Proceedings of the 2004 Winter Simulation Conference
R .G. Ingalls, M. D. Rossetti, J. S. Smith, and B. A. Peters, eds.

POTENTIAL MODELING AND SIMULATION APPLICATIONS OF THE WEB ONTOLOGY LANGUAGE - OWL

Lee Lacy
William Gerber

3505 Lake Lynda Drive, Suite 100
Dynamics Research Corporation
Orlando, FL 32817, U.S.A.

ABSTRACT

The Semantic Web is an evolution of the current world-wide web that provides explicit semantics that enable software applications to better process information representations. The Web Ontology Language – OWL – is a new language for representing information on the Semantic Web. Modeling and simulation (M&S) applications have many information representation challenges. Examples of M&S data include data tables from authoritative data sources, behaviors for computer generated forces, and descriptions of units and entities to be simulated. OWL provides a consistent syntax using the Resource Description Framework (RDF) and predefined constructs with standard semantics. These features enable better information sharing and support reasoning by inferencing systems. OWL is best used for representing object-oriented descriptions of items in a well-defined domain. It could be used in the M&S community to support distributed representations of data, behaviors, descriptions of units and objects to be simulated, and scenarios with initial conditions.

1 INTRODUCTION

The development and use of modeling and simulation applications involves many types of data and information. Developing and reusing this data consumes extensive resources.

New information representation technologies continue to emerge that hold promise for supporting modeling and simulation applications. Using these technologies to standardize information representations may reduce the costs associated with developing and using M&S data. Some of these technologies can be applied to the challenges faced by developers and operators of models and simulations. However, these technologies should be applied to problems that would benefit the most.

2 BACKGROUND

Recent advancements have been made in representing information on the World Wide Web (WWW). The Semantic Web provides new functionality and is supported by the Web Ontology Language - OWL.

2.1 Semantic Web

The “current” web was started by Tim Berners-Lee when he defined the HyperText Markup Language (HTML) and served up HTML pages over the Internet using the HyperText Transfer Protocol (HTTP). The current web is primarily focused on presenting hypertext to human readers that use web browsers.

Berners-Lee envisions the next evolution of the web, the Semantic Web, providing explicit semantics that enable software applications to better process information representations (Berners-Lee 1999).

Semantics are information (meta-data) about the meaning of represented concepts. Explicit semantics are meta-data described by computer-understandable vocabularies called ontologies. By providing explicit semantics, software can perform more sophisticated interpretations of the data.

2.2 Web Ontology Language

The Web Ontology Language – OWL – is a new language for representing ontologies and associated information on the Semantic Web. OWL was released by the World Wide Web Consortium (W3C) in February 2004 as an open standard (W3C 2004). The W3C is the same organization that manages the HTML and XML language standards. OWL evolved from the DARPA Agent Markup Language (DAML) and the European Union’s Ontology Inference Layer (OIL).

OWL extends the W3C’s Resource Description Framework Schema (RDFS) with ontological constructs for describing object-oriented classes, properties, and individuals. OWL uses the RDF/XML syntax for interchanging constructs. The complete language is called OWL Full. The OWL DL and OWL Lite sublanguages are restricted versions of OWL Full that sacrifice expressiveness for performance and simplicity.

3 M&S INFORMATION REPRESENTATION CHALLENGES

The modeling and simulation (M&S) industry has many information representation challenges associated with developing and reusing data. Simulations represent various information about their subject domain in data files that are used to prepare, execute, and debrief simulation executions. Typical M&S information representations include military unit descriptions, entity descriptions (e.g., military platforms), and behaviors for computer generated forces.

A major issue with networking simulations is interoperability. Successful interoperability requires interchange standards so that compliant software knows how to process information. Architectures have been developed with standard vocabularies and protocols for interchanging modeling and simulation-related data (e.g., DIS, SEDRIS, HLA).

Consensus standards for interchanging data are difficult to develop. For example, it is unrealistic to expect everyone to agree to name things the same. However, there is hope that developers can at least agree on how to consistently describe their application's information and data models.

3.1 XML Support for Interoperability

Metalanguages, like XML, support the description of specific Data Interchange Formats (DIFs). XML is now being used widely in the simulation world. A major initiative to support simulation developers with web technologies is the Extensible Modeling Simulation Framework (XMSF), which relies heavily on XML (Brutzman 2002). Major simulation programs (e.g., OneSAF Objective System) have adopted XML as their method of standardizing interchange files (DaCosta 2003).

To successfully benefit from using an XML DIF, all parties must have a complete understanding of the implicit semantics of the data contained in the DIF. This is often accomplished with textual documents that guide the software engineers developing simulation applications.

3.2 Authoritative Data

Modern simulations are heavily dependent on parametric values that make their software data-driven. For example, the Army Materiel Systems Analysis Activity (AMSAA) provides Joint Munitions Effectiveness Manuals (JMEM)-based data for describing the probability of hit odds for various shooter/target pairs. This data is produced by a single organization, but consumed by multiple programs in different ways. AMSAA has begun providing their data in XML.

Simulations need access to easily parseable and interoperable authoritative sources of parametric data. This data is distributed on the servers of Authoritative Data Sources. Producers are unlikely to tailor their data for each consumer. Therefore, interchange standards are required.

3.3 Computer Generated Forces Behaviors

Computer Generated Forces (CGF) behavior development represents a significant investment by the DoD. Research has been conducted on finding ways to share behaviors by representing them as XML data, rather than hard-coding them into software (Bjorkman 2001) (Lacy 2000).

Research has shown the importance of standardizing the vocabularies used in military operations, including the need for a Battle Management Language (BML) (Carey 2002). Standard vocabularies simplify the interfaces between simulations and C4I systems and guide the development of CGF behaviors. Recent efforts have involved using XML to interchange BML data (Hieb 2004).

The OneSAF Objective System (OOS) uses XML to represent CGF behavior compositions and the meta-data for software routines that support the compositions (DaCosta 2002). Standard behavior representations would enable simulation applications to share validated behaviors at varying resolution levels, thereby reducing the cost of behavior development through reuse.

3.4 Domain Descriptions

Authoritative domain descriptions are developed as part of knowledge acquisition / knowledge engineering efforts. These descriptions are needed for accurate representations of the domain. Simulations model entities whose realistic behavior is dependent on parametric descriptions of the entities' attributes. For example, in a virtual simulator for training tank drivers, the maximum speed of the tank must be defined. Objects in the synthetic natural environment are also described. For example, the weight capacity of a bridge might need to be described.

Entities are often task organized into military units. Accurate representation of unit organizations is an important factor in many military simulations. The Unit Order of Battle (UOB) Data Interchange Format was developed to interchange unit organization data (Haddix 1999).

Shareable authoritative descriptions of the domain should reduce the cost of integrating knowledge acquisition / knowledge engineering artifacts into simulations.

3.5 Scenarios

Initial conditions data in simulations often includes unit laydowns (i.e., position, orientation), initial loadings of expendables (e.g., ammunition, fuel), and other parameters that are used to initialize a simulation.

The development of simulation scenarios and associated initial conditions represents a major investment of time and resources prior to an exercise. Standard descriptions of scenarios should reduce the time and effort required to prepare for a simulation execution. This is especially critical in mission rehearsal.

The Military Scenario Development Environment (MSDE) was developed as an XML language for interchanging scenario data (Whittman 2001). Sharing scenario data could significantly reduce the cost and lead time associated with exercise preparation (Lacy 1999).

4 UPGRADING FROM XML TO OWL

XML has long been recommended and used to help interchange simulation data (DaCosta 2002). The advantage of XML markup includes that it is an open standard that is vendor-neutral and supported by COTS.

Although XML solves many format and structure interchange problems, it does not provide explicit semantics. Therefore, it is difficult for consuming software applications to correctly interpret the meaning of the data without extensive programming by software engineers that understand how the data should be interpreted.

XML provides a great deal of flexibility for language designers to organize interchange formats. However, that flexibility leads to potential problems because there are too many ways to represent the same information.

Although the Extensible Markup Language (XML) is beginning to be widely used for interchanging M&S data, it lacks the explicit semantics and standard methods of representation needed.

Semantic Web technologies provide promise for addressing some of the information representation challenges faced by the modeling and simulation industry (Blais 2004) (Lacy 2001) (Lacy 2003).

OWL overcomes the weaknesses associated with XML-only approaches by defining standard methods for representing classes, properties, and individuals. It provides a consistent XML syntax using the Resource Description Framework (RDF) and predefined constructs with standard semantics.

4.1 Inferencing

OWL provides explicit semantics about the data being described. This enables Semantic Web applications to infer new facts that are not explicitly described. For example, if a military equipment ontology states that an M1A1 tank "is-a" tank (subclass relationship), then facts that the software knows about tanks can also be inferred to be true of M1A1s. These features enable better information sharing and support reasoning by inferencing systems.

4.2 Semantic Joins

By identifying common elements in multiple information sources, the sources can be "joined" in much the same way that relational database tables are joined that have common values in record fields.

Data used in modeling and simulation often originates from distributed sources. These sources can be linked if common items can be explicitly stated in an ontology to equal each other. For example, if AMSAA provides data for an "M1A1" and another data provider references an "M1A1 Abrams", these lexically different references can be equated as the same concept in an ontology. This allows the facts provided by both sources to be folded together into a combined data source.

4.3 Dynamic Simulation Construction

Imagine someone in the future using a web-based application to define their requirements for a simulation. Software agents could scour the web for available web services to compose a simulation model. Domain descriptions and parametric data could be harvested from authoritative sources to support the composed simulation. An existing scenario could be found that could be tailored to meet the requirements. All of these activities are possible if simulation web services are described and information is represented using Semantic Web technology.

5 APPROPRIATE AND INAPPROPRIATE USES

A common challenge with new technologies is to identify appropriate applications. Often, technologists have solutions in search of problems. However, the approach should normally be to determine what solutions a new technology provides to an existing problem. Several characteristics can be considered when determining the appropriateness of an information representation for OWL (Lacy 2003).

Application developers must determine whether upgrading their information representations is worth the investment in developing ontologies and marking up their data. Representing information in OWL is most appropriate when OWL's strengths can be leveraged. Similarly, OWL should not be used in situations where its weaknesses adversely affect an application.

5.1 Strengths

OWL's strengths result from its features as an ontology language. OWL provides web-ready object-oriented information representations with an open vendor-neutral language.

OWL was developed to support the Semantic Web. It uses web technology for identifying resources using Uniform Resource Identifiers (URIs) and supports information that is distributed on multiple web servers.

The information representation constructs in OWL supports object-oriented descriptions. OWL supports the definition of classes, individual instances, and property relationships between classes, individuals, and properties.

OWL is an open W3C Recommendation (standard), so there are few potential intellectual property or proprietary licensing issues.

Because OWL is vendor neutral, it is supported by a wide variety of commercial off-the-shelf (COTS) software tools. Using COTS instead of developing custom code also reduces development costs.

Also, representations that need to be read by both humans and computers are good candidates for OWL representations. Software accessing OWL representations can perform semantic joins and inferencing.

5.2 Appropriate Applications

Several characteristics should be considered when determining the appropriateness of an information representation for OWL (Lacy 2003).

Because of its strengths, OWL is best suited for representing a particular class of information representation solutions. Appropriate applications for OWL are characterized by well-defined object-oriented domains that can be described with text, and are described on distributed web servers.

The representations should describe a well-understood domain (e.g., military equipment) so that specific descriptions can be made. Domain concepts must be specific enough to be described by OWL classes.

OWL is well suited for describing object-oriented concepts because of its class, property, and instance constructs. Modern M&S environments are typically implemented using object-oriented methods because of the close correspondence between the problem domain being modeled and conceptual objects in the software solution space. The military equipment domain is an excellent example of a well-defined object-oriented domain. These types of domains can be described with taxonomical relationships in OWL using subclass relationships.

OWL supports text-based data representation, so appropriate applications should be describable with text. Non-textual data can be represented with meta-data described with OWL.

Distributed representations are supported by the web nature of OWL. Applications that provide information on their servers for access by software applications and human readers are good candidates for OWL. Applications can manage their data using proprietary methods, but provide OWL views of the information.

5.3 Weaknesses

OWL suffers from many of the same challenges as XML, as well as challenges related to reasoning.

OWL is a relatively new technology, so there is still limited tool support and few trained practitioners. However, since OWL leverages XML for its syntax and Description Logics (DL) for its theory, it should be a straightforward process to upgrade existing XML or DL applications to support OWL.

As with XML, OWL may have performance issues due to the verbosity of the markup used to describe content. This may be especially problematic for applications with real-time requirements that require access to marked up data.

OWL ontologies primarily supports object-orientation. Therefore, it is often difficult to represent functional type relationships involving verbs or actions (e.g., task descriptions) directly with classes and properties.

Although the reasoning features of OWL are very powerful, it may be difficult to scale reasoning on a web-scale.

5.4 Inappropriate Uses

Because OWL is expressed in XML and is therefore focused on textual descriptions, it does not make sense to use OWL for representing certain types of data. Large datafiles normally represented as binary files such as terrain databases are not good candidates for OWL representations. However, the meta-data used to describe these files are excellent candidates for OWL implementations.

In terms of simulation characteristics, applications should only need the data or information in an off-line mode. It is unrealistic to inference across the web in a real-time system. OWL representations are also appropriate where heterogeneous data is shared between applications. Applications that require real time performance requirements should probably not store their data in OWL files until advancements are made in hardware and software technologies.

6 POTENTIAL M&S USES

Semantic Web technology could be used wherever there is data. However, it may not make sense to use OWL to represent all M&S data. Experimentation and prototyping should be performed to document the return on investment of upgrading to OWL ontologies.

6.1 M&S OWL Applications

Semantic Web technology could be used for a variety of modeling and simulation information representations including:

- Static authoritative domain descriptions,
- Simulation development and composition,
- Dynamic data representation, and
- CGF behaviors (Lacy 2001).

Authoritative domain descriptions often result from KA/KE efforts. Examples include unit organization data and military equipment. The Distributed Interactive Simulation (DIS) enumerations document describes a taxonomy that could easily be migrated into a military equipment ontology. Similarly, fields from relational databases like WARSIM's Equipment Characteristics Database (ECDB) could be used to help identify required ontology properties.

Information needed to support simulation development and composition can be represented using OWL. Examples of this information are the DIFs that support High Level Architecture (HLA) federation composition and the descriptions of web services needed to compose a simulation. OWL-Services (OWL-S) can be used to describe web services that provide M&S support in a Service Oriented Architecture (SOA).

OWL can be used to represent dynamic simulation data. Some information changes as often as every simulation execution. For example, scenario initial conditions files and After Action Review or logger data could be represented in OWL. A large problem in networking simulations is that they are often developed in a "stove-piped" manner and have their own data models and associated schemas. OWL supports publishing a view of the simulation's data that can be interpreted by other systems.

A Defense Modeling and Simulation Office (DMSO) research effort is evaluating the potential of OWL for representing CGF behaviors (Lacy 2003) (Gerber 2004). The focus of this research is on describing the composite behavior representations that are created from combining primitive behaviors. OWL is used to represent the composite behaviors and the meta-data for primitive behaviors. Current research is focused on describing CGF behaviors in OWL.

6.2 Implementing Semantic Web Technology

Although a variety of M&S information representations could be supported by OWL ontologies, it is important to identify the lowest-risk, highest-payoff areas to implement first.

Because of potential performance issues, non-real time applications should be considered first. Candidates for early adoption include areas that have been successfully implemented with XML. These include distributed representations of authoritative data and simulation scenarios.

Existing XML DIFs can be evolved into ontologies. The process for developing each ontology should include specifying requirements, designing the ontology with graphical notation, encoding the ontology into OWL, and testing the resulting ontology. The OWL Lite sublanguage of OWL should be sufficient for encoding M&S data. Using this species of OWL makes information accessible by more tools than OWL DL or OWL Full representations.

The results of developing M&S ontologies should be documented so that others in the community can become aware of their existence and benefit from lessons learned.

6.3 Benefits

Once M&S information is available in OWL, several benefits will be realized. M&S applications will have new features because they will have access to new functionality using the explicit semantics of the information and their ability to inference. Also, the development of applications will be faster and cheaper because reuse will be enabled.

Eventually, software agents will be empowered to perform complex tasks such as composing simulations.

7 SUMMARY

The Semantic Web represents the next generation of the web. OWL is a new language standardized by the W3C. It is used to markup information with explicit semantics.

M&S challenges that can be addressed with Semantic Web technology including representing and sharing authoritative data, CGF behaviors, domain descriptions, and scenarios.

OWL provides the ability to offer new features including inferencing, semantic joins, and eventually dynamic scenario construction.

Based on OWL's strengths at providing web-ready object-oriented information representations with an open vendor-neutral language, it should be used for representing well-defined object-oriented domains that can be described with text, and are described on distributed web servers. However, because of OWL's potential performance issues, it should be used with caution when describing data that is accessed by real-time simulation applications.

Experimentation should be performed involving the migration of M&S XML DIFs to OWL ontologies. The results of this experimentation will help determine the most appropriate uses of this technology that holds so much potential.

REFERENCES

Berners-Lee, T. 1999. *Weaving the Web*. 1st ed. San Francisco: Harper.

Bjorkman, E., J. Tyler, and P. Barry. 2001. Common Human Behavior Representation And Interchange System (CHRIS) After-Action Report, 28 August, 2001, Available online via `<http://www.msiac.dmso.mil/hobm/workshop/CHRIS_AAR5.doc>` [accessed August 19, 2004].

Blais, C., L. Lacy, and S. Vick. 2004. Semantic Web: Implications for Modeling and Simulation System Interoperability. In *Proceedings of the Fall 2004 Simulation Interoperability Workshop* (in press).

Brutzman, D., M. Zyda, J. Pullen, and K. Morse. 2002. Extensible Modeling and Simulation Framework (XMSF) Challenges for Web-based Modeling and Simulation, October 22, 2002. Available online via `<http://www.movesinstitute.org/xmsf/>` [accessed August 19, 2004].

Carey, S., M. Kleiner, M. Hieb, and R. Brown. 2002. Standardizing Battle Management Language – Facilitating Coalition Interoperability. In *Proceedings of the European Simulation Interoperability Workshop*. Paper 02E-SIW-005 (Summer 2002). Available online via

<http://www.alionscience.com/pdf/ Battle_Management.pdf> [accessed August 19, 2004].

DaCosta, B. 2002. XML Support for OneSAF Objective System Behaviors. In *Proceedings of the 11th CGF Conference*, 2002.

DaCosta B., T. Lucas, R. Outar, and D. Helton. 2003. OneSAF Repository Framework: Defining, Storing, and Interchanging XML Data. In *Proceedings of the Spring 2003 Simulation Interoperability Workshop*, March 2003.

Gerber, W. and L. Lacy. 2004. Standard Ontological Behavior Representation to Support Composability (extended abstract). In *Proceedings of the 13th Behavior Representation in Modeling and Simulation Conference*, May 17-20, 2004.

Haddix, F., R. Scrudder, and M. Hopkins. 1999. Unit Order of Battle Toolset. In *Proceedings of the Fall 1999 Simulation Information Workshop*, September 1999.

Hieb, M., J. Pullen, W. Sudnikovich, and A. Tolk. 2004. Extensible Battle Management Language (XBML): A Methodology for Web Enabling Command and Control for Network Centric Warfare. In *Proceedings of the 2004 Command and Control Research and Technology Symposium The Power of Information Age Concepts and Technologies*, Coronado Bay Resort San Diego, California June 15-17, 2004.

Lacy, L., G. Stone, III, and T. Dugone. 1999. Sharing HLA Scenario Data. In *Proceedings of the Fall '99 Simulation Interoperability Workshop*, September 1999.

Lacy, L. and A. Henninger. 2003. Developing Primitive Behavior Ontologies using the Web Ontology Language. In *Proceedings of the Interservice/Industry Training Simulation Conference (I/ITSEC)*, December 2003.

Lacy, L. 2000. Computer Generated Forces Behavior Representation and Reuse Using the eXtensible Markup Language (XML). In *Proceedings of the Simulation Interoperability Workshop*, Fall 2000.

Lacy, L. 2001. Representing Computer Generated Forces Behaviors Using eXtensible Markup Language (XML) Techniques. In *Proceedings of the 10th Conference on Computer Generated Forces*, May 15-17, 2001, Norfolk, VA.

Lacy, L. and J. Pawloski. 2001. Semantic Web Applications for Modeling and Simulation, DMSO Technical Exchange Presentation, July 11, 2001.

Lacy, L. 2003. Data Representation Using Semantic Web Technology, Data Technology Support Working Group Presentation at the MORS Data Workshop, March 26, 2003.

Whittman, R. and C. Harrison. 2001. OneSAF: A Product Line Approach to Simulation Development. In *Proceedings of the European Simulation Interoperability Workshop*, June 25, 2001.

World Wide Web Consortium. 2004. *OWL Overview*, Available on-line via <http://www.w3.org/TR/ owl-features/> [accessed February 10, 2004].

AUTHOR BIOGRAPHIES

LEE W. LACY is the Director of Orlando Operations for Dynamics Research Corporation (DRC). He has worked on major simulation programs including the OneSAF Objective System, WARSIM 2000, and the Close Combat Tactical Trainer (CCTT). He serves as the Vice-Chairman of the Board of Directors of the National Center for Simulation. He is currently the Principal Investigator on DRC's DARPA Agent Markup Language (DAML) semantic web research effort. His research areas include the semantic web, XML, and Computer Generated Forces. He received an M.S. and B.S in Computer Science from the University of Central Florida and is a Certified Modeling and Simulation Professional. His email address is <llacy@drc.com>.

WILLIAM (Bill) J. GERBER, Ph.D., Lt.Col., USAF (Retired) is is a Staff Technical Analyst for modeling and simulation technology at the Dynamics Research Corporation (DRC) Field Office in Orlando, Florida, and is the project manager for the DMSO supported and Air Force Research Laboratory (AFRL) managed contract for developing standard ontological behavior representations to support composability. He received Ph.D. and M.S.Cp.E. degrees in Computer Engineering from the University of Central Florida, an M.S.E. in Nuclear Engineering from the University of California at Los Angeles and a B.S.E.S in both Astronautics and Engineering Sciences from the United States Air Force Academy. He is a Certified Modeling and Simulation Professional. His email address is <wgerber@drc.com>.