F. Fonseca and M. Egenhofer
Knowledge Sharing in Geographic Information Systems
in: P. Scheuerman and F. Bastani (eds.) *Third IEEE International Knowledge and Data Engineering Exchange Workshop*, Chicago IL, IEEE Computer Society Press, November 1999

Knowledge Sharing in Geographic Information Systems

Frederico T. Fonseca
Max J. Egenhofer
National Center for Geographic Information and Analysis
Department of Spatial Information Science and Engineering
University of Maine
fred@spatial.maine.edu
max@spatial.maine.edu

Abstract

The complexity and richness of geospatial data and the difficulty of their representation create specific problems for geographic information systems interoperability. This paper proposes the creation of software components from ontologies as a way to share knowledge and data. These software components are derived from ontologies using an object-oriented mapping. The translation of ontologies into active information system components leads to ontology-driven geographic information systems. The combined use of diverse ontologies and multiple inheritance leads to a discussion of implementation issues, such as the use of Ontolingua as an ontology editor, and CORBA, and Java as object platforms.

1. Introduction

Since Aristotle's theory of substances (objects, things and persons) and accidents (qualities, events and processes) ontology has been used as the foundation for theories and models of the world. Ontology use in Artificial Intelligence (AI) was introduced in [1]. Current research on ontology use can be found throughout the computer science community, in areas such as computational linguistics and database theory. It covers fields ranging from knowledge engineering, information integration, and object-oriented analysis to applications in medicine, mechanical engineering, and Geographic Information Systems.

A Geographic Information System (GIS) is a "computer-based information system that enables capture, modeling, manipulation, retrieval, analysis and presentation of geographically referenced data" [2]. The need to share geographic information is well documented [3-5]. In the past, exchanging geographic information was as simple as sending paper maps or raw data tapes through the mail. Today, there is a huge amount of data gathered about the Earth, computers throughout the world are connected, and the use of GIS has become widespread. Although spatial information systems have been characterized as an integration tool, GIS interoperability is far from being accomplished [6]. Research on interoperability is motivated by the ever-increasing heterogeneity of the computer world. Research on integration of databases can be tracked back to the mid 80s [7] and now interoperability is turning into an integration science [8]. Heterogeneity research in GIS is not an exception, but the complexity and richness of geographic data and the difficulty of its representation raise specific issues for interoperability. We suggest here that sophisticated structures like ontologies are good candidates for representing and abstracting geographic data.

The use of ontology in information system building is thoroughly discussed in [9] and specifically in GIS building in [10, 11]. Ontology playing a software specification role was suggested as early as 1991 in [12]. Also in 1991, Nunes [13] pointed out that the first step to build a next generation GIS would be the building of a systematic collection and specification of geographic entities, their properties and relations. But philosophers and software engineers have different point of views about ontologies. Ontology as an engineering artifact describes a certain reality with a specific vocabulary using a set of assumptions regarding the intended meaning of the vocabulary words. A particular system of categories that reflects a specific view of the world is the philosophical meaning of ontology [9]. Smith [14] notes that since ontology for a philosopher is the science of being, of what is, it is inappropriate to talk of a plurality of ontologies as engineers do. To solve this problem Smith suggests a terminological distinction between referent or reality-based ontology (R-ontology) and elicited or epistemological ontology (E-ontology). R-ontology is a theory about how the whole universe is organized and corresponds to the philosopher point of view. An E-ontology fits the purposes of software engineers and information scientists and is defined as a theory about how a given individual, group, language or science conceptualizes a given domain.

The contribution of this paper is to show how the use of ontologies in GIS development can enable knowledge sharing and information integration beyond all the advantages presented in [10, 11]. The proposed approach provides dynamic and flexible information exchange and allows partial integration of information when completeness is impossible. The remainder of this paper is organized as follows. Section 2 gives an overview of knowledge sharing tools and a foundation for Ontology-Driven Geographic Information Systems. Section 3 introduces characteristics of knowledge and data sharing in GIS, describes the general structure of an ODGIS. Section 4 discusses implementation issues relating to translation of ontologies, object platform, and databases. Section 5 presents conclusions and future work.

2. Related work

Knowledge acquisition and representation are the foundation of AI systems. The cost of building knowledge bases from scratch is high. Neches et al. [15] suggest that it is difficult to lower this cost and it is better to focus research on sharing the acquired knowledge. Sharing is the only way to build qualitative bigger knowledge-based systems because we can rely on previous labor and experience. In this section we review work about tools that support research on ontology building [16], how information systems can be build using ontologies [9], and importance of ontology use for GIS [10, 14].

2.1. Ontology and knowledge sharing

Gruber [16] addresses the problem of portability for ontologies. Ontologies should be expressed in some common formalism in order to be shared. The adopted approach is to translate an ontology specified in a standard, system-independent form into specific computer language representations.

The language to specify the ontologies is called Ontolingua. The syntax and semantics of Ontolingua definitions are based on Knowledge Interchange Format (KIF). KIF is a monotonic first-order predicate calculus with a simple syntax and support for reasoning about relations. The Ontolingua Server [17] extended Ontolingua by providing explicit support for building ontological modules and making an explicit separation between an ontology's presentation and representation. The Ontolingua Server allows multiple users to collaborate in ontology construction in a shared section. It also accepts queries from remote applications. The translation strategy implemented by Ontolingua allows the use of ontology in the development and production phases of a system. The use of an ontology, translated into an active information system component leads to Ontology-Driven Information Systems (ODIS) [9] and, in the specific case of GIS, leads to what we call Ontology-Driven Geographic Information Systems (ODGIS).

2.2. A foundation for ODGIS

Geographical entities are such complex objects that only sophisticated structures such as ontologies can represent or abstract them. [11] present the reasons for building an ontology of geographic kinds. This will enable the understanding of how different information communities exchange geographic information. The study of the ontology of geographic kinds highlights certain characteristic types of distortions that are involved in our cognitive relations to geographic phenomena. Geographic information systems need to manipulate representations of geographic entities, and ontological study of the corresponding entity types, especially those at the basic level, will provide default characteristics for such systems. Entity types present in ontologies can be use to improve the way data is exchanged either in the semantic or the representation aspects. Furthermore, the ontology of the geographic space, of the geographic objects and of the phenomena of the geographic space is different from other ontologies because topology and part-whole relations play a major role in the geographic domain. Topology is important because geographic objects prototypically can be connected or contiguous, scattered or separated, closed or open. A theory of part and whole, or mereology, is important because geographic objects are typically complex and have constituent parts [11]. [18] introduces mereotopology, an alliance of topological methods with the ontological theory of part and whole.

Frank [10] thinks that the use of ontologies will contribute to better information systems by avoiding problems that arise due to the lack of consistency. This is pointed out by Kuhn [19] that asks for spatial information theories that look toward GIS users instead of focusing on implementation issues. Ontology use can also help GIS move beyond the map metaphor that sees the geographic world as layers of independent information that can be put over one another. The several inadequacies of the map metaphor are pointed out in [20]. Nunes [13] says that GIS must evolve from the first conceptualizations of geographic space that were mediated by maps. A map is already a representation of the geographic world. Ontology use can help to implement the real GIS that works with a model of the world instead of being just automated maps [21]. Using ontologies to build GIS applications can help data integration and avoid problems such as inconsistency between ad-hoc ontologies built into the system. However, there is a gap between ontologies and software components. To allow transfer of knowledge from ontologists (the specialists in the area of the application) to software engineers it is necessary to focus on the consistent part of an ontology instead of highlighting differences between ontologies. It is also necessary to exclude the historical and philosophical point of view. Both the engineering and the cognitive views of the world are necessary to produce the small theories that account for the behavior of certain parts of reality. The first is necessary to integrate engineering knowledge into the system and the second to make it understandable to the user at the interface [10]. Ontology plays an essential role in the construction of GIS, since it allows the establishment of correspondences and interrelations among the different domains of spatial entities and relations [11].

Another important point is the definition of who the producers and consumers of ontologies and of ontology-driven information systems are. The concept of the geospatial information community (GIC) as a group of users that share a digital geographic information language and spatial feature definitions was introduced in [3]. [22] revised this concept considering a GIC as "a group of spatial data producers and users who share an ontology of real world phenomena". Guarino [9] agrees considering an ontology as a particular knowledge base that describes facts always true for a community of users. This revised GIC concept is fundamental for ODGIS.

3. Data and knowledge sharing in GIS

The issue of exchanging data and knowledge among GIS users is different from other information systems. The location component in present in geographic information is the basic difference. If we consider a hierarchy of information communities and the correspondent ontologies, reuse of data is done more horizontally and reuse of knowledge is done vertically and horizontally. It is more likely that a GIC needs data of neighboring community than from a distant one. Knowledge sharing is more likely to occur over this vertical axis, through the action of user associations that can gather information and specify high-level ontologies. This kind of knowledge is easily transferable to information communities all over this vertical axis. ODGIS provide data and knowledge exchange allowing upward and downward navigation through a hierarchy of classes derived from ontologies.

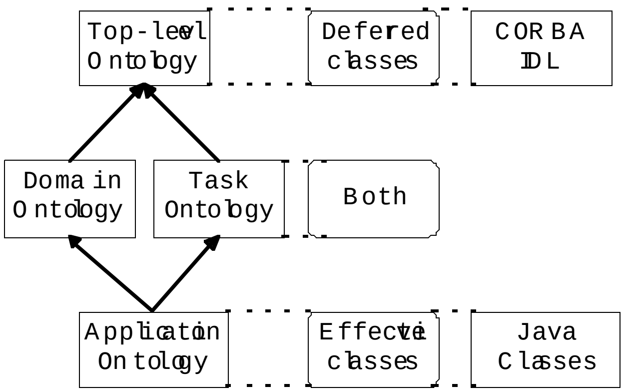
3.1. Ontology-Driven GIS

ODGIS are built using software components derived from diverse ontologies. These software components are classes that can be used to develop new applications. Being ontology-derived, these classes embed knowledge extracted from ontologies. We present here how the system is built and how it is used. The system architecture for an ODGIS is shown elsewhere [23].

Representing geographic entities, either man made features or natural differentiations on the surface of the earth, is a complex task. As Smith [11] argues, they are not merely located in space, they are tied intrinsically to space. For instance, boundaries that seem simple can in fact be very complex; an example is the contrast between country boundaries, which can be somewhat fuzzy, and land parcel limits whose boundaries are crisp. A user that develops an application can make use of the accumulated knowledge of experts that have specified an ontology of boundaries instead of dealing with these complex issues. The same is true for ontologies that deal with geometric representations, land parcels, environmental studies. Users can create new ontologies using the existing ones as a source. The user can specialize an ontology by substituting some of its members or can extend it through the inclusion of new members or relations.

Once the ontologies are specified they can be translated into software components. These components can be seen as frameworks or interfaces. Classes should be provided to perform the activities specified in the ontology. The framework can be used in the conceptual phase of the system. The classes and the framework are necessary to develop applications. The framework comprises the whole ontology and its component parts. The application developer can combine classes from diverse ontologies and create new classes that represent the user needs. This way, a class that represents a land parcel for a specific city can be built from Land parcel component specified in Urban ontology, from polygon specified in Geometry ontology and from crisp boundary specified in Boundary ontology. So the real class is land parcel of city A, but it plays many roles that together give the class its unique characteristic.

The application developer can derive new classes, more specific to the application, called user classes, which are different from more generic ontology classes. The user classes belong to the level of the application ontology. Ontology classes belong to the level of top-level, domain, or task ontologies[9]. Ontology classes are deferred classes and user classes are effective classes [24] (Figure 1).



Ontology classes, extended from [9]

Knowledge and data sharing is accomplished through the use of classes or roles that belong to common ontologies or through conversion of instances of classes up and down in the ontology hierarchy.

We propose here two types of ontology integration. First, before the generation of the translated components, new ontologies can be composed through derivation of existing ontologies or through the insertion of existing ontology references into new ontologies. The second option is to combine the translated components using multiple inheritance to create new classes. The new class plays many roles that correspond to classes used as parents.

4. Implementation issues

We are suggesting here specific tools for implementation. We know that these tools are not the only solutions but we believe that evolution of ontology-driven information systems will use similar tools or an evolution of these same tools.

We use here the structural dimension suggested by [9] when analyzing the impact of ontology on information systems basic components: application programs, databases and user interfaces. On the ODGIS approach the application program will rely on classes derived from ontologies. These classes can be as complex as an entire ontology or as simple as a single ontology component. The application developer is able to browse the ontology that is the origin of these classes. The ontology browser has multiple functions. First, during ontology specification many users can compose an ontology at the same time in a collaborative process. After the ontology has been completed, it is used as an index of geographic entities that are available to the users. These entries in the ontologies point to features stored in spatial databases.

We give here a brief description of an ODGIS architecture. The main components of the architecture are the ontologies, the container, the data warehouses, and the user interface. The architecture includes a coordinator that integrates all other components. The coordinator is also in charge of finding services on the network, and redirecting them to the components. These services can include an ontology-based search engine for geographic information, a query language server, or a framework that can function as a user interface with other embedded services. Such a system deals with instances of classes called objects. These objects are extracted from geographic databases and carry data and operations. One of the most suitable options for implementing interoperable objects or components [25] that need to share both code and data across a heterogeneous network is the use the programming language Java [26].

There are two options for implementing Java objects from ontologies. First the objects can be generated from ontologies specified in an ontology editor such as Ontolingua, which has the ability to create CORBA IDL headers from ontology components. A CORBA IDL header is a skeleton of ontologies and its components, which should be complemented by implementation code written in Java. The second option is to generate Java interfaces from an ontology editor that has this capability. Since Ontolingua is not able to do this, it is necessary to develop an ontology editor to do this kind of work or implement an extension to Ontolingua enabling the translation of ontologies into Java interfaces.

Ontolingua offers a flexible manipulation of ontologies that allows inclusion of entire ontologies or of only some references. A CORBA Interface Definition Language (IDL) [27] header file can be generated by Ontolingua. A CORBA IDL is a skeleton of ontologies and their components that should be complemented by implementation code written in Java [28]. The resulting components are geographic objects. The model for implementing the geographic many-roled objects is suggested in [29]. These objects can be extracted from an object-oriented database that implements roles, such as DOOR [30].

5. Conclusions and future work

This paper proposed the creation of software components from ontologies as a way to share knowledge and information in GIS. These software components were derived from ontologies using an object-oriented mapping. The translation of a geospatial ontology into an active information system component leads to ontology-driven geographic information systems (ODGIS).

The mapping of multiple ontologies to the system classes was done through object-oriented techniques using multiple inheritance. This kind of mapping allows partial integration of information where completeness was impossible. Objects with multiple roles helped the user to avoid the data overload. The extraction of certain roles of selected classes leads to integration of information as Wiederhold [31] pointed out, because we did not combine sources, but selected results derived from them. This paper has demonstrated that ODGIS can play an important role in enabling information and knowledge sharing.

A further study should examine how to incorporate approaches that allow composition of preexisting independently developed ontologies as [32], where a context algebra is proposed to compose diverse ontologies, and the OBSERVER system [33-35], where ontologies are matched using synonym, hyponym and hypernym.

This is an early work and the translation of sophisticated structures like ontologies into classes is a delicate operation. Although systems like Ontolingua [17] suggests that such a mapping can be the answer, further study is necessary to clarify this matter.

6. Acknowledgments

This work was partially supported by the National Science Foundation under NSF grant numbers SBR-9700465 and IIS-9970123. Max Egenhofer's research is further supported by NSF grants IRI-9613646, BDI-9723873, and EIA-9876707; by the National Imagery and Mapping Agency under grant number NMA202-97-1-1023; by the National Institute of Environmental Health Sciences, NIH, under grant number 1 R 01 ES09816-01; by a contract with Lockheed Martin; and by Bangor Hydro-Electric Co.

7. References

[1] Hayes, P., "The Naive Physics Manifesto," in Expert Systems in the Microelectronic Age, D. Michie, Ed. Edinburgh, Scotland: Edinburgh University Press, 1978, pp. 242-270.

[2] Worboys, M. F., GIS - A Computing Perspective. Bristol, PA: Taylor & Francis Inc., 1995.

[3] McKee, L. and Buehler, K., "The Open GIS Guide," . Wayland, MA: Open GIS Consortium, Inc, 1996.

[4] Vckovski, A., "International Journal of Geographical Information Science - Special Issue: Interoperability in GIS," , vol. 12 - Number 4, pp. 297-298, 1998.

[5] Goodchild, M., Egenhofer, M., Fegeas, R., and Kottman, C., Interoperating Geographic Information Systems. Norwel, MA: Kluwer Academic Publishers, 1999.

[6] Vckovski, A., Brassel, K., and Schek, H.-J., "Preface," Interoperating Geographic Information Systems-Second International Conference, INTEROP'99, Zurich, Switzerland, Zurich, Switzerland, 1999.

[7] Batini, C., Lenzerini, M., and Navathe, S. B., "A Comparative Analysis of Methodologies for Database Schema Integration," ACM Computing Surveys, vol. 18, pp. 323-364, 1986.

[8] Wiederhold, G., "Mediation to Deal with Heterogeneous Data Sources," in Interoperating Geographic Information Systems - Second International Conference, INTEROP'99, Zurich, Switzerland, vol. 1580, Lecture Notes in Computer Science, A. Vckovski, K. Brassel, and H.-J. Schek, Eds. Berlin: Springer-Verlag, 1999, pp. 1-16.

[9] Guarino, N., "Formal Ontology and Information Systems," in Formal Ontology in Information Systems, N. Guarino, Ed. Amsterdan, Netherlands: IOS Press, 1998, pp. 3-15.

[10] Frank, A. U., "Spatial Ontology: A Geographical Point of View," in Spatial and Temporal Reasoning, O. Stock, Ed. Dordrecht, The Netherlands: Kluwer Academic Publishers, 1997, pp. 135-153.

[11] Smith, B. and Mark, D. M., "Ontology and Geographic Kinds," International Symposium on Spatial Data Handling, Vancouver, Canada, 1998.

[12] Gruber, T., "The Role of Common Ontology in Achieving Sharable, Reusable Knowledge Bases," Principles of Knowledge Representation and Reasoning, Cambridge, MA, 1991.

[13] Nunes, J., "Geographic Space as a Set of Concrete Geographical Entities," in Cognitive and Linguistic Aspects of Geographic Space, D. Mark and A. Frank, Eds. Norwell, MA: Kluwer Academic Publishers, 1991, pp. 9-33.

[14] Smith, B., "An Introduction to Ontology," in The Ontology of Fields, D. Peuquet, B. Smith, and B. Brogaard, Eds. Bar Harbor, ME: NCGIA, 1998, pp. 10-14.

[15] Neches, R., Fikes, R., Finin, T., Gruber, T., Patil, R., Senator, T., and Swartout, W. R., "Enabling Technology for Knowledge Sharing," AI Magazine, vol. 12, pp. 13-36, 1991.

[16] Gruber, T., "A Translation Approach to Portable Ontology Specifications," Knowledge Systems Laboratory - Stanford University, Stanford, CA, Technical Report KSL 92-71, 1992.

[17] Farquhar, A., Fikes, R., and Rice, J., "The Ontolingua Server: a Tool for Collaborative Ontology Construction," Knowledge Systems Laboratory - Stanford University, Stanford, CA, Technical Report KSL 96-26, 1996.

[18] Smith, B., "On Drawing Lines on a Map," Spatial Information Theory - COSIT 95, Vienna, Austria, 1995.

[19] Kuhn, W., "Metaphors Create Theories for Users," in Spatial Information Theory, vol. 716, Lectures Notes in Computer Science, A. Frank and I. Campari, Eds., 1993.

[20] Kuhn, W., "Are Displays Maps or Views?," Auto-Carto 10, Baltimore, 1991.

[21] Frank, A., "Computer Assisted Cartography—Graphics or Geometry," in Journal of Surveying Engineering, vol. 110, 1984, pp. 159-168.

[22] Bishr, Y. A., Pundt, H., Kuhn, W., and Rdwan, M., "Probing the Concepts of Information Communities - A First Step Toward Semantic Interoperability," in Interoperating Geographic Information Systems, M. Goodchild, M. Egenhofer, R. Fegeas, and C. Kottman, Eds. Norwell, MA: Kluwer, 1999, pp. 165-180.

[23] Fonseca, F. T. and Egenhofer, M. J., "Ontology-Driven Geographic Information Systems," 7th ACM Symposium on Advances in Geographic Information Systems, Kansas City, USA, 1999.

[24] Meyer, B., Object-Oriented Software Construction. 1988, New York, NY: Prentice Hall.

[25] Betz, M., "Interoperable Objects," in Dr. Dobb's Journal, vol. 4, 1994, pp. 22-26.

[26] Lewandowsky, S. M., "Frameworks for Component-Based Client/Server Computing," ACM Computing Surveys, vol. 30, pp. 4-27, 1998.

[27] Orfali, R. and Harkey, D., Client/Server Programming with Java and CORBA, Second Edition. New York: John Wiley & Sons, 1998.

[28] Gosling, J. and McGilton, H., "The Java Language Environment: a White Paper," Sun Microsystems, Mountain View, CA 1995.

[29] Pernici, B., "Objects with Roles," IEEE/ACM Conference on Office Information Systems, 1990.

[30] Wong, R. K., Chau, H. L., and Lochovsky, F. H., "Dynamic Knowledge Representation in DOOR," IEEE Knowledge and Data Engineering Exchange Workshop - KDEX 97, Newport Beach, CA, 1997.

[31] Wiederhold, G., "Foreword: Intelligent Integration of Information," in Intelligent Integration of Information, G. Wiederhold, Ed. Norwell, MA: Kluwer Academic Publishers, 1996.

[32] Wiederhold, G. and Jannink, J., "Composing Diverse Ontologies," 8th Working Conference on Database Semantics (DS-8), Rotorua, New Zealand, 1999.

[33] Mena, E., Kashyap, V., Sheth, A., and Illarramendi, A., "OBSERVER: An Approach for Query Processing in Global Information Systems based on Interoperation across Pre-existing Ontologies," First IFCIS International Conference on Cooperative Information Systems (CoopIS'96), Brussels, Belgium, 1996.

[34] Kashyap, V. and Sheth, A., "Semantic Heterogeneity in Global Information System: The Role of Metadata, Context and Ontologies," in Cooperative Information Systems: Current Trends and Directions, M. P. Papazoglou and G. Schlageter, Eds. London: Academic Press, 1996, pp. 139-178.

[35] Mena, E., Kashyap, V., Illarramendi, A., and Sheth, A., "Domain Specific Ontologies for Semantic Information Brokering on the Global Information Infrastructure," in Formal Ontology in Information Systems, N. Guarino, Ed. Amsterdam: IOS Press, 1998, pp. 269-283.