REPRESENTING GEOTECHNICAL KNOWLEDGE USING ONTOLOGIES INTEGRATED INTO A SPATIAL DECISION SUPPORT SYSTEM

Steven M. Rozic

Department of Earth Sciences, University of Windsor, 401 Sunset Avenue, Windsor, Ontario, Canada N9B 3P4. Tel (519)253-3000 ext.2503; Email rozic2@uwindsor.ca

Phil A. Graniero

Department of Earth Sciences, University of Windsor, 401 Sunset Avenue, Windsor, Ontario, Canada N9B 3P4. Tel (519)253-3000 ext.2485; Fax (519)973-7081; Email graniero@uwindsor.ca

BIOGRAPHY

Steven Rozic: M.Sc. Candidate, University of Windsor. Received B.Sc. in Software Engineering in 2004. Research interests in artificial intelligence and GIS integration, knowledge representation and reasoning, and software engineering methodology.

Phil Graniero: Assistant Professor, University of Windsor since 2000 and Researcher, Great Lakes Institute for Environmental Research since 2001. Research interests in spatial sampling strategies, geospatial and geophysical data acquisition systems, eco-hydrological modeling, wetland hydrology, landscape ecology. Published in journals including Transactions in GIS, Catena, Ecological Applications.

INTRODUCTION

A spatial decision support system (SDSS) is a software tool that can be used to support a domain expert in making complex decisions and assist with problem solving activities where spatial information and their relationships are crucial to the decision management process (Malczewski, 1999). SDSS components can conceptually include a knowledge-based subsystem, which can interact with a database management system (DBMS), external models (e.g. risk assessment models, classification tools, slope behaviour models etc.), and an organizational knowledge base (i.e. human knowledge or experience) (Sprague Jr. and Watson, 1996). The data management and to some extent the external models can be handled easily using the current capabilities of geographical information systems (GIS).

In fact, many SDSS use a GIS to manage the spatial data and many of the spatial analysis operators (Williams, 2003). Although the GIS allows for an expert to do his/her analysis visually and quantitatively, it does not assist the expert in an *intelligent* way; rather, the tool assists the expert in a very low-level, data-centric way.

In order to add the functionality of the knowledge-based subsystem, the idea of an *expert system* – a program that can represent and reason with various kinds of knowledge about the world of interest to assist in solving a problem with a given domain – can be used (Jackson, 1999). The reasoning is often carried out by manipulating representational symbols rather than data, per se. If the GIS were equipped with the reasoning capabilities that an SDSS and expert user could exploit, then the human-machine interaction could be improved during problem solving and decision making. For example, an expert may wish to perform some calculation on sample data points taken from a study area, which is composed of an assemblage of distinct rock masses. He/She may be unaware that the sample points are distributed among different rock masses. Nonetheless, the GIS will compute the result with no recognition of the unrelated subsets. The expert is left to postulate on possible causes for the unrealistic results (if the expert even realizes the results are incorrect). Even upon recognizing the problem, the expert would need to spend considerable time selecting subsets of the data and repeating the analysis. However, if the GIS contained reasoning capabilities, then the system could have recognized that the sample data points were collected in different rock masses and warned the expert of this before executing the calculation. With this knowledge the expert could quickly recognize the problem and correct it so the GIS can manage a series of correct analyses. The expert can then spend the remainder of the calculation time thinking about the spatial problem and not repeating the analysis.

The expert system represents and reasons with various kinds of knowledge by capturing the *semantics*; that is, as unambiguously as possible, the meanings of objects and their relationships in our world of interest (Hakimpour and Timpf, 2002). As of yet there are no generally available GIS tools with coupled expert systems though they are the focus of current research (Graniero, 1999). Ultimately, this new coupling will enhance the expert's ability to find possible solutions to problems within a given domain. The SDSS can then assist the expert at a higher level of problem solving by managing knowledge as well as data. In addition, human-computer interaction can be more efficient.

In order to make use of a combined GIS and expert system, many technical and human issues need to be resolved. For instance, how do we capture the semantics of the world we are interested in? How do we represent these semantics so that humans can understand them? How can we take the human-readable representation and transform it into something that the computer can understand and reason with efficiently and effectively? The key problem in addressing these questions is s*emantic heterogeneity* – the ambiguous nature of terminology used to describe facts or concepts of the world in which we are interested (Bishr, 1998). The proposed solution to remedy the problem is to use *ontologies*, formal descriptions of things and their relationships within a domain of interest (Kuhn, 2002).

The domain of interest in this paper is slope hazard management system. A number of regions in the world have geologic conditions that make them prone to slope failure by landslide. Many of these are in areas where public or private infrastructure would be directly affected; some potentially have direct effects on humans. The Geotechnical In-Situ Sensor Technology (GIST) Network is a research group working to improve the current state of technology available for monitoring unstable slopes. One of the group's goals is to develop an SDSS that can help an expert identify slope risks with complex near-real-time sensor data, or to explore management operations and historical case studies.

OBJECTIVE

For the purpose of this paper a commonly used, broad operational definition for the term ontology is "…a representative vocabulary, typically specialized to some domain or subject matter and as a body of knowledge describing some domain using such a representation vocabulary" (O'Brien and Gahegan, 2004). Based on this operational definition a higher level framework of ontologies can be established to represent a top-level ontology of foundational concepts, a domain ontology to capture vocabulary for a set of objects and terms of reference, a task ontology to capture vocabulary of particular tasks within the domain, and an application ontology to represent the vocabulary needed to solve the application problem currently at hand.

The objective of this paper is to present a candidate system that represents ontologies in a computer useable form and is integrated with a GIS and an expert system environment, more specifically, C Language Integrated Production System (CLIPS) and ESRI's ArcGIS 9.0. In order to create the GIS and expert system environment, the CLIPS programming language was connected with ArcGIS via an ArcAgents user interface extension (D. Ball, GIST Network) which includes a command prompt and a simple application programming interface (API) to allow simple operations such as creating selection sets, adding and removing data, etc. This setup provides the foundational base for the monitoring demonstrations described within this paper.

The structure and operation of the environment for reasoning analysis will be illustrated through several examples of monitoring and evaluating spatially organized phenomena. To begin, a simulated bike race will demonstrate identification of higher-level spatial concepts such as “the peloton”, “the leaders”, and “race stage”. Examples include identifying an individual rider or a pack of riders and their respective acceleration state, velocity, behaviour with respect to a course map and relative to other riders. Furthermore, recognition of concepts that rely on temporal states will be described. For example, a rider manages to veer off course (possibly trying to find an easy shortcut to the finish line), but as the race progresses the deduction that the racer is attempting to cheat no longer becomes feasible when it is discovered that the rider has stopped entirely and therefore the rider has likely crashed.

The paper then illustrates how the simple bike race example is easily transferred to the more complex slope monitoring system which will monitor individual sensors, and identify a pack of sensors as a group monitoring an individual rock mass. In slope monitoring this situation is of great concern when one sensor veers off course (i.e. its normal predicted path of natural movement) or just shuts down (i.e. stops functioning, possibly signalling a slope failure). Observing other potential sensors in the same rock mass might show that in fact the sensor must have malfunctioned because the other sensors appear to be normal, whereas if other sensors simultaneously fail, then it may indicate cable problems. Multiple sensors moving in unison within a rock mass, but differentially from other rock masses may indicate the beginning of a slope failure rather than normal migration.

After the illustrative examples are presented the paper will conclude with a discussion comparing and critically evaluating other potential environments that may suit the needs of a monitoring system of this nature. The paper will conclude with a discussion of future considerations.

REFERENCES

Bishr, Y. (1998) Overcoming the semantic and other barriers to GIS interoperability, International Journal of Geographic Information Science 12(4), 299-314.

Graniero, P. (1999) Decision support modeling helps resource managers, not programmers, Proceedings of the 13th Annual Conference on Geographic Information Systems Vancouver, BC 298-301.

Hakimpour, F. and Timpf, S. (2002) A step towards geodata integration systems, in M. Ruiz, M. Gould and J. Ramon (Eds.), 5th AGILE Conference on Geographic Information Science Palma de Mallorca, Spain.

Kuhn, W. (2002) Modeling the semantics of geographic categories through conceptual integration, in M. Egenhofer and D. Mark (Eds.), GIScience, Springer-Verlag. Lecture Notes in Computer Science.

Jackson, P. (1999) Introduction to Expert Systems, Addison-Wesley Longman Limited.

Malczewski, J. (1999) GIS and Multi-Criteria Decision Analysis, John Wiley and Sons.

O'Brien, J and Gahegan, M. (2005) A knowledge framework for representing, manipulating and reasoning with geographic semantics, in P.F. Fish (Ed.), Developments in Spatial Data Handling 11th International Symposium on Spatial Data Handling, 585 – 598.

Sprague Jr., R. and Watson, H. (1996) Decision Support Systems: Putting Theory into Practice. Prentice-Hall

Williams, H.F.L. (2003) Urbanization pressure increases potential for soils-related hazards, Environmental Geology 44, 933-938.