$		ATR		Cargo Capt.	
		FV	$\neg FV$	FV	$\neg FV$
Truth	FV^*	0.8	0.2	0.95	0.05
	$\neg FV^*$	0.4	0.6	0.1	0.9

F. Characterisation of information

As a matter of convention in this paper, the notions of data, knowledge, evidence and information are all covered by the single term *information*. This is driven by the need to avoid the confusion between the terms and by no means to deny any existence of distinction between these notions. Consequently, "incomplete knowledge", "uncertain evidence", "erroneous data", etc, are all *information* necessarily *imperfect*. The term *uncertainty* may be used sometimes abusively to cover these imperfections as they all induce some uncertainty in the algorithm designer's or decision maker's mind. Moreover, *uncertainty* is considered as the dual of information as classically understood in the field of Generalised Information Theory (GIT) [22].

1) Uncertainty nature: The nature of uncertainty (class UNCERTAINTYNATURE in the URREF ontology) is split into *aleatory* or *epistemic* uncertainty (*e.g.* [23]). Aleatory uncertainty is due to the intrinsic randomness of the phenomenon. For instance, the speed value of a randomly picked-up fishing vessel in a given area at a given time (*e.g.* ϕ_f). Epistemic uncertainty is due to a lack of knowledge. For instance, the speed of a specific fishing vessel (*e.g.* Vessel A) at a given time. While additional information cannot reduce aleatory uncertainty, epistemic uncertainty can be reduced upon the arrival of new information (*e.g.* a radar measurement of Vessel A's speed). As argued in [23], how we deal with the uncertainty of different nature depends on the context and application.

2) Type of information: In [1], the distinction is made between *generic* and *singular* information. Generic information refers to a population of situations such as statistical models, physical rules, logical rules or commonsense knowledge. It is a synthesis of previous knowledge. For instance, the kinematic rule which links the position, speed and heading of a vessel to its new position. Singular information is about the current state of the world such as an observation, a testimony or a sensor measurement. For instance the IR camera analyst's estimation of the type of Vessel A (*e.g.* POI ϕ_l). This distinction is similar to that between knowledge and evidence in the URREF ontology. According to Pearl (as cited in [12]) "Knowledge is understood as "judgments about the general tendency of things to happen," whereas evidence refers to the description of a specific situation".

3) Uncertainty derivation: The way uncertainty measures are derived is usually split into *objective* and *subjective* derivation. This is covered by UNCERTAINTYDERIVATION in the URREF ontology. The uncertainty derivation is objective when it comes from direct observations through physical devices, when it can be assessed in a formal way, *e.g.* via a repeatable derivation process. For instance, the reliability of the IR camera ATR in vessel's classification is objectively derived since it is a direct computation of true positive and false negative. The uncertainty derivation is subjective when it involves some human opinion, judgment or belief. For instance, the initial belief of the decision maker regarding the location of the missing vessel (POI ϕ_a). It is based on his/her personal experience and knowledge of fishermen and of the country, on the geopolitical context, or on some intuition. Note that humans do not always provide subjective information. For instance, a human operator counting the number of vessels in a given area is purely objective, even though it can be erroneous and the human can still have some uncertainty about his counting result.

4) Type of imperfection: Uncertainty is often used as a term to capture the imperfection in information. To avoid confusion, we will use *imperfect information* and distinguish between several types of imperfections such as *uncertainty*, *imprecision*, *error* (falseness), *incompleteness*, *ambiguity*. This is covered by UNCERTAINTYTYPE in the URREF ontology.

5) Scale of measurement: The type of scale over which the information (and associated uncertainty) is expressed also plays a role in the selection of the best framework for representation. The scale refers both to the domain of the variable measured (*e.g.* a set of possible speed values, of possible vessel types) but also to the scale over which the uncertainty is expressed (*e.g.* numerical degree between of $[0;1]$, linguistic quantifier such as *likely*, *probably*, etc). As an example, the types of scale is given by Stevens [24] as *nominal* (*e.g.* the "type of vessel", or "certain vs not certain"), *ordinal* (*e.g.* "slow", fast, very fast", or the sea state as an integer between 1 and 9), *interval* (*e.g.* "date", the location in Latitude-Longitude coordinates) and *ratio* (*e.g.* continuous real scale for "speed" or "length" values). The type of scale is not covered in the URREF ontology.

Table II summarises the characterisation of some information available to solve our specific fusion problem for this maritime use case, referred thus far as ϕ_a to ϕ_l.

IV. COMPARISON OF TWO FUSION RULES THROUGH THE URREF

Acknowledging that the fusion algorithm can be very complex and involve several embedded sub-algorithms, themselves relying on several uncertainty models $\mathcal{M}$, we will consider in the following an "atomic" algorithm which contains only the elements of *uncertainty representation* h and *uncertain reasoning* ρ [25]. We consider here two alternative *uncertainty models* as basis for uncertainty representation and calculus that are probabilities $\mathcal{M}_P$ and belief functions $\mathcal{M}_B$, together with a single associated fusion rule, *i.e.*, Bayes' rule ρ_P for probabilities and Dempster's rule ρ_B for belief functions. Let us assume that the information available is $\Phi_0 = \{\phi_a, \ldots, \phi_d\}$ and $\Phi_t = \{\phi_j, \phi_k\}$ as listed in the blue cells of Table II.

A. Probabilities and Bayes' rule

Let denote by p^X the probability distribution defined over the domain of variable X in the measurement space and by p^Θ the one defined over the decision space. The fusion performed using Bayes' rule is defined as:

$$p_t^\Theta(\theta|\phi_j \wedge \phi_k) \text{ for } \theta \in \{A, B\} \quad (4)$$

for two pieces on singular information. We have thus under the independence assumption[2]:

$$\begin{aligned} &p_t^\Theta(\theta|X_{TA} = FV, X_{TB} = \neg FV) \propto \\ &p_0^X(X_{TA} = FV|\theta)p_0^X(X_{TB} = \neg FV|\theta)p_0^\Theta(\theta) \end{aligned} \quad (5)$$

[2]Meaning that the two sources do not share any information and that the statement of one does not impact the statement of the other.

TABLE II

EXAMPLES OF *input* PIECES OF INFORMATION (POIS) AND THEIR CHARACTERISATION. THE POI ID IN THE FIRST COLUMN IS USED FOR REFERENCE IN THE TEXT. POIS IN BLUE CELLS WILL BE USED IN THE COMPARISON IN SECTION IV. THE POI IN THE RED CELL IS THE *output* OF THE FUSION ALGORITHM.

POI #	Description	Source	Piece of information	Type of information		Uncertainty support	Type of imperfection	Uncertainty nature	Uncertainty derivation		Uncertainty representation notation
ϕ_a	Prior belief about the location of the missing vessel	Decision maker / user	"It is likely that the missing vessel is in A"	Generic	Professional knowledge	Θ	Uncertainty, Imprecision	Epistemic	Subjective	Belief-based	$\eta_{0,dm}^{\Theta}$
ϕ_b	Mapping $\mathcal{X}_T \to \Theta$	Logic	if V_A is not a fishing vessel then B	Generic	Commonsense / Logic	$(\mathcal{X}_T, \Theta)$	Imprecision	Epistemic	Objective	Classical	$\eta_{0,.}^{X,\Theta}$
ϕ_c	ATR's reliability in fishing vessel's classification	Training dataset #2	Confusion matrix (see Table I)	Generic	Statistical knowledge	$(\mathcal{X}, \mathcal{X}^*)$	Uncertainty	Aleatory	Objective	Frequentist	$\eta_{0,DS2}^{X,X^*}$
ϕ_d	Cargo captain's reliability in fishing vessel's identification	Decision maker's estimation	Confusion matrix (see Table I)	Generic	Professional knowledge	$(\mathcal{X}, \mathcal{X}^*)$	Uncertainty	Epistemic	Subjective	Belief-based	$\eta_{0,dm}^{X,X^*}$
ϕ_e	Ratio of fishing vessels for the given area over the last two months	AIS dataset #1	35%	Generic	Statistical knowledge	$\mathcal{X}_T$	Imprecision	Aleatory	Objective	Frequentist	$\eta_{0,AIS}^{X}$
ϕ_f	Speed distribution for fishing vessels	AIS dataset #1	$P(X_S \vert X_T = FV)$	Generic	Statistical knowledge	$(\mathcal{X}_S, \mathcal{X}_T)$	Uncertainty	Aleatory	Objective	Frequentist	$\eta_{0,AIS}^{X}$
ϕ_g	Joint distribution of length and type of vessels	Lloyds	$P(X_S, X_T)$	Generic	Statistical knowledge	$(\mathcal{X}_l, \mathcal{X}_T)$	Uncertainty	Aleatory	Objective	Frequentist	$\eta_{0,Ll}^{(X_1,X_2)}$
ϕ_h	Motion model for a vessel in function of speed and heading	NA	$P_t = f(P_0, v, t)$	Generic	Laws of physics	$(\mathcal{X}_P, \mathcal{X}_S, \mathcal{X}_T)$	Uncertainty	Aleatory	Objective	Propensity	$\eta_{0,.}^{X_1,X_2,X_3}$
ϕ_i	Speed of Vessel A observed	Radar + tracker	5.6 knots	Singular	Measurement data	$\mathcal{X}_S$	None	-	-	-	-
ϕ_j	Type of vessel A observed	IR camera + ATR	Score vector $[0.7; 0.3]$	Singular	Measurement-based	$\mathcal{X}_T$	Uncertainty	Epistemic	Objective	Frequentist	$\eta_{t,ATR}^{X}$
ϕ_k	Type of vessel B observed	Cargo captain	"Vessel B is likely NOT a fishing vessel"	Singular	Observation	$\mathcal{X}_T$,$\eta_{t,CC}^{X}$	Uncertainty, Imprecision	Epistemic	Subjective	Belief-based	$\eta_{t,CC}^{X}$
ϕ_l	Speed of Vessel A observed	IR camera analyst	"Vessel A is quite fast (for a fishing vessel)"	Singular	Observation	$\mathcal{X}_S$	Imprecision	Epistemic	Subjective	Belief-based	$\eta_{t,cam}^{X}$
ϕ_o	Estimated location of the missing vessel	Fusion algorithm	See Table IV	Singular	Measurement-based	Θ	Uncertainty, Imprecision	?	?	?	$\eta_{t,dm}^{\Theta}$

Bayes' rule allows to compute the posterior probability of an event (here A or B) given some observations, considered as certain, based on generic uncertainty. The uncertainty of the source at the time of the observation (*e.g.* the cargo captain is not fully certain that Vessel B is not a fishing vessel) is not considered, and we have rather $p_{t,CC}^{X}(X_{TB} = \neg FV) = 1$. The corresponding pieces of information could be denoted by $\bar{\phi_j} = (X_{TA} = FV)$ and $\bar{\phi_k} = (X_{TB} = \neg FV)$. Rather, prior uncertainty is used. The question is which values should be included in (5) as there are many ways to derive these values:

- $p_0^{\Theta}(\theta)$ is the prior probability regarding both events A and B, before we receive any further information. It can be estimated from (at least) (1) the base rates, that could be the ratio of fishing vessels observed in the area over all the vessels, or the ratio of fishing vessels of size and subtype of the missing vessel over all the fishing vessels (ϕ_e); or (2) the decision maker subjective belief about both events (ϕ_a). As noticed in [26], because prior probabilities refer to *subjective* states of belief, they "may be influenced by base rates and any other information available to the decision maker", and may thus differ from base rates. As we see, the prior may have a huge impact in the estimation of the final belief, and as a consequence its elicitation is crucial. A cautious approach is to consider a uniform prior over Θ as $p_0^{\Theta}(A) = p_0^{\Theta}(B) = \frac{1}{2}$, interpreted either as a *total ignorance* of the decision maker or an *equal chance* assumption. Probability theory is not able to distinguish between these two notions.
- $p_0^{X}(\bar{\phi_j}|\theta)$ and $p_0^{X}(\bar{\phi_k}|\theta)$, for $\theta \in \Theta$, are two likelihood functions, *i.e.* the probability that the source (*i.e.* the ATR algorithm or the cargo captain) declares FV (resp. $\neg FV$) given that the missing vessel is Vessel A (resp. Vessel B). Because we know that the missing vessel is a fishing vessel, we can derive these probabilities from the confusion matrices of the two sources (Table I).

However, the observations FV or $\neg FV$ relate to the hypotheses A and B only through the multivalued mappings (3). Indeed, although $p_0^{X}(\bar{\phi_j}|A)$ and $p_0^{X}(\neg\bar{\phi_j}|A)$ can be respectively the true positive and false positive rates for the ATR, $p_0^X(\bar{\phi_j}|B) = p_0^X(\bar{\phi_j}|\top)$ and $p_0^X(\neg\bar{\phi_j}|B) = p_0^X(\neg\bar{\phi_j}|\top)$ where $\top = FV^* \vee \neg FV^*$ is the tautology. Other said, if the source is not observing the missing vessel, we thus do not know which vessel the source is observing and we should consider only the number of fishing vessels detections regardless the ground truth. We have for instance $p_0^X(\bar{\phi_j}|B) = \frac{0.8+0.4}{2} = 0.6$. See for instance [8] for a detailed computation of the likelihoods. The resulting likelihoods are given in Table III.

TABLE III
LIKELIHOODS EXTRACTED FROM ϕ_c AND ϕ_d FOR $\bar{\phi_j} = (X_{TA} = FV)$ AND $\bar{\phi_k} = (X_{TB} = \neg FV)$.

	ATR		Cargo Capt.	
$p(\bar{\phi}\|\theta)$	$\bar{\phi_j}$	$\neg\bar{\phi_j}$	$\neg\bar{\phi_k}$	$\bar{\phi_k}$
A	0.8	0.2	0.52	0.48
B	0.6	0.4	0.95	0.05

Fusing these pieces of information with Bayes' rule gives thus (with equal priors):

$$\begin{cases} p_{t,F}^{\Theta}(A|\phi_j,\phi_k) \propto 0.8 \times \frac{0.05+0.9}{2} = 0.38 \\ p_{t,F}^{\Theta}(B|\phi_j,\phi_k) \propto \frac{0.8+0.4}{2} \times 0.05 = 0.03 \end{cases} \tag{6}$$

which after normalisation gives:

$$p_{t,F}^{\Theta}(A|\phi_j,\phi_k) = 0.92 \text{ and } p_{t,F}^{\Theta}(B|\phi_j,\phi_k) = 0.08$$

Thus because the Cargo Captain is almost never wrong when he declares that he does not observe a fishing vessel, and because his epistemic uncertainty (although quite high) at the time of his declaration is not considered, then the fusion result is that the missing vessel is in A, with an extremely high certainty.

If we consider the prior ϕ_a that $p_{0,dm}^{\Theta}(A) = 0.6$ (thus $p_{0,dm}^{\Theta}(B) = 0.4$), then the result would be:

$$p_{t,F}^{\Theta}(A|\phi_j,\phi_k) = 0.97 \text{ and } p_{t,F}^{\Theta}(B|\phi_j,\phi_k) = 0.03$$

increasing thus the conclusiveness of the result.

Another way to apply Bayes' rule to fusion considers that each source provides a posterior probability at t, $p_t^{\Theta}(\theta|\phi)$ (*e.g.* [9]). In this case, and if we still consider equal priors, the result is equivalent to apply Dempster's rule with Bayesian BPAs [27]. The results are given in Table IV, in the second column of the Bayes' rule.

B. Belief functions and Dempster's rule

Let denote by m^X the Basic Probability Assignment (BPA) or mass function defined over the domain of variable X in the measurement space and by m^{Θ} the one defined over the decision space. The fusion of ϕ_j and ϕ_k by Dempster's rule is defined as:

$$m_{t,F}^{\Theta} = m_{t,ATR}^{\Theta} \oplus m_{t,CC}^{\Theta} \tag{7}$$

where $m_{t,ATR}^{\Theta}$ and $m_{t,CC}^{\Theta}$ are the singular information at t from the ATR and Cargo Captain respectively. The ATR outputs a score vector of $[0.7; 0.3]$ over $\mathcal{X}_{TA}$ which is directly interpreted as a probability distribution. The Cargo Captain expresses his uncertainty using the term "likely" that we translate as a score of 0.6 on a numerical scale[3]. The probability distribution $p_{t,ATR}^{X_{TA}}$ and $p_{t,ATR}^{X_{TB}}$ being transferred to Θ through the multivalued mapping (rules ϕ_b, (3)), we obtain the two following BPAs:

$$\begin{cases} m_{t,ATR}^{\Theta}(B) = 0.3 \\ m_{t,ATR}^{\Theta}(\Theta) = 0.7 \end{cases} \quad \begin{cases} m_{t,CC}^{\Theta}(A) = 0.6 \\ m_{t,CC}^{\Theta}(\Theta) = 0.4 \end{cases} \tag{8}$$

which clearly express our total ignorance when the observed vessel is not the missing vessel. The result of the combination is then:

$$\begin{cases} m_{t,F}^{\Theta}(A) \propto 0.6 \times 0.7 = 0.42 \\ m_{t,F}^{\Theta}(B) \propto 0.3 \times 0.4 = 0.12 \\ m_{t,F}^{\Theta}(\Theta) \propto 0.7 \times 0.4 = 0.28 \end{cases} \tag{9}$$

with a conflict of 0.18. The pignistic probabilities for A and B are given in Table IV. This direct application of Dempster's rule does not consider the sources' quality. However, this is classically done by discounting ϕ_j and ϕ_k according to the quality of their respective sources. Using discounting factors α_{ATR}^{FV} and $\alpha_{CC}^{\neg FV}$ of 0.8 and 0.9 as extracted from the respective confusion matrices, the discounted BPAs are:

$$\begin{cases} m_{t,ATRd}^{\Theta}(B) = m_{t,ATR}^{\Theta}(B)(1-\alpha_{ATR}^{FV}) = 0.24 \\ m_{t,ATRd}^{\Theta}(\Theta) = m_{t,ATR}^{\Theta}(\Theta)\alpha_{ATR}^{FV} + (1-\alpha_{ATR}^{FV}) = 0.76 \end{cases} \tag{10}$$

$$\begin{cases} m_{t,CCd}^{\Theta}(A) = m_{t,CC}^{\Theta}(B)(1-\alpha_{CC}^{\neg FV}) = 0.54 \\ m_{t,CCd}^{\Theta}(\Theta) = m_{t,CC}^{\Theta}(\Theta)\alpha_{CC}^{FV} + (1-\alpha_{C}^{\neg FV}) = 0.46 \end{cases} \tag{11}$$

The result of their combination is given in Table IV, with a conflict of 0.13.

TABLE IV
SEMANTIC COMPARISON OF BAYES' AND DEMPSTER'S RULES - $\times$ MEANS THAT THE UNCERTAINTY IS CONSIDERED WHILE $\circ$ MEANS THAT IT IS NOT.

Info.	Uncertainty	Prob. & Bayes		BF & Dempster	
		Likel.	Post.	Direct	Disc.
Input					
ϕ_a	$\eta_{0,dm}^{\Theta}$	$\times$	$\circ$	$\circ$	$\circ$
ϕ_b	$\eta_{0,.}^{X,\Theta}$	$\times$	$\circ$	$\times$	$\times$
ϕ_c	$\eta_{0,DS2}^{X}$	$\times$	$\circ$	$\circ$	$\times$
ϕ_d	$\eta_{0,dm}^{X}$	$\times$	$\circ$	$\circ$	$\times$
ϕ_j	$\eta_{t,ATR}^{X}$	$\circ$	$\times$	$\times$	$\times$
ϕ_k	$\eta_{t,CC}^{X}$	$\circ$	$\times$	$\times$	$\times$
Output					
ϕ_o	$\eta_{t,F}^{\Theta}(A)$	0.92	0.77	0.77	0.78
	$\eta_{t,F}^{\Theta}(B)$	0.08	0.23	0.23	0.22
[Bel; Pl]	$\eta_{t,F}^{\eta_t^{\Theta}}$ - A	$\circ$	$\circ$	$[0.51; 0.85]$	$[0.47; 0.87]$
	$\eta_{t,F}^{\eta_t^{\Theta}}$ - B	$\circ$	$\circ$	$[0.15; 0.49]$	$[0.13; 0.53]$

Table IV summarises the semantic comparison of the two rules regarding the information used along both the uncertainty supports and the type of information (generic vs singular). This table highlights known results, that Bayes' rule is not dedicated to process "probable knowledge" [7] but rather exploits generic information about the sources' quality for instance (for likelihood estimations) and priors. Dempster's rule is dedicated to combine uncertain testimonies [13] and

[3]Note that how to translate uncertainty quantifiers into numerical scale is still an open problem.

uses singular uncertainty at t from the sources. The sources' quality is exploited in Dempster's rule through a discounting operation, which results in a higher uncertainty at the output. We also highlight the high conclusiveness of Bayes' rule in this particular case where one source is assumed to be highly reliable. Dempster's rule not only provides less a conclusive output, but the belief-plausibility interval is an estimation of the second-order uncertainty, that Bayes' rule does not provide. Note that even if the BPAs are not Bayesian, the direct application of Dempster's rule with a pignistic probability decision still gives the same values than Bayes' rule with posterior probabilities. This is because the BPAs are simple support function focused on singletons.

The information (and uncertainty) output by the fusion algorithm $\eta^{\Theta}_{t,F}$ is denoted as ϕ_o in Table II. It can be characterised along the same dimensions than the input information, thinking that it may be used by another fusion node. It is not clear however neither how its uncertainty nature should be characterised, nor how should be its uncertainty derivation.

V. Conclusions

The emergence of alternative approaches to probability theory to handle uncertainty has provided a richer choice to the scientists but still with the lack of tools to objectively select the dedicated approach to solve their problem at hand. We illustrating a possible use of the URREF on a practical maritime use case, in which the designer needs to decide which uncertainty representation and reasoning technique for fusion should be used. We revisited some classical works comparing Bayes' and Dempster's rule for fusion through the URREF and highlight the connection between the different elements of the ontology. We focused on the underlying semantics of the two rules discussed in the literature and identify semantic criteria and elements which could complete the URREF ontology. We provided a detailed categorisation of input information along the dimensions of Uncertainty nature, Uncertainty derivation, Information type and Type of imperfection which shows the importance of the derivation of uncertainty values, as it has a direct impact on the interpretation of the output uncertainty. We stress how the uncertainty supports (variables, links between variables, uncertainty expression) crossed with the type of information (generic versus singular) help in clarifying that Dempster's rule does not use generic knowledge but only uncertain singular information, while Bayes' rule relies on generic information. The interpretation of the output of the fusion algorithm (uncertainty derivation and nature) may still be subject to discussion.

Acknowledgments

The author wishes to thank the NATO Allied Command Transformation (NATO-ACT) for supporting the CMRE project on the Data Knowledge and Operational Effectiveness / Maritime Situational Awareness programme.

References

[1] D. Dubois and H. Prade, *Formal representations of uncertainty*, vol. Decision-making - Concepts and Methods, ch. 3, pp. 85–156. ISTE, London, UK & Wiley, Hoboken, N.J. USA, 2009. Invited paper.

[2] P. Smets, "Imperfect information: Imprecision - uncertainty," in *Uncertainty Management in Information Systems. From Needs to Solutions* (A. Motro and P. Smets, eds.), pp. 225–254, Kluwer Academic Publishers, 1997.

[3] R. Y. Wang and D. M. Strong, "Beyond accuracy: What data quality means to data consumers," *Journal of Management Information Systems*, vol. 12, no. 4, pp. 5–34, 1996.

[4] A.-L. Jousselme, P. Maupin, and E. Bossé, "Uncertainty in a situation analysis perspective," in *Proceedings of the 6th Annual Conference on Information Fusion*, (Cairns, Australia), pp. 1207–1214, July 2003.

[5] D. Ongaro and J. Ousterhoutand, "In search of an understandable consensus algorithm," in *USENIX Annual Technical Conference*, (Philadelphia, PA), pp. 305–319, USENIX Association, June 2014.

[6] P. C. G. Costa, K. Laskey, E. Blasch, and A.-L. Jousselme, "Towards Unbiased Evaluation of Uncertainty Reasoning: The URREF Ontology," in *Proc. of the 15th Int. Conf. on Information Fusion*, (Singapore), 2012.

[7] P. Diaconis and S. Zabell, "Some alternative to Bayes's rule," in *Proc. of the Second University of California, Irvine, Conference on Political Economy* (B. Grofman and G. Owen, eds.), pp. 25–38, 1986.

[8] H. Leung and J. Wu, "Bayesian and Dempster-Shafer target identification for radar surveillance.," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 36, no. 2, pp. 432–447, 2000.

[9] S. Challa and D. Koks, "Bayesian and Dempster-Shafer fusion," *Sādhanā*, vol. 29, pp. 145–176, April 2004.

[10] G. Shafer and A. Tversky, "Languages and designs for probability judgment," *Cognitive Science*, vol. 9, pp. 309–339, 1985.

[11] J. Pearl, "Reasoning with belief functions: an analysis of compatibility," *Int. Journal of Approximate Reasoning*, vol. 4, pp. 363–389, 1990.

[12] D. Dubois and H. Prade, "Evidence, knowledge, and belief functions," *Int. Journal of Approximate Reasoning*, vol. 6, pp. 295–319, 1992.

[13] G. Shafer, *A Mathematical Theory of Evidence*. Princeton University Press, 1976.

[14] B. R. Cobb and P. Shenoy, "A comparison of bayesian and belief function reasoning," *Information Systems Frontiers*, vol. 5, no. 4, pp. 345–358, 2003.

[15] B. Ristic and P. Smets, "Target classification approach based on the belief function theory," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 41, no. 2, pp. 574–583, 2005.

[16] A.-L. Jousselme, G. Pallotta, and J. Locke, "A Risk Game to study the impact of information quality on human threat assessment and decision making," Tech. Rep. CMRE-FR-2015-009, NATO STO Centre for Maritime Research and Experimentation, 2015.

[17] A.-L. Jousselme and P. Maupin, "Uncertainty representations for a vehicle-borne IED surveillance problem," in *Proc. of the Int. Conf. on Information Fusion*, (Singapore), July 2012.

[18] J. P. de Villiers, K. Laskey, A.-L. Jousselme, E. Blasch, A. de Waal, G. Pavlin, and P. Costa, "Uncertainty representation, quantification and evaluation for data and information fusion," in *Proc. of the 18th Int. Conf. on Information Fusion*, (Washington, D. C. (USA)), July 2015.

[19] A.-L. Joussselme and P. Maupin, "A brief survey of comparative elements for *uncertainty calculi* and decision procedures assessment," in *Proc. of the 15th Int. conf. on Information Fusion*, 2012. Panel *Uncertainty Evaluation: Current Status and Major Challenges*.

[20] A.-L. Jousselme and G. Pallotta, "Dissecting uncertainty-based fusion techniques for maritime anomaly detection," in *Proc. of the 18th Int. Conf. on Information Fusion*, (Washington, D.C., USA), July 2015.

[21] D. Dubois and T. Denœux, "Conditioning in Dempster-Shafer theory: prediction vs. revision," in *Proc. of the Conference on Belief Functions (BELIEF 2012)* (T. Denoeux and M.-H. Masson, eds.), vol. Belief Functions: Theory and Applications 164 of *Advances in Intelligent and Soft Computing*, (Compiègne, France), pp. 385–392, 2012.

[22] G. J. Klir, "Generalized information theory," *Fuzzy Sets and Systems*, vol. 40, pp. 127–142, 1991.

[23] A. Der Kiureghian and O. Ditlevsen, "Aleatory or epistemic? Does it matter?," *Structural Safety*, vol. 31, pp. 105–112, 2009.

[24] S. S. Stevens, "On the theory of scale of measurement," *Science*, vol. 103, pp. 677–680, June 1946. New Series.

[25] A.-L. Jousselme, A.-C. Boury-Brisset, B. Debaque, and D. Prévost, "Characterization of hard and soft sources of information: A practical illustration," in *Proceedings of the International Conference of Information Fusion*, (Salamanca, Spain), 2014.

[26] J. J. Koehler, "The base rate fallacy reconsidered: Descriptive, normative, and methodological challenges," *Bahavioral and brain sciences*, vol. 19, no. 1, pp. 1–53, 1996.

[27] P. Smets, "Analyzing the combination of conflicting belief functions," *Information Fusion*, vol. 8, pp. 387–412, 2007.