16th International Command and Control Research and Technology Symposium

"Collective C2 in Multinational Civil-Military Operations"

An Ontology for Hypothesis Management in the Maritime Domain

Student Paper

Topic 4: Information and Knowledge Exploitation

Richard Haberlin [Student]

Paulo Cesar G da Costa

Kathryn B. Laskey

George Mason University

Systems Engineering & Operations Research Department
4400 University Dr.
Fairfax, Virginia 22030-4444

(703) – 993 – 9989

[rhaberli;pcosta;klaskey]@gmu.edu

Point of Contact: Richard Haberlin

rhaberli@gmu.edu or (904) – 742 – 7624

An Ontology for Hypothesis Management in the Maritime Domain

ABSTRACT

Situational awareness supports tactical decision making through fusion of information about intelligence, geography, environment, and the geopolitical situation. Advanced decision support systems will provide the decision maker with a number of hypotheses from which the evolving situation may be inferred, limited only by the computational capacity of available computer hardware. Hypothesis Management is needed to control of exponential growth in fusion hypotheses created from incoming data reports delivered by individuals and units connected by a Semantic Services Registry. A Model-Based Systems Engineering Process was applied to design a series of algorithms for a Hypothesis Management Engine that explicitly manage the creation, modification, storage, and filtering of hypotheses. The scenario environment is modeled with the support of a Maritime Domain Ontology, which represents relationships between entities of interest. The effectiveness of the Hypothesis Management Engine is evaluated through simulation of a contextually accurate, randomly generated Hypothesis Knowledge Base which must be updated with incoming track data and queried for inferential reasoning candidates meeting the System Operator's request. This paper summarizes our research results and delineates the planned interaction of the Hypothesis Management Engine with an inferential reasoning system.

I. Introduction

One key requirement of decision support systems is to provide users with a unified view of the operational situation. This unified view is constructed by fusing inputs coming from different information sources. In addition to the updated picture, the concept of situation awareness also requires assessing what this depicted situation means to the system operator. This latter step is directly related to the *Orient* phase of the well-known OODA C2 cycle [12]. Obtaining this assessment involves exploring possible developments of the evolving scenario, including their impacts on the decision maker's goals. These possible developments are called *hypotheses*, each with a given probability to occur and an associated impact. Explicitly representing and reasoning about all can be loosely compared to predicting all the possible sequence of movements in a chess game but at a much larger scale, with each movement generating an exploding number of possible sequences. Hypothesis management algorithms are needed to avoid this unlimited growth while still representing sufficiently many hypotheses for viable decision-making. These algorithms are implemented in a Hypothesis Management Engine.

A Hypothesis Management Engine performs the essential functions of creating, updating, administrating, filtering, and routing hypotheses. It coordinates closely with a Hypothesis Knowledge Base for retrieval and storage of hypotheses, both working and archived. The end result is a set of contextually relevant hypotheses. These hypotheses are built from streaming data, filtered and pruned for computational efficiency, and delivered on demand in response to operator queries. For the maritime domain awareness situation assessment problem, a hypothesis can be thought of as a statement of anticipated action, or as a specifically defined plan of execution in which an actor will conduct an action against a target with a location, time and methodology of his choosing [7].

PROGNOS (PRobabilistic OntoloGies for Net-centric Operation Systems) is a proof of concept system being developed by George Mason University under contract to the Office of Naval Research. The goal of PROGNOS is "to provide consistent high-level fusion of data through knowledge representation and reasoning and enable predictive analysis with principled hypothesis management [13]." Within the system's Knowledge Storage Module is the Hypothesis Knowledge Base, which is used to store each hypothesis created from incoming data. Archived hypotheses are also maintained in the Hypothesis Knowledge Base. The Hypothesis Management Engine is a key component coordinating between the Knowledge Storage Module and Inferential Reasoner.

Hypothesis Collection Framework

Data from organic and non-organic information sources arrive in the Hypothesis Management Engine where they are continuously captured and stored in the hypothesis framework as a collection of 22 attributes representing features relevant to the current environment. This collection is referred to as a *hypothesis vector*. Additionally, every hypothesis vector has an associated weight vector which assigns a credibility value to each of the attribute categories represented by the fields of the hypothesis. This framework of hypothesis vector and associated weight vector will be instantiated as many times as necessary to convey each hypothesis nominated and stored in the Hypothesis Knowledge Base. Figure 1 illustrates a conceptual Hypothesis Knowledge Base framework in which each hypothesis is represented by a 22-attribute hypothesis vector and accompanying weight vector, discussed below.

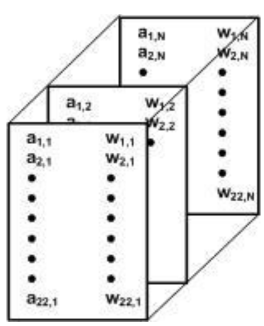

Figure 1 - Hypothesis Knowledge Base

The weight vector is derived from the Source Pedigree of the reporting unit. A Source Pedigree couplet represents the evidential weight assigned to data arriving from a particular source using a specific sensor. A detailed description of the Source Pedigree Ontology and methodology can be found in [4]. Together, the hypothesis vector and weight vector knowledge structures capture the content and strength of each hypothesis. The hypothesis vector describes a specific instantiation of a possible scenario, and the weight vector allows us to update its credibility with incoming data and compare it to others in response to a query.

Query Hypothesis

A query hypothesis is generated by the operator to begin the inferential reasoning process. For the Hypothesis Management Engine, the query hypothesis and its associated priority vector are used to search for and return candidate hypotheses. The random generation process for the query hypothesis in the simulation is discussed in Section IV, below.

The hypothesis framework described above is the structure used to capture and catalogue data available from organic and non-organic collection systems. To realize the decision support available through the inference algorithm, the operator generates a query hypothesis to answer a specific inquiry about the operational environment, using the same framework structure. The Hypothesis Management Engine is called upon to manage the creation, modification, administration, storage and movement of candidate hypotheses to ensure that only attributes and units relative to the current context are presented for inferential reasoning and to maintain computational viability. Associated with the query hypothesis is a priority vector, which allows the operator to prioritize attributes, and aids in the development of candidate hypotheses during the retrieval function described below.

An Example Scenario

A stylized scenario is included to assist in describing the Maritime Domain Ontology and Hypothesis Management Engine. Our example scenario is set in the Mediterranean Sea, Atlantic Ocean, and East Coast of North America. Agents of the terrorist organization Islamic Jihad Group, operating out of Izmir, Turkey, plan to smuggle radiological material into the United States on a bulk cargo vessel, where it will be used to build radiological dispersal devices. They intend to offload the material from the motor vessel *Mustafa Kamal* when it pulls into Baltimore, Maryland. We will refer to this scenario throughout this paper.

Overview of the Paper

Section II is an introduction of the Maritime Domain Ontology, its implementation, and its relationships. Next, Section III provides a brief description of the Hypothesis Management Engine and a detailed examination of the *Process Incoming Data* and *Retrieve Hypotheses* activities. This background is coalesced in Section IV, which describes a MATLAB simulation used to evaluate the algorithm effectiveness. Finally, Section V provides preliminary results and a description of the integration of the Maritime Domain Ontology and Hypothesis Management Engine into the ongoing ONR PROGNOS project. The Maritime Domain Ontology is captured in Protégé Version 4.1.0 (Build 213) [15], and the simulation is programmed in MATLAB Version 7.10.0.499.

II. The Maritime Domain Ontology

Daily, thousands of merchant ships transport millions of tons of material across the World's oceans. Multinational companies own these capital assets and coordinate their transit schedules between coastal nations to maximize time at sea. Because of their global exposure, and their multinational and transient crews, merchant ships are one feasible means to smuggle illicit goods and personnel between nations.

To capture the maritime domain and assist in the situational awareness problem, we have created an ontology of maritime entities that describes the platforms and states of maritime vessels transporting legal and illicit goods across the sea. In its initial form, the ports of departure and arrival are limited to those in the Mediterranean Sea and Atlantic Ocean. In the described scenario, cargo is shipped from departure ports in the Mediterranean to ports on the East Coast of North America.

An ontology defines a common vocabulary for describing entities and relationships within a specific domain for the purpose of sharing a common understanding of the structure of information [11]. One strength of ontologies is the ability to reuse the domain knowledge and structure for subsequent operations. For the maritime domain situation awareness problem, the Hypothesis Knowledge Base is constructed of class instantiations with defined attribute values and additional relationships. In our creation of this particular ontology, reuse of existing ontologies was considered and the following databases were evaluated for potential inclusion:

- Ontolingua ontology library [11]
- DARPA Agent Markup Language (DAML) ontology library [3]
- The United Nations Standard Products and Services Code (UNSPSC) [16]
- RosettaNet Global Supply Chain [14]
- DMOZ Open Directory Project [10]
- FreeBase Open Data Project [6]
- Linked Data Community [8]

Figure 2 illustrates the upper level of the Maritime Domain Ontology. Nodes in yellow represent super-classes and those in white are a selection of sub-classes with class relationships shown by black arcs. Not all sub-classes in the Maritime Domain Ontology are illustrated in the figure. Inter-object relationships are given by green arcs between nodes. For each, an inverse relationship exists that is not shown.

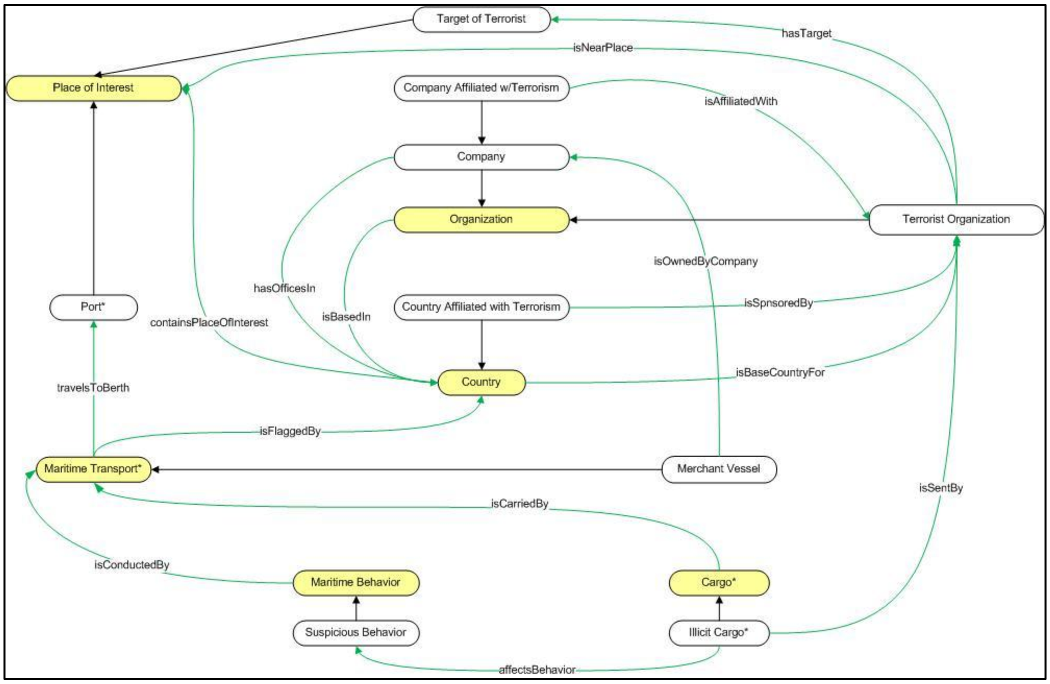

Figure 2 - Maritime Domain Ontology

While this particular version of the Maritime Domain Ontology was created to solve a terrorist and maritime smuggling problem in a restricted domain, the ontology is easily extensible for other bodies of water and modes of transportation. The following section describes the classes and relationships of the Maritime Domain Ontology in the context of the example introduced above.

Ontology Description

Figure 3 illustrates the top two levels of classes in the Maritime Domain Ontology. There are further levels of subclass below that shown in the figure, which provide greater specificity, particularly in cargo types, companies, and military bases. In fact, there are 74 classes specified in the Maritime Domain Ontology, 47 of which are types of legal and illicit cargo.

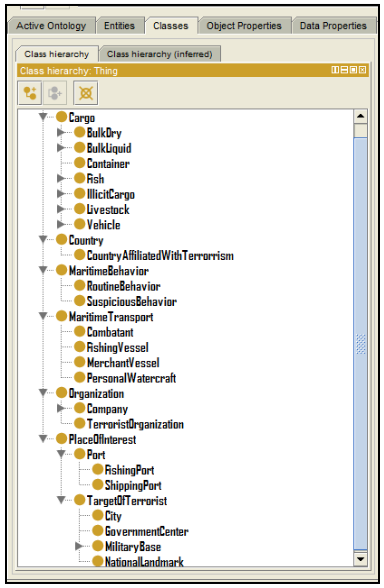

Figure 3 - Maritime Domain Ontology Classes

Descriptions of the six super-classes as they relate to the maritime domain problem are summarized below.

- **Cargo**. Each merchant ship is designed to carry a particular type of cargo, represented by the subclasses of Cargo found in Figure 3. Additionally, in the maritime domain awareness problem, a ship may transport illicit cargo in the form of weapons, components, personnel, drugs, or other contraband. Identifying the declared cargo of each particular vessel will aid in filtering merchant ships when the Hypothesis Knowledge Base is queried by the Hypothesis Management Engine.
- **Country**. All activity not occurring at sea happens in a country. Specifically, countries contain places of interest, have corporate offices of shipping companies, flag maritime transports and may serve as base countries for terrorist organizations.
- **Maritime Behavior**. Professional mariners perform tasks in a manner that maximizes profit and minimizes risk. Routine behavior includes traveling via a great-circle route between ports at the most economical speed for the ship class, avoiding close proximity to land, maintaining a safe distance from other ships, and keeping the ship in good material condition. Deviations from these normality conditions are considered Suspicious Behavior and are likely to be noticed by other mariners.
- **Maritime Transport**. Several types of maritime vessels are included in the Maritime Transport class. With counter-terrorism and counter-smuggling as the underlying themes, merchant vessels, fishing boats, personal watercraft, and the combatants that must interact with them are included.
- **Organization**. Organizations define the social structure of individuals working collectively to accomplish some task. In the Maritime Domain Ontology, they are

located in countries, may consist of or be affiliated with terrorists, and may own/operate maritime transports.

- **Place of Interest**. Places of Interest represent departure and destination points for Maritime Transports and Targets for Terrorist Organizations. Each is located in a country and may have nearby terrorist organizations, increasing the likelihood that this location is used or targeted by the terrorists.

Recall the example scenario in which a merchant ship departing from the Turkish port of Izmir is destined for Baltimore carrying radiological material. The Maritime Domain Ontology super-classes are affected in the following manner:

- **Cargo**: The merchant vessel *Mustafa Kamal* is carrying illicit cargo (CBR Component). Our scenario does not specify what type of merchant vessel this ship is (container, bulk, liquid, etc.), but this information is available for any professionally operated international carrier. It is safe to assume that in addition to the CBR Component, there will be a declared, legitimate cargo bound for Baltimore.
- **Country**: The ship departs from the port city of Izmir, Turkey, and travels to its destination, Baltimore, located in the United States. Islamic Jihad Group is based in Egypt and sponsored by Iran.
- **Maritime Behavior**: During the transit to Baltimore, it is unlikely that the ship will display any suspicious behavior unless the crew is involved in the plot and attempting to avoid authorities. This is not specified in the example description.
- **Maritime Transport**: *Mustafa Kamal* is a commercial merchant vessel. The flag, type, and owner are not specified in the description, but are readily available in a ship registry database. This would be of particular interest if one or more of these attributes correlated with the Islamic Jihad Group.
- **Organization**: Islamic Jihad Group is the terrorist organization believed to be behind the plot to ship CBR components to the United States. From the Maritime Domain Ontology, we know that Islamic Jihad Group is based in Egypt and sponsored by Iran. In this case it can be inferred that they have a cell operating in Izmir, Turkey.
- **Place of Interest**: The departure and destination ports, Izmir and Baltimore, are two obvious places of interest. What is unclear at this time is whether Baltimore is the target as well. As more information is gathered, this may become clearer. A separate hypothesis will be nominated for each reported target, as discussed in Section III.

It is not uncommon that some attributes are not assigned a value. This represents the uncertainty involved in collecting partial information about a particular ship operating within a large group of ships, like the international registry.

Implementing the Ontology

The Maritime Domain Ontology is captured in Protégé Version 4.1.0 (Build 213). The primary use of the Maritime Domain Ontology will be to provide the domain-specific ontology for the PROGNOS inferential reasoning system. By specifying the domain using real-world data, PROGNOS will produce a more realistic output. Specific metrics of represented information in the current build of the Maritime Domain Ontology are found in Table 1.

Table 1 - Maritime Domain Ontology Metrics

Unit	Number	Source
Cargo Classes	47	Subject-matter Expert Data
Country Individuals	78	CIA World Factbook 2010 [2]
Maritime Behavior Properties	41	Subject-matter Expert Data
Maritime Transport Classes	4	Subject-matter Expert Data
Terrorist Org. Individuals	46	U.S. Dept. of State List [5]
Company Individuals	121	Directory of Top Int'l Maritime Shipping Lines & Ship Owners [9]
Port Individuals	73	CIA World Factbook [2]
City Individuals	37	Subject-matter Expert Data
Target Individuals	51	Subject-matter Expert Data

Future builds will expand on this initial baseline to include smaller fishing (only) ports and a more rigorous evaluation of relationships between terrorist organizations and companies.

Relationship Implications within the Maritime Domain Ontology

The Maritime Domain Ontology contains many relationships between classes. Some of these imply increased likelihood that a Maritime Transport individual is associated with terrorism, or is a ship of interest in the maritime domain awareness problem. Some examples that would indicate a suspicious relationship are given below.

- Company Affiliated with Terrorism owning a Maritime Transport
- Country Sponsoring Terrorist Organization flagging a Maritime Transport
- Target of Terrorist is near Terrorist Organization
- Port is near Terrorist Organization
- Company has offices in Country Sponsoring Terrorist Organization

Each of these examples implies increased probability that the Maritime Transport, Company, Target, or Port will be used in the terrorist plot.

III. The Hypothesis Management Engine

The Hypothesis Management Engine performs the essential functions of creating, updating, administrating, filtering and routing hypotheses as sub-activities within the major processes of *Archive Hypotheses, Process Incoming Data*, and *Retrieve Hypotheses*. It coordinates closely with the Hypothesis Knowledge Base for retrieval and storage of hypotheses, both working and archived. The end result is a set of contextually relevant hypotheses built from streaming data that are filtered for computational efficiency and delivered to a model workspace for inferential reasoning [1].

Archive Hypotheses

Units often depart operating areas due to a change of mission only to find themselves back in the same area at a later date. Relational data between entities is not likely to change in the short term and should be maintained to expedite unit situational awareness upon return. The *Archive Hypothesis* activity allows the non-time sensitive attributes of hypotheses to be archived in the Hypothesis Knowledge Base in anticipation of building upon them at a later time. It systematically evaluates each hypothesis stored in the Hypothesis Knowledge Base and removes from each all of the attribute fields associated with spatial and temporal data. This activity results in a database of hypotheses consisting of useful long-term information about relationships between entities and devoid of any spatio-temporal data. Should the unit return to the same operational setting, these hypotheses are available to the Hypothesis Management Engine to build upon with new incoming data. The *Archive Hypothesis* activity is not discussed further in this paper, but the interested reader can find a detailed description of the process and accompanying activity diagram in [1].

Process Incoming Data

The Hypothesis Management Engine continuously creates and updates hypotheses from incoming data. The 22 attributes associated with each hypothesis are binned into four categories, described below and illustrated Figure 4.

- **Identity-based**: The identity attributes of SCONUM, captain, company, departure, and destination are the prioritized means by which maritime transports are identified and updated. Each of the columns in Figure 4 represents a sub-activity performed when activated by the related attribute field in the incoming data frame.
- **Temporal**: Temporal data attributes are those that change continuously with time; e.g. course, speed and position.
- **Behavioral**: Behavioral data attributes are those that identify a professionally run maritime transport from one that may have been negatively influenced by some external organization. The eight behavioral attributes identified in the hypothesis framework are binary fields representing behavioral information about the unit reported by an external observer.
- **Context**: The context attributes set identifies changes to the context of a scenario by changing some element of the "story" behind the hypothesis. Included in this set are

the attributes cargo, target, illicit cargo, and terrorist organization. For example, if a report arrived indicating that *Mustafa Kamal* was bound for New York instead of Baltimore, it is unclear which hypothesized destination is correct, Baltimore or New York. Therefore, duplicate hypotheses are created, differing only in the destination. Because this is performed on each match of the five identity attributes, the Hypothesis Knowledge Base grows exponentially as alternate hypotheses are created.

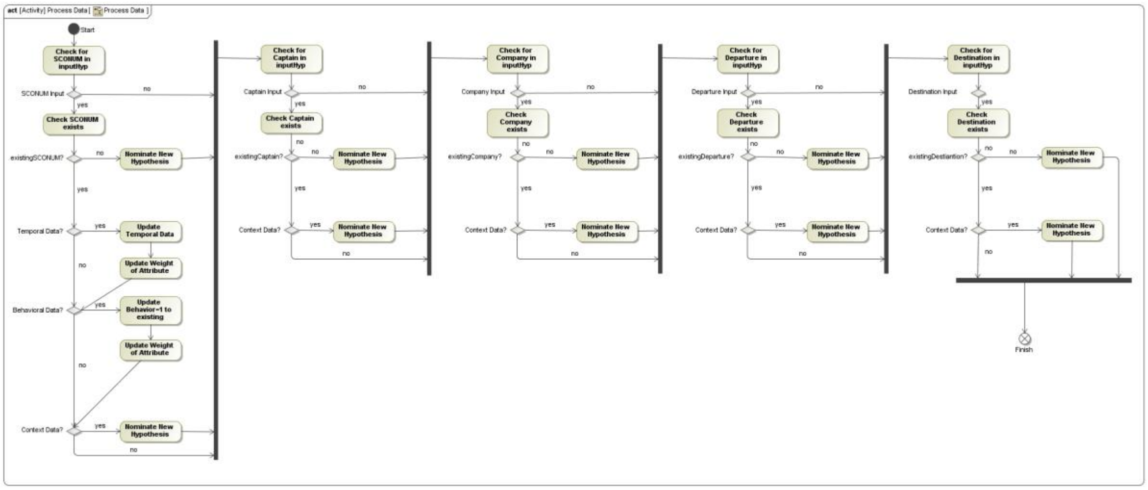

Figure 4 - Process Incoming Data

Evaluations of the identity-based attributes are conducted sequentially to determine if additional potential relationships need to be identified. The following four subsections correspond to the sub-activities of columns within Figure 4. Columns four and five are discussed concurrently for clarity.

SCONUM Evaluation

Most commercial merchants vessels are registered internationally using a Ship's COntrol NUMber (SCONUM). For this reason, incoming data that includes a SCONUM will first be checked against Hypothesis Knowledge Base hypotheses for a matching SCONUM. If no matching SCONUM exists, a new hypothesis is nominated and added to the Hypothesis Knowledge Base. If a match is found, temporal and behavioral data is updated. Finally, if context data is involved, a new hypothesis is nominated for each combination of existing hypothesis and a new context attribute. It is this final function that causes a drastic increase in hypothesis size.

Captain Evaluation

Every maritime transport has a captain, represented by an identification number in the simulation. If that individual is known, a check is run on the Hypothesis Knowledge Base for matching captain identifications. If there is no match, a new hypothesis is nominated. If one or more matches are found, contextual updates are performed. New hypotheses are nominated for contextual combinations, as described above.

Company Evaluation

Commercial merchant carriers and many fishing vessels are owned by large, multinational companies. A process similar to that of the captain evaluation is conducted for companies matching entries existing in the Hypothesis Knowledge Base. Again, unmatched companies trigger a new hypothesis nomination, and existing hypotheses are modified using context data to create multiple combination hypotheses.

Departure/Destination Port Evaluations

Finally, information may be known about the departure and destination ports for maritime transports. Typically, commercial vessels will declare their port of arrival prior to departing their current location. If departure or destination data exists, a search for matching ports is performed on the Hypothesis Knowledge Base hypotheses. Non-matches are nominated as new hypotheses. Matches are updated with context data to create new hypotheses representing other possible combinations of attributes.

Retrieve Hypotheses

In response to an operator query, candidate hypotheses are required for inferential reasoning. The *Retrieve Hypothesis* activity of the Hypothesis Management Engine coordinates with the Hypothesis Knowledge Base for retrieval, filters and prunes the hypotheses within the context of the query, and forwards the filtered hypotheses for inferential reasoning as illustrated in Figure 5, below.

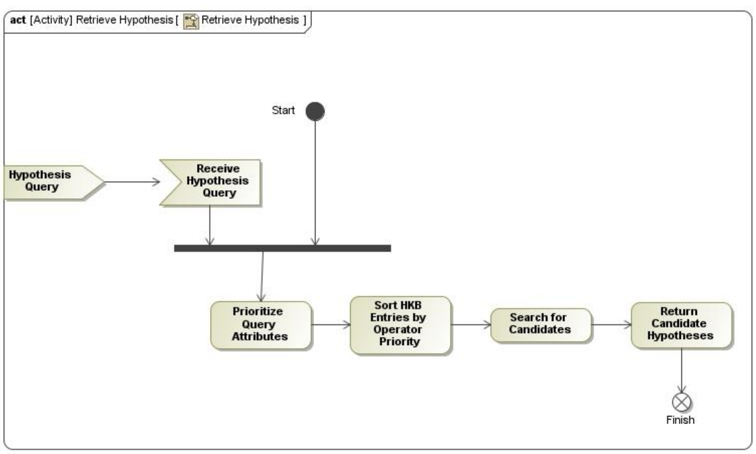

Figure 5 - Retrieve Hypotheses Activity

Query hypothesis data includes the attributes that represent positive or negative information about the query and the entity of interest. Additional query data in the form of a priority vector is used to retrieve the appropriate hypotheses, if they exist, from the Hypothesis Knowledge Base. The activity uses query hypothesis data from the request to iteratively search for and retrieve one or more applicable hypotheses from the Hypothesis Knowledge Base, shown by the code in Figure 6. Returned candidate hypotheses are prioritized by comparison with the priority vector provided by the operator in the query.

```
Sort Query Hypothesis attributes by Priority Vector value
For each priority value above a threshold {
       For each non-zero attribute {
              Check each HKB entry for matching value
              If values match, return hypothesis as candidate}}
```

Figure 6 - Retrieve Hypotheses Pseudo-code

This sub-activity returns one or more working hypotheses. Filtering is accomplished by truncating the search at a different level of the priority vector. Perhaps only the top-three attributes are of importance. In this case, only three iterations through the Hypothesis Knowledge Base are conducted, reducing run-time and candidates returned. The output is a set of filtered hypotheses, which are returned and used by the inference engine. This discrete series of actions is performed at the initiation of each new query.

IV. Simulation

The simulation provides an opportunity to observe the Hypothesis Management Engine in a synthetic environment and gain insight about the computational power that will be required to run the algorithm in a realistic environment. It was programmed using MATLAB version 7.10.0.499, and run on a Dell Inspiron 1570 with 4.0 G RAM and a 1.3 GHz processor. Because of the complexity of the *Process Incoming Data* activity, further simulation would require additional computational resources.

Simulation Methodology

The simulation suggests a scenario in which the Hypothesis Management Engine has been running for a short period of time and gathered a small number of hypotheses (100) in the Hypothesis Knowledge Base. It then receives a varying number of inputs that must be processed for inclusion and for possible relationships. Finally, the Hypothesis Knowledge Base is searched for candidate matches to a randomly generated hypothesis query prioritized by priority weight. The simulation methodology is shown in Figure 7.

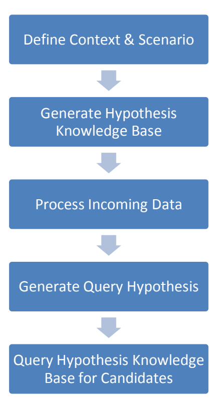

Figure 7 - Simulation Methodology

The context for the simulation is the maritime domain situational awareness problem described in the example scenario. All classes, properties, and individuals available in the Maritime

Domain Ontology were available as randomly-generated inputs for the Hypothesis Knowledge Base, incoming data to be processed, and query hypothesis.

Generate Random Hypothesis Knowledge Base

Using classes, properties, and individuals from the Maritime Domain Ontology, we randomly created a contextually accurate Hypothesis Knowledge Base consisting of 100 entries. Every entry has an opportunity for each of its 22 attribute fields to be included with a 30% probability.

Several of the assumptions for the random Hypothesis Knowledge Base have a strong effect on its density. First, the size of the track pool determines the frequency with which the unit identification numbers are repeated. In this case, a track pool of 10,000 entries is allowed, making duplication unlikely. Similarly, each unit captain has an identification number. The assumed pool of 1000 captains makes duplication of a captain significantly more likely when updating incoming information. The probability of inclusion of 30% discussed above determines if an attribute is to be included in a hypothesis. A greater inclusion would necessarily create a knowledge base of more-dense hypotheses.

Process Incoming Data

The Process Incoming Data activity described in Figure 4 is executed for a varying number of additional data points to identify the growth pattern of the Hypothesis Knowledge Base and simulation run-time. Initial results for input values of between 50 and 300 are summarized in Section V.

The *Process Incoming Data* activity is the most computational stressing and is calculated to be of order $O(N^3)$, where N is the size of the Hypothesis Knowledge Base. Computationally this is the limiting factor, as it is anticipated that a unit will receive hundreds of input data points, each requiring multiple passes through the Hypothesis Knowledge Base.

Generate Random Query Hypothesis

Much like the Hypothesis Knowledge Base, the randomly-generated query hypothesis has a probability that each attribute will be included. For this simulation, a value of 0.7 is used, indicating a high likelihood that each attribute is of interest. For each run of the simulation, a single query hypothesis is created and then executed against the randomly created, 100-entry Hypothesis Knowledge Base and its additional inputs using the *Retrieve Hypotheses* algorithm outlined in Figure 6.

V. Summary of Results

Crewmembers of merchant vessels are regularly multinational and transient. This is one possible way that terrorists or terror organizations can smuggle personnel or material into target countries. Natural log of preliminary run-time and final Hypothesis Knowledge Base size are shown in Figure 8 for a simulation run with P{Inclusion} = 0.30 and a starting Hypothesis Knowledge Base size of 100 entries. The inset shows the direct output data from the simulation.

The linear profile of the Hypothesis Knowledge Base size denotes an exponential growth rate, which is calculated as $O(N^2)$ from the encoded algorithm. This is primarily caused by the creation of context relationships evolving from multiple hypothesis possibilities.

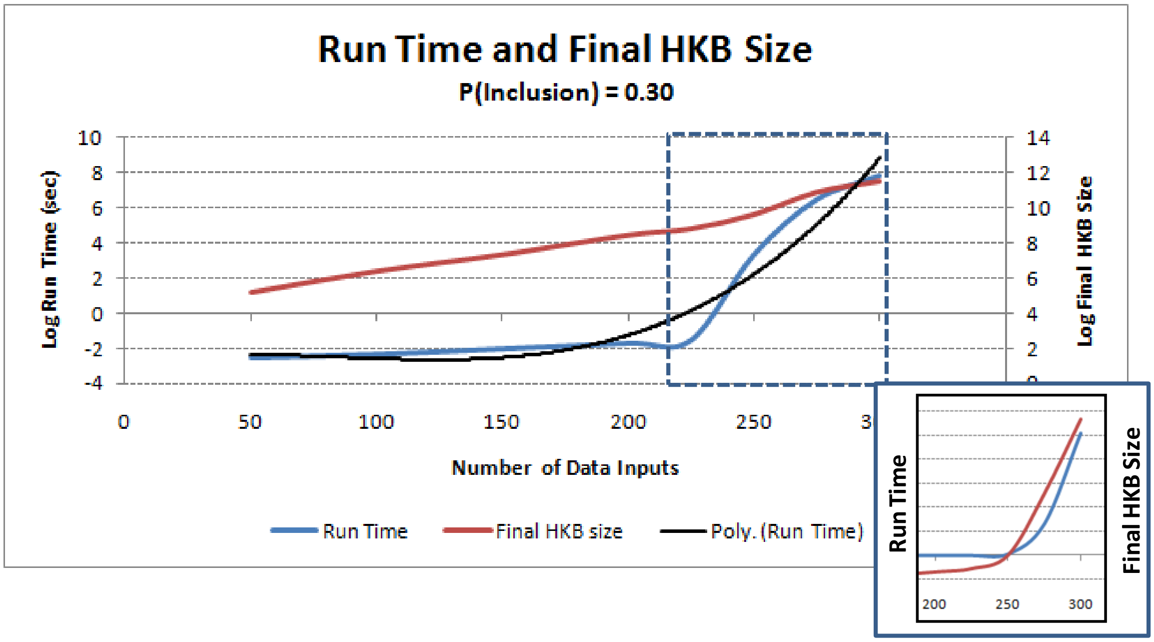

Figure 8 - Run Time and Final HKB Size

Figure 8 also shows an exponential plot for the natural log of run-time. This indicates a higher-level polynomial relationship for processor time, represented by the black trend line. As this is also a function of the processor, we will not comment on it further in this paper.

It is clear that context-variable relationships drive the computational complexity of the Hypothesis Management Engine. Reductions in overhead can only be realistically achieved by reducing the number of context variables for the *Process Incoming Data* activity based on the current environment and intelligence. For example, in a scenario with a known merchant ship name, attributes using the ship name (SCONUM) are of less value. Variable attribute values from the least-known context attribute would be the best to vary.

Follow-on work will add the automated step of prioritizing the Process Incoming Data activity steps based on the priority vector, and the attribute values within the query hypothesis. Context variables in fields of the query hypothesis that are identified will not be processed for additional relationships. This may reduce the complexity considerably.

Interaction of HME with Inferential Reasoning Engine

Recall the primary purpose of the Hypothesis Management Engine is to promote nomination of new hypotheses from incoming data and to retrieve appropriate candidates for inferential reasoning in response to an operator query. The Hypothesis Management Engine algorithms introduced and evaluated in this paper perform these functions using an extensive Maritime Domain Ontology. Further refinement of the data processing methodology will increase computational efficiency and reduce run-time.

VI. References

[1] R. Haberlin, P. C. G. da Costa, and K. B. Laskey, "A Model-Based Systems Engineering Approach to Hypothesis Management," in Proceedings of the 3rd International Conference on Model Based Systems Engineering, 2010.

[2] "CIA - The World Factbook," Central Intelligence Agency - The World Factbook, Jan-2011. [Online]. Available: https://www.cia.gov/library/publications/the-world-factbook/. [Accessed: 18-Jan-2011].

[3] "DAML Ontology Library," DAML Ontology Library, Jan-2011. [Online]. Available: http://www.daml.org/ontologies/. [Accessed: 18-Jan-2011].

[4] R. Haberlin, "Establishing a Source Pedigree Ontology Through Evidential Reasoning," Dec-2010.

[5] "Foreign Terrorist Organizations," US Department of State Foreign Terrorist Organizations, Jan-2011. [Online]. Available: http://www.state.gov/s/ct/rls/other/des/123085.htm. [Accessed: 18-Jan-2011].

[6] "Freebase," Freebase, Jan-2011. [Online]. Available: http://www.freebase.com/. [Accessed: 18-Jan-2011].

[7] R. Haberlin, P. C. G. da Costa, and K. B. Laskey, "Hypothesis Management In Support of Inferential Reasoning," in Proceedings of the 15th International Command and Control Research Symposium (ICCRTS), 2010.

[8] "Linked Data | Linked Data - Connect Distributed Data across the Web," Linked Data, Jan-2011. [Online]. Available: http://linkeddata.org/. [Accessed: 18-Jan-2011].

[9] "Maritime Shipping Companies : Ship Owners," Directory of Top Int'l Maritime Shipping Lines & Ship Owners, Jan-2011. [Online]. Available: http://www.onelasvegas.com/wireless/ship_owners.html. [Accessed: 18-Jan-2011].

[10] "ODP - Open Directory Project," DMOZ Open Directory Project, Jan-2011. [Online]. Available: http://www.dmoz.org/. [Accessed: 18-Jan-2011].

[11] "Ontolingua Home Page," Ontolingua, Jan-2011. [Online]. Available: http://www.ksl.stanford.edu/software/ontolingua/. [Accessed: 18-Jan-2011].

[12] J. Boyd, "Organic Design for Command and Control" [Electronic Version]. Unpublished lecture notes. Retrieved January, 2011, from: http://www.ausairpower.net/JRB/c&c.pdf.

[13] P. C. G. da Costa, K. B. Laskey, and K. Chan, "PROGNOS: Applying Probabilistic Ontologies to Distributed Predictive Situation Assessment in Naval Operations," in Command and Control Research and Technology.

[14] "RosettaNet Home," RosettaNet, Jan-2011. [Online]. Available: http://rosettanet.org/. [Accessed: 18-Jan-2011].

[15] "The Protégé Ontology Editor and Knowledge Acquisition System," Protege, Jan-2011. [Online]. Available: http://protege.stanford.edu/. [Accessed: 19-Jan-2011].

[16] "UNSPSC Homepage," United Nations Standard Products and Services Code, Jan-2011. [Online]. Available: http://unspsc.org/. [Accessed: 18-Jan-2011].

Richard Haberlin is a retired US Naval Flight Officer with extensive experience in anti-submarine warfare and airborne intelligence, surveillance and reconnaissance operations in the Arctic, Atlantic, Mediterranean and Middle East. He is currently a senior analyst for Enterprise Management Solutions Inc. in Alexandria, Virginia. An active member of the Institute for Operations Research and the Management Sciences and the Military Operations Research Society, he holds an M.S. in Operations Research from the Naval Postgraduate School and a B.S. in Ocean Engineering from the United States Naval Academy.

Paulo Cesar G. da Costa is Affiliate Professor of Systems Engineering and Operations Research and Research Assistant Professor of the Center of Excellence in Command, Control, Communications, Computing and Intelligence at George Mason University. Dr. Costa is a retired Brazilian Air Force Flight Officer with extensive experience in electronic warfare, C4I, operations research and military decision support. He teaches courses in decision theory and systems engineering, and has developed PR-OWL, a probabilistic extension of the OWL ontology language. As an invited professor at University of Brasilia, he was a key contributor to the development of UnBBayes-MEBN, an implementation of the MEBN probabilistic first-order logic.

Kathryn Blackmond Laskey is Associate Professor of Systems Engineering and Operations Research and Associate Director of the Center of Excellence in Command, Control, Communications, Computing and Intelligence at George Mason University. Dr. Laskey teaches courses in systems engineering, computational decision theory, and decision support. She has published extensively in the areas of inference, knowledge representation, learning, and information fusion. She developed Multi-Entity Bayesian Networks (MEBN), a language and logic that extends first-order logic to support probability. She was a key contributor to the development of PR-OWL, an upper ontology that allows MEBN theories to be represented in OWL ontologies.