J. Brodeur, "Interoperability of Geographic Information: A Communication –Based Prototype", In **Proceedings of The 8th World Multi-Conference on Systemics, Cybernetics and Informatics – SCI 2004**, Orlando, Florida, 2004, pp. 327-332.

INTEROPERABILITY OF GEOGRAPHIC INFORMATION: A COMMUNICATION PROCESS –BASED PROTOTYPE

(Extended abstract)

Jean Brodeur, Ph.D.
brodeur@rncan.gc.ca

Natural Resources Canada
Centre for Topographic Information
2144 King Street West
Sherbrooke, QC Canada J1J 2E8
Phone: 1-819-564-5600 (x251)

Keywords: Interoperability of geographic information, geospatial repository, ontology communication process, cognition.

Since 1990, municipal, state/provincial, and federal governments have developed numerous geographic databases over the years throughout the World. These geographic databases were defined to fulfill organizations' specific needs. As such, same real world topographic phenomena have been abstracted independently and differently, for instance *vegetation*[1], *trees*, *wooded area*, *wooded area*, *milieu boisé*, *zone boisée* (unknown geometry) (BC Ministry of Environment Lands and Parks (Geographic Data BC), 1992; Natural Resources Canada, 1996; New Brunswick, 2000; OBM, 1996; Québec, 2000; VMap, 1995). Today, information about these geographic phenomena is accessible on the Internet from Web infrastructures specially developed to simplify their access (e.g. NSDI in United States, CGDI in Canada, GDI-DE in Germany). In order to meet their specific needs, users needs to aggregate geographic data coming multiple databases into coherent sets, which still cause many problems.

Early in the nineties, standardization bodies (e.g. OpenGIS Consortium Inc., ISO/TC 211), the research community, and the industry in geographic information undertook the development of interoperability of geographic information to solve problems of sharing and integration. More specifically, they aimed at solving syntactic, structural, and semantic heterogeneities as well as spatial and temporal heterogeneities between geographic data (Bishr, 1997; Charron, 1995; Laurini, 1998; Ouksel and Sheth, 1999; Sheth, 1999). Today, considerable progress has been made regarding syntactic and structural heterogeneities (Egenhofer, 1999; Ouksel and Sheth, 1999; Rodriguez, 2000), (e.g. ISO/TC 211, 2003a; ISO/TC 211, 2003b; ISO/TC 211, 2003c; Open GIS Consortium Inc., 2001). Nevertheless, the semantic issue must also be taken into consideration in the solution to claim complete interoperability of geographic information. A few models, such as the Semantic Formal Data Structure (Bishr, 1997), the Matching Distance model (Rodriguez, 2000) and the *Isis* solution (Benslimane, 2001), have been proposed to increase the interoperability of geographic data. However, we found necessary to develop a conceptual framework for geographic data interoperability (Brodeur, 2004), which is based on the human communication process, cognitive science, and artificial intelligence (e.g. *ontology* (Gruber, 1995)), to better understand where each contribution applies as well as to foster new ones.

[1] Spatial pictograms description: :0D ; :1D ; :2D ; :multiple geometry ; : alternate geometry (see Bédard, Y, et M-J Proulx 2002 Perceptory Web Site. WWW Document, http://sirs.scg.ulaval.ca/Perceptory)

Furthermore, we developed a reasoning methodology called *geosemantic proximity* to evaluate the semantic similarity between geographic abstractions, which is consistent with common spatial analysis methodologies used in the geographic information realm. This approach could also be used with other types of information.

Our conceptual framework develops geographic data interoperability based on a communication process between human beings, thoroughly discussed in the literature (Bédard, 1986; Blake and Haroldsen, 1975; Schramm, 1971; Sowa, 1984; Weiner, 1950), and takes into consideration the semantic issue. In interoperability of geographic data as in human communication, a user agent, which wants information about a given geographic phenomenon (e.g. street) within a specific geographic area, sends a request to a data provider agent using its own vocabulary. When the data provider agent receives the request, he/she/it begins the recognition or the interpretation of the request based on its own knowledge and vocabulary. When he/she/it has recognized the request –i.e. assigned a meaning to the request, then he/she/it gathers the data that answers the request and sends it back to the user agent. When the user agent receives the answer, he/she/it must verify if the answer complies with its original request. If so, then interoperability happens between the two agents. As you can see interoperability is here presented has a bi-directional communication process between the two agents, which can exchange information in an interoperable manner.

In this framework, we use two different types of geographic phenomenon abstractions: geographic concepts (also called *geoConcepts*) and geographic concept representations (also called *geoConceptReps*). *A geoConcept* constitutes an agent's knowledge element (i.e. an element of its ontology), which is able to recognize *geoConceptReps* corresponding to it as well as to generate *geoConceptReps* from itself. A *geoConceptRep* constitutes data about a *geoConcept* that an agent uses to interoperate with another. *GeoConcept* and *geoConceptReps* are made of intrinsic and extrinsic properties. Intrinsic properties describe the literal meaning of a geographic phenomenon whereas extrinsic properties provide meaning because of the interaction of the *geoConcept* or the *geoConceptRep* with other *geoConcepts* or *geoConceptReps*. Together, intrinsic properties and extrinsic properties provide a description of the context of a geographic phenomenon. *Geosemantic proximity* is here defined as the intersection of the context of a *geoConcept* with the context of a *geoConceptRep*. It is developed using a four-intersection matrix between the respective intrinsic and extrinsic properties of the *geoConcept* and the *geoConceptRep* in order to assess their semantic similarity and difference. Accordingly, 16 predicates are derived to qualify the nature of the *geosemantic proximity*.

Based on our conceptual framework, we have developed a prototype called *GsP prototype* using *software agents* (Nwana and Wooldridge, 1996) programmed in Java™, which communicate in XML, to determine the significance of our frameworks as well as of our notion of *geosemantic proximity*. The prototype has been tested using different data product specifications, which provided a list of topographic features of each geographic database with their inherent semantics, serving as application ontologies.

This paper aims at presenting the details of our framework for geographic data interoperability as well as our prototype. It will show the different types of abstractions involved in interoperability and where ontologies and *geosemantic proximity* specifically applies in the context of interoperability. Finally, it will demonstrate how software agents can interoperate using their own vocabulary to express their own abstractions. We foresee that the results of this work can support Web services in the interpretation of requests about geographic information in the user's vocabulary, as well as to support client software and applications in the interpretation of the data delivered in the data provider's vocabulary.

References

BC Ministry of Environment Lands and Parks (Geographic Data BC), 1992. *Digital Baseline Mapping at 1:20,000*. BC Ministry of Environment, Lands and Parks, Victoria, Province of British Columbia.

Bédard, Y., 1986. A Study of the Nature of Data Using a Communication-based Conceptual Framework of Land Information. Ph.D. Dissertation, University of Maine, Orono, Maine.

Benslimane, D., 2001. Interopérabilité de SIG : la solution Isis. *Revue internationale de géomatique*, 11(1), pp. 7-42.

Bishr, Y., 1997. Semantics Aspects of Interoperable GIS. Ph.D. Dissertation, ITC Publication, Enschede, the Netherlands.

Blake, R. H., and Haroldsen, E. O., 1975. *A Taxonomy of Concepts in Communication*. Hastings House Publishers, New York.

Brodeur, J., 2004. Interopérabilité des données géospatiales: Élaboration du concept de proximité géosémantique. Thèse de doctorat, Université Laval, Québec.

Brodeur, J. et al., 2000. Modelling Geospatial Application Databases using UML-based Repositories Aligned with International Standards in Geomatics. In: *Proceedings of Eighth ACM Symposium on Advances in Geographic Information Systems (ACMGIS)*, ACM Press, pp. 39-46.

Charron, J., 1995. Développement d'un processus de sélection des meilleures sources de données cartographiques pour leur intégration à une base de données à référence spatiale. Mémoire de maîtrise, Université Laval.

Egenhofer, M. J., 1999. Introduction: Theory and Concepts. In: *Interoperating Geographic Information Systems*, Goodchild, M. et al. eds. Kluwer Academic Publisher, Boston, Massachusetts pp. 1-4.

Gruber, T. R., 1995. Toward Principles for the Design of Ontologies Used for Knowledge Sharing. *International Journal of Human and Computer Studies*, 43(5-6), pp. 907-928.

ISO/TC 211, 2003a. *ISO 19107:2003 Geographic Information - Spatial Schema*. International Organization for Standardization, Geneva, Switzerland.

ISO/TC 211, 2003b. *ISO 19115:2003 Geographic Information - Metadata*. International Organization for Standardization, Geneva, Switzerland.

ISO/TC 211, 2003c. *WD 19136 Geographic Information - Geography markup language*. International Organization for Standardization, Geneva, Switzerland.

Laurini, R., 1998. Spatial Multi-Database Topological Continuity and Indexing: a Step Towards Seamless GIS Data Interoperability. *International Journal of Geographic Information Science*, 12(4), pp. 373-402.

Natural Resources Canada, 1996. *National Topographic Data Base - Standards and Specifications*. Centre for Topographic Information - Sherbrooke, Sherbrooke, Quebec.

New Brunswick, 2000. *Guide d'utilisation de la Base de données topographiques numériques (BDTN) du Nouveau-Brunswick*. Services Nouveau-Brunswick, Fredericton, New Brunswick.

Nwana, H. S., and Wooldridge, M., 1996. Software Agent Technologies. *BT Technology Journal*, 14(4), pp. 68-78.

OBM, 1996. *Ontario Digital Topographic Database - 1:10,000, 1:20,000 - A Guide for User*. Ministry of Natural Resources, Toronto, Ontario.

Open GIS Consortium Inc., 2001. *Geography Markup Language (GML) 2.0*. Open GIS Consortium Inc., Wayland, Massachusetts.

Ouksel, A. M., and Sheth, A., 1999. Semantic Interoperability in Global Information Systems: A Brief Introduction to the Research Area and the Special Section. *Sigmod Record*, 28(1), pp. 5-12.

Québec, 2000. Base de données topographiques du Québec (BDTQ) à l'échelle de 1/20 000 - Normes de production (Version 1.0)

Rodriguez, M. A., 2000. Assessing Semantic Similarity Among Entity Classes. Ph.D. Thesis, University of Maine, Orono.

Schramm, W., 1971. The Nature of Communication Between Humans. In: *The Process and Effects of Mass Communication*, Schramm, W., and D. F. Robert eds. University of Illinois Press, Champaign-Urbana, IL pp. 3-53.

Sheth, A., 1999. Changing Focus on Interoperability in Information Systems: From Systems, Syntax, Structure to Semantics. In: *Interoperating Geographic Information Systems*, Goodchild, M. et al. eds. Kluwer Academic Publisher, Boston, Massachusetts pp. 5-29.

Sowa, J. F., 1984. Chapter 7: Limits of Conceptualisation. In: *Conceptual Structures: Information Processing in Mind Machine*. Addision-Westley Publishing Company, Reading, Massachusetts pp. 339-351.

VMap, 1995. Vector Map (VMap), Level 1, Military Specifications, Mil-V-89033, U.S. National Imagery and Mapping Agency, Bethesda, MD.

Weiner, N., 1950. *The Human Use of Human Beings: Cybernetics and Society*. Houghton and Mifflin, Boston.