Ontology-based Software for Generating Scenarios for Characterizing Searches for Nuclear Materials*

Richard C. Ward, Alexandre Sorokine, Bob Schlicher Michael Wright, Kara Kruse
Oak Ridge National Laboratory
Oak Ridge, TN 37831
wardrc1@ornl.gov

***Abstract*—A software environment was created in which ontologies are used to significantly expand the number and variety of scenarios for special nuclear materials (SNM) detection based on a set of simple generalized initial descriptions. A framework was built that combined advanced reasoning from ontologies with geographical and other data sources to generate a much larger list of specific detailed descriptions from a simple initial set of user-input variables. This presentation shows how basing the scenario generation on a process of inferencing from multiple ontologies, including a new SNM Detection Ontology (DO) combined with data extraction from geodatabases, provided the desired significant variability of scenarios for testing search algorithms, including unique combinations of variables not previously expected. The various components of the software environment and the resulting scenarios generated will be discussed.**

Keywords-component; ontology, software environment, scenario

I. Introduction

Recently there has been considerable interest in constructing computational systems that utilize ontologies in a multitude of ways [1, 2]. Examples are a semantic-based biosimulation modeling approach [3] that is being built on ontologies of anatomy and the physics of biology and the Gene Ontology (GO) [4] for bioinformatics. Here we present an ontology-based software framework for generating scenarios for a single searcher looking for the presence of special nuclear materials (SNM). Our software, the **ontology-driven scenario generator (ODSG),** will provide a capability to reason detailed scenario descriptions from limited user-input variables and create a multiplicity of scenarios with greater complexity than the initial input. The value to proliferation research is that this approach can be used to generate a wide variety of scenarios, incorporating complexities that were unobtainable from the intuitive heuristics, for testing detection algorithms.

The software system operates by first configuring an end-user application from the SNM Detection Ontology (SNM DO) and other data. Then the user selects scenario variables and ranges as desired. Once the variables are specified, a reverse process constructs the "data" for a series of scenarios using ontologies of data products and simulation models.

Each of the resulting scenarios can be viewed on the screen or encoded into XML or other formats, including KML [5], for further processing, and optionally converted into a human-readable narrative description. With the addition of building heights, elevations of floor levels, searcher, mobile objects, sources and other entities in the scene, the scenarios can be rendered using three-dimensional rendering software such as Blender [6].

II. General Assumptions

The present version of the ODSG software is intended to simulate an urban environment that is traversed by a single searcher on foot carrying a gamma-ray detector in a backpack. Each scenario is generated for an urban setting defined as an area in a city and described by a user-selected set of general descriptors. These general descriptors may include: location type (e.g., "city on the East coast"), the weather (temperature, humidity, etc.), information on the background radiation environment (e.g., possible presence of individuals treated with radioisotopes, presence of man-made objects, industry), hypothesized illicit locations of SNM source, and the general direction and walking time of the searcher carrying the detector.

Further, searching is assumed to be conducted only in the outdoor environment of the city with the searcher walking in non-adaptive patterns based on the shortest path to cross the search area; in this version the presence of a source does not alter the searcher's path. The software design is flexible enough so that future versions could account for teams of searchers and adaptive searching with more complex search protocols.

III. Development of Supporting Ontololgies

ODSG uses multiple ontologies to infer from a general description (a list of user-input variables) to a much more complex detailed description and generates scenarios that are used later to test algorithms of SNM detection. We developed the SNM DO based on a multitude of sources including interviews of subject matter experts (SMEs), field manuals, textbooks, and other sources. SNM DO depicts an SNM detection environment the way it is perceived by the SMEs and outlines elements of the detection environment that may affect sensor readings in the opinion of the SMEs. Fig. 1 shows the general structure of the SNM DO.

In addition to the SNM DO, several other ontologies aimed at depicting the latent background knowledge, were developed. Overall ontology development methodology was based on Basic Formal Ontology (BFO) [7]. Several ontologies were developed that describe geographic data sources, such as TIGER [8], DHS Homeland Security Infrastructure Program (HSIP) [9], and others. Also we developed ontologies for simulation models, such as models for simulating paths of moving objects and pavement and sidewalk configurations. These simulation models were used during scenario generation to substitute for missing or unavailable data. SNM DO was matched with data source and model ontologies using an intermediate ontology based on the entries commonly found in the dataset ontologies and other geographic ontologies such as SWEET [10].

The ontologies were developed using the Simple Ontology Format (SOFT) [11] that provides such capabilities as visualization of ontologies in GraphViz and reasoning over a hierarchy of entities and relations [12]. An example SOFT diagram of portions of the SNM DO is shown in Fig. 2.

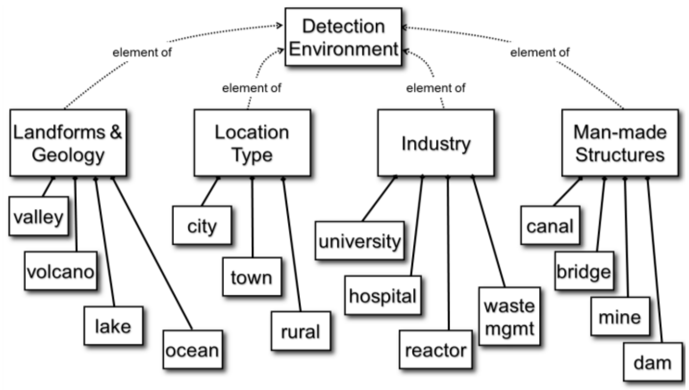

Figure 1. A portion of the special nuclear materials detection ontology (SNM DO). This portion focuses on geographic features.

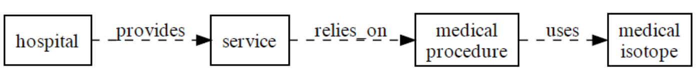

Figure 2. A portion of the SNM DO relating hospitals, procedures, services, medical procedures and isotopes.

IV. SOFTWARE ARCHITECTURE

ODSG system architecture is built around utilizing ontologies in various parts of its data processing cycle. The SNM DO and supporting ontologies are used to configure the interactive scenario generator. At the configuration stage entities from the ontologies are used to link the geodata sources and simulation models and generate the graphical user interface (GUI) of the end-user application. At the scenario generation stage user input is received through the GUI and used to construct the data by either retrieving it from the matching location in the geodatabase or running simulation models if such data are not available.

V. USER INTERFACE

A web interface was developed to capture the user's general input descriptors for the scenario. ODSG is a web-based application whose GUI is generated semi-automatically from the SNM DO and supporting ontologies. Each entity in the SNM DO corresponds to a scenario variable that can be controlled by the end-user through the GUI. Also each SNM DO entity is matched to a corresponding entity in the supporting ontologies that describes geographic datasets and simulation models available for scenario generation. The GUI generator uses these matches to deduce properties of an input variable such type (numeric, enumerated, geographic, etc.) and domain and appropriately formats the GUI elements of that variable.

The end user selects the variables of interest and provides ranges of values for those variables. For example, the user's selection might include geographic region, population, terrain type, presence of major buildings, roads, bridges, etc. associated with the location, and the presence of mobile objects such as people, cars, trucks, etc.

Given the user's input, the program selects a real urban location satisfying those criteria (e.g., East coast city with hospital and university near the scenario center). Following our assumption that the scene is restricted to the plan of the urban landscape provided by maps discussed above, a few of the variables governing scenario generation are: a) the path taken by the individual with the detector (allowed areas of walkable map); b) the types of shielding associated with the buildings or structures - these could be inferred using the ontology from the building type or use (government, school, store, etc.); c) characteristics such as types of soil, types of building materials commonly used, the vegetation present and weather (humidity or rain); d) the presence of or inference of known medical sources in individuals who have been treated or diagnosed using medical radioisotopes; and e) the presence of mobile objects such as cars, pedestrians, etc. Variation in the range of these variables comes partly from inferencing via the ontology and partly from random sampling over assumed typical ranges.

We also use ontologies to reason additional data from existing sources. For example, the possibility of finding anthropogenic radiation sources used in medical treatments can be inferred from the presence of the hospitals of certain types with the search area and thus the presence of treated individuals. Such radiation sources can be detected by the searcher. Fig. 2 illustrates one of these cases - if a hospital in the search area *provides* an oncology service that *relies_on* ventilation/perfusion (V/Q) procedures pulmonary perfusion (that *uses* Tc-99m) and pulmonary ventilation (that *uses* Xe-133), patients exiting this hospital might carry these specific medical isotopes. A SNM detection algorithm must be able to recognize these anthropogenic background sources.

VI. MOBILE OBJECTS

In the scenario generation we had to deal with mobile objects, entities such as the searcher, pedestrians, vehicles, etc. that move through the scene or otherwise change as a function of time. The searcher path is accomplished by weighting each point in a grid on the urban landscape, removing any points that have weights above a defined value (for example buildings, water features, etc.) that the searcher could not traverse, and creating an undirected graph. Using the A*search algorithm [13] a path is computed through this landscape of weighted values that is the minimum path between arbitrarily selected endpoints on roads at the edge of the scene. In addition, we used MASON [14], open source agent-based modeling (ABM) tools, to update and track the objects as they moved through the scene. The ODSG software provides random paths for up to ten pedestrians and ten vehicles.

VII. GEOGRAPHICAL INFORMATION SYSTEMS

ODSG is a web application that uses a PostgreSQL [15] database with the PostGIS extension [16] for most of its storage and data processing needs and Minnesota MapServer [17] for geographic display of the resulting scenarios. Using GIS the track of any mobile object (searcher, pedestrian, vehicles, etc.) is easily visualized and correlated to the text narrative. The approach of combining inferencing from ontologies within a GIS framework to generate scenarios enhances the capability to generate and visualize scenarios for evaluation of SNM detection algorithms. The scenarios were passed to a narrative generator where they are converted into English sentences. In addition, they can be delivered in XML format which could be passed to a 3D-georenderer (Blender) for three-dimensional display of the scene [6], or KML format to be viewed in Google Earth.

VIII. RESULTS

During the system demonstration, ODSG was used to generate about a hundred scenarios using several sets of input variables. In many cases multiple scenarios were generated from the same input data set by using iteration over the permitted ranges of variable values. All scenarios had a single searcher in the scene and many had pedestrians and/or vehicles in the scene, demonstrating the capability of adding mobile objects to the scenario generation.

An example of the capability to generate multiple scenarios from a single input is the sixteen scenarios created from the user input shown in Table I. The user input for “General US Region” is New England. The user has also selected presence in or near the scene of a railway and a port. The GIS map for one of the sixteen scenarios generated (sc0126_005) is shown in Fig. 3. Each scenario displays the searcher path (dark circles) as well the track of three vehicles (squares) passing through the scene. The railway is seen in the bottom portion of Fig. 3. The combination of location and presence of various infrastructures (such as railways and ports) generates multiple output scenarios. This example demonstrates the ease with which a large set of detailed scenarios can be constructed from a much simpler set of generalized user input variables.

Table I. A Portion of the User Input Variables for Example

Number of searchers	1		
Number of pedestrian	0		
Number of vehicles	3		
General US Region	New England		
Type of Detector	Handheld	**Material**	$LaBr_3$
Type of Search	Event-driven	By protocol	
Near search area	Railway	Port	

IX. CONCLUSIONS

Utilizing both domain-specific ontologies and those containing latent-background terminology, we have created a software environment that generates an expanded number of scenarios from a general set of user input variables for purposes of testing algorithms for detection of SNM. The specific ontology developed, the SNM DO, was built using subject-matter expert knowledge of the detection process for searchers on foot in an urban setting. The detailed dependence of the software construction and operation on the ontologies is described and a specific example of the user input variables used to create sixteen scenarios is elaborated. By using ontologies both to configure the software architecture and to drive inferencing based on ontological reasoning, we greatly expanded the number and variety of scenarios generated from a single set of user input. Such applications show the importance of incorporating ontologies into software frameworks for generation of scenarios for activities such as searching for nuclear materials.

Figure 3. Scenario sc0126_005 generated from input in Table 1. The searcher path is shown with dark circles and the vehicle tracks with squares.

ACKNOWLEDGEMENTS

* Research sponsored by U.S. Department of Energy/National Nuclear Security Administration NA-22 Simulation, Algorithms and Modeling Program under contract: OR10-Ontology Demo-PD06.

This manuscript has been authored by UT-Battelle, LLC, under contract DE-AC05-00OR22725 with the U.S. Department of Energy. The United States Government retains and the publisher, by accepting the article for publication, acknowledges that the United States Government retains a non-exclusive, paid-up, irrevocable, world-wide license to publish or reproduce the published form of this manuscript, or allow others to do so, for United States Government purposes.

REFERENCES

[1] Klein, M., D. Fensel, D. van Harmelen, F, and Horrocks, I., The relation between ontologies and XML schemas**,** *Linkoping Electronic Articles in Computer and Information Science*, 2001.

[2] Quix, C., Kensche**,** D.**,** Li, X. (2007) Matching of ontologies with XML schemas using a generic metamode**l**, Lecture Notes In Computer Science, Proceedings of the 2007 OTM Confederated international conference on the move to meaningful internet systems, Vilamoura, Portugal, 2007, Pages: 1081-1098

[3] Neal, M., (2010) Modular semantics-based composition of biosimulation models. Ph.D. Thesis, University of Washington.

[4] Gene Ontology (GO) http://www.geneontology.org/, 2011.

[5] Keyhole Markup Language (KML) http://code.google.com/apis/kml/documentation, 2011.

[6] Roberts, R and P. Pope, private conversations, 2011.

[7] Spear, A. D. (2006) *Ontology for the Twenty First Century: An Introduction with Recommendations*. Saarbrücken, Germany: IFOMIS. Retrieved from http://www.ifomis.org/bfo/documents/manual.pdf

[8] Topologically Integrated Geographic Encoding and Referencing system *(TIGER) http://www.census.gov/geo/www/tiger/*, 2011.

[9] Homeland Security Infrastructure Protection (HSIP) Gold 2011. The Department of Defense, National Geospatial-Intelligence Agency, Office of Americas, North American and Homeland Defense Division, 2011.

[10] Semantic Web for Earth and Environmental Terminology (SWEET) Ontologies, http://sweet.jpl.nasa.gov/ontology/, 2011.

[11] Simple Ontology FormaT (SOFT) http://github.com/sorokine/SOFT (in preparation).

[12] Gansner, E. R., & North, S. C. (2000). An open graph visualization system and its applications to software engineering. *Software – Practice and Experience* **30** (11), 1203--1233.

[13] Hart, P. E.; Nilsson, N. J.; Raphael, B. (1968). "A Formal Basis for the Heuristic Determination of Minimum Cost Paths". *IEEE Transactions on Systems Science and Cybernetics SSC4* **4** (2): 100–107

[14] MASON: A fast discrete-event, multi-agent simulation library, http://cs.gmu.edu/~eclab/projects/mason/.

[15] PostgreSQL: The world's most advanced open source database. (n.d.). Retrieved October 20, 2011, from http://www.postgresql.org/

[16] Team, P. D. (2007). *PostGIS – Spatial extension to PostgreSQL*. http://postgis.refractions.net/.

[17] MapServer, http://mapserver.org/, 2011.