10th International Command & Control Research and Technology Symposium
The Future of C2: Coalition Interoperability

Towards a Formal Ontology for Military Coalitions Operations

Eric Dorion
System Engineering & Architecture Group
System of Systems Section
Defence R&D Canada – Valcartier
2459, Pie-XI North,
Val-Bélair, Québec Canada, G3J 1X5
Eric.Dorion@drdc-rddc.gc.ca

Christopher J. Matheus
Versatile Information Systems, Inc.
5 Mountainview Drive
Framingham MA 01701 USA
cmatheus@vistology.com

Mieczyslaw M. Kokar
Northeastern University
360 Huntington Avenue
Boston, MA, 02115 USA
kokar@coe.neu.edu

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE **JUN 2005**	2. REPORT TYPE	3. DATES COVERED **00-00-2005 to 00-00-2005**

4. TITLE AND SUBTITLE **Towards a Formal Ontology for Military Coalitions Operations**	5a. CONTRACT NUMBER 5b. GRANT NUMBER 5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER 5e. TASK NUMBER 5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Defence R&D Canada - Valcartier,2459, Pie-XI North,Val-Belair, Quebec Canada, G3J 1X5, ,**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S) 11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES
The original document contains color images.

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**		**11**	

Abstract

The goal of this paper is to raise some of the fundamental questions that underpin the development of formal ontologies, especially the ones that are used for systems interoperability. To realize this, the three authors independently collaborated on different aspects of this paper. In this way, questions naturally arose from the review of each others' work. In essence, this paper represents the genesis of the authors' collaboration and constitutes for them a basis for future research.

1 Introduction

Nowadays, it is generally recognized that systems interoperability is enabled only if a strong and shared semantic basis exists. Although the systems may use distinct ontologies, there must be an overlap of semantic concepts for them to exchange meaningful and contextualised information. This overlap in turn must be broad enough so that it covers the operational context that drives the interoperability requirement [1].

In 2003, the Canadian Army asked one of the authors (Eric Dorion) to realize a semantic mapping between the Land Force Command and Control Information System (LFC2IS) and the Global Command and Control System (GCCS) used by the Canadian Navy. LFC2IS is semantically based on the Multilateral Interoperability Programme (MIP) [2] Joint Consultation Command and Control Information Exchange Data Model (C2IEDM)[1] while GCCS is semantically based partly on the Over-The-Horizon Targeting GOLD (OTH-T GOLD) [2] message text format. Given this operational context, an analysis was conducted resulting in a data mapping that was captured in an Excel spreadsheet. From this, the Army developed a software prototype to support interoperability of the two systems. The capabilities of this mapping were demonstrated in the Atlantic Littoral Intelligence Surveillance and Reconnaissance Experiment (ALIX) experiment in the summer of 2004.

Somewhat in parallel to this effort, the two other authors (Chris Matheus and Mitch Kokar), had been working to formalize an OTH-T GOLD subset of the C2IEDM using the OWL Web Ontology Language in order to demonstrate the benefits of capturing the formal semantics of the model. Having been exposed to the first author's work by Erik Chaum, they leveraged the translation captured in the Excel spreadsheet to facilitate the development of their formal ontology.

These parallel efforts resulted in a number of papers [1,4,5] but it is in the conjugation of these efforts that some of the most interesting questions and observations arise. This paper aims at identifying some of these questions. This inquisitive consideration of the

[1] We will refer to the Joint Command and Control Information Exchange Data Model as C2IEDM thoughout the remainder of this text.

problem was made possible through the incessant questioning of the authors' work and was enhanced by their diverse background. On the one hand Eric Dorion is one of Canadian's contributors to MIP and thus has played a role in designing the C2IEDM; on the other hand Chris Matheus and Mitch Kokar have been working with Semantic Web [6] technologies from its early years.

In order to fully grasp the questions at hand, the reader would be invited to capture the semantics of a given domain into a formal ontology; the experience alone sheds significant light on the difficulties involved. Since it is out of the scope of this paper and probably not affordable in terms of time and energy for the reader, the rest of the paper will be structured in such a way that the reader will briefly become an ontology engineer. The questions and observations that arose from this mutual investigation will be brought and discussed in the conclusion following section.

In the next section we provide an introduction to what an "ontology" is and what constitutes a "formal" ontology. We then explore how the mapping between C2IEDM and OTH-T GOLD, as primarily defined in the Excel spreadsheet, using OWL. This section is more technical and requires some knowledge of C2IEDM and OTH-T GOLD, but this hopefully will not impede the clarity of the questions exposed in the last section of the paper.

2 Ontologies

These days, an ontology is generally understood to be an explicit, formal, machine-readable semantic model defining the classes (or concepts) and their possible inter-relations pertinent to a specific domain. The exercise by which we capture the semantics of a domain is termed *ontology engineering.* To construct an ontology one must have an ontology specification language. UML data modelling tools provide one such language but they result in ontologies more appropriate for use in human communication or as the basis for software design. Ontology engineering can also be done such that it results in a *formal* ontology; a formal ontology is one that can be mathematically proven to be self-consistent and can serve as the basis for semantically grounded (i.e., logical) reasoning. It is this latter approach that is assumed by the Semantic Web and the formal languages developed for its realization.

A major driving force for using ontologies has been the emergence of web-enabled agents [7]. These agents can reason about and dynamically integrate the appropriate knowledge and services at run-time based on formal ontologies. Ontologies are also the basis for the Semantic Web, where they are being used to create machine-readable, semantic-descriptions of Web content that can be shared, combined and reasoned about automatically by theorem provers and intelligent agents. As part of its Semantic Web effort, the W3C has developed an XML-based language called the Web Ontology Language (OWL) [8]. OWL is an emerging standard for ontologies and knowledge representations, based on the Resource Description Framework (RDF) [9] and the DARPA Agent Markup Language (DAML) [10], which is the immediate predecessor of

OWL. OWL is a declarative, formally defined language that fully supports specialization/generalization hierarchies as well as arbitrary many-to-many relationships. Both model theoretic and axiomatic semantics have been defined for OWL/DAML providing strong theoretical as well as practical benefits in terms of being able to precisely define what can and cannot be achieved with these languages. The field is relatively young, but several support tools have been developed and many more are on the horizon for creating OWL ontologies and processing OWL documents.

3 Ontology Development

The OWL-based C2IEDM to OTH-T GOLD Interoperability Ontology we developed is shown in Figure 1. Note that this ontology focuses on the subset of C2IEDM needed to capture the information in OTH-T GOLD message, which is why we refer to it as the C2IEDM Track Core Ontology. A complete ontology for C2IEDM would include many more classes and property relationships and would entail a significant amount of additional effort.

At the center of the ontology is the OBJECT-ITEM[2]. This class is used to implement specific instances of objects described in the messages. There are five subclasses of OBJECT-ITEM as shown in the figure, although for our sample data it would be sufficient to just have the subclasses MATERIAL (used to define vessels) and ORGANISATION (used to define military organisations and reporting units). The OBJECT-ITEM class is paralleled by the OBJECT-TYPE class, which also has five subclasses; again only the MATERIAL-TYPE (and its subclass, EQUIPMENT-TYPE) and the ORGANISATION-TYPE (along with its subclasses GOVERNMENT-ORGANISATION-TYPE and MILITARY-ORGANISATION-TYPE) subclasses are needed for our sample data. We suspect that the only other subclasses that might be needed to represent arbitrary OTH-T GOLD track data are PERSON and PERSON-TYPE, which would be needed, for example, to represent Prisoners Of War (POWs).

From an OTH-T GOLD point of view, the key elements of an instance of an OBJECT-ITEM pertaining to a vessel are the vessel's affiliation (e.g., Canada), its type (e.g., Frigate), its status (e.g. hostile), and its position information (e.g., location, heading, speed, etc). This information is captured in associated instances of the OBJECT-ITEM-AFFILIATION, OBJECT-ITEM-TYPE, OBJECT-ITEM-STATUS, and OBJECT-ITEM-LOCATION classes, respectively. Note that all of these classes are referenced by REPORTING-DATA instances. These instances represent "pedigree information" that record the time and source of all updates to an object's attribute and property values. For OTH-T GOLD Track Data, REPORTING-DATA instances need to specify four pieces of information: start date, start time, reporting source and source type code. The start data and time specify when the information was observed and the reporting source identifies

[2] All class and property names used in the C2IEDM Track Core ontology are taken directly from the C2IEDM model whenever possible.

the organization that reported the information. The source type code is intended to identify the type of sensor(s)/system(s) used to detect the reported information.

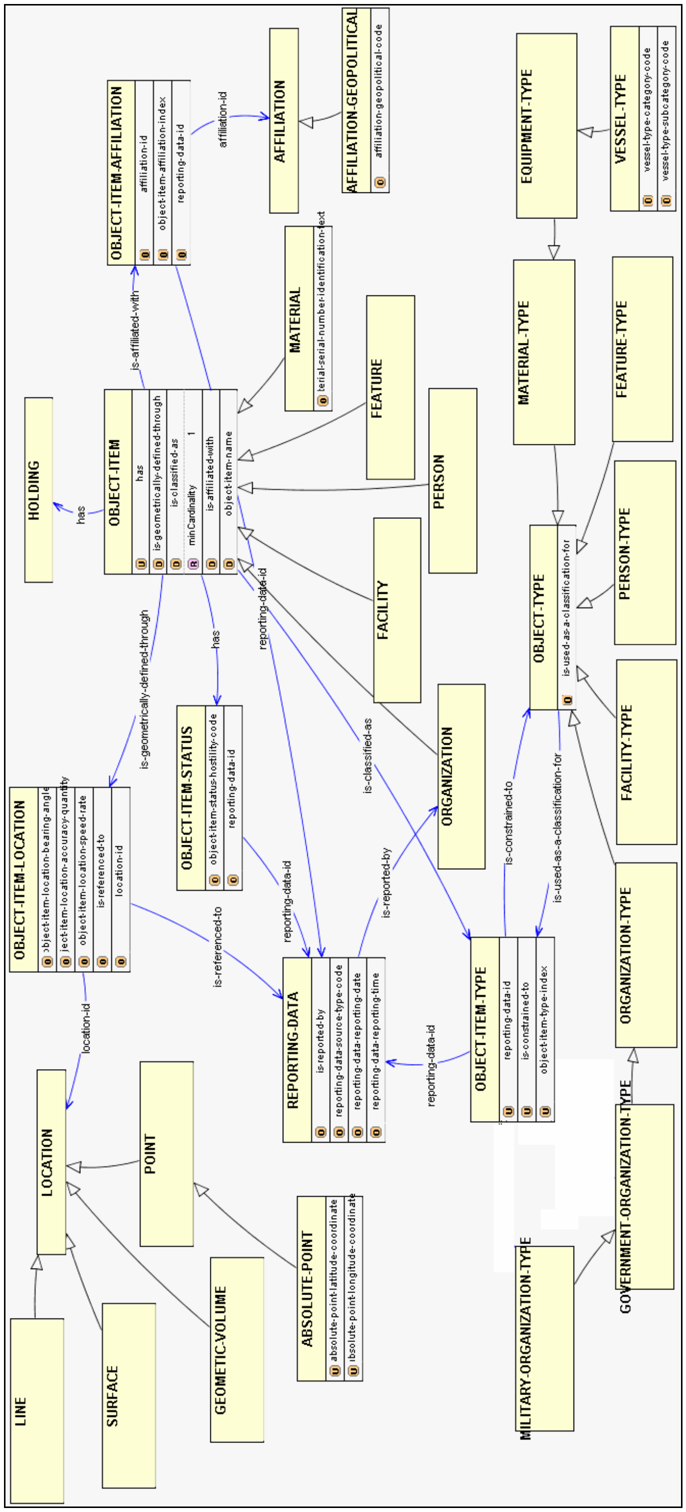

Figure 1: The OWL-based C2IEDM Ontology.

In some cases, there is too much semantic disparity between OTH-T GOLD and the C2IEDM for a complete mapping to be made [1]. A step towards remedying this situation would be to change C2IEDM's reporting-data-reporting-organisation-id to reporting-data-reporting-object-item-id making it possible to be more specific about the platform, sensor or processes that generated the track data. For a look at how pedigree information might be expanded beyond what is currently in the REPORTING-DATA class within the context of C2IEDM and OTH-T GOLD, see [11].

The affiliation of a specific vessel is defined using an instance of the OBJECT-ITEM-AFFILIATION class that references an instance of AFFILIATION. For the sample data all AFFILIATIONS are from the subclass GEOPOLITICAL-AFFILIATION that includes instances for the nationalities of Australia (AS), Canada (CA), Germany (GE), New Zealand (NZ), Spain (SP), United Kingdom (UK), USA (US) and an unspecified Enemy nation symbolized as "SD". The definitions for these AFFILIATION instances are defined in the C2IEDM Object Type Ontology, which is described below.

The OBJECT-TYPE of an OBJECT-ITEM describes the object's inherent characteristics. A specific object may have several OBJECT-TYPE's and the attribution of each type is associated with a REPORTING-DATA instance that defines who observed the specific object type and when it was reported. For OTH-T GOLD Track Data there is always only one OBECT-TYPE for an instance of an OBJECT-ITEM and it is an operational requirement that the OBJECT-TYPE is pre-defined prior the operational deployment; we have pre-defined the OBJECT-TYPES needed for the current data set in the C2IEDM Object Type Ontology as described below. The association between an OBJECT-ITEM instance and its OBJECT-TYPE is achieved through the use of an OBJECT-ITEM-TYPE instance. The C2IEDM model includes an object-item-type-index attribute to distinguish between multiple specifications of an OBJECT-TYPE from different sources. We have maintained this index, as well as others that occur in some of the other classes, in our initial design for the ontology, but it is not clear that it serves any useful purpose that cannot be equally served by the rdf:ID identifier required of all OWL instances; chances are good that we will eliminate indexes from the ontology in the next version. In OWL (and rdf) each object is uniquely identified by its rdf:ID attribute.

The OBJECT-ITEM-STATUS for an OBJECT-ITEM specifies its hostility code status. This code results from the translation of the OTH-T GOLD force-code using the mapping shown in Table 1.

Table 1. OTH-T force-code to C2IEDM object-item-status-hostility-code Mapping

Force Code	object-item-status-hostility-code
00	PENDNG
01	HO
02	PENDNG
03	FR
04	HO

05	PENDNG
06	FR
07	HO
08	PENDNG
09	FR
10	AFR
11	SUSPCT
12	NEUTRL
13	AFR
14	SUSPCT
15	NEUTRL
16	AFR
17	SUSPCT
18	NEUTRL
19	SUSPCT
20	AFR
21	NEUTRL
28	UNK
29	UNK
30	UNK
32	UNK
38	HO
39	FR

Each vessel's location and velocity information for specific times is encoded in an instance of the OBJECT-ITEM-LOCATION class. This class captures latitude and longitude coordinate information in the form of an instance of an ABSOLUTE-POINT, which is a subclass of POINT, which in turn is a subclass of LOCATION. The accuracy of the location information is encoded by a single value in the object-item-location-accuracy-quantity property. The vessel's bearing and speed are captured in the object-item-location-bearing-angle and object-item-location-speed-rate properties.

4 Observations

Section 3 described the process by which an ontology was engineered to capture and formalize the semantic concepts and relations of a specified and circumscribed domain (i.e., the C2IEDM to OTH-T GOLD interoperability data mapping). This section proposes a reflection on certain aspects of this work that will emphasize the inherent difficulties of this endeavor.

4.1 Some semantic concepts elude the ontology engineer.

An ontology is basically a set of semantic concepts and their inter-relationships. Therefore, the ontology engineer's duty is to sufficiently understand the domain to be modeled so that the resulting ontology adheres strictly to the semantics of the domain. Unfortunately, it is often the case where the ontologist fails in that duty. This has usually nothing to do with the competence of the ontologist. It actually stems from the simple fact that no one can be a subject matter expert in every subject. The MIP recognizes this by maintaining an Operational Working Group (OWG) that is responsible to formalize the Information Exchange Requirements (IERs) that pertain to the military coalition operations. These IERs are in turn broken into elementary units of information called Information Content Elements (ICEs). The ICEs are the elements against which the data modelers (ontologists) create the C2IEDM.

Although this prevents some semantic concepts to be captured in a wrongful manner in the C2IEDM, many cases are left uncovered. For example, the C2IEDM has a whole structure that is used to express every possible geometry under the LOCATION entity. Also, a LOCATION can be associated to an OBJECT-ITEM so it is situated in space. The reason it can represent anything is because it is very generic, allowing the representation of points, lines, areas, surfaces, etc. A problem appeared in the MIP Integrated Operational Test and Evaluation Exercise experiment [12] in September 2003 where the coalition Command and Control Information Systems (C2ISs) would not represent correctly military symbols on the screen. The reason for that was that the stored data elements were interpreted differently from one system to another. The MIP Data Modeling Working Group (DMWG) is currently working towards enforcing business rules that will force a single data storage solution for a given military symbol (Figure 2 and [13]).

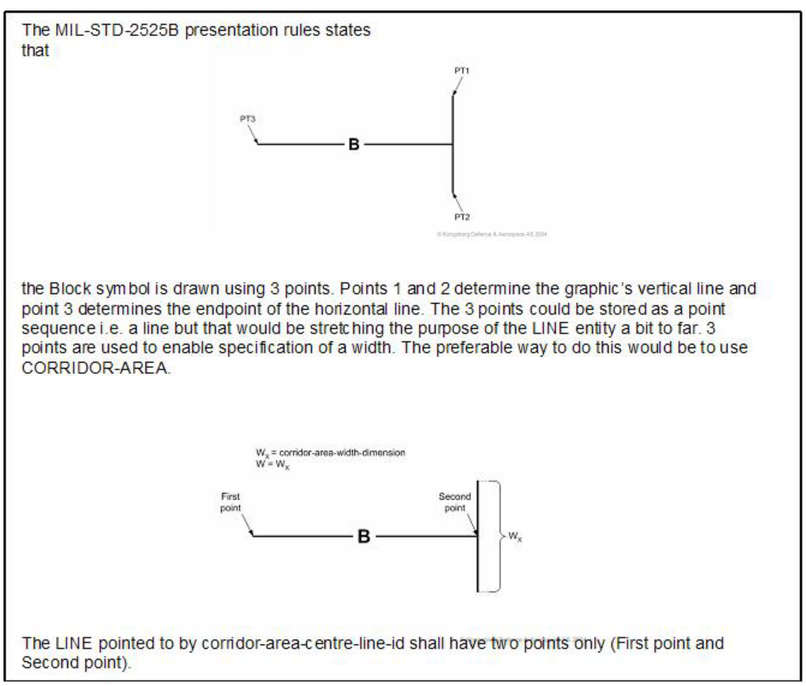

Figure 2: Data storage rules for Military Symbols

This example demonstrates a situation that can be fixed by the standardization of a single shared understanding of the way military symbols will be represented. For ontologists, some cases are trickier when the complete semantics of certain elements are missing or are unrecoverable. Let's take the C2IEDM to OTH-T GOLD country codes mapping for example. For Germany, OTH-T GOLD has 4 different values ("Germany, Federal Republic of", "Germany", "Germany, Berlin" and "Germany, Federal Republic of") while the C2IEDM only has 1 value ("Germany, Federal Republic of"). This automatically forces an imperfect mapping. Under guidance of the military subject matter experts (SMEs) and the driving operational context (the ALIX experiment), the ontologist (Eric Dorion) chose this mapping.

OTH-T GOLD	C2IEDM
German Democratic Republic	Not otherwise specified
Germany	Germany, Federal Republic of
Germany, Berlin	Not otherwise specified
Germany, Federal Republic of	Germany, Federal Republic of

A number of questions and observations can be made about this mapping though. First, it is obvious that the information cannot be pushed back and forth without losing semantics. "Germany" translates to "Germany, Federal Republic of" (left to right in the table) but is re-translated into "Germany, Federal Republic of" instead of "Germany", the initial value. One can argue that this is minor, and it may be true. Nonetheless, there is a semantic loss. The point here is that **the ontologist has to make assumptions on the acceptable level of semantic loss**. In the end, the result relies on him or her knowing enough about the domain in order to make these decisions.

4.2 On the completeness of ontologies.

Some semantic concepts comprised in an ontology seem to be self-evident from the context in which the ontology is used. In the C2IEDM and OTH-T GOLD (or other standard formats), the country code is a semantic concept that is easily associated with the need to attach geopolitical affiliation to "things". Obviously, this assumption to the ontologist is sound. But is it not always accurate. It is easy to build a counterexample where one would build a system component (a "segment" in the GCCS jargon) on top of OTH-T GOLD that tracks history of geopolitical affiliation of persons since World War II. While the segment application model would use the country codes in this context, a correct semantic mapping would consider it on the C2IEDM-based system counterpart. The current data mapping proposed by Eric Dorion (and its OWL counterpart) might prove to be wrong. **Higher level and more abstract semantic concepts must be known in order to succeed in mapping ontologies together.**

4.3 Some semantic concepts pertain only to the tools using that particular ontology.

To an ontologist, it is often very unappealing to include in the ontology anything that has to do with the tools (applications, software module, hardware that store information, etc.) or systems that will use the ontology. The problem is that **the tools are part of the ontology**. The C2IEDM, which is considered to be one of the most generic models has attributes that do not concern the operational requirements (e.g. REPORTING-DATA ent_cat_code is an attribute buried in the physical side of the C2IEDM that's function is only to capture the physical name of the entity referenced by the REPORTING-DATA). On the other hand, capturing tools-specific semantic concepts may be dangerous, especially if this ontology is to be used by other tools[3]. In MIP, a recent proposition was to add a NODE entity to the model that would be used to simplify greatly the database-to-database replication [14]. While the proposition was sound and the arguments convincing, one has to ask himself what is the actual semantic consonance of a NODE in the operational world[4]. After careful consideration, a NODE is something that has a different meaning than what was proposed and that would be used in a different manner. The tool representation of a NODE was to be different than its operational consonance and since that latter prevails (in the MIP community), the proposition must be rejected. Although they are part of the semantic universe of a community of interest, **tools semantic concepts must not interfere with the semantic concepts of the higher order** (human level).

5 Conclusion

Numerous advances have been made towards the formalization and exploitation of ontologies and knowledge in support of coalition interoperability. The MIP lives and breathes the systems-to-systems interoperability problems while the semantic web research initiatives (OWL and the like) make giant steps in improving the collective scientific knowledge by using one of the largest laboratory experiments ever, the world-wide web. As researchers, it is our duty to broaden our comprehension of what ontologies are, what can they be used for and how can they be formalized in an exhaustive way.

[3] The MIP goal is to deliver a solution that will enable system-to-system interoperability. "Tools" using the C2IEDM is a primary focus although the end goal is to support coalition interoperability (human level).
[4] Let's keep in mind that the C2IEDM is an ontology for the exchange of information between coalition partners.

References

1. E. Dorion and Anne-Claire Boury-Brisset. *Information Engineering in Support of Multilateral Joint Operational Interoperability*. CCRTS, San Diego, 2004.
2. Multilateral Interoperability Programme Web Site: http://www.mip-site.org/.
3. U.S. Navy Center for Tactical Systems Interoperability. Operation Specification for Over-The-Horizon Targeting GOLD (OS-OTH-T GOLD), March 2004.
4. C. Matheus, M. Kokar and D. Tribble. SIXA Phase I Project Final Report, ONR STTR Report, February, 2005.
5. C. Matheus, D. Tribble, M. Kokar, M. Cerutti and S. McGirr. Towards a Formal Pedigree Ontology for Level-One Sensor Fusion. 10th International Command and Control Research and Technology Symposium, McClain, Virginia, June 2005.
6. T. Berners-Lee, J. Hendler and O. Lassila. The Semantic Web. Scientific American, May 2001.
7. D. McGuinness, Ontologies and online commerce. IEEE Intelligent Systems, Vol. 16(1), pp. 8–14, 2001.
8. W3C, OWL Web Ontology Language. http://www.w3.org/TR/owl-features/.
9. W3C, Resource Definition Framework. http://www.w3.org/RDF/.
10. http://www.daml.org
11. Matheus, C., Tribble, D., Kokar, M., Cerutti M. and McGirr, S. Towards a Formal Pedigree Ontology for Level-One Sensor Fusion. In Proceedings of the 10th International Command and Control Research and Technology Symposium, McClain, Virginia, June 2005.
12. MIP IOT&E Web Site: http://www.mip-site.org/MIP_IOT.htm.
13. MIP DMWG, CP71144: Location business rules. Representation of task graphics. Greding, Germany, March 2005.
14. MIP, MIP Technical Interface Design Plan (MTIDP). Version 2.2, December 2004, Greding, Germany. http://www.mip-site.org/