Ontology-based Approach for Information Fusion

Anne-Claire Boury-Brisset
Defence R&D Canada - Valcartier
2459 Pie-XI North, Val-Belair, QC, G3J 1X5, Canada
Phone: 1 (418) 844 4000 ext. 4392, Fax: 1 (418) 844 4538
Email: {Anne-Claire.Boury-Brisset}@drdc-rddc.gc.ca

Abstract - Ontologies have received increasing interest in the computer science community and their benefits have been recognized in many areas. In this paper, we discuss the role of ontologies to facilitate information fusion from heterogeneous data and knowledge sources in support of high-level information fusion processes. We review several approaches where ontologies help provide semantic integration of information. We present preliminary work about ontological engineering for level 2 and 3 information fusion that should help semantic integration. Ontology development methods and tools should support the ontological engineering process. To this end, we propose a methodological approach and a flexible environment for ontology management that enables the building of extensible ontologies, and the mapping from ontologies to information sources.

Keywords: ontological engineering, methodology, high-level data fusion, knowledge management, databases.

1 Introduction

Commanders at all levels and types of military organizations require timely and accurate situational awareness of the battlespace as well as prediction of likely intentions of the participants. The techniques being developed for data fusion and resource management in Decision Support Systems for Command and Control are becoming increasingly more sophisticated, particularly through the incorporation of methods for high-level reasoning processes. A fundamental component of these processes is a support database (or databases) containing a priori knowledge that lists expected objects, behaviors of objects, and relationships between objects. Information sources supporting information fusion processes refer to different aspects such as political and geographical knowledge, platforms characteristics, mission guidelines, weapon characteristics, corridors and flight paths, lethality, emitter characteristics, doctrine, etc.

With the objective to implement a scheme to manage and access these information sources in order to achieve a gain in the situational awareness of a commander or crew, we proposed the concept of a knowledge server that provides users (or applications) with a unified access to relevant information from different sources within an integrated environment [1]. As part of this approach, we consider ontologies as a key component to provide a shared understanding of a domain and facilitate knowledge level interoperability among heterogeneous information sources.

The remainder of the paper is organized as follows. In the next section, we introduce the role of ontologies and give examples of some initiatives in the military domain. We then review the role of ontologies to facilitate information fusion from heterogeneous data and knowledge sources, and briefly present different approaches where ontologies help provide semantic integration of information. Then, we provide guidelines to ontology construction and present preliminary work about ontological engineering for level 2 and 3 information fusion that should help semantic integration. Ontology development methods and tools should support the ontological engineering process. We propose a framework and a flexible environment for ontology management that enables the building of extensible ontologies. This environment could be used to facilitate the mapping between the ontology and the mapping from ontologies to information sources. We finally extend the methodology to the design of an ontology-based knowledge server.

2 Ontologies and military models

2.1 Role of ontologies

Ontologies have received increasing interest in the computer science community and their benefits have been recognized in many areas such as knowledge management or electronic commerce. They explicitly encode a shared understanding of some domain that can be agreed among different parties (people or computers). The specification of an ontology comprises a vocabulary of terms, each with a definition specifying its meaning. Ontologies range from controlled vocabularies to highly expressive domain models [2]: integrated data dictionaries designed for human understanding, structured data models suitable for data management, and computational ontologies. Authors agree that there are two essential components of any ontology: a vocabulary of terms, and some specification of meaning for the terms grounded in some forms of
logic. A fundamental distinction between different approaches to ontologies is the manner of specifying the necessary relationships among terms. Formalized logical theories, i.e. computational ontologies provide support for automated reasoning.

Ontologies can be classified into different categories, ranging from general domain-independent knowledge to domain-specific knowledge.

- **Representation or meta-ontologies** conceptualize knowledge representation formalisms.

- **Upper-level ontologies** define general-level descriptive terms that form the foundation for knowledge representation. For example, *space, time or object* are domain independent terms that apply to all domains. CYC and Standard Upper Ontology (SUO) are part of this category.

- **Domain ontologies** represent specific knowledge concepts. For example, *weapon* or *missile* are specific terms of the military domain.

In the Artificial Intelligence community, ontologies constitute the foundation for the design of knowledge-based systems or intelligent agents. One of their main roles is to enable the building of knowledge models that can be reused in a wide range of applications. By doing so, they facilitate knowledge sharing and reuse. Moreover, ontologies facilitate information integration and interoperability between heterogeneous knowledge sources at a high level of abstraction. Information agents make use of ontologies to enable access to heterogeneous knowledge sources. Ontologies can be exploited to index and access semi-structured information sources. They facilitate information retrieval over collections of heterogeneous and distributed information sources. Especially, Internet search engines need domain ontologies to organize information and guide search processes. In the natural language understanding domain, ontologies provide the basis for domain knowledge representation and help identify the semantic categories that are involved in understanding discourse in that domain. Recently, the emerging field of the Semantic Web poses new challenges for ontologies and requires new techniques and technology to address them.

### 2.2 Military ontological models

In the military domain, the importance of ontologies has been recognized for years. A survey of ontologies developed within the military domain is of relevance to benefit from the efforts dedicated to ontology development in this domain. In this section, we point out some of the work described in the literature, mainly in the planning domain.

Common representation of plans has been a subject of interest for a long time. The ARPA/Rome Laboratory Planning Initiative (ARPI) led to the creation of the KRL plan language. Later, as part of the O-Plan project, A. Tate [3] proposed a structure for a plan ontology using new insights gained in the knowledge-sharing community in the US and Europe.

As part of the DARPA Joint Forces Air Component Commander (JFACC) program, an ontology for air campaign planning has been built to represent a wide variety of knowledge content in the air campaign domain [4]. The objectives of this work were to integrate knowledge acquisition and modeling efforts from developed knowledge-based applications, to create a repository for general knowledge about air campaign to use in several applications and to facilitate interoperability and communication between systems with a shared terminology. The JFACC ontology is represented in the knowledge-representation framework LOOM, based on description logics.

The DARPA HPKB (High Performance Knowledge Base) project [5] promotes technologies for developing very large, flexible and reusable ontologies and knowledge bases. In this context, an ontology of military courses of actions (COAs), i.e. outlines of plans, has been developed for a COA critiquing agent [6]. The COA ontology developed within the DISCIPLE-COA project includes objects, features and tasks, all represented as frames according to the Open Knowledge Base Connectivity knowledge model. It imports ontological knowledge from CYC and from the LOOM server, and contains the description of both concrete concepts (e.g. equipment) and abstract concepts (e.g. action or event). This ontology was extended as part of the DARPA Rapid Knowledge Formation (RKF) program to represent the concept of center of gravity (COG) used at the strategic level. As a result, the COG ontology contains many new concepts in different categories such as geographic factors, or economic factors, e.g. electrical-production-capability.

Another significant work is PLANET [7], an ontology for representing plans as a vehicle for knowledge modeling, knowledge sharing and reuse. The ontology defines concepts such as the planning problem context, goals, external constraints, or tasks. PLANET does not include representation for some entities such as agents, resources, time and location. Instead, it is used in combination of Allen’s time relations and the OZONE resource ontology.

Few research has been devoted to analyze high-level data fusion processes from an ontological perspective. W. Johnson and his colleague [8] present an ontological analysis for situation and threat assessment and describe the different types of relations between objects of the
domain. M. Kokar describes a formalization of situation awareness of in [9].

3 Ontologies for information integration

Information integration from heterogeneous sources can be addressed at the structural, syntactic or semantic levels. In this section, we review how ontologies can be used to support the integration of heterogeneous information sources. In this context, ontologies are used to describe the semantics of the information sources in order to make their content explicit. Mechanisms are required to provide mapping between ontologies and to connect ontologies with information sources. 25 different systems for intelligent information integration are reviewed in [10]. Several approaches are possible.

• **Single ontology approach:** one global ontology provides a shared vocabulary for the specification of the semantics. In this case, all information sources are related to the global domain ontology. The global ontology can be a combination of several ontologies.

• **Multiple ontologies approach:** the semantics of an information source is described by its own ontology. There is no common vocabulary, so inter-ontology mapping is needed to identify semantically corresponding terms of different source ontologies, taking into account different views on a domain (for example, different aggregation and granularity of the ontology concepts).

• **Hybrid approach:** information sources are described by local ontologies that are built from a global shared vocabulary that contains basic terms of a domain.

The **single ontology approach** is the simplest one. The drawback of the **multiple ontologies approach** is the lack of a common vocabulary that requires inter-ontology mapping. The advantage is that it facilitates the adding and removing of information sources more easily. The **hybrid approach** constitutes a good compromise. OBSERVER [11] manages local ontologies and provides mechanisms to map different ontologies, whereas TAMBIS [12] proposes a global ontology approach.

Ontology mapping is needed when different ontologies are used to describe different sources within a system. Correspondences have to be established between models. Different approaches are possible. In KRAFT [13], translations between ontologies are performed by mediator agents. It is also possible to relate all ontologies to a top-level ontology, or define a common vocabulary for defining concepts across different ontologies.

Different approaches can be considered to connect ontologies to information sources. A common approach is called **Structure Enrichment**. It consists of building a logical model that resembles the structure of the information structure with additional definition of concepts (e.g. Observer, Kraft). Another approach for the integration of information from the World Wide Web is the use of meta-annotations that add semantic information to an information source (e.g. Ontobroker).

Ontology-based approaches to information integration usually provide a three-layer architecture, exploiting an ontological layer between the presentation layer and the physical layer. At this middleware level, a mediator exploits mapping models and transforms queries into execution plans. Wrappers exploit description of the data sources at the physical layer. By doing so, a unified query language enables a transparent access to multiple diverse data sources.

Ontology-based information integration approaches adopt different languages as ontology representation languages, such as description or terminological languages, or frame-based languages. Among them, OBSERVER and TAMBIS use Description logic. Within the internet community, RDF has been proposed to describe resources, and is being used to facilitate heterogeneous information integration through web portals.

In conclusion, different approaches to information integration and system architectures have been proposed in the literature to meet specific requirements. Situation/Threat assessment and resource management are complex processes that require the exploitation of large amounts of disparate data and knowledge. In the next sections, we examine the information fusion processes in more detail in order to derive an appropriate ontology-based approach to information integration.

As an incremental approach to ontological engineering and system design is desirable in our context, ontological engineering methods and tools are required to build and manage extensible ontologies for information integration.

4 Ontology-based Approach to information fusion

High-level data fusion processes have the following properties:

• They emphasize on symbolic reasoning rather than numeric reasoning:
Multiple types of dynamic and static domain knowledge have to be processed;

There exist numerous constituency-dependency relationships among objects as well as events and activities of interest;

Hierarchical reasoning is required due to vertical organization of military entities and the multiple level of abstraction nature of the reasoning process;

Reasoning in context exploits domain feature databases to facilitate hypothesis management. Thus, it requires a data representation that supports efficient spatial and semantic search.

In other respects, there is a recognized lack of a unifying terminology within the Data Fusion community. Previous work towards the definition of a shared understanding in the fusion domain includes the effort conducted by the Joint Directors of Laboratories (JDL) Data Fusion Working Group. It has resulted in the creation of a Data Fusion Lexicon [14], and a process model for data fusion [15]. New efforts aim at formalizing the concepts related to high-level information fusion with the objective to improve human understanding across the data fusion community (e.g. defence researchers and system developers), and ultimately facilitate communication between distributed fusion systems.

In this context, our objective is to build an ontology for high-level data fusion that captures the main concepts and relationships between concepts at these levels. The ontology should specify both physical and non-physical entities involved in level 2 (situations assessment) and level 3 (threat assessment) information fusion processes. This ontology will constitute the basis to facilitate information integration from heterogeneous sources in support of information fusion processes.

4.1 Ontology development methodology

The development of ontologies is a modeling activity that is complex and time-consuming. Therefore, methodologies have emerged based on experiences gained in the construction of large ontologies. A survey of these methodologies is presented in [16]. These aim at making the development of ontologies more an engineering process rather than an art. The main stages that can be derived from these methodologies consist of the following:

- Definition of the requirements for the ontology: purpose and scope;
- Building informal specification of concepts;
- Encoding: formally represent the concepts and axioms in a language;
- Evaluation of the ontology.

After determining the purpose and scope of the ontology, the next step in the process is to identify the most important concepts in the domain, build a lexicon for these terms, and derive a comprehensive taxonomy of terms of the domain. The use of a mixed top-down and bottom-up approach to ontology development is recommended. The top-down mode may extend the definition of concepts from an existing upper-level ontology, i.e. establish links to upper-level categories that have already been defined within large ontologies (e.g. CYC) or relevant military models (e.g. NATO LC2IEDM data model). The bottom-up approach adds more specific concepts from additional reference sources (e.g. glossaries, terminology or domain databases, etc.).

The semantics of new concepts in the ontology is specified through their definition, their properties, relations with other concepts, and eventually axioms that formally specify definitions and constraints of terms in the domain. Usually, an ontology is decomposed into subdomains organized into different hierarchies of concepts. Top-level concepts being at the top of class hierarchies are sometimes called microtheories, (e.g. Military equipment).

An important aspect in the ontology development process is to explicitly establish relationships that exist between concepts. Some of the relations that can be defined between concepts are:

- Relations that link a concept with more specific concepts (is-a/subsume relation);
- Relations that link a complex object to its constituents (part-of/contains relation);
- Any variety of relations that should be specified. These relations include for example causal, functional dependencies, or temporal relations.

The development of ontologies should be an incremental process, validated by subject matter experts at each stage of the process, and should maximize subsequent reuse and extensibility.

A ontology representation language has to be chosen to encode the ontology. The degree of formality of the ontology is mainly determined by the purpose of the ontology. If the ontology is a framework for communication among people, the representation can be informal, but if the ontology is to be used by software
tools or intelligent agents to support automated tasks, then a more formal representation is required [17].

Different formalisms and knowledge representation languages have been proposed to describe ontologies. Some are limited to describing concepts, attributes, and relations and resemble conceptual models in databases or object-oriented models (ex. UML class models). UML has recently been extended to become a suitable candidate to support ontological engineering [18,19]. Other formalisms use knowledge representation paradigms such as first-order logic, frame-based or description logic. New developments within the Internet community and the semantic Web have led to new ontology languages proposals such as DAML+OIL [20], or OWL (Ontology Web Language) [21].

### 4.2 Ontological engineering for Level 2 and 3 data fusion

Ontological engineering for high-level fusion can be performed by analyzing Level 2 and 3 fusion processes in order to characterize the most important concepts that are part of these processes, and derive a specification for these concepts. This process constitutes a problem-oriented approach to ontology creation.

While level 1 data fusion deals with concrete entities such as emitters, platforms, low-level military units, level 2 and 3 are more concerned with abstract entities (e.g. event, intent, or goal). The output of Level 1, i.e., information about individual objects, is aggregated into a composite tactical picture at Level 2. This concerns the situation assessment issue and leads to a more symbolic representation of the environment and the relationships among the entities and the events in it. Situation assessment focuses on relational information to determine the meaning of a collection of entities. Moreover, environmental information is also taken into account in this analysis.

Figure 1 illustrates some important elements of Level 2 fusion that should be further refined and specified in the ontology. Relationships are not specified here, but typical relationships between entities include constituency-dependency, causal, temporal relations or geometrical proximity.

At the highest level of data fusion is the threat assessment level, or impact assessment (i.e. level 3), that projects the current situation into the future and infers about the impact of the assessed situation, the vulnerability and the force capabilities.

Figure 1. Elements of Situation Assessment

Threat assessment aims at determining engagement outcomes as well as assessing an enemy’s intent based on knowledge about enemy doctrine, level of training, political environment, etc. The focus is on intent, capability and opportunity [22]. These elements are depicted in Figure 2.

Figure 2. Elements of Threat Assessment

### 5 Ontology building environment

Ontology engineering tools provide functions for editing, browsing and visualizing ontologies through user-friendly interfaces. Several tools have been proposed from the Artificial Intelligence (AI) community for the last decade. New environments are now proposed in close relation to new ontology languages, in particular to develop the Semantic Web.

In a project aimed at building ontologies to facilitate information exchange between participants from different nations in the context of coalition operations, we have designed a collaborative ontology building environment to facilitate the capture of concepts [23]. Given the
objective of collaboratively build and implement coalition ontologies, the requirements for this environment were:

- To support the methodological process of building shared ontologies. In particular, an incremental development of ontologies was envisioned.
- To facilitate the ontology development process by non-specialists, and validation by military people, with services enabling ontology browsing and search for concepts.
- To provide the appropriate knowledge representation formalism to represent ontologies, depending on the expressiveness required, according to the role of the ontology.
- To provide functionalities related to collaborative work in order to facilitate the building of ontologies by different people in different locations (possibility to create groups of users with specific access rights, and provide a shared space to discuss the ontology under development).
- To provide import-export facilities in order to be able to reuse existing ontologies and to export the ontology being built in a specific formalism (e.g. XML or RDF Schema).

As a result, the OntoCINC Server is a web-based collaborative environment that enables different people to develop a common ontology. It provides a flexible mechanism to specify a meta-model that represents both an ontology description (concepts, attributes, relations) and collaborative aspects to facilitate discussion about the ontology under development (by adding issues, decision and related-questions properties to concepts).

This tool could be a candidate to encapsulate a high-level fusion ontology. The environment could be adapted to meet specific requirements for the modeling of high-level fusion concepts. Depending on the level of formality required, a specific meta-model could be defined as a foundation to specify the ontology. Concepts could be associated both with informal definitions from a data fusion lexicon and formal semantics. Furthermore, an export module could be integrated within the environment to translate the ontology into an appropriate formal language (e.g. DAML).

This kind of environment facilitates incremental ontology development, from taxonomy definition to formal ontology building.

6 Knowledge server design approach

Level 2 and 3 information fusion processes require a significant amount of a priori database information to support the component analyses. The KNOWMES project (KNOWledge Management and Exploitation Server) aims at providing an ontology-based knowledge server that exploits a priori databases containing heterogeneous information in support of Situation/Threat Assessment and Resource Management (STA/RM) [1].

In the next sections, we provide some guidelines of a methodological approach for the design of an ontology-based knowledge server and describe the different types of data sources managed by the server.

6.1 Methodology

The proposed methodological approach for ontology-based information fusion comprises two main stages that are interrelated.

1. Construct the ontology

- Analyze all the processes that the server will support;
- Identify all the domain objects involved in the process;
- Construct the ontology. Follow the development process presented in section 4.1 and reuse existing ontological components when possible (e.g. planning, resource, space and time ontologies);
- Make sure that the ontology covers the domain of the data sources to be exploited by the server.

2. Identify relevant data sources and their structure

- Identify all data sources to be exploited by the server (databases, knowledge bases, doctrine documents, etc.);
- Provide mechanisms to relate objects from the ontology to the information sources. This should resolve both structural and semantic heterogeneity so that unified queries can be performed through the ontological layer.

Support databases (either relational or object-oriented) have their underlying data models. Databases or knowledge bases containing instances of concepts from the ontology can be easily managed and exploited, as they directly reflect the ontology model. In particular, object-oriented databases facilitate the mapping of ontologies to
sources, e.g. a hierarchy of military resources and their instances. Geographic information systems use specific data models to represent geo-referenced information. Appropriate mappings should be applied to ontological components (spatial, geometric or geographic objects).

Free-text and semi-structured documents (doctrine, CIA World Fact Book, and other relevant military sources) are available in a form that cannot be immediately related to an ontology. Meta-annotation is needed to specify the semantics of these information sources using concepts of the ontology. One solution is to convert the documents in XML/RDF format, based on the document structure (DTD or XML/RDF Schema), tagged with ontological markup.

The global process is cyclic because both the ontology and the data sources evolve over time. The adding of new sources is likely to affect the ontology. The scope of the ontology is extended when new concepts have to included, and the level of granularity increases if sub-concepts are added within an existing hierarchy of concepts. Each time, the mapping between the ontology and the sources has to be updated accordingly.

Consequently, the challenge is to manage extensible ontologies as well as evolving underlying resources. Research is conducted to support the automatic creation of semantic markup by using a pre-existing ontology. Moreover, ontology learning is an active research area that aims at semi-automatically enriching an ontology by using text processing or machine learning techniques. Both techniques need further investigation.

6.2 Support databases

Information sources supporting information fusion processes are in different forms and formats. Major categories of databases required for level 2 and 3 data fusion are provided in [24, p. 362]. Representational techniques to support data fusion processes are usually spatial and object-oriented. However, high-level information fusion processes also exploit military doctrine, procedures, or lessons learned, usually expressed in semi-structured textual format. Consequently, three different types of data sources are currently managed by the knowledge server: an object-oriented database, a GIS database, and an XML database.

The Object-Oriented Database contains a priori data/knowledge to support the different fusion levels. It has been chosen due to its capability to organize objects and relationships between objects [25], and to manage them effectively. Low-level multi-source data fusion contains information about what can be measured, e.g. entity information, sensor information as well as geopolitical information. Situation and threat assessment deals with what can be deduced. Related data include behavior of objects (predefined patterns) and relations between objects (formation, communication scheme). Resource management is about what can be managed. The associated database contains force resources, e.g. weapon and sensor capabilities.

The GIS database contains geographical and topological information such as terrain elevation, roads, buildings, or commercial corridors.

The XML database contains documents of relevance for high-level fusion processes, such as doctrine, standard operational procedures, or rules of engagement.

7 Conclusion

In this paper, we presented an ontological approach to high-level information fusion, and its application to heterogeneous information integration for the design of a knowledge server supporting STA/RM processes. The building of an ontology is time-consuming and should be supporting by methodologies and tools. In this context, we proposed a methodological approach to ontological engineering and to information integration.

As ontologies promote knowledge reuse and sharing, we should benefit from previous work conducted in related domains, for example in the military planning domain. Other research projects conducted at DRDC-Valcartier include the COP21 TD that aims at building a situational awareness knowledge portal. The proposed conceptual framework relies on the concepts of ontologies and contextual services. The work presented herein and other ongoing initiatives should benefit from each other.

References


[15] N. Steinberg, C. L. Bowman, and F. E. White, Revision to the JDL data fusion model, Joint NATO/IRIS Conference, Quebec City, October 1998.


