Reasoning-Ready Sensor Data

James D. McCarthy[1], Phil A. Graniero[1]

[1]University of Windsor, Department of Earth and Environmental Sciences, 401 Sunset Ave., Windsor, Ontario, Canada, N9B 3P4
Telephone: 1-519-253-3000 x2485
Fax: 1-519-973-7801
Email: {mccartr, graniero}@uwindsor.ca

1. Introduction

Sensor webs bring environmental observations to decision makers through the use of automated data collection and dissemination methods; however the collected data has a ‘data-centric’ focus. To enable advanced problem solving on the data using reasoning and inference methods the data we use must be more ‘information-centric’. Therefore, we must transform our data into a representation that supports these analysis methods. One such representation is an ontology that defines all of the concepts and relationships in the domain of interest. The transformation of data into an ontological form can be time-consuming, requiring detailed knowledge of the target domain. The decision maker should not be expected to handle the conversions themselves; aside from the time-consuming aspect, incorrect or incomplete transformation introduces the potential for loss of information and semantics. The objective of this study was to create an automated transformation from ‘raw’ spatially explicit sensor data to ontologically and semantically rich information.

Common data representation standards were explored and classified according to their typical role and potential usage in a sensor web monitoring scenario. The classification resulted in the following groups of representations:

Data Description Languages: these representations provide a detailed description of data with little attempt to provide meaning. They do not explicitly store information related to relationships between the concepts. Examples include SensorML (Botts 2005), Observations and Measurements (Cox 2006), and GIS metadata.

Conceptual Ontologies: these representations describe the concepts and relationships related to a domain, structured as an ontology. They are not typically used for machine-based reasoning. Examples include XML Schema (W3C 2004a) and OWL (W3C 2004b) when used for representation only.

Operational Ontologies: these representations describe an ontology in a format suitable for reasoning engines such as CLIPS (Riley 2006). These languages are syntactically different than Conceptual Ontologies, but contain the same information.

Sensor web data can be encoded in several different formats such as SensorML and O&M, the Sensor Web Language (Nickerson and Lu 2004), and proprietary formats. Formats can be specific to the application using the data and the environmental constraints of the given collection scenario, and it is typical to transform data at various stages of collection and distribution to accommodate these constraints. In stages where bandwidth and power supplies are small, messages must be short. The messages can carry more information where bandwidth and power consumption are not a limiting factor.

Our approach to move from a data-centric perspective to an information-centric perspective involves transforming data supplied by a data source into an operational ontology in order to prepare the data for reasoning within an application (fig. 1). We focus on increasing the amount of information that can be derived from sensor measurements after they have been made in order to enable reasoning and improve spatial analysis capabilities, while making use of common methods of working with sensor data to minimize the impact on existing monitoring workflows.

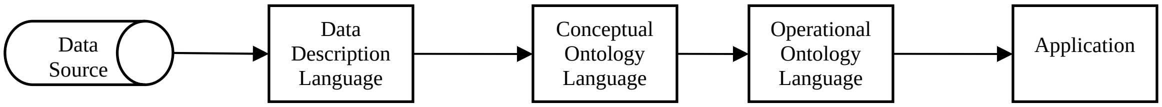


Figure 1. Generalized transformation steps for preparing reasoning-ready data.

2. The Data Representation Chain

We constructed an engine to automate this transformation chain, which can be used to specify particular representation languages and to manage the transformation process by applying appropriate processing tools at each stage. Below, we describe the particular languages and processes we chose to implement in our proof-of-concept (fig. 2).

Figure 2. Language-specific transformation steps in our proof-of-concept.

The Open Geospatial Consortium (OGC 2007) has created a set of encodings for discovering, serving, describing, and archiving sensor observations as well as descriptions of the sensors themselves. These XML-based data description languages can be used to create archived documents, but they are more typically generated by a server upon request. Our architecture uses OGC's Sensor Observation Service (SOS) (Na and Priest 2006) as the specification for serving sensor observations. Observations are stored in an SOS database that can be queried from the client side. We assume that field constraints have already been addressed in the data delivery process; for example, see the Sensor Web Automation Network (SWAN) (Graniero et al., 2007). The results generated from these queries are SensorML documents (descriptions of the characteristics of the sensors making the measurements) and O&M documents (descriptions of the measurements). When used in tandem, SensorML documents provide the context for the data stored in O&M documents. It is this contextual relationship we wish to capture in the ontology.

The first transformation takes a SensorML or O&M document and converts it from a data description language to a conceptual ontology stored in OWL. The resulting OWL document contains instances that conform to an ontology created to model the sensor web domain. Relating the instances in the sensor web ontology provides all of the information necessary for reasoning on the sensors and the measurements they make. The conversion to OWL is achieved using a set of XSLT (eXstensible Stylesheet Language Transformation) (W3C 1999) templates that map the concepts expressed in the OGC specifications to concepts in the ontology, parsing the document recursively until all of the concepts in the input document have a representation in the resulting OWL document and all of the relationships between the concepts have been created.

Once the resulting OWL document is validated against the base ontology and is confirmed as OWL-DL (a species of OWL which is guaranteed to be computationally complete and decidable), the instances of the conceptual ontology are converted into an operational ontology in CLIPS, a language geared towards the creation of rule-based expert systems. This conversion is done using Java, leveraging the Protégé ontology editor (Protégé 2007). One benefit of the two-step conversion process is to separate the structural changes from the syntactic changes. The conversion to OWL handles the structural changes while maintaining an XML-based syntax. The conversion to CLIPS is then simply a syntactic change, moving from an XML notation to a set of CLIPS instances. If both of these changes were undertaken at one time, the potential for mistakes and loss of semantic information would be much greater. It also makes it easier to translate the common conceptual ontological representation into a variety of syntactical variants to support the input requirements of a wide range of monitoring applications.

The target application is a slope hazard monitoring example built in REASON (Rozic 2006), a spatial decision support framework. REASON uses ArcAgents (Harrap and Ball, described in Rozic 2006), a bridge between a CLIPS reasoning engine and ArcGIS. The system's knowledge is partitioned into several ontologies, including a domain ontology and the CLIPS version of our sensor web ontology. In this system we can request data from an SOS database and automatically convert it into CLIPS code that contains the instances of specific sensors and their measurements. The CLIPS instances are tied to both the features stored in the GIS layers and the sensor ontology, enabling reasoning and spatial analysis on both the sensor data and the GIS data. In the example application, a slope model (Hutchinson et al., 2007) was loaded into the GIS (fig. 3) along with a domain ontology that describes the geotechnical concepts related to slope monitoring.

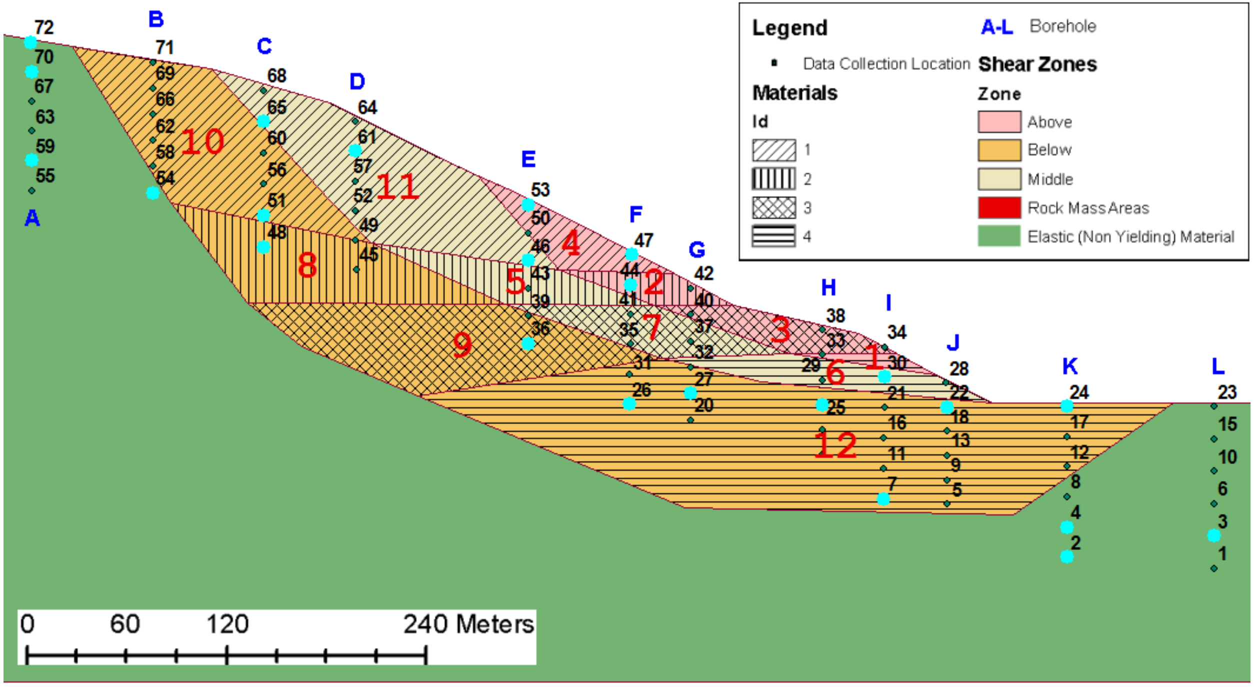
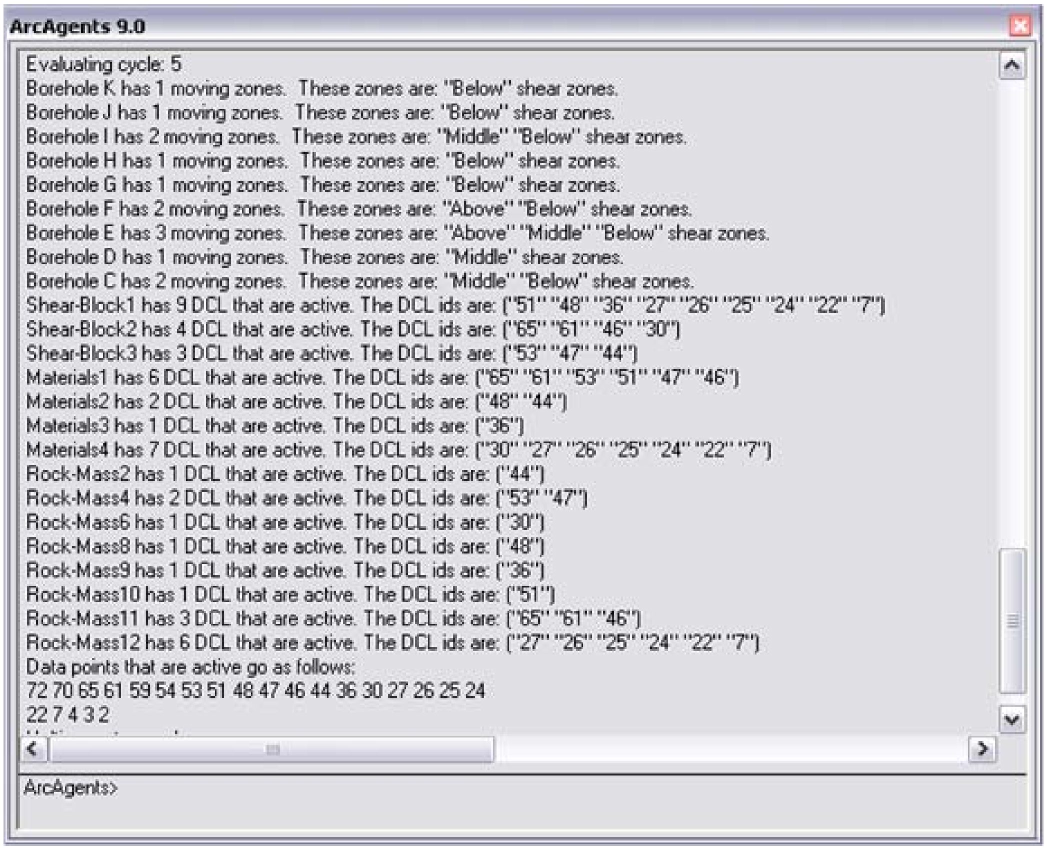


Figure 3 - Slope model within the GIS showing embedded sensors

Using this information along with the sensor ontology, the system can reason on incoming measurements from inclinometers in boreholes, identifying active zones within the rock mass and evaluating the potential for hazardous motion within the slope (fig. 4).



```
Evaluating cycle: 5
Borehole K has 1 moving zones.  These zones are: "Below" shear zones.
Borehole J has 1 moving zones.  These zones are: "Below" shear zones.
Borehole I has 2 moving zones.  These zones are: "Middle" "Below" shear zones.
Borehole H has 1 moving zones.  These zones are: "Below" shear zones.
Borehole G has 1 moving zones.  These zones are: "Below" shear zones.
Borehole F has 2 moving zones.  These zones are: "Above" "Below" shear zones.
Borehole E has 3 moving zones.  These zones are: "Above" "Middle" "Below" shear zones.
Borehole D has 1 moving zones.  These zones are: "Middle" shear zones.
Borehole C has 2 moving zones.  These zones are: "Middle" "Below" shear zones.
Shear-Block1 has 9 DCL that are active. The DCL ids are: ("51" "48" "36" "27" "26" "25" "24" "22" "7")
Shear-Block2 has 4 DCL that are active. The DCL ids are: ("65" "61" "46" "30")
Shear-Block3 has 3 DCL that are active. The DCL ids are: ("53" "47" "44")
Materials1 has 6 DCL that are active. The DCL ids are: ("65" "61" "53" "51" "47" "46")
Materials2 has 2 DCL that are active. The DCL ids are: ("48" "44")
Materials3 has 1 DCL that are active. The DCL ids are: ("36")
Materials4 has 7 DCL that are active. The DCL ids are: ("30" "27" "26" "25" "24" "22" "7")
Rock-Mass2 has 1 DCL that are active. The DCL ids are: ("44")
Rock-Mass4 has 2 DCL that are active. The DCL ids are: ("53" "47")
Rock-Mass6 has 1 DCL that are active. The DCL ids are: ("30")
Rock-Mass8 has 1 DCL that are active. The DCL ids are: ("48")
Rock-Mass9 has 1 DCL that are active. The DCL ids are: ("36")
Rock-Mass10 has 1 DCL that are active. The DCL ids are: ("51")
Rock-Mass11 has 3 DCL that are active. The DCL ids are: ("65" "61" "46")
Rock-Mass12 has 6 DCL that are active. The DCL ids are: ("27" "26" "25" "24" "22" "7")
Data points that are active go as follows:
72 70 65 61 59 54 53 51 48 47 46 44 36 30 27 26 25 24
22 7 4 3 2
```

```
ArcAgents>
```

Figure 4 - The results from one cycle in the slope monitoring application

3. Concluding Remarks

The knowledge base upon which conclusions are drawn can be improved by enhancing the domain ontology with case studies of other similar slopes and their behaviours, providing more examples for the reasoning engine to draw upon, and since the knowledge base is continually updated with new sensor data it is constantly learning more about the domain. The system can be used for hazard simulation, testing, and monitoring; validation of the quality and potential utility of sensor data; and investigation of the problem space through reasoning, inference, and spatial analysis. These are only possible, however, if measurements from conventional sensor webs can be correctly and consistently transformed into a reasoning-ready form. If the system is used for real-time hazard evaluation, the process must be automated. The transformation process described here provides the reasoning-ready sensor data that enables these analyses, providing a decision maker with detailed, relevant information about their problem domain.

4. Acknowledgements

This project was conducted with the generous financial support of GEOIDE, Project TDMDSD08.

5. References

Botts M, 2005, *OpenGIS Sensor Model Language (SensorML). Open Geospatial Consortium Best Practices Document.*

Cox S, 2006, *Observations and Measurements. Open Geospatial Consortium Best Practices Document.*

Graniero PA, Weiler M, Nickerson BG, Arp PA and Jabeur N, 2007, SWAN – Sensor Web Automation Network: An integrated sensor web design and deployment infrastructure for watershed monitoring, *GEOIDE Annual Scientific Meeting*, Halifax, Canada.

Hutchinson DJ, Diederichs MS, Carranza-Torres C, Harrap R, Rozic S and Graniero P, 2007, Four dimensional considerations in forensic and predictive simulation of hazardous slope movement, In: *Proceedings of the 1st Canada-US Rock Mechanics Symposium*, Vancouver, Canada, 8pp.

Na A, Priest M, 2006, *OpenGIS Sensor Observation Service Implementation Specification.*

Nickerson BG and Lu J, 2004, A language for wireless sensor webs. In: *Proceedings of the Communication Networks and Services Research Conference, CNSR 2004*, Fredericton, New Brunswick, Canada, 293-300.

OGC, 2007, *Open Geospatial Consortium, Inc.* Available at http://www.opengeospatial.org

Protégé, 2007, *The Protégé Ontology Editor and Knowledge Acquisition System.* Available at http://protege.stanford.edu

Riley G, 2006, *CLIPS: A Tool for Building Expert Systems.* Available at http://www.ghg.net/clips/CLIPS.html

Rozic SM, 2006, *Representing spatial and domain knowledge within a spatial decision support framework.* M.Sc. Thesis, University of Windsor.

W3C, 1999, *XSL Transformations (XSLT) Version 1.0. W3C Recommendation 14 November 1999.* Available at http://www.w3.org/TR/xslt/

W3C, 2004a, *XML Schema Part 0: Primer Second Edition. W3C Recommendation 28 October 2004.* Available at http://www.w3.org/TR/xmlschema-0/

W3C, 2004b, *OWL Web Ontology Language Overview. W3C Recommendation 10 February 2004.* Available at http://www.w3.org/TR/owl-features/