URREF: Uncertainty representation and reasoning evaluation framework for information fusion

PAULO C. G. COSTA
ANNE-LAURE JOUSSELME
KATHRYN B. LASKEY
ERIK BLASCH
VALENTINA DRAGOS
JUERGEN ZIEGLER
PIETER DE VILLIERS
GREGOR PAVLIN

Uncertainty management is a key aspect of any information fusion (IF) system. Evaluation of how uncertainty is dealt with within a given IF system is distinct from, although closely related to, evaluation of the overall performance of the system. This paper presents the Uncertainty Representation and Reasoning Evaluation Framework (URREF), which is developed by the ISIF Evaluation of Techniques for Uncertainty Representation Working Group (ETURWG) for evaluating the uncertainty management aspects of IF systems. The paper describes the scope of the framework, its core element—the URREF ontology, the elementary fusion process it considers, and how these are related to the subjects being evaluated using the framework. Although material about the URREF has been previously published elsewhere, this work is the first to provide a comprehensive overview of the framework, establishing its scope, core elements, elementary fusion process considered, and relationship between these and the subjects they are designed to evaluate. We also briefly describe a few use cases of the framework, discussing how URREF can be applied in their evaluation.

Manuscript received October 22, 2018; revised May 5, 2019; released for publication May 13, 2019.

Refereeing of this contribution was handled by Sten Andler.

Authors' addresses: P. C. G. Costa, K. B. Laskey, George Mason University, Fairfax, VA, USA (E-mail: {pcosta,klaskey}@gmu.edu). A.-L. Jousselme, NATO STO Centre for Maritime Research and Experimentation, La Spezia, IT (E-mail: anne-laure.jousselme@cmre.nato.int). E. Blasch, Air Force Office of Scientific Research, Arlington, VA, USA (E-mail: erik.blasch.1@us.af.mil). V. Dragos, ONERA—The French Aerospace Lab, Palaiseau, France (E-mail: valentina.dragos@onera.fr). J. Ziegler, Competence Centers ISR, IABGmbH, Ottobrunn, Germany (E-mail: jziegler@iabg.de). P. de Villiers, University of Pretoria, Pretoria, South Africa, Council for Scientific and Industrial Research, Pretoria, South Africa (E-mail: pieter.devilliers@up.ac.za). G. Pavlin, D-CIS Lab, Thales Research and Technology, Delft, The Netherlands (E-mail: gregor.pavlin@d-cis.nl).

I. INTRODUCTION

Evaluating how well an Information Fusion (IF) system performs requires defining the relevant criteria to be assessed and testing the IF system's fusion algorithm, data model, and architecture against that criteria. Empirical evaluation techniques are effective when assessing the latter two, but face a major limitation when addressing the former. More specifically, they often require embedding some uncertainty representation and its associated reasoning scheme within the fusion method, which serves as an enabler and becomes often the subject of evaluation itself. Inherently, it is not a trivial problem to isolate the uncertainty representation from either its reasoning scheme or the fusion algorithm, which prevents an effective assessment of the IF system since current methods cannot capture the impact of these in the overall IF system's performance. The work described in this paper focuses on addressing this limitation, providing a principled method for evaluating how the uncertainty representation and reasoning aspects of an Information Fusion impact its overall performance.

IF applications typically must deal with information that is incomplete, imprecise, inconsistent and otherwise in need of a sound methodology for representing and managing uncertainty. Complex and dynamic use cases make such tasks even more difficult, as apparently minor differences in how uncertainty is handled may drastically affect the output of the IF process. In short, it is fair to state that uncertainty management is a key aspect in most—if not all—IF systems. Despite this importance, the IF community still does not have a standardized framework for evaluating how uncertainty is represented and managed in IF systems. IF systems typically perform uncertainty reasoning to achieve their goals, which means they would benefit from a framework to evaluate how well they are performing on it.

The lack of an uncertainty evaluation framework for IF systems tends to be more widely acknowledged at higher levels of the Joint Directors of Laboratories (JDL) model [1]–[3]. More specifically, Low-Level Information Fusion (LLIF) systems (i.e., below JDL level 2) tend not to represent semantics explicitly. Semantics is commonly understood among theoreticians and algorithm developers, and is typically implicitly encoded in algorithms through devices such as variable naming conventions. LLIF systems tend to rely exclusively on probability theory as the paradigm for uncertainty representation and reasoning. This is justified by the typically large amount of available data, which justifies the use of statistical models to address the fusion problems at hand. Tools and techniques for evaluating probabilistic inference systems are well-understood. In contrast, because of the complexity and variety of semantic categories for High-Level Information Fusion (HLIF), applications usually require making semantics explicit and accessible to formal reasoning tools. Furthermore, HLIF systems make use of a variety of theories and methods to represent and reason with uncertainty. For example, deciding whether three different radars receiving echos from the same location are seeing one, two, or three tracks is a problem for which uncertainty is well understood and for which standard evaluation methods are well established. On the other hand, deciding whether the incoming fighter formation poses a danger to a radar installation may involve a multiplicity of sources of uncertainty, and may require consideration of complex semantic concepts such as enemy doctrine, the spatial configurations associated with hostile and innocuous formations, how danger should be defined, and the like.

Uncertainty analysis is even more critical for systems relying on multiple types of data and different uncertainty paradigms. Soft data is unstructured and intrinsically ambiguous [4], and tracking its uncertainty [5] often requires explicit semantics [6]. Heterogeneous fusion combines data of different natures, and uncertainty propagation for heterogeneous fusion still lacks a well-established and widely agreed upon theoretical foundation [7].

Clearly, a system that can reason about these and other HLIF problems must consider complex semantics, and may be required to employ multiple uncertainty formalisms (e.g., a fuzzy membership function might be used to transform verbal danger categories into a quantitative representation, which might be combined with a probability distribution on events leading to different levels of danger). The design of a HLIF system would definitely benefit from an uncertainty evaluation framework that would guide the selection of the most suitable uncertainty representation and reasoning technique. An ability to compare uncertainty handling approaches would enable exploitation of semantically rich representations to help assess its performance when facing an uncertain input. With the emergence of alternative uncertainty theories in addition to probabilities (see for instance [8] for a survey) came the question of which approach is the best suited for uncertainty handling in a specific problem setting. The question has been addressed both theoretically (e.g., [9]–[12]) and in practical implementation of fusion solutions (e.g., [13]–[15]). Handling uncertainty in fusion problems is indeed a major challenge for 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 varieties, etc. Acknowledging the existence of different types or facets of information quality provides partial answers (e.g., [16]–[18]). Nevertheless a deep understanding of the different uncertainty representation and reasoning techniques, their underlying mathematical frameworks, and associated hypotheses and semantics, is necessary to guide a fusion system's designer in making informed choices about the most suitable technique to the problem at hand. Such a deep understanding provides clearer explanations of the algorithms to the user for an improved synergy between the human and the machine [19].

The International Society of Information Fusion (ISIF) recognized this problem, and created a working group to address it. The ISIF Evaluation of Techniques for Uncertainty Representation Working Group (ETURWG) [20], [21] was created in the ISIF Board of Directors meeting just after the Fusion 2011 conference (Chicago, IL, USA) to specifically address this issue. The ETURWG's main goals are (1) to establish features required for any quantitative uncertainty representation to support the exchange of soft and hard information in a net-centric environment; (2) to develop a set of use cases involving information exchange and fusion requiring reasoning and inference under uncertainty; and (3) to define evaluation criteria supporting principled comparisons among different approaches applied to the use cases. As of this writing, the group has convened 104 general meetings spanning its 7 years of activities, and resulted in 43 peer-reviewed articles on the subject. The group's website[1] provides comprehensive information about its activities, including agendas and minutes of the meetings, datasets used, documentation on case studies and discussions, as well as a large amount of information related to the research efforts by the group.

This paper provides an overview of the Uncertainty Representation and Reasoning Evaluation Framework (URREF). It not only updates but also substantially enhances a similar paper published in the Proceedings of the Fusion 2012 conference [22]. After this brief introduction, Section II provides an overview of recent and current efforts in evaluating uncertainty in IF systems. Section III introduces the framework, which supports assessment of the impact of uncertainty representation on a fusion system. This is followed by a section covering the relationship between the framework elements and the subjects it is evaluating. Section V presents a brief description of case studies applying the framework. The final section contains discussion and conclusion.

II. EVALUATING UNCERTAINTY IN FUSION SYSTEMS

The evaluation of how uncertainty is dealt with within a given IF system is distinct from, although closely related to, the evaluation of the overall performance of the system [23], [24]. Figure 1 shows the elements of a generic IF model. The figure distinguishes between processes associated with low-level and high-level IF, a distinction dating to the seminal fusion model developed by the Joint Directors of Laboratories (JDL) [1]–[3]. Evaluation criteria and associated metrics for the overall system include the effects of the uncertainty representation, but there are also effects of other aspects of the fusion system that can affect the performance of the system. These are more encompassing in scope than

[1] http://eturwg.c4i.gmu.edu, free registration required for full access

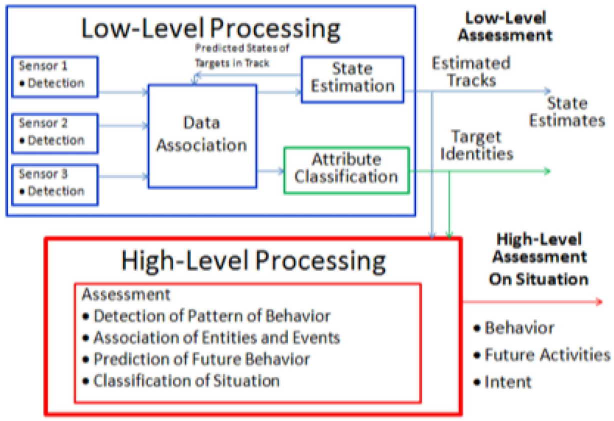


Fig. 1. The principal processing components of the IF process include both high level and low level processing components. Low level fusion processes include detection, association, state estimation and attribute classification, whereas high level fusion processes include behavioural pattern estimation, association, behaviour prediction and situation classification.

those focused on the uncertainty handling within the system. Metrics focused on uncertainty handling should address the contribution of uncertainty handling to the overall system performance.

For example, fusion-system-level metrics include timeliness, accuracy and confidence. Clearly, different choices in uncertainty representation approaches will affect the achievable timeliness, accuracy, and confidence of a system, and therefore must be considered when evaluating both the system's performance as a whole and the specific impact of the uncertainty handling approach. Yet, when evaluating timeliness (or any other system-level metrics), one will likely find some factors not directly related to the handling of uncertainty itself, such as object tracking and classification report updates (i.e., Level 1 fusion), situation and threat assessment relative to scenario constraints (i.e., Level 2/3 fusion), overall system architecture (e.g., centralized, distributed, etc.), data management processes and feedback/input control processes (i.e., Level 4 fusion considerations), and user-machine coordination based on operating systems (i.e., Level 5 fusion), and others.

The IF community envisions effortless interaction between humans and computers, seamless interoperability and information exchange among applications, and rapid and accurate identification and invocation of appropriate services. As the complexity of fusion solutions grows, we end up with a mixture of components handling different types of uncertainties, often by using different methods.

Here, the term "uncertainty" is intended to encompass a variety of aspects of imperfect knowledge, including incompleteness, inconclusiveness, vagueness, ambiguity, and others. The term "uncertainty reasoning" is meant to denote the full range of methods designed for representing and reasoning with knowledge when approaches based on Boolean algebra (e.g. propositional logic) are not applicable (e.g. when Boolean truth-values are unknown, unknowable, or inapplicable. Commonly applied approaches to uncertainty reasoning include probability theory, fuzzy logic, subjective logic, Dempster-Shafer theory, DSmT, and numerous other methodologies.

The problem of representing and reasoning with complex and heterogeneous data was addressed by a working group of the World Wide Web Consortium [25]. The working group's findings are relevant to the challenge considered in this paper. Information fusion under uncertainty is an intrinsic requirement for many of the problems in the World Wide Web domain. A full realization of the World Wide Web as a source of processable data and services demands formalisms capable of representing and reasoning under uncertainty.

- Automated agents are used to exchange Web information that in many cases is not perfect. Thus, a standardized format for representing uncertainty would allow agents receiving imperfect information to interpret it in the same way as were intended by the sending agents.
- Data often are intrinsically uncertainty-laden. Examples include weather forecasts or gambling odds. Canonical methods for representing and integrating such information are necessary for communicating it in a seamless fashion.
- Non-sensory collected information is also often incorrect or only partially correct, raising concerns related to trust or credibility. Uncertainty representation and reasoning helps to resolve tension amongst information sources having different confidence and trust levels.
- Dynamic composability of Web Services will require runtime identification of processing and data resources and resolution of policy objectives. Uncertainty reasoning techniques may be necessary to resolve situations in which existing information is not definitive.
- Information extracted from large information networks such as the World Wide Web is typically incomplete. The ability to exploit partial information is very useful for identifying sources of service or information. For example, that an online service deals with greeting cards may be evidence that it also sells stationery. It is clear that search effectiveness could be improved by appropriate use of technologies for handling uncertainty.

These problems all require IF, both low and high level. They bear an obvious relationship to the kinds of problems found in the sensor, data, and IF domain.

III. UNCERTAINTY REPRESENTATION AND REASONING FRAMEWORK

This section describes an evaluation framework to support assessment of how the choice of uncertainty

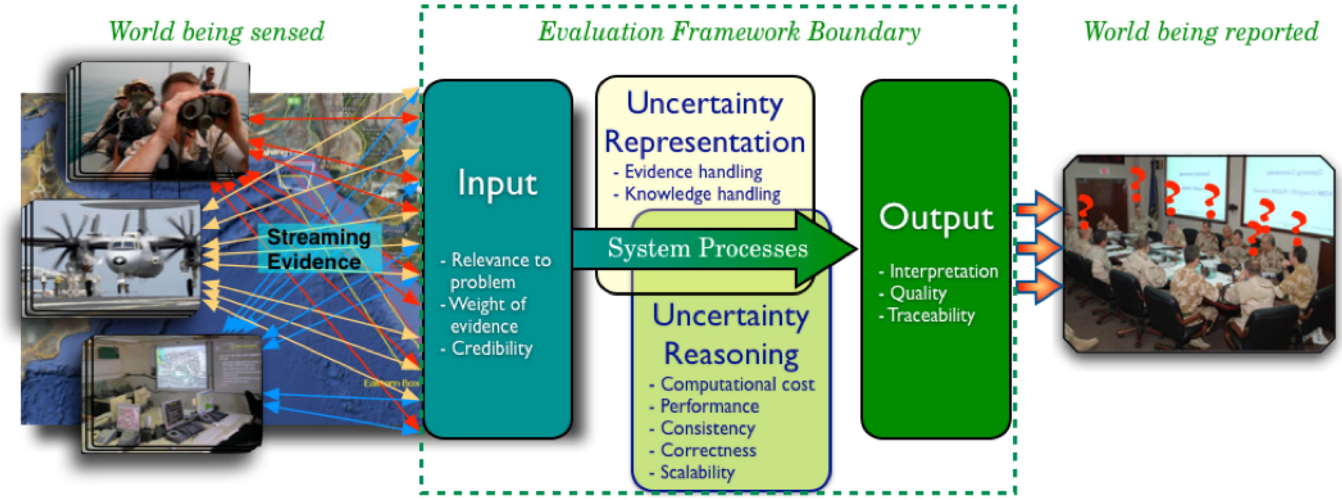


Fig. 2. URREF Boundary. This figure depicts the world being sensed on the left, the role of uncertainty representation and reasoning within a fusion system in the center, and the world being perceived on the right. The evaluation framework boundary encompasses the fusion system input, uncertainty representation, uncertainty reasoning and the fusion system output. Everything inside the evaluation framework boundary is known as the Uncertainty Representation and Reasoning Framework (URREF). The uncertainty representation and uncertainty reasoning are the primary subjects of evaluation, whereas the input and output are secondary subjects of evaluation.

representation and reasoning impacts the performance of an IF system. The scope of the framework is the main focus of the first sub-section, which is followed by an overview of its main component, the URREF ontology. Finally, an elementary fusion process is presented, as a means to identify the primary evaluation subjects of the evaluation methodology envisioned for the framework.

A. The URREF Scope

The basic idea behind the framework is to analyze an abstract fusion system and define its input data and output products. In a hypothetical IF system of the future, the uncertainty representation approach would be "plug-and-playable." That is, one might run the system with a Bayesian approach, then switch to a Dempster-Shafer approach, and then a Fuzzy Random Set approach. Alternatively, one might use a combination of uncertainty reasoning methods, as best suited for different aspects of the problem. The input data are the same in each case, as are the output products (but not necessarily the specific contents of the output products). Figure 2 depicts the uncertainty representation and reasoning evaluation framework (URREF) and its role in the overall fusion process.

There are two elements in the picture that are exogenous to the evaluation framework, named in the picture as "World being sensed" and "World being reported." Between these two external elements, the boundary of the evaluation framework encompasses the way uncertainty is handled when data is input to the system, during the processes that occur within it, as well as when the final product is delivered to the IF system's users. The uncertainty representation and uncertainty reasoning are the primary subjects of evaluation, whereas the input and output are secondary subjects of evaluation.

The first external element refers to the events of interest to the IF system that happen in the world and are perceived by the system sources. Note that the implicit definition of sources in this case encompasses anything that can capture information and send it to the system. That is, both hard sources (e.g., imaging, radar, video, etc.) and soft sources (HUMINT reports, software alerts, etc.) are considered external to the evaluation system with respect to their associated sensorial capabilities, while the way they convey their information is within the scope of the system [24], [26], [27].

This reflects an important distinction between the evaluation of an IF system and the evaluation of its handling of uncertainty. To illustrate the distinction, consider the Input element in Figure 2. This element addresses the system's ability to represent uncertainty as an intrinsic part of the information being captured. As an example, information regarding trust of the input from a given sensor is important to evaluating how the overall system handles uncertainty, although it might not be as critical for its overall performance. A key question for evaluating uncertainty representation is what the uncertainty characteristics of the input data are, and how they affect the use of different uncertainty schemes. On the other hand, the format of the input might be important to evaluation of system interoperability, but is not included in Figure 2 because it does not relate to uncertainty handling. In general, the elements inside the evaluation framework boundary in the figure are important to evaluation of uncertainty handling, but not necessarily to evaluation of other aspects of fusion system performance. Likewise, elements that are critical to overall evaluation but not important to uncertainty handling are not included here.

In the ideal system model, having the appropriate data characteristics is critical. If the characteristics do not span the range of uncertainty techniques, then the model may not give meaningful results about the operationally significant differences between the techniques. Correctly identifying the desired input data characteristics will shape the future development of use cases and modeling data sets for those use case.

Once information is in the IF system, it will be processed to generate the system's deliverable that requires uncertainty characterization and reporting in the Output step. This process involves fusion techniques and algorithms that are directly affected by the uncertainty handling technique being used, as well as its impact on the system's inferential process. In this case, the URREF evaluation criteria focus on aspects that are specific to the way uncertainty is considered and handled within the fusion process. This is not an evaluation of the system's performance as a whole. We want to understand how the uncertainty representation affects system performance, and whether different uncertainty representation schemes are more or less robust to variations in the remaining parts of the IF system architecture. We want to focus specifically on the uncertainty representation aspects, and attempt, as best as possible, to separate those aspects from the overall system performance and architecture concerns.

After the information is fused and properly treated, then it is conveyed to the system's users. In Figure 2, these are represented by an image depicting decision-makers who would likely be supported by the IF system in their tasks. The URREF output step involves the assessment of how information on uncertainty is presented to the users and, therefore, how it impacts the quality of their decision-making process.

B. The URREF Ontology

The word "framework" in URREF's name reflects the conclusion we reached during the early ETURWG meetings, as we discussed how uncertainty in IF systems should be evaluated. From the very beginning, it became clear to us that we were not developing a tool to measure a set of metrics related to uncertainty in a given system. After all, because uncertainty is embedded in practically all aspects of the process, each application would have so many nuances that designing a "one-size-fits-all" evaluation tool would either be too specific for use in diverse IF systems, or too generic to be useful. In other words, we soon realized that what was needed to move the state-of-the-art in uncertainty evaluation was not a monolithic evaluation program or tool, but a set of standards, best practices, guidelines, and other development tools that provides coherent and consistent support for those tasked with evaluating uncertainty in information systems. We call this set an evaluation framework.

The reasons behind this view of URREF as a framework instead of a system, program, or tool, also implied that the diversity and complexity of the IF systems to be evaluated would require this framework to be flexible and adaptable enough to be used by developers with distinct backgrounds and requirements. We soon realized that defining common terminology was an enormous challenge, as a given term might have different meanings to different people, whereas a common idea might be given different names by different people. Designing a "mother of all evaluation taxonomies" was not an option, as it would be useless to various use cases, such as existing systems with already established semantics. Thus, when designing the framework we were naturally inclined to adopt ontology as a knowledge representation technique, as an ontology provides embedded support for reasoning and allows for explicit semantics that could be aligned, adapted, or reused when developing evaluation systems.

Designing an ontology for URREF proved to be a tall order though. Within the ETURWG we have people with distinct backgrounds, so it was natural to see some "semantic misalignment" regarding concepts such as data quality, accuracy, precision, etc. These differences in understanding proved to be challenging to deal with, but an accurate preview of the challenges that arise when using a framework that invariably includes concept definitions that may not fully match the views of different users. Not surprisingly, it took a considerable amount of time to arrive at a stable version of the URREF ontology, and while all in the group would prefer one or more specific concepts to be defined in a different way, the group agreed that the current version of the ontology is sufficient to support the evaluation of uncertainty in IF systems consistently and coherently. Most of the concepts used have been drawn from seminal work in related areas such as uncertainty representation (e.g., [27]–[35], evidential reasoning (e.g., [36]–[38]), and performance evaluation (e.g., [9], [39]–[41]). We now describe the main aspects of the URREF ontology, including its classes, properties, and key concepts. The reader would benefit from actually accessing the files, and even following the work of the ETURWG group. In addition to the information provided in the group's website, as indicated earlier in this paper, the ontology itself can be downloaded or opened directly from an ontology editor (e.g., Protégé [42]) via its official URL.[2] Alternatively, cloning the group's GitHub repository[3] would provide access to not only the current version of the ontology but also previous versions, references, and other related working documents.

Figure 3 depicts the main classes of the URREF ontology, which were identified by the ETURWG group as pertinent to the evaluation of uncertainty within an IF system. These classes represent concepts meant to be sufficient to support the design of evaluation processes that follow the same semantic constraints and

[2]http://eturwg.c4i.gmu.edu/files/ontologies/URREF.owl

[3]https://github.com/paulocosta-gmu/urref/tree/master

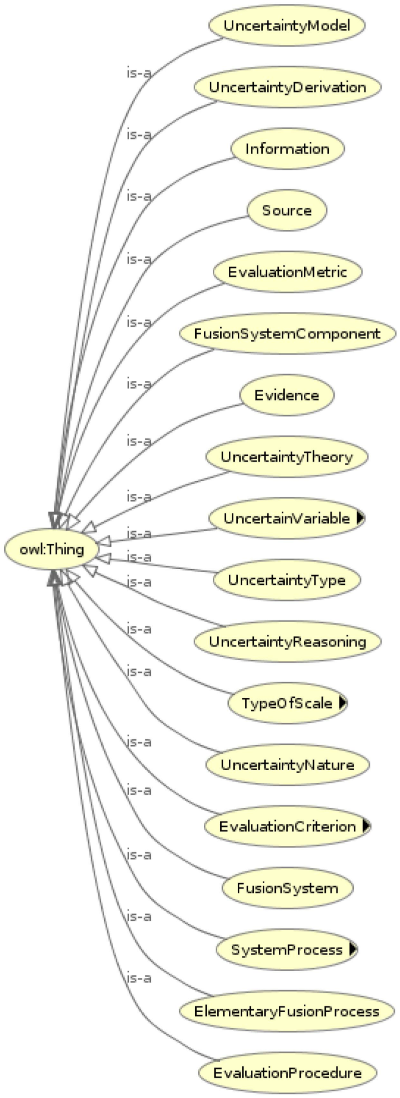


Fig. 3. Main classes of the URREF ontology

abide by the same principles of mathematical soundness. To emphasize the pragmatic aspect of the work of the ETURWG, it can be noted that these concepts capture the main aspects the group agreed upon when developing the use cases described in Section V. In fact, a brief comparison between these concepts and those of the first version of the ontology (cf. [22]) will show that many classes had to be added as a result of both the evolving discussions and the requirements elicited from the use cases.

The eighteen main classes of the URREF criteria focus on aspects that are specific to the way uncertainty is considered and handled within the fusion process. Figure 3 was built using the Protégé OWLviz plugin.[4] The classes are depicted as collapsed at the first level. Classes with a small black arrow head at the right have subclasses which can be shown in an expanded view. One example is the class *TypeOfScale*, which is depicted in its entirety in Figure 4. Its individuals correspond to

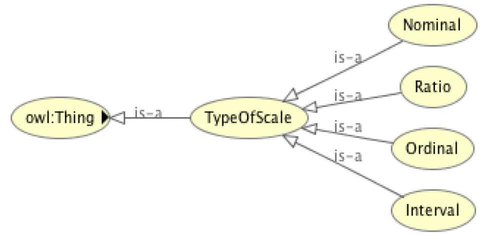


Fig. 4. URREF *TypeOfScale* class

specific scales used in quantifying the metrics employed when evaluating an IF system according to a given criteria, and its subclasses aggregate the types of quantification adopted. For instance, assume the precision of a given sensor (i.e., using the subclass *Precision* as evaluation criterion) would be evaluated using

$$u_{pre} = \sum_{t=1}^{n} L(r_t, a_t), \tag{1}$$

where n is the number of measurement trials, and L is a loss function with parameters r for reported value and a for actual value. In this case, the range of the loss function will dictate which type of scale should be used in that evaluation (e.g., a loss function returning a ratio between the two parameters would be classified under the associated type of scale). In the URREF framework, this class provides a way of mapping evaluation subjects and criteria chosen to the potential metrics and associated quantification types that can be used in a given evaluation.

While the type of scale defines how to quantify the metrics used to assess a given criterion in an evaluation, the *EvaluationMetrics* class defines what metric is being used (i.e., what is) the parameter being assessed. In the example of Eq. (1), the criterion being assessed is performance and the formula itself can be seen as the metric used to assess that criterion. Currently, the ontology only includes examples from NATO's Standardization Agreement 2511 (STANAG 2511) effort, which incorporates categories of reliability and credibility. Reliability has traditionally been assessed for physical machines to support failure analysis. Source reliability of a human can also be assessed. Credibility is associated with a machine process or human assessment of collected evidence for information content [43]. As the group work progresses, further standards are likely to be included as well.

Another example is the *EvaluationCriterion* class, depicted in Figure 5 and is at the core of any evaluation procedure. Not surprisingly, it is the larger class of the URREF ontology and the one with more levels. When looking at its main sub-classes, the more detail-oriented readers would be able to establish a parallel between these subclasses and the items within the Evaluation Framework Boundary framework depicted in Figure 2. More specifically, the Uncertainty Representation and Uncertainty Reasoning boxes can be mapped directly

[4]https://github.com/protegeproject/owlviz

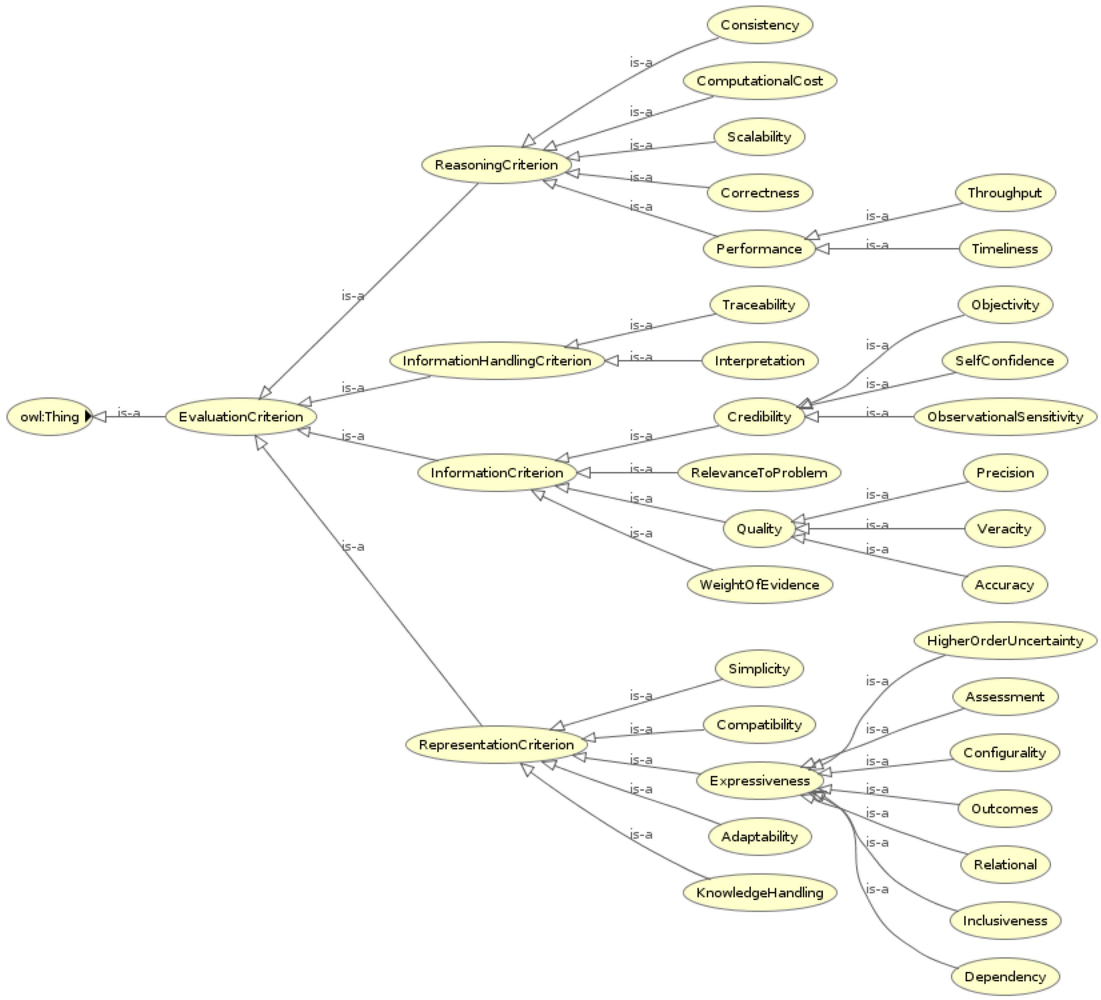


Fig. 5. *EvaluationCriterion* class

to the equally named sub-classes, while the classes *InformationHandlingCriterion* and *InformationCriterion* can be associated with the flow of information between Input and Output boxes.

The above classes form the structure of the URREF ontology, and were meant to collectively support the evaluation of uncertainty of an IF system. This is the third version of the URREF ontology, and at the time of this writing the group is now focusing on the case studies, which provide the necessary testbed for its ideas—and might force changes in the above classes. This approach privileges the pragmatism of having a good solution against having an "ideal" but unattainable solution. For instance, a definitive reference would involve having universally accepted definitions and usage for terms such as "Precision." This is unfeasible in any field of research that is not tightly controlled by a unique authoritative entity. The ETURWG approach also takes into consideration that more important than naming a concept is to ensure that it is represented clearly and distinctly within the ontology so to ensure the consistency of the latter.

Ontology reasoning requires axioms and properties to be defined, formally exposing the relationships between the above concepts that ultimately drive the logical reasoning that makes ontologies a very flexible and powerful technique. As an example, the object property *HasDerivationOfUncertainty* is used to map individuals of class *Evidence* (i.e., the domain of the property *hasDerivationOfUncertainty*) to individuals of class *UncertaintyDerivation* (i.e., the range of the property). The reasoner would use this relationship between these classes to support queries, automated classification, and other features the URREF could provide to its users.

A comprehensive description of the URREF ontology, with its classes, properties, and other elements is not within the scope of this paper. For a comprehensive overview of the URREF ontology, interested readers should refer to the ETURWG Github repository and the ETURWG website already mentioned in this paper.

C. The URREF Elementary Fusion Process

The elements of the Uncertainty Representation and Reasoning (URR) techniques to be assessed and compared will be referred within the URREF framework as

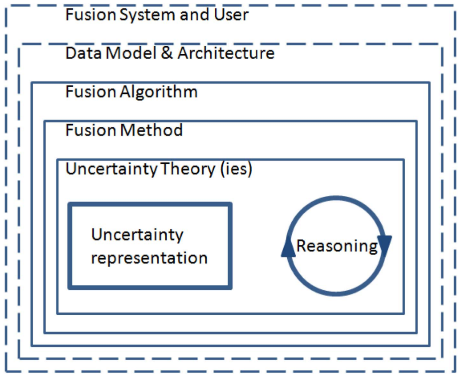


Fig. 6. An approximate hierarchy of fusion system components as possible evaluation subjects.

evaluation subjects. Owing to the complex and multiple connections between elements it seems difficult (if at all possible) to separate the uncertainty representation (e.g., an instantiated probability distribution) from its associated reasoning scheme (e.g., Bayes' rule), from its underlying uncertainty theory or mathematical framework (e.g., probability theory), from an underlying semantic representation (e.g., possible worlds, Ontology Web Language (OWL)), from the fusion method, from the fusion algorithm processing information (e.g., a specific implementation possibly involving some approximation), from a higher-level fusion system possibly including some human interaction.

Figure 6 illustrates some system components to assess and which interact to build a complete fusion system. As far as the URREF is concerned, the elements of an Uncertainty Representation and Reasoning scheme are the main evaluation subjects (thick lines in Figure 6), while the uncertainty theory, fusion method and fusion algorithm are of secondary focus. It is not the main purpose of the URREF to address the assessment of the fusion system nor the data model nor the architecture (dotted lines in the figure). Empirical evaluation techniques often require embedding some uncertainty representation and its associated reasoning scheme within the fusion method, which serves as an enabler and becomes often the subject of evaluation itself. Inherently, it is not a trivial problem to isolate the uncertainty representation from either its reasoning scheme or the fusion algorithm which may implement other contributing aspects, albeit minor.

For each evaluation subject, a series of evaluation criteria of interest is then defined in the URREF ontology [22] (see Section IV). It happens that the same criterion applies to different subjects with thus possible different associated metrics (or measures). For instance, *Accuracy* can be a quality criterion of information and of a source of information.

The fusion method is further detailed here by defining a generic procedure that highlights the main elementary constructs of uncertainty representation and reasoning that are the primary URREF evaluation subjects to be further defined in Section IV. The fusion method may be very complex, involving possibly several uncertainty representations, combination or inference rules, possibly framed in different uncertainty theories. Here, we abstract away complexities that are inessential to our purpose to obtain a simple, albeit quite general, fusion method aimed at clarifying the information flow. The result can be considered as an "atomic" fusion process.

The elementary constructs of a fusion process are shown in Figure 7, and illustrated with corresponding human intelligence fusion and multiple radar fusion examples in Table I:

① S is a source of information;

② ϕ is a piece of information provided by (or extracted from) S. It can be as simple as a measurement but could also be a natural language statement, a probability distribution, or in general a piece of information with some uncertainty already represented in a specific uncertainty theory;

③ h is the uncertainty representation process by which ϕ is transformed into a dedicated mathematical function conveying some notion of uncertainty. The process h is typically the choice of the solution designer who selects the way incoming information may be converted into a mathematical object. It can be learned from data when available or it can be general to all POIs, specified by type of source, by type of information, etc. Prior information on source's quality (e.g., reliability), source's self-confidence in statement, contextual information, comparison with other POIs, etc, may be captured by h;

④ $h(\phi)$ is the instantiated mathematical representation as built by h and expresses either the self-assessment of the source, an external assessment by the designer based on prior source's quality knowledge or an aggregation of both;

⑤ ρ is the inference process which transforms $h(\phi)$ into another $h_{\oplus}(\phi)$ within the same uncertainty theory. At this point, a series of POIs from other sources $\{h(\phi)\}_{i=1,\ldots,N}$ are combined, where other POIs are deduced, predicted, revised, etc;

⑥ $h_{\oplus}(\phi)$ is the resulting piece of information built from $h(\phi)$ and other related information;

⑦ l is the decision process which transforms $h_{\oplus}(\phi)$ to provide the decision, i.e., the output information ϕ';

⑧ ϕ' is the information output, to be possibly sent other systems. It can be a formal representation, i.e., an uncertainty function (such as a probability distribution), or a single measurement estimated after the decision process (soft versus hard decision). It can thus contain or not contain some uncertainty;

⑨ the reasoning process is $l \circ \rho$;

⑩ the Atomic Decision Procedure (ADP) is $l \circ \rho \circ h$.

TABLE I

Elementary fusion process constructs illustrated at the hand of a) a human intelligence fusion example and b) a multiple radar centralized fusion example

Element	Example
① S	a) Human observer b) Radar sensor
② ϕ	a) Human report, b) Radar range velocity measurement
③ h	a) Convert a natural language statement to a belief function over locations, b) convert a range and angle measurement and associated Root Mean Square Error (RMSE) error value to a Gaussian distribution with mean and variance
④ $h(\phi)$	a) Belief function b) Gaussian probability distribution
⑤ ρ	a) Dempster's combination rule (combine multiple reports) b) Bayes' rule (combine multiple measurements form different radars)
⑥ l	a) Maximum of plausibility rule b) Find expected value of posterior distribution
⑦ ϕ'	a) Element with maximum plausibility (or complete plausibility distribution over singletons) b) Expected value of the posterior distribution

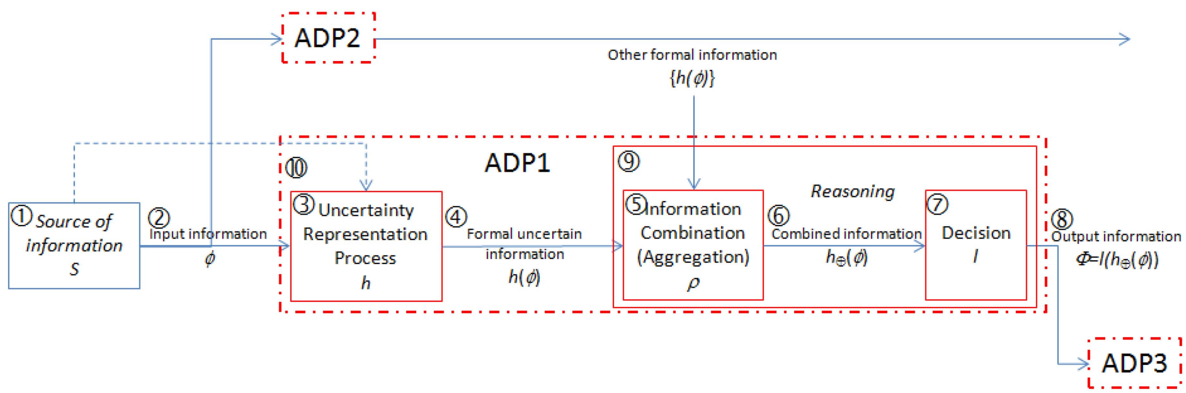


Fig. 7. Basic information flow and evaluation subjects.

Figure 7 illustrates this process and depicts each of the above 10 items in its appropriate place in the process.

As further detailed in [44], the method can distinguish between:

a) information processors (providing POIs): Elements ①, ③, ⑤, ⑦, ⑨;
b) the provided: Elements ②, ④, ⑥, ⑧;
c) the pairs (process; output information): (①,②); (③,④); (⑤,⑥); (⑦,⑧); (⑨,⑧); (⑩,⑧)=(①,②)

From an algorithmic standpoint, we may want to assess each of the 10 items above. However, based on the following observations some simplifications arise:

- Each information processor can be assessed through the information it provides, so it is natural to consider the pairs (processor; output information);
- The pair (①,②), (source; input information), is defined as a secondary evaluation subject and its previous characterization should be considered in the assessment of the primary subjects (see Section IV);
- In some cases, the reasoning process ($l \circ \rho$) may be considered as a whole, without separating the combination from the decision.

Thus the most important pairs (i.e., primary subjects) are:

- (③,④)—the uncertainty representation process h together with its output;
- (⑨,⑧)—the reasoning process together with its output;
- (⑩,⑧)—the pair (representation, reasoning) together with its output.

IV. URREF EVALUATION SUBJECTS

Following the previous detailed description of an elementary fusion process, this section defines the different evaluation subjects and identifies the corresponding criteria of the URREF ontology.

Definition 1 (**Evaluation subject**) An *Evaluation Subject* is an item which can be assessed through the Uncertainty Representation and Reasoning Evaluation Framework according to the criteria defined in the URREF ontology.

Evaluation subjects correspond to design choices to assess for an enlightened solution design. The identification of the evaluation subjects helps to better specify and communicate the goal of the URREF ontology but also better focus the effort on the primary subjects that are uncertainty representations and reasoning schemes embedded in fusion algorithms. In the following we thus specify what is understood by "uncertainty representation" and by "reasoning."

The Joint Directors of the Laboratory (JDL) or updated version of the Data Fusion Information Group (DFIG) model fusion model (e.g., [45]) is a functional description of a series of fusion problems organized along levels. In order to solve these problems, a modeling step is required which isolates the real world entities and processes (RWEPs [46]) of interest, identifies the corresponding (uncertain) variables, possible sources of information, makes some assumption of the world's dynamics and states, represents the underlying uncertainty and finally designs the reasoning scheme by either merging, updating, revising information for an estimation (or prediction) of the variables states.

DEFINITION 2 (**Fusion problem**) A *fusion problem* corresponds to some unknown states or dynamics of the real world and for which several sources of information are available. Fusion problems typically correspond to the different levels of the JDL/DFIG model and encompass as subclasses for instance tracking, target classification, anomaly detection, threat assessment and resource management.

Note that the notion of source depends on the modeling and does not necessarily mean several sensors. Features in a classification problem could be considered as "sources." A fusion problem is solved by a fusion method.

DEFINITION 3 (**Fusion method**) A *fusion method* is a set of rules encoding a solution to the fusion problem at hand, involving several sources of information. It implements some uncertainty representations and reasoning schemes.

For instance, a Kalman filter is a fusion solution to a multi-sensor filtering problem in tracking applications. It implements an updating scheme involving a prediction step followed by a revision step within a probabilistic framework [47]. A naive Bayes classifier is a fusion solution to a classification problem, which is implemented as a naive Bayes (i.e. probabilistic) model where features (possibly provided by different sources) are assumed to be independent, followed by a maximum *a posteriori* (MAP) decision rule.

DEFINITION 4 (**Uncertain variable**) An *uncertain variable* represents a feature of the real world for which the state is unknown, partially known or uncertain. It describes the fusion problem and its state has to be estimated by the fusion method.

The concept of uncertain variable generalizes the one of random variable itself representing a random phenomenon (and generally expressed by a probability distribution), to encompass the cases of epistemic uncertainty where uncertainty is not due to the variability of the phenomenon, but to a lack of knowledge. We can define thus two types of variables relative to the nature of uncertainty (see class *UncertaintyNature* of the URREF ontology [22]): Random variable and epistemic variable.

For instance, in a Kalman filter the uncertain (random) variables correspond to the position and the speed of the target at time t and $t+1$, usually gathered into (random) state vectors $\mathbf{x}_t$ and $\mathbf{x}_{t+1}$, but also to the measurements received by the sensors represented by a state vector $\mathbf{y}_t$. In a vessel classification problem, the uncertain (epistemic) variable would be the class of the specific vessel observed.

The primary purpose of the URREF is to assess how uncertainty is handled in a given fusion method, with a specific focus on the uncertainty representation and the reasoning components. In a formal uncertainty handling, both components abide to rules and constraints defined by the *uncertainty theory* considered.

DEFINITION 5 (**Uncertainty theory**) An *uncertainty theory* is a set of axioms and rules describing uncertainty representation and reasoning. Two components can be distinguished, although possibly strongly connected:

1) The *representation* which defines *uncertainty relations (or functions)* through established sets of axioms;
2) The *reasoning* which defines *inference (or belief change) rules* to manipulate uncertainty functions and create new ones.

Uncertainty functions and inference rules can be assigned different semantics.

Examples of quantitative uncertainty theories are probability theory, evidence theory, fuzzy sets theory, random sets theory, possibility theory, and imprecise probability theory. Some qualitative theories are possibilistic logic, fuzzy logic or probabilistic logic.

A Kalman filter is framed into probability theory which itself defines probability functions to convey uncertainty notions. Probability functions must satisfy the three axioms of $P(\emptyset) = 0$ for the impossible event, $P(\Omega) = 1$ for the certain event and $P(A) + P(\bar{A}) = 1$ for any event (where $\emptyset$ denotes the empty set, Ω denotes the universe and $\bar{A}$ denotes the complement event of A). The most classical inference rule is Bayes' rule which defines the posterior probability of an event based on the occurrence of another one as $P(A \mid B) = P(B \mid A)P(A)/P(B)$. Several interpretations (or Uncertainty-Derivations [22]) can still be assigned to probability values, roughly either objective (e.g., frequentist) or subjective (e.g., degree of belief).

DEFINITION 6 (**Uncertainty relation**) An *uncertainty relation* is a mathematical or logical object conveying some notion of uncertainty. It can be an *uncertainty function* if each subset of the frame is related to a value between 0 and 1 or a binary relation such as an accessibility relation in modal logic.

The uncertainty relation covers uncertainty functions such as probability functions but also equivalence relations between states defining for instance rough sets. Uncertainty relations are the core representation of uncertainty, and express how much or how we or/and the sources are uncertain. They are defined over sets of variables, which themselves represent *what* we are uncertain about.

DEFINITION 7 (**Uncertainty Modeling Scheme**) An *Uncertainty Modeling Scheme* (UMS) is a theoretical concept that provides a mapping between (i) domain independent mathematical concepts and (ii) classes of fusion problems. A UMS

(1) introduces types of *uncertain variables* and the types of *relations* between these variables that are relevant for the modeling of a specific type of problem;
(2) provides *semantics* for a selection of uncertain relation types;
(3) formulates *assumptions* about the represented problem type;
(4) defines uncertainty functions over these variables.

For example, the UMS defining representations used by Kalman Filters introduce random variables representing the states of a dynamic process and observations. Moreover, it relates covariance matrices to the normally distributed process dynamics and observations, respectively. This model is based on the assumptions that the represented dynamic processes are linear and normally distributed. The UMS for causal Bayesian Networks associates basic conditional probabilities with uncertain causality. This model assumes Markov property, conditional independence theoretically captured by d-separation concepts and Markov Blankets. A UMS typically corresponds to a specific type of reasoning scheme. A UMS represents a theoretical basis for the solution of a specific use case (see Def. 8).

DEFINITION 8 (**Uncertain Domain Model**) An *Uncertain Domain Model (UDM)* is an artifact defined through (i) a set of uncertain variables and (ii) uncertainty relations which encode some assumptions about the real-world dynamics and states in a specific application. An UDM is a specific instantiation of a representation of the uncertainty associated with a specific real-world problem itself framed into an uncertainty theory and thus constrained by the rules and axioms. Such framing is provided by a suitable UMS (see Def. 7).

UMS defines the form of h and ρ, i.e. types of variables and functions in combination with a suitable uncertainty theory. The UDM defines the specific constellations of the variables and specific parameters used in h and ρ. The UMS supports theoretical analysis that facilitates (i) comparison of uncertainty representations and reasoning in a class of applications and (ii) an evaluation of the adequacy of a specific technique in a specific application (use case). The evaluation of a UDM supports the engineering process in the development of a specific fusion solution. An uncertain domain model could be the graphical part of a Bayesian network together with the instantiated joint probability distribution defining uncertainty over the set of variables. An uncertain domain model describes uncertainty about states of the variables and relations between variables and expresses thus some assumptions about either

(1) uncertain knowledge of possible states and dynamics of the world (generic knowledge/information/uncertainty);
(2) uncertain evidence about the current state of the world (singular information/uncertainty).

Although it is more common to associate singular evidence to a source of information, generic knowledge can also itself be derived from some source. For instance, a statistical model representing the maritime traffic and linking kinematic variables through some (possibly conditional) probability distributions (e.g. see [48]) can be interpreted as an uncertainty function derived from a specific AIS dataset covering a particular area during a given period of time, the source of this model.

DEFINITION 9 (**Uncertainty reasoning scheme**) An *uncertainty reasoning* scheme encodes some inference under uncertainty aiming at solving the fusion problem, by means of rules defined for several *uncertainty functions*.

For instance, Bayes' rule can be used "both for prediction from observations and revision of uncertain information" [49]. It can be used as a merging (fusion) rule performing a conjunction (product) of likelihoods provided by different sources. Dempster's rule itself encodes merging of (singular) testimonies for independent sources [50]. The combination rules have also different semantics and maybe thus dedicated to solve different types of problems (e.g., [49]).

DEFINITION 10 (**Source (of information)**) A *source of information* is any entity providing some piece of information.

A source of information is a relative notion and covers anything from where information can be extracted, i.e. a dataset, a database, an image, a video, a witness, etc, or the device providing it, i.e. a radar, a camera, an expert, etc. It can provide either generic or singular information.

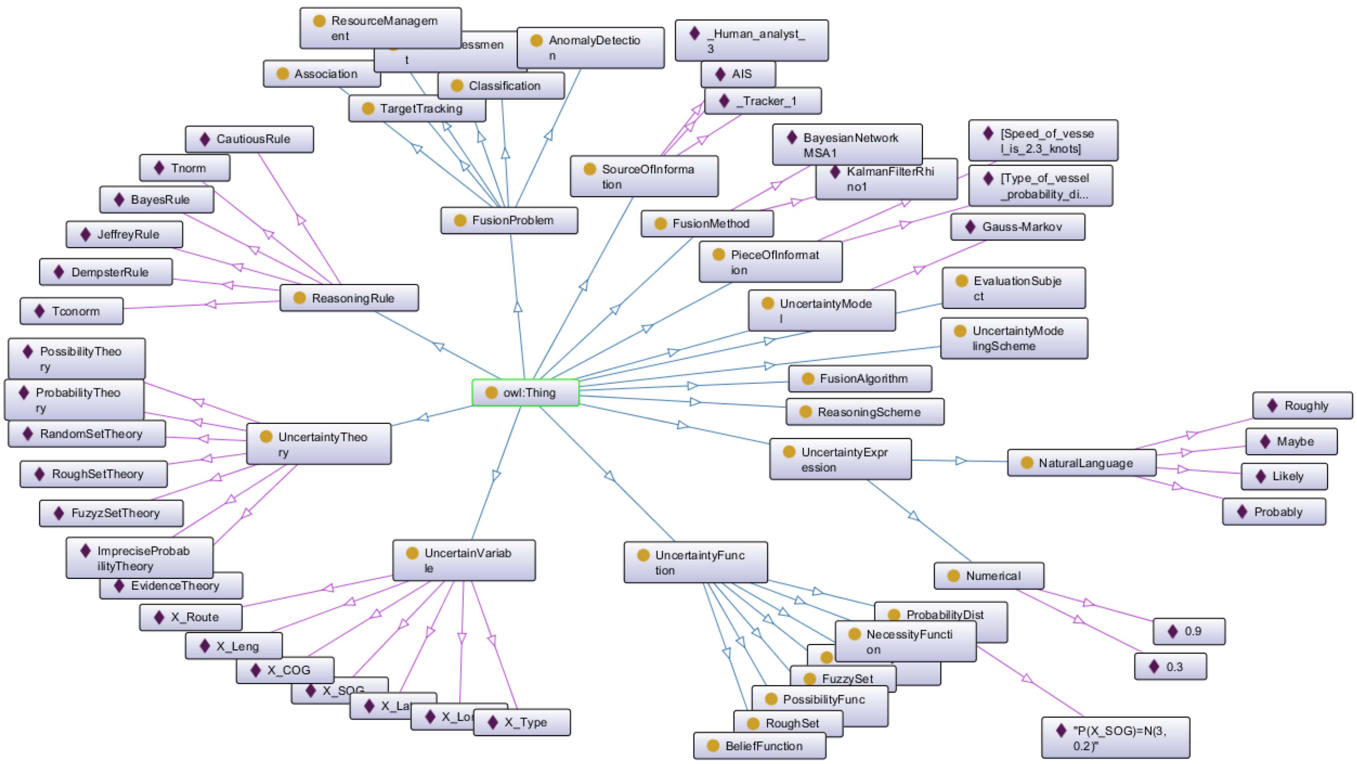


Fig. 8. URREF evaluation subjects, with sample instances (purple diamond bullets)

DEFINITION 11 (**Piece of information**) A *piece of information* is an item possibly conveying some information, and provided by a source.

The term "piece of information" is used in this paper in its most general meaning covering other notions such as evidence, knowledge and/or data. A piece of information can be as simple as a measurement (on the scale of real numbers) but could be a fact (i.e., an observation, known to be true), an uncertain statement already modeled into a given mathematical formalism (i.e., a probability distribution), an unstructured statement in natural language, etc.

Figure 8 lists the URREF evaluation subjects. Elements within rectangles with yellow circle bullets are classes. Examples of instance for each class are provided in rectangles with purple diamond bullets. The meaning of the relationship is displayed on arrows. N-ary relationships are displayed with blue arrows containing a triangle.

We identify the *primary evaluation subjects* of the URREF as:

- the **uncertainty representation**, which is either instantiated or theoretical: a particular probability distribution or probabilities in general; it may include instantiated uncertainty representations of processes in the real-world and how those processes are observed;
- the associated **reasoning** (or calculus) that comprises the combination, conditioning, updating, inference, decision, transformation rules. The calculus may be assessed while instantiated within a fusion method or theoretically, regardless any application or algorithm, focusing on the semantics for instance (e.g., Bayes' rule in general).

In URREF, the first is represented by the classes *UncertaintyTheory and UncertaintyModel*, while class *UncertaintyReasoning* represents the latter.

It is expected that a preliminary assessment of theoretical objects, either uncertainty representations or reasoning rules, is performed in the initial design phase (inception phase [51]), relying mainly on the literature and on the expertise of the fusion method designer. This pre-screening should provide guidance on the selection of appropriate models or reasoning schemes to be implemented which best suit the fusion problem at hand as far as uncertainty handling is concerned. In a second step, the assessment of instantiated representations and reasoning schemes should be assessed through a specific implementation of the fusion solution in a fusion algorithm, processing data. Then, output data analysis should provide some assessment on the implemented uncertainty handling method.

Secondary evaluation subjects of the URREF encompass other elements which either support or can be derived from the assessment of the primary subjects, but which are not the main concern of the URREF ontology:

- the **fusion method**, making use of instantiated uncertainty representations embedding **pieces of information** ϕ built according to a specific uncertainty representation process h and associated calculus $l \circ \rho$, and implemented by the **fusion algorithm**;

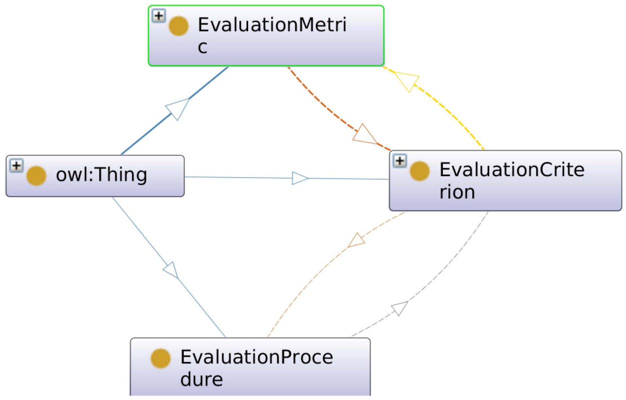


Fig. 9. In an EVALUATION PROCEDURE, EVALUATION SUBJECTS are assessed by EVALUATION CRITERIA, which are measured by EVALUATION METRICS.

- the **source** of information which provides the different and which quality may impact the whole fusion process. It can be expected that an uncertainty representation is able to properly capture and handle the meta-information about the source quality;
- the **pieces of information** input, processed and output throughout the process. Input and output information are only two special cases but others can be considered provided by internal steps such as for instance the aggregated information. The information assessment is at the core of the assessment of the uncertainty representation and reasoning. However, the development of such information quality criteria is not currently the main purpose of the ETURWG;
- the **uncertainty theory** (or framework) for uncertainty representation and reasoning (e.g., probability, fuzzy set, belief function theories). It can be assessed either theoretically, based on axioms, properties and original semantics as reported in the literature or through the assessment of the output provided by a specific **fusion algorithm** implementing the **fusion method** and specific instantiated uncertainty representations.

The fusion algorithm may be assessed either as a whole (assessing only the output) or through its different components that are the instantiated uncertainty representation (process and output information), and instantiated calculus (process and output information). Equivalently, the uncertainty theory can be assessed considering the theoretical uncertainty representation (i.e., general uncertainty function such as a probability or a belief function) on the one hand or/and the theoretical calculus apparatus (i.e., the set of reasoning tools available to this framework) on the other hand.

For each evaluation subject, there exists a corresponding set of evaluation criteria within the ontology, as illustrated in Figure 9. The quality of the source is assessed by *QualityCriterion*, the provided are assessed by *InformationCriterion*, the uncertainty representation part of the fusion method is assessed through *RepresentationCriterion* and the reasoning part is assessed through *ReasoningCriterion*.

A. Source criteria

Criteria about the source of information are necessary to characterize information input to the fusion process (other said, output by the source). The use of these criteria is rather informative than "judgmental." We assume that these initial assessments are known prior to processing the information and the question is *if* and *how* the fusion method, and especially the uncertainty representation and reasoning scheme are able to handle the different source quality dimensions. They are directly linked to the criteria on expressiveness (i.e., class *ExpressivenessCriterion*). As such, the source is a secondary evaluation subject and impacts the other subjects.

B. Information Criteria

Pieces of information (POIs) appear at different steps of the fusion process and include in particular, input data, measurement or declaration before any modeling of uncertainty (i.e., input information or dataset), the instantiated uncertainty representation (after uncertainty has been modeled), aggregated information (after the combination or inference process) and output information to be consumed by the user. Each of these POIs should be characterized according to the same subset of criteria although the expectations in their respect may differ. For instance, it is not expected that the input information be precise, nor true. Yet, it would be expected at the output. Also, comparing pieces of information at several steps of the process provides assessment of relevance (if one has an impact on the other one). Therefore, the same set of evaluation criteria should be used to assess input information, uncertain information (after h), combined information, and output information. If the same measure is used to capture this criterion, only the values (and the user's expectations) may change, not the criteria themselves. For input information, the assessment is rather a characterization, while for the other POIs during the process, the assessment criteria can be turned to optimization criteria to further tune the algorithm (e.g., maximize the Accuracy).

C. Representation Criteria

The Representation criteria (class *RepresentationCriterion*) are aimed at assessing the primary subject of evaluation within the URREF. Unsurprisingly, expressiveness is the main one. Indeed, at the inception phase [51], i.e. before any instantiation of an uncertainty representation, we are interested in the expressive power provided by its underlying uncertainty theory. This is a *prior* (theoretical) assessment driven by the problem at hand which mainly relies on analyzes of (1) the axiomatic constraints of the framework and (2) the current literature about the development of the approaches and tools to support the representation of concepts of interest as identified within the expressiveness list of criteria. The instantiated uncertainty representation should also be assessed along with the subset of criteria. An instantiated uncertainty representation is a piece of information and as such, will be assessed using the information criteria described above.

D. Reasoning criteria

This subset of criteria is so far not very detailed within the URREF ontology of criteria. Several interrelated elements must be considered:

a) the calculus and mathematical apparatus of the uncertainty theory, i.e., the set of reasoning tools available within this mathematical framework,
b) a particular instantiation of use of one of these rules, and
c) the fusion method making use of this apparatus.

For a more detailed analysis, these three subjects should be clearly distinguished, although the same criteria may be applicable and relevant to all of them. For instance, if we consider the *Consistency* criterion:

a) a particular rule of combination could be assessed according to its theoretical ability to provide consistent results,
b) a specific use of the rule which relies on other elements such as the universe of discourse selected or the type of uncertainty function to be combine, could be assessed according to the consistency criterion, and
c) a method embedding the rule with the uncertainty function and associated universe of discourse within a higher-level reasoning scheme (e.g., nearest neighbors approach, back-propagation) may also be assessed according to the same criterion of consistency.

V. CASE STUDIES

The URREF framework and its ontology component were developed through an iterative process, an essential part of which was to apply the framework to of a set of use cases. The use cases were selected to reflect a range of considerations relevant to evaluation of uncertainty representation within the context of an overall fusion application. Applying the framework to use cases grounds the ideas in concrete application areas, and helps to uncover requirements that emerge as the framework is applied to a concrete problem.

The requirements of the use cases in development are the main driver dictating what properties are needed within the URREF ontology. As such, the work on developing these use cases has been generating new insights and requirements for the URREF (e.g., [51]–[55]). The three use cases are described briefly below, with emphasis on how URREF was applied to the use case, what was learned through this process, and how the framework evolved in response to applying it to the use case.

A. Maritime Domain Awareness

We consider a use case of maritime surveillance where a harbor area is monitored by a set of sources mixing sensors and humans: *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. 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., must be known with a very high confidence to the Watch Officer.*

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 vessel's captain). The radar covers the whole area, the Infra-Red (IR) camera covers only the area around Track *A*, a cargo vessel is in the vicinity of Track *B* but too far from Track *A* for visual identification, and 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/or *false* (the value provided by the source does not correspond to the true value). Consequently, when combining the different POIs, the Watch Officer may face conflicting information.

In order to solve that fusion problem, several solutions can be designed. In [56], we illustrated how the URREF can support the designer in the decision of which uncertainty representation and reasoning method for fusion should be used. Two different fusion methods are compared: One framed into *probability theory* using Bayes' rule, and another one framed into *evidence theory* using Dempster's rule. The URREF criteria defined in classes *UncertaintyType*, *UncertaintyDerivation* and *UncertaintyNature* are used to categorize the input information highlighting the importance of the derivation of uncertainty values, as it has a direct impact on the interpretation of the output uncertainty. We stressed how the elements supporting uncertainty (e.g., variables, links between variables, uncertainty expression) crossed with the type of information (generic knowledge versus singular evidence) help in clarifying that Dempster's rule does not use generic knowledge but uncertain singular information (evidence), while Bayes' rule relies on generic information (knowledge).

B. Counter Rhinoceros Poaching Decision Support

The rhino poaching use case involves a decision support system that directs the patrol effort of the rangers to the areas with elevated risk of poaching [57], [58]. The central part of such a system is a set of Bayesian threat models, each with context evidence instantiated to correspond to a specific area or cell. A threat model is implemented as a Bayesian Network (BN) that captures the correlations between various context factors influencing the poaching (facilitators/inhibitors) as well as observable phenomena that might indicate an imminent threat. The system outputs a probability heat map that indicates the suitability for poaching at a specific point in space and time. The first attempt at applying the URREF ontology to the counter rhino poaching decision support system is presented in [59]. Given information in such a probability heat map, the rangers can position scarce resources distributed over large surface areas, such that the chance of preventing poaching is improved. Thus, the decision support system for counter rhino poaching operations covers all of the components of the OODA loop. The use of URREF concepts is demonstrated in [60] with reference to the OODA loop applied to the rhino decision support use case. Additional sources of information include human intelligence (HUMINT) reports of the field operations as well as the current status of the international rhino trafficking agencies.

Uncertainty may enter into a fusion system during both the design/modeling and routine operational phases. Selective application of the URREF to the anti-rhino poaching use case is demonstrated to characterize uncertainty during the design/modeling phase in [46] and during routine fusion system operation in [51], [60]. In particular, the URREF criteria are applied within the context of a fusion system development and deployment life cycle, as demonstrated on a high level context driven fusion approach to tracking poachers [51].

C. Cyber Threat Models

Systems for threat analysis enable users to understand the nature and behavior of threats and to undertake a deeper analysis for detailed exploration of threat profile and risk estimation. Models for threat analysis require significant resources to be developed and are often relevant to limited application tasks. In the Cyber Threat Use Case we presented and discussed a model for cyber threats which comprises an expert model and its translation into a Bayesian network (BN) as a tool for the development of practical scenarios for cyber threats analysis [61]. The BN for cyber threats is automatically generated from the expert model, highlighting vulnerabilities of systems along with threat-specific patterns, actors, actions and indicators [62]. For this use case, the goal of using the URREF ontology was to capture the quality of the knowledge. While the expert model was created manually by domain experts, by following a time consuming and expensive process, the BN was created thanks to an automatic procedure. Thus, the resulting models have different characteristics and granularity levels, and the question of their accuracy has to be addressed. For this purpose, the main URREF class considered for analysis was *RepresentationCriterion*, a general class regrouping several criteria explaining how uncertainty is characterized, captured and stored during modeling and representation stages, and introducing the most specific concepts of *Simplicity*, *Adaptability* and *Expressiveness* [52]. To analyze the model underlying the cyber threat application, *Simplicity* and *Expressiveness* criteria were considered. *Simplicity* is important since the expert model has to be created manually; *Expressiveness* is regarded to assess whether the knowledge encoded in the models is sufficient. Moreover, metrics were defined for those criteria, based on the characteristics of the models created (number of nodes in the model, density of connections). Several experiments carried out with different configurations of the model showed how the quality level of the knowledge representation, as captured by means of *Simplicity* and *Expressiveness*, is impacted by parameters of the model but also a complementary evolution of those criteria, as increasing the *Simplicity* goes hand in hand with decreases in *Expressiveness*. Future work is planned to carry out a complete assessment of knowledge representation using URREF criteria, to apply them to different BNs of different sizes and granularities, and to correlate the criteria for knowledge representation with other criteria of the URREF ontology.

VI. DISCUSSION AND CONCLUSION

Evaluation of IF systems presents intrinsic challenges due to the complexity of fusion systems and the sheer number of variables influencing their performance. In LLIF systems, the impact of uncertainty representation is well understood, and generally quantifiable. However, at higher levels of IF the approach chosen for representing uncertainty has an overall impact on system performance that is hard to quantify or even to assess from a qualitative viewpoint. This issue was recognized by the Fusion community when creating the ETURWG, with the main goal of providing an unbiased framework for evaluating the impact of uncertainty in IF systems. From the beginning, it became clear that the various approaches and technical considerations demand a common understanding that is only achievable by a formal specification of the contrasting semantics and pragmatics involved. As a result, the group developed the methodology for evaluation, the elements of the framework supporting it, a set of formal definitions of the distinct subjects under evaluation, as well as the linkage between these and the key aspects of the framework. As explained in this work, URREF is not a system or software application that can be "directly applied" to a use case. Yet, the use cases described here were essential for the group to achieve an understanding of all the nuances and idiosyncratic aspects of the process of evaluating techniques that are fundamentally different in their assumptions and views of the world. They provided the grounding for establishing the URREF concepts and mechanisms needed to mitigate the effects the underlying assumptions of each theory have in biasing the design of evaluations—each usually geared towards the strengths of one technique at the expense of the others. URREF does not completely remove the subjectivity and biases involved in evaluating uncertainty representation techniques, but is a strong step towards that direction.

ACKNOWLEDGMENT

Work in the ETURWG was commissioned by and had the full support from the International Society of Information Fusion, which vibrant community was the main driver of this volunteer effort by researchers from all over the world. Authors Kathryn Laskey and Paulo Costa would like to recognize the support of the US Army Research Office via Agreement #W911NF-11-1-0176 during the beginning of the ETURWG activities. During the last seven years, many people actively participated in the ETURWG effort, and undoubtedly left their contribution to the framework. Although we do not have the space to cite all of them, we would like to mention some who have devoted a reasonable amount of their time in the group. The authors are specially grateful for the contributions from Sten Andler, Mark Locher, Matt Roberts, Amandine Belenger, Dafni Stampouli, Gavin Powell, Max Kruger, Brian Ulcliny, Audun Josang, Alta de Waal, Claire Laudy, Simon Maskell, Kellyn Rein, Joe Steinhauer, David Hall, and Spandana Jagtap.

REFERENCES

[1] A. N. Steinberg, C. L. Bowman, and F. E. White
"Revisions to the JDL data fusion model,"
in *Sensor Fusion: Architectures, Algorithms, and Applications III*, B. V. Dasarathy, Ed., vol. 3719. Orlando, FL, USA: SPIE, Mar. 1999, pp. 430–441. [Online]. Available: http://link.aip.org/link/?PSI/3719/430/1.

[2] J. Llinas, C. Bowman, G. Rogova, A. Steinberg, E. Waltz, and F. White
"Revisiting the JDL Data Fusion Model II,"
in *In P. Svensson and J. Schubert (Eds.), Proceedings of the Seventh International Conference on Information Fusion (FUSION 2004)*, Stockholm, Sweden, Jul. 2004.

[3] E. Blasch and P. Hanselman
"Information fusion for information superiority,"
in *Proceedings of the IEEE 2000 National Aerospace and Electronics Conference. NAECON 2000. Engineering Tomorrow (Cat. No.00CH37093)*, Oct. 2000, pp. 290–297.

[4] V. Dragos
"Shallow semantic analysis to estimate humint correlation,"
in *Information Fusion (FUSION), 2012 15th International Conference on*. IEEE, 2012, pp. 2293–2300.

[5] ———
"Assessment of uncertainty in soft data: a case study,"
in *Information Fusion (FUSION), 2014 17th International Conference on*. IEEE, 2015, pp. 1–8.

[6] ———
"An ontological analysis of uncertainty in soft data,"
in *Information Fusion (FUSION), 2013 16th International Conference on*. IEEE, 2013, pp. 1566–1573.

[7] V. Dragos, X. Lerouvreur, and S. Gatepaille
"A critical assessment of two methods for heterogeneous information fusion,"
in *Information Fusion (Fusion), 2015 18th International Conference on*. IEEE, 2015, pp. 42–49.

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

[9] P. Walley
"Measures of uncertainty in expert systems,"
Artificial Intelligence, vol. 83, pp. 1–58, 1996.

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

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

[12] G. J. Klir
"Probabilistic versus possibilistic conceptualization of uncertainty,"
in *Proceeding of the first international Symposium on Uncertainty modeling and analysis*, 1990, pp. 38–41.

[13] 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.

[14] 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.

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

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

[17] 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.

[18] 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, July 2003, pp. 1207–1214.

[19] D. Ongaro and J. Ousterhout
"In Search of an Understandable Consensus Algorithm," in *Proceedings of the 2014 USENIX Conference on USENIX Annual Technical Conference*, ser. USENIX ATC'14. Berkeley, CA, USA: USENIX Association, 2014, pp. 305–320. [Online]. Available: http://dl.acm.org/citation.cfm?id=2643634.2643666.

[20] "ISIF.org/Working Groups—Evaluation of Techniques for Uncertainty Representation Working Group."
[Online]. Available: http://isif.org/evaluation-techniques-uncertainty-representation-working-group/.

[21] "Evaluation of Techniques for Uncertainty Representation Working Group Website."
[Online]. Available: http://eturwg.c4i.gmu.edu/.

[22] P. C. G. Costa, K. B. Laskey, E. Blasch, and A.-L. Jousselme
"Towards Unbiased Evaluation of Uncertainty Reasoning: The URREF Ontology,"
in *Proceedings of the Fifteenth International Conference on Information Fusion*. Singapore: International Society for Information Fusion, 2012.

[23] 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*.

[24] P. C. G. Costa, R. N. Carvalho, K. B. Laskey, and C. Y. Park
"Evaluating uncertainty representation and reasoning in HLF systems,"
in *14th International Conference on Information Fusion*, Jul. 2011, pp. 1–8.

[25] K. J. Laskey and K. B. Laskey
"Uncertainty Reasoning for the World Wide Web: Report on the URW3-XG Incubator Group,"
in *Proceedings of the Fourth International Conference on Uncertainty Reasoning for the Semantic Web—Volume 423*, ser. URSW'08. Aachen, Germany, Germany: CEUR-WS.org, 2008, pp. 108–116. [Online]. Available: http://dl.acm.org/citation.cfm?id=2889804.2889814.

[26] P. C. G. Costa, K. Chang, K. Laskey, T. Levitt, and W. Sun
"High-level fusion: Issues in developing a formal theory,"
in *2010 13th International Conference on Information Fusion*, Jul. 2010, pp. 1–8.

[27] A.-L. Jousselme, P. Maupin, and E. Bossé
"Quantitative Approaches,"
in *Concepts, Models and Tools for Information Fusion*. Artech House, Inc., Mar. 2007, vol. Chapter 8, p. 376.

[28] A. P. Dempster
"A generalization of Bayesian inference,"
Journal of the Royal Statistical Society, vol. 30, pp. 205–247, 1968.

[29] L. A. Zadeh
"Fuzzy logic and approximate reasoning,"
Synthese, vol. 30, pp. 407–28, 1975.

[30] ———
"Fuzzy Sets as a Basis for a Theory of Possibility,"
Fuzzy Sets and Systems, vol. 1, pp. 3–28, 1978.

[31] J. Pearl
Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference,
1st ed. Morgan Kaufmann, Sep. 1988.

[32] G. Shafer
"The Dempster-Shafer theory,"
New York, NY, USA, pp. 330–331, 1992.

[33] J. Pearl
Causality: models, reasoning, and inference.
Cambridge University Press, Mar. 2000.

[34] J. Dezert and F. Smarandache
"DSmT: A New Paradigm Shift for Information Fusion,"
in *Proceedings of Cogis'06 Conference*, Paris, France, Mar. 2006.

[35] A. Jøsang
Subjective Logic: A Formalism for Reasoning Under Uncertainty,
ser. Artificial Intelligence: Foundations, Theory, and Algorithms. Springer International Publishing, 2016. [Online]. Available: //www.springer.com/gp/book/9783319423357.

[36] D. Schum
Evidential Foundations of Probabilistic Reasoning.
New York: Wiley, 1994.

[37] J. B. Kadane and D. A. Schum
A Probabilistic Analysis of the Sacco and Vanzetti Evidence, 1st ed. Wiley, May 1996.

[38] D. A. Schum
"Inference Networks and the Evaluation of Evidence: Alternative Analyses,"
arXiv:1301.6737 [cs, stat], Jan. 2013, arXiv: 1301.6737. [Online]. Available: http://arxiv.org/abs/1301.6737.

[39] K. C. Chang, Z. Tian, S. Mori, and C. Chong
"MAP track fusion performance evaluation,"
in *Proceedings of the Fifth International Conference on Information Fusion. FUSION 2002. (IEEE Cat.No.02EX5997)*, vol. 1, Jul. 2002, pp. 512–519 vol.1.

[40] N. UzZaman and J. F. Allen
"Temporal evaluation,"
in *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies: short papers—Volume 2*, ser. HLT '11. Stroudsburg, PA, USA: Association for Computational Linguistics, 2011, pp. 351–356. [Online]. Available: http://dl.acm.org/citation.cfm?id=2002736.2002809.

[41] E. Blasch, P. Valin, and E. Bosse
"Measures of effectiveness for high-level fusion,"
in *Information Fusion (FUSION), 2010 13th Conference on*, Jul. 2010, pp. 1–8.

[42] M. A. Musen
"The Protégé Project: A Look Back and a Look Forward,"
AI matters, vol. 1, no. 4, pp. 4–12, June 2015. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4883684/.

[43] E. Blasch, K. B. Laskey, A.-L. Jousselme, V. Dragos, P. C. Costa, and J. Dezert
"URREF reliability versus credibility in information fusion (STANAG 2511),"
2013.

[44] 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.

[45] A. N. Steinberg and C. L. Bowman
"Rethinking the JDL data fusion levels,"
in *National Symposium on Sensor and Data Fusion*, 2004.

[46] 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 *Information Fusion (Fusion), 2015 18th International Conference on*. IEEE, 2015, pp. 50–57.

[47] S. Benferat, D. Dubois, and H. Prade
"Kalman-like filtering and updating in a possibilistic setting,"
in *Proc. 14th European Conf. on Artificial Intelligence (ECAI 2000)*. Berlin, Germany: IOS Press, 2000, pp. 8–12.

[48] B. Ristic, B. La Scala, M. Morelande, and N. Gordon "Statistical analysis of motion patterns in AIS Data: Anomaly detection and motion prediction," in *Proc. of the 11th Conference on Information Fusion*, Cologne, Germany, 2008.

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

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

[51] G. Pavlin, A.-L. Jousselme, J. P. De Villiers, P. C. G. Costa, and P. de Oude "Towards the Rational Development and Evaluation of Complex Fusion Systems: a URREF-Driven Approach," in *Proceedings of the Twenty-First International Conference on Information Fusion (FUSION 2018)*, Cambridge, UK, Jul. 2018.

[52] V. Dragos, J. Ziegler, and J. P. de Villiers "Application of URREF criteria to assess knowledge representation in cyber threat models," in *Information Fusion (Fusion), 2018 21th International Conference on*. IEEE.

[53] M. Kruger "Experimental Comparison of Ad Hoc Methods for Classification of Maritime Vessels Based on Real-life AIS Data," in *Proceedings of the Twenty-First International Conference on Information Fusion (FUSION 2018)*, Cambridge, UK, Jul. 2018.

[54] A. de Waal and K. Yoo "Latent Variable Bayesian Networks Constructed Using Structural Equation Modelling," in *Proceedings of the Twenty-First International Conference on Information Fusion (FUSION 2018)*, Cambridge, UK, Jul. 2018.

[55] E. Camossi and A.-L. Jousselme "Information and Source Quality Ontology in Support to Maritime Situational Awareness," in *Proceedings of the Twenty-First International Conference on Information Fusion (FUSION 2018)*, Cambridge, UK, Jul. 2018.

[56] A.-L. Jousselme "Semantic criteria for the assessment of uncertainty handling fusion models," in *Proc. of the 19th Int. Conf. on Information Fusion*, Heidelberg, GE, July 2016.

[57] H. Koen, J. P. De Villiers, G. Pavlin, A. de Waal, P. de Oude, and F. Mignet "A framework for inferring predictive distributions of rhino poaching events through causal modelling," in *Information Fusion (FUSION), 2014 17th International Conference on*. IEEE, 2014, pp. 1–7. [Online]. Available: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=6916121.

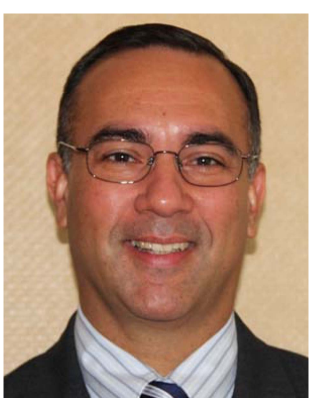
[58] H. Koen, J. P. De Villiers, H. Roodt, and A. de Waal "An expert-driven causal model of the rhino poaching problem," *Ecological Modelling*, vol. 347, pp. 29–39, 2017.

[59] J. P. de Villiers, G. Pavlin, P. Costa, K. Laskey, and A.-L. Jousselme "A URREF interpretation of Bayesian network information fusion," in *Information Fusion (FUSION), 2014 17th International Conference on*. IEEE, 2014, pp. 1–8.

[60] J. P. de Villiers, A.-L. Jousselme, A. de Waal, G. Pavlin, K. Laskey, E. Blasch, and P. Costa "Uncertainty evaluation of data and information fusion within the context of the decision loop," in *Information Fusion (FUSION), 2016 19th International Conference on*. IEEE, 2016, pp. 766–773.

[61] T. Kiesling, M. Krempel, J. Niederl, and J. Ziegler "A model-based approach for aviation cyber security risk assessment," in *Availability, Reliability and Security (ARES), 2016 11th International Conference on*. IEEE, 2016, pp. 517–525.

[62] J. Ziegler and B. Haarmann "Automatic generation of large causal bayesian networks from user oriented models," in *Proceedings of the 6th Workshop on Sensor Data Fusion (INFORMATIK 2011: Informatik schafft Communities)*. Citeseer, 2011.

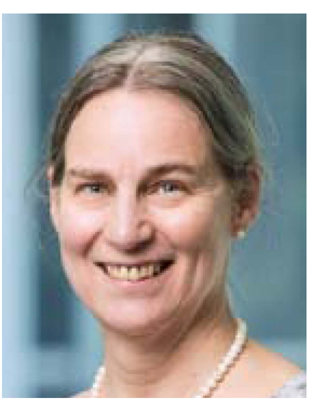


Paulo C. G. Costa is an Associate Professor of Systems Engineering and Operations Research at George Mason University, and Associate Director of the C4I & Cyber Center's Radio and Radar Engineering Laboratory.

His teaching and research interests comprise the areas of probabilistic ontologies, multi-sensor information fusion, Bayesian reasoning, predictive analysis, cybersecurity and decision theory.

He is a former fighter pilot with extensive experience in tactical and operational planning, and an expert in requirements engineering for complex systems, such as intelligent transportation and health-care support systems. Dr. Costa is an IEEE senior member, a member of the International Council of Systems Engineering, and currently serves as President of the International Society for Information Fusion (ISIF—tenures 2019 and 2020).

Anne-Laure Jousselme received her PhD degree from the Electrical Engineering Department of Laval University in Quebec City (Canada) and the Institut National Polytechnique de Grenoble (France) in 1997. Formely with Defense Research and Development Canada (DRDC), she is now with the NATO STO Centre for Maritime Research and Experimentation (CMRE) in La Spezia (Italy), where she conducts research activities on reasoning under uncertainty, high-level and hard & soft information fusion, information quality assessment and serious gaming applied to maritime situational awareness and anomaly detection. She is area editor of the International Journal of Approximate Reasoning and associate editor of the Perspectives on Information Fusion magazine. She is a member of the Boards of Directors of the International Society of Information Fusion (ISIF) where she serves as VP membership and of the Belief Functions and Applications Society (BFAS) where she serves as Secretary. She serves on program committees of the International Conference of Information Fusion and the International Conference on Belief Functions. She was Tutorial Chair of FUSION 2007 in Quebec City (CA), International Co-chair of FUSION 2015 in Washington and Technical Co-chair of FUSION 2019 in Ottawa (CA). She was general Chair of the Canadian Tracking and Fusion Conference (CTFG) in 2014 in Ottawa (CA) and Local Organizer of the International Conference of Scalable Uncertainty Management (SUM) in 2015 in Quebec City (CA).



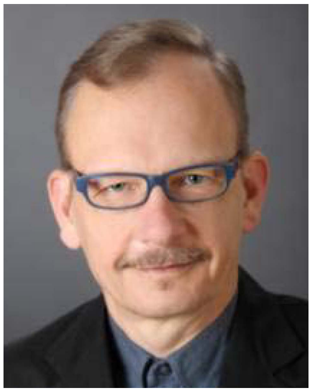
Kathryn Blackmond Laskey, Ph.D., is Professor of Systems Engineering and Operations Research at George Mason University and Associate Director of the Center of Excellence in Command, Control, Communications, Computing and Intelligence (C4I Center). She teaches and performs research on multisource information fusion, decision theoretic knowledge representation and reasoning methodology, data analytics, and decision support. A major focus of her research has been knowledge representation and reasoning for higher level multi-source fusion to support situation awareness and decision support. She has performed research in diverse application areas, including modeling the emplacement of improvised explosive devices, detecting insider threats, predicting aircraft delays, managing terrorist risk at public facilities, and planning military engagements. Dr. Laskey developed multi-entity Bayesian networks (MEBN), a language and logic that extends classical first-order logic to support probability. She was a key contributor to the development of the PR-OWL language for representing uncertainty in OWL ontologies. She serves on the ISIF Board of Directors and has is co-founder and active participant in the ISIF Evaluation of Techniques for Uncertainty Management Working Group (ETURWG). She serves on the Board of Directors of the Washington Metropolitan Area chapter of INCOSE and is past board chair of the Association for Uncertainty in Artificial Intelligence. Dr. Laskey served on several boards and committees of the United States National Academy of Sciences.

Erik Blasch is a program officer at the Air Force Research Laboratory (AFRL)—Air Force Office of Scientific Research (AFOSR) in Arlington, VA. Previously he was he was a principal scientist at the AFRL Information Directorate in Rome, NY, USA (2012–2017), exchange scientist to the Defence Research and Development Canada (DRDC) in Valcartier, Quebec (2010–2012), and Information Fusion Evaluation Tech Lead for the AFRL Sensors Directorate—COMprehensive Performance Assessment of Sensor Exploitation (COMPASE) center in Dayton, OH (2000–2009). Additional assignments include USAF Reserve Officer Col supporting intelligence, acquisition, and space technology. He was an adjunct associate professor in Electrical and Biomedical Engineering (2000–2010) at Wright State University and the Air Force Institute of Technology (AFIT) teaching classes in signal processing, electronics, and information fusion as well as research adjunct appointments at the Univ. of Dayton (2001–2014), Binghamton University (2012–2017), and Rochester Institute of Technology (2015–2017).

Dr. Blasch was a founding member of the International Society of Information Fusion (ISIF), (www.isif.org), 2007 President, and Board of Governors (2000–2010). He served on the IEEE Aerospace and Electronics Systems Society (AESS) Board of Governors (2011–2016), distinguished lecturer (2012–2018), co-chair of 5 conferences, and associate editor of 3 academic journals. He has focused on information fusion, target tracking, robotics, and pattern recognition research compiling 800+ scientific papers and book chapters. He holds 25 patents, received 33 team-robotics awards, presented 60+ tutorials, and provided 9 plenary talks. His co-authored books include High-Level Information Fusion Management and Systems Design (Artech House, 2012), Context-enhanced Information Fusion (Springer, 2016), Multispectral Image Fusion and Colorization (SPIE, 2018), and Handbook of Dynamic Data Driven Applications Systems (Springer 2018).

Dr. Blasch received his B.S. in Mechanical Engineering from the Massachusetts Institute of Technology ('92) and Masters' Degrees in Mechanical ('94), Health Science ('95) and Industrial Engineering (Human Factors) ('95) from Georgia Tech and attended the University of Wisconsin for a MD/PhD Neuroscience/Mechanical Engineering until being call to military service in 1996 to the United States Air Force. He completed an MBA ('98), MS Econ ('99), and PhD ('99) in Electrical Engineering from Wright State University and is a graduate of Air War College ('08). He is the recipient of the IEEE Bioengineering Award (Russ-2008), IEEE AESS magazine best paper Award (Mimno-2012), Military Sensing Symposium leadership in Data Fusion Award (Mignogna-2014), Fulbright scholar selection (2017), and 15 research/technical and team awards from AFRL. He is an American Institute of Aeronautics and Astronautics (AIAA) Associate Fellow, Society of Photonics and Industrial Engineers (SPIE) Fellow, and Institute of Electrical and Electronics Engineers (IEEE) Fellow.

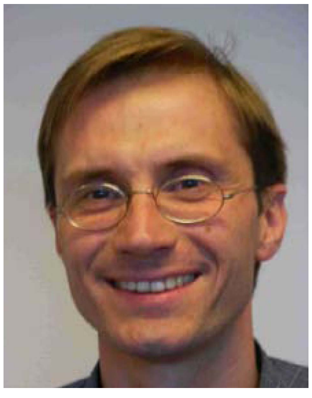
Dr. Valentina Dragos is a research scientist, member of the Department of Information Modeling and Systems at ONERA, The French Aerospace Lab in Palaiseau, France. Valentina received Master and PhD degrees in Computer Science from Paris V University and area of research interest is in artificial intelligence, with emphasis on natural language processing, semantics technologies and automated reasoning. Since joining ONERA in 2010, Valentina has been active in several national and EU projects focused on crisis management, maritime surveillance and cyber terrorism. For those projects, her contributions addressed various topics such as: semantic interoperability for command and control systems, heterogeneous information fusion, exploitation of semantic data (HUMINT, OSINT) for situation assessment and analysis and exploration of social media. Valentina is currently involved in NATO Research Task Groups, focusing on social media exploitation for operations in the information environment.



Juergen Ziegler is a Senior Technical Manager for Information Fusion at Industrieanlagenbetriebsgesellschaft mbH. He is a member of the department Competence Centers ISR.

His interests comprise the areas of Situational Awareness using methods of higher-level information fusion with applications in identification, Cyber Situational Awareness, medical diagnostics and reconnaissance. He is an expert in applications of Bayesian networks. One of the main issues of his work is model building with a focus on automatic generation of models, automatic generation of Bayesian Networks, assessment of the quality of knowledge models and ergonomic aspects of model generation. Another focus of his work is interoperability of data exchange and data fusion for situational pictures. He was one of the main authors of a STANAG about identification (STANAG 4162 Edition 3).

Pieter de Villiers is an associate professor at the University of Pretoria, South Africa and was a principal researcher at the Council for Scientific and Industrial Research (CSIR) until October 2017. He obtained his Bachelors and Masters degrees at the University of Pretoria, South Africa, and a PhD in 2008 at the University of Cambridge, UK, in statistical signal processing (particle filtering). From 2010 until 2018 he was performing research into data fusion at the Radar and Electronic Warfare competency at the CSIR. His research interests include data fusion, target tracking, Bayesian inference, nonlinear filtering, pattern recognition, graphical models and machine learning. Pieter has been regularly attending the International Conference of Information Fusion since 2010 and his ISIF activities include membership of the technical program committees, a tutorial selection committee and acting as session chairs over the years. He is the general co-chair for the 23rd International Conference on Information Fusion to be held in 2020 in South Africa. Pieter is a member of the official ISIF Evaluation Techniques for Uncertainty Representation and Reasoning Working Group (ETURWG). He is also a guest editor for a special issue at the Journal of Advances in Information Fusion (JAIF).



Gregor Pavlin received the M.Sc. degree in theoretical engineering and the Ph.D. degree in computer science from Graz University of Technology, Austria in 1995 and 2001, respectively. He has extensive industrial experience in safety critical software systems as well as complex AI-driven solutions. His current research interests are (i) robust algorithms and architectures supporting distributed probabilistic AI, (ii) machine learning and (iii) interoperability in complex service oriented processing systems. Since 2006 he has been a senior researcher and project manager at a corporate research lab of the Thales Group in Delft, the Netherlands. Between 2006 and 2015, he was also a part-time visiting researcher at the Intelligent Autonomous Systems lab, University of Amsterdam. He also has an extensive experience with the coordination of European and national collaborative projects. He served in the organizing committee of the Fifth International Symposium on Intelligent Distributed Computing in Delft (IDC 2011) and is also a member of the organizing committee of the 23rd International Conference on Information Fusion to be held in South Africa in 2020. He is also a member of the official ISIF Evaluation Techniques for Uncertainty Representation and Reasoning Working Group (ETURWG).