PROGNOS: APPLYING PROBABILISTIC ONTOLOGIES TO DISTRIBUTED PREDICTIVE SITUATION ASSESSMENT IN NAVAL OPERATIONS

Paulo Cesar G. da Costa*
Kathryn Blackmond Laskey
KC Chang

Center of Excellence for C4I
The Volgenau School of Information Technology and Engineering
George Mason University
4400 University Drive
Fairfax, VA 22030-4444
(703) 993-1644
[pcosta; klaskey; kchang]@gmu.edu

*Point of Contact: Paulo C G Costa
pcosta@gmu.edu and/or (703) 835-0751
Abstract

Achieving agile Command and Control in Maritime Operations requires composeability - the ability to construct responses “on the fly” to queries about a given situation, by discovering and drawing upon the appropriate resources from among the vast collection of resources existing on a distributed network. Although net-centric architectures such as FORCEnet provide the necessary connectivity and computational power needed to achieve fast and adaptive decision cycles, the sheer volume of data creates informational and cognitive bottlenecks that hinder agility. To address these limitations, new approaches bridging the gap from data interchange to knowledge interchange are needed, enabling C2 systems to produce a dynamic, comprehensive, and accurate battlespace picture. This is the main focus of PROGNOS, a system for Predictive Naval Situation Awareness currently being developed at George Mason University’s C4I Center. PROGNOS will integrate four state-of-the-art enabling technologies into a distributed system architecture that represents domain knowledge as a modular collection of probabilistic ontologies, combine these “knowledge nuggets” dynamically into complex situation models, and apply theoretically sound, computationally efficient hypothesis management and inference to combine evidence and background knowledge to reason about the current situation. PROGNOS will also interoperate with other FORCEnet systems by interacting via semantically enabled services.

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

Facing asymmetric threats in a network centric environment, modern military systems confront increasingly demanding challenges in information integration. Key requirements include interoperability with diverse systems, incorporation of a wide variety of traditional and non-traditional types of data coming from geographically dispersed sources, and processing huge volumes of noisy, incomplete and uncertain data in a timely manner. There is a driving need to provide dependable situational awareness for decision-makers within an environment such as FORCEnet [1] (depicted in Figure 1).

Advances in connectivity and computation are by themselves insufficient to meet the challenge. The sheer volume of data creates informational and cognitive bottlenecks. Incompatible formats and semantic mismatches necessitate tedious and time-consuming manual processing at various points in the decision cycle. As a result, massive amounts of potentially relevant data remain unexploited, and decision makers’ limited cognitive resources are too often focused on low-level manual data integration rather than high-level...
reasoning about the military situation. New approaches are needed to bridge the gap from data interchange to knowledge interchange, to free human operators from information overload and low-level manual tasks, and to provide them with actionable, decision-relevant information.

To address these challenging issues, the PROGNOS project combines state-of-the-art research on Multi-Entity Bayesian Networks [2-4], Probabilistic Ontologies [5-10], Spatio-Temporal Hypothesis Management [11-12], and Efficient Distributed Hybrid Inference [13-15] to develop new approaches to automated information integration, with application to predictive situation awareness in the maritime domain.

2. Enabling Technologies

PROGNOS is based on a synergic combination of the four enabling technologies. Figure 2 shows how these elements are integrated in a distributed predictive situation awareness system. A brief description of the enabling technologies follows.

2.1 Multi-Entity Bayesian Networks (MEBN): Combining Probability and Logic

Tasks at higher levels of the JDL fusion framework, such as the level 3 task of predicting threat behavior, require reasoning about complex situations in which entities of different types are related to each other in diverse ways. This is particularly true in an asymmetric warfare where the threats are elusive, secretive, and decentralized entities that often appear unconnected and engage in stealthy behavior that is difficult to predict. Automated methods for reasoning about such complex situations require expressive representation languages. Military situations are inherently uncertain, and the available data are inevitably noisy and incomplete. It is essential to be able to represent and reason with uncertainty. Recent years have seen rapid advances in the expressive power of probabilistic languages (e.g., [2, 16-21]). These new languages permit representation of and reasoning about highly complex situations. PROGNOS uses multi-entity Bayesian networks, a computational logic that combines the expressive power of first-order logic with the ability of Bayesian networks to represent and reason with uncertainty, as its logical basis [2]. A MEBN domain model implicitly represents a joint probability distribution over situations involving unbounded numbers of entities interacting in complex ways.

2.2 Probabilistic Ontologies

Ontologies provide the “semantic glue” to enable knowledge sharing among distinct systems cooperating in data rich domains. An ontology specifies a controlled vocabulary for representing entities and relationships characterizing a domain. Ontologies facilitate interoperability by standardizing terminology, allow automated tools to use the stored data in a context-aware fashion, enable intelligent software agents to perform better
knowledge management, and provide other benefits of formalized semantics. However, effective multi-INT fusion requires reasoning under uncertainty, and traditional ontology formalisms provide no principled, standardized means to represent uncertainty. Interest is growing in combining semantic technology with probabilistic reasoning (e.g., [5-10, 22-24]). Probabilistic ontologies ([5-10]) provide a principled, structured, sharable formalism for describing knowledge about a domain and the associated uncertainty and could serve as a formal basis for representing and propagating fusion results in a distributed system. The PR-OWL [25-26] probabilistic ontology language is founded in MEBN logic and has the expressive power to represent any first-order Bayesian theory. PR-OWL provides the necessary mathematical basis for information fusion and prediction services in net-centric environments. PROGNOS will employ PR-OWL within a distributed fusion and prediction architecture to enable approximate Bayesian inference on problems of greater complexity than previously possible.

2.3 Spatio-Temporal Hypothesis Management

As noted above, recent work on combining probability with first-order logic has greatly expanded the range of problems that can be tackled by automated fusion systems. However, for problems of the scale required for maritime 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 [11]. Methods from the multi-target tracking literature can be generalized to search over the vast number of hypotheses [12]. PROGNOS will employ a MCMC hypothesis management (MC2HM) module to nominate, refine, and prune hypotheses.

2.4 Efficient Distributed Hybrid Inference

As reports about a naval 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 sub-processes, each representing aspects of the problem relevant to its own function. PROGNOS will apply Multiply-Sectioned Bayesian networks (MSBN) [25], 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 random 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. PROGNOS will apply the HMP-BN algorithm [13-14], and efficient approximate inference method based on distributed message passing in hybrid discrete and continuous Bayesian networks. HMP-BN uses the unscented transformation [26] to
approximate arbitrary continuous transformations of arbitrary continuous distributions. The unscented transformation has been shown to be more accurate than traditional linearization methods.

3. An Architecture for Distributed Predictive Situation Assessment

Work on PROGNOS is integrating the abovementioned enabling technologies in a distributed system architecture. In this architecture, domain knowledge is represented as MEBN fragments, or MFrags, which define a joint probability distribution over situation variables (see Figure 2). Typically, MFrags represent relatively small, modular “knowledge nuggets” that are instantiated and combined to construct a complex situation model.

As streaming evidence arrives, the system matches evidence to existing hypotheses and/or nominates new hypotheses via MC2HM, generating an approximation to the posterior distribution of hypotheses given evidence. In the conceptual view of Figure 2, the hypothesis management process passes results to the inference process, which builds a Bayesian network to predict future events.

Figure 3 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 [7], facilitating widespread usage of Web Services for efficient resource sharing in uncertain open and distributed environments such as FORCEnet.

The conceptual view of Figures 2 and 3 will be implemented according to the architecture depicted in Figure 4, 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.

3.1 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 minimum 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.

3.2 The Knowledge Storage Module
MEBN logic represents the world as comprised of entities that have attributes and are related to other entities. Random variables represent features of entities and relationships among entities. Therefore, 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).

3.3 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 MFrags. 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 MFrags 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 that pertains to a given type of mission or domain about which PROGNOS needs to reason. In other words, it is a library of POs that are mostly 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.

3.4 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 are going to be 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’s maintenance and simulation drills.

3.5 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.

4. Illustrative Example: Anticipating and Preventing a Terrorist Incident
To illustrate the above concepts, we present a scenario in which the USS Carney (DDG64), part of Combined Task Force (CTF) 150, is conducting Maritime Security Operations (MSO) in the North
Arabian Sea, supporting Operation Enduring Freedom. MSO complements the counter-terrorism and security efforts of regional nations and seeks to disrupt violent extremists’ use of the maritime environment as a venue for attacking targets or for supporting its own logistic needs. The diagram in Figure 5 summarizes the situation being depicted in this example.

The FORCEnet sensor network available to CTF 150 is quite capable, but the dense naval traffic near Karachi makes spotting an illegal ship difficult. By means of its Knowledge Exchange Module, USS Carney’s PROGNOS receives information from diverse sources and monitors hundreds of ships sailing in the surrounding area. As part of its hypothesis management cycle, PROGNOS requests inferences regarding ships within a 100 NM radius from USS Carney that might be involved in illicit activities.

Less than an hour before USS Carney PROGNOS’ request was sent, in Lahore (Northern Pakistan), two intelligence analysts (AN1 and AN2) have collaborated to detect and prevent an attempted attack on a high-profile meeting. Although the two analysts were using different knowledge fusion systems, their collaboration was triggered by the arrest of a Lahore resident (P) attempting to depart for Karachi by plane when a canine unit detected explosive residue. P was declared a person-of-interest by AN2, initiating an automatic interaction adding P to AN1’s social network. This uncovered relationships with other agents and led to the halt of a terrorist attack plan [7]. We extend this scenario into the naval domain by postulating a link in AN1’s social network between P and S, owner of a small fishing dhow in Karachi. S is declared a person-of-interest regarding the terrorist plan. AN2, who has been monitoring the system for information related to the conference, makes a query for the current locations of persons-of-interest. He finds a SIGINT report of a call received by a cell phone owned by S. This phone is geo-located in the North Arabian Sea within 100NM of the USS Carney.

The request by USS Carney for information about potential illicit activities is matched, through SOA discovery, with the SIGINT report flagged by AN2 as associated with a person-of-interest. The PROGNOS system aboard the USS Carney invokes the service and receives a report indicating that S is connected to a participant in a suspected impending terrorist attack and is located within 100NM of the USS Carney. Upon receipt of the report, PROGNOS’ Knowledge Exchange Module translates the inference as a PR-OWL Finding that can be understood by the Reasoning Module. Since PROGNOS was performing a hypothesis management cycle (based on both its HMPO and MPO) involving the specific task of maritime target assessment, it brings into play a suitable PO stored in the Task-Specific PO Library, containing MFrags designed to capture knowledge relevant to Maritime Security missions. The report about S, now in the form of a MEBN finding in
PR-OWL format, can then be used as input to PROGNOS’ reasoning process. The results of reasoning are then used to update the status of the entity in PROGNOS’ Knowledge Storage Module representing the dhow owned by S.

In this example, it is fair to assume that additional inferences or data are also arriving at USS Carney’s PROGNOS in response to its request. These updates continuously feed its Reasoning Module to provide the most timely and accurate picture of the situation, including inferences about which targets among those hundreds are more likely to be of interest given the USS Carney’s mission. The result of this process is a comprehensive situational awareness view that includes PROGNOS’ prediction of the mission specific relevance for each target.

This problem requires reusable patterns of knowledge about events in space and time, how agents own and use objects, social interactions among agents, etc. For example, individuals are usually at the same location as their cell phones, may call each other to coordinate activities, and make plans with other agents in their social networks. In an operational system, these types of reasoning would make use of available ontologies. PR-OWL allows the user of such an ontology to add probabilistic information to represent uncertain relationships. In the example, it is not assumed that AN1 and AN2 are using PROGNOS. Indeed, they would be expected to deal with many different types of data coming from diverse systems. However, PROGNOS’ modular architecture allows for the incorporation of task-specific knowledge repositories such as a module designed to support interoperation with the social network analysis system used by AN2. In particular, if AN2 is using a social network ontology, PROGNOS could incorporate that ontology as one of its task-specific ontologies. Figure 6 illustrates a social network MFrag that could be used had we assumed AN1 and AN2 were using a MEBN/PR-OWL-based system such as PROGNOS. In the figure, each MFrag is a template for a fragment of a Bayesian network, and can be instantiated repeatedly.

5. System Integration Approach

PROGNOS integrates four state-of-the-art technologies into a distributed system architecture for predictive naval situation awareness. The heart of the system, the Reasoning Module, combines expressive probabilistic knowledge representation with efficient inference to enable reasoning about complex naval scenarios.

The reasoning system will be based on UnBBayes-MEBN, an open source, Java-based MEBN implementation that provides much of the necessary capability. UNBBayes-MEBN will be...
enhanced to include additional functionality such as hypothesis management and hybrid discrete-continuous inference. PROGNOS development will be guided through the development of increasingly demanding case studies as a means to ensure scalability and relevance to current operational needs.

PR-OWL Probabilistic Ontologies, also currently being implemented via a Java-based GUI, provide a convenient solution for both knowledge representation and exchange, both internal to PROGNOS and as a major driver ensuring its interoperability features as well. The format’s compliance with present web standards facilitates its use in Knowledge Exchange Module’s external interactions. Implementation within a SOA framework is the subject of ongoing research. Although the topic has been treated in research papers [7, 8, 9], the required compliance to FORCEnet standards suggests that issues such as data formats, communications protocols, and other interoperability issues should be taken into consideration during PROGNOS development as well.

PR-OWL is also the format of choice to implement the Knowledge Storage Module and the Knowledge Management Module, since its structure is already being developed in UnBBayes-MEBN. In both cases, ensuring a seamless interoperability from the ground up with data protocols being developed within the DoD and NATO’s environment is a major goal for this project. As an example, the implementation of the JC3IEDM protocol in the form of OWL ontologies has been subject to intense research (e.g. [29, 30, 31]) and PROGNOS would mostly benefit from these results.

Finally, the Simulation Module will also be implemented in Java, in a way to generate both random and pre-defined tracks within a time-controlled fashion.

6. Closing Remarks

Although a project still in its initial phases, PROGNOS presents some interesting approaches to applying collaborative technologies to network centric operations. It tackles a complex problem with distinct technologies in a parallel and synergic effort. The ideas behind its implementation have the potential of providing insights to researchers and practitioners in the field of predictive situation awareness.

References


