15TH INTERNATIONAL COMMAND AND CONTROL RESEARCH AND TECHNOLOGY SYMPOSIUM
"THE EVOLUTION OF C2: WHERE HAVE WE BEEN? WHERE ARE WE GOING?"

PROBABILISTIC ONTOLOGIES AND PRAGMATICS FOR COMPLEX SYSTEMS INTEGRATION

Paulo Cesar G da Costa
Bernard Zeigler
Michael Hieb
Kathryn B. Laskey

George Mason University
System Engineering & Operations Research Department
4400 University Drive
Fairfax, VA 22030-4444
(703) 993-9989
[pcosta; bzeigle1, mhieb, klaskey]@gmu.edu

Point of Contact: Paulo Cesar G. da Costa
pcosta@gmu.edu and/or (703) 993-9989

Abstract

Information systems are increasingly pervasive in our lives, and are becoming increasingly sophisticated as technology evolves. This has resulted in an ever-growing complexity that makes interoperability a major issue. We argue that addressing the interoperability problem requires consideration of at least three distinct aspects of data: syntax, semantics, and pragmatics. The syntactical dimension, which provides the link between the data and its representation, is already well understood. The semantic dimension provides the link between the data and its meaning. Ontology engineering addresses this aspect. Technologies and tools for ontology engineering are under active development. Current-generation semantic technology lacks a principled way of representing uncertainty, a requirement for many domains of interest. The pragmatic dimension provides the link between the data and its use. Pragmatic frames, in close relationship with ontologies, address this dimension both in its static and dynamic aspects.

This paper presents our research on pragmatics and probabilistic ontologies, which are integrated in a coherent, mathematically sound computational environment via the System Entity Structure (SES) architecture. The key objective is to present a principled theoretical framework for seamless integration through a scalable high-level knowledge fusion methodology.

I Motivation: The Knowledge Integration Challenge

Networked systems are becoming ubiquitous across a wide array of human activities, growing ever more complex with each passing year. Complexity increases not as an end in itself, but as a side effect of success. New capabilities are implemented, new technologies are added, scope is broadened, specialization increases, larger problems are tackled -- and complexity grows. Because of complexity creep, success can be a mixed blessing. An intricate but capable system is a blessing in its ability to support complex user operations, but it can become a curse when the need arises to interoperate with other equally complex systems.

As an example, disaster relief operations engage players such as governmental agencies, private companies, NGOs, military, law enforcement volunteers, and many others whose numbers and diversity make command and control (C^2) a great challenge. In this environment, a key issue is to provide information technology support for rapid and accurate decision-making, which involves finding, integrating, and continuously converting vast amounts of heterogeneous data into actionable information. Effective use of these resources can lead to more informed decisions and contribute to the success of the overall operation. However, this can only be achieved by addressing the challenges of: (1) enabling interoperability among diverse systems and data repositories; (2) incorporating a wide variety of traditional and non-traditional types of data coming from geographically dispersed sources; and (3) providing the ability to process massive volumes of noisy, incomplete and uncertain data in a timely manner.

Advances in connectivity and computation alone are insufficient to meet these challenges. The sheer volume of data creates informational and cognitive bottlenecks. Incompatible formats and semantic mismatches demand 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' precious cognitive resources are too often focused on low-level manual data integration rather than high-level reasoning about the situations to be addressed. New approaches are needed to bridge the gap from data interchange to knowledge interchange, to free human decision-makers from information overload and low-level manual tasks, and to provide them with actionable, decision-relevant information.

Our approach is to achieve true high-level knowledge interchange by leveraging the unexplored synergistic potential of current state-of-the-art approaches to interoperable systems. More specifically, the main thrust of the research is to merge probabilistic ontologies with pragmatic frames so to establish a consistent framework that can be applied to many domains of application. Synthesizing these leading edge technologies is by itself a major challenge with a great potential to advance science. A second, equally ambitious goal is to move towards a general theory of high-level knowledge integration.

Research on the subject of information fusion has focused primarily on specific application areas. The bulk of research effort has concentrated on lower-level data alignment (e.g. multi-sensor data fusion, syntactic protocols, distributed simulation, etc), on semantic mapping solutions (e.g. Semantic Web approaches, specialized semantic mapping solutions, etc), or other topics that do not fully address the fundamentals of high-level knowledge integration. This gap has been recognized and there have been some efforts to address it (e.g. [1]). However, research on this subject is still in its infancy.

In other words, although some of the above issues have been addressed in an isolated fashion, more has to be done towards a solution for the problem of automated high-level knowledge fusion. This is a gap that has been identified in applications of complex networked systems, but has not been addressed in its fundamental aspects with the academic rigor necessary to produce a full-fledged theory and scientific discipline. Our aim is not only to bring about important theoretical and practical advances through our research, but in addition, to foster the development of a rigorous scientific discipline of complex systems integration.

Following this initial motivation on the problem of Knowledge Integration for complex endeavors, section 2 provides an overview of the current research regarding interoperability. In section 3 we present the main concepts, methods, and technologies behind our approach, emphasizing the relevant issues in merging probabilistic ontologies with pragmatic frames. Then, section 4 illustrates how our approach would be applied to a complex endeavor, by means of a case study inspired by the disaster relief operations following hurricane Katrina. Section 5 provides a brief overview of our general approach for integrating the complex suite of technologies we are dealing with. Section 6 is a short discussion of our research.

II. BACKGROUND AND RELATED RESEARCH

II.1 Theories of Interoperability

There are some proposed theories that focus on assessing the interoperability level of distinct systems. An example is the two-level theory of interoperability by Dahmann et al. [2] or Tolk's Levels of Conceptual Interoperability Model (LCIM) [3]. These theories build up from key research ideas that shaped the field of data alignment. The first of these ideas came from the work by Zeigler in 1976 [4] on the Discrete Event System Specification (DEVS), which is a mathematical formalism for specifying and composing components into systems. DEVS is used to describe components across a spectrum that ranges from mathematical expressions and mathematical approximations to discrete approximations and discrete interpolation, with the discrete aspects of the DEVS spectrum executable on a digital computer. The descriptive range allows DEVS to cope with the specification of components across several levels.

Dahmanm [2] asserted that interoperability occurs in levels, while Tolk [3] and Turnitsa [5] further lend support to the idea of interoperability levels by specifying the LCIM classification. There, each increase in the level of interoperability gives rise to a corresponding increase in the density of information that is exchanged. For example, the shift from the syntactic to the semantic to the pragmatic level of interoperability brings a shift from "Which format?" to "What is going on?" to "How can/should I respond?" Each level presupposes and requires the levels below it. At the syntactic level, the systems are able to interchange data; at the semantic level, they attach the same meaning to the data being interchanged; the pragmatic level, additional context or information is exchanged to enable appropriate action to be taken. Tolk *et al.* [6] provides an in depth review into each of the LCIM complementary theories of model based data engineering, process engineering and constraint-assumption engineering.

Page *et al*. [7] address the issue from a process-centric perspective and propose a framework for the broader simulation interconnection problem based on the ideas of capability maturity models (CMM) [8, 9]. They argue that the problem should be seen from three distinct dimensions: composability, interoperability, and integratability. Composability between two models is achieved when they share compatible objectives and can be directly related with the LCIM's conceptual interoperability level. Along the same lines, interoperability is the ability of two models to exchange data or services at run-time, and is directly related to the LCIM's Dynamic, Pragmatic, Semantic and Syntactic levels. Finally, integratability is the ability of configuring and modifying a set of components to make them interoperable and possibly composable, and is directly related to the LCIM's technical level.

While other models of interoperability such as the Levels of Information Systems Interoperability Model (LISI) [10] or the NATO Reference Model of Interoperability (NMI) [11] exist, these models address only limited specification of higher levels of interoperability. The LCIM, by considering alignment and harmonization of data, provides the required theoretical framework to explore interoperability at a higher level.

Although our research is not specifically aimed at defining taxonomies for interoperability, the LCIM levels are incorporated into our framework for high-level information fusion using probabilistic ontologies (POs) [12]. As an example, our use of pragmatics and POs addresses all the LCIM levels (syntax, semantics, pragmatics, and dynamic).

II.2 Ontologies and Uncertainty

As people started to realize that syntax-based solutions were not enough to satisfy the increasing need for interoperable systems, they started to look for semantics as the silver bullet to satisfy their goals. As a result, ontology engineering became a major topic of interoperability research. Since its adoption in the field of Information Systems, the term ontology has been given many different definitions. A common underlying assumption is that the formal foundation for knowledge representation and reasoning would be classical logic [13]. However, classical logic provides no consistent support for plausible reasoning, which we see as a major requirement to semantic interoperability applications.

The *de facto* standard for developing ontologies is OWL, a W3C recommendation [14]. OWL has its roots in its web language predecessors (i.e. XML, RDF), and in traditional knowledge representation formalisms that have historically not considered uncertainty. Examples of these formalisms include Frame systems [15] and Description Logics, which evolved from the so-called "Structured Inheritance Networks" [16]. This historical background somewhat explains the lack of support for uncertainty in OWL, a serious limitation for applications in uncertainty-laden domains such as genetics or medicine. In fact, virtually all of the current ontology formalisms are based on classical logic, and SW languages such as OWL provide no consistent support for uncertainty representation or plausible reasoning.

This lack of support for uncertainty can perhaps be justified in closed-world systems designed to perform well-defined tasks, for which clear and unambiguous vocabularies can be constructed. But the Semantic Web vision as described in Benners-Lee *et al*. [17] requires heterogeneous systems to interoperate in an open world. Inevitably, vocabularies that are adequate for a single stand-alone application break down when required to interoperate with systems employing different vocabularies originally tailored to different tasks. Inevitably, there is incomplete and partial overlap of terminology and concepts. The same phenomenon occurs in virtually all complex endeavors in which two or more systems must interact within an environment similar to the one faced by SW applications.

Probabilistic ontologies enable representation of knowledge in domains characterized by uncertainty. As such, they improve the quality of service descriptions, enable more thorough analysis of service composition opportunities, and provide a theoretically sound methodology for semantic mapping under uncertainty.

III ENABLING HIGH-LEVEL KNOWLEDGE FUSION

As previously mentioned, there is no established scientific theory of, and no general-purpose, theory-based methodology for high-level information fusion. Therefore, our vision on how to enable high-level fusion relies on a multi-disciplinary approach. The component technologies underlying our approach are described below.

III.1 Multi-Entity Bayesian Networks

Multi-entity Bayesian Networks (MEBN) [18] is a probabilistic logic with first-order expressive power. MEBN was developed to meet the representational and computational challenges inherent in higher-level multi-source fusion and situation awareness. Specifically, MEBN can represent degrees of plausibility for any hypothesis that can be expressed in first-order logic. Its basis in directed graphical models gives it a natural representation for cause and effect relationships. Its built-in capability for context-specific independence provides a natural way to represent contextual factors important for hypothesis management, such as conditions under which a hypothesis can be pruned because it has little or no impact on conclusions of interest. MEBN also supports a natural representation for essential categories of uncertainty for general situation awareness, such as uncertainty about entity existence (*i.e.*, is a report a false alarm); uncertainty about the type of entity; and uncertainty about functional relationships (*e.g.*, which entity gave rise to a report). Its basis in Bayesian theory provides a natural theoretical framework for learning with experience. Its graphical representation supports an intuitive interface for specifying probabilistic ontologies. Finally, its modular representation formalism supports adaptability, by allowing changes to be made to parts of an ontology without affecting other parts or other ontologies, and composability, by allowing problem-specific models to be constructed "on the fly," drawing only from those resources needed for the specific problem.

MEBN represents the world as made up 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 (MFrags) organized into MEBN Theories (MTheories). An MFrag represents a conditional probability distribution of the instances of its resident random variables (RVs) given the values of instances of their parents in the fragment graphs and given the context constraints. RVs are graphically represented in an MFrag either as resident nodes, which have distributions defined in their home fragment, or as input nodes, which have distributions defined elsewhere. Context nodes are the third type of MFrag nodes, and represent conditions assumed for definition of the local distributions. Figure 1 depicts the generic structure of an MFrag.

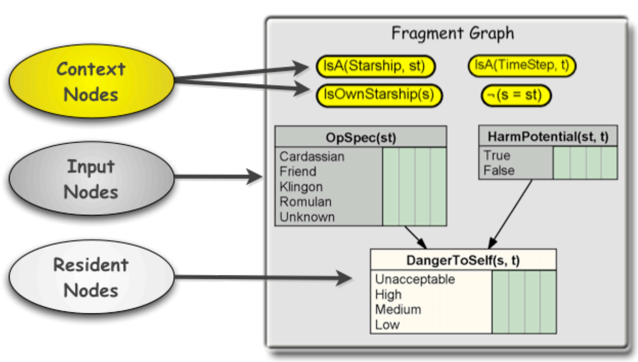

Figure 1 – The basic components of an MFrag

Typically, MFrags are small, because their main purpose is to model "small pieces" of domain knowledge that can be reused in any context that matches the context nodes. This is a very important feature of the logic for modeling complex situations — the knowledge representation version of the "divide and conquer" paradigm for decision-making. Decomposition is accomplished by modeling a complex situation as a collection of small MFrags, each representing some specific element of a more complex situation. The additional advantage of MEBN modeling is the ability to reuse these "small pieces" of knowledge, combining them in many different ways in different scenarios. A coherent collection of MFrags is called an MTheory. An MTheory represents a joint probability distribution for an unbounded, possibly infinite number of instances of its random variables. This joint distribution is specified implicitly through the local and default distributions within each MFrag, together with the conditional independence relationships implied by the fragment graphs. Figure 2 depicts the Starship MTheory, a toy example devised as part of the original work on PR-OWL probabilistic ontologies [12], which we describe in the next section.

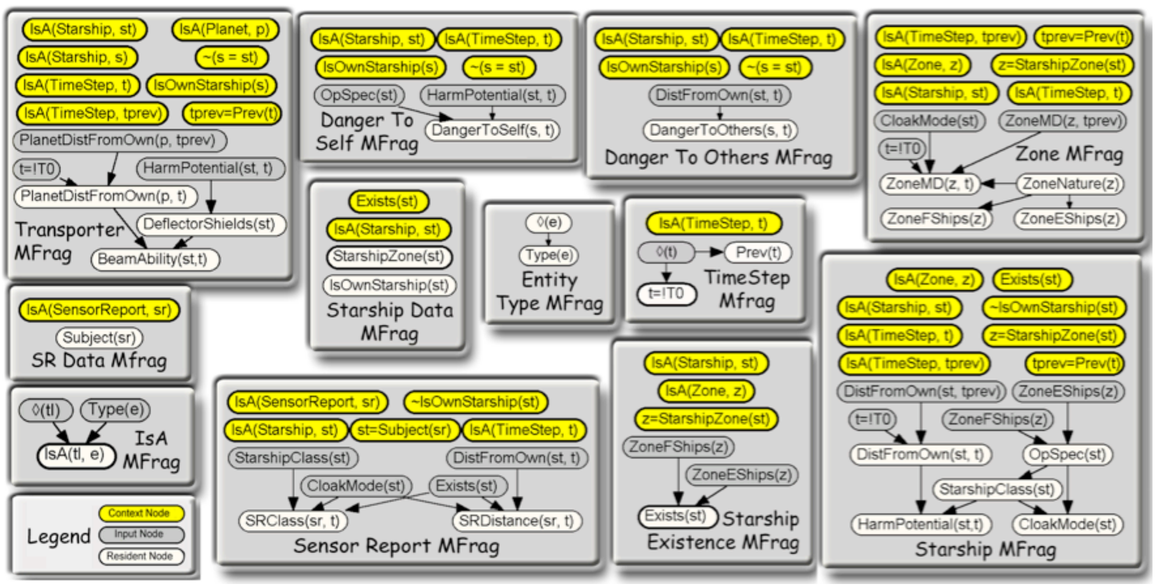

Figure 2 – The starship MTheory

Automated methods for reasoning about such complex situations require expressive representation languages. Many real-world operational scenarios in which high-level knowledge fusion is performed typically involve a high degree of uncertainty, and the available data are inevitably noisy and incomplete. It is essential to be able to represent and reason with uncertainty. A MEBN domain model implicitly represents a joint probability distribution over situations involving unbounded numbers of entities interacting in complex ways.

Recent years have seen rapid advances in the expressive power of probabilistic languages (e.g., [18-26]). They have the potential to trigger major advances in the ability to represent and reason about complex situations. In comparing MEBN with other expressive probabilistic languages [19-26], only BLOG [25], Markov logic [20-22], and MEBN have full first-order expressive power. BLOG is based on a text-style programming language for specifying complex probability models, whereas MEBN provides a natural graphical interface for specifying probabilistic relationships and distributions. In Markov logic, a probability distribution is defined by attaching numerical weights to first-order formulas. A probability distribution can be specified for an existing first-order knowledge base simply by attaching a weight to each formula in the knowledge base [20]. The result is an undirected graphical model in which nodes are ground atoms (hypotheses about specific domain entities) and arcs connect atoms that appear in groundings of the same formula in the knowledge base. MEBN's directed representation and support for context-specific independence allow it to represent cause and effect relationships, to support predictions about the effects of interventions [27], and to support computationally efficient hypothesis management methods. Furthermore, MEBN can represent both discrete and continuous distributions, an essential capability for fusion systems. Markov logic does not share these representational advantages. MEBN is the logic behind the probabilistic ontology language, PR-OWL.

III.2 PR-OWL Probabilistic Ontologies

Ontologies provide the "semantic glue" to enable knowledge sharing among distinct systems cooperating in data rich domains such as Predictive Analysis. 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 higher-level knowledge 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., [28-36]). Probabilistic ontologies 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 probabilistic ontology language [12, 28-32] is founded in MEBN logic and has the expressive power to represent any first-order Bayesian theory. PR-OWL provides the representation power required for information fusion and prediction services in net-centric environments. PR-OWL ontologies interoperate with the non-probabilistic part of ontologies written in the World Wide Web Consortium's standard ontology language OWL, thus facilitating interoperability with other semantically aware systems. In this proposed research, PR-OWL is used to design a distributed high-level fusion framework that performs approximate coherent Bayesian reasoning on problems of greater complexity than previously possible. UnBBayes-MEBN [37, 38] is an open source, java-based graphical editor for PR-OWL ontologies being developed in conjunction with the University of Brasilia. Figure 3 shows the current UnBBayes-MEBN graphical interface for developing MTheories displaying an MFrag of the above-mentioned Starship MTheory.

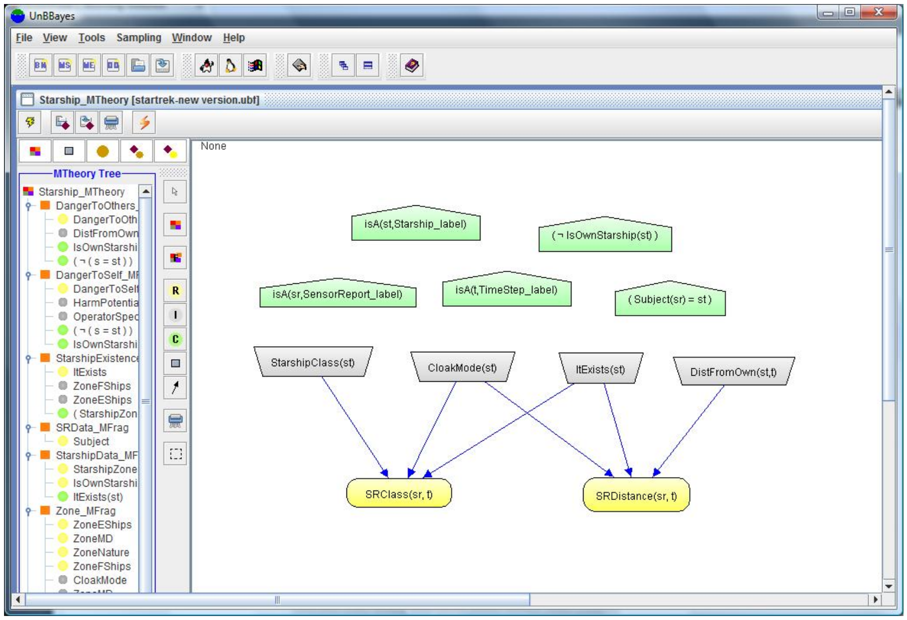

Figure 3 – UnBBayes-MEBN Graphical Editor

III.3 Spatio-Temporal Hypothesis Management

As noted above, recent work on combining probability with first-order logic has expanded the range of problems that can be tackled by automated fusion systems. However, for problems of the scale required for high-level information fusion, 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 [39]. Methods from the multi-target tracking literature can be generalized to search over the vast number of hypotheses [40]. For example, Markov Chain Monte Carlo Data Association (MCMCDA) is a novel algorithm to do recursive hypothesis formation and management [41]. It has a strong theoretical grounding as an approximation to the optimal Bayesian solution, and has been shown to work well in practice. We intend to develop a MCMC hypothesis management (MC2HM) in our approach as a means to ensure, among other things, a scalable solution.

III.5 Pragmatic Frames

Zeigler and Hammond define the idea behind pragmatics is that the consumer's use of the information should determine the description mechanism, or ontology, used by the producer. Then, the developer of the ontology, also called data engineer, has the task of tuning the ontology to the pragmatic frame [42, page 3].

Pragmatics is defined as the use of metadata in relation to metadata structure and context of application. In other words, pragmatics uses metadata to convey context and its relation to meaning, and pragmatic frames uses such context information to disambiguate meaning. The idea was based on Speech-Act theory [43-46] and focuses on elucidating the intent of the semantics constrained by a given context. For example, suppose that I say: "I see the plane." There is no context here to determine whether the word "plane" refers to a flat space defined by at least 3 points, an airplane, or a wood working tool, which are all valid semantic values for the word "plane." It does not make sense to examine, in detail, the low level semantics of attributes of the word "plane" when an examination of the use of the plane will obviate further examination of the details.

Pragmatic frames are a means to convey Pragmatics through an ontological framework. Basically, they are used to delineate a data engineer's domain of interest and relate the ontology as being adequate or not to this domain. That is, an ontology supports (or is applicable to) a pragmatic frame if the world states (or state changes) that it can describe include those that are needed by the frame. Further, an ontology is minimal for a frame if it supports only that frame, not a larger one, and two ontologies are pragmatically equivalent in a pragmatic frame if there is a one-to-one correspondence of their world state descriptions such that corresponding descriptions are used in the same manner within the pragmatic frame of interest.

Pragmatic equivalence is an important concept. Even though world state descriptions generated by the ontologies may differ, the manner in which they are processed downstream leads to the same results. For example, messages sent within one ontology might not differ from those of a second except in numerical precision. Consider corresponding number strings that are the same only up to a given number of significant digits. Pragmatic equivalence holds if both strings are treated equally by downstream processors. We say that this difference is absorbed within, or modulo, the pragmatic frame. Of course, another frame may treat these strings differently, leading to pragmatic inequivalence in this frame.

Finally, pragmatic frames can address both static and dynamic situations. Static pragmatic frames focus on the comparisons to determine the degree of similarity of two frames, subtrees, or trees. Dynamic pragmatics refers to the change in state of the pragmatic frame due to a continuously occurring change of context, a discrete-time context change, or a discrete event.

IV TYPICAL APPLICATION SCENARIO

To illustrate the potential synergy of merging the above concepts in a consistent framework, a scenario is presented in which four distinct systems need to co-operate within a disaster relief operation such as the aftermath of hurricane Katrina, the disaster relief operations after the Asian tsunami triggered by the Sumatra-Andaman earthquake in 2004, or the 2010 Haitian earthquake. In such situations, many failures and mistakes occur because of interoperability gaps in communications and C^2 structure. Such errors can lead to loss of material and lives. However, after the emergency peak and as the situation evolves, the communications and C^2 structures adapt to the circumstances and become more reliable. At this point, the relief teams and their major support apparatus start the long process of providing relief to the victims in the form of medical assistance, food distribution, and other subsistence enablers. The lack of a proper infrastructure, combined with the large volume of assets being coordinated, imposes a great stress to the IT systems supporting the endeavor.

In order to replicate this kind of environment, relevant aspects of the systems involved in the complex endeavor should be captured, and their respective impacts on their operations properly assessed. One of the first tasks of our research is to develop such model. To that end, a rigorous assessment of the conditions in the aftermath of a disaster relief is required. For the sake of brevity, we focus here on the problem of food distribution.

Each system has to deal with a vast amount of information and has a specific data structure that optimizes its operations. For example, if system A was designed to deal with the transportation of perishable items, it will need to represent information about perishability (e.g. expiration dates, transport temperatures, etc). Similarly, if system B's focus is on cataloging and distributing donation items (not only perishable food) it may use metrics comparable to A's, but with different intentions and assumptions over both the semantics (e.g. 25 degrees is meant to be the storage temperature for item XYZ) and pragmatics (e.g. level of precision required for a specific use for that temperature information). Dealing with these minor distinctions is usually expensive and mostly infeasible from a computational standpoint. Ultimately, high-level fusion must often be assigned to humans. Ontologies might help to overcome the semantic barrier, but there is still the issue of fitness of the information to the task at hand. Pragmatic frames address this issue.

However, the combination of ontologies and pragmatic frames still cannot overcome the former's lack of a standard representation of uncertainty, which is ubiquitous in virtually all real world issues in complex endeavors. In the above toy example, suppose System C, another system involved in the food distribution process, receives information about item XYZ but doesn't "understand" temperature in degrees Fahrenheit. The consequences of this information mismatch will depend on system C's behavior. One possible outcome is that the system computes XYZ's storage temperature as 25 degrees Celsius (77 degrees Fahrenheit), causing a container of XYZ to rot. A less catastrophic output would be that the system assumes XYZ's storage temperature information is missing, incomplete, or incompatible with its purposes, raising a flag for human intervention. However, human experts are a scarce resource in complex endeavors. Therefore, a probabilistic expert system might be used to replicate the behavior of a human expert. A probabilistic ontology with statistical regularities about storage temperatures would support the design of systems capable of performing plausible inferences on, for example, the average storage temperature for items such as XYZ. Obviously, many additional types of uncertainty can arise in complex environments. This reinforces the need for a representational scheme that is capable of both representing and reasoning in the presence of incomplete, ambiguous, or dissonant information. The combination of POs and pragmatic frames is depicted in Figure 4.

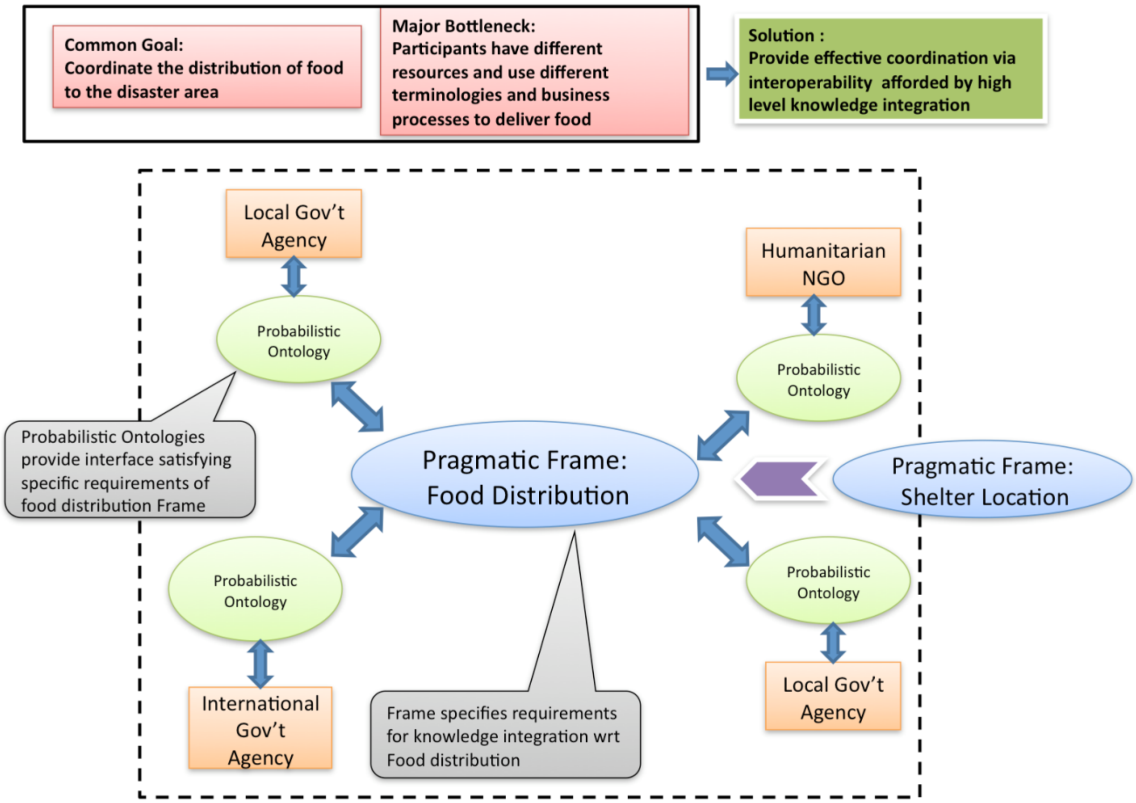

Figure 4 – Typical complex endeavor scenario and proposed knowledge integration approach

As it can be seen in the figure, each participant system will be aligned a probabilistic ontology while sharing a common pragmatic frame conveying the pragmatics of food distribution. In this case, each system has a different level of participation in the food distribution operations, but all are involved and must provide relevant information. The role of the food distribution frame is to facilitate communication by establishing pragmatic equivalence of the ontologies in the frame. The use of probabilistic ontologies instead of conventional ontologies improves the efficacy of this process by enabling representation of and reasoning with incomplete knowledge. In the simplified scenario depicted above, the output of probabilistic mapping combined with the pragmatic frame information would enable a plausible reasoner to infer the posterior likelihood of a scale mismatch given the sharp differences between the storage temperature reported and the statistical distribution for similar occurrences. Clearly, this example barely touches the inherent complexity of such interactions and a more realistic scenario would include various frames covering a large spectrum of relief activities. Yet, it does illustrate one of the many aspects of such interactions in which we take for granted that a human supervisor would "catch" eventual inconsistencies, and in an environment in which time and resources are very limit this reinforces the need for a flexible, adaptive scheme for establishing pragmatic equivalence.

V THE TECHNICAL INTEGRATION STRATEGY

Our approach to integrating the technologies involves developing support for pragmatic frames in the PO editor and devising the necessary adaptations to the System Entity Structure (SES) framework [47], the Component-based System Modeling and Simulation environment (CoSMoS) [48], and the reasoning aspects of MEBN / PR-OWL.

The current formulation of a pragmatic frame relies on the SES as an implementable semantic and pragmatic ontology. The axiomatic formulation of the SES allows expressions to be represented in a formal mathematical/logical language. Zeigler and Hammonds [42] show that formal representations allow translations that claim to be equivalent to be examined rigorously. In our research, we aim to extend this formulation to include support for probabilistic ontologies as a way of keeping the advantages of the current SES framework while incorporating the benefits of principled representation of uncertainty and plausible reasoning.

The Component-based System Modeling (CoSMo) offers a foundation to represent families of mixed simulatable and non-simulatable models such as DEVS models as well as UML and XML Schemas. The development of CoSMo is inspired by databases, SES, UML, and Computer Supported Collaborative Work (CSCW). A key concept of CoSMo is to visually and incrementally specify a family of state-based, hierarchical component models that can be automatically translated to executable code for target simulators. Based on this framework, a robust, scalable, integrated modeling and simulation environment called CoSMoS is developed [48,49]. It affords verification and validation of a family of models with support for built-in logical consistency for cellular, discrete-event, and discrete-time model types. Its concept and design is based on visual, persistent modeling with an underlying end-to-end process for specifying and executing parallel DEVS models using DEVS-Suite simulator [50, 51]. In fact, CoSMoS can be used as a testbed for experimenting and evaluating time-based interoperability among ontologies represented in MEBN.

V DISCUSSION AND FUTURE WORK

To assess the feasibility of the integration approach described above, a proof of concept system is planned to apply the techniques of section 3 to a complex issue such as the disaster relief problem. The major components include a MEBN reasoner with a set of probabilistic ontologies and pragmatic frames, a knowledge base, and a description of a formal theory on high-level information fusion. All these components are under development and at a stage in which is premature to claim statistical significance of its results, but their integration into a proof of concept system is already being tested in agent-based simulation environments (e.g. see high-level fusion work on PROGNOS [52]. The purpose of this paper is to present our theoretical framework for the general problem high-level knowledge fusion, which is in its own right an important contribution to the C2 community.

REFERENCES

[1] Toth, G.; Kokar, M.; Liggins, M. (2009) Directions for Higher-Level Fusion Research: Needs and Capabilities. Panel discussion at the 12th International Conference on Information Fusion, July 8, 2009, 14:20-17:00, Seattle, WA, USA.

[2] Dahmann, J.; Salisbury, M.; Barry, P.; Turrell, C.; and Blemberg, P. (1999) HLA and Beyond: Interoperability challenges, in *Information and Security*, pp. 25-42, 1999.

[3] A Tolk and J Muguira, "Levels of Conceptual Interoperability Model," Proceedings of the Simulation Interoperability Work, 2003.

[4] Zeigler, B.P.; Praehofer, H.; Kim, T.G. (2000) *Theory of Modeling and Simulation*, 2nd ed. (original 1976). New York, USA: Academic Press, 2000 & 1976.

[5] Turnitsa, C. (2005) Extending the Levels of Conceptual Interoperability Model, in IEEE Summer Computer Simulation Conference, IEEE CS Press.

[6] Tolk, A.; Diallo, S.; King, D. R.; Turnitsa, D.C. (2009) A Layered Approach to Composition and Interoperation in Complex Systems, in *Computer Systems in Knowledge Based Environments*. pp. 41-74. Berlin Heidelberg: Springer-Verlag.

[7] Page, E. H., Briggs, R., and Tufarolo, J. A. (2004) Toward a Family of Maturity Models for the Simulation Interconnection Problem. Proceedings IEEE Spring Simulation Interoperability Workshop, IEEE CS Press.

[8] Humphrey, W.S. (1989) *Managing the Software Process*, SEI Series in Software Engineering. Addison-Wesley Publishing Company.

[9] Systems Security Engineering Capability Maturity Model (SSE-CMM) Project (2003) Systems Security Engineering Capability Maturity Model (SSE-CMM) Model Description Document Version 3.0. 15 June, 2003.

[10] [C4ISR Interoperability Working Group (1998) Levels of Information Systems Interoperability (LISI).

[11] NATO Allied Data Publication 34 (ADatP-34) (2003) NATO C3 Technical Architecture (NC3TA) Version 4.0, March 2003.

[12] Costa, P.C.G. (2005) *Bayesian Semantics for the Semantic Web*. Doctoral Thesis, School of Information Technology and Engineering, George Mason University. Fairfax, VA, USA.

[13] Genesereth, M.R., Nilsson, N.J. (1987) Logical Foundations of Artificial Intelligence, Morgan-Kaufman, 1987

[14] Patel-Schneider, P. F., Hayes, P., & Horrocks, I. (2004, February 10). OWL Web Ontology Language - Semantics and Abstract Syntax. W3C Recommendation. Retrieved December 05, 2004, from http://www.w3.org/TR/owl-semantics/.

[15] Minsky; M.L. (1975) Framework for Representing Knowledge. In *The Psychology of Computer Vision*. P. H. Winston (Eds.), pages 211-277. New York, NY: McGraw-Hill.

[16] Brachman, R.J. (1977) What's in a Concept: Structural Foundations for Semantic Networks. International Journal of Man-Machine Studies, 9(2), 127-152.

[17] Berners-Lee, T., Hendler, J., & Lassila, O. (2001). The Semantic Web, in *Scientific American Digital* (pp. 7): Scientific American, Inc.

[18] Laskey, K.B. (2007) MEBN: A Language for First-Order Bayesian Knowledge Bases, *Artificial Intelligence*, 172(2-3).

[19] Getoor, L. and Taskar, B. (2007). *Introduction to Statistical Relational Learning*. Cambridge. Cambridge, MA, MIT Press.

[20] Domingos, P. and M. Richardson (2007) Markov Logic: A Unifying Framework for Statistical Relational Learning. In Lise Getoor and Ben Taskar, eds. *Introduction to Statistical Relational Learning.* MIT Press, Cambridge, MA.

[21] Richarson, M. and Domingos, P. (2006) Markov Logic Networks. *Machine Learning* (62), 107-136, 2006.
[22] Singla, P. and Domingos, P. (2007) Markov Logic in Infinite Domains. Proceedings of the Twenty-Third Conference on Uncertainty in Artificial Intelligence. AUAI Press.
[23] Heckerman, D., Meek, C., and Koller, D., (2004). Probabilistic Models for Relational Data. MSR-TR-2004-30. Redmond, WA: Microsoft Corporation
[24] Kersting, K. and De Raedt, L., (2001). Adaptive Bayesian Logic Programs. Proceedings of the Eleventh International Conference on Inductive Logic Programming (ILP 2001), Springer-Verlag.
[25] Milch, B. Marthi, B., Russell, S., Sontag, D. Ong, D., and Kolobov, A. (2007) "BLOG: Probabilistic Models with Unknown Objects". In Lise Getoor and Ben Taskar, eds. *Introduction to Statistical Relational Learning*. MIT Press, Cambridge, MA.
[26] Pfeffer, A. (2000) *Probabilistic Reasoning for Complex Systems*. Stanford, CA, Stanford University.
[27] Pearl, J. (2000). *Causality: Models, Reasoning and Inference*. Cambridge University Press.
[28] Costa, P. C. G., and Laskey, K.B. (2006) PR-OWL: A Framework for Probabilistic Ontologies, Proceedings of the Conference on Formal Ontologies and Information Systems, (FOIS 2006), November 9-11, 2006, Baltimore, MD, USA.
[29] Costa, P. C. G., Laskey, K.B., Wright, E.J., and Laskey, K.J. (2007) Probabilistic Ontologies: The Next Step for Net-Centric Operations. Proceedings of the 12th International Command and Control Research and Technology Symposium. June 19-21, 2007, Newport, RI, USA: CCRP publications (awarded as the best student paper of the modeling and simulation track).
[30] Laskey, K.B., Costa, P. C. G., Wright, E.J., and Laskey, K.J. (2007) Probabilistic Ontology for Net-Centric Fusion, Proc. of the Tenth International Conference on Information Fusion.
[31] Laskey, K.B., Costa, P.C.G. and Janssen, T. (2008) Probabilistic Ontologies for Knowledge Fusion, Proceedings of the Eleventh Annual Conference on Information Fusion. http://ite.gmu.edu/~klaskey/papers/LaskeyCostaJanssen_POFusion.pdf
[32] Costa, P. C. G., Laskey, K. B., and Laskey, K. J. (2005) PR-OWL: A Bayesian Ontology Language for the Semantic Web. Proceedings of the Workshop on Uncertainty Reasoning for the Semantic Web, International Semantic Web Conference, http://ceur-ws.org/Vol-173.
[33] Helsper, E.M., & van der Gaag, L.C. (2001) Ontologies for Probabilistic Networks: A Case Study in Oesophageal Cancer. Paper presented at the Thirteenth Dutch-Belgian Artificial Intelligence Conference. Amsterdam, The Netherlands.
[34] Laskey, K.J., Laskey, K.B., Costa, P., Kokar, M., Martin, T. and Lukasiewicz, T., eds. (2008) W3C Uncertainty Reasoning for the World Wide Web Incubator Group Report, World Wide Web Consortium, http://www.w3.org/2005/Incubator/URW3/XGR-urw3-20071002.
[35] Sharma, R., D. Poole, et al. (2007) A System for Ontologically- Grounded Probabilistic Matching. Proceedings of the Fifth Annual Bayesian Modeling Applications Workshop, Vancouver, Canada.
[36] Koller, D., Levy, A.Y., and Pfeffer, A. (1997). P-CLASSIC: A Tractable Probabilistic Description Logic. Paper presented at the Fourteenth National Conference on Artificial Intelligence (AAAI-97), July 27-31. Providence, RI, USA.
[37] Carvalho, R.N., Santos, L. L., Ladeira, M., and Costa, P.C.G. (2007) A Tool for Plausible Reasoning in the Semantic Web using MEBN. In Proceedings of the Seventh International Conference on Intelligent Systems Design and Applications, 381-386. IEEE Press.
[38] Costa, P.C.G.; Ladeira, M.; Carvalho, R.N.; Laskey, K.B.; Santos, L.L.; and Matsumoto, S. (2008) A First-Order Bayesian Tool for Probabilistic Ontologies. In Proceedings of the 21st Florida AI Research Symposium (FLAIRS).
[39] 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.

[40] Stone, L. D., C. A. Barlow, et al. (1999). *Bayesian Multiple Target Tracking*. Boston, MA, Artech House.

[41] Oh, S., S. Russell, et al. (2004). Markov Chain Monte Carlo Data Association for General Multiple-Target Tracking Problems. Proceedings of the 43rd IEE Conference on Decision and Control. Atlantis, Paradise Island, Bahamas.

[42] Zeigler, B.P.; and Hammonds, P.E. (2007) *Modeling and Simulation-Based Data Engineering: Introducing Pragmatics into Ontologies for Net-Centric Information Exchange*. Burlington, MA: Academic Press.

[43] Searle, J. (1969) *Speech Acts*. Cambridge, MA: Cambridge University Press.

[44] Searle, J. (1975) Indirect Speech Acts, in P. Cole and J.L. Morgan (eds.) *Syntax and Semantics*. Vol. 3: Speech Acts (pp.59-82). New York, NY: Academic Press.

[45] Searle, J. (1979) The Classification of Illocutionary Acts. *Language in Society*, 8: 137-151.

[46] Harnish, R.M. (1994) Mood, Meaning and Speech Acts. In S.L. Tsohatzidis (ed.*) Foundations of Speech Act Theory* (pp. 409-449). London & New York: Routledge.

[47] Zeigler, B. P. (1984) Multifaceted Modeling and Discrete Event Simulation. London: Academic Press.

[48] Sarjoughian, H.S.; and Elamvazhuthi, V. (2009) CoSMoS: a visual environment for component-based modeling, experimental design, and simulation. SimuTools 2009: 59.

[49] CoSMoS (2009) http://sourceforge.net/projects/cosmosim/, Visited August 28, 2009.

[50] Kim, S.; Sarjoughian, H.S.; Elamvazhuthi, V. (2009) "DEVS-Suite: A Simulator for Visual Experimentation and Behavior Monitoring", High Performance Computing & Simulation Symposium, Proceedings of the Spring Simulation Conference, March, San Diego, CA

[51] DEVS-Suite (2009) http://sourceforge.net/projects/devs-suitesim/, Visited August 15, 2009.

[52] Costa, Paulo C. G.; Chang, Kuo-Chu; Laskey, Kathryn B.; Levitt, Tod; and Sun, Wei (2010) High-Level Fusion: Issues in Developing a Formal Theory. Proceedings of the 13th International Conference on Information Fusion (Fusion 2010). July 26-29, 2010, Edinburgh, UK.