CHAPTER 12

Geo-Ontologies

Frederico Fonseca and Gilberto Câmara

12.1 Introduction

In order to understand how people see the world and how, ultimately, mental conceptualizations of the comprehended geographic features are represented in a computer system, we must develop abstraction paradigms. The result of the abstraction process is a general view of the process from the real object to its computer representation. Different levels of abstraction allow the development of specific tools for the different types of problems at each level. Creating a solid conceptual model is at the foundation of system design practice. Lately, ontologies were brought into the discussion on modeling. For instance, Guarino (1998, p. 10) says that "every (symbolic) information system (IS) has its own ontology, since it ascribes meaning to the symbols used according to a particular view of the world." Wand and Weber (2004, p. v) argue that since theories of ontology are tools that help us describe a specific world (i.e., the target of an IS), "our information systems will only be as good as our ontologies."

The subject of ontology also is an important field of research in geographic information science (Mark 1993; Frank 1997; Smith and Mark 1998; Bittner and Winter 1999; Rodríguez et al. 1999; Bishr and Kuhn 2000; Câmara et al. 2000; Frank 2001; Kuhn 2001; Kavouras et al. 2005). Ontologies have been used as a means of knowledge-sharing among different user communities, thus improving interoperability among different geographic databases. Information can be integrated based primarily on its meaning by integrating ontologies that are linked to sources of geographic information. The use of an ontology, translated into an active, information-system component, leads to Ontology-Driven Information Systems (ODIS) (Guarino 1998) and, in the specific case of Geographic Information Systems (GIS), it leads to what is called Ontology-Driven Geographic Information Systems (ODGIS) (Fonseca and Egenhofer 1999). In GIS, the use of ontologies is diverse. They can be used to deal with aerial images (Câmara et al. 2001) or with urban systems (Fonseca et al. 2000), for instance. Ontologies are theories that use a specific vocabulary to describe entities, classes, properties and functions related to a certain view of the world. They can be a simple taxonomy, a lexicon or a thesaurus, or even a fully axiomatized theory (Gruber 1995; Guarino and Giaretta 1995).

12.2 Ontology and ontologies

The two uses of the term that we need to contrast are: 1) the way the word is used in Philosophy; and 2) the most current use of the term in Artificial Intelligence, Computer Science, and Information Systems. In Philosophy, Ontology is the basic description of entities in the world, the description of what would be the truth, and the term is used with an uppercase O. Guarino (1998) considers the philosophical meaning of ontology to be a particular system of categories that reflects a specific view of the world. Smith (1998) 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 software 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's 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.

Researchers that use the philosophical meaning of ontology are resorting to the theory of Ontology, the ontology methods, the tools and theories developed within the philosophical discipline of Ontology to find the basic constructs of information systems. They are investigating what information systems are as a concept. From their findings they are able to draw the primitives that conceptual models should use if we are to build better information systems. On the other hand, when Guarino is talking about ontology-driven information systems, the ontologies he is referring to are computational ontologies or ontologies of the second kind we discussed above, ontologies in the original Artificial Intelligence (AI) sense. These are real artifacts that explain a domain. Guarino calls them "engineering artifacts."

Indeed, even the more restricted term computational ontology has been used with more than one meaning in the literature. In this section we review some of the meanings and argue that computational ontologies are theories that explain a domain. In this we agree with Guarino and Giaretta (1995), Smith (2003), and Wand and Weber (2004).

Guarino and Giaretta (1995) recommend that we restrain ourselves to the meaning of the term ontology which aims towards a theory instead of the simple specification of particular epistemic states. Analyzing ontology as a theory, they say (p. 30) that "an ontological theory differs from an arbitrary logical theory (or knowledge base) by its semantics, since all its axioms must be true in every possible world of the underlying conceptualization." Here they are trying to clarify a common use of the term ontology. Gruber (1995) gave a definition of ontology as a "specification of a conceptualization" based on Genesereth and Nilsson's (1987) work. One of the interpretations of Gruber's definition that Guarino wants to avoid is that a conceptualization would define a state of affairs. Guarino (1998) uses an example of the relations among a set of blocks on a table. In Gruber's definition, an ontology would specify for instance that block A is over block B and block C is on the side of block A. Guarino says that the problem with this notion of conceptualization is that it refers to common relations in the blocks' world, i.e., extensional relations. These relations depict a specific state of affairs. In this case they are reflecting a specific arrangement of blocks on the table. Guarino thinks we need to address the meaning of these relations instead of the current situation on the table. He says that an ontology should describe intensional relations such as the meaning of above for instance. Guarino summarizes with the definition 'C = <D, W, R>' in which C is a conceptualization, D is a domain, W is a set of relevant state-of-affairs or possible worlds, and R is a set of conceptual relations on the domain space <D, W>. After clarifying what a conceptualization is, he gives a new definition of an ontology:

> *…an ontology is a logical theory accounting for the intended meaning of a formal vocabulary, i.e. its ontological commitment to a particular conceptualization of the world. The intended models of a logical language using such a vocabulary are constrained by its ontological commitment. An ontology indirectly reflects this commitment (and the underlying conceptualization) by approximating these intended models.* (p. 7)

Smith (2003) says that in the current context of research on information sharing an ontology is seen as a dictionary of terms expressed in a canonical syntax. In this use it is implied that ontology is a common vocabulary shared by different IS communities. Smith then gives a definition of an IS (or computational) ontology: "an ontology is a formal theory within which not only definitions but also a supporting framework of axioms is included (perhaps the axioms themselves provide implicit definitions of the terms involved)" (Smith 2003, p. 158).

Wand and Weber (2004) say that although many ontologies restrict themselves to be more a taxonomy than a theory, they still have predictive and explanatory tones. They say "if phenomena are classified correctly according to the theory, humans will be better able to understand and predict the phenomena and thus work more effectively and efficiently with the phenomena" (p. iv). Nevertheless, ontologies as theories of Ontology or as computational ontologies are important models of the world and the activity of ontology engineering has been compared to modeling.

12.3 What's Special about Spatial

"What is special about spatial?" (Anselin 1989; Egenhofer 1993) or what is special about geo-ontologies? A geo-ontology has to provide a description of geographic entities, which can be conceptualized in two different views of the world (Couclelis 1992; Goodchild 1992). The field view considers spatial data to be a set of continuous distributions while the object view conceives the world as occupied by discrete, identifiable entities. Representing geographic entities–either constructed features or natural variation on the surface of the earth–is a complex task. These entities are not merely located in space, they are tied intrinsically to space. They take from space some of its structural characteristics (Smith and Mark 1998). A geo-ontology is different from other ontologies because topology and part-whole relations play a major role in the geographic domain. Geographic objects can be connected or contiguous, scattered or separated, closed or open. They are typically complex and have constituent parts (Smith and Mark 1998). The topological and containment relations between objects have led to the introduction of mereology (Husserl 1970), which describes the relation between parts and wholes. For a review of mereology see Simons (1987) and Casati and Varzi (1999). Smith (1995) introduced mereotopology, which extends the theory of mereology with topological methods.

The development of ontologies of the geographic world (i.e., geo-ontologies) is important to allow the sharing of geographic data among different communities of users. Nevertheless, before we share digital data it is necessary to collect and organize it. Conceptual schemas are built in order to abstract specific parts of the real world and to represent schematically what data should be collected and how they must be organized. In the next sections we review the most recent work on geo-ontologies and geographic data models, in order to gain insight on how the distance between ontologies and conceptual schemas can be shortened.

Spatial databases intend to be a representation of geographic space. But what exactly constitutes geographic space? The most widely accepted conceptual model for geographic information science considers that geographic reality is represented as either fully definable entities (objects) or smooth, continuous spatial variation (fields). The object model represents the world as a surface occupied by discrete, identifiable entities, with a geometrical representation and descriptive attributes. The field model views the geographic reality as a set of spatial distributions over the geographic space. As some authors have already pointed out (Couclelis 1992), the field and object models have an underlying common notion, which is the implicit reliance on Cartesian (or absolute) space as an a priori frame of reference for locating spatial phenomena. In this view, Cartesian space is simply a neutral container within which all physical processes occur. The primitive notion on a Cartesian space is the idea of georeferenced location. Each entity of space is associated to one or more locations on Earth, and spatial relations are derived from the location. The alternative to absolute space is to consider a relative notion of space (Couclelis 1997), constituted through the spatial relations arising among geographic entities. In the framework of relative space, the primitive notion is that of the spatial relation between entities. Spatial interaction models and location-allocation models used in transportation are examples of applications that use the relative notion of space.

Current GIS technology embodies an absolute view of space, since the most common geometric representations available in GIS – such as grids, triangulated irregular networks (TINs) and planar vector maps – are all based on the notion of a georeferenced location. It is therefore not surprising that the notions of objects and fields, as defined in the current GIS literature, can be generalized into a single formal definition.

Nunes (1991) pointed out that the first step in building a next-generation GIS would be the creation of a systematic collection and specification of geographic entities, their properties, and relations. Ontology playing a software specification role was suggested by Gruber (1991). Wiederhold (1994) suggested the use of ontologies as the common point among diverse user communities. 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 (Smith and Mark 1998). Frank (1997) believes that the use of ontologies will contribute to better information systems by avoiding problems such as inconsistencies between ontologies implicitly embedded in GIS, conflicts between the ontological concepts and the implementation, and conflicts between the common-sense ontology of the user and the mathematical concepts in the software. Harvey (1999) warns that bringing fundamental semantic concerns early into the design process is important. Bittner and Winter (1999) say that the usual role of ontologies in modeling spatial uncertainty is to support object extraction processes. Kuhn (1993) asks for spatial information theories that look toward GIS users instead of focusing on implementation issues. Another semantic approach to integrate geographic information is GeoCosm (Ram et al. 2001), a web-based prototype to integrate autonomous distributed heterogeneous geospatial data. They employ a canonical model that integrates diverse conceptual schemas. An ontology is used to help in solving conflicts among information sources.

Fonseca (2001) proposed a framework that uses ontologies as the foundation for the integration of geographic information. By integrating ontologies that are linked to sources of geographic information, Fonseca created a mechanism that allows geographic information to be integrated based primarily on its meaning. Since the integration may occur across different levels, he also created the basic mechanisms for changing the level of detail. The use of an ontology, translated into an information system component, is the basis of Ontology-Driven Geographic Information Systems (ODGIS).

12.4 Geo-Ontologies

A geo-ontology describes entities, semantic relations, and spatial relations:

1. entities can be assigned to locations on the surface of the Earth;
2. semantic relations between these entities include, e.g., hypernymy—relation of class to subclass, hyponymy—relation of subclass to class, mereonomy—part of a whole, and synonymy—same as; and
3. spatial relations between entities (e.g., adjacency, spatial containment, proximity and connectedness).

A geo-ontology also has two basic types of concepts: concepts that correspond to physical phenomena in the real world and concepts that correspond to features of the world that we create to represent social and institutional constructs. We call the first type physical concepts and the second type social concepts. It is important to note that both result from human conventions. As discussed in the literature, although the description of physical features may vary according to cultural and social conventions, they represent variations on the surface of the Earth rather than social conventions per se (Frank and Mark 1991; Mark and Egenhofer 1994; Mark et al. 1999; Smith and Mark 2003).

The physical concepts can be further subdivided into:

- Concepts that are associated with individual geographic objects, each of which has a clearly defined boundary such as qualitative differentiations or spatial discontinuities in the physical world. These are equivalent to the notion of bona fide objects (Smith and Mark 1998). Examples: lake and mountain.
- Concepts that are assumed to be continuous in space (fields). Examples: temperature, slope, pollution and population density.

The social and institutional concepts can be further subdivided into:

- Concepts describing individual objects created by institutional and legal conventions. These are equivalent to the notion of fiat objects or non-naturally demarcated geographical entities of Smith and Mark (1998). Examples: parcel and borough.

- Concepts that are assumed to be continuous over space and represent socially agreed conventions. Examples: social exclusion, infant mortality, homicide rate and human development.

A concept in a geo-ontology is defined by a name, a definition and a set of attributes. A geo-ontology is a set of terms and a set of semantic and spatial relations between terms. The set of semantic relations is created by the semantic components present in the definitions of the terms of a geo-ontology. For instance, the definition of a stream being a flow of water in a channel or bed as a brook, rivulet, or small river and the definition of a creek being a small, often shallow or intermittent tributary to a river would lead to the consequent semantic relation of hyponymy that a creek is a stream.

Another example is that the definition of a basin being a region drained by a single river system and the definition of a valley being an extensive area of land drained or irrigated by a river system would lead to the consequent semantic relation similarity that a basin is similar to a valley.

The set of spatial relations is created by the spatial components present in the definitions of the terms of a geo-ontology. The spatial nature of the terms generates spatial relations between terms in a geo-ontology. For instance, the definition of affluent as a stream or river that flows into a larger one leads to the consequent spatial relation that an affluent is connected to streams.

A second example of spatial relations is in the definition of a valley as an elongated lowland between ranges of mountains, hills or other uplands, often having a river or stream running along the bottom that leads to the consequent spatial relation of mountains being adjacent to valleys.

As the above examples show, a geo-ontology has to take into consideration not only semantic relations such as synonymy, similarity, mereonomy and hyponymy, but also spatial relations such as adjacency, spatial containment and connectedness. Given that both semantic and spatial relations are conceptual components of geo-ontologies, defining these relationships plays a critical role in the integration of geographic information.

References

Anselin, L. 1989. What Is Special About Spatial Data? Alternative Perspectives on Spatial Data Analysis. Santa Barbara, Calif.: NCGIA Technical Report.

Bishr, Y. A. and W. Kuhn. 2000. Ontology-based modelling of geospatial information. Pages 24–27 in *Proceedings of the 3rd AGILE Conference on Geographic Information Science*, Helsinki, Finland, May 2000. Edited by A. Ostman, M. Gould and T. Sarjakoski.

Bittner, T. and S. Winter. 1999. On ontology in image analysis in integrated spatial databases. In *Integrated Spatial Databases: Digital Images and GIS - Lecture Notes in Computer Science 1737*, edited by P. Agouris and A. Stefanidis, 168–191. Berlin: Springer-Verlag.

Câmara, G., M. Egenhofer, F. Fonseca and A.M.V. Monteiro. 2001. What's in an image? Pages 474–488 in *Spatial Information Theory. Foundations of Geographic Information Science*. Edited by D. R. Montello. Proceedings, International Conference COSIT 2001, Morro Bay, Calif., September 19-23, 2001. Lecture Notes in Computer Science Volume 2205/2001. Berlin: Springer.

Câmara, G., A. Monteiro, J. Paiva and R. Souza. 2000. Action-driven ontologies of the geographical space: Beyond the field-object debate. Pages 52–54 in GIScience 2000—Program of the First International Conference on Geographic Information Science, Savannah, Ga., October 28–31, 2000, chaired by M. Egenhofer and D. Mark.

Casati, R. and A. Varzi. 1999. *Parts and Places*. Cambridge, Mass.: MIT Press.

Couclelis, H. 1992. People manipulate objects (but cultivate fields): Beyond the raster-vector debate in GIS. In *Theories and Methods of Spatio-Temporal Reasoning in Geographic Space*, edited by A. U. Frank, I. Campari and U. Formentini, 65–77. Lecture Notes in Computer Science 639, New York: Springer-Verlag.

Couclelis, H. 1997. From cellular automata to urban models: New principles for model development and implementation. *Environment and Planning B: Planning and Design* 24:165–174.

Egenhofer, M. 1993. What's special about spatial?-Database requirements for vehicle navigation in geographic space. *SIGMOD RECORD* 22(2):398–402.

Fonseca, F. 2001. *Ontology-Driven Geographic Information Systems*. Ph.D. Thesis. Orono: University of Maine.

Fonseca, F. and M. Egenhofer. 1999. Ontology-driven geographic information systems. Pages 14–19 in *Proceedings, 7th ACM Symposium on Advances in Geographic Information Systems*, held in Kansas City, Mo., November 1999. Edited by C. B. Medeiros. <http://www.spatial.maine.edu/~fred/fonseca_acmgis.pdf> Accessed 12 February 2007.

Fonseca, F., M. Egenhofer, C. Davis and K. Borges. 2000. Ontologies and knowledge sharing in urban GIS. *Computer, Environment and Urban Systems* 24(3):232–251.

Frank, A. 1997. Spatial ontology. In *Spatial and Temporal Reasoning*, edited by O. Stock, 135–153. Dordrecht, The Netherlands: Kluwer Academic.

Frank, A. 2001. Tiers of ontology and consistency constraints in geographical information systems. *International Journal of Geographical Information Science* 15(7):667–678.

Frank, A. and D. Mark. 1991. Language issues for GIS. In *Geographical Information Systems, Volume 1: Principles*, edited by D. Maguire, M. Goodchild and D. Rhind, 147–163. London: Longman.

Genesereth, M. R. and N. J. Nilsson. 1987. *Logical Foundations of Artificial Intelligence*. Los Altos, Calif.: Morgan Kaufmann.

Goodchild, M. 1992. Geographical data modeling. *Computers and Geosciences* 18(4):401–408.

Gruber, T. 1991. The role of common ontology in achieving sharable, reusable knowledge bases. Pages 601–602 in *Proceedings of the 2nd International Conference on Principles of Knowledge Representation and Reasoning (KR'91)*, Cambridge, Mass., April 22-25, 1991. Edited by J. F. Allen, R. Fikes and E. Sandewall. San Francisco: Morgan Kaufmann Publishers.

Gruber, T. R. 1995. Toward principles for the design of ontologies used for knowledge sharing. *International Journal of Human Computer Studies* 43(5/6):907–928.

Guarino, N. 1998. Formal ontology and information systems. In *Formal Ontology in Information Systems*, edited by N. Guarino, 3–15. Amsterdam, The Netherlands: IOS Press.

Guarino, N. and P. Giaretta. 1995. Ontologies and knowledge bases: Towards a terminological clarification. In *Towards Very Large Knowledge Bases: Knowledge Building & Knowledge Sharing*, edited by N.J.I. Mars, 25–32. Amsterdam, The Netherlands: IOS Press.

Harvey, F. 1999. Designing for interoperability: Overcoming semantic differences. In *Interoperating Geographic Information Systems*, edited by M. Goodchild, M. Egenhofer, R. Fegeas and C. Kottman, 85–98. Norwell, Mass.: Kluwer Academic.

Husserl, E. 1970. *Logical Investigations*. Translated by J. N. Findlay from the second German edition of *Logische Untersuchungen*. London: Routledge and Kegan Paul - Humanities Press.

Kavouras, M., M. Kokla and E. Tomai. 2005. Comparing categories among geographic ontologies. *Computers and Geosciences* 31(2):145–154.

Kuhn, W. 1993. Metaphors create theories for users. In *Spatial Information Theory*, edited by A. Frank and I. Campari, 366–376. Lectures Notes in Computer Science 716. Berlin: Springer-Verlag.

Kuhn, W. 2001. Ontologies in support of activities in geographical space. *International Journal of Geographical Information Science* 15(7):613–631.

Mark, D. 1993. Toward a theoretical framework for geographic entity types. In *Spatial Information Theory*, edited by A. Frank and I. Campari, 270–283. Lectures Notes in Computer Science 716. Berlin: Springer-Verlag.

Mark, D. and M. Egenhofer. 1994. Calibrating the meanings of spatial predicates from natural language: Line-region relations. Pages 538–553 in *Sixth International Symposium on Spatial Data Handling (SDH '94)*, Edinburgh, Scotland, September 1994. Edited by T. Waugh and R. Healey. <http://www.spatial.maine.edu/~max/Calibration.pdf> Accessed 12 February 2007.

Mark, D., B. Smith and B. Tversky. 1999. Ontology and geographic objects: An empirical study of cognitive category. Pages 283–298 in *Spatial Information Theory-Cognitive and Computational Foundations of Geographic Information Science*. Edited by C. Freksa and D. Mark. Proceedings, International Conference COSIT 1999, Stade, Germany. Lecture Notes in Computer Science Volume 1661. Berlin: Springer-Verlag.

Nunes, J. 1991. Geographic space as a set of concrete geographical entities. In *Cognitive and Linguistic Aspects of Geographic Space*, edited by D. Mark and A. Frank, 9–33. Norwell, Mass.: Kluwer Academic.

Ram, S., V. Khatri, L. Zhang and D. D. Zeng. 2002. GeoCosm: A semantics-based approach for information integration of geospatial data. Pages 152–165 in *Revised Papers from the HUMACS, DASWIS, ECOMO, and DAMA on ER 2001 Workshops*. Edited by H. Arisawa, Y. Kambayashi, V. Kumar, H. C. Mayr and I. Hunt. 21st International Conference on Conceptual Modeling, Yokohama, Japan, November 27-30, 2001. Lecture Notes in Computer Science; Volume 2465. London: Springer-Verlag.

Rodríguez, A., M. Egenhofer and R. Rugg. 1999. Assessing semantic similarity among geospatial feature class definitions. Pages 1–16 in *Interoperating Geographic Information Systems*. Edited by A. Vckovski, K. Brassel and H.-J. Schek. Second International Conference, INTEROP'99, Zurich, Switzerland, March 10-12, 1999. Lecture Notes in Computer Science 1580. Berlin: Springer.

Simons, P. 1987. *Parts: An Essay in Ontology*. Oxford: Clarendon Press.

Smith, B. 1995. On drawing lines on a map. Pages 475–484 in *Spatial Information Theory—a Theoretical Basis for GIS*. Edited by A. Frank and W. Kuhn. Proceedings, International Conference Cosit'95, Semmering, Austria, September 21-23, 1995. Lecture Notes in Computer Science 988. Berlin: Springer Verlag.

Smith, B. 1998. An introduction to ontology. In *The Ontology of Fields*, edited by D. Peuquet, B. Smith and B. Brogaard, 10–14. Santa Barbara, Calif.: National Center for Geographic Information and Analysis.

Smith, B. 2003. Ontology. In *The Blackwell Guide to the Philosophy of Computing and Information*, edited by L. Floridi, 155–166. Malden, Mass.: Blackwell.

Smith, B. and D. Mark. 1998. Ontology and geographic kinds. Pages 308–320 in *Proceedings of the Eighth International Symposium on Spatial Data Handling*, held in Vancouver, British Columbia, Canada. Edited by T. K. Poiker and N. Chrisman. Burnaby, British Columbia: International Geographical Union.

———. 2003. Do mountains exist? Towards an ontology of landforms. *Environment and Planning B: Planning and Design* 30(3):411–427.

Wand, Y. and R. Weber. 2004. Reflection: Ontology in information systems. *Journal of Database Management* 15(2):iii–vi.

Wiederhold, G. 1994. Interoperation, mediation and ontologies. Pages 33–48 in *International Symposium on Fifth Generation Computer Systems (FGCS94)*; Workshop on Heterogeneous Cooperative Knowledge-Bases, Tokyo, Japan, December 1994. Edited by K. Yokota. Tokyo: ICOT>