Algebraic Structures for Spatial Ontologies

Gilberto Câmara[1] , Frederico Fonseca[2], Antonio Miguel Monteiro[1]

[1]Image Processing Division, National Institute for Space Research, Av. dos Astronautas, 1758 - 12227-001 - São José dos Campos , SP, Brazil

{gilberto, miguel}@dpi.inpe.br

[2] School of Information Sciences and Technology, Pennsylvania State University
University Park, PA 16802, USA

fredfonseca@ist.psu.edu

Introduction

The idea of *ontology* has proven to be an important field of research in geographical information science (Smith and Mark 1998), and it is expected that the use of ontologies will improve interoperability among different geographical databases (Fonseca and Egenhofer 1999). Further advances in the field require abstract specifications of spatial ontologies, to derive useful properties that can be used on the implementation of interoperability frameworks based on ontologies. This paper describes on-going work towards achieving such abstract descriptions, where the aim is to propose algebraic structures that support the concepts of *equivalence* and *transformations* between spatial ontologies.

An Algebraic Framework for Interoperability

We propose a set of definitions that aim can be used as a basis for algebraic formulation for *simple spatial ontologies*, with an emphasis on interoperability of geographical data. The basic assumptions are:

(a) There are no polysemous words in the ontology (i.e., all words have a single meaning);

(b) There is only one possible connection between two concepts on this simple ontology;

(c) All concepts are *spatial concepts* (which refer to phenomena which are spatially referenced).

Definition 1. **A concept** is a pair c = [d, V], where d is a string of characters (containing a definition of the concept) and V is a set of values that the concept can be associated with (the **domain** of the concept).

Definition 2. **A connection** is a characteristic relation between two spatial concepts that is captured on the ontology (including the empty relation). In our model, the set of connections is Ω = {*Synonymy, Hyponymy (IS_A), Mereonomy (PART_OF),* $\varnothing$}. The concept of "part-of" is spatially aware and means spatial containment.

Definition 3. An ontology is a relation O = [S, Π], where S = {s_1, …, s_n} is a set of *spatial concepts* and Π : S x S → Ω is a mapping from an ordered pair of spatial concepts to the set of connections Ω. Note that $\Pi(s_i, s_j) \neq \Pi(s_j, s_i)$ in general.

An example would be an ontology consisting of the set of concepts S = {borough, parcel, piece of land, park}, and the mapping Π defined as:

- Π(parcel, piece of land) = Synonymy.
- Π(park, parcel) = IS_A.
- Π(parcel, borough) = PART_OF.

Definition 4. The merging operation ⊕ between maps Π_1 and Π_2 is such that $\Pi_1 \oplus \Pi_2 = \Pi_3$ where Π_3 is defined by:

- $\forall s_i, s_j \in S_1 \land s_i, s_j \notin S_2, \Pi_3(s_i, s_j) = \Pi_1(s_i, s_j)$

- $\forall s_i, s_j \notin S_1 \land s_i, s_j \in S_2, \Pi_3(s_i, s_j) = \Pi_2(s_i, s_j)$

- $\forall s_i \in S_1 \land s_j \in S_2, s_i, s_j \notin S_1 \cap S_2, \Pi_3(s_i, s_j) = \Pi_{aux}(s_i, s_j)$, where Π_{aux} must be provided to link two concepts that were not related before the operation.

  $\forall s_i \in S_1 \land s_j \in S_2 / s_i, s_j \in S_1 \cap S_2,$

- if $\Pi_2(s_i, s_j) = \Pi_1(s_i, s_j), \Pi_3(s_i, s_j) = \Pi_1(s_i, s_j)$

  else $\Pi_3(s_i, s_j) = f_{merge}(\Pi_1(s_i, s_j), \Pi_2(s_i, s_j))$

The merging operation thus requires two auxiliary operations: (a) linking concepts from the two ontologies; (b) resolving ambiguities between concepts whose relation is defined differently in each ontology.

Definition 5. **The ontology composition operation** is a mapping Θ between two ontologies $O_1 \Theta O_2 = O_3$, where $O_3 = [S_3, \Pi_3]$ is defined by the set $S_3 = S_1 \cup S_2$ and by the mapping $\Pi_3 = \Pi_1 \oplus \Pi_2$.

Proposition 1. The composition operation is associative and commutative. (The proof of this proposition is straightforward from Definition 5).

Proposition 2. **There exists an identity ontology I**, such that $O_1 \Theta I = O_1$. This ontology is the empty set of concepts.

Proposition 3. **The set of ontologies constitute an Abelian monoid** under the composition operation Θ and the identity ontology I.

Based on the above definitions and propositions, we can now propose some important properties of ontologies, which are relevant to interoperability considerations.

Definition 6. **A mapping Ψ between two ontologies O_1 and O_2** is such that $\forall s_i \in S_1, \exists s_j \in S_2 / \Psi(s_i, s_j) \in \Omega$. For simplicity, we denote such operation by **$O_2 = \Psi(O_1)$**. Thus, establishing a relation between ontologies requires a mapping between their concepts.

Definition 7. **Two ontologies O_1 and O_2 are equivalent** if there is a mapping Ψ such that $\forall s_i \in S_1, \exists\, s_j \in S_2\, /\Psi(s_i, s_j) = Synonymy$. Strong equivalence therefore implies synonymy for all concepts.

Definition 8. **An ontology O_2 is a generalization of an ontology O_1** when $\forall s_i \in S_1, \exists\, s_j \in S_2\, /\Psi(s_i, s_j) = Synonymy \vee \Psi(s_i, s_j) = IS\_A$. In other words, for every concept in the first ontology, there is a synonymous or a more general concept in the second ontology. For example, consider the ontology O_1, whose set of spatial concepts consists of land cover types $S_1 = \{$multi-layered forest, forest with shrubs, grassland with sparse trees, grassland with sparse shrubs$\}$, taken from the FAO land cover classification system (Gregorio and Jansen 1998), and an ontology $S_2 = \{$forest, grassland$\}$. The mapping Ψ between S_1 and S_2 is

- Ψ(multi-layered forest, forest) = IS_A.

- Ψ(forest with shrubs, forest) = IS_A.

- Ψ(grassland with sparse trees, grassland) = IS_A.

- Ψ(grassland with sparse shrubs, grassland) = IS_A.

Proposition 4. Two ontologies O_1 and O_2 are interoperable if O_1 and O_2 are equivalent or if O_2 is a generalization of O_1. This proposed definition of interoperability is more general than earlier work on spatial ontologies, which considers similarity between words, and makes the implicit hypothesis that similar concepts in different ontologies should be synonymous (Rodriguez, Egenhofer, and Rugg 1999).

Conclusions

This paper presents an abstract specification of spatial ontologies, which allows formal definitions of the concepts of *transformation, equivalence, generalization and interoperability* between ontologies. This abstract specification is based on a simplified spatial ontology where all concepts are spatially-referenced, and where the possible connections between concepts are *synonymy*, IS_A (*hyponymy*) and PART_OF (*mereonomy*). Further work on the topic will investigate the following issues:

- Development of computational tools based on the specifications.

- Extensions of the interoperability definition to include the mereonomy conditions.

- Extension of the definition of spatial ontologies to include properties of spatial concepts, in the sense that "an owner is a property of a parcel".

Acknowledgements

References

Fonseca, Frederico, and Max Egenhofer. 1999. Ontology-Driven Geographic Information Systems. Paper read at 7th ACM Symposium on Advances in Geographic Information Systems, at Kansas City, MO.C. Bauzer-Medeiros (ed), ACM Press, N.Y., 1999.

Gregorio, A., and L. Jansen. 1998. Land Cover Classification System: Classification Concepts and User Manual. Food and Agriculture Organization of the United Nations (FAO/UN).

Rodriguez, Andrea, Max Egenhofer, and R. Rugg. 1999. Assessing Semantic Similarity Among Geospatial Feature Class Definitions. In *Interoperating Geographic Information Systems - Second International Conference, INTEROP'99, Zurich, Switzerland*, edited by A. Vckovski, K. Brassel and H.-J. Schek. Berlin: Springer-Verlag.

Smith, Barry, and David Mark. 1998. Ontology and Geographic Kinds. Paper read at International Symposium on Spatial Data Handling, at Vancouver, Canada.T. Puecker and N. Chrisman (ed), 1998.