

High Level Fusion and Predictive Situational Awareness with Probabilistic Ontologies

Paulo Cesar G. Costa

Center of Excellence in C4I
George Mason University, MS 4B5
Fairfax, VA 22030-444 U.S.A.
pcosta@c4i.gmu.edu

Kathryn B. Laskey

Center of Excellence in C4I
George Mason University, MS 4B5
Fairfax, VA 22030-444 U.S.A.
klaskey@c4i.gmu.edu

Kuo-Chu Chang

Center of Excellence in C4I
George Mason University, MS 4B5
Fairfax, VA 22030-444 U.S.A.
kchang@c4i.gmu.edu

Rommel N. Carvalho

Center of Excellence in C4I
George Mason University, MS 4B5
Fairfax, VA 22030-444 U.S.A.
rommel.carvalho@gmail.com

Abstract - *The change of focus in modern warfare from individual platforms to the network has caused a concomitant shift in supporting concepts and technologies. Greater emphasis is placed on interoperability and composeability. New technologies such as SOA and semantically aware systems have come into the spotlight. This paper argues that just as the problem space demands interoperability of diverse technologies, so must the solution space. In other words, not only are new approaches needed, but they must also come together as a seamlessly interoperable technological tool set. This can be accomplished only via a consistent multi-disciplinary approach. In this paper, we present some of the major requirements of today's Predictive Situation Awareness Systems (PSAW), propose our approach as a coordinated mix between state-of-the-art research efforts, and present the architecture for enabling our approach.*

Keywords: probabilistic reasoning, naval predictive situational awareness, web services, Bayesian networks, MEBN, PR-OWL, probabilistic ontologies, distributed hybrid inference, spatio-temporal hybrid analysis.

1 Introduction

Knowledge fusion aims to produce a dynamic, comprehensive, and accurate battlespace picture for the warfighter that integrates tactical data from multiple intelligence sources. Until the recent past, battlespace information systems have usually been self-contained solutions merging data from a finite array of sensors, sending their data via proprietary protocols, and following predefined schemas. As technology evolved at an ever-increasing pace, and as increasing bandwidth and connectivity made possible the exchange of enormous volumes of data, old-style stovepipe systems became obsolete. The new concept of NetCentric Operations changes the focus from individual platforms to the network. This change in focus has in turn caused a shift in supporting concepts and technologies. Greater emphasis is placed on requirements such as interoperability and composeability. New concepts and

technologies such as SOA and semantically aware systems have come into the spotlight.

A recurring phenomenon during times of fundamental conceptual and technological change is the attempt to solve new problems with old tools. This is analogous to hanging onto old doctrinal frameworks in the face of new technologies. As a specific example, it is akin to employing stealth fighters with Vietnam-era tactics. Just as new weapons technologies demand changes in doctrine, advances in information processing capability require fundamentally new fusion technologies for delivering value to the warfighter. These new technologies must be developed with the new realities in mind, and must be tailored to exploit the new capabilities. When faced with today's increasingly complex problem of distributed knowledge fusion, one has to realize that just as the problem space demands interoperation of diverse areas of knowledge, so must the solution space. In other words, not only are new approaches needed, but they must also act as a seamless technological whole. This can be accomplished only through a consistent, scientifically principled, multi-disciplinary approach.

In section 2 we address the requirements for a predictive situation awareness (PSAW) system and the major issues that must be faced when attempting to meet the requirements. Our approach of combining a suite of technologies into an integrated technological tool set is introduced in section 3. Finally, section 4 presents PROGNOS (**PR**obabilistic **Ontolo**Gies for **Net**-centric **Operation** Systems), a system that applies our multi-disciplinary approach to address the problem of predictive situational awareness within the maritime operations domain.

2 PSAW System Requirements

Information in the battlefield is derived through reports coming from diverse sources, each with its own distinct syntax, and each having different semantics. There are many kinds of uncertainty involved in this process, e.g., noise in sensors; incorrect, incomplete, or deceptive hu-

man intelligence; missing or incorrectly recorded data. Because of this, it is essential to have a coherent, consistent, and principled means to represent different kinds of uncertainty, and to communicate uncertainty among the systems performing PSAW. Furthermore, lack of understanding of cause and effect mechanisms in the world as represented by the systems' models is a major source of interoperability difficulties. Effective interoperability requires understanding the relationship between reports from different systems and the events reported upon. Thus, the representational framework has to be expressive enough to capture not only the uncertainty but also the subtle and intricate relationships among entities of different types.

In addition, distributed PSAW involves dealing with large amounts of data of diverse types. This generates a requirement for automated integration of diverse data types (e.g., geolocation, feature, and identification) from multiple intelligence sources, in a consistent and timely manner. In a distributed environment, this demands not only better algorithms, but also the ability to represent and communicate semantic information, including information about uncertainty.

Finally, because modern PSAW systems must operate in a distributed environment, we assume a SOA architecture that enables interoperation through a series of information exchanges, usually dealing with data and/or software modules residing at different physical locations and controlled by different organizations, and producing real-world effects as a result of those interactions.

A brief analysis of the above high-level requirements led us to the conclusion that in order to fulfill its objectives a distributed PSAW system must:

- a. Provide interoperable methodologies for propagating uncertainty through the integration process to characterize and distinguish situational conditions for predictive analysis and impact assessment under various behaviors and environments.
- b. Have a rigorous mathematical foundation and efficient algorithms to combine data from diverse sources for reliable predictive situation assessment.
- c. Include automated techniques to reduce users' information processing load and provide timely actionable knowledge to decision makers.
- d. Operate in a distributed environment where data sources may be geographically dispersed, lack common syntax and semantics, have distinct owners and comply with diverse exchange policies.

The inherent complexity implied by each of the above items makes it clear that there is no and there will not be a silver bullet to address them. Also, applying effective techniques from distinct areas of knowledge without a coordinated approach would have the potential to

create a complex, "Hydra" system whose management would be nightmarish. Thus, in this paper we propose to develop a consistent multi-disciplinary approach, which main components we describe in the next section.

3 A Multi-Disciplinary Approach

We have already argued in this venue [1] that semantically aware systems are essential to distributed knowledge fusion, and that probabilistic ontologies can provide semantic awareness while also establishing the framework for representing and reasoning with uncertainty in a principled way. Thus, to address item "a" from the previous section our multi-disciplinary approach includes a probabilistic ontology language, PR-OWL [2, 3]. PR-OWL's basis in Bayesian first-order logic, MEBN [4, 5], ensures logical coherence. Our PSAW efforts include incorporating needed enhancements to the representation framework, and applying the framework to develop PR-OWL ontologies representing key aspects of PSAW knowledge.

To address item "b", the representational framework of MEBN/PR-OWL functions as a basis for the development of mathematically rigorous and computationally efficient algorithms for Spatio-Temporal Hypothesis Management [6] and Efficient Hybrid Inference [7-9]. These efficient approximate inference methods are essential to the engineering success of our predictive situational awareness framework.

Items "c" and "d" demand not only the representational framework and efficient reasoning algorithms above cited, but also the development of computational tools to ensure usability and scalability of the overall solution as well as interoperability among its components. We are addressing both items through a series of additions and improvements to the Bayesian package UnBBayes-MEBN [10, 11], which will be used as the main platform for our technological suite.

3.1 The Representational Framework

Bayesian probability provides a mathematically sound representation language and formal calculus for rational degrees of belief, which gives different agents the freedom to have different beliefs about a given hypothesis. This provides a compelling framework for representing uncertain, incomplete knowledge that can come from diverse agents. Bayesian Networks (BNs) provide a means of parsimoniously expressing joint probability distributions over many interrelated hypotheses. A Bayesian network consists of a directed acyclic graph (DAG) and a set of local distributions. Each node in the graph represents a random variable. A random variable denotes an attribute, feature, or set of hypotheses about which an agent may be uncertain. The graph represents direct qualitative dependence relationships; the local distributions represent quantitative information about the strength of those dependencies. The graph and the local distributions together represent a joint probability distribution over the random vari-

ables denoted by the nodes of the graph. We have argued elsewhere (e.g. [12]) that most complex problems cannot be addressed by BNs due to their limited attribute-value representation. That is, each problem instance in a BN is limited to the same fixed number of attributes, with only the evidence values changing from problem instance to problem instance. MEBN overcomes this limitation by adding first-order expressive power. Anything that can be expressed in first-order logic (FOL) can be assigned a probability by MEBN logic. It represents the world as consisting of entities that have attributes and are related to other entities. Knowledge about the attributes of entities and their relationships to each other is represented as a collection of MEBN fragments (MFrag) organized into MEBN Theories (MTheories). An MFrag represents a conditional probability distribution for instances of its resident random variables given their parents in the fragment graph and the context nodes. An MTheory is a set of MFrag that collectively satisfies consistency constraints ensuring the existence of a unique joint probability distribution over instances of the random variables represented in each of the MFrag within the set. MEBN semantics integrates the standard model-theoretic semantics of classical first-order logic with random variables as formalized in mathematical statistics.

MEBN's FOL expressiveness is the logical basis for our PSAW systemic approach. An MFrag can be roughly compared to a BN template that can be instantiated as many times as needed to build a Situation-Specific Bayesian Network (SSBN) in response to a system's query. In a generic system concept, knowledge about entities and their respective attributes is encoded as a set of MFrag. When a query is posed to the system (e.g. in a PSAW problem), a MEBN algorithm is run to build a SSBN for answering the query. The set of MFrag encoding the knowledge of a given domain is called an MTheory and is stored in PR-OWL format.

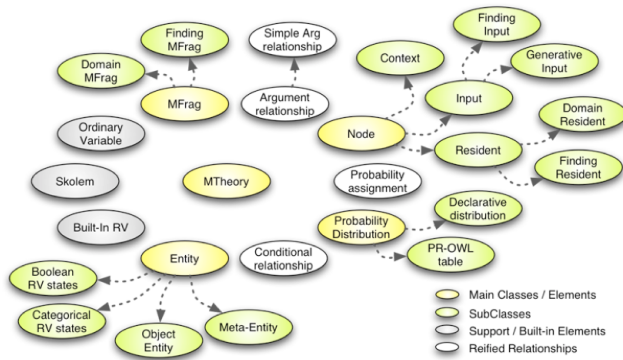


Figure 1. PR-OWL Classes.

Among other useful features, MEBN logic provides the expressive power to represent and reason about hypothetical entities. Uncertainty about whether a hypothesized entity actually exists is called existence uncertainty. Efficient approximate inference about existence uncertainty, as provided by the Spatio-Temporal Hypothesis Management algorithms, enables predictive situational awareness

for situations in which the number of entities involved is unknown. This is an essential capability for PSAW.

PR-OWL [2] is an OWL upper ontology for representing MTheories. Its classes and properties allow the ontology engineer to specify MFrag while maintaining compatibility with the widespread, W3C recommended OWL ontology language. The main classes of PR-OWL are depicted in Figure 1, and a complete explanation of the format can be found in [2] or at <http://www.pr-owl.org>.

As an ontology language, PR-OWL supports logical reasoning. Thus, PR-OWL ontologies can represent the knowledge needed for tasks such as inferring the class structure from the asserted properties. As pointed out in [13], this is in contrast with the Object Oriented (OO) modeling paradigm, in which instances are created as members of some class, and their behavior is specified by the class structure. As such, changing the class structure of an OO model would have a direct impact on the whole system's behavior, whereas in an OWL or in a PR-OWL ontology the class structure is inferred from its asserted properties, which also dictate the class membership of its instances. While the OO paradigm works well in closed systems with centralized control of the data schema, the ontology engineering paradigm is more suitable to situations in which such control either does not exist or is difficult to enforce. Therefore, the inherent flexibility of PR-OWL's data modeling paradigm is a perfect fit for a distributed PSAW system, in which information may come from many systems, possibly following distinct data schemas and semantics.

One limitation for using OWL in PSAW systems is its lack of support for principled uncertainty representation and, consequently, for plausible reasoning. As an example from the naval domain, suppose a USS Destroyer sends a query regarding ships within a 100 NM radius that might be involved in illicit activities. Using its own system, with its own data structure and internal logic, intelligence agency I2 reports back about a dhow whose owner is within the same social network as a person suspected of plotting a terrorist attack. Intelligence reports and sensor information are usually incomplete and plagued with uncertainty, and thus would not be sufficient for a logical reasoner to make any conclusions about whether the dhow is involved in an illicit operation. In our example, a probabilistic reasoner can use the same information to provide an update on the chances of this dhow being involved or not in an illicit operation. In addition, it is reasonable to assume that additional inferences or data on this dhow and on other ships as well will also be arriving at the Destroyer's system in response to the request, and will allow for an almost continuous update on the respective chances that each vessel within the stated radius is involved in illicit activity. In other words, expert knowledge stored in a probabilistic-capable ontology format allows for data coming from various sources to be almost continuously updated, resulting in automated support for comprehensive situational awareness. Most tasks in a PSAW system

require reusable patterns of knowledge about events in space and time. In an operational system, this type of reasoning would make use of available “legacy” deterministic ontologies by the sources. PR-OWL allows the user of such an ontology to add probabilistic information to represent uncertain relationships.

3.2 Achieving Sufficient Inferential Power

The combination of probability with first-order logic within the MEBN/PR-OWL framework described above is part of the state-of-the-art research that has greatly expanded the range of problems that can be tackled by automated fusion systems. However, for problems of the scale required for predictive analysis, exact evidential reasoning is generally intractable. Traditional fusion systems cope with complexity by decomposing the problem into hypothesis management and inference. Hypothesis management produces an approximate model that achieves tractability by combining similar hypotheses and/or pruning unlikely hypotheses and tracks. For the higher-level fusion problems considered here, the concept of a track must be generalized to a complex spatio-temporal entity that is related to and interacts in varied ways with other evolving spatio-temporal entities.

An expressive Bayesian logic such as MEBN permits the expression of sophisticated hypotheses about unbounded numbers of entities and their interrelationships. In a given situation, a situation-specific Bayesian network (SSBN) can be constructed from the generic MEBN domain model to reason about the actual entities involved. In general, there will be uncertainty about the number of entities in the situation, their relationships to each other, their past and future behavior, and the association of reports to entities. Hypothesis management for MEBN domain models must be appropriately generalized to apply to complex interacting spatio-temporal entities [6]. Methods from the multi-target tracking literature can be generalized to search over the vast number of hypotheses [14]. In our approach, we employ a MCMC [15] hypothesis management (MC2HM) module to nominate, refine, and prune hypotheses.

In an efficient distributed hybrid inference scenario, as reports about a situation arrive, the predictive situation awareness system begins an interleaved process of hypothesis management and predictive inference. Conceptually, we can think of hypothesis management and model construction as producing a Bayesian network for reasoning about a given situation. In a network-centric architecture, the inference task would be distributed among geographically dispersed and functionally distinct subprocesses, each representing aspects of the problem relevant to its own function. Our approach employs Multiply-Sectioned Bayesian networks (MSBN) [16], a computational architecture for distributed inference in large Bayesian networks.

The prediction problem involves reasoning in space and time, and requires both discrete and continuous ran-

dom variables, which may not be Gaussian. This poses a computational challenge, because traditional Bayesian network inference algorithms are limited to discrete random variables or to linear Gaussian continuous random variables. We apply the HMP-BN algorithm [7-8], and efficient approximate inference method based on distributed message passing in hybrid discrete and continuous Bayesian networks. HMP-BN uses the unscented transformation [9] to approximate arbitrary continuous transformations of arbitrary continuous distributions. The unscented transformation has been shown to be more accurate than traditional linearization methods.

3.3 Improving UnBBayes-MEBN

In order to realize the approach we have described in the sections above, we teamed up with the developers at the University of Brasilia and started a series of improvements to the MEBN reasoner they have developed with our support, UnBBayes-MEBN [10, 11]. In this section, we comment on the most recent work on the platform to make it suitable for using within a PSAW system based on our approach.

The most far-reaching modification we are currently implementing to UnBBayes-MEBN is to make it a full ontology editor. At this point, the package can be seen as a MEBN tool that can perform reasoning (i.e. build a SSBN over an MTheory upon the receipt of a query and then apply a BN algorithm to it) and save the model in PR-OWL format. This is a powerful capability, but it does not provide the ability to build a complete ontology with properties, classes, and restrictions other than those included in the MTheory. The desired end state is to have a tool capable of not only representing and reasoning with uncertainty, but to also provide an interface and support to build both the deterministic and probabilistic aspects of a probabilistic ontology. As we have stated earlier, merging the ability to provide support for both types of knowledge representation and reasoning in a single environment is no minor achievement, and one that is the current subject of many research efforts. Our vision of how to achieve this end state is set forth in [2], and the basic concept is illustrated in Figure 2.

The figure depicts a plugin for the OWL Protégé editor (available from <http://protege.stanford.edu/>), being used to construct an MFragment using a GUI similar to a standard BN editor. The idea of such a plugin is to hide from users the complex constructs required to convey the many details of a probabilistic ontology, such as the reified relationships, composite random variable term constructions (with or without quantifiers and Exemplar constants), and others. In the figure, an MFragment was selected from the combo box in the top of the viewing area, thus information about its nodes is displayed in a graphical format that allows the user to build more nodes, edit or view the existing ones. When a node is chosen, such as the node ZoneEShips(z) in the picture, it appears highlighted (a red box around it) and all its data is shown in

The screenshot shows the UniBay software interface. The main window displays a Bayesian network diagram titled "AttackPlanIFrag". The diagram includes the following nodes and their relationships:

- Green nodes (Labels):**
 - `isAgntPrsnOffset_label`
 - `isAgntPrsn_label`
 - `isAgntVenue_label`
 - `isAO/PrsnOffset_label`
- Yellow nodes (Igg):**
 - `PoliticalImportanceIgg`
 - `isActivePlanIgg`
 - `isEscapeAWIgg`
 - `isSnatchIgg`
 - `isWeaponSupplierIgg`
 - `SupplierVpnIgg`
- Pink nodes (Igg):**
 - `isTargetIgg`
 - `isAgntOfIgg`

The directed edges (dependencies) are as follows:

- `isAgntPrsnOffset_label` points to `isTargetIgg`.
- `isAgntPrsn_label` points to `isTargetIgg`.
- `isAgntVenue_label` points to `isTargetIgg`.
- `isAO/PrsnOffset_label` points to `isTargetIgg`.
- `PoliticalImportanceIgg` points to `isTargetIgg`.
- `isActivePlanIgg` points to `isTargetIgg`.
- `isActivePlanIgg` points to `isAgntOfIgg`.
- `isEscapeAWIgg` points to `isAgntOfIgg`.
- `isSnatchIgg` points to `isAgntOfIgg`.
- `isWeaponSupplierIgg` points to `isAgntOfIgg`.
- `isAgntOfIgg` points to `SupplierVpnIgg`.
- `SupplierVpnIgg` points to `isTargetIgg`.

The left pane shows the "Theory Tree" with the following structure:

- LahoreBombing
 - AttackPlanIFrag
 - SuppliesIggOffset
 - PoliticalImportance
 - isAgntOf
 - isAgntPlan
 - isTarget
 - isEscapeIgg
 - isSnatch
 - isWeaponSupplier
 - LocationIggOffsetIgg
 - HUMINTLocationIgg
 - MissingCarReport
 - AgentIgg
 - isEscapeIgg
 - isWeaponSupplier
 - LocationIggOffsetIgg
 - SocialNetworkIgg
 - isSnatch
 - isWeaponSupplier
 - isWeaponSupplier
 - isAgntOf
 - isAgntOf
 - LocationConstraintIgg
 - LocationConstraint

The bottom status bar indicates "Ready".

Another important modification to the current UnBBayes-MEBN is support for continuous random variables. PSAW systems operate within an environment in which data from interacting spatio-temporal entities is represented in the form of continuous variables (such as spatial coordinates, time periods, or continuous attributes such as mass or length). Uncertainty about these quantities must be represented using continuous distributions. Thus, having the ability to perform Bayesian reasoning with continuous variables is an essential feature any successful PSAW system must provide. This problem has been already addressed in our group (e.g. [7]) and is now being implemented in UnBBayes-MEBN for application to problems in PSAW.

A final aspect of our work on improving UnBBayes-MEBN is related to the quality of the merging process. More specifically, when data that is reliable and accurate is combined with inaccurate or biased data (especially if the uncertainties or variances of the data are unknown [19]), the result could be worse than what would be obtained by tasking the most appropriate sensor in a sensor suite, thus defeating the purpose of data fusion per se. To address that, we are developing a new evaluation module in UnBBayes that allows a user to evaluate the classifica-

This evaluation module is based on the Fusion Performance Model (FPM) [21]. The focus is in on classification performance as described in [22]. This module is very valuable for a decision maker to analyze trade-off between performance and costs and to select proper sensor suites according to requirements and constraints.

4 Use Case: PROGNOS

In the conceptual view of Figure 3, the hypothesis management process passes results to the inference process, which builds a Bayesian network to predict future events.

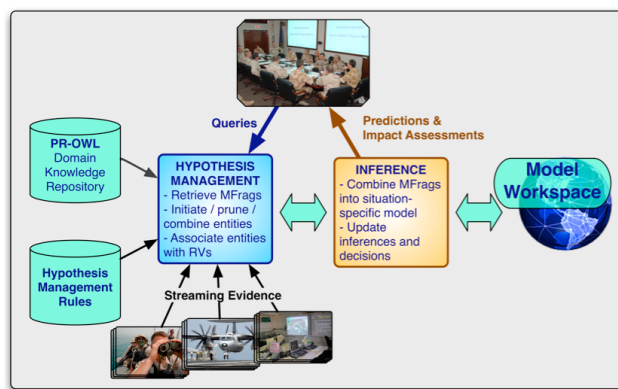


Figure 4 shows a broader concept for employing a MEBN/PR-OWL-based system in a distributed net-centric SOA. The bar represents the loosely coupled relationship between service consumers and providers. PROGNOS architecture uses probabilistic ontologies to fill a key gap in semantic matching technology [23], facilitating wide-

spread usage of Web Services for efficient resource sharing in uncertain open and distributed environments.

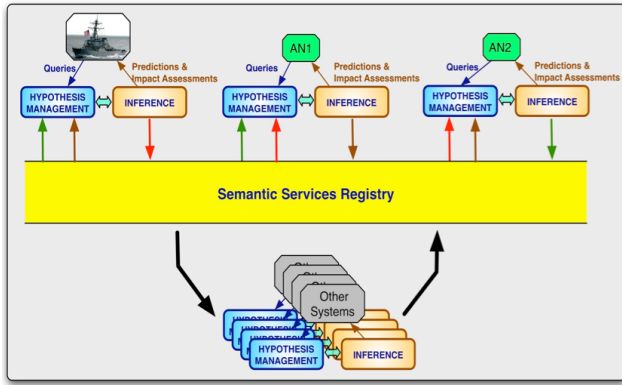


Figure 4. Distributed Predictive Situation Assessment and Impact Assessment.

The conceptual view of Figures 3 and 4 will be implemented according to the architecture depicted in Figure 5, which shows the major components of the PROGNOS system. According to this architecture, each FORCENet platform (e.g., a ship) would have its own system that receives information from the platform's sensors and from its FORCENet peers. It is assumed that these inputs provide a fairly precise tactical view in which the geographical position of the entities surrounding the platform is known and well discriminated. The platform is also a peer in FORCENet and exchanges data and knowledge as services with its peers

The high level architecture depicted in the diagram was devised to provide a scalable, easily maintainable system with five independent modules. We now present each module at a greater level of detail.

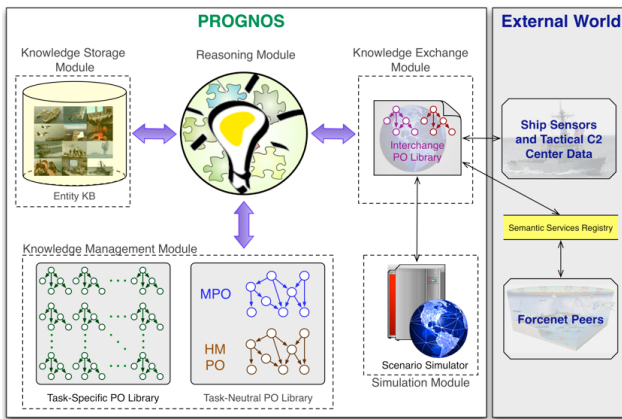


Figure 5. Distributed Predictive Situation Assessment and Impact Assessment – Component Architecture.

4.1 The Reasoning Module

The reasoning module is the heart of the PROGNOS system, responsible for performing all of its reasoning services. It is composed of a MEBN reasoner that interacts

with the other modules and coordinates the execution of SSBN construction, which includes interleaved hypothesis management and inference within the constructed SSBN. In response to a query, the MEBN Reasoner relies on the system's Knowledge Management Module to define the information necessary to answer the query. Then, it starts the SSBN construction process that will include successive accesses to the Knowledge Storage Module to retrieve all available information pertinent to the process and to support a continuous cycle of hypothesis formation, evaluation, and pruning that will run until it succeeds in creating the SSBN required to answer that query given the information at hand. During this process, external sources of knowledge may be queried via the Knowledge Exchange Module, which provides an advanced interface between the system and the external world. Finally, for training, evaluation, or other specific purposes, this interaction may be simulated via the Simulation Module.

4.2 The Knowledge Storage Module

A MEBN-based system needs to have a means of keeping track of the entities it is reasoning about. In PROGNOS, this task is performed by the Knowledge Storage Module, which has the Entity KB as its major component. There, every track and its respective data are stored within a schema based on and dynamically linked to the PROGNOS system's MPO (Main Probabilistic Ontology).

4.3 The Knowledge Management Module

If the reasoning module is the heart that runs and coordinates the system's algorithms, then the Knowledge Management Module can be seen as the brains of the system, which is responsible for understanding the situation at hand and defining how to proceed in face of a situation. The module contains a set of probabilistic ontologies that capture domain knowledge in the form of MFragments. There are two distinct libraries, one comprised of POs representing task-dependent knowledge and the other containing two specific POs with knowledge that applies to any task. The latter is called Task-Neutral PO Library, and includes the Main Probabilistic Ontology (MPO), which captures concepts that are routinely used by the system (e.g. properties of entities, naval terms and possible meanings, relationships between those, etc). The second PO of the Task-Neutral PO Library, the Hypothesis Management PO (HMPO), is focused on MFragments capturing the knowledge used in the Hypothesis Management process. It is kept separate from the MPO to facilitate maintenance and scalability. The other set of POs is the Task-Specific PO Library, which contains probabilistic ontologies pertaining to particular types of mission or domain about which PROGNOS needs to reason. In other words, it is a library of POs that are used in support of specific mission types, and can thus be upgraded or modified to reflect changes in the specific task-related concepts without requiring changes in the MPO, HMPO, or other system resources.

4.4 The Simulation Module

This module consists of the Scenario Simulator, which generates tracks in order to simulate the situations depicted in the case studies supporting the analysis. Basically, it sends geographical data (coordinates, known or probable) and status (friend, foe, unknown, etc.) of fictitious entities that used to evaluate the system's response. In the deployed PROGNOS system, this module would be connected to the system via the Knowledge Exchange Module and can be reconfigured to support system maintenance and simulation drills.

4.5 The Knowledge Exchange Module

PROGNOS continuously exchanges knowledge with the platform's sensors and tactical C2 system, the Simulation Module, FORCEnet peers, and other networked systems. This module, whose main component is the Interchange PO Library, manages all those connections. Internal exchanges between the Reasoning Module and the platform's sensors and tactical C2 system, or the Simulation Module are performed via a direct link using a common protocol. External exchanges, in the majority of the cases, will be performed between PROGNOS and peers using a common SOA standard throughout FORCEnet. However, there will be cases in which the system might need to exchange knowledge with non-FORCEnet peers that do not conform to SOA standards. For those situations, PROGNOS relies on a set of interchange POs to support interoperability. As an example, if exchanging information with a JC3IEDM compliant system, PROGNOS would base its messages on a JC3IEDM PO, while interchange with other systems might either require a specifically built PO or may be managed by a general interchange PO. In any case, all should be part of the Interchange PO Library.

5 Hypothesis Management

Hypothesis management is one of the most complex research challenges posed to the PROGNOS team, and we have been developing a new approach for the problem of both managing existing hypotheses and creating new ones as the situation evolves. During this initial period, we have developed a preliminary design for the HM engine, which is depicted in figure 6.

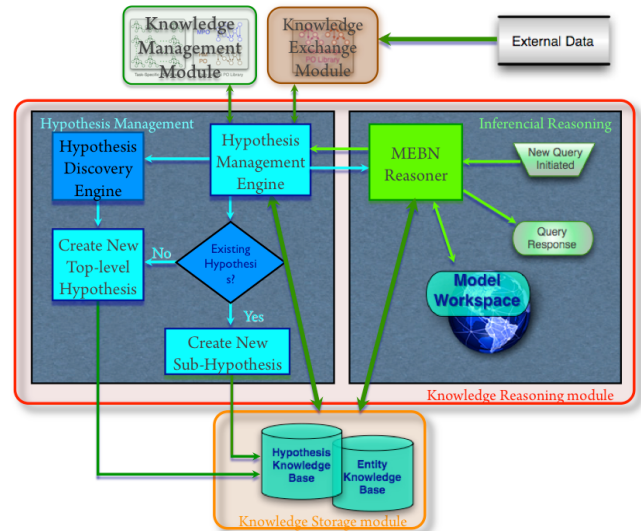


Figure 7 – Hypothesis management engine preliminary concept

HM produces an approximate model that achieves tractability by combining similar hypotheses and/or pruning unlikely hypotheses and tracks. In the diagram above, incoming data or new queries trigger a new cycle of the hypothesis management process. Both are assessed with respect to relevance and impact to existing hypothesis, either creating a new hypothesis or extending an existing one.

In other words, new data or queries are analyzed in conjunction with the knowledge management POs (i.e. existing knowledge within the system), and can not only refine the knowledge on existing hypothesis but also trigger the hypothesis discovery engine to create new ones.

6 Conclusions

This paper addresses a complex problem with requirements spanning diverse areas of knowledge. Although the approach presented here is still in the research stage and thus is a work in progress, the ideas behind its implementation have the potential to provide insights to researchers and practitioners in the field of predictive situation awareness.

References

- [1] K.B. Laskey, P.C.G. Costa, and T. Janssen, "Probabilistic Ontologies for Knowledge Fusion", 11th International Conference on Information Fusion (Fusion 2008), Cologne, Germany, June 30 – July 08, 2008.
- [2] P.C.G. Costa, *Bayesian semantics for the Semantic Web*, Doctoral Dissertation, Volgenau School of Information Technology and Engineering, George Mason University, Fairfax, VA, 2005.

- [3] P.C.G. Costa, and K.B. Laskey, "PR-OWL: A framework for probabilistic ontologies", Proceedings of the Fourth Conference on Formal Ontologies and Information Systems (FOIS 2006), B. Bennet and C. Felbaum (eds.), Baltimore, MD, USA, November 9-11, 2006, pp. 237-249.
- [4] K.B. Laskey, "MEBN: A Language for First-Order Bayesian Knowledge Bases", *Artificial Intelligence*, 172(2-3), 2007.
- [5] E.J. Wright, S. Mahoney, K.B. Laskey, M. Takikawa, and T. Levitt, "Multi-Entity Bayesian Networks for Situation Assessment", *Proceedings of NSSDF*, 2002.
- [6] Laskey, K.B., Mahoney, S.M. and Wright, E. (2001) Hypothesis Management in Situation-Specific Network Construction, *Uncertainty in Artificial Intelligence: Proceedings of the Seventeenth Conference*, San Mateo, CA: Morgan Kaufmann.
- [7] W. Sun, *Efficient Inference for Hybrid Bayesian Networks*. Doctoral Dissertation. Volgenau School of Information Technology and Engineering, GMU, Fairfax, VA, 2005.
- [8] W. Sun, and KC Chang "Hybrid Message Passing for General Mixed Bayesian Networks", *Proc. of the 10th International Conf. on Information Fusion*, Quebec, Canada, 2007.
- [9] W. Sun, and KC Chang, "Unscented Message Passing for Arbitrary Continuous Bayesian Networks". *Proceedings of the 22nd AAAI Conference on Artificial Intelligence*, Vancouver, Canada, 2007.
- [10] R.N. Carvalho, L.L. Santos, M. Ladeira, and P.C.G. Costa, "A Tool for Plausible Reasoning in the Semantic Web using MEBN". In *Proceedings of the Seventh International Conference on Intelligent Systems Design and Applications*, 2007, IEEE Press, pp. 381-386.
- [11] P.C.G. Costa, M. Ladeira, R.N. Carvalho, K.B. Laskey, L.L. Santos, and S. Matsumoto, "A First-Order Bayesian Tool for Probabilistic Ontologies". In Proceedings of the 21st International Florida Artificial Intelligence Research Society Conference (FLAIRS-21), Palo Alto: AAAI Press, 2008.
- [12] K.B. Laskey, and P.C.G. Costa, "Uncertainty Representation and Reasoning in Complex Systems", Chapter in *Complex Systems in Knowledge based Environments: Theory, Models and Applications*. Tolk, Andreas; Jain, Lakhmi C. (eds.), Springer, New York, NY, USA, 2009.
- [13] D. Allemang, and J. Hendler, *Semantic Web for the working ontologist: effective modeling in RDFS and OWL*. Elsevier, Burlington, MA, USA, 2008.
- [14] L.D. Stone, Barlow C.A., and T.L. Corwin, *Bayesian Multiple Target Tracking*. Artech House, Boston, MA, 1999.
- [15] Oh, S., S. Russell, et al. (2004). Markov Chain Monte Carlo Data Association for General Multiple-Target Tracking Problems. *Proceedings of the 43rd IEEE Conference on Decision and Control*. Atlantis, Paradise Island, Bahamas.
- [16] Y. Xiang. A probabilistic framework for cooperative multi-agent distributed interpretation and optimization of communication. *Artificial Intelligence*, 87(1-2):295-342, 1996.
- [17] M. Henrion, "Propagating Uncertainty in Bayesian Networks by Probabilistic Logic Sampling," in J. Lemmer and L. Kanal, editors, *Uncertainty in Artificial Intelligence*, North-Holland, Amsterdam, 1986, pp. 317-324.
- [18] S. L. Lauritzen and D. J. Spiegelhalter, "Local Computations with Probabilities on Graphical Structures and their Application in Expert Systems", *Journal Royal Statistical Society B*, 50, 1988.
- [19] D. Hall and J. Llinas, "An introduction to multisensor data fusion," *Proceedings of the IEEE*, vol. 85, 1997, pp. 6-23.
- [20] Rommel Carvalho and KC Chang, "A Performance Evaluation Tool for Multi-Sensor Classification Systems," submitted to *the 12th International Conf. on Information Fusion*, Seattle, 2009.
- [21] KC Chang and Y. Song, "Performance modeling for multisensor data fusion," *SPIE*, Orlando, FL, USA, 2003.
- [22] KC Chang, E. Sivaraman, and M. Liggins II, "Performance modeling for multisensor tracking and classification," *SPIE*, Orlando, FL, USA, 2004.
- [23] P.C.G. Costa, K.B. Laskey, E.J. Wright, and K.J. Laskey, "Probabilistic Ontologies: The Next Step for Net-Centric Operations", in *Proceedings of the 12th International Command and Control Research and Technology Symposium*. June 19-21, 2007, Newport, RI, USA, CCRP publications.