M. Andrea Rodríguez
Isabel F. Cruz
Max J. Egenhofer
Sergei Levashkin (Eds.)

LNCS 3799

GeoSpatial Semantics

First International Conference, GeoS 2005
Mexico City, Mexico, November 2005
Proceedings

Springer

Lecture Notes in Computer Science 3799

M. Andrea Rodríguez   Isabel F. Cruz
Max J. Egenhofer   Sergei Levashkin (Eds.)

GeoSpatial Semantics

First International Conference, GeoS 2005
Mexico City, Mexico, November 29-30, 2005
Proceedings

Springer

Volume Editors

M. Andrea Rodríguez
Universidad de Concepción, Departamento de Ingeniería y Ciencias de Computación
Edmundo Larenas 215, Concepción, Chile
E-mail: andrea@udec.cl

Isabel F. Cruz
University of Illinois at Chicago, Department of Computer Science
851 S. Morgan St. (M/C 152), Chicago, IL 60607, USA
E-mail: ifc@cs.uic.edu

Max J. Egenhofer
University of Maine, Department of Spatial Information Science and Engineering
348B Boardman Hall, Orono, ME 04469-5711, USA
E-mail: max@spatial.maine.edu

Sergei Levashkin
UPALMZ, Centro de Investigación en Computación
Building 07738, Mexico City, Mexico
E-mail: sergei@cic.ipn.mx

Library of Congress Control Number: 2005935535

CR Subject Classification (1998): H.2, H.2.8, H.3, I.2, H.4, C.2

ISSN 0302-9743
ISBN-10 3-540-30288-3 Springer Berlin Heidelberg New York
ISBN-13 978-3-540-30288-9 Springer Berlin Heidelberg New York

Springer is a part of Springer Science+Business Media

springeronline.com

Printed in Germany

Typesetting: Camera-ready by author, data conversion by Scientific Publishing Services, Chennai, India
Printed on acid-free paper SPIN: 11586180 06/3142 5 4 3 2 1 0

Preface

GeoS 2005 was the 1st International Conference on Geospatial Semantics. It was held in Mexico City, November 29 and 30, 2005.

Within the domain of geographic information science (GIS), semantics has become one of the most prominent research themes over the last few years. Such concepts as ontology-driven geographic information systems and the geospatial Semantic Web have fuelled a plethora of research in such areas as geo-ontologies and semantic similarity. These topics complement the traditional focus in GIS research, which has dealt primarily with geometric entities, their spatial relations, and efficient data structures. Geospatial semantics are expected to play an increasingly important role for next-generation spatial databases and geographic information systems, as well as for specialized geospatial Web services.

GeoS 2005 was organized in order to provide a forum for the exchange of state-of-the-art research results in the areas of modeling and processing of geospatial semantics. Of particular interest were contributions that addressed theories for geospatial semantic information; formal representations for geospatial data; models and languages for geo-ontologies; alignment and integration of geo-ontologies; integration of semantics into spatial query processing; similarity comparisons of spatial datasets; ontology-based spatial information retrieval; ontology-driven GIS; geospatial Semantic Web; and multicultural aspects of spatial knowledge.

This volume contains 19 papers, which were selected from among 42 submissions received in response to the Call for Papers. Each submission was reviewed by three or four Program Committee members and 15 long and 4 short papers were chosen for presentation. Authors of papers included in this volume come from 11 different countries, highlighting the breadth of the international research community that focuses its attention on geospatial semantics. The program was rounded off with an invited keynote by Jerry Hobbs, and poster presentations.

We are indebted to many people who made this event happen. The members of the Program Committee offered their help with reviewing submissions. Our thanks go also to Miguel Torres, Marco Moreno, Rolando Quintero, and Giovanni Guzmán, who formed the Local Organizing Committee and took care of all the logistics. The Centro de Investigación en Computación, Mexico City, Mexico, was the local host and co-sponsored GeoS 2005. Finally, we would like to thank all the authors who submitted papers to GeoS 2005.

November 2005

M. Andrea Rodríguez
Isabel F. Cruz
Max J. Egenhofer
Sergei Levashkin

Organization

Organizing Committee

General Chair	Sergei Levashkin Centro de Investigación en Computación, Mexico City, Mexico
Program Chairs	M. Andrea Rodríguez Universidad de Concepción, Chile Isabel F. Cruz University of Illinois at Chicago, USA Max J. Egenhofer University of Maine, USA
Local Organization	Centro de Investigación en Computación, Mexico City, Mexico Miguel Torres (Chair) Marco Moreno Rolando Quintero Giovanni Gúzman

GeoS 2005 Program Committee

Naveen Ashish	NASA Ames Research Center, USA
Brandon Bennett	University of Leeds, UK
Stefano Borgo	Laboratory for Applied Ontology, Italy
Gilberto Câmara	Instituto National de Pesquisas Espaciais, Brazil
Clodoveu Davis Jr.	Pontificia Universidade Católica de Minas Gerais, Brazil
Frederico Fonseca	Pennsylvania State University, USA
Andrew U. Frank	Technical University Vienna, Austria
Christian Freksa	University of Bremen, Germany
Mark N. Gahegan	Pennsylvania State University, USA
Adolfo Guzmán-Arenas	Instituto Politécnico National, Mexico
Stephen C. Hirtle	University of Pittsburgh, USA
Kathleen Hornsby	University of Maine, USA
Christopher B. Jones	Cardiff University, UK
Marinos Kavouras	National Technical University of Athens, Greece
Craig Knoblock	University of Southern California, USA

Werner Kuhn	University of Muenster, Germany
David M. Mark	University at Buffalo, USA
Marco Painho	ISEGI, Universidade Nova de Lisboa, Portugal
Sudha Ram	University of Arizona, USA
Jose Ruiz-Shulcloper	Centro de Aplicaciones de Tecnologías Avanzadas, Cuba
Shashi Shekhar	University of Minnesota, USA
Amit Sheth	University of Georgia, USA
Barry Smith	Saarland University, Germany, and University at Buffalo, USA
Stefano Spaccapietra	Swiss Federal Institute of Technology, Lausanne, Switzerland
Stephan Winter	University of Melbourne, Australia
Michael F. Worboys	University of Maine, USA

GeoS 2005 Additional Reviewers

B. Aleman-Meza	University of Georgia, USA
Yao-Yi Chiang	University of Southern California, USA
Scott Farrar	University of Bremen, Germany
Dan Goldberg	University of Southern California, USA
Margarita Kokla	National Technical University of Athens, Greece
Cedric du Mouza	CNAM Paris, France
Menakshi Nagaraja	University of Georgia, USA
Konstantinos Nedas	University of Maine, USA
Matthew Perry	University of Georgia, USA
Laurent Prevot	Laboratory for Applied Ontology, Italy
Mohammad Reza Kolahdouzan	University of Southern California, USA
Snehal Thakkar	University of Southern California, USA
Christelle Vangenot	EPFL Lausanne, Switzerland
Huiyong Xio	University of Illinois at Chicago, USA

Sponsoring Institutions

Instituto Politécnico Nacional (IPN), COFAA-IPN, CGPI-IPN, Mexico
Centro de Investigación en Computación (CIC), Mexico
Consejo Nacional de Ciencia y Tecnología (CONACYT), Mexico

Table of Contents

Theories for the Semantics of Geospatial Information

Formal Representations for Geospatial Data

Similarity Comparison of Spatial Data Sets

Ontology-Based Spatial Information Retrieval

Geospatial Semantic Web

Short Papers

Comparing Representations of Geographic Knowledge Expressed as Conceptual Graphs*

Athanasios Karalopoulos, Margarita Kokla, and Marinos Kavouras

National Technical University of Athens,
15780, Zografos Campus, Athens, Greece
akaralop@mail.ntua.gr, {mkokla, mkov}@survey.ntua.gr

Abstract. Conceptual Graphs are a very powerful knowledge and meaning representation formalism grounded on deep philosophical, linguistic and object oriented principles [1], [2]. Concerning geographic knowledge representation and matching, the study and analysis of geographic concept definitions plays an important role in deriving systematic knowledge about concepts and comparing geographic categories in order to identify similarities and heterogeneities [4]. Based on the proposed algorithm for the representation of geographic knowledge using conceptual graphs, we also present a method that takes into consideration the special structure of conceptual graphs and produces an output that shows how much similar two geographic concepts are and hence which concept is semantically closer to another. For producing the conceptual graph representation of any geographic concept definition we follow two steps, tagging and parsing, while for measuring the similarity between two geographic ontologies we apply proper modifications to the Dice coefficient that is mainly used for comparing binary structures.

1 Introduction

Conceptual Graphs are a powerful knowledge and meaning representation formalism grounded on deep philosophical, linguistic and object-oriented principles [1], [2]. They provide extensible means of capturing and representing real-world knowledge. Fundamental studies about Conceptual Graphs and some of their applications in the field of Knowledge Representation are found among others in [3].

Concerning geographic knowledge representation, the study and analysis of geographic concept definitions plays an important role in the attempt to derive systematic knowledge about concepts and compare geographic categories in order to identify semantic similarities and heterogeneities [4]. Therefore, the exploitation of effective methods for the representation of geographic definitions forms the basis of the research for analyzing geographic concepts in order to structure their meaning and extract semantic information.

The purpose of the present research is to develop an algorithm for the representation of geographic knowledge using conceptual graphs and then, based on the

* This work extends the use of conceptual graphs in geographic knowledge representation as first introduced in [18]. It also addresses the issue of comparison.

M.A. Rodríguez et al. (Eds.): GeoS 2005, LNCS 3799, pp. 1 – 14, 2005.

proposed methodology and the special features and structures of conceptual graphs, to describe a well-defined process for comparing two geographic concept definitions in order to quantitatively measure their semantic similarity. The comparison process takes into consideration the structure of the corresponding conceptual graphs and produces an output that shows how much similar two geographic concepts are and hence which concept is semantically closer to another.

By introducing an algorithm that takes a geographic concept definition as input and produces the corresponding conceptual graph representation, we achieve to break many limitations and obstacles in the extraction of semantic information from definitions of geographic concepts. Furthermore, we provide alternative deterministic means of facilitating semantic interoperability since the similarity between geographic ontologies depends on specific results of the introduced method for comparing geographic ontologies.

2 Related Work

During the last years, research has been done in order to represent and extract information about geographic concepts. Approaches on geographic knowledge representation include methodologies that are based on analyzing geographic concept definitions and finding effective representations. These can be found among others in [5] and [6].

Conceptual Graphs are a diagrammatic and expressive way of knowledge representation that was firstly introduced for the representation of contents of natural language texts. According to the conceptual graph theory [7], a conceptual graph is a network of concept nodes and relation nodes. The concept nodes represent entities, attributes, or events (actions) while the relation nodes identify the kind of relationship between two concept nodes.

Conceptual Graphs are formally defined by an abstract syntax that is independent of any notation, but the formalism can be represented in either graphical or character-based notations. In the graphical notation, concepts are represented by rectangles, conceptual relations by circles and the arcs that link the relations to the concepts are represented by arrows. The linear form is more compact than the graphical and it uses square brackets instead of boxes and parentheses instead of circles.

Research into establishing comparison methods for similarity measurement between two conceptual graphs is included in [8] and [9]. The main goal of the proposed approaches is to determine whether a query graph is completely contained in any given conceptual graph.

On the other hand, in many text-oriented applications, comparison methods for text representations are proposed and implemented. For instance, in [10] different types of coefficients are introduced for similarity measurement of various data structures and representations. Among them, the Jaccard coefficient, the Cosine coefficient and the Dice coefficient are mainly used for comparing binary structures not only because their results are widely accepted, but also because they are very simple.

Our algorithm for similarity measurement is based on the Dice coefficient, which is calculated using the following formula:

$$S_{D1, D2} = 2\,C\,(D_{1,2}) \,/\, (C\,(D_1) + C\,(D_2)) \quad (1)$$

$C\,(D_{1,2})$ is the number of terms that the two representations (D_1 and D_2) have in common and $C\,(D_1)$, $C\,(D_2)$ is the total number of terms in D_1, D_2 respectively. Its simplicity and normalization are the main reasons for taking it as the basis for our proposed algorithm.

After adopting proper modifications to the above formula due to the special structure and content of conceptual graphs representing geographic concept definitions, we propose a comparison methodology that measures similarity quantitatively and can be used as a matching criterion for similarity measurement between two geographic ontologies.

3 Unfolding Concept Definitions

Every geographic concept definition is usually given by a few sentences that contain two types of information: the *genus* and the *differentia*. The *genus*, or hypernym, specifies the class in which the concept is subsumed and contains information that is frequently used for concept taxonomy construction. On the other hand, the *differentia* specifies how different that concept is from the other concepts in the same class. It is a set of attributive adjectives and prepositional phrases that differentiates words with the same genus. It can also provide the purpose, the location, the look and many other aspects of general knowledge through the existence of one or more sub-clauses, each one giving a different kind of general information.

For example, Table 1 shows the genus and the differentia of the definition: *A Canal is a long and narrow strip of water made for boats and irrigation.* This definition of the concept *Canal* appears in the lexical database WordNet [11].

Table 1. Genus and Differentia of the geographic concept *Canal*

Genus	*Strip*
Differentia	*Long, narrow* (attributive adjectives) *Of water* (prepositional phrase) *Made for boats and irrigation* (sub-clause)

Moreover, we consider that every definition of a geographic concept consists of two parts: the *main* and the *secondary* part. The *main* part of the definition is the clause that contains the genus, its attributive adjectives and the prepositional phrases describing the genus, while the *secondary* part contains the given sub-clauses, which further describe the geographic concept.

The main part consists of the *determinant* section, which follows the general form [{article}+{concept name}+{is}], and the *attributes* section. The *attributes* section is the descriptive clause of the *main* part that contains the genus, the attributive adjectives and the prepositional phrases. The *attributes* section has the general form: [{attributive adjective}*+{genus}+{prepositional phrase}*], where the asterisk declares one-or-many. Table 2 shows the above parts in the definitions of the concept *Canal*.

Table 2. *Canal* Definition' s main and secondary parts

Main part	Determinant section	Attributes section
	A Canal is	*A long and narrow strip of water*
Secondary part	*Made for boats and irrigation* (sub-clause)	

The *secondary* part of a definition contains one or more clauses that provide a particular kind of information (purpose, location, etc.). Each sentence in the secondary part contains a reserved phrase (for example: used for, located at, made for etc.) that indicates the semantic relation of the provided information [4]. In the above example, the *secondary* part contains only one sentence ('*made for boats and irrigation'*) in which the deserved phrase '*made for*' declares that the sentence describes the purpose of the described concept.

4 Representation Algorithm

The proposed methodology transforms the definition of a geographic concept into the corresponding conceptual graph without losing any of the information contained in the definition. The representation algorithm consists of two main steps: *tagging* and *parsing*. In the first step, we follow appropriate rules to tag every word of the concept definition. In the second step, we apply a deterministic algorithm in order to parse the tagged definition and create the corresponding conceptual graph.

Alshawi [13] was the first who developed the idea of using a hierarchy of phrasal patterns to identify formulas in concept definitions. Later on, other researchers [14], [15] proposed the method of parsing the definition first, and then doing a search to locate defining formulas and use some heuristics to find the words involved in the relations. This paper is based on the last approach. We parse a geographic definition sentence before we transform it into a conceptual graph and then perform further steps at the graph level.

We separately tag and parse the *main* and the *secondary* part of a geographic concept definition. In that way, we produce two conceptual graphs, one corresponding to the *main* part of the definition and the other to the *secondary* one. By joining them, we result in the complete conceptual graph representation of the geographic concept.

4.1 Tagging

Every definition is made of tokens. Table 3 summarizes the chosen parts of speech (tags) that we associate with the words of the *main* and the *secondary* part of the geographic concept definition. The difference between 'vb' and 'v' tags is that 'vb' always belongs to the *determinant* section of the *main* part and represents the special verb that introduces the definition of the geographic concept.

Concerning the *determinant* section, which always consists of an {article}, the {concept name} and the verb {is} (for example: 'A Canal is'), it is tagged using the abbreviations 'art', 'n' and 'vb'. Therefore, the tagging step for the *determinant* section of *Canal* produces the output: '{A (art) Canal (n)} {is (vb)}'.

Table 3. Tags used in the first step of the algorithm

Article	Noun	Verb "be"	Verb	Adjective	Preposition	Conjunction	Reserved Phrase
Art	n	vb	v	adj	prep	conj	rp

As regards the *attributes* section, which contains the genus, the attributive adjectives and one or more prepositional phrases, it is classified into the general form [{attributive adjective}*+{genus}+{prepositional phrase}*]. Consequently, it is tagged using the abbreviations 'adj' for all attributive adjectives, 'n' for the genus and 'prep', 'n' for the prepositional phrases. Thence, the tagging process on the *attributes* section of '*Canal*' produces: '{a (art) long (adj)} {and (conj)} {narrow (adj)} {strip (n)} {of (prep)} {water (n)}'.

Finally, for the *secondary* part of a concept description, which contains one or more sentences, we apply the tagging process in each one of them. The abbreviation for the reserved phrase is 'rp' (made for, used for, located at, etc.) while the rest words of the *secondary* part are usually tagged with the abbreviations 'n', 'adj' and 'conj'.

The *tagging* step for the given definition of *Canal* results in: '{A Canal (n)} {is (vb)} {a long (adj)} {and (conj)} {narrow (adj)} {strip (n)} {of (prep)} {water (n)} {made for (rp)} {boats (n)} {and (conj)} {irrigation (n)}.

4.2 Parsing

The *parsing* process in the introduced methodology is an algorithmic procedure consisting of three phases. In the first phase, we *parse* the tagged *determinant* and *attributes* sections of the *main* part of the definition in order to create the corresponding conceptual graph. Next, we apply *parsing* rules in all clauses that belong to the tagged *secondary* part of the definition ending in the creation of the corresponding conceptual graph for each clause. Finally, we combine the previously created conceptual graphs in a single one that represents the entire geographic concept definition.

Parsing *Determinant* and A*ttributes* Sections (*Main* Part)

The conceptual graph of the tagged *determinant* section (*{article (art) concept name (n)}{is (vb)}*) always follows the general form of Figure 1. The concept type *{genus}* refers to the genus contained in the *attributes* section of the tagged *main* part. Figure 2 shows the conceptual graph for the representation of the *determinant* in the phrase *'A Canal is a …strip…'*.

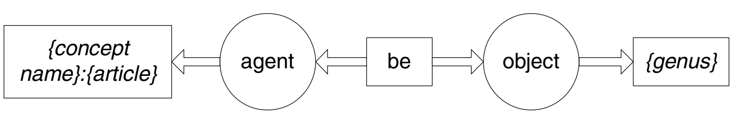


Fig. 1. Conceptual graph representing the *determinant* section

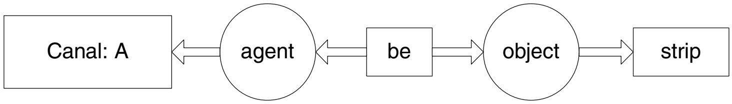


Fig. 2. Conceptual graph for *Canal*'s *determinant*

Concerning the attributive adjectives (tagged with '*adj'*) in the *attributes* section, we define one concept type for each one of them, which is connected to the genus concept type via a concept relation of type '*atr'* (Figure 3).

Moreover, for every tagged prepositional phrase, we introduce a conceptual relation of type '*preposition'* which is also connected to the genus of the definition and to the graph that corresponds to the remaining terms of the prepositional phrase. In general, a tagged prepositional phrase consists of one preposition (tagged with *'prep'*), one or more attributive adjectives (tagged with '*adj'*) and nouns ('*n'*): *{preposition}{attributive adjectives}*{noun}**. The attributive adjectives (if exist) characterize the noun (for example: 'a strip of water' or 'a strip of cold water'). Figure 4 contains the general form of the conceptual graph corresponding to the prepositional phrase of type *{preposition}{attributive adjective}{noun}*.

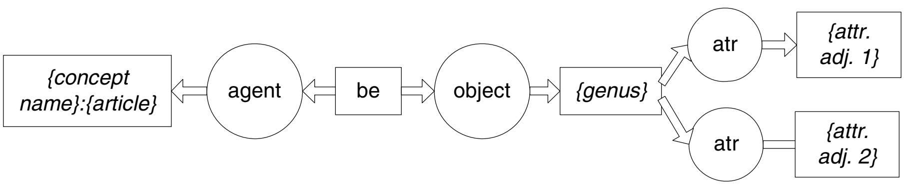


Fig. 3. Conceptual graph general form for attributive adjectives

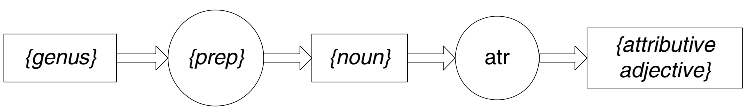


Fig. 4. Conceptual graph general form for every prepositional phrase

Therefore, for the given definition of *'Canal'*, the *main* part is represented as follows.

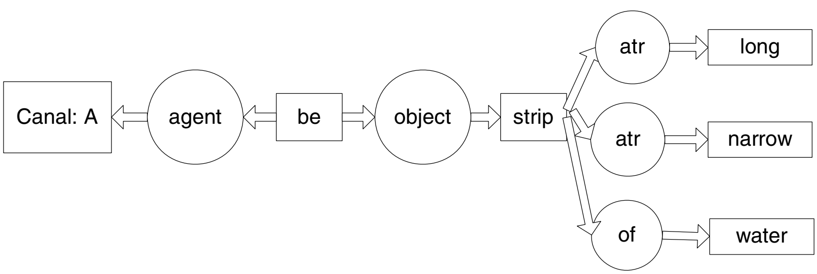


Fig. 5. *Canal*'s *main* part conceptual graph

Parsing *Secondary* Part

Every sentence in the *secondary* part, as resulted from the tagging process, consists of a reserved phrase that reveals the sentence's semantic relation type and the remaining part providing the information itself or value of the relation (for example 'made for boats and irrigation'). In the parsing procedure, the tagged reserved phrase is transformed into the corresponding concept type (for example *'made for'*). This concept is related to the genus concept via a concept relation of type *'agent'* and to the concept types that correspond to other structural elements of the sentence via a concept relation of type *'object'*.

Figure 6 shows the general conceptual graph representation form of a definition's *secondary* part. We consider that the general type of every sentence in the *secondary* part is: *{reserved phrase}({attributive adjectives}{information})**, where the *'information'* is represented with the concepts *'info 1'*, *'info 2'*.

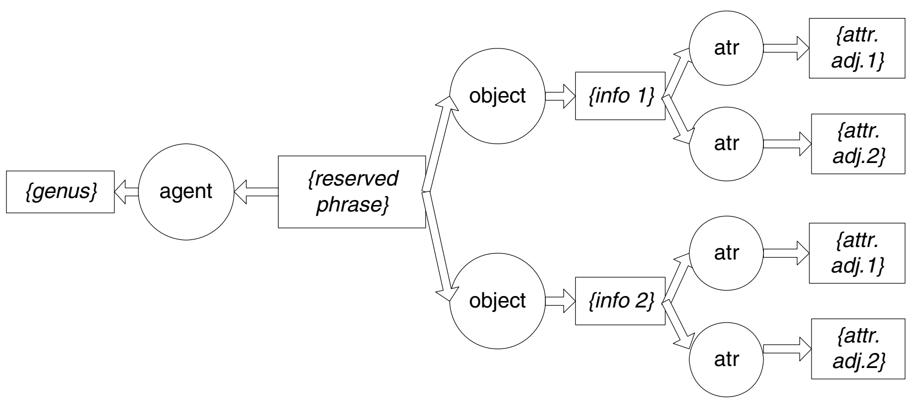


Fig. 6. Conceptual graph general form for the *secondary* part

Figure 7 shows the representation of the *secondary* part of the 'C*anal'* definition.

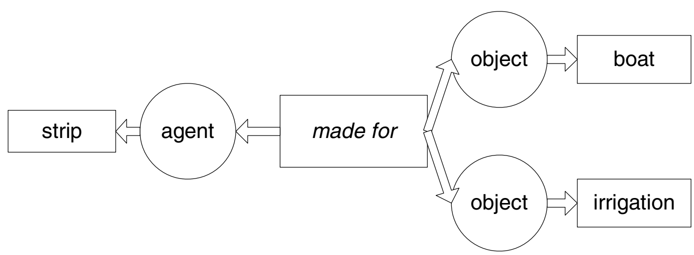


Fig. 7. Conceptual graph representation for *Canal*'s *secondary* part

The above step draws from the methodology for analyzing definitions and extracting information in the form of semantic relations which was introduced by [15] and further pursued by [16] and [17]. This approach consists in the syntactic analysis of definitions and in the application of rules, which examine the existence of certain syntactic and lexical patterns. Patterns take advantage of specific elements of definitions, in order to identify a set of semantic relations and their values based on the syntactic analysis.

Combination

The combination of the conceptual graphs corresponding to the *main* and *secondary* parts of a geographic concept definition produces the integrated representation of the definition. It is the simplest step in the overall procedure since both of the two graphs contain the common concept '*genus*'.

Figure 8 represents the conceptual graph corresponding to the output of the parsing method for the *main* and the *secondary* parts of the definition: *'A Canal is a long and narrow strip of water made for boats and irrigation'*.

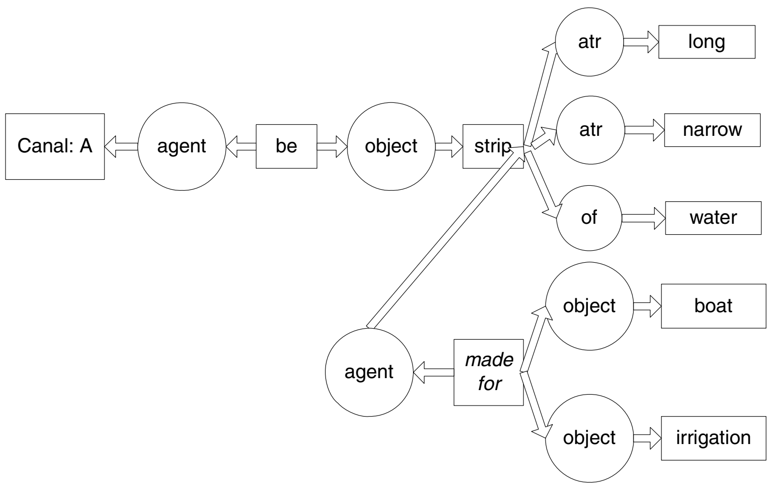


Fig. 8. Conceptual graph representing *Canal'* s concept definition

5 Comparison Algorithm

Analyzing geographic concept definitions constitutes an effective way for revealing and capturing geographic knowledge. Based on the proposed algorithm for representing geographic knowledge using conceptual graphs, we introduce a straightforward methodology for the semantic comparison of two geographic concepts.

The procedure takes as input two geographic concept definitions and follows the next steps:

1. Builds the corresponding representations of the given definitions (CG1 and CG2).
2. Determines the 1…n intersections of CG1 and CG2 (I1, I2, … In).
3. Applies a well-defined formula in each intersection that relatively measures how similar the two conceptual graphs are, in order to produce a real number between 0 and 1 that shows the value of similarity between the two geographic concepts according to our algorithm.
4. Summarizes the outputs of the previous step in order to produce the overall similarity value.

In the next paragraphs, we describe the proposed comparison methodology along with an illustrative example that semantically compares concepts *Sea* and *Lake*. The definitions of these concepts, as they appear in the lexical database WordNet are:

- *Sea*: A large body of salt water partially enclosed by land.
- *Lake*: A body of fresh water surrounded by land.

5.1 Building Conceptual Graph Representations CG1 and CG2

For transforming the definitions of the two geographic concepts into the corresponding conceptual graphs, we follow the introduced representation algorithm. Applying the two steps, tagging and parsing, in every part of the given definitions, we construct the conceptual graphs CG1 and CG2 shown in Figures 9 and 10.

In this step, it is necessary to find synonyms and hypernyms for category terms and concepts. Reference ontologies, dictionaries or thesauri may provide this information, however human intervention may also be necessary at this phase.

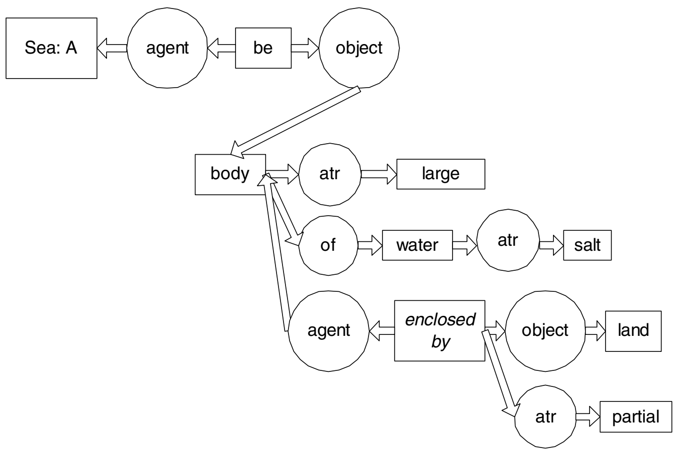


Fig. 9. Conceptual graph representation CG1 of '*Sea'* definition

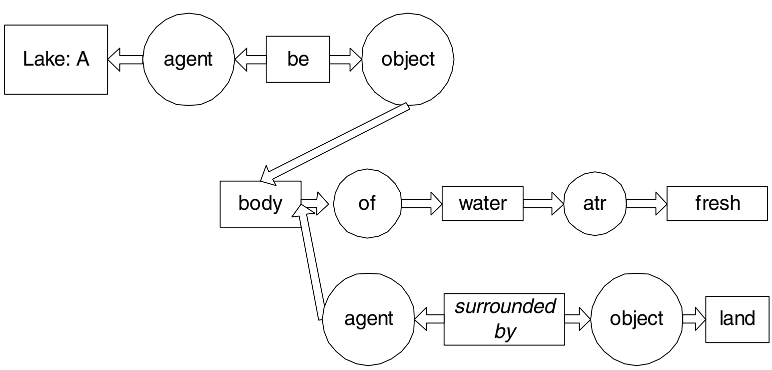


Fig. 10. Conceptual graph representation CG2 of '*Lake'* definition

For the purpose of our running example, we used WordNet and Merriam-Webster online. For example, concepts “enclosed by” and “surrounded by” are synonymous and therefore they represent the same concept.

Exploring the two definitions and analyzing their corresponding representations (CG1 and CG2), we conclude that they have the same genus or hypernym (‘*body’*), which means that they subsume in the same class. But, concerning their *differentia*, which specifies how different a concept is from another concept in the same class, we notice that the concept ‘*Sea’* is characterized by the attributive adjective ‘*large’*, the prepositional phrase ‘*of water’* and a single sub-clause (‘*enclosed by land’*) which describes further the concept, while ‘*Lake’* is characterized by the prepositional phrase *‘of water’* and the sub-clause ‘*surrounded by land*’. The attributive adjective ‘*fresh’* refers to the noun *‘water’*.

The next table summarizes the differences in every part of the given definitions.

Table 4. Genus and differentia for *‘Sea’* and *‘Lake’*

	Definition: *Sea*	Definition: *Lake*
Genus	Body	Body
Main part	Large, of water	Of fresh water
Secondary part	Enclosed by land	Surrounded by land

5.2 Determining Intersections I1, I2, …In of CG1 and CG2

After comparing CG1 and CG2, we determine their intersections depending on their structure, concept nodes and relation nodes. We name the corresponding conceptual graphs I1, I2, … In.

Every intersection I consists of all concept types that appear both in CG1 and CG2 and all relations that relate these concepts and appear in both CG1 and CG2. When an intersection consists of a single concept node, then there are not any relation nodes.

Therefore, comparing the conceptual graph representations of definitions ‘*Sea*’ and ‘*Lake*’, we build the intersections I1 and I2.

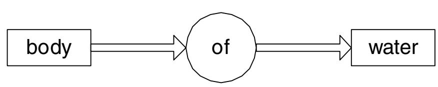


Fig. 11. Intersection I1 of CG1 and CG2

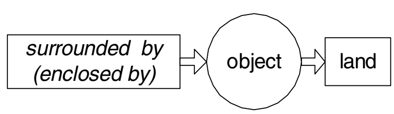


Fig. 12. Intersection I2 of CG1 and CG2

It is important to mention that we never consider the intersection of Figure 13 because it is common to all conceptual graphs that represent geographic concept definitions according to the introduced methodology.

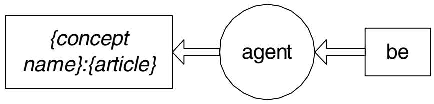


Fig. 13. Common intersection

5.3 Applying the Proposed Similarity Formula in I1, I2…In

To determine how similar CG1 and CG2 are, based on each of their intersections, we apply a deterministic formula that produces a number between 0 and 1. 1 indicates that CG1 and CG2 are semantically equivalent, while 0 indicates that they are completely different.

Moreover, because the similarity between two geographic concepts represented using conceptual graphs depends on both the concept types that they have in common and their position in CG1 and CG2, it is essential to construct a similarity measure that depends on both of these characteristics.

In the comparison algorithm, we adopt and properly reform the Dice coefficient in order to measure the similarity between CG1 and CG2 (where CG1 and CG2 represent geographic concepts). The proposed coefficient is analogous to the Dice coefficient but it also depends on what kind of concepts the two graphs have in common. For example, two geographic definitions that share the same genus are more similar than two entities that have in common only one or more attributive adjectives.

Therefore, if CG1 and CG2 are conceptual graphs that represent the definitions of two geographic concepts, I is any of their intersection and:

- C_{CG1} and C_{CG2} represent the number of concept nodes in CG1 and CG2.
- $C_{I\text{-}GENUS} = 1$ when I contains the common genus of CG1 and CG2 (if exists) and 0 otherwise.
- $C_{I\text{-}MAIN}$ is the number of concept nodes of I that also belong to the main part of CG1 and CG2.
- $C_{I\text{-}SEC}$ is the number of concept nodes of I that also belong to the secondary parts of CG1 and CG2.

Then the conceptual similarity measure S_C of CG1 and CG2 based on their intersection I is calculated as follows:

$$S_C = 2(W_{GENUS}*C_{I\text{-}GENUS} + W_{MAIN}*C_{I\text{-}MAIN} + W_{SEC}*C_{I\text{-}SEC})\ (C_{I\text{-}GENUS} + C_{I\text{-}MAIN} + C_{I\text{-}SEC}) / (C_{CG1}+C_{CG2}) \quad (2)$$

Where:

- $W_{GENUS} = 0.5$, is the weight of the common genus in CG1 and CG2 (if exists).
- $W_{MAIN} = 0.3$ / (total number of concept nodes of CG1 and CG2 belonging to their main part), is the weight of every concept node in I that belongs to both main parts of CG1 and CG2.

- W_{SEC} = 0.2 / (total number of concept nodes belonging to the secondary parts of CG1 and CG2), is the weight of every concept node in I that belongs to both secondary parts of CG1 and CG2.

Assigning different weights to concepts of I, depending on their position in CG1 and CG2, we achieve to relate the value of S_C not only with the total number of concepts that the two conceptual graphs have in common in intersection I, but also with the exact position of every concept in I in both definition representations. This means that the proposed similarity measure is higher for two definitions that have a number of common concepts belonging to their main parts than two definitions that share the same number of common concepts but in their secondary parts.

The selected values ensure that the weight of the common genus (if exists) is always bigger than the weight of any other concept the two graphs have in common and that the weight of any common concept belonging to both main parts of CG1 and CG2 is always bigger than the weight of any common concept belonging to both secondary parts of the two graphs. In case that CG1 and CG2 are exactly the same (i.e. they have the same genus and the same main and secondary parts), the similarity measure equals 1.

Therefore, applying the above formula for the calculation of S_C for I1 and I2, we are able to measure the semantic similarity between the geographic concepts '*Sea*' and '*Lake*' based on their intersections:

$S_{C(I1)} = 2\ (0.5 * 1 + (0.3/5)*1 + (0.2/5)*0)\ (1 + 1 + 0) / (7 + 5) = 0.186.$

$S_{C(I2)} = 2\ (0.5 * 0 + (0.3/5)*0 + (0.2/5)*2)\ (0 + 0 + 2) / (7 + 5) = 0.026.$

5.4 Estimating the Similarity Measure from $S_{C(I1)}, S_{C(I2)},\ldots S_{C(In)}$

The exact value of the proposed similarity measure for two geographic concepts expressed by conceptual graphs is the sum of $S_{C(I1)}, S_{C(I2)}\ldots S_{C(In)}$.

Consequently, the corresponding value for concepts '*Sea*' and '*Lake*' is: 0.186 + 0.006 = 0.212. From this result, it is obvious that CG1 and CG2 are semantically similar and that they do have concepts in common. In case there were a greater number of common concepts (especially if they belonged to the main parts of the two graphs), this value would be higher.

6 Conclusions and Further Work

The present research focuses on the representation of geographic concept definitions using conceptual graphs and the development of a comparison methodology that is based on the proposed representation method.

Developing a straightforward and easy-to-implement process for transforming a structured geographic concept definition into the corresponding conceptual graph representation breaks many limitations and obstacles in the extraction of semantic information from definitions of geographic concepts and facilitates the implementation of an interoperable geographic environment.

Moreover, the comparison algorithm, based on the structure and content of the graphs expressing geographic concepts, produces as output a similarity value between 0 and 1, which shows how much two concepts are semantically close to each other.

The present work is the first step towards establishing methodologies for identifying and representing similarities between concepts in geographic ontologies. The next step involves the extension of the introduced algorithm in order to allow measuring the similarity between two geographic concept definitions according not only to the conceptual similarity of their representations, but also to their relational similarity. This is very important because of the bipartite nature of conceptual graph representations (concepts and relations).

Furthermore, we are going to incorporate characteristics which ensure that the semantic similarity is measured not only quantitatively but also qualitatively and that the similarity algorithm also takes into account the heterogeneities between two conceptual graphs that represent geographic concept definitions.

References

1. Sowa, J. *Conceptual Structures: Information Processing in Mind and Machine*. Addison-Wesley, 1984.
2. Sowa, J. *Knowledge Representation: Logical, Philosophical and Computational Foundations*. Brooks Cole Publishing Co., 2000.
3. Tepfenhart, W., Cyre, W. Conceptual Structures: Standards and Practices. *In 7th Intl. Conf. Conceptual Structures*. Springer-Verlag, Blacksburg, VA, 1999.
4. Kavouras M., Kokla M., Tomai E. Determination, Visualization and Interpretation of Semantic Similarity among Geographic Ontologies. In: Gould, M., Laurini, R., Coulondre, S. (eds): *Proceedings of the 6th AGILE Annual Conference on Geographic Information Science*. Presses Polytechniques et Universitaires Romandes pp. 51-56, 2003.
5. Kuhn, W. Modeling the Semantic of Geographic Categories through Conceptual Integration. In Egenhofer, D., Mark, D. (eds): *Proceedings of the Second International Conference GIScience*. Springer, pp. 108-118, 2002.
6. Kokla, M., Kavouras, M. Extracting Latent Semantic Relations from Definitions to Disambiguate Geographic Ontologies. In Zavala, G. (eds): *GIScience 2002 Abstracts.* University of California Regents, pp. 87-90, 2002.
7. Conceptual Graph Standard. NCITS.T2 Committee on Information Interchange and Interpretation, 2002.
8. Ellis, G., Lehmann, F. Exploiting the Induced Order on Type-Labeled Graphs for fast Knowledge Retrieval. In Tepfenhart, W., Dick, J., Sowa, J. (eds): *Conceptual Structures: Current Practices*. Lecture Notes in Artificial Intelligence. Springer-Verlag, 1994.
9. Huibers, T., Ounis, I., Chevallet, J. Conceptual Graph Aboutness. In Ekland, P., Ellis, G., Mann, G. (eds) *Conceptual Structures: Knowledge Representation as Interlingua*. Lecture Notes in Artificial Intelligence. Springer, 1996.
10. Rasmussen, E. Clustering Algorithms. In Frakes, W., Baeza-Yates, R. (eds.): *Information Retrieval: Data Structures and Algorithms*. Prentice Hall, 1992.
11. WordNet 2. A Lexical Database for the English Language. *Cognitive Science Laboratory*, Princeton University.
12. Alshawi, H. Analysing the Dictionary Definitions. In Boguraev, B., Briscoe, T. (eds): *Computational Lexicography for Natural Language Processing*. Longman Group UK Limited, pp. 153-170, 1989.
13. Klavans, J., Chodorow, S., Wacholder, N. From Dictionary to Knowledge Base via Taxonomy. In *Proceedings of the 6th Annual Conference of the UW Centre for the New OED*: Electronic Test Research, pp.110-132, 1990.

14. Montemagni S., Vanderwende, L. Structural Patterns vs. String Patterns for Extracting Semantic Information from Dictionaries. In *Proceedings of the 14th COLING*, pp. 546-552, 1992.
15. Jensen, K., Binot, J. Disambiguating prepositional phrase attachments by using on-line dictionary definitions. *Computational Linguistics*, 13 (3-4): 251-260, 1987.
16. Ravin, Y. Disambiguating and Interpreting Verb Definitions. In Jensen, K., Heidorn E., Richardson, D (eds): *Natural Language Processing: The PLNLP Approach*. Kluwer Academic Publishers, 1993.
17. Vanderwende, L. The Analysis of Noun Sentences using Semantic Information Extracted from on-line Dictionaries. PhD Thesis, Faculty of the Graduate School of Arts and Sciences. Georgetown University, Washington D.C., 1995.
18. Karalopoulos A., Kokla M, Kavouras M. Geographic Knowledge Representation Using Conceptual Graphs. *7th AGILE Conference on Geographic Information Science*, Crete, Greece, 2004.

Ontology Ontogeny: Understanding How an Ontology Is Created and Developed

Hayley Mizen, Catherine Dolbear, and Glen Hart

Ordnance Survey, Research and Innovation, Romsey Road,
Southampton SO16 4GU United Kingdom
{Hayley.Mizen, Catherine.Dolbear,
Glen.Hart}@ordnancesurvey.co.uk

Abstract. This paper describes the development of a systematic method for creating domain ontologies. We have chosen to explicitly recognise the differing needs of the human domain expert and the machine in our representation of ontologies in two forms: a conceptual and a logical ontology. The conceptual ontology is intended for human understanding and the logical ontology, expressed in description logics, is derived from the conceptual ontology and intended for machine processing. The main contribution of our work is the division of these two stages of ontology development, with emphasis placed on domain experts themselves creating the conceptual ontology, rather than relying on a software engineer to elicit knowledge about the domain. In particular, this paper concentrates on the creation of conceptual ontologies and analyses the success of our methodology when tested by domain experts.

1 Introduction

Ordnance Survey, the national mapping organisation for Great Britain, is investigating the potential benefits of introducing a Topographic Semantic Reference System to improve the integration of topographic and thematic data. The ultimate purpose is to enable machine understanding, which in turn provides the potential for data and service interoperability. An ontology is an important component of a semantic reference system, and we are therefore researching the nature of such ontologies and methods to create them. This paper describes our current work on developing a methodology to create domain ontologies. In part we have titled the paper "Ontology Ontogeny" to emphasis our interest in the development of ontologies, ontogeny being the development processes an animal undergoes from egg to adult; but in part we just thought it too good a conjunction of terms with similar roots to miss.

Section 2 provides background to the research, explaining our motivation and placing the research in context. We provide a brief review of other approaches to ontology construction in Section 3 and outline our own views on the structure of ontologies in Section 4. In Section 5 we describe our own methodology and in Section 6 provide an analysis of its success to date. Finally, Section 7 contains our closing observations and suggestions for future research directions.

M.A. Rodríguez et al. (Eds.): GeoS 2005, LNCS 3799, pp. 15 – 29, 2005.

2 Background

Ordnance Survey has the challenge of enabling third parties to integrate their data with the topographic data that it provides. In order for an organisation to complete its business related tasks it is frequently necessary for multiple data sources to be combined (integrated) and used together in a structured way. As there may be differences in semantics as well as in the structure of these datasets, the data must be adapted to fit the task, often with compromises being made. Currently, the cost of these integration and adaptation activities is a major barrier to the adoption and efficient exploitation of complex datasets. An important aspect of this integration process is the recognition of semantic differences between datasets. Often these differences are missed due to incomplete documentation, but more importantly mistakes occur because of misunderstanding due to assumptions made at the domain level. These mistakes may be costly: subtle differences in semantics may result in data being improperly integrated, which may not be noticed until after operational decisions are made.

We are investigating whether technologies currently applied to the development of the Semantic Web, particularly ontologies, may facilitate the capture of domain knowledge in such a way as to detect errors in data integration, or, due to the explicit nature of the semantics, prevent them occurring at all. Ultimately this technology could enable such integration and adaptation to occur "on the fly" – making the Semantic Web a reality. Given that this cannot be fully achieved in the near or medium term, our general approach is an incremental one. Manual processes will be systematically automated, eventually enabling some fully automated processes and services and others which are significantly automated, but still require some manual input. We are therefore initially placing an emphasis on ontologies being used as an aid to largely manual processes.

In order to increase the understanding and acceptance of the technology of Ontologies within Ordnance Survey, we have taken the notion of Semantic Reference Systems as proposed by Werner Kuhn [1] and broadened its definition. Whilst Kuhn describes such systems in terms of top level ontologies that provide grounding for other ontologies, we use it to also encompass what we term foundational domain ontologies. These are ontologies that are intended to establish de facto semantics for a particular topic area. In the case of Ordnance Survey, it would be to establish a Topographic Semantic Reference System. Kuhn rightly states that a Semantic Reference System is more than just an ontology: it must also support the transformations between domains. At this stage though, our research is limited to the development of the ontological component.

We see a Topographic Semantic Reference System as complementary to the existing Coordinate referencing system (The British National Grid) and the developing Feature Referencing System (OS MasterMap®) [2]. Its purpose will be to provide a common semantic definition of the principal topographic concepts applicable to this country, which will assist users of Ordnance Survey data to automatically conflate and adapt it with their own data. In order to build such a system however, we must first understand the necessary structure of the ontology and how it will be constructed.

3 Previous Approaches

The creation of an ontology is usually viewed as a knowledge acquisition task, which, as defined by Kidd [3], involves eliciting, analysing and interpreting human expert knowledge, and transferring this knowledge into a suitable machine representation. Many other ontology methodologies are based around a similar structure, or contain similar design criteria, but all differ slightly and not one has become a formal or even de-facto standard. Uschold's methodology and Fernàndez-López and Gómez-Pérez's METHONTOLOGY are believed to be the most representative [4]. Both methodologies propose initial modelling phases that develop an implicit shared understanding and explicit informal human-readable glossaries before structuring the information in a logical ontology. Uschold and King first define their classes precisely and unambiguously using natural language which are structured as a semi-formal hierarchy before building a logical ontology [5]. METHONTOLOGY further develops a more systematic method for domain conceptualisation. It provides a set of tasks for assisting the ontology modeller in capturing and structuring the information required for a logical ontology using a series of tables, a "Data Dictionary", and a series of concept trees [6]. Some of these representations however, are clearly specific to their domain of Chemistry and would not be suitable for a geographic ontology. In other existing methodologies, the processes of knowledge capture and formal coding have been carried out at the same time (for example, [7] and [8]). However, we support the approach of Uschold and King [6] and Gómez-Pérez et al.[9], who advocate the use of separate stages in ontology development.

The most popular methodologies [6] and [9] promote the creation of concept trees and sub-groups of similar classes. These promote an early dependence on the structures of formal languages and encourage the ontology modellers to group classes under familiar headings that in some cases do not represent the true logic underlying the relationship. This is particularly true for sub-sumption relationships, for example in a topographic ontology, concepts may be unnecessarily divided under "natural" and "man-made" branches in a hierarchy. We believe an ontology should also be much more than a taxonomy, and in fact, we discourage the use of hierarchies altogether, as they decrease the potential for inference and reuse by creating dependency between concepts. Under the umbrella of risk management, outside the world of academia, we have found that not all domains have a clear classification structure and cannot always be divided into small bounded modules. We have yet to look further into overcoming difficulties found with ontology modularisation and scalability and have identified this as an area of future research. More detailed reviews and discussions of ontology methodologies can be found in [5] and [10].

Knowledge representation is procedural and people find it difficult to describe exactly how they carry out these procedures or tasks. As the expert becomes more competent in their activity, the more automatic their use of knowledge becomes, and the less accessible it is to the knowledge engineer [11]. Past approaches in the AI community as part of the development of expert systems have tended to view knowledge elicitation as a preliminary to the more serious business of encoding knowledge in a software language. Rather than placing emphasis on the importance of knowledge elicitation from a domain expert, our strategy is instead to provide the domain expert with a set of clear and systematic steps that enable them to author a first-stage or "conceptual" ontology themselves.

4 Our Approach to Ontology Construction

While our methodology may be broadly applicable to the construction of any type of ontology, we are focusing on the development of domain ontologies in particular. A domain ontology is a formalisation of the knowledge in a subject area (domain) such as topography, ecology, biology etc, and differs from other types of ontology such as the task ontology (a formalisation of the knowledge necessary to solve a specific problem or task abstracted above the level of a specific situation or organisational context).

Each ontology can be thought of as a pair of two linked ontologies: a conceptual ontology and a logical ontology. The conceptual ontology is intended to be primarily for human consumption: it attempts to balance the need for maximal formality of the ontology whilst retaining clear human comprehension. It is a means for domain experts to capture domain knowledge, which encourages them to record and describe their ideas explicitly in a standard structure. It should be free from the constraints of the logical ontology, and should not be influenced by the structures or rules that description logics present. The logical ontology provides a machine interpretable representation, typically using a derivative of first order logic such as description logic and is produced by an ontology expert familiar with languages such as the W3C standard language for representing ontologies; OWL (Web Ontology Language). It is generated from the conceptual ontology and, as we have found, information will be lost during this translation due to the inability of description logics to represent the true complexity of a conceptual ontology[1]. We have considered the possibility of including an intermediate stage between the conceptual ontology and the OWL ontology, where information is transformed into a more expressive logic such as First Order Logic to achieve a more complete representation. The advantage of the SHOIN(D) logic on which OWL is based is however in the tractability of its reasoning. .We believe a split between these two ontologies is important, given the difficulty most people have in comprehending description logics and their inability to fully express the full richness of a domain. We emphasise that the conceptual ontology should be constructed and verified by the domain expert themselves, rather than the ontology engineer, and cite this as an advantage of our two-stage methodology.

Conceptualising a domain before processing it in a logical ontology can play a more significant role that simply collating information to be modelled. When separated from the formalisms of logical modelling, the structure can be used by domain experts themselves to record their knowledge and interpretations of their domain. In some instances, the domain expert may not have any existing complete documentation of their domain, in which case these stages of conceptualisation and knowledge capture are a useful mechanism for exposing domain information. While ontology experts' modelling techniques tend to pre-empt the knowledge structure imposed by description logics and ontology languages such as OWL, we assume that the domain experts are unfamiliar with ontologies and their rigorous structures. Instead of communicating the methodology using jargon familiar only to ontology engineers, we use common terms that can be easily understood by our target audience. For example,

[1] Information loss also occurs during the creation of a conceptual ontology but this is less easily measured.

instead of using terms like "classes", "properties", and "attributes" we use the words "concepts", "relationships" and "characteristics". Our methodology is presented using a systematic structure, similar to the task-based structure used by Gómez-Pérez et al.[9], but is additionally supported by illustrations, examples, and written guidelines. A systematic task list promotes the use of a standard ontology structure and ensures the ontologies are produced consistently, which maximises the potential for interoperability between different ontologies.

5 Method for Constructing a Conceptual Ontology

Our approach is to provide domain experts with a comprehensive and systematic set of criteria and guidelines to assist them through the entire conceptual ontology life-cycle. The methodology is still being developed, and we describe the basic skeleton of tasks for building a domain conceptual ontology only, supported by examples from the flood risk management ontology. The methodology comprises four main tasks: deciding on the requirements and content of the ontology; populating a knowledge glossary and constructing a set of triples (relationships between concepts); evaluating the ontologies; and finally, documentation of the conceptual ontology.

Stage 1- Preparatory

Task 1: Identifying the Requirements

At the very onset of modelling the domain knowledge, the domain expert formulates a set of requirements for the ontology. This will provide the modeller (the domain expert) with a clear focus for ontology content and scope. It can be used throughout the ontology life-cycle as an evaluation tool. The criteria for identifying the requirements are similar to that identified by both Uschold and King [6] and Grüninger and Fox [7]. Primarily, the modeller records their definition of an ontology, their purpose for building it (which determines which type of ontology they produce), the scope of the intended ontology (based on the purpose), and a set of competency questions. We advise that the scope should be contained and restricted in size, so that ontologies produced are manageable and consistent. If the scope is large (e.g. the domain of topography) then the modeller may wish to sub-divide the domain into further domain ontologies (hydrology, urban areas, etc.), and integrate the modules together when they are all complete. The competency questions will differ depending on which type of ontology is being built. For domain ontologies, the competency questions are formulated so that they can be used to check at each stage of ontology construction whether the correct relationships have been created between the concepts, and whether the relationships created sufficiently describe the domain. To define competency questions, some pre-conceptions about which concepts are core to describing the domain are required. Generic examples include, *"Does the ontology sufficiently describe the domain to a level of granularity suitable for the purpose? Do all concepts have at least one link to another concept?"*. Examples specific to a hydrology domain ontology within the topographic field would be: *"Have I sufficiently described the essence of being a"River" in terms of its relationships to its characteristics and links to other concepts? Have I made the distinctions clear in the relationships describing "River" and "Stream?"*

Task 2: Collecting the Data

Here, we acquire the input knowledge base needed to construct the conceptual model, based on the purpose, scope and competency questions. When appropriate, the modeller should reuse other ontologies that also suit the purpose of the ontology they are building. We are currently developing our research for reusing single concepts and sets of concepts and relationships from other conceptual ontologies, and the reuse of full conceptual ontologies.

The modeller should identify any documentation that captures the knowledge they wish to be in the ontology. The information must be suited to the purpose, be within scope, and be true to their representation of the domain in question. Where documentation is not available or sufficient, the ontology will be built using the domain expert's knowledge of the domain. Either manually or through using semi-automated data mining programmes, the modeller should extract the semi-structured sentences that contain information required to be in the ontology. These should contain important descriptor terms such as "and", "or", "sometimes", and "not"; terms that describe probability: "must", "likely", "might", "maybe", "sometimes"; and terms that describe possibility, including "usually" and "typically". It should then be verified that these sentences are complete within themselves, and complete in terms of recording all necessary information required. The aim is to reduce ambiguity by restructuring sentences, but ensure information is not lost. The sentences are then validated against the goals or purpose. It is well understood [12] that the linguistic and logical meanings of "and" and "or" are different. By recording these semi-structured sentences, our methodology provides the logical ontology modeller with a documentation trail so that he or she can check back to understand exactly which of the two possibilities the domain expert meant.

Stage 2: Populating a Knowledge Glossary

The first step in capturing and structuring the domain knowledge is to populate a knowledge glossary. Comparisons can be drawn with the "Data Dictionary" and the "Tables of attributes" proposed by Gomez-Perez et al. [9], but the glossary is more suitable for an audience less familiar with "classes" and "attributes". We have used common natural language for the glossary headings and provide guidelines to assist the domain experts in identifying the correct information. Table 1 provides an example of two concepts from the flood risk domain ontology populated in a knowledge glossary.

Table 1. Knowledge Glossary

Term	Synonym term	Natural language text definition	Linguistic term	Conceptual ontology term	Core / Sec.	Core concepts chars	Value and units	Rules, constraints and assumptions
Flood risk map	Flood map	A map classifying risk into risk levels applicable to different areas.	Noun	Concept	Core	Has scale Shows risk level	Scale: 1:25000 to 1:100000	Scale is for regional maps
Is an input of		A relationship term to describe the link between two concepts, where one is used in the creation of the other.	Verb	Relationship	Core			Has inverse relationship (has input)

The information required for the glossary is extracted from the semi-structured sentences and enhanced by the domain expert. The modeller is encouraged to record the linguistic definition of a term (e.g. noun, verb) as an intermediate step to identifying which terms are concepts in the ontology and which are relationship terms or characteristics (attributes). The nouns are more likely to be concepts and verbs are most likely to be relationship terms. Defining the terms and recording these is a useful means for the domain expert to clarify their definition and interpretation of the term and its use within the ontology. The definitions will also be used in later stages of the methodology to identify relationships to other terms. The "core concepts" which are key to describing the domain are distinguished from the "secondary concepts" which either describe aspects of the core concepts or have differentiating relationships with them. This is useful for later stages of modelling. Secondary concepts are not members of the domain under consideration, but are necessary to enable concepts in the domain to be related to other domains. For example in the case of hydrology a core concept "River" could define a relationship to a secondary concept "Field" that would rightly belong to a different domain. Core concepts are vital to the ontology and are presumed to have the most relations to other concepts. They should be described within the ontology not only by their relations to other concepts, but also by their relation to their attributes (e.g. has size, has location), or as we term them in the conceptual ontology methodology, "characteristics". The domain expert is encouraged to identify these using the semi-structured sentences and their own knowledge, and will use this information to explicitly describe the core concepts by their wholes and parts in the conceptual ontology. Characteristics of secondary concepts are not required in the conceptual ontology. The domain expert uses the glossary to record any assumptions, rules or restrictions governing the use of the definition, the characteristics or values within the ontology to reduce the assumptions made when creating the network of relationships between concepts and to avoid information loss at this early stage in development.

We appreciate that not all the knowledge required for the ontology will be captured from the semi-structured sentences and domain expert's knowledge, and that the glossary will undoubtedly be added to when the ontology is developed further. However, when the modeller is content with the information they have captured, the glossary should be validated against the purpose and scope set in the requirements stage. We are currently developing more efficient techniques than populating a table for composing the glossary and more formally testing the content of the glossary against the semi-structured sentences.

Stage 3: Creating a Semantic Network of Triples

The next stage is to use the information captured in the knowledge glossary to construct a concept network that describes the domain in question. A concept network visualises an ontology as nodes (concepts) and links (relationships between concepts). This is much more than Gomez-Perez's "Concept Classification Trees" [9] which organise domain concepts in taxonomies. Our approach limits the use of hierarchical relationships that can encourage the creation false groupings of concepts or unnecessary divisions between groups of concepts (e.g. the division of "natural" and "man-made" concepts in a traditional topographic object classification), although these are not completely prohibited. Instead, we argue that richer inference can be achieved if the concepts are defined within themselves and through a range of relationships to other concepts (i.e. concept-to-concept relations and concept-characteristic-relations), so the shape and form of a semantic net is more comparable to a lattice than a hierarchy.

We have adapted Gruber's five design criteria to reflect our own interpretations [13]. These criteria should be used throughout the ontology life-cycle to enforce consistency and coherence. The modified criteria are:

1. Clarity: Definitions should be expressed unambiguously to ensure the intended meanings are comprehensible. They should represent the modellers interpretation of their domain.
2. Coherence: Relationships should be consistent with definitions.
3. Extendibility: It should be possible to add new terms without the revision of existing definitions accepting the addition of new relationships.
4. Minimal encoding bias: The choice of terms should not be made purely for convenience or implementation.
5. Minimal ontological commitment: Secondary concepts should be described using the weakest model only. These do not need to be described in terms of their characteristics. Gruber suggests that all terms should be defined using the weakest model, thus making as few claims as possible. But although this maximises reusability, if ontologies are to be integrated through techniques such as semantic similarity, identification of matches between concepts will be essential. Core concepts should therefore be described additionally by their wholes and parts through relations to their characteristics although these should be both necessary and sufficient for the purpose and scope.

We specify a number of rules for creating a concept network to enforce consistency of the ontologies, including the following:

a. The modeller should work bottom-up, building the ontology with the most specific concepts which can then be generalised when necessary (identifying super-ordinates), to prevent groups of concepts being grouped under hierarchies or false semantics. Membership of a concept to another should be created instead by inference.
b. Multiple inheritance should only be created when the concept can inherit all of the characteristics of both super-ordinate concepts.
c. We advise only creating hierarchies when necessary for describing the domain, where the sub-ordinate inherits all the characteristics of its super-ordinate plus other characteristics, or when the ontology needs to move between different levels of granularity. The modeller should consider whether an alternative relationship can be used instead.
d. If new concept or relationship terms (i.e. those that are not already in the glossary) are needed when building the concept network they should be validated against the scope, goal or purpose, and added to the glossary before adding them to the conceptual ontology; this will ensure the term is used consistently with its definition.
e. If information can not be captured in the concept network, it should be recorded as semi-structured sentences or as an example for the logical ontology modeller who will attempt to include this information in the logical ontology.
f. If concepts or small groups of concepts are found to have no links into the rest of the concept network, the modeller should review their inclusion in the semantic net. If their inclusion is not suited to the scope or description of the domain they should be disregarded.

The domain expert should choose which method of representation both suits their ontology and their personal preference. To date we have used two methods of visualising the concept network: using network diagrams for graphically displaying links between concepts (Figure 1 illustrates an example from the flood risk management ontology), and creating a list of "conceptual ontology triples" where the concepts and relationships are recorded as subject-predicate-object. Both can be difficult to manage if the scope of the ontology is large, and the former does not facilitate the capture of "restrictions, assumptions and constraints". Cyclicity and repeated triples are also difficult to manage in a list of triples. Similarly with the glossary, we are developing more sophisticated tools for capturing the triples using a user-friendly interface.

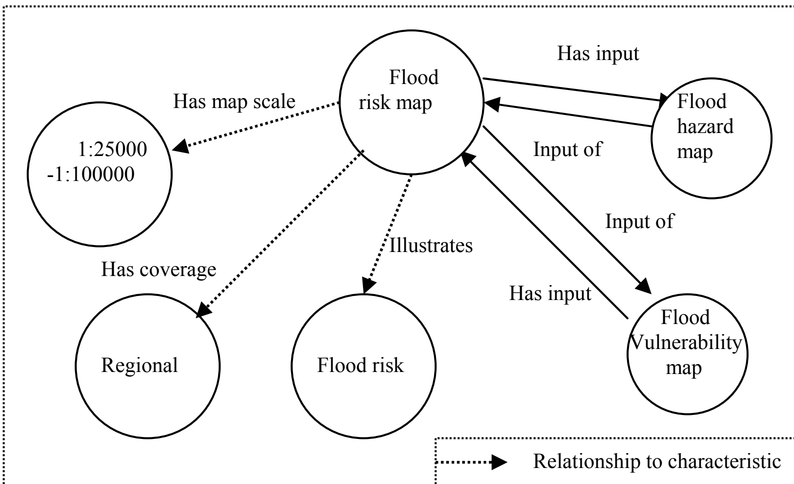


Fig. 1. Concept network for concept "Flood Risk Map"

The domain expert should use the information captured in the knowledge glossary, plus their own knowledge to complete a concept network by completing the following tasks systematically:

Task 1: Create the links between the core concepts and their characteristics. Additional characteristics that were not captured in the glossary can be added if suitable to the purpose and scope. The modeller is likely to use the relationship term "has" to create the link between a concept and its characteristic. This should be specialised where possible to explicitly describe the link. For example, we would say "Flood Event, Has Location, Location", instead of "Flood Event, Has, Location".
Task 2: Identify links between different core concepts using the most suitable relationship term that explicitly defines the type of link. For the topographic domain ontology, we found these to primarily be mereologic (part of), topographic (next to), and affordance relationships.
Task 3: Using the "equivalent to" relationship term, add in links between synonym concepts. These concepts must share a full set of characteristics.

Task 4: Create the links between core concepts and secondary concepts. These should be relationships that describe the core concepts in terms of their relation to other things that are not their characteristics.
Task 5: Create a relations network between the relationship terms. Similarly to the concept networks the modeller can produce the relation network using graphical network diagrams where the relationship terms are the nodes. The relation network should identify which relationship terms are sub-ordinates, which have an inverse relationship (e.g. "has part" and "part of"), which are transitive (e.g. "is input of") etc. If any new relationship terms are added to the model they should be added to the glossary first. The relations network is then used to identify which relationships are missing, incomplete, or inconsistent in the concept network. It is common for ontology modellers to record relationships uni-directionally so it is likely that all inverse relationships will have to be added to the concept network.

Stage 4: Evaluation of the Conceptual Ontology

The modeller should firstly check whether all information captured in the glossary has been captured as triples or restrictions and constraints in the concept network, or has been recorded as information loss. Secondly they should check that the information captured in the concept and relations networks has been captured in the glossary. If there is information missing from the glossary further checks should be made against the scope and purpose. The modeller can now evaluate their conceptual ontology against the following criteria:

- Logical consistency: Checks are made for cyclicity, repetitions, and missing triples. The competency questions can be used to identify core concepts and triples that have not been captured.
- Conceptual accuracy: The domain expert should agree with the information that has been captured as triples, in that it represents his/her own interpretation of the domain, task or application.
- Minimal ontological commitment: Only those relationships suited to the purpose and within scope have been created, i.e. the core concepts are well defined by their explicit relationships to other concepts and relations to their characteristics. Secondary concepts have only been used in the ontology to describe the core concepts.
- Clear differentiation between ontologies: The concepts and relationships captured in should be suited to the ontology type created (i.e. a domain ontology does not contain concepts more suitable to a task ontology).
- Vagueness has been handled well: the modeller has attempted to capture probability, possibility, uncertainty and fuzziness within the conceptual ontology.
- Information loss is recorded.

Stage 5: Documentation of Conceptual Model

The conceptual ontology documentation must include the knowledge glossary, the concept and relationship networks, recorded information loss, and any defined rules and assumptions made throughout the modelling process.

6 Analysis

Our methodology for conceptual ontologies was exposed within the European Sixth Framework project Orchestra [14] where it was accepted in November 2004 as the standard for constructing the risk management domain ontologies. Feedback from the domain experts, and our own experiences in using the methodology, has enabled us to identify obstacles within the methodology that occur in real situations outside of the academic bubble, and has subsequently been used to further develop the methodology. We discuss the main obstacles found when building the five risk management ontologies (for flooding, earthquakes, coastal zone, forest fire and systemic risk) here.

6.1 Problems with Scalability

The domain experts massively underestimated the amount of time required to produce an ontology and consequently built their ontologies based on a large scope (planning and preparation phases of risk management). The resulting conceptual ontologies were consequently a mix of both domain and task ontology concepts and relationships that jumped between levels of granularity and which were incomplete and inconsistent. This identifies three major problems in the methodology: firstly, it does not provide guidelines for limiting the ontology to a small scope in order to produce smaller, more manageable ontologies; secondly, the guidelines for separating concepts into those that are suitable for either domain or task ontologies are unclear; and thirdly, there are no guidelines for modularising the ontology so that it can either be produced by various people at the same time, or broken down into sub-domains for later partial reuse. The solution to the first problem is fairly trivial and can be solved immediately by encouraging the domain expert to define a small, contained and restricted scope at the outset of the ontology modeling phase to ensure that ontology construction is manageable and is more likely to be complete. The second and third however, require further thought. We believe the processes for constructing a domain and task ontology should differ, but we have yet to produce full task ontologies through which we can refine the existing method to distinguish between these different processes or develop a new methodology specifically for task ontologies. When reviewed, the Orchestra partners' ontologies were found to contain more task based concepts than domain ones. We have begun to develop more technical approaches to solving the third problem, for example using a tool suitable for conceptual modeling that is similar to the Protégé version control system the author of the information input can be tagged. To avoid missing concepts that lie between obvious boundaries, the competency questions could be used to check whether all the required concepts and relationships have been captured.

6.2 Recording Triples

The domain ontologies produced in Orchestra included many concept-concept relationships, but included limited numbers of concept-characteristic relationships where core concepts are described by their wholes and parts. In most cases the level of explicit detail required by the conceptual ontology was not captured within the risk management conceptual ontology triples. The types of relationships recorded were generic and ambiguous; for example, "Rainfall causes Flood", from which the logical ontology would not then infer that it was specifically that it is *"heavy rainfall that causes a river to burst its banks which then causes a flood",* which is the true logical relationship. The tabular format for recording the triples was not the most effective means of encouraging the level of detail required from a conceptual ontology. It also proved difficult to identify loops of iterated relationships, repeated triples, or missing triples. This has prompted us to develop more efficient and effective means of capturing and structuring this information which include the use of text mining tools to extract concepts from documents, along with developing our own tools to facilitate the authoring of the domain ontology "triples". The intention is that the tool would take the domain expert through the steps of the conceptual ontology methodology up to the triples stage. The triples could be stored as either RDF or as simplified OWL concepts, whilst retaining the distance between the domain expert and the restrictions of OWL. This would of course not be full OWL as most of the knowledge would still be in natural language in annotation which would require further methods for transforming it into a complete logical representation. We are also developing a toolset of common ontologies that describe spatial relations, shapes (e.g. lines and polygons), time, and other relationship terms that can be reused to produce the Ordnance Survey full topographic ontology, or by others producing geographic ontologies.

6.3 Dealing with Information Loss

We encouraged domain experts to record any information that they could not model as triples either against the relevant triples in a column labelled "restrictions" in the triples table, or as semi-structured sentences. We evaluated the information loss to identify common areas across domains where information could not be captured as triples.

The primary cause of information loss was in the recording of fuzzy or uncertain relationships. It is common to find that domain experts do not have an explicit model of the conditions under which a relationship is true. This is part of the well-known knowledge elicitation problem and therefore it is difficult for domain experts to record information at the level of detail required. Our solutions to common issues are:

1. Quantified uncertainty and probability (e.g. one flood in 100 years). In these circumstances we record the probability as a concept within the ontology.
2. Where an instance has characteristics of more than one class (e.g. a section of a floodplain containing a number of different vegetation types). In the conceptual ontology we record *"Floodplain, has cover, Grassland and Shrubs and…"*, which would be added to the logical ontology as *"Floodplain, has cover, a number of: grassland, shrubs…"*
3. Where there is a lack of information (e.g. a flood is less likely to occur when the river banks are high). The solution to this is to use a scale of categories that can be assigned meanings (e.g. high – low; less likely – probable – more likely).

Another common area of information loss occurred in domains which attempt to model comparisons that are numerical and based on inexact relationships. For example within the earthquake risk domain, many of the concepts in "risk assessment" require comparisons to be made between the hazard (the demand) and the vulnerability/resistance of the elements at risk (the capacity). This type of relationship cannot be modelled in the triples format. Similarly, the occurrence of induced events depends on inexact relationships between the causative and consequent event. In addition, information loss occurred when domain experts attempted to model triples that have conditions (e.g. an "if…then…else" statement) and tasks and processes. These issues suggest a conceptual ontology should comprise more than a glossary and a set of triples.

6.4 Evaluating Ontologies

Although we have identified the domain ontologies produced within Orchestra as being incomplete and inconsistent, our set of criteria was insufficient for a robust evaluation, as we have no means of formally testing the logical consistency of the conceptual ontologies using the competency questions. We intend to incorporate this feature into the tools we are developing for recording the triples more effectively. We have since identified that the evaluation criteria will also vary depending on who is using the ontology. The ontology producer would want their conceptual ontology to be logically consistent, agree with purpose and scope, have well defined concepts, and contain reused concepts and relationships only originating from authoritative sources; and in these cases a logical ontology modeller is often required to second the evaluation to ensure logical consistency, until there are more formal means of testing this. Someone who intends to reuse an ontology, in addition to looking for the producer's requirements, would want to reuse an ontology produced using the de facto standard, in a format compatible with theirs, and would perform checks to ascertain whether the ontology has reused ontologies from credible sources or from companies with similar interests to their own, hence, evaluation would be suited to check for this criteria.

The domain experts reported that the methodology was very systematic. This assisted them in consistently recording the required information in a structure that was common across the five risk management domains, which enhanced the potential for interoperability. Although not all were complete and consistent (primarily caused by the problems with scalability) the risk management conceptual ontologies reflected the domain experts' true interpretation of their own domains. The information was captured without being constrained by the description logic representation of ontology languages such as OWL, a common limitation of promoting codification in early stages of ontology development. Our approach clearly demonstrated the benefits of separating conceptualisation of the domain, which is captured in a conceptual ontology, to the stages of formalising the domain in a logical ontology. The mere process of capturing their knowledge more formally has also enlightened the domain experts about details within their data. Previously undocumented relationships and assumptions have become explicit, and areas of similarity across the five risk management domains have been identified, which will facilitate future interoperability research.

7 Conclusions and Further Work

The primary output of this research is the robust testing of our proposed methodology for assisting domain experts to construct ontologies themselves: an exercise which has not been reported in the literature before. Our approach successfully demonstrated the benefits of splitting ontology construction into two separate stages: conceptual ontology modelling and logical ontology modelling. As a consequence, the resulting domain ontologies for risk management and hydrology reflected the domain experts' interpretation of their own domain within a structure suitable for transformations into a logical ontology but without the common restrictions and compromises forced by description logic formalities. The ontologies were also found to be more expressive (that is, they were more than hierarchies or taxonomies) than many previous attempts by domain experts to develop ontologies described in the literature. Evaluation of the ontologies and feedback on the domain experts' experiences was useful for identifying future developments in the methodology. It firstly illustrated where further detailed explanation was needed and secondly it identified the areas for further research. These include the development of tools for assisting the domain expert in recording the conceptual triples, for example, to identify cyclicity and facilitate formal testing through the use of competency questions. Another area of further research concerns ontology modularity, in order to facilitate scalability and conceptual and logical ontology reuse.

Acknowledgments

This work has been partially funded by the European IST FP6 project Orchestra (IST FP6-511678).

This article has been prepared for information purposes only. It is not designed to constitute definitive advice on the topics covered and any reliance placed on the contents of this article is at the sole risk of the reader.

References

1. Kuhn, W. Implementing Semantic Reference Systems, In, *Proceedings of the 6Th Agile conference on Geographic Information Science*, pp. 63-72, 2003.
2. Ordnance Survey MasterMap ® http://www.ordnancesurvey.co.uk/oswebsite/products/osmastermap/
3. Kidd, A. *Knowledge Acquisition for Expert Systems, A Practical Handbook*, Plenum Press. 1987.
4. Cristiani, M., Methodologies for the Semantic Web. *AIS SIGSEMIS Bulletin*, 1(2): 102-136. 2004.
5. Uschold, M. and King, M., A Methodology for Building Ontologies, In *IJCA195 Workshop on Basic Ontological Issues in Knowledge Sharing*, Montreal, 1995.
6. Fernàndez-López, M., Gómez-Pérez, A., and Jurista, N., *METHONTOLOGY: From ontological art towards ontological engineering*, Stanford, 1997.
7. Grüninger, M., and Fox, M. Methodology for the Design and Evaluation of Ontologies In *IJCA195 Workshop on Basic Ontological Issues in Knowledge Sharing*, Montreal, 1995.
8. Noy, N. F., and McGuinness, D.L., *Ontology Development 101: A guide to creating your first ontology*, 2001.
9. Gómez-Pérez, A., Fernàndez-López M., and De Vicente, M., Towards a method to conceptualize domain ontologies. In *Working notes of the workshop on Ontological Engineering, ECAI'96*, pp. 41-52. ECCAI, 1996.
10. Hemsley-Flint, F. Towards interoperability between ecological and topographical data through the application of ontologies, First year report, unpublished, 2005.

11. Bainbridge, L. Asking Questions and Accessing Knowledge, In *Future Computing Systems*, 1: 143-150, 1986.
12. Gruber T. Towards principles for the design of ontologies used for knowledge sharing. *International Journal of Human-Computer Studies* 43: 907-928. 1995.
13. Rector A, Drummond, N, Horridge M, Rogers, J, Knublauch,,H., Stevens, R, Wang, H and Wroe, C. OWL Pizzas: Common errors & common patterns from practical experience of teaching OWL-DL. In *Proceedings of the Engineering and Knowledge Management Workshop*, 2004.
14. Orchestra Open Architecture and Spatial Data Infrastructure for Risk Management IST FP6-511678 http://www.eu-orchestra.org/ 2004.

Representing the Meaning of Spatial Behavior by Spatially Grounded Intentional Systems

Christoph Schlieder

Laboratory for Semantic Information Technology,
Bamberg University, Germany
christoph.schlieder@wiai.uni-bamberg.de

Abstract. The problem of interpreting the trajectories of a person (user) moving in a spatial environment is fundamental for the design of any location-based application. We argue that in order to correctly assign a meaning to the spatial behavior encoded by the trajectory, it is necessary to express the meaning in terms of the user's intentions, more specifically, the goals that the user intends to achieve. Along the trajectory, these intentions will change frequently because the user's initial goal is decomposed into sequences of subgoals. The paper proposes a representational formalism and a reasoning mechanism for knowledge about an agent who acts according to changing intentions: *spatially grounded intentional systems*. An objective consists in making the representation as expressive as possible without running into a behavior interpretation problem that is computationally intractable. The approach is shown to be sufficiently expressive to model the interaction between intentions and behavior in a location-based game, CityPoker.

1 Introduction

Location-based applications are found in a variety of contexts ranging from tourist information (Abowd et al, 1997) and interactive geo-art (Hull, Clayton and Melamed, 2004) to mixed-reality games (Flintham et al. 2003). Basically, these applications implement information services which run on mobile devices such as PDAs or Smartphones. The trajectory of the user moving with the device in the spatial environment is analyzed in order to determine what information to display. The challenge of designing location-based applications consists in solving this interpretation problem by exploiting knowledge about how user behavior relates to the spatial environment in the specific problem domain. Since this knowledge is strongly domain-dependent, each new application makes it necessary to spend considerable effort on modeling (1) a spatial environment, (2) different types of spatial behavior, and (3) their interaction.

Interpreting spatial behavior – whether by a piece of software or by a human observer – is a process of assigning meaning. Opinions diverge about how this meaning should be formally described. In the simplest case, meaning is construed as classification, as a mapping onto one of a few possible behavioral types (Laube and Imfeld, 2002). More complex data structures may be used to describe the compositional nature of meaning (Patterson et al. 2004). However, there is a central assumption shared by the approaches to location-based computing: meaning depends

M.A. Rodríguez et al. (Eds.): GeoS 2005, LNCS 3799, pp. 30 – 44, 2005.

on spatial context. The same type of behavior can be interpreted differently if it occurs at different places. This is very similar to the basic intuition behind activity-based spatial ontologies which characterize spatial entities by the actions they afford (Jordan et al. 1998; Kuhn, 2001). Applying this principle to regions of space amounts to define regions in terms of what you can do there and what you cannot do there.

Current research on location-based information services explores the relation between behavior and environment from at least two perspectives. On the one hand one looks for ways to represent spatial behavior such that knowledge about the environment can be inferred. Ashbrook and Starner (2002), for instance, analyze GPS-tracks of a user to learn which locations in a city are significant to that user. On the other hand one tries to model the spatial domain knowledge in a way that it can be used to interpret user behavior. The second perspective is adopted in this paper.

We present a representational formalism and a reasoning mechanism for knowledge about an agent who acts according to changing intentions in a spatial environment: *spatially grounded intentional systems*. The main body of the paper is organized as follows. In section 2 a fundamental issue for intention modeling, the room crossing problem, is described and a location-based game, CityPoker, is presented as scenario for studying the analysis of user intentions. Minimal requirements for the representation of intentions in the agent are derived in section 3 from a modeling of the scenario. Spatially grounded intentional systems are introduced as a representational framework and shown to meet the requirements in section 4. The paper concludes with discussing related work and a perspective on future work (section 5).

2 Modeling the Spatial Specificity of Behavior

2.1 The Room Crossing Problem

A concrete example helps to illustrate how the relation between the user's behavior and the spatial environment is typically conceptualized in the field of location-based computing. It is drawn from Hull, Clayton and Melamed (2004) who describe an authoring tool for applications known as "mediascapes" – applications that deliver or capture digital media depending on the spatial context of the user. The tool supports a visual programming approach in which the designer specifies an application by drawing regions on a 2D-map of the environment. Each region is automatically associated with two events: the entering and leaving of the user. Other events may be specified manually. The designer then decides how the system should react when a specific event occurs. Finally, a piece of code is generated in the markup-language of the authoring tool. In the code example shown below, the application plays an audio file as soon as the user enters the region (Fig. 1).

This type of conceptualization of the behavior-environment relation has two implications for the design of appropriate representational formalisms. The first is the structuring of the environment by spatial regions (defined by the <region> tag in the code example). Location-based applications typically consider regions being sets of points that satisfy additional topological and geometric constraints such as path-connectedness. By set-theoretic inclusion the regions form a partonomy which constitutes a primary source of spatial knowledge for the location-based application.

```
<region name=”northwest”>
    <circle          x=”123”          y=”456”
range=”20”/>
    <onEnter>
        <playMedia media=”harp”/>
    </onEnter>
```

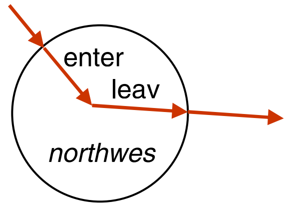


Fig. 1. Location-based presentation of digital media (adapted from Hull & al., 2004)

The second, less obvious, implication is the direct mapping of the user’s spatial behavior onto information services (specified by the <onEnter> tag in the code example). Directly mapping behavior onto services raises problems which have been discussed by Schlieder and Werner (2003) in the context of a museum information system. Of immediate practical consequence is the *room crossing problem*. It arises from the difference between a visitor entering a room in a museum with the intention to cross it and a visitor entering the same room with the intention to visit it. The location-based application should react differently in the two cases. However, with directly mapping behavior onto services, since there is only one service that the entering behavior can be mapped to, the system will necessarily provide the same information service both times.

The room crossing problem occurs at almost all levels of granularity in the partonomy with particularly annoying consequences at the level of atomic regions (i.e. regions not containing other regions of the partonomy) since the user enters and leaves them frequently. For a museum information service, the regions defined around individual exhibits are considered atomic. Obviously, a visitor who quickly moves through a room unintentionally crossing various atomic regions should not be bombarded with information about exhibits.

Location-based application designers generally solve the room crossing problem by adding code that sets conditions in the mapping which depend on the behavior history of the user or on data from context sensors. Schlieder and Werner (2003) have shown that a conceptually simpler solution of the problem can be achieved by explicitly modeling user intentions and by decomposing the behavior-service mapping into a behavior-intention mapping and a subsequent intention-service mapping.

However, the proposal of using intentions as mediating representational level, or, which amounts to the same, the proposal of interpreting spatial behavior in terms of user intentions still has a major gap: a description of how to represent and compute the behavior-intention mapping is missing. It is this gap that we are now closing.

2.2 A Location-Based Game as Scenario

Location-based games which are played in geographic space, on a playing field that extends beyond vista space, provide an ideal scenario for studying the interaction of environment, behavior and intentions. The main advantage lies in the fact that the set of user intentions is completely determined by the rules of the game. This situation contrasts with that of daily life activities but it offers the possibility to study the interaction of low-level and high-level intentions because the underlying process of planning moves in the game is well-understood.

CityPoker was conceived as a GPS-game in 2004 by the author and his team at the Laboratory for Semantic Information Technology, Bamberg University. The game confronts two players who move around a city in search of cards with which they can improve their poker hand. When the game ends after a preset time limit (e.g. 2 hours) the player with the highest hand wins. CityPoker is played with a deck of 20 cards from which five cards are dealt to each player at the beginning of the game. The remaining 10 cards are hidden in pairs at 5 geographic locations called caches. Fig. 2 shows a typical initial situation of the game.

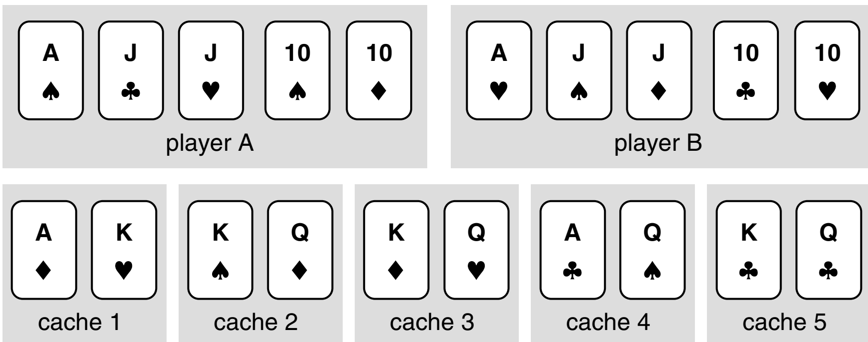


Fig. 2. Starting situation in the CityPoker game

During the game, a player may visit each cache once but is not forced to do so. At the cache, the player changes one of his cards in turn for one of the 2 cards in the cache. The card that is discarded becomes available to the opponent unless, of course, the opponent has already visited the cache. CityPoker is a full-information game which means that at any moment in time each player knows which cards the opponent possesses as well as which cards are hidden in which cache. This implies communication between players and is one of the reasons why it is attractive to design a CityPoker assistance system running on Smartphones.

GPS is used by the players to find the caches. However, the location of the caches is not specified exactly. The players obtain location information by answering a multiple choice quiz: "*Cache 1 is located at N 49° 53,XXX - E 10° 53,YYY. Get more information by correctly answering the following question. Which pope is buried in Bamberg cathedral? (a) Clemens II, XXX=535 YYY=595 (b)… (c)…*" A player not knowing the answer will have to search at all three potential cache locations. Additional help is provided by a perceptual hint such as *"The cache is near a tree"*. Finding a cache almost always involves some kind of spatial search which makes time a crucial element in the game. Note, for instance, that for the situation shown in Fig. 2 there is a winning strategy ("keep the ace, get the four kings") for a player who is capable of reaching 4 caches before the opponent reaches even the first cache. A deeper analysis of the game is out of the scope of this paper and will be published elsewhere. The spatial behavior of the players is given by their movements in geographic space as recorded by the GPS receivers (Fig. 3).

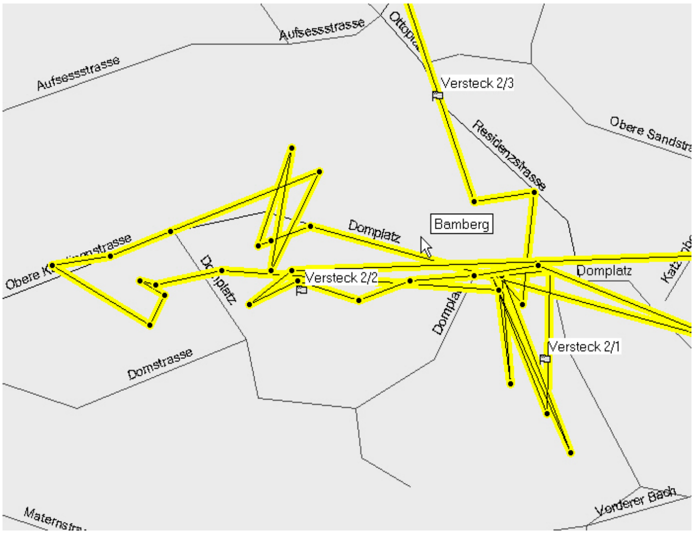


Fig. 3. GPS track of a player at three potential cache locations

Table 1. Spatial regions in the CityPoker game

primary region	secondary regions	tertiary regions
playing field	search region 1	potential cache 1/1
		potential cache 1/2
		potential cache 1/3
	search region 2	potential cache 2/1
	. . .	
	search region 5	potential cache 5/3

Spatial regions with a special significance for the players are (1) the *playing field* which confines the game to some part of geographic space, e.g. the historic center of the city, (2) a *search region* for each of the 5 caches, defined by coordinate ranges, e.g. N 49° 53,000 - N 49° 53,999, (3) a total of 15 = 3 · 5 *potential cache* locations specified by more precise coordinates, e.g. N 49° 53,535. Because of measurement imprecision also potential caches are to be considered spatial regions rather than points. The resulting partonomy constitutes a strict hierarchy:

3 Interpreting Behavior in Terms of Intentions

3.1 Intentions in the CityPoker Scenario

Spatial behavior is analyzed on the basis of GPS track data describing how the two players move in geographic space. Using spatio-temporal properties of the tracks, it is possible to distinguish different types of spatial behavior. For the automatic analysis, the GPS tracks are segmented at the boundaries of the spatial regions and, if necessary, at additional points. Behaviors are defined in a way that they can be recognized using simple spatio-temporal characteristics. In the example we are using the rather limited set of behaviors listed in Tab. 2.

Table 2. Spatial behaviors in the CityPoker game

behavior	Spatio-temporal characteristics
orienting	Not changing position or moving very slowly
riding	moving fast along a shortest path in the street network
searching	moving slowly in random patterns
other	any other type of behavior

Table 3. User intentions in the CityPoker game

region type	intention	state of affairs aimed at
playing field	*ChooseCard*	best possible decision is taken on which card to get next from which cache
	GetCard	search region of the card is reached in minimum time and the card is found
search region	*AnswerQuiz*	quiz is answered correctly
	ReachCache	potential cache corresponding to the answer is reached and card is found
potential cache	*FocusSearch*	region to be searched is further restricted by exploiting the perceptual hint
	SearchCard	spatial search is successful, i.e. the card is found at the cache

Table 4. Intention-service mapping for the CityPoker game

intention	information service offered
ChooseCard	show textual information about the state of the game listing which cards players and caches possess
GetCard	display large scale map showing search regions, current position and position of the opponent
AnswerQuiz	show quiz question associated with search region as text
ReachCache	display map at medium scale showing potential caches, current position and position of opponent (if applicable)
FocusSearch	show hint giving a perceptual cue as text
SearchCard	display small scale map showing additional features in the vicinity of the cache, e.g. trees, benches, …

Interpreting the spatial behavior of a player in terms of intentions (behavior-intention mapping) helps the assistance system for CityPoker to determine which information to display to the player (intention-service mapping). Since the game is played using bicycles, the possibilities to interact with the system while moving around are extremely limited. Correctly anticipating the user's information needs is thus of great utility. Another task which profits from interpreting a player's behavior is the design of an intelligent software opponent for single player variants of the game. Because of its simplicity, we use the information presentation task as our running example.

An assistance system for players of CityPoker should be able to distinguish a number of user intentions. In general, the set of intentions is determined by analyzing the information needs of the user. A useful heuristic consists in considering the regions of the partonomy one by one and asking what intentional actions of the user could possibly be supported by an information service. Generally, there is no need to distinguish intentions that cannot be supported by different services. The result for our domain is shown in Tab. 3.

Note that not every type of intentional action is possible everywhere. For instance, the final spatial search for a card (*SearchCard*) may only occur within a potential cache region. Note also that no intention to cross a potential cache region or a search region has been modeled because players generally cross them unintentionally and no specific information service is available that could support the intention if it occurred. Thus, if a crossing room problem arises it will have to be solved differently. Information services supporting a player having one of the intentions from Tab. 3 are listed in the intention-service mapping shown in Tab. 4.

3.2 Requirements for Modeling Intentions

A sequence of spatial behaviors such as the one shown in Fig. 4 constitutes the input to the analysis of player behavior. In order to keep the example within a manageable size, the sequence chosen is less complex than most sequences found in CityPoker games. Furthermore, the track is rendered in schematic form only. Segmenting the track yields the 12 subtracks shown. Using the spatio-temporal criteria of Tab. 2, the subtracks are assigned to one of the 3 possible spatial behaviors (or 4 behaviors when the "other" behavior is counted). The result of this categorization is listed in Tab. 5 which also indicates the most specific spatial region associated with the behavior. The last column of Tab. 5 lists what we would expect as correct interpretation of the behavior in terms

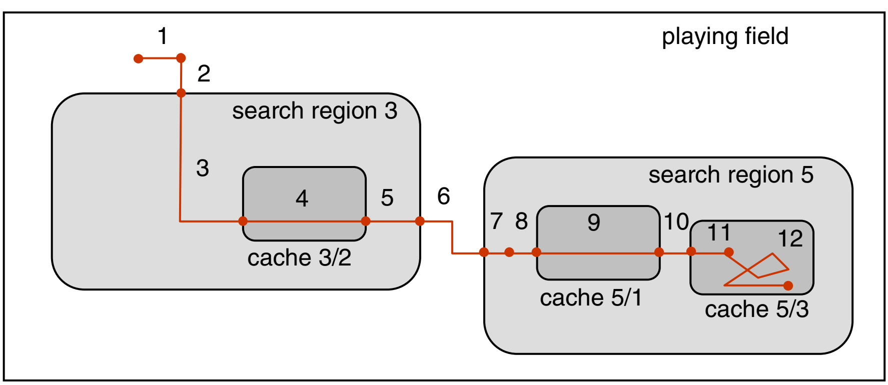


Fig. 4. A possible sequence of behaviors in CityPoker

Table 5. Interpretation of behavior in terms of intentions

sub-tracks	behavior sequence	spatial sequence	intention sequence
1	orienting	playing field	*ChooseCard*
2	riding	playing field	*GetCard*
3	riding	search region (3)	*GetCard*
4	riding	potential cache (3/2)	*GetCard*
5	riding	search region (3)	*GetCard*
6	riding	playing field	*GetCard*
7	orienting	search region (5)	*AnswerQuiz*
8	riding	search region (5)	*ReachCache*
9	riding	potential cache (5/1)	*ReachCache*
10	riding	search region (5)	*ReachCache*
11	orienting	potential cache (5/3)	*FocusSearch*
12	searching	potential cache (5/3)	*SearchCard*

of intentions. In other words, the interpretation task consists in computing the entries in the forth column from those in the other columns.

Before analyzing the example in more detail, we formally define what it means to interpret spatial behavior in terms of intentions.

Definition: Let **R** denote a set of spatial regions, **B** a set of behaviors, and **I** a set of intentions, all three sets being finite. A *behavior sequence* is a sequence $B = b_1, \ldots, b_n$ of behaviors from **B**. Each behavior b_k occurs in a region r_k and the sequence $R = r_1, \ldots, r_n$ is called the *spatial sequence* associated with B. An *intention sequence* is a sequence $I = i_1, \ldots, i_n$ of intentions from **I**.

In our example, **R** = {playing field, … , potential cache 5/3}, **B** = {orientation, riding, searching}, and **I** = {*ChooseCard*, …, *SearchCard*} as specified by Tab. 1-3. The behavior sequence constitutes the input for the analysis of the user's behavior. We formally define what type of output to expect from the analysis:

Definition: An *interpretation* m: $\mathbf{B}^n \rightarrow \mathbf{I}^n$ maps behavior sequences of length n onto intention sequences of the same length. If an interpretation m maps $B = b_1, \ldots, b_n$ to $m(B) = i_1, \ldots, i_n$ we say for every index k that the intention i_k expresses the *meaning* of behavior b_k in the context of m.

Although the behavior sequence listed in Tab. 5 is rather simple, it provides counterexamples which entail a basic negative result about interpreting behavior, namely that we cannot expect to find an interpretation of behaviors that is independent of the position of the behaviors in the sequence. To be more precise, it is generally not possible to express the interpretation m: $\mathbf{B}^n \rightarrow \mathbf{I}^n$, $b_1, \ldots, b_n \rightarrow i_1, \ldots, i_n$ as a mapping of behaviors m^*: $\mathbf{B} \rightarrow \mathbf{I}$ with the property $m^*(b_k) = i_k$ for $k = 1, \ldots, n$. The reason is that we may find indexes k and l for which $b_k = b_l$ but $i_k \neq i_l$. For instance, in Tab. 5 we find $b_2 =$ riding $= b_{10}$ but $m(b_2) =$ *GetCard* $\neq$ *ReachCache* $= m(b_{10})$. Even if the spatial sequence associated with the behavior sequence is taken into account, i.e. the type of region where a behavior occurs, we cannot generally find an interpretation that is independent of the positions of the behaviors. For instance, in the behavior sequence of Tab. 5 we find $b_3 =$ riding $= b_8$ and $r_2 =$ search region $= b_8$ but $m(b_3) =$ *GetCard* $\neq$ *ReachCache* $= m(b_{10})$. We summarize these observations in form of a requirement:

Requirement 1: The representational formalism describing the behavior-intention mapping must be more expressive than a look-up table in the two arguments behavior and region.

To put it differently, the spatial behavior of players in the CityPoker scenario shows an inherent complexity that makes it necessary to look at the behavior history in order to assign the correct interpretation. How many steps in the behavior sequence constitute an adequate temporal context? Unfortunately, in many cases no fixed-sized context will serve the purpose. This is because of the recursively nested structure of the partonomy for which it is often not possible to specify a depth bound. Intentions interpret behaviors occurring in regions of a certain type. In many problem domains, if the user crosses a chain of nested regions $r_1 \subseteq r_2 \ldots \subseteq r_k$ this gives rise to a sequence of k entering behaviors followed by a sequence of k leaving behaviors. Generally speaking, it may be necessary to take behavior contexts of the type $b_1^k\, b_2^k$ into account for which the number of occurrences of behavior b_1 and b_2 must match. It is well known that no finite state mechanism (e.g. finite automaton, regular grammar) is able to analyze such patterns for arbitrary length k. In other words, for a number of location-based applications an even stronger requirement is needed:

Requirement 2: The representational formalism describing the behavior-intention mapping must be more expressive than a finite state mechanism.

One expects to find a trade-off between the expressiveness of the representational formalism and the computational costs of solving the behavior interpretation problem. Since location-based applications need to interpret behavior on-the-fly, it seems reasonable to restrict expressiveness in a way that the interpretation problem can be solved in polynomial time. This could be considered a kind of third requirement.

4 Spatially Grounded Intentional Systems

4.1 A Rule-Based Agent Model

Different architectures for software agents that perceive, reason and act in space have been proposed in the fields of autonomous robotics, agent systems, and cognitive modeling. It is in the last field that the most sophisticated architectures for a detailed simulation of mental processes have been developed. Architectures used for cognitive modeling purposes are generally conceived as rule-based systems, that is, they represent the mental state of the agent including percepts and actions by a pattern of symbols and use rules to generate the next mental state. A well-known example is the ACT-R architecture (Anderson and Lebiere, 1998). Because cognitive adequacy is the primary goal, cognitive modeling architectures employ complex pattern matching mechanisms and combine the rule-based approach with non-declarative forms of knowledge representation such as neural network models of long-term memory. These architectures certainly satisfy requirement 1 and even requirement 2 from section 3.2. However, they are not suited for out purposes because of the complexity of the associated behavior interpretation problem.

In the context of location-based applications, a tractable computational approach to behavior analysis is a key concern. Therefore, an agent model based on a restricted type of production systems is certainly more adequate. We briefly remind that a *production system* (T,N,P,S) consist of a set T of terminal symbols, a set N disjoint from T of non-terminal symbols, a set of productions P which in the most general case are of the form $u \rightarrow v$ with $u,v \in (N \cup T)^*$, and a specific start symbol $S \in N$. For many modeling purposes, productions of a more restricted type are sufficient. A restricted class meeting requirement 2 and sufficiently expressive for most location-based application domains consists of context-free productions. These productions have on their left hand side only a single non-terminal, i.e. they have the form $u \rightarrow v$ with $u \in N$, $v \in (N \cup T)^*$. It is well-known that a context-free production system can be transformed into an equivalent one in which all productions are of (context-free) form $u \rightarrow v$ with $u \in N$, $v = aw$ where $a \in T$ and $w \in N^*$ (Greibach normal form).

As the agent's spatial behavior is directly observable while the mental states are not, a straightforward modeling choice consists in representing behavior by terminal symbols and mental states (intentions) by non-terminal symbols of the production system. This is comparable to the modeling choice made in computational linguistics where the visible constituents of a sentence, namely the words, are represented by terminals while the constituents that a particular syntax theory introduces as explanatory constructs (e.g. "noun phrase") are represented by non-terminals. To clearly distinguish between both types of symbols, we will write strings denoting non-terminals with a capital first letter and print them in italics. A typical production rule reads:

MentalState1 → action *MentalState2… MentalStateN*

This rule expresses that *MentalState1* of the agent could have the consequence that the agent performs the action and then adopts *MentalState2… MentalStateN.* The right hand side of the rule does not specify a necessary consequence as there may be several rules having the same left hand side.

While from the agent's first person perspective a distinction between percepts and actions makes sense, this is different for a location-based application acting as external observer. Sensors only provide behavioral data. Based on that data, the location-based application has to decide whether a certain type of behavior (e.g. approaching and orienting towards an object) qualifies as perception or action. Therefore, we need not further subdivide the set of behaviors **B**. In the context of the interpretation problem as it is defined in section 3.2, only a single type of mental state, namely intentions from **I** are relevant. Production rules in Greibach normal form are of the type:

Intention1 → behavior *Intention2… IntentionN*

These production rules can be used to model a simple type of planning which decomposes the higher-level *Intention1* by first acting and then following *Intention2… IntentionN* in turn. Of course, production rules need not be stated in normal form. For instance, we may have:

Intention1 → entering-behavior *Intention1* leaving-behavior

Such a recursive production rule can be used to describe the behavior sequence arising from crossing a chain of nested regions.

Definition: Let **B** denote a set of behaviors, and **I** a set of intentions, both sets being finite. An *intentional system* is a production system A = (**B**, **I**, P,S) with terminals **B**, non-terminals **I**, a set of productions P, and start symbol $S \in$ **I** also known as the highest-level intention of the agent.

A simple model of a CityPoker player is obtained by changing slightly **B** and **I** from Tab. 2 and Tab. 3. To keep the example simple, we assume that behavior is always correctly recognized **B' = B – {**other**}**, and we add a highest-level intention **I' = I** ∪ {*PlayCityPoker*}. Without the assumption of correct recognition a larger set of productions would be needed. A = (**B'**, **I'**, P, *PlayCityPoker*) is an intentional system with the following set of production rules P = {R1, …, R9}:

(R1) *PlayCityPoker* → *ChooseCard GetCard*

(R2) *ChooseCard* → orienting

(R3) *GetCard* → riding *GetCard*

(R4) *GetCard* → *AnswerQuiz ReachCache*

(R5) *AnswerQuiz* → orienting

(R6) *ReachCache* → riding *ReachCache*

(R7) *ReachCache* → *FocusSearch SearchCard*

(R8) *FocusSearch* → orienting

(R9) *SearchCard* → searching

4.2 Computational Interpretation of Behavior

Given an intentional system with production rules in context-free form, the behavior interpretation problem is tractable since it can be solved by applying classical parsing techniques such as a chart parser based on the Early algorithm. For instance, we can solve the interpretation problem for the behavior sequence given in Tab. 5 by using the intentional system A = (**B'**, **I'**, P, *PlayCityPoker*) specified in the last section. The corresponding parse tree is shown in Fig. 5.

Actually, the parse tree provides more information than needed. The interpretation of a behavior (terminal) is given by the intention (non-terminal) from which it directly derives. For instance, the meaning of the 7th behavior in the sequence, orienting, is expressed by the intention *AnswerQuiz*. In Fig. 5 intentions expressing meanings of behaviors are marked by boxes. However, the other intentions also provide valuable information: they describe how high-level intentions decompose into lower-level intentions. At some moment, the *GetCard* intention is overridden by the *ReachCache* intention. Although the *GetCard* intention persists, the more specific *ReachCache* intention is the one that the user actively follows at that moment. Consequently, it is this intention which the location-based application should support with an adequate information service. By proceeding in the way described, all instances of the room crossing problem occurring in the scenario are solved. For example, the riding behavior in the potential cache region 3/2 is interpreted by the *GetCard* intention, whereas the same behavior in the potential cache region 5/1 is interpreted by the *ReachCache* intention.

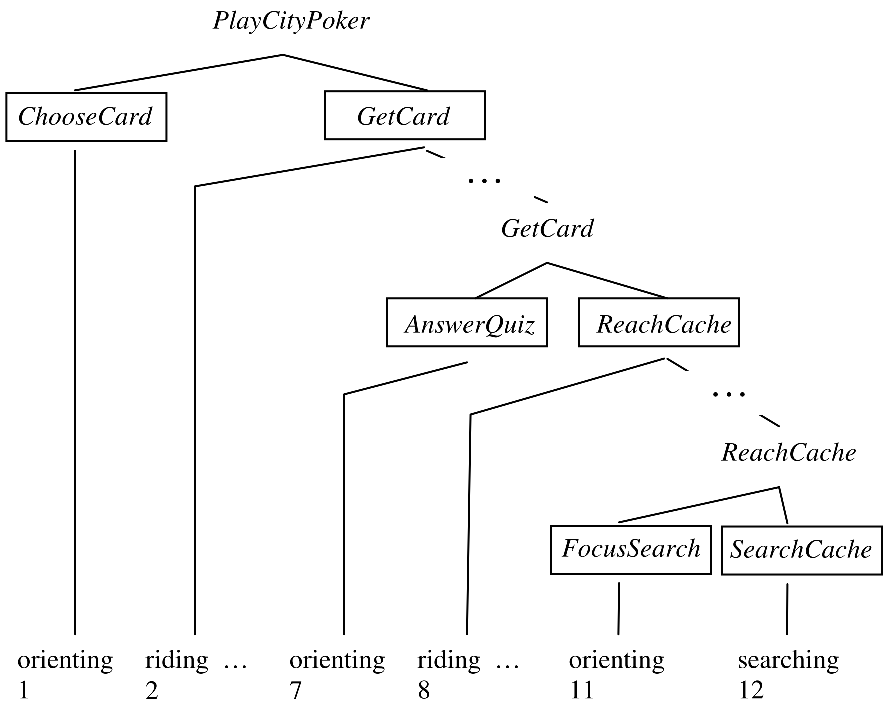


Fig. 5. Behavior interpretation problem as parsing problem

Note that not every set of production rules leads to a deterministic parsing process. Generally, there is more than one rule which can be applied at each processing stage. By exploiting the spatial specificity of intentional behavior there is a simple way to reduce rule conflicts and thereby speed up the parsing process. We assign the production rules to regions in space where they are valid. An obvious choice is to relate the validity of rules to the regions of the partonomy which define possible behaviors and intentions. The parser applies a rule only if it is valid in the region associated with the next behavior to be processed. This leads to the following definition.

Definition: Let **R** denote a set of spatial regions, **B** a set of behaviors, and **I** a set of intentions, all three sets being finite. A *spatially grounded intentional system A* = (**B**, **I**, *P*, *S, G*) is an intentional system (**B**, **I**, *P*, *S*) together with a relation G⊆*P*×**R** describing the regions in which a production rule is applicable.

In our running example, the behavior sequence of Fig.4 and Tab. 5, the spatial grounding of the production rules is specified as shown below (Fig. 6). Using the spatially grounded intentional system, the parsing process that produces the interpretation of Fig. 6 is deterministic. Although spatial grounding does not completely eliminate rule conflicts, it proves to be a valuable approach to reduce them significantly.

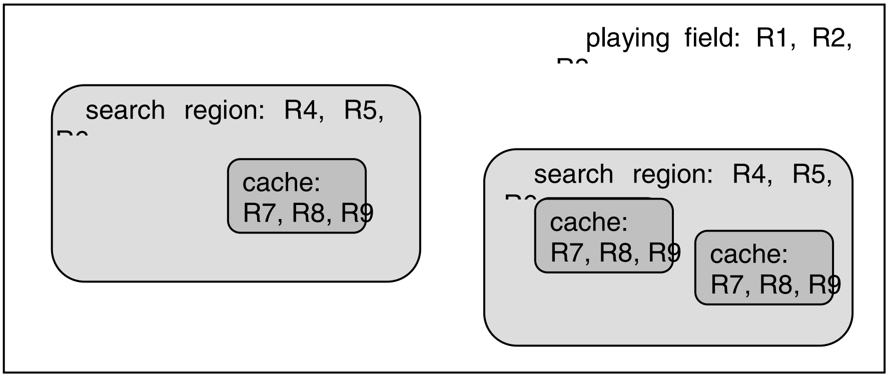


Fig. 6. Spatial grounding of production rules

5 Discussion of the Results in the Context of Related Work

We have seen in the CityPoker example that spatially grounded intentional systems provide a way of modeling how spatial behavior relates to user intentions in location-based applications. A major advantage of the approach lies in its expressiveness. The meaning assigned to the trajectory of the user is described by the hierarchy of goals encoded in the parse tree. This sharply contrasts with a flat representation that interprets the trajectory as a sequence of elementary behaviors. It has been shown that the expressiveness of context-free production rules is needed for handling goal-directed behavior in an environment which is structured by a spatial partonomy. Another characteristic of the approach is to spatially ground the production system, that is, to restrict the region within which a rule may be applied to interpret behavior. Because of the spatial grounding of the rule set it becomes possible to automatically perform the interpretation of behavior by efficient parsing processes.

Laube and Imfeld (2002) approach the problem of assigning meaning to spatial behavior from a different application perspective, namely that of tracking animal locomotion. Animals can be modeled as moving point objects producing trajectories in geographic space not very different – at first sight – from those of human users of location-based services. However, goal-driven behavior is most probably less complex in animals than humans. If true, this would mean that animal behavior can be interpreted by referring to a more limited repertoire of possible intentions than human behavior. Also, one would expect animals being involved in less complex problem solving processes resulting in rather shallow parse trees with little hierarchical nesting of subgoals. In other words, grounded intentional systems can be used to model the spatial behavior of animals but they are designed to provide the expressiveness required for analyzing complex human behavior. For animal locomotion, simpler models such as the Markov model proposed by Laube and Imfeld (2002) may provide a more adequate balance between expressiveness and computational complexity.

Different models of human agents acting in space have been proposed by researchers interested in semantic problems of geographic information science. A typical example is Frank et al. (2001) who describe a multi-agent system simulating – in one of their scenarios – travelers moving in an airport. Common to such approaches is the focus on simulation. The objective consists in producing complex spatial behavior which is why complex agent models are adopted. Unfortunately, this makes it difficult to deal with the reverse problem, namely that of automatically interpreting behavior in terms of the mental states of the agent model. For the same reason, the fine-grained models of goal-directed behavior that have been proposed by cognitive psychologists such as the ACT-R architecture (Anderson and Lebiere, 1998) or the BDI architecture used in mulit-agent systems research (Rao and Georgeff, 1991) prove to be inadequate for describing user intentions in the context of location-based services.

While the agent models considered in simulations in geographic information science are generally too complex, those proposed by research on location-based applications are rather too simple. Ashbrook and Starner (2002), for instance, use a second order Markov model to predict user behavior, i.e. with a temporal context limited to two steps in the behavior sequence. The focus of this work and similar work by Patterson et al. (2004) is on behavior prediction by probabilistic models. However, the ambiguity of behavior interpretation which is expressed by the fact that the same behavior sequence might be interpreted by different parse trees cannot be rendered adequately by these approaches.

Nevertheless, behavior interpretation could profit from modeling probabilities. Spatially grounded intentional systems handle the uncertainty involved with interpretation only by producing alternative parse trees. Although this is a very useful way to render the structural ambiguity inherent to the interpretation task, it does not provide an additional mechanism for addressing the problem of recognition indeterminacy at the level of basic behaviors. Our current research addresses the question of how approaches to probabilistic parsing could be used on this problem.

Another research issue concerns complex partonomies with overlapping regions and regions that belong to multiple superregions. Such partonomies give rise to rule conflicts: a production rule *R1* associated with region *A* and a production rule *R2* with region *B* are both applicable in the non-empty intersection of *A* and *B*. Conflict resolution schemes are currently explored that define the order in which rules should be applied. It seems plausible that the probabilistic parsing approach can be extended to incorporate conflict resolution schemes but further research is necessary to clarify the issue.

Acknowledgments

I wish to thank Klaus Stein, Peter Kiefer, and Sebastian Matyas for their comments on earlier versions of the paper. My discussions with Andrew Frank about the role of activities in ontological modeling helped to shape the ideas expressed in this paper.

References

1. Abowd, G., Atkeson, C., Hong, J., Long, S., Kooper, R., & Pinkerton, M. Cyberguide. A mobile context-aware tour guide. *Wireles. Netwworks*, 3: 421-433, 1997.
2. Anderson, J., & Lebiere, C. *The atomic components of thought*. Mahwah, NJ: Erlbaum, 1998.
3. Ashbrook, D., & Starner, T. Using GPS to learn significant locations and predict movement across multiple users. *Personal and Ubiquitous Computing*, 7: 275–286, 2002.
4. Flintham, M., Anastasi, R., Benford, S. D., Hemmings, T., Crabtree, A., Greenhalgh, C. M., Rodden, T. A., Tandavanitj, N., Adams, M., and Row-Farr, J. Where on-line meets on-the-streets: experiences with mobile mixed reality games. In *Proc. Conference on Human Factors in Computing Systems – CHI 2003*, pp. 569-576, ACM Press: New York, 2003.
5. Frank, A., Bittner, S., and Raubal, M. Spatial and Cognitive Simulation with Multi agent Systems, In *Proc. Int. Conference on Spatial Information Theory - COSIT 2001*, pp. 124-139, Berlin: Springer-Verlag, 2001.
6. Hull, R., Clayton, B., Melamed, T. Rapid authoring of Mediascapes. In N. Davis et al. (Eds.) *Proc. International Conference on Ubiquitous Computing - UbiComp 2004*, pp. 125-142, Berlin: Springer, 2004.
7. Jordan, T., Raubal, M., Gartrell, B., and Egenhofer, M. An affordance-based model of place in GIS. In: Chrisman and Poiker (eds.), *Proc. 8th Int. Symposium on Spatial Data Handling - SDH 1998*, pp. 98-109, Vancouver: IUG, 1998.
8. Kuhn, W. Ontologies in support of activities in geographical space. *International Journal of Geographical Information Science*, 15 (7): 613-631, 2001.
9. Laube, P., and Imfeld, S. Analyzing relative motion within groups of trackable moving point objects. In M. Egenhofer et al. (Eds.) *Proc. Conference on Geographic Information Science - GIScience 2002*, pp. 132-144, Berlin: Springer, 2002.
10. Patterson, D., Liao, L., Gajos, K., Collier, M., Licvic, N. Olson, K., Wang, S., Fox, D., & Kautz, H. Oportunity knocks: A system to provide cognitive assistance with transportation services. In N. Davis et al. (Eds.) *Proc. 6th International Conference on Ubiquitous Computing - UbiComp 2004*, pp. 433-450, Berlin: Springer, 2004.
11. Rao, A. and Georgeff, M. Modeling rational agents within a BDI-architecture. In R. Fikes et al. (Eds.), *Proc. of Knowledge Representation and Reasoning – KRR 1991,* pp. 473-484. San Mateo, CA: Morgan Kaufmann Publishers, 1991.
12. Schlieder, C., & Werner, A. Interpretation of Intentional Behavior in Spatial Partonomies. In Freksa et al (eds.) *Spatial Cognition III*, pp. 401-414, Berlin: Springer, 2003.

Processes and Events in Dynamic Geo-Networks

Antony Galton[1] and Michael Worboys[2]

[1] School of Engineering, Computer Science and Mathematics,
University of Exeter, Exeter EX4 4QF, UK
A.P.Galton@exeter.ac.uk

[2] National Center for Geographic Information and Analysis,
University of Maine, Orono, ME 04469, USA
worboys@spatial.maine.edu

Abstract. Traditional spatial information systems hold only a single state of the 'real world'. However, geographic phenomena have not only static but dynamic characteristics. The work described in this paper contributes to the general research effort toward a generic ontology of dynamic geographic-scale phenomena and its application to the provision of modeling, analysis, and retrieval of data in a spatio-temporal GIS. These issues are addressed in this paper with reference to *dynamic geo-networks*, that is, networks embedded in a (2-dimensional) geographic space. After an introductory and motivational section, the basic ontological categories of events and states are discussed. The paper develops these ideas in the context of flows in dynamic geo-networks, and goes on to discuss the possible kinds of causal relations. The paper concludes with an overview of the results and pointers to further research directions.

1 Introduction

Traditional spatial information systems hold only a single state of the 'real world'. This state is almost always the most recent in time for which the data were captured. Interactions with the system therefore are 'timeless', in that only information contained in the single state can be retrieved. However, geographic phenomena have not only static but also dynamic characteristics. Geographic information systems are now beginning to have some temporal functionality, and a *spatio-temporal information system* manages information that is both geospatially and temporally referenced. While truly spatio-temporal information systems are still in the research arena, GISs are beginning to be extended so that they can offer some practical temporal functionality.

There are many potential application domains for spatio-temporal systems, including environmental change monitoring, transportation, socioeconomic and demographic applications, health and epidemiology, multimedia, governance and administration, crisis management, and defense. In addition to these more traditional spatio-temporal application areas, the increased use of real-time, mobile and *in situ* sensors is leading to many new potential applications for spatio-temporal data models and systems.

M.A. Rodríguez et al. (Eds.): GeoS 2005, LNCS 3799, pp. 45–59, 2005.

Spatiotemporal database research can be divided into two broad categories: that dealing with change (e.g. administrative boundary evolution, environmental change) and that dealing with moving objects (vehicles, ships, people). Although movement is a kind of change, focus on movement of objects leads to a rather different emphasis in terms of the research issues. For example, with movement, themes such as construction of suitable data structures for trajectories and prediction of future positions become dominant. A recent European project focused on movement is Chorochronos [14]. The emphasis of the present work is on more general change.

The work described in this paper contributes to the general research effort toward a generic ontology of dynamic geographic-scale phenomena and its application to the provision of modeling, analysis, and retrieval of data in a spatio-temporal GIS. There is a parallel here with the early days of static GIS, and the process of construction of basic primitives to form a generic framework around which systems could be developed. While conceptual modeling methodologies such as entity-relationship analysis [5], extensions to the relational database model [6], semantic data modeling [4,17], and the object-oriented approach [2,19,3] formed the basis for all information systems, GIS needed geometric and topological types and operators that were sufficiently universal to be deployed by a diverse range of geospatial systems, but specialized enough to be useful for the GIS community. These were developed first by the GIS research community as a collection of generic spatial data types (e.g., [15,8]), or collections of classes and operations in relational [21] and object-oriented settings [7,23,22].

2 Conceptual Models for Spatio-temporal Information Systems and Motivation for the Research

In this section we briefly review examples of conceptual modeling approaches for spatio-temporal information systems, indicate their limitations, and use these to motivate our own work. While there has been considerable progress in the underlying data structures and indexes for temporal, and to a lesser extent, spatio-temporal databases (for a review, see [1]), there has been much slower progress with general methods for the conceptual modeling of temporal and spatio-temporal phenomena.

As an example, consider the temporal extended entity relationship (TEER) model [10]. TEER does not add new syntactical constructs to the extended entity relationship model [9] but extends the meaning of existing constructs by adding a temporal dimension. In the TEER model, each entity has associated with it a temporal element that gives the lifespan of the entity. This lifespan can be a single time interval or a union of disjoint time intervals. Each attribute of an entity is then allowed to vary functionally over the lifespan of the entity (i.e., each attribute has associated with it a function from the entity's lifespan to a range of values). In a similar way, relationships have lifespans associated with them, and each relationship attribute can vary temporally over the relationship lifespan, constrained by the lifespans of the participating entities. Also, subclasses have a constrained temporal relationship with their superclasses. In summary, TEER extends EER by allowing attributes to act as *temporal fluents*, but there is no notation for expressing processes or events.

Another approach to modeling spatio-temporal information is an extension of the object model provided in the TRIPOD project [13] to include the construction of object histories. TRIPOD provides both a model and an implementation for a spatiotemporal database. The underlying data model consists of a collection of temporal snapshots, and takes as primitive the notion of *history*, which models the change of an object, its attributes (also objects), or the relationships that it participates in. Most types can be extended to histories, which provide a sequence of instant or interval timestamped values of the type, thus showing the evolution through time of the instances of the object, attribute, or relationship. TRIPOD moves beyond TEER by providing not just a conceptual model, but also a logical model and implementation, but again there is no notation for explicitly expressing processes or events.

Previous models therefore take objects as the fundamental components of the system, and model change by considering time-varying properties of these objects. For example, a transportation system would be constructed around roads and their time-varying attributes, such as traffic flow rates. The limitation of these approaches is the absence of explicit ways of talking about events and processes, except in so far as events and process affect attributes of objects. The only models which approach the modeling of events in a spatio-temporal system are concerned with the creation, mutation, and destruction of objects [16], and so still have an object bias.

The research question addressed by this paper is how to provide for processes and events in conceptual models of geo-phenomena. This question is addressed in this paper with reference to *dynamic geo-networks*, that is, networks embedded in a (2-dimensional) geographic space. We have chosen to work with networks because although embedded in 2-space and exhibiting many of the properties and behavior of fully spatial structures, they are essentially one-dimensional, and therefore have a simpler structure than more general 2-space entities.

The kinds of application domains we are abstracting from include transportation, utilities, and communication networks. In all these domains, the underlying framework is a network along whose links there are flows. We assume that at a given time, within each link the flow is homogeneous, although we will extend this later to account for "seepage" along a link. In each such application domain, specific theories are well advanced, for example, in the field of transportation, theories seek to describe interactions between the mobile components, e.g., vehicles, vehicle operators, and pedestrians, and the static infrastructure, the highway and its control devices, such as signage, lights, and markings (for example, see [11]). However, what is currently missing is a generic approach to the modeling of the dynamic aspects of geo-networks that provide means of representing not only objects and their temporally varying properties, but also the processes and events in which they participate. The definitions of objects, processes, and events and relationships between these entities provides a rich framework upon which dynamic applications can be modeled.

3 Processes and Events

It is generally agreed that the key concepts required for modeling dynamic phenomena include *object*, *state*, *process* and *event*, but there is little consensus as to exactly how these should be defined. The view taken in this paper is as follows.

Objects and processes have in common the possibility of undergoing change. Change in an object is the most familiar kind, as when, say, a set of traffic lights show red at one time and green at a later time. But a process can also undergo change: for example, the flow of traffic over a bridge is a process, and this can be fast at some times and slow at others. Thus it makes sense to speak of the current *state* of an object or process.

Events differ very significantly in this respect from both objects and processes: it does not make sense to speak of an event as undergoing change. This is because an event is a completed episode of history, and the properties it has, it possesses timelessly. For example, a road accident occurred at a certain location, at a certain time, involving particular cars which were damaged in particular ways. All of these are properties which the event has without temporal qualification; it does not make sense to say that any of them change, and the same goes for all genuine properties of an event.[1]

Formally, objects and processes are distinguished from events by the fact that whereas the attributes of objects and processes include both variable (i.e., time-dependent) and constant (i.e., time-independent) properties, the attributes of events are all time-independent. A time-independent attribute may take one of the forms $P(x)$ or $f(x) = v$; here P is a predicate applying to an entity x, and f is a function whose value for entity x is v. Note that time does not feature as an argument in these expressions. A time-dependent attribute, on the other hand, takes one of the forms $P(x, t)$ and $f(x, t) = v$, in which the holding or not holding of predicate P when applied to x, as well as the value of function f for argument x, may vary according to the time given as the second argument. Some examples to illustrate these ideas are:

Time-independent properties of event *accident37*:

$$location(accident37) = intersection(king\_street, union\_street)$$
$$Involves(accident37, car45)$$
$$time(accident37) = t123$$
$$type(accident37) = collision$$

Time-independent properties of processes:

$$type(roadworks12) = pipelaying$$
$$direction(traffic\_flow335) = northeast$$

[1] By 'genuine' property we mean intrinsic as opposed to relational properties. Of course an event can 'change' from being the most recent car accident to the second most recent one, but this is a change in the relation of the event to the present: it changes because the present changes, not because the event does.

Time-dependent properties of processes:

$location(roadworks12, t) = king\_street.section15$
$speed(traffic\_flow335, t) = 37mph$
$phase(traffic\_light\_cycle85, t) = phase3$

The general picture we are painting is as follows. Objects, states, and processes belong to the 'snapshot' view of the world, the world at one time. We can look at the snapshot and say: here we see such-and-such objects, in such-and-such states, undergoing such-and-such processes. Snapshots are continually renewed as time passes, i.e., the snapshot differs from one moment to the next, and this means that the elements present in a snapshot must be capable of undergoing change. By contrast, events belong to the fixed historical record. This is not renewed in the sense of being replaced by a new record, but only in the sense of being incrementally added to as time passes: as events occur, they are added to the record, but once there they are fixed for all time.[2]

The events that are added to the record are generated by the processes that exist in the continually evolving snapshot. For example, if a process p comes into existence (meaning that it is present in a snapshot for the first time), then the event 'process p started' is added to the record. When the process stops, we get the event 'process p stopped', and at that point we can also add to the record the complete episode which consists of the lifetime of p from its inception to its termination, whose attributes include such things as its duration and the magnitude of the resulting change (e.g., if p is a motion process, the distance traveled).

The distinction between the evolving sequence of snapshots and the fixed historical record is reminiscent of the SNAP/SPAN ontologies of Grenon and Smith [12], with the crucial difference that we regard processes as belonging in the snapshots whereas in the SNAP/SPAN framework they are placed in the SPAN ontology. For us, a snapshot is something dynamic in that it incorporates ongoing processes (states of change), whereas a SNAP ontology, as we understand it, only incorporates the static properties of the world at one time.

4 Networks and Flows

In this paper a network is defined as a directed multigraph embedded in a surface, with flows associated with the links. The elements of this definition are explained as follows.

A *directed multigraph* is a set of nodes and links such that each link is associated with an ordered pair of nodes; it is a multigraph because we allow more than one link to be associated with the same ordered pair of nodes. Formally, we have a triple $\langle N, L, nodes\rangle$, where $nodes : L \rightarrow N \times N$ is a

[2] Do not confuse this with the fact that *our* records may undergo changes, e.g., through the correction of errors. By the 'fixed' historical record we mean the ideally correct record of what actually happens.

function associating node-pairs with links (intuitively, $nodes(l) = \langle n_1, n_2 \rangle$ tells us that link l connects node n_1 to node n_2). We shall also use the notation $nodes(l) = \langle innode(l), outnode(l) \rangle$.

A directed multigraph is an abstract entity which is not inherently spatial in nature. To 'spatialize' it, we must consider it as embedded in space, which in geographical contexts usually means the two-dimensional space of the earth's surface or some delimited portion thereof. Let S be the space in question, then the embedding is specified by a function loc which (1) maps each node of the multigraph onto a location in S, and (2) maps each link l of the multigraph onto a curve segment whose endpoints are at $loc(innode(l))$ and $loc(outnode(l))$. When we speak of the length of a link, we mean, of course, the length of this curve segment.

The spatially embedded multigraph is a purely static entity. To make it into a dynamic network, we add to each link one or more *flows*, where a flow is conceived as an ongoing movement along the link from its in-node to its out-node—a process in the sense that this term was used earlier. We do not specify exactly what it is that is moving: it may be discrete objects such as vehicles, or it may be some fluid such as water or blood. We do, however, assume that there is some measure of *amount* for whatever is moving through the network, in terms of which we can speak of its density and rate of flow. We require the possibility of associating more than one flow with a given link, so that, for example, we can distinguish between the flow of trucks, the flow of cars, and the flow of bicycles along a given stretch of road: each of these flows has a different *flow-type*.

The flow of type T associated with link l will be denoted $l.flow(T)$. This is a process possessing time-varying attributes such as:

– *speed*: the speed at which the flow passes through the link (measured in distance/time).
– *density*: the amount of flow per unit distance along the link (measured in quantity/distance — the exact meaning of 'quantity' will depend on the nature of the flow, e.g., number of vehicles, volume of water).
– *throughput*: the amount of flow passing through the link in unit time (measured in quantity/time).

The syntax of these attributes is illustrated by $speed(l.flow, t) = s$. Clearly $throughput(l.flow(T), t) = speed(l.flow(T), t) \times density(l.flow(T), t)$.

Here we have assumed that the flow attributes are uniform along the length of a link. In reality this may not be the case for two reasons:

– The speed and density of the flow may vary along the length of the link (e.g., in a stretch of road between two sets of traffic signals, the flow near the ends is typically both slower and denser than in the central portion, even if their product, the throughput, is uniform);
– A link may allow *seepage*, i.e., the loss or gain of flow elements *along* the link (e.g., water leaking from a water pipe is a loss; but cars entering a residential street from private driveways situated along its length represents a gain).

To handle seepage adequately we should replace *throughput* by *input* and *output*, related by

$$output(l.flow(T), t) = input(l.flow(T), t) + seepage(l.flow(T), t).$$

(Here seepage *into* the flow is regarded as positive.)

If we want to retain the idealized assumption of uniformity of flow along a link, and still allow for seepage, this can be done by introducing an extra notional node in the middle of the link which can function as a sink or source: then all the seepage is regarded as concentrated at that node, and manifests itself as the difference between the throughputs of its incoming and outgoing flows (i.e., the two links into which the original link was divided).

A flow along a link will generally have a maximum possible throughput, determined by the capacity of the link for that type of flow and general considerations relating to the nature of the flow elements. For example, in a road, there will be a maximum speed defined by the speed limit (in practice rather higher than this!), and at a given speed, there will be a maximum density determined by the average separation of vehicles traveling at that speed (in practice rather less than the officially sanctioned braking distances). Where there is more than one type of flow along a link, these will be as it were in competition with each other for the capacity afforded by the link; the details of this will in general be complicated, and we defer consideration of this to a later occasion.

Nodes may be classified according to the number of incident links. Exceptionally, we might allow a node with only one link. If there is to be a non-zero flow along the link then the node must act as a sink or a source depending on the direction of flow. If there is flow both towards and away from the node, then there must be two separate links to carry those flows (even if, physically, they occupy the same channel, e.g., a narrow road carrying two-way traffic).

If two or more links are mapped to the same pair of nodes (i.e., $nodes(l_1) = nodes(l_2) = \ldots$), then we shall call them *parallel* links. An example would be in a road network, where there are two or more lanes in the northbound direction between one junction and the next; each of these lanes would be modelled as a separate link, and those links will all be parallel. In general, there will be seepage between these parallel links. Likewise, if $nodes(l_1) = \langle n_1, n_2 \rangle$ and $nodes(l_2) = \langle n_2, n_1 \rangle$ we shall say that l_1 and l_2 are *antiparallel* links. A maximal set of links any two of which are either parallel or antiparallel will be called a *linkage*. A linkage incorporates all the links in either direction between a given pair of nodes.

A node which is incident to exactly two linkages corresponds to a control point in the network, that is, some device or circumstance which has the effect of controlling the flow attributes on the incident links. An example would be the start of a speed-limit zone in a road network. Since the flow attributes on either side of the speed-limit sign are different, we must place a node at that point to maintain the uniform flow condition on individual links.

Nodes incident to three or more linkages correspond to junctions or intersections. Assuming there is no seepage at the node itself, then we can say that the sum of the outputs of a given type from all its incoming links is equal to the sum of the inputs of that type into all its outgoing links.

If we examine the flow *at a node* in detail, we can see that it can be quite complicated. For example, let the incoming links be $l_1, l_2, \ldots, l_m$, and the outgoing links $l'_1, l'_2, \ldots, l'_n$. Then each pair $\langle l_i, l'_j \rangle$ represents a potential component of the flow, i.e., a flow out of l_i and into l'_j. At a four-way road intersection, for example, with incoming links $l_1, \ldots, l_4$ and outgoing links $l'_1, \ldots, l'_4$, where l_i and l'_i are antiparallel, then the flow of any given type across the intersection actually has twelve components

$$\langle l_1, l'_2 \rangle, \langle l_1, l'_3 \rangle, \langle l_1, l'_4 \rangle, \langle l_2, l'_1 \rangle, \langle l_2, l'_3 \rangle, \langle l_2, l'_4 \rangle,$$
$$\langle l_3, l'_1 \rangle, \langle l_3, l'_2 \rangle, \langle l_3, l'_4 \rangle, \langle l_4, l'_1 \rangle, \langle l_4, l'_2 \rangle, \langle l_4, l'_3 \rangle$$

If the traffic is dense, it will not be possible for all these flows to proceed simultaneously without collisions occurring. The purpose of traffic lights is to successively enable various subsets of these twelve flows in turn, in such a way that no flows that cross each other are enabled together.

5 Events in a Dynamic Network

We may distinguish three broad classes of events in networks.

- Changes to the structure of the network itself, for example
  - the introduction of a new link
  - the removal (or permanent closure) of a link
  - the creation of a new node dividing an existing link or linkage into two
  - the creation of a new node unconnected with any existing link
  - the removal of a node and all its incident links
  - the removal of a node consequent on a one-to-one merger of its ingoing links with its outgoing links

  After any of these changes, we are in a sense dealing with a different network. Such changes can be expected, in general, to have an impact on the flows even in those parts of the network which have not been changed.
- Changes which do not affect the structure of the network itself but which may affect the flows in the networks; for example
  - introduction of an obstruction in a link, reducing flow along that link
  - temporary closure of a link
  - removal of an obstruction
  - reopening a temporarily closed link
  - introduction of an obstruction at a node which reduces or blocks one or more flow components through that node
- Finally, there are *flow events*, the changes that occur in the flows themselves (often as a result of events of one of the previous two kinds).
  - creation of a new flow on a link
  - removal of a flow from a link
  - a flow comes to a halt (i.e., the speed becomes zero)

- a flow 'dries up' (i.e., the density becomes zero)
- a flow starts up again from a halted state
- a flow starts up again from a dried-up state

Here we may also think about increases or decreases in flow attributes such as speed, density, and throughput. These should be regarded as processes rather than events; but temporally bounded episodes of such increases or decreases may be singled out as events, for example a sudden decrease resulting from the introduction of an obstacle into a link. In particular, we may be interested in peaks and troughs, i.e., events in which some flow attribute attains a local maximum or minimum value. A peak will be preceded by an episode of increasing value and followed by an episode of decreasing value, and *vice versa* in the case of troughs.

6 Causal Relations Amongst Events, States and Processes

In modeling the evolution of a dynamic system such as one of our networks, it is hard to avoid bringing in the notion of causality. An accident on the road causes an obstruction which causes reduced traffic flow. The accident is an event, the obstruction is an object or a state, and the reduced traffic flow is a state or a process. It seems that causal relations can exist between entities of various different types. In this section we attempt to provide a clear account of these causal relations which will, we hope, be adequate for the purpose of modeling network phenomena.

It should be noted first that 'causes' is not by any means the only causal relation we are interested in. Worboys and Hornsby [24] list the following event-event relationships: initiation, perpetuation (or facilitation), hindrance (or blocking), and termination. They also list event-object relationships: creation, sustaining in being, reinforcement, degradation, destruction, splitting and merging. Our intention here is to examine some of these relationships more closely, particularly from the point of view of events, states, and processes in networks, and to derive a more systematic classification of the most important relationships.

Consider the following scenario: at time t_1, the traffic flow at a particular position P on a road network is *high*; at time t_2, a little later than t_1, an accident occurs near P, resulting in an obstruction in the road; at time t_3, a little later than t_2, the traffic flow at P is *low*.

A close analysis of this scenario reveals the following events:

- E_1 is the occurrence of the accident.
- E_2 is the road's becoming obstructed.
- E_3 is the reduction of the flow at P from *high* to *low*.

Each of these events occurs at a definite time; for simplicity we shall assume they all occur at t_2 (this is a matter of choosing the temporal granularity of the representation appropriately). In addition, we can identify the following states:

- S_1 is the state of the road being clear
- S_2 is the state of the traffic flow being high
- S_3 is the state of the road being obstructed
- S_4 is the state of the traffic flow being low

These states are all time-dependent, that is, they will hold at some times and not at others. For example, S_1 holds over the interval (t_1, t_2) whereas S_2 holds over the interval (t_2, t_3). Note that a state may be the state of an object (e.g., the road) or of a process (e.g., the traffic flow).

What are the causal relations amongst these various states and events? We advocate the following interpretation:

- State S_1 *enables* S_2. (Note that it would be incorrect to say that it *causes* it: the road being clear cannot itself generate the traffic required for high flow!)
- States S_1 and S_2 are *terminated by* events E_2 and E_3 respectively. This is essentially a matter of definition: a state S is naturally terminated by an event E which is defined as the coming into existence of a state incompatible with S.
- Event E_1 *causes* event E_2, which in turn causes event E_3.
- Events E_2 and E_3 *initiate* states S_3 and S_4 respectively. This is again a matter of definition.
- State S_3 *perpetuates* (or *maintains*) state S_4.

Figure 1 shows these relations in diagrammatic form, using rectangular boxes to represent states, and circles to represent events.

From this we can see that each of the causal relations is specific to particular combinations of types, as follows:

- Event-Event: *causes*
- Event-State: *initiates*, *terminates*
- State-State: *enables*, *perpetuates*

In addition, we may introduce

- State-Event: *allows*, *prevents*

These causal relations, shown diagramatically in figure 2, may be explained as follows. Here, for simplicity, we assume that all events are *punctual*, that is, they occur at an instant rather than over an interval; the explanations would need to be adjusted to allow for events which take place over an interval, i.e., *durative* events.

- 'E_i causes E_j' means that E_j occurs as a result of E_i's occurring. We do not attempt to define 'as a result of'—see the voluminous philosophical literature on this subject (for example, [20]).
- 'E initiates S' means that if E occurs at t then (1) S must hold over an interval beginning at t, and (2) there is an interval ending at t throughout which S does not hold.

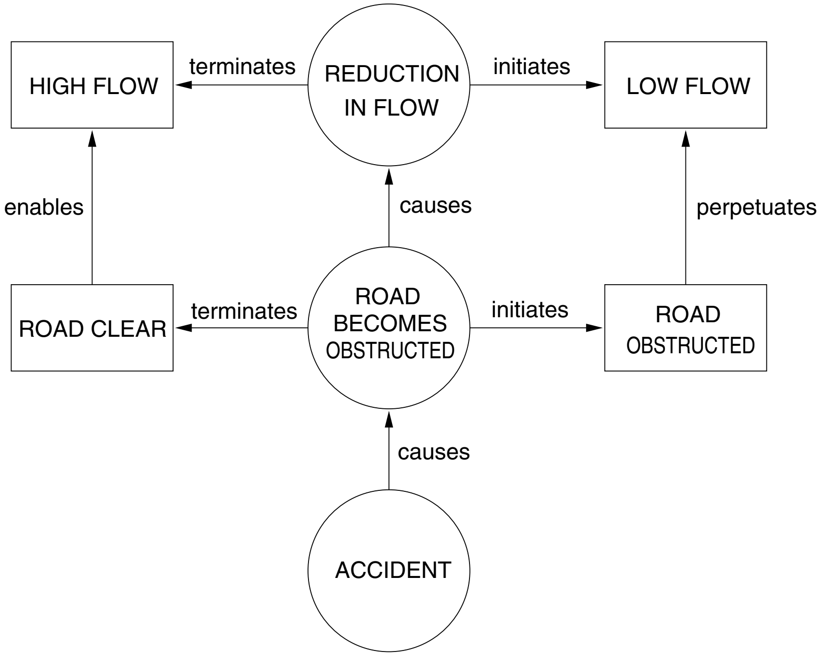


Fig. 1. Diagram of event–state relationships

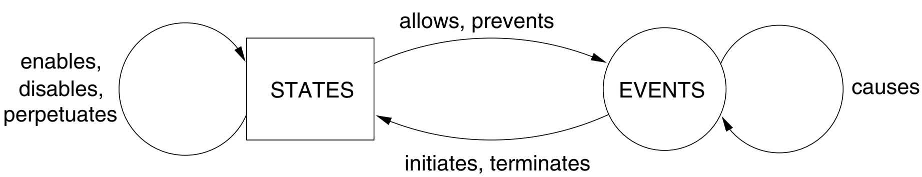


Fig. 2. Kinds of causal relationships between states and events

- 'E terminates S' means that if E occurs at t then (1) S holds over an interval ending at t, and (2) there is an interval beginning at t throughout which S does not hold.
- 'S_i enables S_j' means that if S_i holds at t then it is possible for S_j to hold at t.
- 'S_i disables S_j' means that if S_i holds at t then it is *not* possible for S_j to hold at t.
- 'S_i perpetuates S_j' means that if S_i and S_j both hold at t, and S_i holds throughout an interval i which starts at t, then S_j will also hold throughout i.
- 'S allows E' means that if S holds at t then it is possible for E to occur at t.
- 'S prevents E' means that if S holds at t then it is not possible for E to occur at t.

It should be emphasized that we are not here making hypotheses about the way in which English words such as 'prevents' and 'enables' are actually used; rather we are stipulating how they *should* be used in the technical context of modeling dynamic networks in an information system. The terms have, however, been chosen so as to accord as nearly as possible to normal English usage.

These type-restrictions may seem to be too narrow: it is easy to come up with examples which are *prima facie* exceptions to them. However, we believe that in every case, a careful analysis will show that our type-restrictions are valid. We illustrate this as follows:

1. Can a state cause an event? Example: the presence of ice on the road caused the car to skid. In fact the presence of ice cannot possibly *cause* the car to skid; the immediate cause must be some maneuver by the driver, e.g., braking. Thus the true picture is that the state of iciness *allowed* the skidding event, and the braking event *caused* it. Of course, we could say that the true cause of the skidding is the conjunction of the ice on the road and the braking; but we believe that our account is simpler and entirely adequate: an event causes another event *in the context of* a state which allows it.
2. Can one event prevent another event? Example: The accident prevented John from arriving at the station on time. This statement leaves out a number of intermediate states and events which are required for a full analysis. The correct picture is as follows:
   - The accident (event) *causes* the road to become obstructed (event)
   - The road's becoming obstructed (event) *causes* the reduction in the traffic flow (event)
   - The road's becoming obstructed (event) *initiates* the road's being obstructed (state)
   - The reduction in the traffic flow (event) *initiates* the low traffic flow (state)
   - The road's being obstructed (state) *perpetuates* the low traffic flow (state)
   - The low traffic flow (state) *prevents* John's timely arrival (event)
3. Finally, can a state cause a state? Example: The presence of ice on the road caused the road to be dangerous. Again, a fuller analysis reveals what is going on here:
   - The road's becoming icy (event)[3] *causes* the road to become dangerous (event)
   - The road's becoming icy (event) *initiates* the road's being icy (state)

[3] 'The road's becoming icy' perhaps refers ambiguously to either an event or a process. Here we mean the event, i.e., the completed transition from a state in which the road is not icy to a state in which the road is icy. This transition may be accomplished by means of a process of incremental accumulation of ice on the road surface; and this process might also be described as the road's becoming icy.

– The road's becoming dangerous (event) *initiates* the road's being dangerous (state)
– The road's being icy (state) *perpetuates* the road's being dangerous (state)

The reader may find it helpful to draw diagrams for each of these examples, in a similar manner to Figure 1.

7 Conclusions and Further Work

In this paper we have attempted to identify some of the key general concepts which will be needed to underpin any truly event- and process-oriented model of dynamic geospatial systems. We have concentrated on networks as essentially one-dimensional entities embedded in a two-dimensional space, and have proposed a three-tier representation of these in terms of (1) an abstract graph-like structure, (2) a spatial embedding of this, and (3) a system of flows utilizing the embedded structure.

We then turned our attention to the events which may occur in networks of this kind, carefully distinguishing events, which are delimited episodes in the unchanging historical record, from processes, which are dynamic phenomena existing in the present and subject to change as time passes. We classified network events into changes in the structure of the network itself, temporary disruptions to the flow-carrying capacity of the network, and changes in the flows. An important issue here is how to represent the causal dependencies amongst the various events that can occur and the time-varying states of the objects and processes present in the network. We proposed a set of terms for the various causal relations required here, which we hope may become accepted as a standard terminology in this area.

While we believe that the above work represents a real advance in our understanding of the dynamic aspects of spatial networks, much remains to be done. In particular, the concepts introduced informally in this paper will need to be formalized rigorously in a mathematical (logical or set-theoretic) setting. Then further studies will be needed to show how the general ideas developed here can be applied to specific domains such as traffic (specifically the Integrated Transport Network developed by the Ordnance Survey of Great Britain [18]) or communications systems. Finally, it is evident that some of these ideas will be applicable to wider geospatial domains, beyond our specific focus on networks, and there is much scope for further work in exploring these possibilities.

Acknowledgments

Antony Galton's work is supported by the Engineering and Physical Sciences Research Council of Great Britain under EPSRC grant number EP/C014502/1. Mike Worboys' work is supported by the National Science Foundation under NSF grant numbers IIS-0429644, BCS-0327615, and the Ordnance Survey of Great Britain.

References

1. T. Abraham and J. Roddick. Survey of spatio-temporal databases. *GeoInformatica*, 3:61–69, 1999.
2. G. Booch. *Object-Oriented Analysis and Design with Applications*. Addison-Wesley, Reading, MA, 2nd edition, 1993.
3. G. Booch, J. Rumbaugh, and I. Jacobson. *The Unified Modeling Language User Guide*. Addison-Wesley, Reading, MA, 1998.
4. M. Brodie, J. Mylopoulos, and J. Schmidt, editors. *On Conceptual Modeling*. Springer, Berlin, 1984.
5. P. P-S. Chen. The entity-relationship model—toward a unified view of data. *ACM Transactions on Database Systems*, 1(1):9–36, 1976.
6. E. Codd. Extending the relational database model to capture more meaning. *ACM Transactions on Database Systems*, 4(4):397–434, 1979.
7. M. J. Egenhofer and A. Frank. Object-oriented modeling for GIS. *URISA Journal*, 4:3–19, 1992.
8. M. J. Egenhofer and J. R. Herring. A mathematical framework for the definition of topological relationships. In K. Brassel and H. Kishimoto, editors, *Proc. Fourth International Symposium on Spatial Data Handling*, pages 803–813, Zurich, 1990.
9. R. Elmasri and S.B. Navathe. *Fundamentals of Database Systems*. The Benjamin Cummings Publishing Company, 3 edition, 2000.
10. R. Elmasri and G. Wuu. A Temporal Model and Query Language for ER databases. In *Proceedings of the Sixth International Conference on Data Engineering*, pages 78–83, 1990.
11. N.H. Gartner, C.J. Messer, and A. Rathi. Traffic flow theory: A state of the art report. Technical report, US Department of Transportation, Federal Highway Administration, www.tfhrc.gov/its/tft/tft.htm.
12. P. Grenon and B. Smith. SNAP and SPAN: Towards dynamic spatial ontology. *Spatial Cognition and Computation*, 4(1):69–104, 2004.
13. T. Griffiths, A. Fernandes, N. Paton, S.-H. Jeong, N. Djafri, K.T. Mason, B. Huang, and M. Worboys. Tripod: A spatio-historical object database system. In R. Ladner, K. Shaw, and L. Abdelguerfi, editors, *Mining Spatio-Temporal Information Systems*, pages 127–146. Amsterdam: Kluwer, 2002.
14. R. H. Güting, M. Böhlen, C. Jensen, N. Lorentzos, M. Schneider, and M. Vazirgiannis. A foundation for representing and querying moving objects. *ACM Transactions on Database Systems*, 25(1):881–901, 2000.
15. R. H. Güting and M. Schneider. Realm-based spatial data types: The ROSE algebra. *VLDB journal*, 4(2):243–286, 1995.
16. K. Hornsby and M.J. Egenhofer. Identity-based change: A foundation for spatio-temporal knowledge representation. *International Journal of Geographical Information Science*, 14(3):207–204, 2000.
17. R. Hull and R. King. Semantic data modeling: Survey, applications, and research issues. *ACM Computer Surveys*, 19:201–260, 1987.
18. Ordnance Survey of Great Britain. *OS MasterMap User Guide*. Romsey Road, Southampton, UK, 2005.
19. J. M. Smith and D. C. P. Smith. Database abstractions: Aggregation and generalization. *ACM Transactions on Database Systems*, 2(2):105–133, 1977.
20. E. Sosa and M. Tooley, editors. *Causation*. Oxford Readings in Philosophy. Oxford University Press, 1993.

21. J.W. van Roessel. Design of a spatial data structure using the relational normal forms. *International Journal of Geographical Information Systems*, 1:33–50, 1987.
22. M.F. Worboys. Object-oriented approaches to geo-referenced information. *International Journal of Geographical Information Systems*, 8(4):385–399, 1994.
23. M.F. Worboys, H. M. Hearnshaw, and D. J. Maguire. Object-oriented data modelling for spatial databases. *International Journal of Geographical Information Systems*, 4(4):369–383, 1990.
24. M.F. Worboys and K. Hornsby. From objects to events: GEM, the geospatial event model. In M. Egenhofer, C. Freksa, and H. Miller, editors, *Third International Conference on GIScience*, volume 3234 of *Lecture Notes in Computer Science*, pages 327–344. Berlin: Springer Verlag, 2004.

A Qualitative Trajectory Calculus and the Composition of Its Relations

Nico Van de Weghe[1], Bart Kuijpers[2], Peter Bogaert[1], and Philippe De Maeyer[1]

[1] Department of Geography, Ghent University,
Krijgslaan 281 (S8), B-9000 Ghent, Belgium
{nico.vandeweghe, peter.bogaert, philippe.demaeyer}@ugent.be
[2] Theoretical Computer Science Group, Hasselt University,
Agoralaan, Gebouw D, B-3590 Diepenbeek, Belgium
bart.kuijpers@uhasselt.be

Abstract. Continuously moving objects are prevalent in many domains. Although there have been attempts to combine both spatial and temporal relationships from a reasoning, a database, as well as from a logical perspective, the question remains how to describe motion adequately within a qualitative calculus. In this paper, a Qualitative Trajectory Calculus (QTC) for representing and reasoning about moving objects in two dimensions is presented. Specific attention is given to a central concept in qualitative reasoning, namely the composition of relations. The so-called composition-rule table is presented, which is a neat way of representing a composition table. The usefulness of QTC and the composition-rule table is illustrated by an example.

1 Introduction

In the last two decades, spatial and temporal information have received significant attention in various fields of research, ranging from geography and geometry to artificial intelligence and computer science. Qualitative calculi have been proposed, both in the temporal (e.g. [1],[2]) and the spatial (for an overview: see [3]) domain. The mentioned formalisms are especially suited to express relationships between entities. This type of formalism has gained wide acceptance as a useful way of abstracting from the real world. Only in recent years, the attention has extended to applications that involve spatio-temporal data. Nevertheless, both from the database area [e.g. 4-8] as from the research domain of qualitative reasoning [e.g. 9-13] movements of objects have been studied.

In the widely used Region Connection Calculus (RCC) [14] and 9-Intersection Model [15], topological relationships between two regions are defined. Apart from some limiting cases, such as a car accident and a predator catching a prey, where moving objects *meet*, mobile objects are represented by use of *disconnected from* (DC) in RCC and *disjoint* in the 9-Intersection Model. So, a limitation of these formalisms is that all DC relations are undifferentiated. This approach ignores some important aspects of reasoning about continuously moving physical objects. For example, given two trains on a railroad, it is of the utmost importance to know their movement with respect to each other in order to detect whether or not they would crash in the near future. Therefore, a challenging question remains: "how do we

M.A. Rodríguez et al. (Eds.): GeoS 2005, LNCS 3799, pp. 60 – 76, 2005.

handle changes in movement between moving objects, if there is no change in their topological relationship?" With this in mind, the Qualitative Trajectory Calculus (QTC) is presented in [16]. QTC is a language for representing and reasoning about movements of objects in a qualitative framework, able to differentiate between groups of disconnected objects. In this paper, we specifically study the 81 relations of the so-called QTC Double-Cross, or QTC_C for short.[1] This calculus is partly based on the Double-Cross Calculus introduced by Freksa and Zimmermann [20]. We discuss the reasoning power of QTC_C and apply the important reasoning technique of composition tables, originating from the domain of temporal reasoning [1]. Since a composition table encodes all possible compositions of relations for a specific calculus, a simple table look-up operation can replace complex theorem proving [21]. This is why composition tables are very useful from a computational point of view [22,23]. Besides the simple look-up mechanisms, composition tables play an important role when working with incomplete information and larger inference mechanisms as exemplified in Section 4. It is not surprising that composition tables have found their way in the domain of qualitative spatial reasoning [24-27]. As composition table look-up forms an integral part of temporal and spatial reasoning calculi [28], it will have its importance in spatio-temporal reasoning, and thus when studying moving objects. In this paper, instead of the full composition tables for QTC_C, composition rules to generate composition tables are presented. These rules can be implemented in information systems in order to generate composition tables automatically, which is highly preferable due to the extent of the tables; the 81 QTC_C-relations generate a matrix composed of 6561 (81×81) entries.

This paper is organized as follows. In Section 2, QTC_C is defined. In Section 3, we discuss the composition-rule table. In Section 4, we show how both QTC_C and the composition-rule table can be used for reasoning with incomplete knowledge about moving objects. Note that we did not intend to present a formal background of the calculus, neither did we intend to make a comparison to other calculi and conceptual approaches dealing with orientation and/or motion, such as [4-13,29-43].[2]

2 QTC Double-Cross (QTC_C)

We assume continuous time for QTC. Depending on the level of detail and the number of spatial dimensions, different types of QTC are defined in [16]. In general, QTC makes comparisons between positions of two objects at different moments in time. The movement of the first object (called *k*) with respect to the second object (called *l*) is studied by comparing the distance between *l* at the current time point (denoted *t*) and *k* during the period immediately before the current time point (denoted t^-), with the distance between *l* at *t* and *k* during the period immediately after the current time point (denoted t^+). In addition, the movement of *l* with respect to *k* is studied by comparing the distance between *k* at *t* and *l* at t^-, with the distance between

[1] For a description and an illustration of how QTC can be extended to movements along (road) networks and how QTC_C has to be used during longer periods containing multiple QTC_C relations, we refer, respectively, to [17, 18], and [19].

[2] For a formal axiomatization of QTC and a confrontation with several other calculi, see [16].

k at t and l at t^+. Each object can move away from or towards the other, or can be stable with respect to the other. These three possibilities result respectively in the qualitative values of +, − and 0. In QTC Basic or QTC_B, only this changing of distance is of importance. The calculus QTC_C considers additionally the direction in which an object is moving with respect to the line segment between the two objects. QTC_C is partly based on the Double-Cross Calculus introduced by Freksa and Zimmermann [20,26,44,45]. Their central research question was: "*Consider a person walking from some point a to point b. On his way, he is observing point c. He wants to relate point c to the vector ab*" [45,p.51]. Freksa and Zimmermann propose a double-cross induced by two reference points: the positions of the observer at t_1 (point *a* in Fig. 1a) and the point where the observer is walking towards (point *b* in Fig. 1a). Through these pinpoints, the reference line (*RL*) is defined. Also through these pinpoints and perpendicular to *RL*, the first perpendicular reference line ($RL\perp1$) and the second perpendicular reference line ($RL\perp2$) are defined. The three lines (*RL*, $RL\perp1$, and $RL\perp2$) form a double-cross and distinguish six 2D regions, six 1D infinite half lines, one 1D line segment between the two reference points, and the two 0D reference points themselves (Fig. 1b). This way, they define a set of fifteen basic relations that can be utilized to navigate using qualitative spatial information. Based on the front/back dichotomy and the left/right dichotomy, the position of the observed point *c* can be described in terms of these fifteen relations. For example, in Fig. 1c, *c* is localized in zone 14. A major goal of this calculus was to find a natural and efficient way to deal with incomplete knowledge, e.g. if it is not possible to decide whether the third point is behind or in front of the second point.

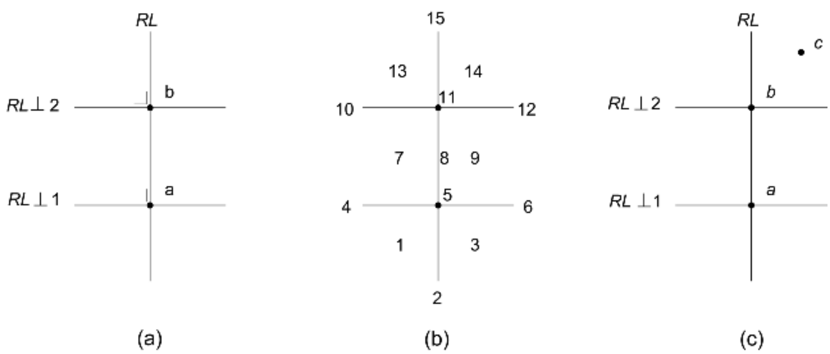


Fig. 1. The Double-Cross Calculus (Source: [44])

Worth mentioning is the difference between the approach of Freksa-Zimmermann and the QTC calculi. In the Double-Cross Calculus (Fig. 2a), the location of a moving point from t_1 ($k|t_1$) to t_2 ($k|t_2$) with respect to a static point ($l|t_1 = l|t_2$) is described. The movement of k results in a vector of which the beginning and the end serve as pinpoints for the double-cross. The double-cross forms the reference frame for the calculus. However, QTC_C (Fig. $2b_1$ and b_2) examines the movement of two objects k and l with respect to each other, between t_1 and t_2. Both movements are represented via a vector

(Fig. $2b_2$), degenerated to a point if an object is not moving (Fig. $2b_1$). The origins of these vectors serve as pinpoints for the double-cross, being the reference frame for the calculus. The Double-Cross Calculus only considers a single movement (Fig. 2a), in which one of both objects is moving. QTC_C supports single movements (Fig. $2b_1$) as well as dual movements in which both objects move (Fig. $2b_2$).

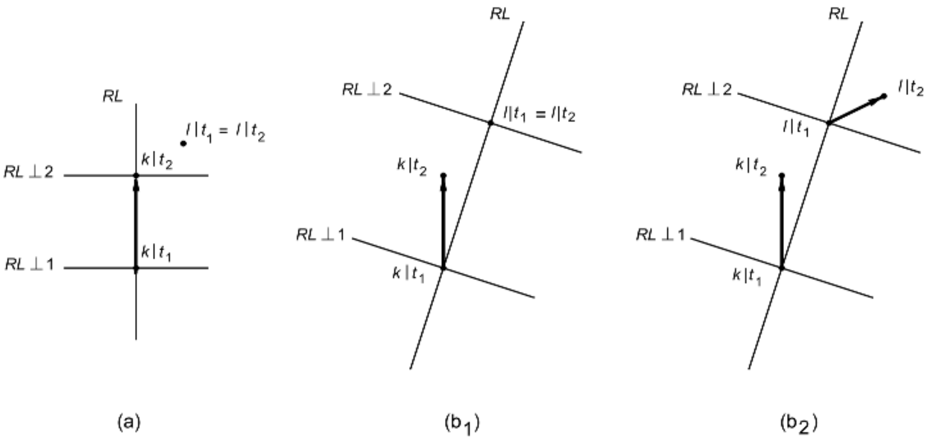


Fig. 2. Double-Cross Calculus versus QTC_C

Hence, a two-dimensional movement is presented in QTC_C using the following conditions (C):

Assume: objects k and l
RL_{kl}: the directed reference line from k to l

C1. Movement of the first object, with respect to the first perpendicular reference line at time point t (distance constraint):

$-$[3]: k is moving towards l:
$\exists t_1\ (t_1 < t \wedge \forall\ t^-\ (t_1 < t^- < t \rightarrow d(k|t^-, l|t) > d(k|t, l|t))) \wedge$
$\exists t_2\ (t < t_2 \wedge \forall\ t^+\ (t < t^+ < t_2 \rightarrow d(k|t, l|t) > d(k|t^+, l|t)))$
+: k is moving away from l:
$\exists t_1\ (t_1 < t \wedge \forall\ t^-\ (t_1 < t^- < t \rightarrow d(k|t^-, l|t) < d(k|t, l|t))) \wedge$
$\exists t_2\ (t < t_2 \wedge \forall\ t^+\ (t < t^+ < t_2 \rightarrow d(k|t, l|t) < d(k|t^+, l|t)))$
0: k is stable with respect to l:
all other cases

C2. The movement of the second object wrt the second perpendicular reference line at time point t can be described as in condition 1 (C1) with k and l interchanged.

C3. Movement of the first object with respect to the directed reference line from k to l at time point t (side constraint):

[3] We write – here, because there is a decrease in distance between both objects. If there is an increase in distance, we write +. If the distance remains the same, we will write 0.

−: k is moving to the left side of RL_{kl}:

$\exists t_1 \ (t_1 < t \wedge \forall\, t^- \ (t_1 < t^- < t \rightarrow k$ is on the right side of RL_{kl} at $t)) \wedge$

$\exists t_2 \ (t < t_2 \wedge \forall\, t^+ \ (t < t^+ < t_2 \rightarrow k$ is on the left side of RL_{kl} at $t))$

+: k is moving to the right side of RL_{kl}:

$\exists t_1 \ (t_1 < t \wedge \forall\, t^- \ (t_1 < t^- < t \rightarrow k$ is on the left side of RL_{kl} at $t)) \wedge$

$\exists t_2 \ (t < t_2 \wedge \forall\, t^+ \ (t < t^+ < t_2 \rightarrow k$ is on the right side of RL_{kl} at $t))$

0: k is moving along RL_{kl}:

all other cases

C4. The movement of the second object wrt the directed reference line from l to k at time point t can be described as in condition 3 (C3) with k and l interchanged.

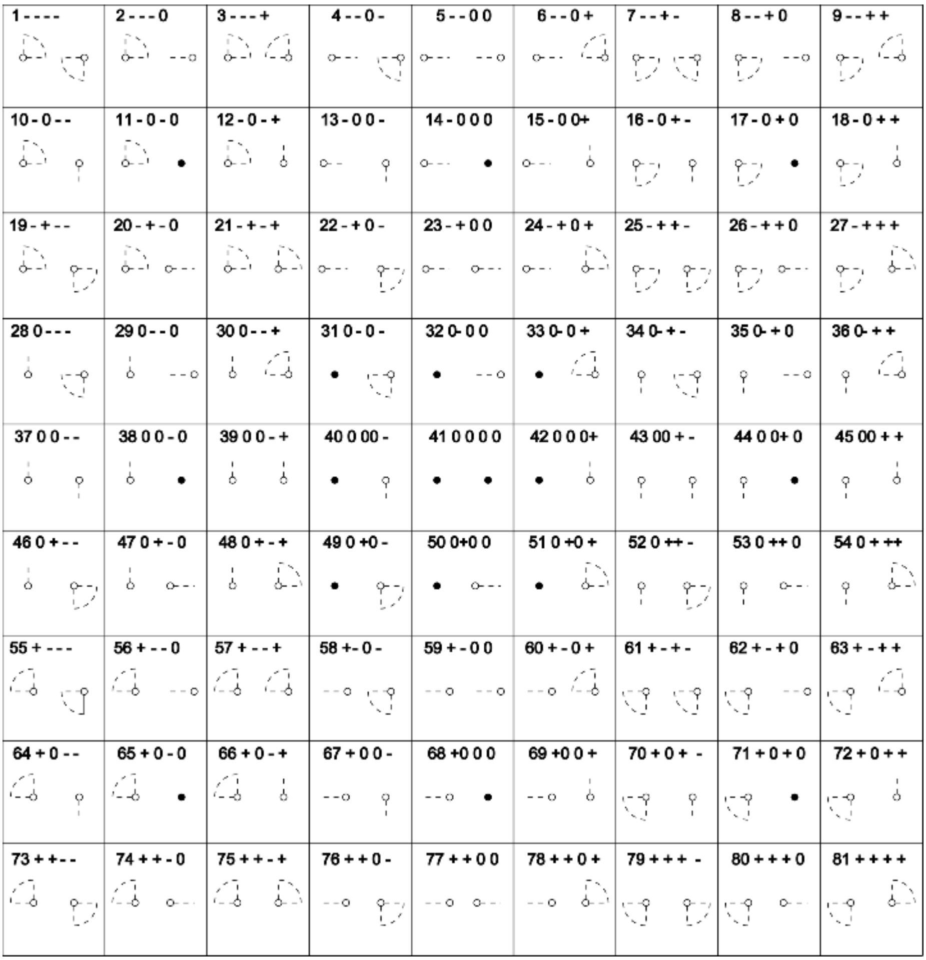


Fig. 3. QTC_c-relation icons

We can represent a trajectory by a label consisting of four characters, each one giving a value respectively for the four conditions above. There are 81 (=3^4) QTC_C-relations. Despite the fact that qualitative reasoning typically deals with a small number of relations, we do believe that the 81 relations form a good foundation for mimicking human reasoning, since they are only based on two constraints resulting in, what we prefer to call, the towards/away-from dichotomy (for the distance constraint) and left/right dichotomy (for the side constraint). Note that in fact QTC_C is a combination of two Double-Cross relations, and that, in [45], Zimmermann and Freksa argue that it has been studied in experimental psychology that humans tend to use rectangular reference systems. As a result, there will be a clear difference in the movements represented by the 81 QTC relations. Each QTC_C-relation can be represented by a so-called relation icon (Fig. 3). The left dot represents the position of k and the right dot the position of l. A dot is filled if the object can be stationary, and open if the object cannot be stationary. Important is that the disk quarters are, topologically spoken, open. I.e., the movement of k in relation $(- - - 0)_C$ in Fig. 3 can be from k to every point on the curved part of the quarter part excluding the horizontal and the vertical line segment. On the other hand, the movement of l in this relation can only be from l straight to k, which is along the dashed line drawn from the open right dot.

3 Composition for QTC_C

A composition table is a central issue in qualitative reasoning. The idea behind a composition table is to compose a finite set of new facts and rules from existing ones, i.e., if two existing relations $R_1(k,l)$ and $R_2(l,m)$ share a common object (l), they can be composed into a new relation set $R_3(k,m)$, depicted by:

$$R_1(k,l) \otimes R_2(l,m) = R_3(k,m)$$

A composition table contains the set of compositions that are possible between all relations in a certain calculus; the left column containing R_1, the top row containing R_2, and the other entries containing $R_3 = R_1 \otimes R_2$.

3.1 Central Concepts

In this section, we focus on the two central concepts laying at the basis of the so-called composition-rule table for QTC_C. This table, which is generated by use of diagrammatic reasoning, is a compacted representation of a composition table. Let us consider both central concepts in detail:

a) Which rotation do we need, such that l of R_2 matches l of R_1?
Generating the composition relation $R_3(k,m)$, means finding out how k moves with respect to m and vice versa, based on $R_1(k,l)$ and $R_2(l,m)$; or in other words: based on the movement of k with respect to l (and vice versa), and the movement of l with respect to m (and vice versa). As said before, the movements are represented via relation icons. In order to be able to compose $R_1(k,l)$ and $R_2(l,m)$ by use of diagrammatic reasoning, we combine the relation icons representing R_1 and R_2. This will be done in two steps. In the first step, the relation icon of $R_2(l,m)$ is translated onto the relation icon of $R_1(k,l)$, in such a way that both origins of the vector

representing the movement of l in $R_1(k,l)$ and the vector representing the movement of l in $R_2(l,m)$ match. In the second step, the relation icon representing R_2 is rotated in such a way that the vector representing the movement of l in R_2 matches the vector representing the movement of l in R_1. Depending on the 6561 (81×81) composition combinations that are possible in QTC_C, one finds eight basic rotation[4] possibilities that can be classified in two groups:

- crisp rotations (0°, 90°, 180°, 270°): only one rotation angle is possible.
- range rotations (0°-90°, 90°-180°, 180°-270°, 270°-360°): there is a range of rotations, over which the second relation has to be rotated.

Notwithstanding that this is not the full spectrum of rotation possibilities, all others can be generated by combining multiples of these eight basic possibilities.

b) How is *k* moving with respect to *l* in R_1, and how is *m* moving with respect to *l* in R_2?

After the matching of the relation icons, the composition of $R_1(k,l)$ and $R_2(l,m)$ can be generated, by studying how k is moving with respect to l in R_1 and how m is moving with respect to l in R_2. This question can be answered by taking the appropriate characters from the QTC_C labels. Based on the definition of QTC_C, one can say:

- the movement of k with respect to l in R_1 can be found in the QTC_C label representing $R_1(k,l)$. The first character of this label represents the movement of k with respect to $RL\perp 1$. The third character of this label represents the movement of k with respect to RL_{kl}.
- the movement of m with respect to l in R_2 can be found in the QTC_C label representing $R_2(l,m)$. The second character of this label represents the movement of m with respect to $RL\perp 2$. The fourth character of this label represents the movement of m with respect to RL_{ml}.

3.2 Diagrammatic Reasoning

By combining the two central concepts with the diagrammatic reasoning process explained in this section, one can generate all 6561 compositions. Fig. 4-8 contain specific basic rotation possibilities. 'Specific', because only the rotation possibilities each time a new diagrammatic reasoning technique had to be used, are handled. The rotation possibility 'the rotation angle is 270°' is for example not handled since it is analogous with the rotation possibility 'the rotation angle is 90°'. Each possibility consists of a starting situation (a), representing $R_1(k,l)$ and $R_2(l,m)$ after the first part of the matching process, i.e., after translation. Each possibility also contains one (b) or multiple (b_1, b_2, etc.) composition results. Also the double-cross, needed to determine $R_3(k,m)$, is represented. The first and the second character in each zone of the first cross (centre at k) stand for respectively the first and the third character that the QTC_C-relation will get if the velocity vector of k is inside the specific zone. The first and the second character in each zone of the second cross (centre at l) stand for respectively the second and the fourth character that the QTC_C-relation will get if the velocity vector of l is

[4] Just as in trigonometry, we take anti-clockwise angles as being positive.

inside the specific zone. Because of the visualization aspect, only the labels not containing 0 and having no overlap with the velocity vectors are represented. Let us describe the diagrammatic reasoning process of the specific basic rotation possibilities:

Basic rotation possibility: 0° (Fig. 4). If one needs to combine two relations that can be matched without rotation, things are quite straightforward. Let us give an example. Fig. 4a shows R_1 and R_2 after translation: $R_1(k,l) = (- + 0\ 0)_C$ and $R_2(l,m) = (- + 0\ 0)_C$. Fig. 4b contains the composition result $R_3(k,m)$, which could be determined for this rotation possibility without rotation. The labels in the cross centered at k, are – (standing for the first character of $R_3(k,m)$) and 0 (standing for the third character of $R_3(k,m)$). The labels in the cross centered at m, are + (standing for the second character of $R_3(k,m)$) and 0 (standing for the fourth character of $R_3(k,m)$). Thus, $R_3(k,m) = (- + 0\ 0)_C$.

To let de reader become familiar with the technique of diagrammatic reasoning used here, let us give a second example of this rotation possibility:

- Fig. 4a': $R_1(k,l) = (- + - 0)_C$ and $R_2(l,m) = (- 0\ 0 -)_C$.
- Fig. 4b': $R_3(k,m) = (- 0 - -)_C$.

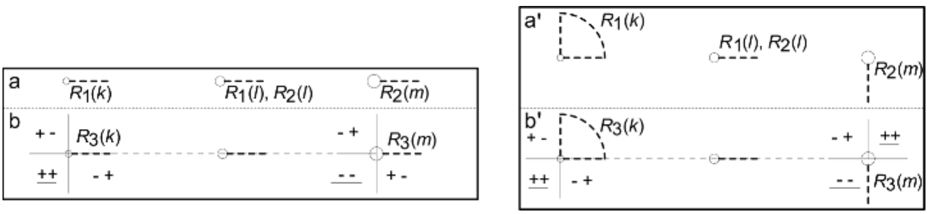


Fig. 4. Composition rules for basic rotation possibility: 0°

There are 9 options for the labels of k in R_1 and there are 9 options for the labels of m in R_2: (– –), (– 0), (– +), (0 –), (0 0), (0 +), (+ –), (+ 0), and (+ +). As a result, there are 9×9=81 compositions for this rotation possibility in QTC_C, as well as for the other rotation possibilities.

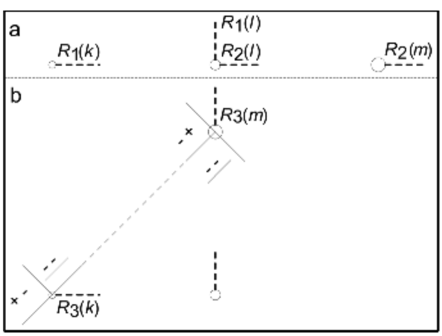


Fig. 5. Composition rules for basic rotation possibility: 90°

Basic rotation possibility: 90° (Fig. 5). In contrast to the previous rotation possibility, a rotation is needed here, which will of course have its consequences for the orientation of the double-cross as can be seen in the example in Fig. 5b. The figure shows R_1, R_2 and R_3, or in other words $R_1 \otimes R_2 = R_3$, or in QTC labels:

$$(-\ 0\ 0\ +)_C \otimes (-\ +\ 0\ 0)_C = (-\ +\ +\ +)_C$$

Basic rotation possibility: 180° (Fig. 6). There is no big difference between this rotation possibility and the former one. Worthwhile mentioning is that when the distance between *l* and *m* varies, different situations may occur. I.e., when the distance between *l* and *m* varies between:

- is smaller than the distance between *k* and *l* (Fig. $6b_1$);
- is equal to the distance between *k* and *l* (Fig. $6b_2$);
- is larger than the distance between *k* and *l* (Fig. $6b_3$).

In Fig. $6b_2$, the position of *l* after the matching process is equal to the position of *m* after the matching process. Since QTC_C only studies objects having the RCC-relation DC, the relation in Fig. $6b_2$ can be ignored. Note also the big difference in orientation of the double cross between Fig. $6b_1$ and Fig. $6b_3$.

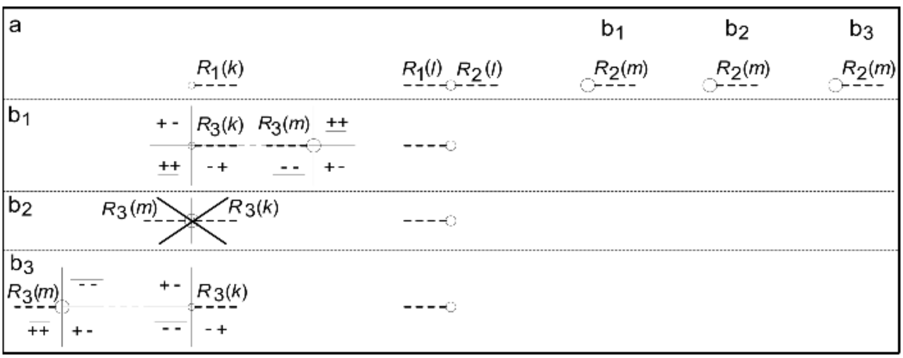


Fig. 6. Composition rules for basic rotation possibility: 180°

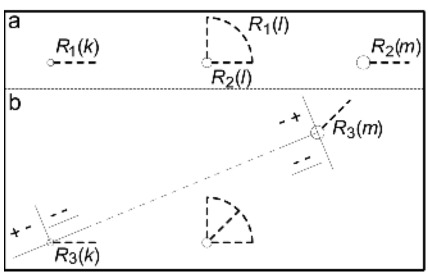


Fig. 7. Composition rules for basic rotation possibility: 0°-90°

Basic rotation possibility: 0°-90° (Fig. 7). Although all angles between 0° and 90° may be used as rotation angle, there is only one result.

Basic rotation possibility: 90°-180° (Fig. 8). There are five different options.[5]

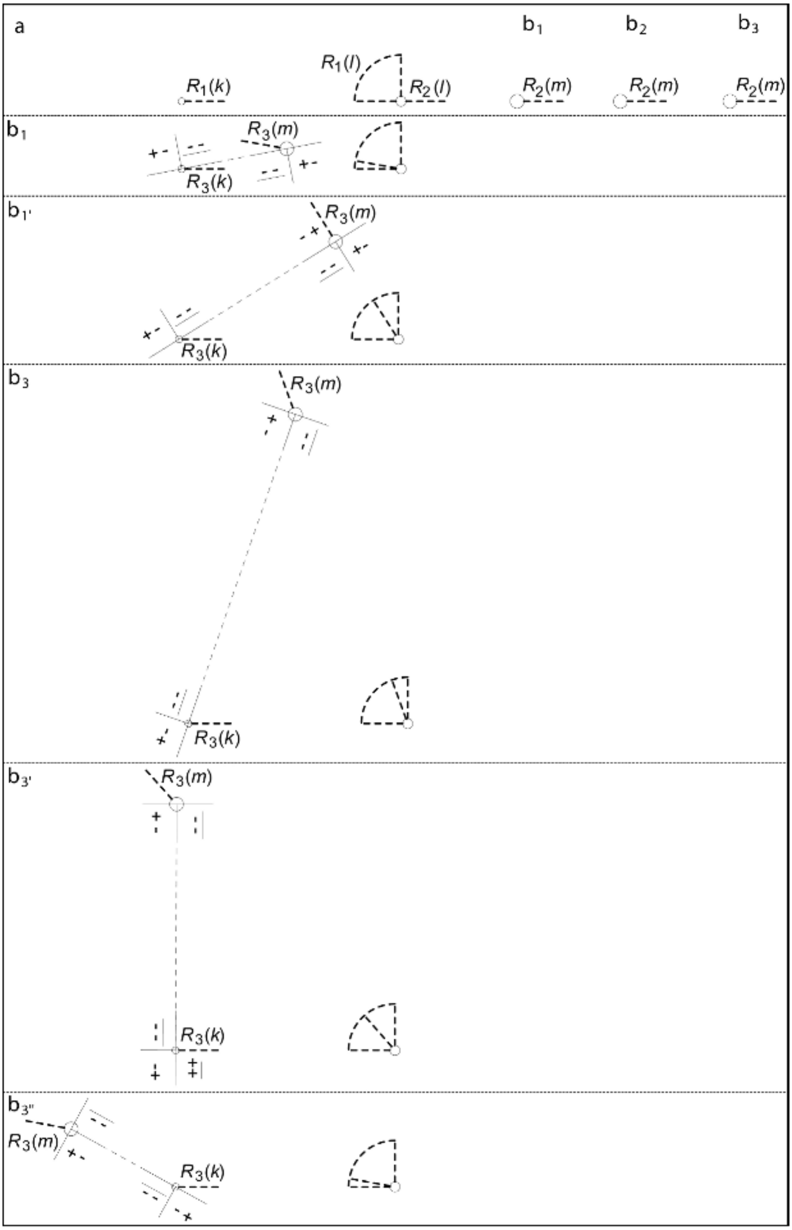


Fig. 8. Composition rules for basic rotation possibility: 90°-180°

[5] Due to space limitations, we do not go in detail on the diagrammatic reasoning process that determines the landmark values. A description of this process can be found in [16].

3.3 Composition-Rule Table

As described in the previous section, all 6561 compositions of QTC_C can be read from the diagrams, generated by diagrammatic reasoning. 1296 of the 6561 relations are invalid, due to the impossibility of inference between a moving point and a stationary point is impossible. 964 of the remaining 5265 relations are strong 964 and 4301 are weak. It is highly preferable to construct a compact table in which the compositions can be found by a simple table look-up. Such a compact table has been called the composition-rule table and will be worked out in this section. Although this so-called composition-rule table is not a traditional composition table containing all entries, this table does give all information contained in a composition table. In addition, it forms a basis of how to implement the composition rules in a practical information system.[6]

Table 1. Composition-rule table for QTC_C[7]

			X	0°		90°		180°				270°		0°-90°		90°-180°										180°-270°										270°-360°	
R_1C_1	R_1C_3			C_1	C_3	C_1	C_3	C_1	C_3	C_1	C_3	C_1	C_3	C_1	C_3	C_1	C_3	C_1	C_3	C_1	C_3	C_1	C_3	C_1	C_3	C_1	C_3	C_1	C_3	C_1	C_3	C_1	C_3	C_1	C_3	C_1	C_3
0	0	0	X	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
–	0	R	X	–	0	–	+	–	0	+	0	–	–	–	+	–	+	–	+	–	+	0	+	+	+	+	–	0	–	–	–	–	–	–	–	–	–
–	–	TR	X	–	–	–	U	–	–	+	+	U	–	–	U	–	U	–	U	–	U	–	+	U	+	+	U	+	–	U	–	U	–	U	–	U	–
0	–	T	X	0	–	–	–	0	–	0	+	+	–	–	–	–	–	–	–	–	–	–	0	–	+	+	+	+	0	+	–	+	–	+	–	+	–
+	–	TL	X	+	–	U	–	+	–	–	+	+	U	U	–	U	–	U	–	U	–	–	–	–	U	U	+	+	+	+	U	+	U	+	U	+	U
+	0	L	X	+	0	+	–	+	0	–	0	+	+	+	–	+	–	+	–	+	–	0	–	–	–	–	+	0	+	+	+	+	+	+	+	+	+
+	+	BL	X	+	+	+	U	+	+	–	–	U	+	+	U	+	U	+	U	+	U	+	–	U	–	–	U	–	+	U	+	U	+	U	+	U	+
0	+	B	X	0	+	+	+	0	+	0	–	–	+	+	+	+	+	+	+	+	+	+	0	+	–	–	–	–	0	–	+	–	+	–	+	–	+
–	+	BR	X	–	+	U	+	–	+	+	–	–	U	U	+	U	+	U	+	U	+	+	+	+	U	U	–	–	–	–	U	–	U	–	U	–	U
R_2C_2	R_2C_4			C_2	C_4	C_2	C_4	C_2	C_4	C_2	C_4	C_2	C_4	C_2	C_4	C_2	C_4	C_2	C_4	C_2	C_4	C_2	C_4	C_2	C_4	C_2	C_4	C_2	C_4	C_2	C_4	C_2	C_4	C_2	C_4	C_2	C_4
0	0	0	X	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
+	0	R	X	+	0	+	+	–	0	+	0	+	–	+	+	–	+	0	+	+	+	+	+	+	+	+	–	+	–	+	–	0	–	–	–	+	–
+	+	TR	X	+	+	U	+	–	–	+	+	+	U	U	+	–	U	–	+	U	+	U	+	U	+	+	U	+	U	+	U	+	–	U	–	+	U
0	+	T	X	0	+	–	+	0	–	0	+	+	+	–	+	–	–	–	0	–	+	–	+	–	+	+	+	+	+	+	+	+	0	+	–	+	+
–	+	TL	X	–	+	–	U	+	–	–	+	U	+	–	U	U	–	–	–	–	U	–	U	–	U	U	+	U	+	U	+	+	+	+	U	U	+
–	0	L	X	–	0	–	–	+	0	–	0	–	+	–	–	+	–	0	–	–	–	–	–	–	–	–	+	–	+	–	+	0	+	+	+	–	+
–	–	BL	X	–	–	U	–	+	+	–	–	–	U	U	–	+	U	+	–	U	–	U	–	U	–	–	U	–	U	–	U	–	+	U	+	–	U
0	–	B	X	0	–	+	–	0	+	0	–	–	–	+	–	+	+	+	0	+	–	+	–	+	–	–	–	–	–	–	–	–	0	–	+	–	–
+	–	BR	X	+	–	+	U	–	+	+	–	U	–	+	U	U	+	+	+	+	U	+	U	+	U	U	–	U	–	U	–	–	–	–	U	U	–

The composition-rule table (Table 1) contains a top heading ('X 0° 90°…270° 360°') and two sub-headings ('R_1C_1 R_1C_3 $R_3C_1R_3C_3$…$R_3C_1R_3C_3$' and 'R_2C_2 R_2C_4 $R_3C_2R_3C_4$…$R_3C_2R_3C_4$'). The body of the table consists of two parts, an upper part belonging to the first sub-heading and a lower part belonging to the second sub-heading. The top heading shows which column stands for which rotation that has to be made in order that l in R_2 matches l in R_1 (e.g. X: both relations cannot be matched via a rotation; 90°-180°: the range between 90° and 180° matches both relations). Apart from the X-column, every column contains at least one sub-heading

[6] In a way, the concept of the composition-rule table that is presented here for QTC_C, can be compared with the concept of the condensed composition table for the Interval Calculus (the Interval Calculus is presented in [1], its compaction is presented in [2]). For QTC_C, there is a compression from 6561 to 306 entries. For the Interval Calculus, there is a compression from 169 to 7 entries. Both compression results are below 5% (4.7% for QTC_C and 4.1% for the Interval Calculus).

[7] Due to space limitations, R_3 is removed for all sub-headings in the table.

('R_3C_1 R_3C_3' for the upper part and 'R_3C_2 R_3C_4' for the lower part). In the columns 'R_3C_1' and 'R_3C_3', respectively the first and the third character of the composition label can be found. If all qualitative variables are possible for one character, then the character is represented by 'U' standing for the universal set of qualitative values; if multiple combinations of characters are possible for a specific rotation possibility, then multiple columns need to be presented. In the columns 'R_3C_2' and 'R_3C_4', respectively the second and the fourth character of the composition label can be found. The rows of the upper part differentiate between the movements of k with respect to l (the first and third character in R_1). The rows of the lower part differentiate between the movements of m with respect to l (the second and fourth character in R_2). The abbreviations next to the characters stand for these movements: T (Top), TL (Top-Left), L (Left), BL (Bottom-Left), B (Bottom), BR (Bottom-Right), R (Right), and TR (Top-Right). Let us explain the composition-rule table by use of the examples for the first 2 rotation possibilities, given in section 3.2:

Basic rotation possibility: 0°. In the examples with 'rotation angle: 0', no rotation has to be made in order that l in R_2 matches l in R_1. Thus, we select the column labeled '0°'. In the first example, $R_1C_1 = –$, and R_1C_3 is 0. In other words, the movement of k with respect to l is R, which can be found in the upper part of the table. We thus need to select the row labeled '– 0 R'. The intersection between column '0°' and row '– 0 R' gives: $R_3C_1 = –$ and $R_3C_3 = 0$. Still in the first example, $R_2C_2 = +$ and $R_2C_4 = 0$. In other words, the movement of m with respect to l is R, which can be found in the lower part of the table. We thus need to select the row labeled with '+ 0 R'. The intersection between column '0°' and row '+ 0 R' gives: $R_3C_2 = +$ and $R_3C_4 = 0$. Combining the solutions from the upper part and the lower part of the table, results in $R_3(k,m) = (– + 0\ 0)_C$. Let us write down the second example in a shorter way:
Rotation constraint:

R_2 needs no rotation to match R_1. Thus, select column '0°'.

Character constraints:

Movement of k in R_1: TR. Thus, select row '– – TR' in upper part.
Movement of m in R_2: B. Thus, select row '0 – B' in lower part.
Thus, $R_3(k,m) = (– 0 – –)_C$.

Basic rotation possibility: 90°. In this example $(–\ 0\ 0\ +)_C \otimes (–\ +0\ 0)_C$, the composition-rule table shows:
Rotation constraint:

R_2 has to be rotated over 90° to match R_1. Thus, select '90°' column.

Character constraints:

Movement of k in R_1: R. Thus, select row '– 0 R' in upper part.
Movement of m in R_2: R. Thus, select row '+ 0 R' in lower part.
Thus, R_3 contains $(– + + +)_C$.

At first sight, this possibility is far more complex since there are five different options. Because every option stands for a specific rotation angle or rotation range, the first option of the upper part needs to be combined with the first option of the lower part, the second option of the upper part with the second option of the lower part, etc. Thus, we may not take the cross-product. Besides this, the same methodology as before can be used.

4 The Example of 'Puzzling the Past'

Scientists have researched many evolutions of phenomena, often dealing with moving objects. Frequently, a jigsaw puzzle has to be constructed without knowing on beforehand which part will enter the picture and at what time. Therefore, such reconstructions are very complicated to represent, implement, analyze, visualize, etc. This section shows that QTC_C is well suited for this kind of interesting research, since this qualitative calculus can handle incomplete knowledge and composition rules have been constructed for QTC_C[8]. We consider the example of geomorphologic research performed by several teams.

4.1 Initial Research

Suppose two scientific teams are doing research on a site, independent of each other. Both have been asked to describe the movement of object k with respect to object n (R_3). The problem for both teams is that they cannot find data, in order to determine directly the movement of k with respect to n. Therefore, both teams need other data to infer an (incomplete) answer. The following data has been found:

Team 1: $R_1(k,l) = (- + + -)_C$, $R_2\ (l,n) = (- - - +)_C$.
Team 2: $R_1(k,m) = (- + - +)_C$, $R_2(m,n) = (\,- + - +)_C$.

In order to determine $R_3(k,n)$, the composition-rule table (Table 1) can be used.

Team 1. In order to compose $R_1(k,l) = (- + + -)_C$ and $R_2(l,n) = (- - - +)_C$, the second relation needs a rotation between 180° and 360°. Thus, there are three basic rotation possibilities: 180°-270°, 270°, and 270°-360°.

180°-270°. Select column '180°-270°'. $R_1C_1 = -$, and $R_1C_3 = +$. In other words, the movement of k with respect to l is BR, which can be found in the upper part of the table. Therefore, select the row labeled '– + BR'. The intersection between column '180°-270°' and row '– + BR' gives the qualitative values for the first and the third character of the composition label, having five options:

option 1: $R_3C_1 = \text{U}$ and $R_3C_3 = -$,
option 2: $R_3C_1 = -$ and $R_3C_3 = -$,
option 3: $R_3C_1 = -$ and $R_3C_3 = \text{U}$,
option 4: $R_3C_1 = -$ and $R_3C_3 = \text{U}$,
option 5: $R_3C_1 = -$ and $R_3C_3 = \text{U}$.

$R_2C_2 = -$ and R_2C_4 =s +. In other words, the movement of n with respect to l is TL, which can be found in the lower part of the table. Therefore, select the row labeled '– + TL'. The intersection between column '180°-270°' and row '– + TL' gives the qualitative values for the second and the fourth character of the composition label:

option 1: $R_3C_2 = \text{U}$ and $R_3C_4 = +$,
option 2: $R_3C_2 = \text{U}$ and $R_3C_4 = +$,
option 3: $R_3C_2 = \text{U}$ and $R_3C_4 = +$,

[8] Note that analogous reasoning processes have been worked out for the temporal domain [2, 46].

option 4: R_3C_2 = + and R_3C_4 = +,
option 5: R_3C_2 = + and R_3C_4 = U.

Combining the solutions from the upper part and the lower part of the table, results in:

option 1: $(\text{U U} - +)_\text{C}$,
option 2: $(- \text{U} - +)_\text{C}$,
option 3: $(- \text{U U} +)_\text{C}$,
option 4: $(- + \text{U} +)_\text{C}$,
option 5: $(- + \text{U U})_\text{C}$.

270°. Dual reasoning as for 'rotation possibility: 180°-270°', results in option 6: $(- \text{U U} +)_\text{C}$.
270°-360°. Dual reasoning as for 'rotation possibility 180°-270°', results in option 7: $(- \text{U U} +)_\text{C}$.

The disjunction of all results gives:

$$(\text{U U} - +)_\text{C} \cup (- \text{U} - +)_\text{C} \cup (- \text{U U} +)_\text{C} \cup (- + \text{U} +)_\text{C} \cup (- + \text{U U})_\text{C} \cup$$
$$(- \text{U U} +)_\text{C} \cup (- \text{U U} +)_\text{C} =$$
$$(\text{U U} - +)_\text{C} \cup (- \text{U U} +)_\text{C} \cup (- + \text{U U})_\text{C}.$$

Team 2. In order to compose $R_1(k,m) = (- + - +)_\text{C}$ and $R_2(m,n) = (- + - +)_\text{C}$, the second relation needs a rotation between -90° and 90°. Thus, there are three basic rotation possibilities: 270°-360°, 0°, and 0°-90°.

270°-360°. Option 1: $(\text{U} + - \text{U})_\text{C}$
0°. Option 2: $(- + - +)_\text{C}$
0°-90°. Option 3: $(- \text{U U} +)_\text{C}$

The disjunction of all results gives:

$$(\text{U} + - \text{U})_\text{C} \cup (- + - +)_\text{C} \cup (- \text{U U} +)_\text{C} =$$
$$(\text{U} + - \text{U})_\text{C} \cup (- \text{U U} +)_\text{C}.$$

4.2 Follow-Up Research

Both research teams get a different incomplete result. However, the real answer must be a subset of the incomplete answer. Suppose a third and a fourth team want to do further research on this site and can use the results of Team 1 and Team 2. Suppose Team 3 is convinced that Team 1 and Team 2 were correct. This would mean that Team 3 takes the conjunction of both former results as being the new incomplete relation to which the correct answer will certainly belong. Suppose Team 4 doubts the correctness of the data gathered by Team 1 and Team 2, but does not know which team would have had the best results. Team 4 could take the disjunction of both former results as being the set to which the correct result has to belong.

Team 3. $((\text{U U} - +)_\text{C} \cup (- \text{U U} +)_\text{C} \cup (- + \text{U U})_\text{C}) \cap ((\text{U} + - \text{U})_\text{C} \cup (- \text{U U} +)_\text{C})$

Team 4. $((\text{U U} - +)_\text{C} \cup (- \text{U U} +)_\text{C} \cup (- + \text{U U})_\text{C}) \cup ((\text{U} + - \text{U})_\text{C} \cup (- \text{U U} +)_\text{C})$

Team 5. Finally, a new team (Team 5) finds a new methodology and can detect directly: $R_3(k,n) = (- + - +)_C$.

Thus, one can see that the incomplete answers of both Team 1 and Team 2, and thus also the answers of Team 3 and Team 4, contained the correct answer.

5 Concluding Remarks

The example of 'puzzling the past', in which a jigsaw puzzle of a configuration of moving points is represented, needs further investigation, since we strongly believe that this example forms a basis for implementation of incomplete spatio-temporal knowledge in information systems. QTC_C and the concept of the composition-rule table can be used in a variety of research domains, such as geomorphology, geology, archaeology, and biology. In complex researches, there is a huge number of anchor points, teams, measurements per team, updates, etc. Such assessments will become complex, but we are convinced that implementation of QTC_C in an information system can lead to interesting results for this widespread but difficult kind of reconstruction processes.

Acknowledgments

This research is funded by the Research Foundation - Flanders, Research Project G.0344.05.

References

1. Allen, J.F. Maintaining Knowledge about Temporal Intervals, *Communications of the ACM*, 26 (11): 832-843, 1983.
2. Freksa, C. Temporal Reasoning Based on Semi-Intervals, *Artificial Intelligence*, 54: 199-227, 1992.
3. Cohn, A.G. and Hazarika, S.M. Qualitative Spatial Representation and Reasoning: An Overview, *Fundamenta Informaticae*, 46 (1-2): 1-29, 2001.
4. Wolfson, O., Xu, B., Chamberlain, S., and Jiang, L. Moving Object Databases: Issues and Solutions, *Proc. of SSDBM*, Capri, Italy, pp. 111-122, 1998.
5. Erwig, M., Güting, R.H., Schneider, M., and Vazirgiannis, M. Spatio-Temporal Data Types: An Approach to Modelling Objects in Databases, *Geoinformatica*, 3 (3): 269-296, 1999.
6. Moreira, J., Ribeiro, C., and Saglio, J.-M. Representation and Manipulation of Moving Points: An Extended Data Model for Location Estimation, *Cartography and Geographic Information Systems*, 26 (2): 109-123, 1999.
7. Nabil, M., Ngu, A., and Shepherd, A.J. Modelling and Retrieval of Moving Objects, *Multimedia Tools and Applications*, 13 (1): 35-71, 2001.
8. Pfoser, D. Indexing the Trajectories of Moving Objects, *IEEE Data Engineering Bulletin*, 25 (2): 3-9, 2002.
9. Muller, P. A Qualitative Theory of Motion Based on Spatiotemporal Primitives, In: Cohn, A.G., Schubert, L., and Shapiro, S. (Eds.), *Proc. of KR*, Trento, Italy, pp. 131-142, 1998.
10. Galton, A. *Qualitative Spatial Change*, University Press, Oxford, UK, 409 pp, 2000.

11. Hornsby, K. and Egenhofer, M. Identity-Based Change: A Foundation for Spatio-Temporal Knowledge Representation, *International Journal of Geographical Information Science*, 14 (3): 207-224, 2000.
12. Claramunt, C. and Jiang, B. An Integrated Representation of Spatial and Temporal Relationships between Evolving Regions, *Geographical Systems*, 3 (4): 411-428, 2001.
13. Hazarika, S.M. and Cohn A.G. Qualitative Spatio-Temporal Continuity, In: Montello, D.R. (Ed.), *Proc. of COSIT*, Morro Bay, USA, *LNCS*, (2205), Springer-Verlag, pp. 92-107, 2001.
14. Randell, D., Cui, Z, and Cohn, A.G. A Spatial Logic Based on Regions and Connection, In: Nebel, B., Swartout, W., and Rich, C. (Eds.), *Proc. of KR*, San Mateo, USA, pp. 165-176, 1992.
15. Egenhofer, M. and Franzosa, R. Point-Set Topological Spatial Relations, *International Journal of Geographical Information Systems*, 5 (2): 161-174, 1991.
16. Van de Weghe, N. *Representing and Reasoning about Moving Objects: A Qualitative Approach*, PhD thesis, Belgium, Ghent University, Faculty of Sciences, Department of Geography, 268 pp., 2004.
17. Van de Weghe, N., Cohn, A.G., Bogaert, P., and De Maeyer, Ph. Representation of Moving Objects along a Road Network, *Proc. of Geoinformatics*, Gävle, Sweden, pp. 187-197, 2004.
18. Bogaert, P., Van de Weghe, N. and De Maeyer, Ph. Description, Definition and Proof of a Qualitative State Change of Moving Objects along a Road Network, In: Raubal, M., Sliwinski, A., and Kuhn, W. (Eds.), *Proc. of the Münster GI-Days. Geoinformation and Mobility, from Research to Applications*, Münster, Germany, pp. 239-248, 2004.
19. Van de Weghe, N., Cohn, A.G., De Maeyer, Ph., and Witlox, F., 2005. Representing Moving Objects in Computer-Based Expert Systems: The Overtake Event Example, *Expert Systems with Applications*, 29 (4). (*Accepted for publication*)
20. Freksa, C. and Zimmermann, K. On the Utilization of Spatial Structures for Cognitively Plausible and Efficient Reasoning, *Proc. of the Conf. on Systems, Man, and Cybernetics*, Chicago, USA, pp. 261-266, 1992.
21. Vieu, L. Spatial Representation and Reasoning in Artificial Intelligence, In: Stock, O. (Ed.), *Spatial and Temporal Reasoning*, Kluwer, Dordrecht, Netherlands, pp. 5-41, 1997.
22. Bennett, B. *Logical Representations for Automated Reasoning about Spatial Relationships*, PhD thesis, UK, University of Leeds, School of Computer Studies, 211 pp., 1997.
23. Goyal, R.K. *Similarity Assessment for Cardinal Directions Between Extended Spatial Objects*, PhD thesis, USA, University of Maine, Graduate School, Spatial Information Science and Engineering, 167 pp., 2000.
24. Randell, D.A. and Cohn, A.G. Modelling Topological and Metrical Properties of Physical Processes, In: Brachman, R., Levesque, H., and Reiter, R. (Eds.), *Proc. of KR*, Toronto, Canada, pp. 55-66, 1989.
25. Egenhofer, M. Reasoning about Binary Topological Relations, In: Gunther, O. and Schek, H.-J., *Proc. of the Symposium on Large Spatial Databases*, Zurich, Switzerland, *LNCS*, (525), Springer-Verlag, pp. 143-160, 1991.
26. Zimmermann, K. and Freksa, C. Enhancing Spatial reasoning by the Concept of Motion, In: Sloman, A., Hogg, D., Humphreys, A., Ramsay, A., and Partridge, D. (Eds.), *Proc. of AISB*, Birmingham, UK, pp. 140-147., 1993.
27. Frank, A.U. Qualitative Spatial reasoning: Cardinal Directions as an Example, *International Journal of Geographical Information Science*, 10 (3): 269-290, 1996.
28. Gooday, J.M. and Cohn, A.G. Conceptual Neighbourhoods in Temporal and Spatial Reasoning, In: Rodriguez, R. (Ed.), *Proc. of the ECAI Spatial and Temporal Reasoning Workshop*, Amsterdam, Netherlands, 1994.

29. Mukerjee, A. and Joe, G. A qualitative Model for Space, *Proc. of AAAI*, Los Altos, USA, pp. 721-727, 1990.
30. Jungert, E.The Observer's Point of View: an Extension of Symbolic Projections, In: Frank, A.U., Campari, I., and Formentini, U. (Eds.), *Proc of the Conf. on Theories and Methods of Spatio-Temporal Reasoning in Geographic Space*, Pisa, Italy, *LNCS*, (639), Springer-Verlag, 179-195, 1992.
31. Hernández, D. *Qualitative Representation of Spatial Knowledge*. *LNAI*, (804), Springer-Verlag, 202 pp., 1994.
32. Schlieder, C. Reasoning about Ordering, In: Frank, A.U. and Kuhn, W. (Eds.), *Proc. of COSIT*, Semmering, Austria, *LNCS*, (988), Springer-Verlag, 341-349, 1995.
33. Hernández, D. and Jungert, E. Qualitative Motion of Point-Like Objects, *Journal of Visual Languages and Computing*, 10: 269-289, 1999.
34. Musto, A., Eisenkolb, A., Röfer, T., and Stein, K. Qualitative and Quantitative Representations of Locomotion and their Application in Robot Navigation, *Proc. of IJCAI*, San Francisco, USA, 1067-1073, 1999.
35. Sogo, T., Ishiguro, H. And Ishida, T. Acquisition of Qualitative Spatial Representation by Visual Observation, *Proc. of IJCAI*, San Francisco, USA, 1054-1060, 1999.
36. Fernyhough, J.H., Cohn, A.G., and Hogg, D.C. Constructing Qualitative Event Models Automatically from Video Input, *Image and Vision Computing*, 18 (2), 81-103, 2000.
37. Nabil, M., Ngu A., and Shepherd A.J. Modelling and Retrieval of Moving Objects, *Multimedia Tools and Applications*, 13 (1), 35-71, 2001.
38. Hornsby, K. and Egenhofer, M. Modelling Moving Objects over Multiple Granularities, *Annals of Mathematics and Artificial Intelligence*, 36 (1-2), 177-194, 2002.
39. Stolzenburg, F., Obst, O., and Murray, J. Qualitative Velocity and Ball Interception, In: Jarke, M., Köhler, J., and Lakemeyer, G. (Eds.), *Proc. of KI*, Aachen, Germany, *LNCS*, (2479), Springer-Verlag, 283-298, 2002.
40. Du Mouza, C. and Rigaux, P. Multi-Scale Classification of Moving Object Trajectories, *Proc. of SSDBM*, Santorini Island, Greece, 307-316, 2004.
41. Dylla, F. and Moratz, R. Exploiting Qualitative Spatial Neighborhoods in the Situation Calculus, In: Freksa, C., Knauff, M., Krieg-Brückner, B., Nebel, B., and Barkowsky, T., *Proc. of Spatial Cognition IV*, Frauenchiemsee, Germany, *LNCS*, (3343), Springer-Verlag, 304-322, 2004.
42. Parent, C., Spaccapietra, S., and Zimanyi, E. Spatio-Temporal Conceptual Models: Data Structures + Space + Time, *Proc. of ACM GIS*, Kansas City, USA, 26-33, 1999.
43. Worboys, M. Event-Oriented Approaches to Geographic Phenomena, *International Journal of Geographical Information Science*, 19 (1), 1-28, 2005.
44. Freksa, C. Using Orientation Information for Qualitative Spatial Reasoning, In: Frank, A.U., Campari, I., and Formentini, U. (Eds.), *Proc. of the Conf. on Theories and Methods of Spatio-Temporal Reasoning in Geographic Space*, Pisa, Italy, *LNCS*, (639), Springer-Verlag, 162-178, 1992.
45. Zimmermann, K. and Freksa, C. Qualitative Spatial Reasoning Using Orientation, Distance, and Path Knowledge, *Applied Intelligence*, 6 (1), 49-58, 1996.
46. Kulpa, Z. Diagrammatic Representation for a Space of Intervals, *Machine Graphics and Vision*, 6 (1), 5-24, 1997.

Modeling Noteworthy Events in a Geospatial Domain

Stephen Cole and Kathleen Hornsby

Department of Spatial Information Science and Engineering,
National Center for Geographic Information and Analysis,
University of Maine, Orono ME 04469-5711, USA
{scole, khornsby}@spatial.maine.edu

Abstract. This paper presents an approach for modeling semantics associated with *occurrents* in geospatial domains. Occurrents correspond to what is commonly thought of as a happening or activity in the real world. We describe a modeling approach where representations of occurrents are modeled as classes of *events.* Additional semantics are gained by modeling specialized subclasses of event classes as *derived events*. Significant occurrents are modeled as *noteworthy events*, i.e., happenings or activities in a domain that require intervention, for example, an automated notification that a noteworthy event has been detected. The representation is extended to treat *event sequences* that capture a variety of occurrent-based semantics, modeling both routine and unexpected occurrents as experienced, for example, by moving entities, such as vessels in a harbor.

1 Introduction

Geospatial entities are constantly experiencing change. For example, vehicles and pedestrians in a city interact and adjust to each other, the movement of air traffic is affected by weather patterns, and the movement of animals in an ecosystem changes with respect to the presence of predators. To facilitate modeling these interactions, static geospatial data models must evolve to spatio-temporal data models that support these kinds of dynamic experiences. There is now a well-developed literature on this evolution, (see for example, [1. 2, 3, 4]). A recent focus of spatio-temporal modeling is the incorporation of dynamic happenings or *occurrents* that take place in a domain into the geospatial data model. An occurrent refers to what is commonly thought of as a process, activity, or happening in the real world, such as a traffic accident or running a race [5]. These entities unfold themselves over time in successive phases or temporal parts [6]. Occurrents can be contrasted with *continuants*, referring to entities that endure over time, for example, a building, land parcel, or rock. In this paper, an approach for modeling occurrents and sequences of occurrents in dynamic geospatial domains is presented. We describe a modeling approach for representing the semantics associated with significant occurrents as *noteworthy events*. Noteworthy events capture happenings or activities in a domain that require intervention, for example, a system-generated notification that evidence of a significant occurrent has been detected in the database.

A harbor, including all the waters, waterways, vessels and port facilities located within it, serves as a prototypical geospatial domain for presenting the modeling constructs described in this paper. A harbor is highly dynamic with vessels moving through the harbor and interacting with each other. These interactions can be between

M.A. Rodríguez et al. (Eds.): GeoS 2005, LNCS 3799, pp. 77 – 89, 2005.

vessels, as they meet in the harbor, or between a vessel and the harbor's environment and port facilities, for example tying up to a pier or sailing through a rainstorm. As a result of these interactions, a vessel visiting the harbor will experience many occurrents. Thus a realistic representation of a harbor must model the occurrents that are known to take place, along with continuant entities that exist there, for example, the various zones of the harbor.

The remainder of this paper is structured as follows: Section 2 describes related research on spatio-temporal and event-based geospatial modeling, as well as event notifications. Section 3 describes a geospatial model of a harbor. The geospatial model is extended through an event-based approach presented in Section 4, where categories of events including derived events and noteworthy events are discussed. Methods for modeling sequences of events as well as the higher order semantics associated with sequences are discussed in Section 5. Section 6 presents conclusions and future work.

2 Related Work

There has been much interest in both the computer science and geographic information science research community in being able to model dynamic or time-varying phenomena. The literature on temporal and spatio-temporal data models and query languages is extensive (see, for example, [7, 8, 9, 3]). This work has been extended to modeling scenarios of change; particularly change with respect to the existence of objects over time [2, 10]. There have been similar advancements in spatio-temporal indexing and analysis [11, 12, 13].

Most of this research has focused on modeling continuants. Recently, however, there has been an interest in developing *occurrent-based approaches* (or the term 'event-based approaches' is sometimes used) and research on modeling the dynamic aspects of geospatial domains has highlighted how a solely continuant-centered perspective misses the contribution of dynamic aspects by modeling happenings or occurrents implicitly [14]. Grenon and Smith [5] have presented a formal ontology for describing geospatial objects and geospatial events and processes. Their work on SNAP and SPAN ontologies treat continuants and occurrents respectively, presenting a framework for handling both kinds of entities and allowing for a representation of the relations that can exist between these ontologies. The Geospatial Event Model (GEM) [15] presents a formal model for reasoning about continuants, occurrents, and settings (extents in space and time) in geospatial domains, emphasizing the central role of occurrents. In addition, continuants and occurrents are shown to be important for reasoning about and understanding geospatial object-based and field-based representations [16].

This paper focuses on developing a spatio-temporal model for dynamic geospatial domains that include moving entities, such as automobiles, planes and ships. Some earlier work considers movement over spatio-temporal paths, or geospatial lifelines [17] or trajectories [11], typically from a location-based representation of moving objects (see for example, [18]) or a combination of location-based and geometry-based methods [19]. We take a different approach; one that is based on the semantics of occurrents and sequences of occurrents that take place in a domain. In this paper, movement is modeled through sequences of events that capture the dynamic experiences of an entity as it travels through space. Using events to model movement enables the inclusion of other event-based application functions such as *event-condition-action* rules [20] and gives a foundation for developing event notification services [21, 22].

3 A Geospatial Model of a Harbor

A geospatial model of a harbor comprises classes of continuant entities modeled as objects. These objects represent naturally occurring entities, such as islands and rivers, and man-made entities, like vessels, piers and aids to navigation. *Fiat objects* represent entities such as traffic separation schemes and precautionary areas. The fiat object class *zone* is used as a basis for the geospatial representation of the harbor by partitioning the harbor along fiat boundaries into a finite set H of ρ zones. Examples of the zones in a harbor include, fishing and recreation zones, designated anchorages, and inbound and outbound lanes of Traffic Separation Schemes (Figure 1). An offshore zone is included to represent the waters outside the harbor limits. Some zones within a harbor are further distinguished as *destination zones D* (i.e., zones that may be declared as the intended destination of a vessel entering or leaving the harbor). Thus, the set of destination zones is a subset of harbor zones. Examples of *destination zones* include berths within the harbor where vessels can be moored, such as a shipyard, petroleum terminal, ferry landing, and designated anchorage, as well as the offshore zone. Since vessels typically arrive and depart from a harbor, every *destination zone* is also a *departure zone*. For example, if a vessel arriving at the harbor from sea intends to moor alongside a dock in the shipyard, its *departure zone* is the offshore zone and its *destination zone* is the shipyard.

Vessels in the harbor are continuant entities modeled as objects. Attributes of vessel objects include *identity*, *name*, *location* (the vessel's point coordinates), *currZone* (the harbor zone currently containing the vessel's point coordinates), *oldZone* (the harbor zone the vessel exited when entering the *currZone*), and *destination* (the vessel's intended destination zone). The attributes *identity*, *name*, and *location* are vessel dependant, this information is assumed to be available automatically from the vessel via Automated Vessel Identification Systems [23].

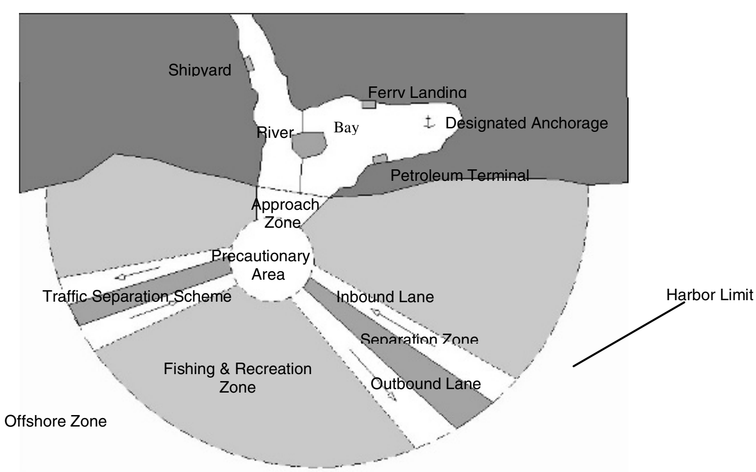


Fig. 1. Partitioning of a harbor into *zones*

4 Extending the Geospatial Model with Occurrents

Along with continuant entities modeled as objects, a harbor is associated with *occurrents*, where a *harbor occurrent* refers to a happening or activity that takes place in one of the harbor's zones. In the *event-based* modeling approach presented in this paper a dynamic domain's continuants and occurrents are modeled as classes of objects and events respectively. A UML class diagram depicts event classes that model some prototypical harbor occurrents (Figure 2). All harbor event classes have the attributes *zone*, specifying the harbor zone where the event took place, and *occurTime*, a time stamp

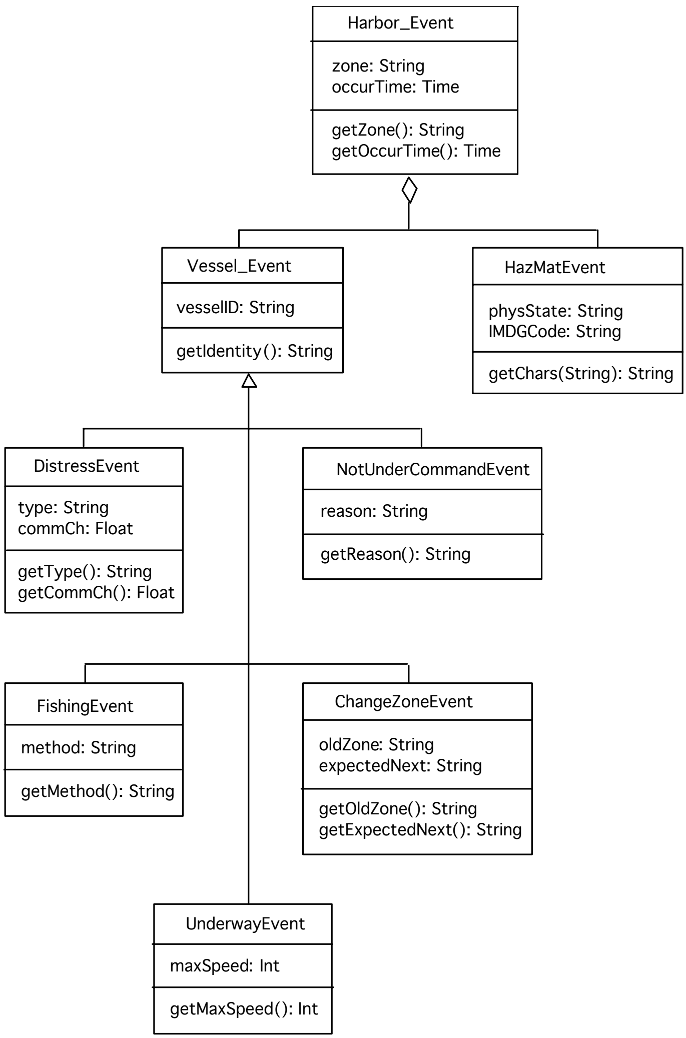


Fig. 2. UML representation of harbor event classes

indicating when the event took place. In this paper, we adopt the perspective that is common in many database approaches, where events are assumed to be instantaneous [22], i.e., events do not have any duration and signal a change in state of an object. This subset of event classes could be further extended with other classes of events, including those that model natural occurrents, such as wind and tide events.

An instance of an event class is simply referred to as an *event*. For example, vessel WK1 exits the eastern fishing and recreation zone and enters the inbound lane of the southeast traffic separation scheme at time t_1. This movement across a zone boundary is modeled as event ***cze$_1$***, an instance of the *ChangeZoneEvent* class (Figure 3). Events are distinguishable from each other by their respective attribute values (i.e., the set of attribute name:value pairs defining an instance of an event class is unique to any particular event). Due to the ephemeral nature of events, the attribute values of an event are immutable.

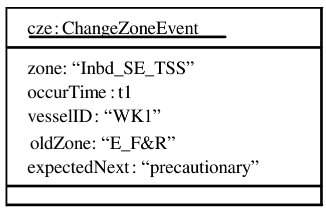


Fig. 3. UML representation of event *cze*

4.1 Defining Derived Events

Including more events that correspond to different perspectives of a domain in a spatio-temporal model enables a richer and more realistic data model. *Derived events* provide a means for increasing the range of events modeled. The notion of a derived event is based on an event meeting a set of conditions, similar to the event-condition-action rules of active DBMS [20]. For example, when a vessel triggers a *ChangeZoneEvent* by entering any harbor zone from a departure zone (i.e., the *oldZone* attribute value of a *ChangeZoneEvent* corresponds to the name of a departure zone) the derived ***DepartEvent*** *dep* is identified and an event representing the *DepartEvent* is created. Departure zones are defined as elements of set *D*, the subset of zones in a harbor that may be declared as a destination. (As defined in section 3, the elements of set *D* may be either destination zones or departure zones). Therefore, an instance of the *ChangeZoneEvent* class *cze* is a *DepartEvent dep* if $cze.oldZone \in D$.

Conversely, the derived ***ArrivalEvent*** *arr* marks the end of a vessel's traversal through the harbor. The *ArrivalEvent arr* is based on an instance of the *ChangeZoneEvent cze*, where $cze.zone \in D$.

Derived events, therefore, are specialized subclasses of the event class on which they are based. In this paper, all derived events are subclasses of the *ChangeZoneEvent* class. In subsequent sections of the paper, we will show the relevance of derived events for modeling dynamic geospatial domains.

4.2 Defining Noteworthy Events

The occurrents associated with a domain provide the building blocks necessary to model the domain's spatio-temporal aspects, for example, the movements of vessels traversing a harbor. Many of these occurrents, while necessary to construct a realistic model, happen routinely. Other occurrents are *significant* and it is desirable to distinguish those events that model such *significant occurrents* in order to develop information systems that are capable of managing them. Significant occurrents, modeled as *noteworthy events*, are dynamic happenings that require intervention in the domain, such as the response that is necessary in the case of a serious traffic accident, or further monitoring of a situation, such as a flood watch during periods of high water levels. Identifying an event as being noteworthy triggers some outside action, for example, a notification prompting a user to devise and initiate an appropriate response.

A Priori* Noteworthy Event Classes.** In an event-based model of a domain, certain event classes are, by definition, noteworthy. These event classes are referred to as *a priori* noteworthy event classes, i.e. every instance of this class is noteworthy wherever and whenever it may take place. In a harbor domain, for example, three event classes are defined as *a priori* noteworthy event classes, the ***HazMatEvent, ***DistressEvent*** and ***NotUnderCommandEvent***. *HazMatEvents* model the case when any substance that poses a threat to the ecosystem is released into the environment, such as would happen if a rail car of chlorine derailed, or a storage tank at the petroleum terminal overflowed, or a vessel ran aground and punctured a fuel tank. *DistressEvents* represent cases where vessels declare that outside assistance is required to prevent the loss of life or property as the result of, for example, a collision or fire on board. *NotUnderCommandEvents* model cases where vessels are unable to avoid collision by maneuvering as required by law, for example, a mechanical failure resulting in the loss of propulsion or steering. These occurrents are always noteworthy in the harbor management domain because user intervention is required (e.g. to dispatch the necessary assistance to contain and clean up a spill, or provide life-saving assistance) or, in the case of the *NotUnderCommandEvent*, notification is necessary to maintain the user's situational awareness. One possible method for modeling *a priori* noteworthy events includes adding a *noteworthy* attribute to the event's class definition.

Noteworthy Derived Events. Noteworthy events may also be distinguished by comparing the events with which an object is associated with the events with which an object is *expected* to be associated. For example, in the harbor domain, modeled expected vessel movements are used as the basis for identifying an ***UnexpectedDestinationEvent***, a subclass of the ChangeZoneEvent class. Vessels are assumed to declare their intended destination to a harbor's vessel traffic manager when approaching the harbor limits or getting underway from a berth in the harbor. Vessels are expected to complete their traversals by entering their intended destination zone. If any other destination zone is entered, however, an instance of the *UnexpectedDestinationEvent* class is identified. We refer to this type of event as a *noteworthy derived event*. In general, noteworthy events of this type are derived events where the associated base event's attribute value(s) violates a constraint(s) and, due to that violation, some external action (e.g. a notification) is triggered.

Another example of a noteworthy derived event class is the ***UnexpectedZoneEvent***, also a subclass of the ChangeZoneEvent class. An *UnexpectedZoneEvent* describes cases where vessels deviate from a customary route and enter a zone other than their *expectedNext* zone (i.e., $cze.zone \neq cze'.expectedNext$ where cze' is the *ChangeZoneEvent* preceding *cze* in a vessel's traversal of the harbor). For example, vessel WK2 with a current zone, approach zone, and an intended destination, ferry landing, is assigned an expected next zone, the bay. If, however, the next *ChangeZoneEvent* associated with WK2 happens in the river rather than the bay as expected, an *UnexpectedZoneEvent* is identified.

In dynamic domains, *unexpected occurrents* will take place along with routine, planned, or expected events. These types of occurrents may happen commonly or rarely. In either case, unexpected occurrents must be accounted for if they are to be managed. Incorporating methods for identifying unexpected events in information systems that model such occurrents provides the foundation for the development of computational tools for information systems, for example, event notification services.

5 Event Sequences

The notation ${}^{v}_{ID}e^{zone}_{time}$ is used to represent events, where *e* is an abbreviation of the event's class name, *ID* is the identifier of the vessel involved in the event, *zone* is the zone where the event took place, and *time* is a time stamp specifying when the event happened. For example, a *ChangeZoneEvent cze* associated with vessel QRS upon entering the approach zone at time *t* is represented as ${}_{QRS}cze^{approach}_{t}$. The derived *ArrivalEvent*, introduced in section 4.1 as a specialized subclass of the *ChangeZoneEvent* class, models vessel RST arriving at its destination, the shipyard, at time *t* and is represented as ${}_{RST}arr^{shipyard}_{t}$.

The *v* component of this notation is reserved for the attribute name:value pair used to identify derived events in those cases where the defining condition is based on an attribute other than *ID*, *zone*, or *time*. The noteworthy derived *UnexpectedDestinationEvent* (i.e., a *ChangeZoneEvent* that happens in a destination zone other than the vessel's declared destination) is associated with vessel RST. This vessel declared the shipyard as its destination, but instead, entered the ferry landing at time *t*. This event is represented as, ${}^{dest:shipyard}_{RST}ude^{ferry\_landing}_{t}$.

A domain's occurrents take place in an ordered progression. To model such a progression, an ordering > on a set of events *O* is defined as $\forall e, e' \in O$ where $e > e'$ if the *occurTime* (i.e., the time stamp indicating when an event happened) of *e* precedes the *occurTime* of e'. A sequence of events, or *event sequence E*, is defined as a set of events $\left\{ {}^{v}_{ID}e^{zone}_{t1}, {}^{v}_{ID}e^{zone}_{t2}, \ldots, {}^{v}_{ID}e^{zone}_{tn} \right\}$ where, ${}^{v}_{ID}e^{zone}_{t1} > {}^{v}_{ID}e^{zone}_{t2} > \ldots > {}^{v}_{ID}e^{zone}_{tn}$. For any two timestamps, t_j and t_k, one and only one of the following relations must hold: either t_j precedes t_k, or t_k precedes t_j, or t_j equals t_k [24]. In this work it is assumed that for any two events, one precedes the other. This modeling approach will be extended in future work to include coincidental events or other temporal combinations.

In the harbor model, a vessel's traversal of a harbor is represented as an event sequence referred to as a *transit* $_{ID}T$, where *ID* is the *identity* (i.e., *vesselID*) of the vessel associated with the transit. A transit models all the occurrents experienced by a particular vessel while it is traversing the harbor, (e.g., a sequence of ChangeZoneEvents) formally defined as $_{ID}T = \{e \in E \mid e.vesselID = ID\}$. For example, the event sequence that models a ferry's traversal of the harbor, starting from outside the harbor limits in the offshore zone and ending upon reaching its destination the ferry landing (Figure 4), is a transit consisting of five ChangeZoneEvents, and is represented as:

$$_{FRY}\,cze_{t1}^{inbd\_SE\_TSS} > {}_{FRY}\,cze_{t2}^{precautionary} > {}_{FRY}\,cze_{t3}^{approach} > {}_{FRY}\,cze_{t4}^{bay} > {}_{FRY}\,cze_{t5}^{ferry\_landing}$$

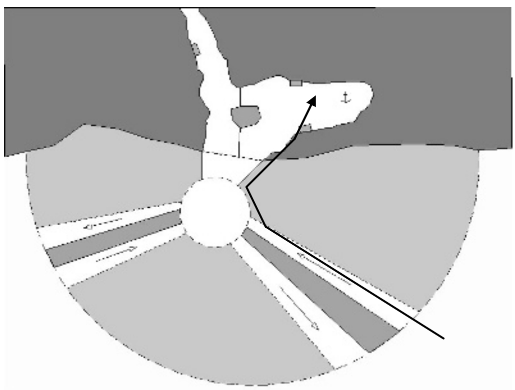


Fig. 4. Transit from offshore to the ferry landing

Transits in their simplest form consist of an ordering of only *ChangeZoneEvents*. The representation of a transit (or any other domain event sequence) can be extended, however, to incorporate different types of events, especially noteworthy events that carry special domain semantics and trigger a notification when they take place in a domain. Including specialized subclasses of events or *derived events* in event sequences, such as transits, results in a richer more semantically meaningful depiction of the modeled domain. In the following subsections, we discuss derived events and their role in event sequences. We also examine noteworthy events in sequences and how noteworthy sequences arise from scenarios in dynamic geospatial domains.

5.1 Derived Events in Sequences

Representing a transit as a sequence of *ChangeZoneEvents* is a rather low-level semantic description of the movement of a vessel. Ideally, we would like this representation to be closer to reality and incorporate more of the semantics associated with the vessel's movement. *ChangeZoneEvents* can be translated into semantically richer derived or noteworthy events that in turn can be included in an event sequence to provide a more realistic representation of a vessel's traversal through the harbor. For example, the transit of vessel OT1, whose intended destination was the oil terminal, consists of five *ChangeZoneEvents* that occur in the following zones: upon entering the inbound lane of the west traffic separation scheme, in the precautionary area, the approach zone, the bay, and finally the anchorage. This sequence of events is represented as:

$$ {}^{\mathit{oldZone:\ offshore}}_{OT1}cze^{\mathit{inbd\_W\_TSS}}_{t1} > {}_{OT1}cze^{\mathit{precautionary}}_{t2} > {}_{OT1}cze^{\mathit{approach}}_{t3} > {}_{OT1}cze^{\mathit{bay}}_{t4} > {}^{\mathit{dest:oil\_terminal}}_{OT1}cze^{\mathit{anchorage}}_{t5} $$

Applying the conditions defined in section 4.1 for derived events further distinguishes cze_{t1} as a *DepartEvent,* and cze_{t5} as an *ArrivalEvent*. A derived event captures the same attributes as its base event plus the semantics associated with its defining condition (e.g., that the *oldZone* attribute value is a departure zone). Therefore, the sequence can be revised as:

$$ {}^{\mathit{oldZone:\ offshore}}_{OT1}dep^{\mathit{inbd\_W\_TSS}}_{t1} > {}_{OT1}cze^{\mathit{precautionary}}_{t2} > {}_{OT1}cze^{\mathit{approach}}_{t3} > {}_{OT1}cze^{\mathit{bay}}_{t4} > {}^{\mathit{dest:oil\_terminal}}_{OT1}arr^{\mathit{anchorage}}_{t5} $$

Applying the constraints defined in section 4.2 identifies the *ArrivalEvent* as an *UnexpectedDestinationEvent* (i.e., a noteworthy derived event) since the constraint that an *ArrivalEvent* must take place in the vessel's intended destination zone was violated. Therefore the sequence can be further refined to include the semantics represented by the noteworthy *UnexpectedDestinationEvent* and written as:

$$ {}^{\mathit{oldZone:\ offshore}}_{OT1}dep^{\mathit{inbd\_W\_TSS}}_{t1} > {}_{OT1}cze^{\mathit{precautionary}}_{t2} > {}_{OT1}cze^{\mathit{approach}}_{t3} > {}_{OT1}cze^{\mathit{bay}}_{t4} > {}^{\mathit{dest:oil\_terminal}}_{OT1}ude^{\mathit{anchorage}}_{t5} $$

Including derived events in event sequences results in a more semantically meaningful representation of the occurrents experienced by moving entities in a domain. Additionally, this more meaningful description of movement requires no more events than the original sequence on which it is based.

5.2 Noteworthy Events in Sequences

We have discussed how to model the occurrents in a geospatial domain as events, and how events can be ordered to form event sequences that model, for example, a vessel's traversal through a harbor. In this section, we discuss the semantics associated with noteworthy events in sequences.

Transits may include *a priori* noteworthy events, such as a *NotUnderCommandEvent* or a *DistressEvent.* For example, vessel KW1 loses steering and upon completing necessary repairs, continues towards its destination. This portion of the traversal of KW1 is represented by:

$$ \ldots > {}_{KW1}nuc^{\mathit{precautionary}}_{t3} > {}_{KW1}uwy^{\mathit{precautionary}}_{t4} > \ldots $$

This sequence shows an *a priori* noteworthy *NotUnderCommandEvent* nuc_{t3}, modeling the loss of steering, and a subsequent *UnderwayEvent* uwy_{t4}, indicating the vessel's ability to maneuver to avoid collision has been restored. Event sequences may also contain noteworthy derived events, such as *UnexpectedDestinationEvents* or *UnexpectedZoneEvents*. For example, vessel BLK was expected to enter next the precautionary zone during a transit. However, the next *ChangeZoneEvent* takes place in the south fishing and recreation zone, triggering a noteworthy derived *Unexpected-ZoneEvent*:

$$\ldots > {}^{expectedNext:precautionary}_{BLK}\,cze^{inbd\_W\_TSS}_{t2} > {}_{BLK}\,uze^{S\_F\&R}_{t3} > \ldots$$

Event sequences, therefore, represent a mixture of event semantics, modeling both routine and unexpected occurrents in a domain.

5.3 Noteworthy Event Sequences

An event sequence containing at least two noteworthy events is defined as a *noteworthy event sequence*. Modeling noteworthy sequences provides a foundation for analyzing the entire sequence for unusual or suspicious trends, as opposed to considering individual noteworthy events in isolation. Treating the sequence as a whole is also useful for confirmation that the dynamics of a domain are progressing normally. As an example, consider the following subsequence of a transit:

$$\ldots > {}_{LNG1}\,nuc^{inbd\_W\_TSS}_{t2} > {}_{LNG1}\,uwy^{inbd\_W\_TSS}_{t3} >$$

$${}^{expectedNext:approach}_{LNG1}\,cze^{precautionary}_{t4} > {}^{expectedNext:precautionary}_{LNG1}\,uze^{E\_F\&R}_{t5} > \ldots$$

This transit contains two noteworthy events, *NotUnderCommandEvent nuc* and the derived noteworthy *UnexpectedZoneEvent* uze_{t5}. As a result of highlighting this sequence as noteworthy, users can be alerted to consider the entire event sequence in determining what remedial actions are warranted. By considering the entire sequence it appears that the repairs necessitated by the *NotUnderCommandEvent* were inadequate, since an *UnexpectedZoneEvent* uze_{t5} takes place after LNG1 resumed its traversal of the harbor (i.e., after the *UnderwayEvent* uwy_{t3} took place). If noteworthy events are considered in isolation underlying trends may be missed, resulting in an inappropriate response.

5.4 Additional Semantics Associated with Event Sequences

A sequence of events can be useful in identifying unexpected movement patterns in a domain. For example, some transits may be incomplete or represent an aborted traversal of the harbor, as when a vessel enters the harbor waters and then leaves the harbor again without ever reaching its intended destination. For example, the ferry FRY1 has an intended destination of the ferry landing and its transit consists of the following events: a *DepartEvent* on entering the inbound lane of a traffic separation scheme from the offshore zone, a *ChangeZoneEvent* on entering the precautionary area, where the expected next zone is the approach zone, an *UnexpectedZoneEvent* on entering the outbound lane of a traffic separation area, and an *UnexpectedDestinationEvent* on entering the offshore zone:

$$ {}^{oldZone:\ offshore}_{FRY1}\, dep^{inbd\_W\_TSS}_{t1} > {}^{expectedNext:\ approach}_{FRY1}\, cze^{precautionary}_{t2} > $$
$$ {}_{FRY1}\, uze^{outbd\_SE\_TSS}_{t3} > {}^{dest:\ ferryLanding}\, ude^{offshore}_{t4} $$

This sequence contains an *UnexpectedDestinationEvent*, i.e., this vessel's intended destination is the ferry landing, yet the transit ended in the offshore zone (Figure 5a). In addition, the transit started from and ended in the same zone. This characteristic transit is referred to as an ***aborted transit***. An aborted transit $T_{aborted}$ is defined as a transit T where $\exists\, e, e' \in T$ such that e is a *DepartEvent dep*, e' is an *UnexpectedDestinationEvent ude*, and *e.oldZone=e'.zone*., meaning a transit containing an *UnexpectedDestinationEvent* such that the departure zone (*e.oldZone*) and destination zone (*e'.zone*) are the same. A different example of this type of movement pattern is FRY1 departing the ferry landing for sea (i.e. an intended destination of offshore), but unexpectedly returning to the ferry landing instead (Figure 5b).

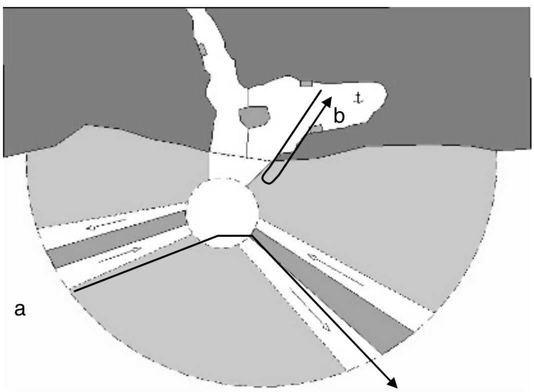


Fig. 5. Aborted transits: (a) an aborted transit where a vessel with an intended destination of the ferry landing exits the harbor waters without reaching its destination and (b) a vessel departs the ferry landing for sea, but unexpectedly returns to the ferry landing

6 Conclusions and Future Work

A realistic data model of a dynamic geospatial domain must include representations of the occurrents that happen as well as the continuants that exist in the domain. In this paper, these entities are modeled as classes of *events* and *objects* respectively. The range and semantic meaning of the events in the model can be extended with *derived events*, defined as specialized subclasses of an event class. *Noteworthy events* have additional higher order semantics associated with them in that they are exceptional as opposed to routine in nature. Such noteworthy events may be defined *a priori* (i.e., an event class is designated as noteworthy and every instance of it is noteworthy) or in terms of a derived event. Examples of noteworthy events modeling the significant occurrents of a harbor domain include, *NotUnderCommandEvents*, *UnexpectedDestinationEvents*, and *UnexpectedZoneChangeEvents*. Incorporating events into data models not only provides a more realistic account of a dynamic geospatial domain, but also provides the means for modeling objects' movements as sequences of events. In this paper, we also present some of the semantics associated with event sequences as a whole, for example, noteworthy sequences and aborted transits. For these cases, examining an entire sequence of events reveals certain semantics that would have been missed if events were considered singly. Although the domain described in this paper is a harbor, we believe that the concepts presented in this paper are general and hold for many other domains, including, for example, road transportation or environmental modeling.

Future topics for research involve investigating how combinations of two or more events form distinctive patterns based on the relative values of attributes common to the events, as well as the semantics associated with the patterns. In addition, implementations of event notification services based on automated responses to noteworthy events need further development.

Acknowledgments

This research is supported by a grant from the National Geospatial-Intelligence Agency under grant number NMA201-00-1-2009.

References

1. Worboys, M. A unified model of spatial and temporal information. *Computer Journal*, 37: 26–34, 1994.
2. Hornsby, K. and Egenhofer, M. Identity-based change: A foundation for spatio-temporal knowledge representation. *International Journal of Geographical Information Science*, 14: 207–204, 2000.
3. Peuquet, D. *Representations of Space and Time*, New York: Guilford, 2002.
4. Galton, A. Desiderata for a spatio-temporal geo-ontology. In Kuhn, W., Worboys, M., and Timpf, F. (eds.) *Spatial Information Theory: Foundations of Geographic Information Science* (COSIT 2003). Lecture Notes in Computer Science Vol. 2825. Springer, Berlin, pp. 1-12, 2003.
5. Grenon, P. and Smith B. SNAP and SPAN: Towards dynamic spatial ontology. *Journal of Spatial Cognition and Computation*, 4: 69-104, 2004.
6. Bittner, T., Donnelly, M. and Smith, B.: Endurants and perdurants in directly depicting ontologies, *AI Communications*, IOS Press, 17: 247-258, 2004.
7. Snodgrass, R., (ed.). *The TSQL2 Temporal Query Language*. Kluwer Academic Publishers, 1995.
8. Abraham, T. and Roddick, J. Survey of spatio-temporal databases. *GeoInformatica*, 3: 61-99, 1999.
9. Yuan, M. Representing complex geographic phenomena in GIS. *Cartography and Geographic Information Science*, 28: 83-96, 2001.
10. Worboys, M. Modeling changes and events in dynamic spatial systems with reference to socio-economic units, In *Life and Motion of Socio-Economic Units*, Frank, A., Raper, J., and Cheylan, J. (eds.), New Zealand, MacMillan. 8: 129-138, 2001.

11. Pfoser, D., Jensen C., and Theodoridis, Y. Novel approaches in query processing for moving object trajectories. In *Proceedings of the 26th International Conference on Very Large Data Bases* (VLDB 2000), Cairo, Egypt, Morgan Kaufmann Publishers Inc., pp. 395-406, 2002.
12. Song, Z., and Roussopoulos, N. K-Nearest neighbor search for moving query point. In *Proceedings of the 7th Intl. Symposium on Advances in Spatial and Temporal Databases*, Springer-Verlag, Berlin, Heidelberg, New York, pp. 79–96, 2001.
13. Papadias, P., Tao, Y. Kalnis, P. and Zhang, J. Indexing spatio-temporal data warehouses. In *18th International Conference on Data Engineering*, San Jose, CA, IEEE Computer Society, pp. 166-175, 2002.
14. Worboys, M. Event-oriented approaches to geographic phenomena. *International Journal of Geographical Information Science*, 19: 1-28, 2005.
15. Worboys, M. and Hornsby, K. From objects to events: GEM, the geospatial event model. In: Egenhofer, M., Freksa, C., and Miller, H. (eds.) *Proceedings of the Third International Conference on Geographic Information Science* (GIScience 2004) Adelphi, MD, Lecture Notes in Computer Science Vol. 3234. Springer, New York, pp. 327-344, 2004.
16. Galton, A. Fields and objects in space, time, and space-time. *Journal of Spatial Cognition and Computation*, 4: 39-68, 2004.
17. Hornsby, K. and Egenhofer, M. Modeling moving objects over multiple granularities. *Annals of Mathematics and Artificial Intelligence*, Kluwer Academic Press, 36: 177-194, 2002.
18. Wolfson, O. Moving objects information management: The database challenge, In: Halevy, A., Avigdor, G. (eds.) *Proceedings of the 5th Workshop on Next Generation Information Technologies and Systems* (NGITS'2002), Caesarea, Israel, June 25-26, 2002, Lecture Notes in Computer Science Vol. 2382. Springer, Berlin, Heidelburg, New York, pp. 75-89, 2002.
19. Xie, R. and Shibasaki, R. A unified spatio-temporal schema for representing and querying moving features. *SIGMOD Record*, 34: 45-50, 2005.
20. Dayal, U., Buchmann, A., McCarthy, D. Rules are objects too: A knowledge model for an active, object-oriented database system. In Dittrich (ed.), *Advances in Object-Oriented Database Systems*, Springer, New York , pp. 129-143, 1988.
21. Carzaniga, A., Rosenblum D., Wolf, A. Achieving scalability and expressiveness in an internet-scale event notification service. In *Proceedings of the 19th Annual ACM Symposium on Principles of Distributed Computing* (PODC 2000), Portland, OR, ACM Press, pp. 219-227, 2002.
22. Hinze, A. and Voisard A. A flexible parameter-dependent algebra for event notification services. Technical Report Number tr-b-02-10, Freie Universitat Berlin, pp. 1-17, 2002.
23. USCG (2003) Universal shipborne automatic identification system transponder. United States Coast Guard. (http://www.navcen.uscg.gov/marcomms/ais.htm), 2005.
24. Galton, A. *Qualitative Spatial Change*, Oxford: Oxford University Press, 2000.

Measuring Semantic Similarity Between Geospatial Conceptual Regions

Angela Schwering[1, 2] and Martin Raubal[2]

[1] Ordnance Survey of Great Britain,
United Kingdom
[2] Institute for Geoinformatics,
University of Muenster, Germany
{angela.schwering, raubal}@uni-muenster.de

Abstract. Determining the grade of semantic similarity between geospatial concepts is the basis for evaluating semantic interoperability of geographic information services and their users. Geometrical models, such as conceptual spaces, offer one way of representing geospatial concepts, which are modelled as n-dimensional regions. Previous approaches have suggested to measure semantic similarity between concepts based on their approximation by single points. This paper presents a way to measure semantic similarity between conceptual regions—leading to more accurate results. In addition, it allows for asymmetries by measuring directed similarities. Examples from the geospatial domain illustrate the similarity measure and demonstrate its plausibility.

1 Introduction

Semantic similarity measurements between concepts are the basis for establishing semantic interoperability of information services. To ensure successful communication between *geographic* information services and their users, it needs to be determined how similar their used *geospatial* concepts are. There exist various approaches to measure such similarity between concepts, depending on the concepts' types of representation. A common approach to representing concepts is based on geometrical models, where concepts are modelled as n-dimensional regions. Semantic similarity between concepts has previously been determined by approximating the regions through points and then measuring the distances between them. Such approximation inevitably leads to a loss of information and is therefore an inaccurate measure of similarity between concepts. In this paper, we present an approach of measuring semantic similarity between conceptual regions instead of their pointwise estimates. Such method improves the quality of the measurements by enhancing the accuracy of its results.

For the formal representation of conceptual regions we utilize Gärdenfors' idea of conceptual spaces—sets of quality dimensions within a geometrical structure [1]. Concepts can then be represented as n-dimensional regions in a vector space. The measurement of semantic similarity between conceptual regions is based on applying previously defined distance measures, such as given by instances of the Minkowski metric, to vectors forming the convex boundaries of the concepts whose similarities

M.A. Rodríguez et al. (Eds.): GeoS 2005, LNCS 3799, pp. 90 – 106, 2005.

get evaluated. This way, a semantic distance function between two conceptual regions can be established. Furthermore, with the presented approach it is possible to account for asymmetries of similarity judgements, i.e., concepts are judged to be more similar to their superconcepts than vice versa. Examples from the geospatial domain are used throughout the paper to illustrate the semantic similarity calculations and their interpretations: we represent concepts such as mountain, hill, and lowland in a conceptual space according to their shapes through the dimensions 'height' and 'width'.

Section 2 introduces formal conceptual spaces and gives an overview of geometrical similarity measures. Section 3 describes the semantic similarity measure for conceptual regions and thereby applies the Euclidean distance function for calculation of similarity values. In section 4 we demonstrate how the proposed measure accounts for the fact that similarity judgements may be asymmetric. Section 5 focuses on the illustration and interpretation of distance values in different topological configurations (e.g., meeting and overlapping) of conceptual regions. The final section provides conclusions and directions for future work.

2 Related Work

This section defines formal conceptual spaces and introduces geometrical similarity measures.

2.1 Formal Conceptual Spaces

The idea of a conceptual space was introduced by Peter Gärdenfors as a framework for representing information at the conceptual level [1]. Such representation rests on the foundation of cognitive semantics [2], asserting that meanings are mental entities—mappings from expressions to conceptual structures, which themselves refer to the real world. Conceptual spaces can be utilized for knowledge representation and sharing, and support the paradigm that concepts are dynamical systems [3]. According to Gärdenfors, a conceptual space is a set of quality dimensions with a geometrical or topological structure for one or more domains. A domain is represented through a set of integral dimensions, which are distinguishable from all other dimensions. For example, the colour domain is formed through the dimensions hue, saturation, and brightness. Concepts cover multiple domains and are modelled as n-dimensional regions. Every object or member of the corresponding category is represented as a point in the conceptual space. This allows for expressing the similarity between two objects as the distance between their points in the space. Recent work by Gärdenfors deals with the idea of representing actions and functional properties in conceptual spaces [4].

In [5], a methodology to formalize conceptual spaces as vector spaces is presented. Formally, a conceptual vector space is defined as $\mathbf{C}^n = \{(c_1, c_2, \ldots, c_n) \mid c_i \in \mathbf{C}\}$ where the c_i are the quality dimensions. A quality dimension can also represent a whole domain and in this case $c_j = \mathbf{D}^n = \{(d_1, d_2, \ldots, d_n) \mid d_k \in \mathbf{D}\}$. The fact that vector spaces have a metric allows for the calculation of distances between points in the space. This can also be utilized for measuring distances between concepts, although it requires their approximation by “prototypical points.” In order to calculate these so-called *semantic distances* between instances of concepts all quality dimensions of the space must be represented in the same relative unit of measurement. Assuming a normal distribution, this is ensured by calculating the z scores for these values, also called z-transformation [6]. For specifying different contexts one can assign weights to the quality dimensions of a conceptual vector space. This is essential for the representation of concepts as dynamical systems. In this case $\mathbf{C}^n$ is defined as $\{(w_1c_1, w_2c_2, \ldots, w_nc_n) \mid c_i \in \mathbf{C}, w_j \in \mathbf{W}\}$ where **W** is the set of real numbers.

2.2 Geometrical Similarity Measures

There exist a number of approaches to assess semantic nearness in a conceptual space with quite differing philosophies: some focus on angle or length difference, and others on the distance between vectors. Following Jones and Furnas [7] we choose a geometric representation with iso-similarity contours to demonstrate the semantic differences of the similarity functions: Moving an object along a contour line—analogous to contours in topographic maps—does not have an effect on its similarity value.

Euclidian- and City-Block Distance Measure

The most common way of measuring similarity in conceptual spaces is the Minkowski metric (equation 1) which measures semantic distance in analogy to spatial distance. The Minkowski metric is a generic formula: r=1 results in the city-block distance and r=2 in the Euclidian distance. According to the city-block metric the distance equals the sum of the absolute distances of each dimension and the Euclidian distance is computed as the square root of the sum of the dimension-wise squared differences [8]. Similarity s is a linear decaying function of the semantic distance d [9, 10].

$$d(q,c) = \left[\sum_{i=1}^{n} |q_i - c_i|^r \right]^{1/r} \tag{1}$$

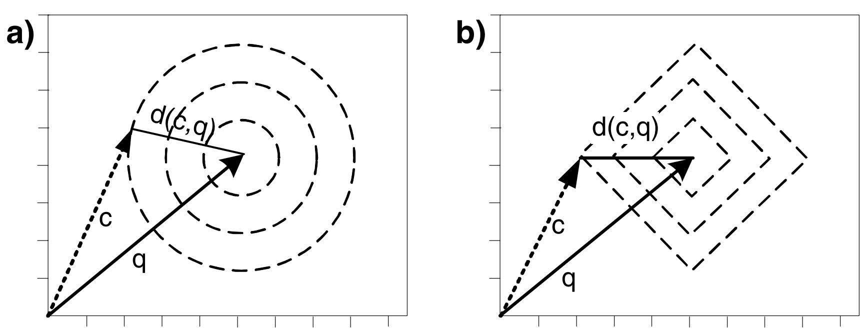


Fig. 1. Geometric comparison of the Euclidian (a) and the city-block distance measure (b)

Figure 1 shows the iso-similarity contours for the Euclidian and city-block metrics: Vectors along one contour line all have the same similarity to the query vector. In the two-dimensional figures the Euclidian similarity contours are circular and the city-block contours are quadratic.

Johannesson and Gärdenfors demonstrated in experimental studies—subjects had to rate the similarity between different mollusc shells and beetles—the usability of the Minkowski metric—especially the different underlying assumptions when Euclidian and city-block metrics are applied—within conceptual spaces [1, 11-13].

Cosine Similarity Measure

The cosine measure (equation 2) is a normalized inner product of two vectors: The inner product is divided by the product of the Euclidian vector lengths.

$$s(q,c) = \frac{\sum_{i=1}^{n} q_i * c_i}{\sqrt{\sum_{k=1}^{n} (q_k)^2} * \sqrt{\sum_{j=1}^{n} (c_j)^2}} \tag{2}$$

Because of the Euclidian length normalization, all vectors having the same direction are transformed into the same unit vector regardless of their length (figure 2). Therefore all vectors on the same radiating line—the iso-similar contour—have the same similarity. Only the angle of separation influences the similarity value: The greater the angle between two vectors the lower is their similarity. In the two-dimensional representation of figure 2 the iso-similarity contours are lines with symmetrical similarity values on both sides of the query vector, but in an n-dimensional space the contours are cone-shaped. The cosine similarity measure is bounded from zero to one [7]. It is used, for example, to determine similar terms in a concept space [14]. The pseudo-cosine measure shown on the right side of figure 2 is similar to the cosine measure but normalized by the city-block length of the vectors.

This set of similarity measures was chosen because they are most frequently used in cognitive spaces, but there exists a variety of other similarity measures for vector

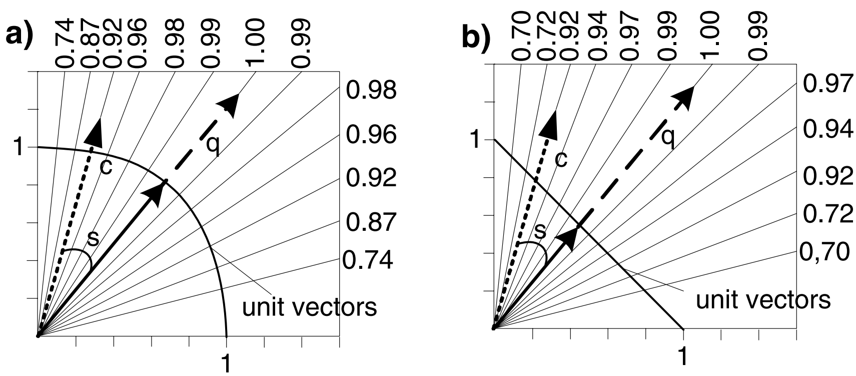


Fig. 2. Geometric comparison of the Cosine (a) and Pseudo-Cosine similarity measure (b) [7]

spaces such as the dice measure, overlap measure and the Jaccard measure [7]. Besides the mentioned similarity measures for vector spaces there exist a number of other approaches to measure semantic similarity such as the feature matching model [15], Matching-Distance Similarity Model [16, 17], Distance [18] and the transformational model [19]. However, these are based on different representational models. Here, we consider only geometric similarity measures based on conceptual spaces.

Previous distance measures for concepts in conceptual spaces first reduce concepts to a single point such as the balance point or centre point—often representing their prototypes [5, 20]. Then similarity measures for instances are applied [21]. By reducing concepts to single points or instances, the expressiveness as well as the significance of the distance measure are reduced. Neither a balance point nor the centre point can fully represent the semantics of a concept. The semantic similarity measure proposed in the following section overcomes this shortcoming.

3 Semantic Similarity Between Geospatial Concepts

Current similarity measures confine themselves to estimating the similarity between instances, i.e. between points[1] in the vector space. The semantic similarity measure between geospatial concepts presented here is based on the similarity measure of instances, but measures the distance between concepts represented as convex regions in space. Similarity gets calculated in a two-step process: At first the concepts are reduced to a set of vector pairs. This way we transform the convex regions of concepts into a format to which the original similarity measures can be applied. Then the similarity measure is used for this set of vector pairs to calculate the similarity value. After stating some preliminary assumptions, section 3.2 explains how to calculate the vector pairs between concepts and section 3.3 focuses on the calculation of the similarity value.

3.1 Preliminary Assumptions

Before describing the similarity measurement procedure we need to introduce some preliminary assumptions: A concept is modelled as a convex region in an n-dimensional space, i.e. the region is continuous, completely closed and the hull of the region is convex. Extreme distance values occur for position vectors whose end-points lie on the hull of the region. Figure 4 illustrates that the distances measured from a vector in the inside of query concept 'hill' (vector q_i) lie between the maximum and minimum distance values measured somewhere on the hull of the region (here vectors q_3 and q_6). Therefore it is sufficient to consider only the hull of the region representing query concept 'hill' to estimate the distance values.

The considerations about the similarity measure start with the simplifying model that the convex hull consists of a set of discrete points. Later we transfer the findings to a continuous function. We further assume that all concepts have the same dimensions.

---

[1] A point is represented by a vector in a vector space. We use the term 'point' to underline the difference between instances modelled as points and concepts modelled as regions.

Like the similarity measures in section 2.2 the illustration is simplified by representing concepts in a two-dimensional space. In the text we explain the findings for n-dimensional spaces.

3.2 Calculating Vector Pairs Between Concepts

The hull of query concept Q is formed by the endpoints of a set of position vectors $S_Q=\{q_1, q_2, \ldots, q_n\}$. The first step of the distance calculation between Q and a concept C aims at defining for each vector q_i in S_Q one or several corresponding vectors of C. All vectors form vector pairs with q_i and these are the basis for the similarity calculation. The identification which vector pairs reflect best human similarity measurement is an important question. Figure 3 illustrates three different strategies.

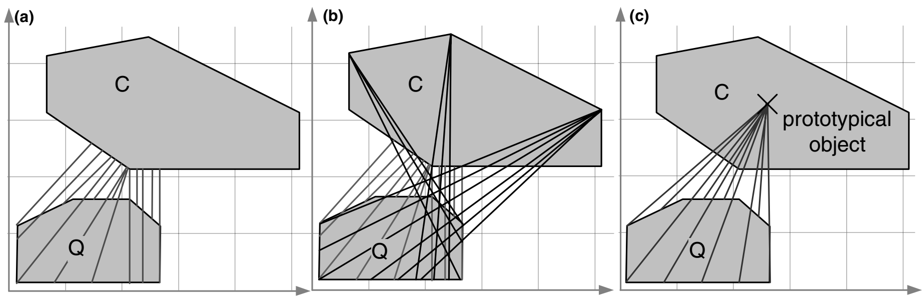


Fig. 3. Different strategies can be applied to identify for each vector q_i one or several corresponding vectors of a concept C: (a) searching for the vector with minimum distance, (b) searching for the vectors with either minimum or maximum distance, or (c) defining a prototypical object as reference object for the similarity calculations

Strategy (a) is inspired by the idea that humans intuitively assess similarity by comparing a concept Q with that instance of the other concept C which is most similar to Q. Therefore the corresponding vector has the minimum distance to a vector q_i.

Strategy (b) supposes that not only the most similar, but also the most dissimilar instances of a concept influence human similarity assessment. Therefore there exist two corresponding vectors of C: one with minimum and another with maximum distance to a vector q_i.

Strategy (c) assumes that there exists a reference instance—e.g. a prototype of the concept—which is used as corresponding vector for all vectors q_i. In this case the similarity measure is only influenced by the shape of the conceptual region Q and concept C is reduced to one prototypical point.

Extensive human subject tests are required to substantiate the choice for one strategy. Such investigation is important, but lies outside the scope of this paper which focuses on developing a formal procedure for measuring similarity between conceptual regions.

$$nearVec(q_i) = \min_{\forall x \in C}\left(\left|f_c(x) - q_i\right|\right),\ q_i \in S_Q \tag{3}$$

For the following calculations we apply strategy (a). Therefore we compute for each vector q_i in S_Q the vector c in C with the minimum distance to q_i according to the following formula (equation 3). Concept C is given by the function $f_c(x)$.

$$\left(q_i, nearVec(q_i)\right),\ q_i \in S_Q \tag{4}$$

This vector pair (equation 4) consisting of vector q_i and the nearest vector in concept C *nearVec(q_i)* specifies the corresponding vectors between which the semantic distance is measured.

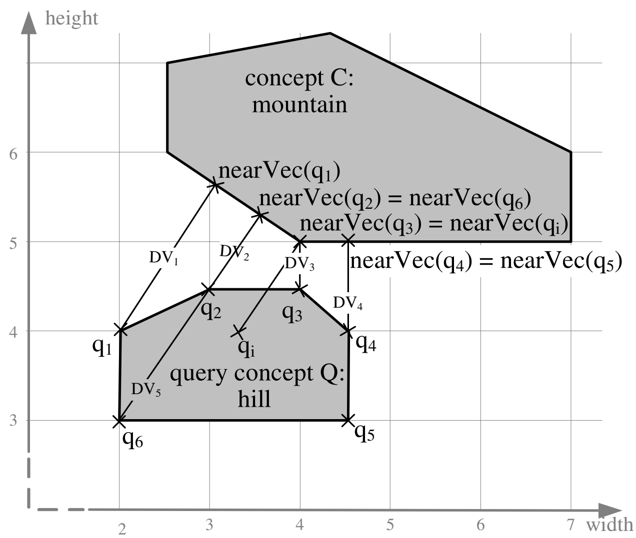


Fig. 4. Computation of corresponding vectors from query concept Q to concept C[2]

Figure 4 illustrates the procedure: for each point q_i on the hull of query concept 'hill' one searches for the vector of concept 'mountain' with the minimum distance as illustrated by the difference vector DV_i. The vector pair determination depends also on the applied similarity measure: For cosine measures strategy (a) aims at minimizing angle size and therefore computes for each vector q_i in S_Q the position vector c of C with the smallest angle difference.

[2] For demonstration purposes we represent the concepts in the example by two dimensions only. For a complete description though, more dimensions such as 'shape' etc. are needed. The dimensions of the conceptual space 'width' and 'height' are measured in some standardized unit. For the calculation of semantic distances it is required to represent all dimensions in the same relative unit of measurement. The original, non-standardized units—in this example width and height can be measured in meters or kilometres—are standardized by the z-transformation [6].

3.3 Applying Euclidian Distance Measure to Calculate Distance Value

All existing similarity measures explained in section 2 can be applied to the vector pairs introduced in section 3.2. This paper focuses on the most commonly used semantic similarity measure in conceptual spaces: the Euclidian distance measure.

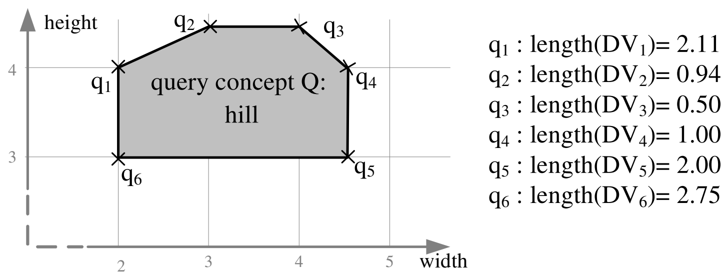


Fig. 5. Distances illustrated as values for each point q_i

The Euclidian distance between two vectors in an n-dimensional space is measured by calculating the absolute difference vector DV_i. The length of vector DV_i is the semantic distance value for point q_i. Therefore each vector q_i of the hull of Q has a distance value to concept C. The hull of an n-dimensional region is an (n-1)-dimensional object in an n-dimensional space.

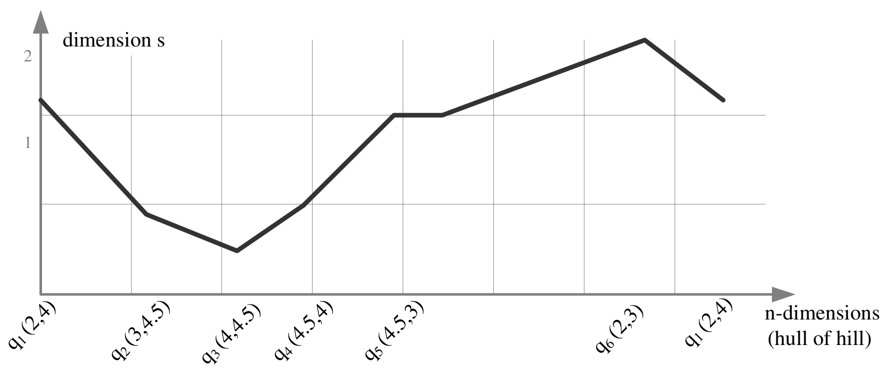


Fig. 6. Distances illustrated as distance function $DV(q_i)$

The distance values can be represented on a dimension s. For better understanding we illustrate the distance values in two different ways: In figure 5 dimension s is represented by assigning the distance values to each point of the hull of query concept

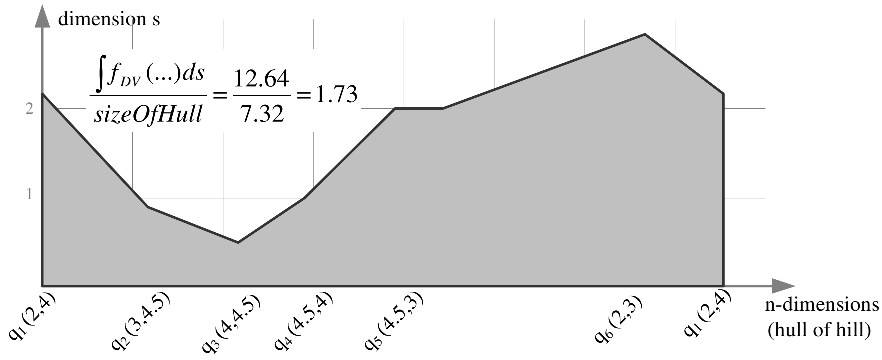


Fig. 7. Transformation of the pointwise Euclidian distance values into one distance value

'hill'. Figure 6 shows the dimension s as a function of the hull of concept 'hill'. Both figures illustrate the same fact using different representations.

The Euclidian distance measure evaluates the length of the distance vectors. Distances near to zero indicate that concept 'hill' lies very close to concept 'mountain'. A small value on dimension s stands for high similarity. The minimum distance is for vector (4,4.5) and the maximum distance is for vector (2,3).

$$\frac{\int f_{DV}(...)ds}{sizeOfHull} \tag{5}$$

Currently with the finite set of vectors q_i the semantic distance can only be approximated. For a continuous computation we use the integral over the semantic distance function. Since the integral depends not only on the value of dimension s, but also on the size of the concept's hull—the bigger the hull the greater is the interval on the hull-dimensions—a normalization factor such as the size of the hull is required. Figure 7 shows the computation of the integral with respect to distance dimension s to estimate the distance value (equation 5). The semantic distance from concept 'hill' to concept 'mountain' is 1.73.

4 Asymmetric Similarities

Psychologists found that the perceived similarity between two stimuli is not necessarily symmetric: non-prominent objects are more similar to a prominent object than vice versa [20]. In 1977, Amos Tversky [22-24] proposed a similarity measure allowing for asymmetric similarities. Geometric similarity measures are based on multidimensional spaces and assume metric properties such as minimality, symmetry and triangle inequality between items. The inability to deal with asymmetric similarities between objects and concepts is probably the most heavily criticized aspect of geometric similarity measures and was the reason for various extensions of conceptual spaces [25, 26].

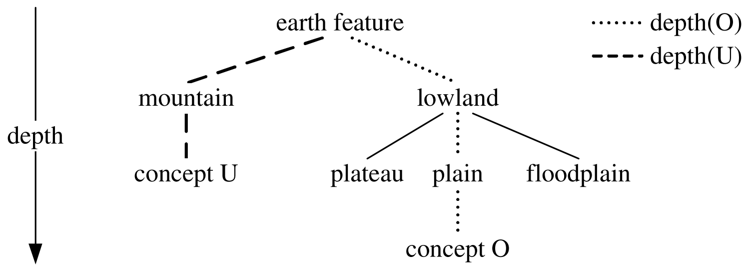


Fig. 8. Depth of concepts in the concept hierarchy (see also [16])

In the geospatial domain, Rodriguez and Egenhofer developed the Matching-Distance Similarity Measure (MDSM) which accounts for asymmetries in similarity assessment [16, 17, 27, 28]: One component of the MDSM is Tversky's Feature Matching Model, which becomes an asymmetric measure depending on the choice of parameters α and β in the contrast model [22, 23].

$$\alpha(U,O)=\begin{cases}\dfrac{depth(U)}{depth(U)+depth(O)}, depth(U)\leq depth(O)\\ 1-\dfrac{depth(U)}{depth(U)+depth(O)}, depth(U)>depth(O)\end{cases} \quad (6)$$

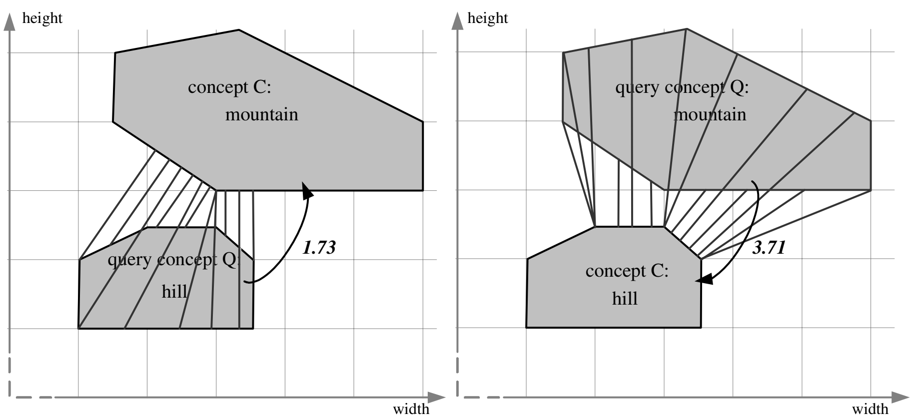


Fig. 9. Asymmetric semantic distance between two concepts

Following the idea that people perceive similarity from a subconcept to its superconcept greater than the similarity from the superconcept to the subconcept, and that the superconcept is commonly used as a basis for the similarity judgment, Rodriguez and Egenhofer developed a formula (equation 6) to compute these parameters by building the ratio of the concepts' depth in the ontology (figure 8). In general, concepts deeper in the hierarchy are smaller—in terms of size of their represented conceptual regions—than their superconcepts, which have a greater degree of generalization.

The semantic distance measure proposed in this paper reflects the above described observation: concepts with a high degree of specialization covering a small region in the conceptual space tend to be more similar to general concepts than vice versa. Figure 9 shows how the similarity measure between concepts works and illustrates the effect of asymmetry. On the right hand side the query concept Q is much broader than the compared concept C. Therefore the semantic distance is greater and the similarity value is smaller than in the figure on the left side.

5 Illustration and Interpretation of the Distance Value

The following examples illustrate the results of the proposed semantic similarity measure for different topologic configurations and give an interpretation of the semantic distance values. Disjoint concepts were already discussed in section 3. Here we focus on meeting, overlapping, inside/containing and covering/covered by concepts. We refer to Egenhofer's definition of the topologic operators 'disjoint', 'meet', 'overlap', 'inside', 'contains', 'covers' and 'covered by' [29, 30].

5.1 Meeting and Overlapping Concepts

Figure 10 shows two meeting concepts 'steep face' and 'mountain' and their semantic distance function. For the interval where the conceptual regions 'steep face' and 'mountain' meet, the semantic distance is zero. The length of the interval is the same as the length of the contact. From this distance function it is not possible to distinguish whether 'steep face' meets 'mountain' from outside—the interiors of 'steep

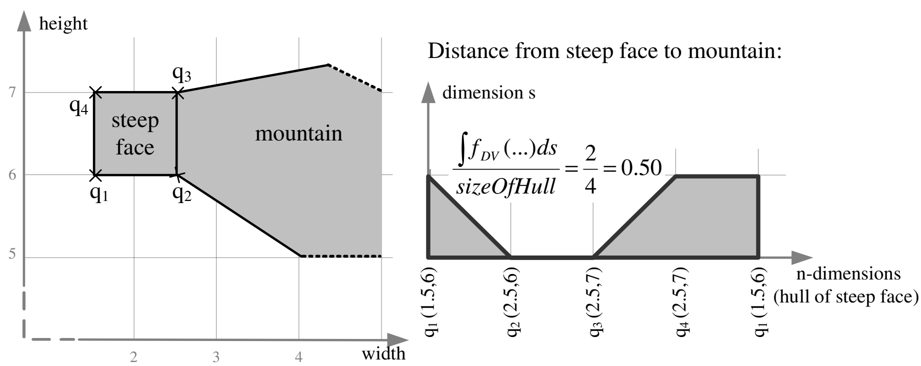


Fig. 10. Meeting concepts and their semantic distance function

face' and 'mountain' do not intersect, but their boundaries do—or one covers the other, i.e. their interiors and boundaries intersect (see also section 5.2).

Figure 11 shows two overlapping concepts with their semantic distance function. The semantic distance between overlapping concepts is zero for the whole overlap. Therefore we cannot resolve the difference between meeting and overlapping concepts purely from the distance function. However, from the distance function one can gather information about the topology of both concepts:

1. If the distance function becomes zero, both concepts either meet (figure 10), overlap (figure 11), the query concept covers the other concept ('plateau' covers 'lowland' in figure 12), or it is inside another concept ('plain' is inside 'lowland').
2. If the distance function does not become zero at any time and every vector in C is also vector in Q, then concept Q contains concept C ('lowland' contains 'plain' in figure 12). If such a vector does not exist, then concepts Q and C are disjoint.

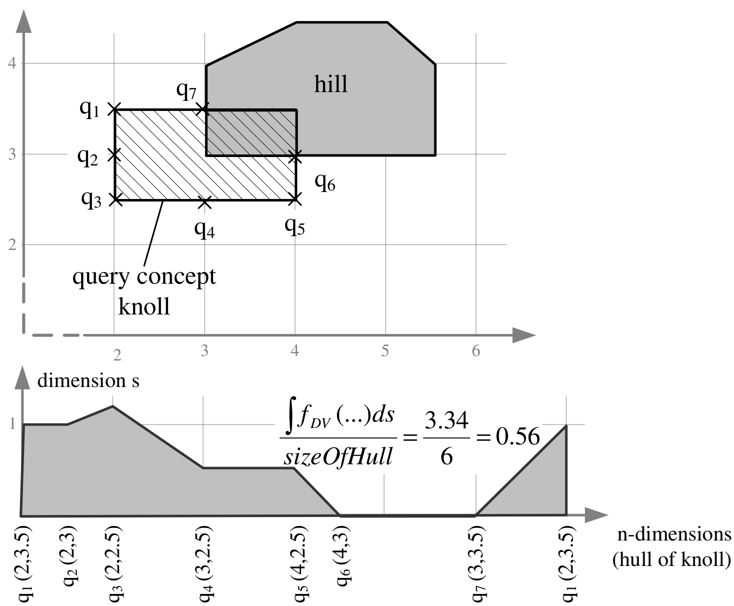


Fig. 11. Overlapping concepts and their semantic distance function

To estimate the difference between meeting, overlapping, covering concepts, and concepts being inside other concepts, an additional measure computing the degree of overlap is required to refine the similarity values. The ratio of the overlapping and non-overlapping parts of the region is a good indicator for the degree of semantic overlap and therefore also for the similarity. The greater the overlap and the less the non-overlapping parts, the higher is the similarity between both concepts. A brute-force algorithm can be used for computing overlap of convex hulls: the plain SWEEP algorithm is also applicable in the 3-dimensional case (for detailed explanation see [31, 32]). Such computation is important, but outside the scope of this work.

5.2 Concepts Being Inside or Covered by Other Concepts

Figure 12 shows concepts inside or covering other concepts: 'plain' lies inside 'lowland', i.e. there is a complete overlap. Semantically interpreted this means that 'plain' is a subconcept of 'lowland'. If a concept is inside another concept, the distance values from the superconcept to the subconcept are always greater zero (for covering concepts the distance is greater or equal to zero). The distance values from the subconcepts 'plain' respectively 'plateau' to their superconcept 'lowland' are zero. The overlap measure can be used for further distinction.

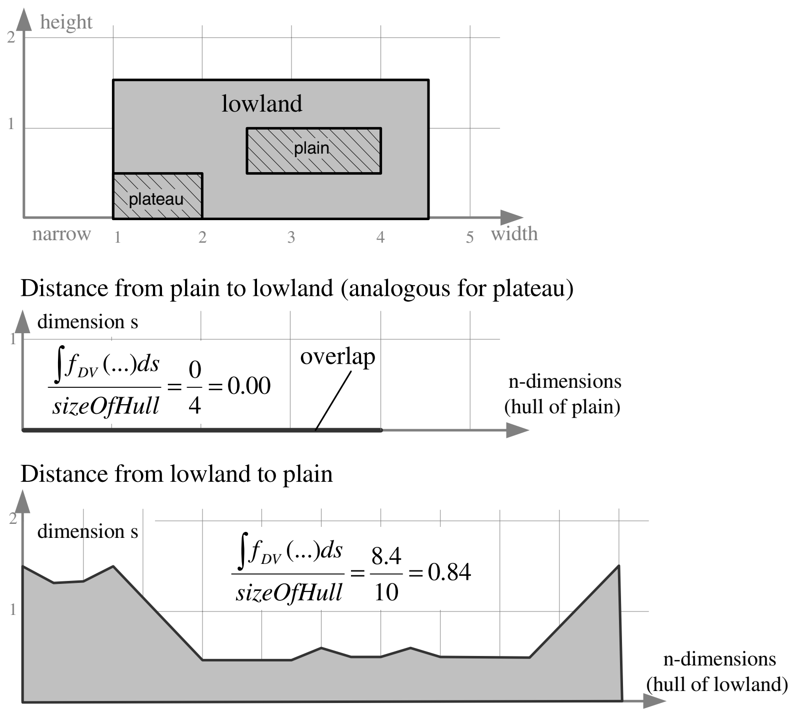


Fig. 12. Overlapping concepts with different topologic configurations

Figure 13 gives an overview of the concepts discussed above in one conceptual space. The given semantic distance values are based on the described similarity measure and do not include additional refinements of an overlap measure. These semantic distances must be transformed into a similarity value according to a similarity function, for example a linear decaying function of these semantic distances.

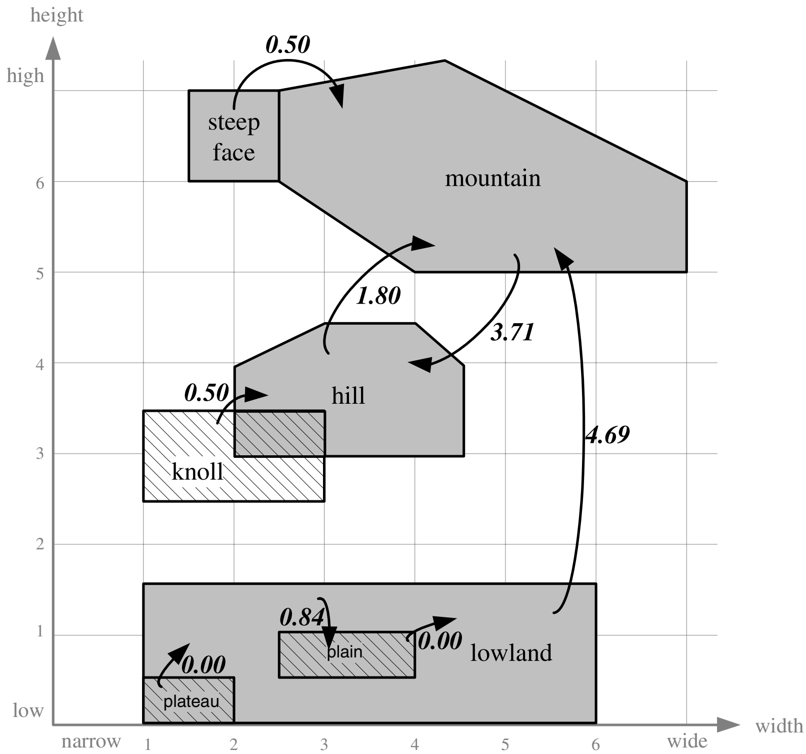


Fig. 13. Similarity values between concepts with different topologic relations in one conceptual space

6 Conclusions and Future Work

This paper develops a geometric similarity measure between concepts modelled as conceptual regions in a conceptual space. Previous approaches reduce concepts to a prototypical point and use these as input for pointwise similarity measures. The similarity between concepts is therefore reduced to the similarity of their prototypes. Such reduction of regions to single points inevitably leads to a loss of information. These measures neither account for the shape of the conceptual regions, nor for their size. The similarity measure presented in this paper includes the whole conceptual region of a query concept for similarity calculation. Shape, size and distances of a concept to another concept influence the similarity measure. Moreover, this directed similarity measure accounts for the fact that people's similarity judgments are asymmetric.

The paper leads to different directions for future research:

1. Geospatial concepts are often complex with non-obvious dimensions. We simplified the concept description in the example by representing only two dimensions. The underlying quality dimensions for a concept, its values on a dimension and the

dependencies between dimensions can be identified by human subject tests (e.g. [13]). As well multi-dimensional scaling can be used to identify potential dimensions used by humans to judge similarity.

2. Here we make the simplifying assumption that both concepts are described by the same, independent dimensions. However, many concepts are represented by different numbers of dimensions. Future research needs to investigate whether it is feasible to either leave out different dimensions and consider only common ones, or whether missing dimensions have a negative impact on the similarity of concepts. Sometimes, different dimensions can be mapped to each other (see for example the mapping of RGB to HSB colours in [5]). Dependencies between dimensions may be discovered in human subject tests—e.g. [13]—which leads to non-orthogonal axes in the representation.
3. Since the determination of vector pairs is a unidirectional process—for each vector of the hull of Q the corresponding vector of C is determined—the size and shape of Q has a great influence on the similarity function. However, vectors in C that do not belong to a vector pair have no effect. When applying the Euclidian distance measure, the part of C being far away from Q does not influence the similarity at all. Future work must investigate empirically whether it is justifiable to consider only the part of a concept C with the minimum distance. Other strategies must be investigated (see section 3.2). We propose to include the distribution function of instances of the concept in the similarity measure, e.g. consider only that part of a concept with the density of instances larger than a given marginal value.
4. People's similarity judgments are highly dependent on their tasks and the general context. Future work needs to compare the calculated similarity values with results from human subject tests using different scenarios. Differences in similarity values for different contexts could be represented in conceptual spaces by assigning weights to the quality dimensions.

Acknowledgments

We would like to thank Michael Lutz, Ulrich Streit and the 3 anonymous reviewers for their valuable comments.

References

1. Gärdenfors, P. *Conceptual Spaces: The Geometry of Thought.* Cambridge, MA: MIT Press, 317 pp., 2000.
2. Lakoff, G. Cognitive Semantics, in *Meaning and Mental Representations (Advances in Semiotics)*, U. Eco, M. Santambrogio, and P. Violi, Editors. Indiana University Press: Bloomington, Indianapolis, USA., pp. 119-154, 1988.
3. Barsalou, L. Situated simulation in the human conceptual system. *Language and Cognitive Processes*, 5(6): 513-562, 2003.
4. Gärdenfors, P. Representing actions and functional properties in conceptual spaces, in to appear: *Body, Language and Mind*, T. Ziemke and J. Zlatev, Editors. 2004.
5. Raubal, M. Formalizing Conceptual Spaces, in *Formal Ontology in Information Systems, Proceedings of the Third International Conference (FOIS 2004)*, A. Varzi and L. Vieu, Editors. IOS Press: Amsterdam, NL, pp. 153-164, 2004.

6. Devore, J. and R. Peck. Statistics - The Exploration and Analysis of Data. *4th ed. Pacific Grove*, CA: Duxbury, pp. 713, 2001.
7. Jones, W.P. and G.W. Furnas. Pictures of Relevance: A Geometric Analysis of Similarity Measures. *Journal of American Society for Information Science,* 38(6): 420-442, 1987.
8. Suppes, P., et al. Foundations of Measurement - *Geometrical, Threshold, and Probabilistic Representations*. Vol. 2, San Diego, California, USA: Academic Press, Inc. 493 pp., 1989.
9. Attneave, F. Dimensions of Similarity. *American Journal of Psychology*, 63: 516-556, 1950.
10. Melara, R.D., L.E. Marks, and K.E. Lesko. Optional processes in similarity judgments. *Perception & Psychophysics*, 51(2): 123-133, 1992.
11. Johannesson, M. Combining Integral and Separable Subspaces. in *Twenty-Third Annual Conference of the Cognitive Science Society*. Lawrence Erlbaum, 2001.
12. Johannesson, M. The Problem of Combining Integral and Separable Dimensions. Lund University Cognitive Studies: Lund, Sweden, 16 pp., 2002.
13. Johannesson, M. Geometric Models of Similarity. Lund University Cognitive Studies. Vol. 90. Lund, Sweden: Lund University, 171 pp., 2002.
14. Pike, W. and M. Gahegan. Constructing Semantically Scalable Cognitve Spaces. in *Conference on Spatial Information Theory (COSIT03).* Kartause Ittingen, Switzerland: Springer-Verlag, Berlin, 2003.
15. Tversky, A. and S. Sattath, Additive Similaritiy Trees. *Psychometrika*, 42(3): 319-345, 1977.
16. Rodríguez, A. and M. Egenhofer, Determining Semantic Similarity Among Entity Classes from Different Ontologies. *IEEE Transactions on Knowledge and Data Engineering*, 2003. 15(2): 442-456, 2003.
17. Rodríguez, A. and M.J. Egenhofer, Comparing Geospatial Entity Classes: An Asymmetric and Context-Dependent Similarity Measure. *International Journal of Geographical Information Science*, 18(3): 229-256, 2004.
18. Rada, R., et al., Development and application of a metric on semantic nets. *IEEE Transactions on systems, man, and cybernetics*, 19(1): 17-30, 1989.
19. Hahn, U., N. Chater, and L.B. Richardson, Similarity as transformation. *Cognition*, 87: 1-32, 2003.
20. Rosch, E., Principles of Categorization, in *Cognition and Categorization,* E. Rosch and B. Lloyd, Editors. Lawrence Erlbaum Associates: Hillsdale, New Jersey, pp. 27-48, 1978.
21. Goldstone, R.L. and A. Kersten, Concepts and Categorization, in *Comprehensive Handbook of Psychology*, A.F. Healy and R.W. Proctor, Editors. Wiley: New Jersey, pp. 599-621, 2003.
22. Tversky, A., Features of Similarity. *Psychological Review*, 84(4): 327-352, 1977.
23. Tversky, A. and I. Gati, Studies of Similarity, in *Cognition and Categorization*, E. Rosch and B. Lloyd, Editors. 1978, Lawrence Erlbaum: Hillsdale, NJ, pp. 79-98, 1978.
24. Tversky, A. and I. Gati, Similarity, Separability, and the Triangle Inequality. *Psychological Review*, 89(2): 123-154, 1982.
25. Johannesson, M., Modelling Asymmetric Similarity with Prominence. *British Journal of Mathematical and Statistical Psychology*, 53: 121-139, 2000.
26. Krumhansl, C.L., Concerning the Applicability of Geometric Models to Similarity Data: The Interrelationship Between Similarity and Spatial Density. *Psychological Review*, 85(5): 445-463, 1978.
27. Rodríguez, A., M. Egenhofer, and R.D. Rugg, Assessing Semantic Similarities Among Geospatial Feature Class Definitions, in *Interoperating Geographic Information Systems, Interop '99*, A. Vckovski, K. Brassel, and H.-J. Schek, Editors. Springer: Berlin-Heidelberg-New York. pp. 189-202, 1999.

28. Rodríguez, A., Assessing Semantic Similarity Among Spatial Entity Classes, in Spatial Information Science and Engineering. PhD Thesis. University of Maine: Maine, 168 pp., 2000.
29. Egenhofer, M. and R. Franzosa, Point-set topological spatial relations. *International Journal of Geographical Information Systems*, 5(2): 161-174, 1991.
30. Egenhofer, M.J., A Formal Definition of Binary Topological Relationships, in *Third International Conference on Foundations of Data Organization and Algorithms (FODO)*, W. Litwin and H. Schek, Editors. Springer: Paris, France, pp. 457-472, 1998.
31. Goodman, J.E. and J. O'Rourke, Handbook of Discrete and Computational Geometry. *1st Ed. ed. Discrete Mathematics and Its Applications*. Boca Raton: CRC Press, 1997.
32. de Berg, M., et al., Computational Geometry - *Algorithms and Applications*. Vol. 1. Berling, Germany: Springer, 365 pp., 1977.

Using Semantic Similarity Metrics to Uncover Category and Land Cover Change

Ola Ahlqvist

Department of Geography,
The Ohio State University,
1036 Derby Hall, 154 North Oval Mall,
Columbus, OH 43210-1361, USA
ahlqvist.1@osu.edu

Abstract. Analysis of geographic data that uses a nominal measurement framework is problematic since it limits the possible analytic methods that can be applied. Land cover change analysis is an example of this where both the actual change analysis as well as classification changes over time can be problematic. This study illustrates the use of semantic similarity metrics on parameterized category definitions, and how these metrics can be used to assess land cover change over time as a degree of perceived change with respect to the original landscape state. It also illustrates how changes of the categories, the classification system, over time can be analyzed using semantic similarity measures.

1 Introduction

Recent research on geographic data ontologies has illustrated that semantic inconsistencies and uncertainty need to be dealt with in a controlled way (c.f. Fonseca et al., 2002, Visser et al., 2002). Comber et al., (2004) specifically noted that the use of land cover classifications in landscape analysis is severely hampered by semantic uncertainty. Another challenge with land cover data is that the classes often are a of mix land use and land cover. This paper demonstrates an approach to account for semantic differences in a post classification land use land cover change analysis. The suggested representation is introduced together with a U.S. land use and land cover dataset used for the demonstration. The methodology to quantify and assess semantic change is then described. Changes in both the land cover nomenclatures and the actual landscape change is illustrated through a worked experiment.

2 Theoretic Background and Data Description

2.1 Uncertain Conceptual Spaces and Semantic Similarity

Over the last fifteen years or so there have been several examples of suggested approaches to model and represent concepts in a geographic information analysis context (Nyerges 1991, Livingstone and Raper 1994, Usery 1996, Bishr 1998, Mennis 2003). In this experiment I will use the idea of *Uncertain Conceptual Spaces*

M.A. Rodríguez et al. (Eds.): GeoS 2005, LNCS 3799, pp. 107 – 119, 2005.

Ahlqvist (2004) based on the cognitive theory of conceptual spaces (Gärdenfors, 2000). In essence, an uncertain conceptual space is a multidimensional attribute space made up of a collection of defining attribute *domains*, such as vegetation cover, temperature, shape, or location. A *property* is then defined as a point or fuzzy region in a low dimensional subspace, for example the interval of tree cover percentage values that help identify a forest from a non-forest. Moreover, for any concept definition, each property of that concept is assigned a certain importance, or *salience*, in relation to other properties of the concept. This enables us to declare some properties as more important than others for defining a concept.

The uncertain conceptual space is formally defined as a collection, or set, of property definitions. Each property definition is represented as a set of values from a certain domain, for example the interval of tree crown cover values. To represent the semantic uncertainty we often find in concept definitions such as "forest", the uncertain conceptual space implement rough fuzzy set constructs (Dubois and Prade, 1990). Fuzzy and rough set theories were generalized by Dubois and Prade (1990) into rough fuzzy sets, a joint representation for vague and resolution limited information. Ahlqvist et al (2003) recently demonstrated a geographic application of rough and fuzzy data integration. The chosen representation thus combines an explicit representation of two important types of semantic uncertainty; *indiscernibility* related to lack of information granularity and *vagueness* related to vague definitions of information categories.

There is also a wealth of research working towards assessing and measuring semantic similarity of categories, for some examples see Hahn and Chater (1997) and Jones et al. (2003). Two common approaches to estimate similarity use either the proportion of shared features (Tversky, 1977) or the psychological distance between related properties (e.g. Nosofsky, 1986). A number of formalizations of these and other similitude measures can be found in Bouchon-Meunier et al. (1996).

The way semantic similarity is derived is not of primary importance in this work since any quantitative measure of semantic similarity could be used for the change assessment. Thus, without going into the details of the formal calculations of the similarity metrics, Fig. 1 below illustrates the general idea of how the used representation and similarity assessment works. Again, a detailed account of this is available in Ahlqvist (2004).

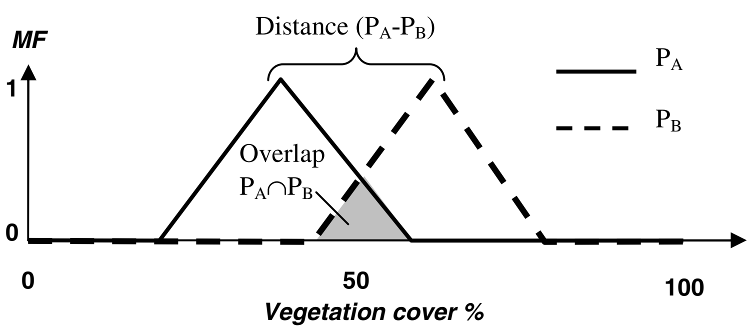


Fig. 1. Illustration of the general idea of attribute property definitions as fuzzy membership functions and the measurement of semantic distance and overlap between two properties

In the context of land use and land cover change assessment, many land use land cover classes are defined based on the amount of vegetation cover. In Fig. 1 an attribute domain has been defined as percentage values of vegetation cover. The two functions P_A and P_B outlined are attribute values or *properties* of two land cover class definitions, class A and B. The functions should be understood as fuzzy membership functions that indicate membership values to a class as a function of vegetation cover percentage. There may be several attribute domains but this example will be limited to one defining domain. Any area with a vegetation cover between 25 and 55% will have some amount of membership in class A with the highest membership around 40%. Thus, areas with 40% vegetation cover would be good examples of this class. Class B is defined to be areas with 45 to 75% vegetation cover with the highest membership around 60%. In this way vagueness of categories such as "Low intensity residential" can be captured by the representation in the form of graded membership.

To calculate the semantic similarity metric we follow the distance based approach (Nosofsky, 1986) and employ a dissimilarity measure using a Euclidean distance metric. The difference P_A-P_B is formally estimated using the fuzzy *dissemblance* index (Kaufman and Gupta, 1985) that calculates the distance between two membership functions. If several domains are involved in the class definitions a weighted sum of the distances is calculated.

The semantic *overlap* metric measures the overlap or intersection of the fuzzy functions. Again, if several domains are involved in the class definitions a weighted sum of the overlaps is calculated and this corresponds to a weighted measure of satisfiability (Bouchon-Meunier et al., 1996) following the shared feature approach (Tversky, 1977).

Both semantic similarity metrics are formally defined to apply a certain *perspective* using one of the classes as a reference. This is motivated by the idea of context effects such as the concept asymmetry we can find in statements such as "a hospital is more similar to a building than a building is to a hospital" (Rodriguez and Egenhofer 2004). Both similarity metrics are also defined to take on values in the range [0,1] where a large distance or a large overlap would have values close to 1, and a small distance/overlap would have values close to 0.

2.2 Data Description and Land Cover Category Parameterization

National Land Cover Data (NLCD) from 1992 and 2001 for Chester County, PA, south-west of Philadelphia on the U.S east coast was downloaded from the USGS seamless server (http://seamless.usgs.gov/) and processed in ArcGIS v.9. The metadata was of special importance as a source for the parameterization of land cover classes. The accompanying metadata includes brief descriptions of each land cover class in the datasets.

The NLCD classification contains 21 (1992), and 16[1,2] (2001) different land cover classes with a spatial pixel resolution of 30 meters. The NLCD was produced as a cooperative effort between the U.S. Geological Survey (USGS) and the U.S. Environmental Protection Agency (US EPA) to produce a consistent, land cover data layer for the conterminous U.S. using early 1990s Landsat thematic mapper (TM) data.

Table 1. Land use land cover classes used in the 1992 and 2001 data sets

1992 Land use land cover classes	2001 Land use land cover classes
11 Open Water	11. Open Water
12 Perennial Ice/Snow	12. Perennial Ice/Snow
21 Low Intensity Residential	21. Developed, Open Space
22 High Intensity Residential	22. Developed, Low Intensity
23 Commercial/Industrial/Transportation	23. Developed, Medium Intensity
	24. Developed, High Intensity
31 Bare Rock/Sand/Clay	31. Barren Land (Rock/Sand/Clay)
32 Quarries/Strip Mines/Gravel Pits	
33 Transitional	
41 Deciduous Forest	41. Deciduous Forest
42 Evergreen Forest	42. Evergreen Forest
43 Mixed Forest	43. Mixed Forest
51 Shrubland	52. Shrub/Scrub [1]
61 Orchards/Vineyards/Other	
71 Grasslands/Herbaceous	71. Grassland/Herbaceous[1]
81 Pasture/Hay	81. Pasture/Hay
82 Row Crops	82. Cultivated Crops
83 Small Grains	
84 Fallow	
85 Urban/Recreational Grasses	
91 Woody Wetlands	90. Woody Wetlands[2]
92 Emergent Herbaceous Wetlands	95. Emergent Herbaceous Wetlands[2]

Table 1 demonstrates that the land cover categories used in the national datasets have undergone some changes, both in terms of naming of similar classes as well as removing and adding classes from the 1992 mapping to the 2001 mapping. Obviously this will create problems for anyone wanting to compare these two datasets to look for changes in the landscape over time.

Further examination of the detailed descriptions in the metadata of some similar looking classes reveal quite significant differences in the way they were defined in 1992 and in 2001.

National Land Cover Data (NLCD) 1992

21. Low Intensity Residential - Includes areas with a mixture of constructed materials and vegetation. Constructed materials account for 30-80 percent of the cover. Vegetation may account for 20 to 70 percent of the cover. These areas most commonly include single-family housing units. Population densities will be lower than in high intensity residential areas.

---

[1] Additional shrub and grassland classes are used in Alaska.
[2] Additional wetland classes are used in coastal areas.

National Land Cover Data (NLCD) 2001

22. Developed, Low Intensity - Includes areas with a mixture of constructed materials and vegetation. Impervious surfaces account for 20-49 percent of total cover. These areas most commonly include single-family housing units.

A requirement for a semantic similarity assessment in the Uncertain Conceptual Spaces approach is that a detailed, formal description of each category in the classification systems can be provided. To help with the formal definition of the land cover classes in this experiment I used attribute domains suggested by the parameterized Land Cover Classification System (LCCS), developed by the Food and Agriculture Organization (FAO) of the United Nations (Di Gregorio and Jansen, 1998). A similar, fully worked example can be found in Ahlqvist (2005).

3 Experiment

The following experimental description will first give a few examples of the parameterization of the 1992 and 2001 land cover categories. I will then describe the two different types of analysis matrices, the change matrix and the semantic similarity matrix, and how they are used together to generate maps of semantic landscape change.

3.1 Parameterization

The parameterization was performed manually by evaluating the text description of each class and its relation to other classes. The understanding gained of the classes is then described in terms of a unified set of attribute domains. The two examples below (Table 2 and 3) illustrate the parameterization of the same two land cover categories as in the previous example from the 1992 and 2001 classification respectively. They demonstrate the used attribute domains, how property values can be assigned as intervals over the domain, and the assignment of different salience as detailed below.

In the Table 2 and 3 examples bold face values indicate attribute values with membership 1 to the class, for example the 1992 class 21 "Low intensity residential" (Table 2) is defined to be areas that have 0-20 % water cover (membership=1 within that interval), if water is present it can be of any form (ice or water), it should be covered by 30-80% impervious materials, and so on. For the analysis, each class is defined in a text file forming a collection of concept description text files that are provided as input to the similarity assessment algorithm.

An important feature of the attribute parameterization is the possibility to use fuzzy and rough set definitions. This is not apparent in the examples here mostly because the class definitions in the documentation are "artificially crisp", i.e. they use very exact class limits in the definition text although most people would recognize their graded character. Another feature of the chosen parameterization framework is the assignment of "salience" to attributes. Each attribute typically has varying importance to the definition. For residential categories attributes **Impervious surface, Vegetation Cover, Development, Surface Type** are important characteristics mentioned either ex

Table 2. Example parameterization of the NLCD 1992 class 21 Low Intensity Residential. Bold face indicating attribute values with membership value 1 to the class, plain style values indicate 0 membership in the class.

Domain	Scale	Range
Water Cover	Ratio	[**0 10 20** 30 40 50 60 70 80 90 100]
Water Phase	Nominal	{**Ice, Water**}
Impervious surface	Ratio	[0 10 20 **30 40 50 60 70 80** 90 100]
Vegetation Cover	Ratio	[0 10 **20 30 40 50 60 70** 80 90 100]
Development	Nominal	{**Residental**, Commercial, Mining}
Surface Type	Nominal	{Earthen material, **Constructed**}
Tree Cover	Ratio	[**0 10 20 30 40 50 60 70** 80 90 100]
Tree Height	Ratio	[**0 10 20 30 40 50 60 70 80 90 100**]
Deciduous Pct.	Ratio	[**0 10 20 30 40 50 60 70 80 90 100**]
Evergreen Pct.	Ratio	[**0 10 20 30 40 50 60 70 80 90 100**]
Shrub Cover Pct.	Ratio	[**0 10 20 30 40 50 60 70** 80 90 100]
Woody Tenure	Nominal	{**(Semi)Natural, Cultivated/Planted**}
Grass/Herb Tenure	Nominal	{**(Semi)Natural, Cultivated/Planted**}
Grass/Herb Cover Pct.	Ratio	[**0 10 20 30 40 50 60 70** 80 90 100]
Crop	Nominal	{**RowCrops, SmallGrains, Fallow, Hay, Grass**}
Water Persistence	Nominal	{Permanent, Periodically, Waterlogged}

Table 3. Example parameterization of the NLCD 2001 class 22 Developed, Low Intensity. Bold face indicating possible values given membership value 1, other values given membership 0.

Domain	Scale	Range
Water Cover	Ratio	[**0 10 20** 30 40 50 60 70 80 90 100]
Water Phase	Nominal	{**Ice, Water**}
Impervious surface	Ratio	[0 10 **20 30 40 50** 60 70 80 90 100]
Vegetation Cover	Ratio	[0 10 20 30 40 **50 60 70 80** 90 100]
Development	Nominal	{**Residental, Commercial**, Mining}
Surface Type	Nominal	{Earthen material, **Constructed**}
Tree Cover	Ratio	[**0 10 20 30 40 50 60 70 80** 90 100]
Tree Height	Ratio	[**0 10 20 30 40 50 60 70 80 90 100**]
Deciduous Pct.	Ratio	[**0 10 20 30 40 50 60 70 80 90 100**]
Evergreen Pct.	Ratio	[**0 10 20 30 40 50 60 70 80 90 100**]
Shrub Cover Pct.	Ratio	[**0 10 20 30 40 50 60 70 80** 90 100]
Woody Tenure	Nominal	{(Semi)Natural, **Cultivated/Planted**}
Grass/Herb Tenure	Nominal	{(Semi)Natural, **Cultivated/Planted**}
Grass/Herb Cover Pct.	Ratio	[**0 10 20 30 40 50 60 70 80** 90 100]
Crop	Nominal	{RowCrops, SmallGrains, Fallow, Hay, **Grass**}
Water Persistence	Nominal	{Permanent, Periodically, Waterlogged}

plicitly or implicitly by the textual class description and all the rest are less so. Thesalience can be continuously graded on a scale from 0 to 1 but in this experiment I have only use binary values 0 and 1 to switch importance on an off, indicated as grey shadowed rows (1) or white rows (0) in Tables 2 and 3. The use of either 0 or 1 as salience values is admittedly an arbitrary one and these values could be more fine-tuned to match the user's or producer's understanding of each class. Methods to elicit graded weights of importance can be found in Saaty (1990).

Table 4. Land use land cover change matrix from 1992 to 2001 for Chester county, PA. For detailed information on class codes, see Table 1.

Change	**2001**													
1992	**11**	**21**	**22**	**23**	**24**	**31**	**41**	**42**	**43**	**81**	**82**	**90**	**95**	**Total**
11	5464	378	428	313	151	163	1593	60		1627	776	1186	735	**5464**
21	68	61684	24922	10513	2400	2989	13041	1684		17408	6183	1819	476	**61684**
22	6	1078	1817	1608	538	36	180	21		433	155	41	28	**1817**
23	392	5794	7366	9696	4793	326	1962	60		2813	2438	614	316	**9696**
32	71	253	664	1473	768	289	113			117	487	105	111	**1473**
33	51	562	1391	1370	446	340	619	12		14474	2557	121	40	**14474**
41	935	21768	14767	4360	703	10050	438228	8514	21	137162	83631	21395	3343	**438228**
42	357	4012	2365	540	72	1740	9329	4614	5	29612	6558	1265	479	**29612**
43	91	10783	4219	679	96	2247	24541	3082	6	45861	12266	1530	416	**45861**
81	558	31724	23188	6341	988	3694	96742	3741	12	519116	187549	7184	2530	**519116**
82	100	5676	6556	3085	831	704	4160	66		60602	65693	587	653	**65693**
85		3290	144	36	2	2	93	4		404	63	11	3	**3290**
91	13	37	17	5		64	2297	8		969	651	2133	24	**2297**
92	223	243	353	193	38	77	1352	66		1617	1235	585	213	**1617**
Total	**5464**	**61684**	**24922**	**10513**	**4793**	**10050**	**438228**	**8514**	**21**	**519116**	**187549**	**21395**	**3343**	**519116**

3.2 Change Analysis–The Change Matrix

A straightforward post-classification change analysis was produced by overlaying the 1992 and 2001 land cover datasets to produce a change image where each pixel holds information on what class it had in 1992 and in 2001. From the change image a contingency matrix can be extracted where counts of pixels for each combination of land cover class change is recorded. In a contingency or change matrix for identical classification systems from two times the major diagonal would hold the number of pixels that have not changed over time and any off diagonal entry would indicate the amount of change from the row category to the column category.

However, in this experiment, and in many applied situations, differences in class definition and other class changes makes the role of the major diagonal less clear and it is often hard to tell what a change from say "85 Urban/Recreational Grasses" to "21 Developed, Open Space" mean. They could be the same but still get assigned to these different categories that even appear in different branches of the two-tiered classification system, because of classification system changes from 1992 to 2001. Even for closely matching categories there are slight changes in definition that makes a straightforward analysis complicated. So to do a change analysis from this overlay requires substantial post analysis work.

3.3 Change Analysis–The Semantic Similarity Matrix

The parameterized class definitions from the previous section were used to produce a cross product of pair wise semantic similarity calculations for all land cover classes in the change analysis. Two different similitude matrices were produced; one that holds the *dissemblance*, between categories (Table 5), and one that holds the *overlap* between categories (Table 6).

The dissemblance and overlap matrices can in themselves be used to analyze class relationships between the two classification systems. To find the most similar catego-

Table 5. Dissemblance matrix holding pair wise estimates of semantic distance between land cover classes in the 1992 and 2001 land cover data sets

Dissemblance	2001												
1992	**11**	**21**	**22**	**23**	**24**	**31**	**41**	**42**	**43**	**81**	**82**	**90**	**95**
11	0	0.77	0.77	0.77	0.77	0.77	0.77	0.77	0.77	0.77	0.77	0.16	0.16
21	0.75	0.57	0.12	0.51	0.55	0.76	0.74	0.74	0.74	0.74	0.74	0.74	0.74
22	0.79	0.71	0.35	0.52	0.5	0.79	0.82	0.82	0.82	0.84	0.84	0.82	0.82
23	0.75	0.60	0.53	0.51	0.53	0.75	0.75	0.75	0.75	0.76	0.76	0.75	0.75
32	0.81	0.92	0.86	0.82	0.82	0.03	0.85	0.85	0.85	0.88	0.88	0.88	0.88
33	0.81	0.41	0.29	0.11	0.01	0.82	0.85	0.85	0.85	0.66	0.66	0.88	0.88
41	0.54	0.50	0.51	0.53	0.55	0.55	0.02	0.44	0.23	0.53	0.53	0.23	0.29
42	0.54	0.50	0.51	0.53	0.55	0.55	0.44	0.02	0.23	0.53	0.53	0.23	0.29
43	0.50	0.46	0.47	0.48	0.51	0.51	0.23	0.23	0.02	0.49	0.49	0.11	0.2
81	0.70	0.52	0.51	0.52	0.55	0.71	0.70	0.70	0.70	0.09	0.46	0.74	0.64
82	0.70	0.52	0.51	0.52	0.55	0.71	0.70	0.70	0.70	0.46	0.46	0.74	0.64
85	0.76	0.24	0.47	0.24	0.29	0.77	0.76	0.76	0.76	0.58	0.59	0.79	0.71
91	0.64	0.72	0.72	0.74	0.77	0.77	0.59	0.59	0.59	0.74	0.74	0.05	0.22
92	0.57	0.69	0.68	0.68	0.69	0.69	0.58	0.58	0.58	0.66	0.66	0.30	0.07

Table 6. Overlap matrix holding pair wise estimates of semantic overlap between land cover classes in the 1992 and 2001 land cover data sets

Overlap	2001												
1992	**11**	**21**	**22**	**23**	**24**	**31**	**41**	**42**	**43**	**81**	**82**	**90**	**95**
11	1	0.71	0.71	0.71	0.71	0.71	0.71	0.71	0.71	0.71	0.71	0.85	0.85
21	0.04	0.71	0.80	0.83	0.71	0	0.50	0.50	0.50	0.30	0.30	0.31	0.31
22	0.50	0.71	0.71	0.74	0.96	0.33	0.17	0.17	0.17	0.01	0.01	0.01	0.01
23	0.20	0.71	0.56	0.78	0.75	0.13	0.37	0.37	0.37	0.18	0.18	0.19	0.19
32	0.58	0	0	0.19	0.48	0.90	0.19	0.19	0.19	0.01	0.01	0.01	0.01
33	1	0.82	0.82	0.84	0.95	0.90	0.19	0.19	0.19	0.82	0.82	0.01	0.01
41	0.90	0.89	0.84	0.79	0.78	0.89	1	0.78	0.89	0.91	0.91	0.95	0.90
42	0.90	0.89	0.84	0.79	0.78	0.89	0.78	1	0.89	0.91	0.91	0.95	0.90
43	0.90	0.89	0.84	0.79	0.78	0.89	0.80	0.80	1	0.91	0.91	0.95	0.90
81	0.78	0.92	0.81	0.81	0.73	0.62	0.65	0.65	0.65	1	0.89	0.45	0.76
82	0.78	0.90	0.78	0.78	0.69	0.62	0.65	0.65	0.65	0.89	1	0.45	0.76
85	0.71	1	0.85	0.91	0.86	0.56	0.59	0.59	0.59	0.91	0.82	0.41	0.70
91	0.64	0.73	0.60	0.49	0.44	0.60	0.83	0.83	0.83	0.63	0.63	0.92	0.83
92	0.80	0.77	0.72	0.68	0.60	0.66	0.78	0.78	0.78	0.73	0.73	0.84	0.95

ries we would look for combinations of a low dissemblance and a high overlap. One such example would be the “Deciduous forest” categories where dissemblance is 0.025 and overlap is 1.

3.4 Change Analysis – The Semantic Similarity Maps

In the dissemblance matrix we find for each 1992 category its semantic distance to all 2001 categories. We see e.g. that the 1992 class 21 “Low Intensity Residential” has a small distance (0.124) to the 2001 class 22 “Developed, Low Intensity”, but larger distance to most other classes, and that indicates the closest match. From a land cover change perspective, any pixels that changed from class 21 in 1992 to class 22 in 2001 should be regarded as a minor or no change. Thus, the values from each cell in the dissemblance matrix were used to replace the class change information in the original change image to generate a semantic distance change image (Fig. 2, left).

Cell values from the overlap matrix were also used to generate a change image in terms of overlap (Fig. 2, right). The images in Fig. 2 illustrate the degree of land cover change from a semantic similarity perspective. Each cell shows how “semantically different” the state in 2001 is compared to the original state in 1992. Dark areas signify the biggest difference and white areas signify the smallest difference. The very dark pixels that show up as “shadows” on the edges of shaded areas are mostly related to image registration errors, but they appear as ‘big’ changes.

Generally, the developed areas in the center of the images show larger differences compared to the surrounding landscape in the image peripheries that consists mostly

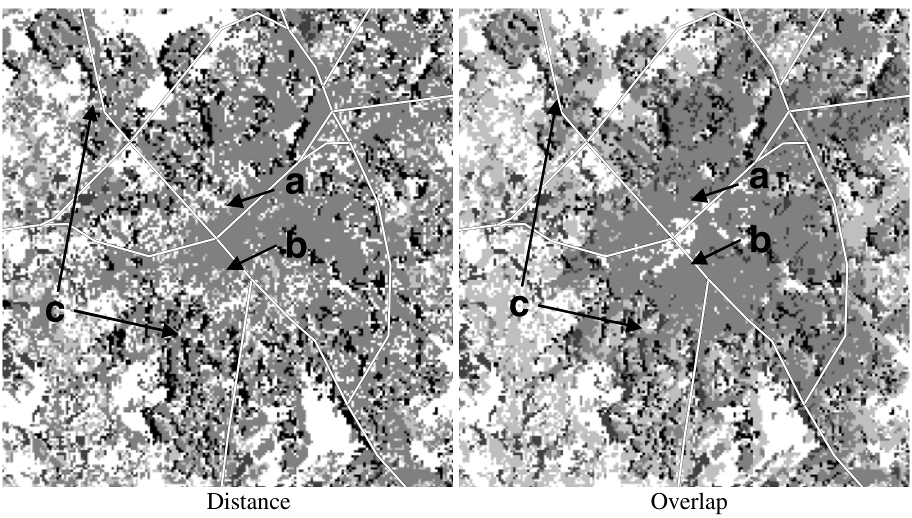


Fig. 2. Land cover change in eastern Chester County, PA measured as semantic distance and overlap between land cover classes. Dark shades indicate large difference/little overlap and lighter shades small difference/large overlap between the classes recorded for 1992 and 2001. The downtown district (a) and areas indicated by (b) and (c) are discussed in the text.

of agricultural or forested areas. This is most likely due to the changes in definition of the "developed/residental" categories from the 1992 to the 2001 classification. For example the 2001 "Developed" categories also include commercial development that had its own category in 1992. Also some changes to the limits of impervious surface and vegetation cover percentages have changed from 1992 to 2001 (see examples in Table 2 and 3). To that end the semantic change analysis does not pick up an actual change in the landscape, but identifies the change in definition between the past and present class used to describe the same land cover.

However, the two measures address different aspects of semantic similarity between two classes and they can be used together to refine the change assessment based on four general class relationship cases (figure 3).

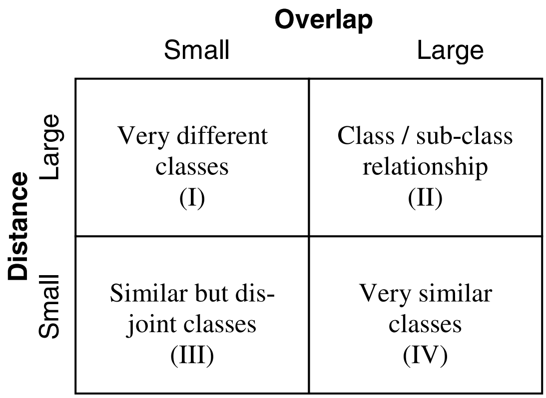


Fig. 3. Four general semantic class relationship cases based on combination of small/large semantic distance and overlap values respectively

A low overlap together with a large distance would indicate two very different classes (Fig.3, I). A high overlap together with a large distance would indicate that one class is a sub-class of the other (II), a low overlap with a small distance would indicate similar but disjoint classes (III). The classes "Developed, low intensity" and "Developed, medium intensity" would be an example of this case. And finally a high overlap together with a small distance would indicate two very similar classes (Fig. 3, IV). The 1992 and 2001 forest categories are very similar and have similitude metrics accordingly as we can see from tables 5 and 6. An analysis of change would then probably not see case (II) as very disturbing, rather something that is probably related to definitional changes, whereas case (I and III) would be more important events.

The three areas indicated by a, b and c in Fig. 1 roughly identify three of these different types of change. Area (a) is the city center and has a fairly large dissemblance value (~0.5) and overlap is very high (~0.9). The difference in this central area then is probably not so much a change in cover as it is in changed definition. Many of the pixels in this area have changed from "High intensity residential" in 1992 to "Developed, High intensity" in 2001. Obviously similar categories but the high overlap together with moderate dissemblance indicates that the old definition can be viewed as a less general subclass of the new category mostly because of the inclusion of commercial areas in the 2001 "Developed" category.

Surrounding the inner city district (Fig. 2, b) is a region with smaller dissemblance values (0.1-0.5) and smaller overlaps (0.7-0.8). These are mostly areas that have changed from a low intensity residential to developed areas of medium to high intensity. This would correspond to the general case where classes are very similar in their definition but still differ slightly, corresponding roughly with case III in Fig. 3.

Some areas where bigger changes have occurred are indicated by (c) in Fig. 2. These are places where new developments, mostly residential, have emerged on previous farmland or forest land. In these areas the distance metric is large and the overlap metric is small which together indicate a big shift in the perceived landscape, corresponding to the general case of two very different classes (Fig. 3, I).

4 Discussion

In summary, the suggested method serves two major purposes in the context of land use land cover change assessment. First, it enables users to follow graded changes in the landscape based on quantitative similarity metrics. Second, it enables users to evaluate category changes over time, such as to what degree an old land cover class is similar to currently used land cover classes. It is unclear to what degree this method can separate actual change from changes in category definition without user input. In the experiment some reasoning and overall knowledge of the area was required to interpret the results. However, this should be compared to the current situation where changes in land use land cover taxonomies make change assessments probably as user intensive if not more, and with no explicit recognition of the semantic differences.

Although this experiment did not use the full possibility of defining fuzzy membership intervals, this feature is probably one of the most important of the suggested approach to represent categories. Nevertheless, the representation was used to evaluate graded overlaps and distances between categories and thus enabled the similarity assessment to identify both slight and major differences between categories/landscape changes. The use of fuzzy and rough set constructs to define nominal categories is in a sense an implementation of the ideas to provide a possibility to compute with words put forward by Openshaw (1996). The similarity between categories, similar to most quantitative assessments, is calculated more precisely if the category attributes are defined over ordinal or ratio or scalar domains. Nevertheless, the use of rough fuzzy set constructs in the category representation makes it possible to include nominal attributes in the definition as well and calculate upper and lower approximations of the similarity measures (Ahlqvist, 2004).

In the experiment we only saw dissemblance and overlap values for pairs of categories from different taxonomies but it is equally possible to produce a matrix that holds all possible combinations including those within classification systems. These fully extended semantic similarity matrices can then be used to explore category relationships within a classification system. Such analyses can reveal how well classes are separated or if there may be a risk of confusion between some classes, a situation the may be problematic from a data accuracy perspective (Ahlqvist and Gahegan, in press). The complete matrices will also have the extra feature of containing similarity estimates in "both directions" for any pair of categories, for example how much class A overlap with B as well as how much B overlaps with A. This addresses the concept asymmetry idea introduced briefly in the theoretical background. For example, the residential classes detailed in **Table 2** and **Table 3** had a high salience for the attributes that deal with vegetation cover, amount of impervious surfaces and type of development. A forest class on the other hand would have high salience values for tree cover, tree height and type of vegetation attributes. Comparing a residential class with a forest class would look at the forest class from the perspective of the residential class by using salience weights corresponding to the residential class. It is therefore important to define values for all attributes that have a salience value greater than zero in any class definition included in the analysis. A separate analysis of these different "views" is likely to give additional insights into the relationship between classes.

The concept of a semantic change map obviously opens a new window onto "mind spaces", where the close connection between a geographic category definition and its spatial expression is articulated. There are a number of other potential visualizations to explore this type of semantic representation. One such example is to enhance the map visualization of the semantic space by combining the two similarity metrics in one bi-variate colored map to reveal different combinations of distance and overlap simultaneously in one image. Another example, recently demonstrated by Ahlqvist et al. (2004), is to visually explore the semantic similarity matrices using large bi-variate colored matrix displays and scatter plots.

Future research will address development of a visual interface for knowledge elicitation and concept definition. In this work we intend to extend previously developed interfaces for web based Delphi discussions to enable collaborative concept negotiation and definition. A natural extension to further this direction is to provide the class definition text files in standardized markup languages such as OWL (http://www.w3.org/2004/OWL). Additional user studies are also necessary to verify the credibility of formal category definitions.

References

1. Ahlqvist, O. A parameterized representation of uncertain conceptual spaces, *Transactions in GIS*, 8(4): 493-514, 2004.
2. Ahlqvist, O. Using uncertain conceptual spaces to translate between land cover categories, *International journal of geographical information science*, 19(7): 831-857, 2005.
3. Ahlqvist, O and Gahegan, M. *Photogrammetric Engineering and Remote Sensing*, Probing the relationship between classification error and class similarity, in press,
4. Ahlqvist, O., Keukelaar, J., and Oukbir, K. Rough and fuzzy geographical data integration, *International journal of geographic information science*, 17(3): 223-234, 2003.
5. Bishr, Y. Overcoming the semantic and other barriers to GIS interoperability, *International Journal of Geographic Information Science*, 12(4): 299-314, 1998.
6. Bouchon-Meunier, B., Rifqi, M., Bothorel, S. Towards general measures of comparison of objects, *Fuzzy Sets and Systems*, 84: 143-153, 1996.
7. Comber, A., Fisher, P., and Wadsworth, R. Assessment of a semantic statistical approach to detecting land cover change using inconsistent data sets. *Photogrammetric Engineering and Remote Sensing*, 70: 931-938., 2004.

8. Di Gregorio, A., and Jansen, L.J.M. Land Cover Classification System: Classification Concepts And User Manual, FAO, Rome, 1998.
9. Dubois, D., and Prade, H. Rough fuzzy sets and fuzzy rough sets, *Int.J.General Systems*, 17:191-209, 1990.
10. Fonseca, F.T., Egenhofer, M.J., Agouris, P., and Câmara, G. Using ontologies for integrated geographic information systems, *Transactions in GIS*, 6(3): 231-257, 2002.
11. Gärdenfors, P. *Conceptual spaces : the geometry of thought*, MIT Press, Cambridge, Mass., 307pp., 2000.
12. Hahn, U., and Chater, N. Concepts and Similarity, *Knowledge, Concepts and Categories*, (Lamberts, K., and Shanks, D., editors), East Sussex, Psychology Press, pp. 43-92, 1997.
13. Jones, C.B., Alani, H., and Tudhope, D. Geographical Terminology Servers – Closing the Semantic Divide, in: Duckham et al. (Eds.) *Foundations of geographic information science*, Taylor & Francis, London, 257 p, pp. 205-222, 2003.
14. Kaufman, A., and Gupta, M.M. *Introduction to fuzzy arithmetic*, Van Nostrand Reinhold Company, New York, 351pp., 1985.
15. Livingstone D., and Raper J. Modelling environmental systems with GIS: theoretical barriers to progress, in: Worboys, M.F. (Ed), *Innovations in GIS 1*, Taylor & Francis, London, pp. 229-240, 1994.
16. Mennis, J.L. Derivation and implementation of a semantic GIS data model informed by principles of cognition, *Computers, Environment and Urban Sytems*, 27, pp.455-479, 2003.
17. Nosofsky, R.M. Attention, similarity,and the identification-categorization relationship, *Journal of Experimental Psychology: General*, 115: 39-57, 1986.
18. Nyerges, T.L. Geographic information abstractions: conceptual clarity for geographical modeling, *Environment and Planning A*, 23: 1483-1499, 1991.
19. Openshaw, S., Commentary: Fuzzy logic as a new scientific paradigm for doing geography, Environment and Planning A, 28:761-768, 1999.
20. Power et al. Describe a hierarchical fuzzy pattern matching based on ad hoc rules for combinations of crisp landcover classes, 2001.
21. Rodríguez, M.A., and Egenhofer, M.J. Comparing geospatial entity classes: an asymmetric and context-dependent similarity measure, *International Journal of Geographical Information Science*, 18: 229-256, 2004.
22. Saaty T. L. *The analytic hierarchy process: planning, priority setting, resource allocation.* Pittsburgh, PA, RWS Publications, 1990.
23. Tversky, A. Features of similarity, *Psychological review*, 84: 327-352, 1977.
24. Usery, E.L. A conceptual framework and fuzzy set implementation for geographic features, in: in: P.A. Burrough and A.U. Frank (Eds), *Geographic, Objects, With Indeterminate Boundaries*, Taylor & Francis, London, 345 pp., pp. 71-85, 1996.
25. Visser U., Stuckenschmidt H., Schuster G., Vogele T. Ontologies for geographic information processing, *Computers and Geosciences*, 28(1): 103-117, 2002.

Measuring Arrangement Similarity Between Thematic Raster Databases Using a QuadTree-Based Approach

Denis J. Dean

Remote Sensing/GIS Program, Colorado State University,
113 Forestry Building, Fort Collins, Colorado 80523-1472
denis@cnr.colostate.edu

Abstract. Measuring the degree of similarity between thematic raster databases is a common task widely used in remote sensing accuracy assessment, spatial model validation, and many other geospatial tasks. However, conventional similarity measures look only at point-to-point similarity; they are not designed to evaluate the similarity of shapes and arrangements of features within the databases being compared. This study proposes a technique of assessing arrangement similarity based on a comparison of quadtree representations of the maps being evaluated. Empirical assessment shows that the technique produces results that agree strongly with subjective evaluations of the similarity of artificial raster databases produced by a survey of map users.

1 Introduction

Measuring the degree of similarity between two maps, or the functionally equivalent task of measuring similarity between a map and one or more ground observations, is arguably the most fundamental quantitative measurement made in the field of geospatial science. Without the ability to measure similarity, it is impossible to perform any kind of empirical map accuracy assessment (e.g., assessment of the classification accuracy of remotely sensed data, assessment of the accuracy of the outputs of predictive spatial models, etc.). This impossibility arises from the fact that any empirical accuracy assessment procedure must by definition involve measuring the degree of similarity between a map and some (presumably highly accurate) reference data. Since knowledge of map accuracy is fundamental to the analysis of any data a map may contain, it follows that similarity measures are fundamental to map analysis, and thus to geospatial science.

Conventional techniques for measuring map similarity are based on point-to-point comparisons. These techniques typically compare the value of some variable *V* recorded at point *P* on a first map to the value of *V* recorded at *P* on either a second map or as measured on the ground. In some cases, point *P* may be defined as a pixel or raster cell rather than a true geometric point, but regardless of how points are defined, comparisons of this sort are fundamentally point-to-point processes.

When point comparisons of this sort are conducted at a representative sample of locations from across a map, a reasonable estimate of overall point accuracy can be obtained. In cases where variable *V* is continuous (e.g., elevation), this estimate can be summarized using some central tendency measure (such as root mean square error)

M.A. Rodríguez et al. (Eds.): GeoS 2005, LNCS 3799, pp. 120 – 136, 2005.

of the differences between the values recorded in the two maps at each sampling point. In cases where *V* is discrete (e.g., land cover categories), the measure of similarity recorded at each sample point is binary; either the data from the two maps agrees with one another or it does not. In these cases, tabular results are produced, resulting in the error matrices widely used in accuracy assessment of remotely sensed data [1].

Map comparisons such as these can do an excellent job of describing the point-to-point similarity of two maps. However, it is easy to demonstrate that point-to-point similarity is only one aspect of overall similarity. Consider the case of two checkerboard maps whose values (represented by colors) are out of sync (Figure 1). In this case, there is no point where the two maps share identical values (e.g., the upper left corner of the left map is yellow while the upper left corner of the right map is red). Thus, a point-to-point comparison would conclude that the two maps are completely dissimilar. However, a visual examination reveals that the two maps are highly similar in that the features they contain (i.e., the individual squares that comprise the checkerboards) are identical in size and shape and are arranged in identical patterns. This implies that the full breadth of similarity cannot be captured by point-to-point similarity alone; some measure of similarity of size, shape and pattern of map features is also required. For the sake of brevity, we will term this second aspect of similarity as *arrangement similarity*.

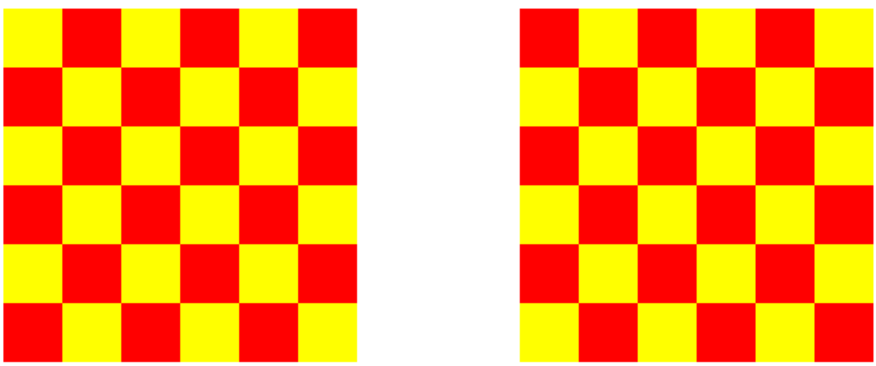


Fig. 1. Raster databases with no point-to-point similarity but obvious arrangement similarity

The primary goal of this study was to develop and validate a quantitative, objective way of measuring arrangement similarity between thematic raster databases. In the course of achieving this goal, a secondary goal of developing a way of creating pseudorandom thematic raster databases with controlled amounts of arrangement similarity was also established. How both of these goals were achieved will be described in the remaining sections of this document.

Obviously, this limited arrangement similarity metric does not address the need for equivalent measures for vector databases, or for raster databases containing continuous data. However, we felt that starting with a limited metric was appropriate, because our results might produce insights that would facilitate the development of metrics applicable to other areas. Furthermore, thematic raster databases are quite common (e.g., classified remotely sensed data, outputs of predictive spatial models, etc.), so even the limited metric developed here would have wide applicability.

2 Previous Research

Psychologists have conducted a tremendous amount of research into the subject of how humans perceive similarity between images. A great deal of this research involves color perception (e.g., [2]) and thus falls outside the topic of arrangement similarity. Other research explains the perception of similarity in two-dimensional photographs of three-dimensional scenes based on how objects in the scenes occlude one another, and how they are affected by perspective (e.g., [3]). These studies are also not applicable to maps, especially in light of the fact that [4] found that humans process information from maps in fundamentally different ways than they do information from other types of images. However, some psychological research deals with how image characteristics other than color, perspective and occlusion impact human similarity perception, and thus is directly applicable to this study ([5], [6]). Empirical studies in this area indicate that up to a point, humans are generally quite good at ignoring dissimilar stray points scattered randomly throughout two otherwise similar images (this sort of dissimilarity is often termed "static"). However, when the total amount of static exceeds some threshold (i.e., the signal to noise ratio drops to below some critical value), human ability to recognize similarity diminishes quickly. Humans also appear to be quite good at recognizing image similarity when larger features are simply broken down into multiple smaller features – i.e., humans recognize similarity when a feature is shown as a single entity in one image and as multiple smaller entities (that collectively make up the larger entity) in another image. Where humans have the most difficulty recognizing similarity is when feature boundaries are obscured by dissimilarities. Thus, humans are more likely to recognize arrangement similarity between two images featuring *X%* random static then they are to identify arrangement similarity between images with *X%* dissimilarity confined to regions near boundaries between features shown in the images.

Researchers from a variety of other fields have occasionally delved into the topic of arrangement similarity. [7] and [8] recognize the need for measuring arrangement similarity (although they do not use that term) when evaluating land use /land change models, and they present a number of techniques to address the issue. Many of these techniques are based on a cell aggregation approach, which unfortunately blurs feature boundaries by iteratively averaging cell values with values from neighboring cells. Other techniques presented by these authors involve comparisons to null models, and are only appropriate in situations where meaningful null models can be developed. [9] attempted to simultaneously evaluate point-to-point and arrangement similarity of maps showing the occurrences of various species of wildlife. They concluded that their technique provided more meaningful appraisals of map similarity than did conventional point-to-point similarity measures. Unfortunately, while their results were promising, their approach was somewhat arbitrary, in that it involved the *a priori* selection of a scale at which arrangement similarity was to be measured. This is somewhat similar to the combined point-to-point and arrangement similarity assessment technique developed by [10], which involved summarizing multiple point-to-point analyses conducted at points separated from one another by random distances (somewhat akin to semivariogram analysis). [10] used their approach to assess the level of fragmentation present in a single map, but it is not difficult to imagine a modified version of this approach that assesses similarity between multiple maps.

[11] developed a novel approach to evaluate what may be viewed as arrangement similarity that involved looking at images in a vector format. Unfortunately, they did not extend their technique to include evaluations of the similarity of entire maps; instead, they limited their technique only to the assessment of individual objects.

Geospatial scientists have occasionally attempted to quantify the degree of arrangement similarity between maps. [12] presented a combined point-to-point and arrangement similarity assessment technique that was based on a series of gestalt rules that produced intuitively appealing results, but it was never evaluated in any objective way. [13], [14] and [15] all proposed similarity assessment approaches based upon the concepts of fuzzy logic. The justifications for each of these techniques vary, but it is possible to view all of them as being ways of assessing combined point-to-point and arrangement similarity. Each approach involves changing the concept of raster cell class membership from a binary phenomenon (i.e., either cell *X* is a member of class *Y*, or it is not) to a probabilistic phenomenon (i.e., the probability of cell *X* belonging to class *Y* is *Z*). Each author develops a different technique for constructing class membership functions, and none use membership functions that incorporate any measure of arrangement similarity. However, if such membership functions were developed, these techniques could be used to assess arrangement similarity.

[16] and [17] developed and evaluated the CLC (*C*ombined *L*ocation and *C*lassification) method for accuracy assessment of multitemporal raster databases acquired via remote sensing. As its name implies, the CLC is another combined metric. It functions by iteratively evaluating images shifted in relation to one another, and thus is mostly intended to account for discrepancies in registration between otherwise identical remotely sensed images. The evaluation described by [17] shows that the CLC technique works well in this context, but the technique was never intended to address the more general problem that is the subject of this study.

3 Methods

The methodology adopted for this study was to develop and validate a technique for creating pseudo-random thematic raster databases with controlled amounts of arrangement similarity, and then to use this technique to generate databases to validate the proposed arrangement similarity metric. This section will describe how databases containing controlled amounts of arrangement similarity were created (subsection 3.1), how the proposed quadtree-based arrangement similarity metric operates (subsection 3.2) and compares to the Kappa statistic (subsection 3.3), and finally, how the arrangement similarity metric was evaluated (subsection 3.4).

3.1 Pseudorandom Raster Databases with Controlled Arrangement Similarity

The midpoint displacement method (MPDM), presented by [18], is an established technique for creating pseudorandom raster databases containing continuous data and exhibiting controlled amounts spatial autocorrelation. The MPDM requires as inputs (1) the number of rows and columns of cells to be placed in the raster databases to be created, and (2) the value of an *H* parameter, which sets the target amount of spatial autocorrelation to be present in the raster database produced by the MPDM. As output, the MPDM produces a raster database where each cell contains a pseudorandom real number. At high *H* values, cell values are highly correlated with the values in nearby cells (thus resulting in a database with high spatial autocorrelation); at low *H* values, cell values are less correlated with nearby cell values (thus producing databases with low spatial autocorrelation).

It is a trivial matter to convert MPDM outputs into raster databases containing thematic data. This can be accomplished via reclass operations as described in any GIS text (e.g., [19]). This study used *equal area reclassing* (EAR) to accomplish this task. In EAR, the range of values from the MPDM output database are divided into *n* intervals (where *n* is an integer value chosen by the analyst). Interval widths are set so that each interval encompasses an equal number of MPDM raster cells, thereby producing a thematic database of *n* categories, each covering identical areas.

This study used the MPDM/EAR approach to produce pseudorandom thematic databases with controlled amounts of spatial autocorrelation. However, we needed to be able to introduce controlled amounts of arrangement dissimilarity into these databases. To accomplish this, we developed a technique termed the *edge correlated error* (ECE) method. This technique was founded upon the psychological studies mentioned previously that found that humans have difficulty recognizing similarities between images when the boundaries of common features differ between the images.

The ECE method starts with two identical copies of a single MPDM/EAR database. One copy is set aside as the *base map*; this database will remain unchanged throughout the ECE process. A controlled amount of arrangement dissimilarity (relative to the base map) will be introduced into the second copy, which is termed the *permeated map*. In addition to these two databases, the ECE method requires as inputs the number of cells into which dissimilarity will be introduced, and an edge correlation (*EC*) measure, whose function will be described shortly.

As its name implies, the ECE method creates dissimilarities whose locations are correlated with edges (category boundaries). To accomplish this, the ECE algorithm starts by constructing a series of distance databases, one for each category in the base map. For the distance database representing category *X*, each cell contains the distance from the cell back to the nearest cell containing category *X*.

Once these distance databases are created, the remainder of the ECE algorithm proceeds iteratively. Each iteration introduces dissimilarity into a single cell of the permeated map, and iterations continue until the desired number of dissimilarities are created. Each iteration starts by randomly selecting a *dissimilarity group*. If the base map contains *c* categories, the dissimilarity introduced into any given cell must fall into one of $c(c-1)$ unique dissimilarity groups. To understand this, consider a dissimilar raster cell that falls into category C_1 in the base map and category C_2 in the permeated map. This is an example of what might be termed C_1C_2 dissimilarity. Alternatively, a C_1 cell from the base map might fall into category C_3 in the permeated map, thereby becoming an instance of C_1C_3 dissimilarity. Simple combinatorics reveals that with *c* categories, there are $c(c-1)$ such dissimilarity groups.

The dissimilarity group chosen for any given iteration of the ECE algorithm may be denoted as $C_B C_P$ (i.e., a cell from category C_B in the base map that falls into category C_P in the permeated map). The algorithm also randomly selects a variate from a Gaussian distribution with a mean of zero and a standard deviation of *EC*. We will term the absolute value of this variate the *maximum allowable distance* (MAD). The algorithm then proceeds by repeatedly selecting random cells until a cell that meets all three of the following criteria is found:

1. The cell has not be impacted by previous iterations of the ECE algorithm.
2. The cell is classified in category C_B in the base map.
3. The cell is no farther than the MAD from a cell falling into category C_P in the base map.

Once a cell meeting these criteria is found, its value in the permeated map is set to C_P, and the ECE algorithm iterates on to find another cell to make dissimilar[1]. Iterations continue until the predefined number of cells into which dissimilarity is to be introduced is reached.

The ECE method's use of the *EC* parameter controls the degree to which error cells introduced into the permeated map are correlated with feature edges. At low *EC* values, the random variates extracted from the Gaussian distribution are likely to be close to zero. Since the absolute value of these variates determines the MAD from a feature boundary at which a dissimilarity can be introduced into the permeated map, small *EC* values correspond to permeated maps where dissimilarities tend to be highly correlated with feature boundaries. Obviously, large *EC* values equate to larger MADs, and thus produce permeated maps where dissimilarities are not highly correlated with feature boundaries.

Examples of databases created by the ECE algorithm are shown in Figure 2. The base map shown in part A of the figure was created via the MPDM/EAR technique, with $H = 0.9$ and containing 100 rows and 100 columns of raster cells. The remaining databases show the impact of altering the ECE's *EC* and total number of error cells inputs.

3.2 Measuring Arrangement Similarity Via Comparisons of Quadtrees

Quadtrees are an alternative method of representing thematic raster databases. In many instances, quadtrees can reduce the memory needed to store a database and shorten the amount of time needed to conduct various analyses ([20], [21]). In essence, quadtrees recursively break the database into quarters until a point is reached where all of the cells in the portion of the database being represented contain a single value. At this point, the common value is represented in the quadtree via a single scalar number. The result is the inverted tree-like structure depicted in Figure 3.

---

[1] It is entirely possible that *no* raster cell meets all three of the criteria for cell selection. To account for this possibility, as implemented the ECE algorithm only randomly evaluates a limited number of cells for possible selection. If it fails to find a suitable cell within that limit, new $C_B C_P$ and maximum allowable distance values are selected, and the search for a suitable cell begins anew.

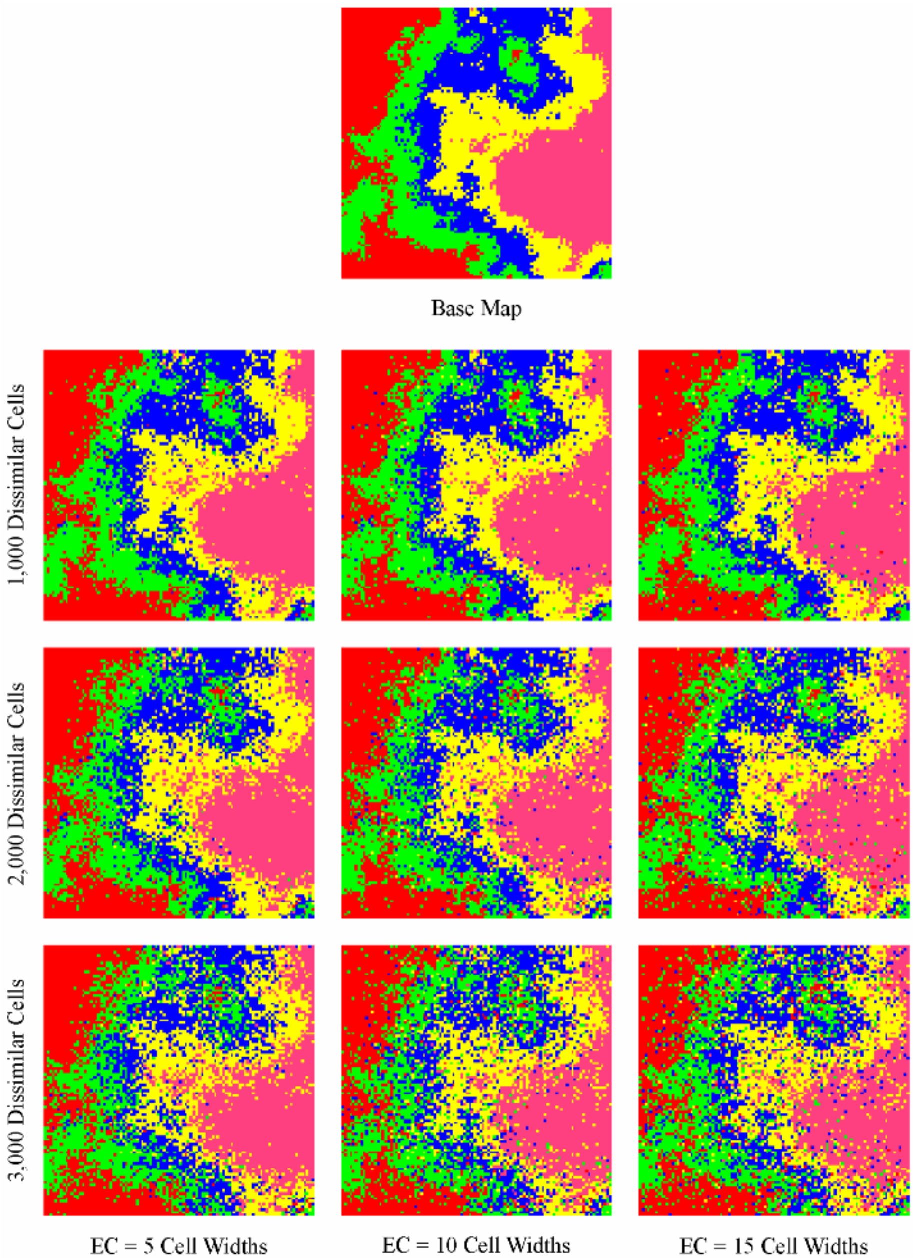


Fig. 2. Examples of ECE outputs

Each four-pronged unit that comprises a quadtree is termed a *node*. Each of the four elements in any given node is either a *pointer* to another node or a *leaf*, i.e., a scalar value describing the values recorded in the raster cells represented by the leaf.

The horizontal rows of nodes that make up the various levels of the quadtree are termed *tiers*, and the number of tiers in a tree determines the *depth* of the tree. The topmost tier in the tree (i.e., the tier containing a single node) is referred to as tier 1, the next tier is tier 2, and so on. Finally, any path from the topmost node in a tree to a leaf is termed a *branch* of the tree.

By its very nature, a quadtree is a multi-scale representation of a raster database. Nodes near of the top of the tree represent large portions of the database while nodes farther down represent smaller portions (i.e., the topmost node represents the entire database, nodes in the second tier represent quarters of the database, third tier nodes represent sixteenths of the database, and so on). Furthermore, if one ignores the leaf values and focuses exclusively on the arrangement of nodes, a quadtree is a complete representation of the arrangement of features within a raster database. Thus, if two raster databases have identical arrangements, they will have identical quadtrees, regardless of the values recorded at any given point in the databases. Returning to the two databases shown in Figure 1 (which have absolutely no point-to-point similarity), we can see that the arrangement of their quadtree representations are identical (Figure 4), thereby indicating that the databases have perfect arrangement similarity.

Thus, arrangement similarity between two raster databases can be measured by evaluating the similarity of their quadtree representations. Any such evaluation will by definition be a multiscale comparison of database similarity, because the trees themselves are multiscale representations of the databases. This is the foundation of the *Q*uadtree-*B*ased *A*rrangement *S*imilarity (QBAS) index.

The QBAS index is simply an average of all of the tier-to-tier comparisons of the trees. Each tier-to-tier comparison is treated as being equally important (i.e., comparisons at each scale are treated as being equally important), so the QBAS index can be defined as shown in equation (1).

$$QBAS = \frac{1}{Max(D_1, D_2)} \times \sum_{i=1}^{Max(D_1,D_2)} TC_i \qquad (1)$$

Where:

D_1	=	The depth of quadtree number 1.
D_2	=	The depth of quadtree number 2.
$Max(X,Y)$	=	The maximum of X and Y.
TC_i	=	The results of a comparison of tier i in quadtree 1 to tier i in quadtree 2.

Recall that in a fully populated quadtree (i.e., a tree where no branches end prematurely; all branches extend all the way through the tree and end in the tree's bottommost tier), tier i will have $4^{(i-1)}$ nodes. Thus, the first (topmost) tier has $4^{(1-1)} = 1$ node, the second tier has at most $4^{(2-1)} = 4$ nodes, the third tier has a maximum of $4^{(3-1)} = 16$ nodes, and so forth. Given this, a tier-to-tier comparison for tier i, which is nothing more then the average of the comparisons of each pair of corresponding nodes from the two databases, can be produced using equation (2).

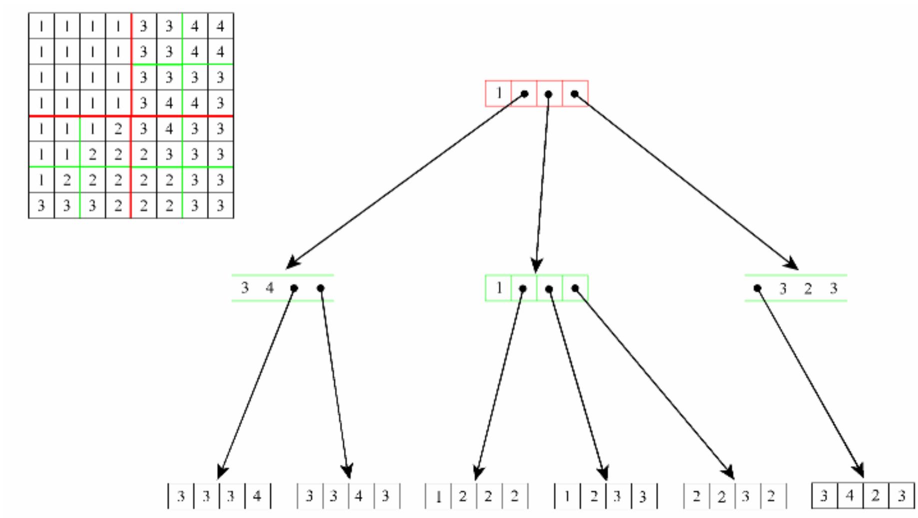


Fig. 3. An example of a quadtree. The elements in each node are arranged as upper left, upper right, lower left and lower right (thus the '1' in the topmost node represents the upper left quarter of the database). Note that the raster database contains a total of 64 scalar values that represent the database while the quadtree requires only 40 scalar values and pointers to represent the same database.

$$TC_i = \frac{1}{4^{(i-1)}} \sum_{j=1}^{4^{(i-1)}} NC_j \tag{2}$$

Where NC_j is the result of comparing node j in quadtree A to the corresponding node j in quadtree B. In order to be meaningful, only nodes representing identical portions of databases can be compared; for example, when considering nodes from the second tier of the quadtrees, the node representing the upper left corner of database A can only be meaningfully compared to the node representing the same portion of database B. Given this method of identifying nodes to be compared, NC_j values can be produced using a series of three simple rules:

1. If neither of the nodes involved in the comparison exist (i.e., the branches leading to the portion of the database represented by the nodes being compared end at leaf nodes in tiers prior to the tier containing the node being analyzed), the nodes are omitted from any further calculation.
2. If one of the two nodes involved in the comparison exists but the other does not, the nodes have no similarity, and thus are assigned a NC_j value of 0.00.
3. If both nodes being compared exist, they are assigned an NC_j value based on the degree of similarity of their four elements. This evaluation looks only at whether or not corresponding elements are of the same type. Thus, if corresponding elements are both pointers or both leafs, the elements are considered similar; if one element is a pointer and the other is a leaf, they are considered dissimilar.

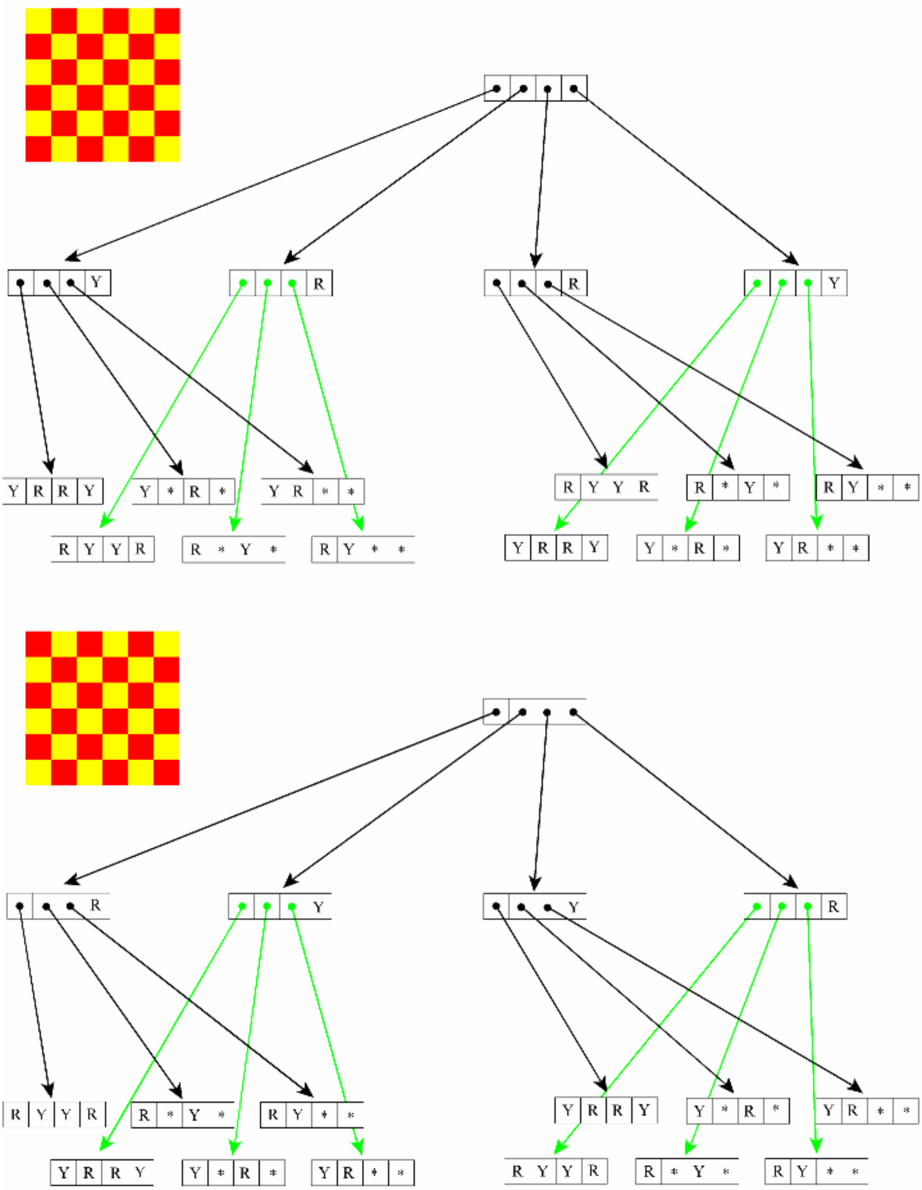


Fig. 4. Quadtree representations of the two databases from Figure 1. The arrangement of nodes within the two trees is identical, implying that the databases have perfect arrangement similarity. Further note that since the number and rows and columns in the databases is not a power of two, some tree elements contain placeholders (denoted with asterisks '*').

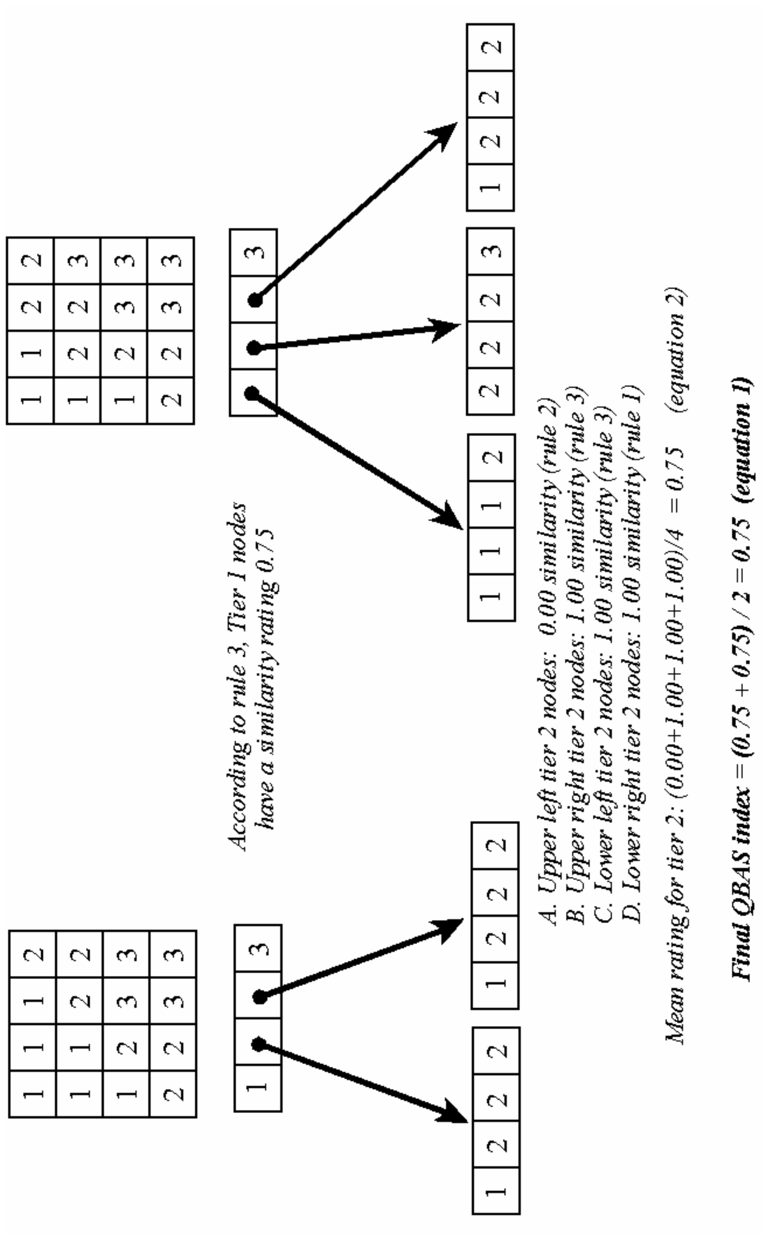


Fig. 5. Example of computing the QBASindex

Thus, if none of the four corresponding elements in the two nodes being compared are similar, the nodes are assigned a similarity of 0.00, if one of the four elements are similar the nodes are assigned a similarity value of 0.25, and so on up to the point where all four of the nodes' corresponding elements are judged as being similar, in which case the nodes are assigned a similarity values of 1.00.

Note that in cases where quadtrees representing databases with numbers of rows and columns that are not powers of two are being compared, some elements may be placeholders that represent nonexistent portions of the database (Figure 4). In these cases, rule 3 is simply extended to include placeholders. Thus, two elements are considered similar if they are both leafs, both pointers, or both placeholders; they are dissimilar if they represent any two of these categories.

The QBAS index ranges from 0.0 to 1.0, where higher values correspond to greater degrees of arrangement similarity. A complete QBAS computation example is provided in Figure 5.

3.3 QBAS Versus the Conventional Kappa Statistic

It is a trivial matter to show that the QBAS index differs from the conventional Kappa statistic (which is described in detail in many remote sensing texts, including [1]). This is hardly surprising, because Kappa is actually a measure of point-to-point similarity corrected to account for the biases that may occur if the categories present in the database cover unequal portions of the map. Kappa is not a measure of arrangement similarity, although it is sometimes erroneously presented as such.

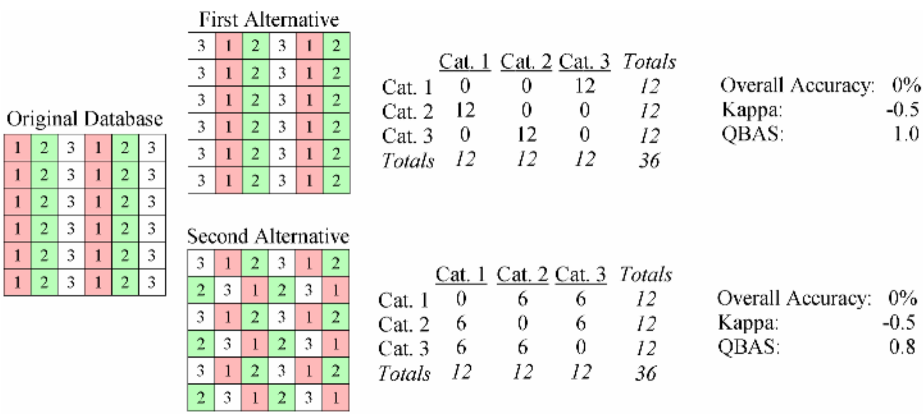


Fig. 6. Comparison of QBAS and the conventional Kappa statistics

Figure 6 shows an example of an original database and two alternatives to which the original is compared. Both pairings (original vs. first and original vs. second) have no point-to-point similarity, and this fact is accurately reported by the overall accuracy figure. Conventional Kappa analysis indicates that both pairs have equal “Kappa similarity,” but even a casual examination of the figure reveals that the first alternative shares a great deal of arrangement similarity with the original (i.e., both databases are composed of vertical stripes of equal-valued cells) while the second does not. The QBAS index reflects this fact, but the Kappa statistic does not.

The Kappa statistic is computed using only error table row and column totals, the total number of cells falling along the tables' main diagonals, and the total number of cells represented in the tables. Since all of these values are identical for both error tables shown in Figure 6, it is mathematically inevitable that the tables would produce identical Kappa values. This reveals the insensitivity of Kappa analysis to database characteristics that directly impact arrangement similarity.

3.4 Evaluating QBAS

The QBAS index has sound theoretic foundations, but empirical study is needed to determine if it actually functions as intended. To that end, QBAS was tested via survey techniques.

Two hundred and eighty eight (288) pairs of base and permeated maps were created using the ECE method. All ECE databases contained 100 rows and 100 columns of raster cells. These pairs of maps represented all possible combinations of:

- *H* values 0.1, 0.2, ..., 0.9 (for a total of eight different *H* values),
- Numbers of categories 3, 5, 7 and 9 (four different numbers of categories),
- Numbers of dissimilar cells 1000, 2000, and 3000 (three numbers of dissimilar cells),
- *EC* coefficients of 5, 10 and 15 cell widths (three different *EC* coefficients)

Thus, the 288 pairs of maps represented all possible combinations of the 8×4×3×3 = 288 parameter values just described. Each pair was printed side-by-side on an 8½-by-11 inch sheet of paper, which allowed each map to appear as a 5-by-5 inch square. Ten printed copies of each pair were produced, resulting in 288×10 = 2,880 printed pages. These printed pages were arranged into 144 bound volumes, with 20 pages per volume. The pages were arranged in a fashion developed by [22] designed to eliminate any biases caused by the absolute or relative order of the pages.

The 144 volumes were evaluated by an equal number of community members from northern Colorado region. Survey takers were asked (via a set of standardized written instructions) to judge the degree of arrangement similarity of each of the 20 pairs of maps shown in their volume on a continuous ten point scale. In addition to writing their ratings on each page of their volumes, respondents were encouraged to make notes describing any problems they encountered in formulating their ratings. Once the ratings were complete, these notes were reviewed, and any pages (or sometimes entire volumes) where the notes indicated that the reviewers were not following the instructions were eliminated from further analysis.

Each respondent's remaining ratings were then normalized by subtracting each respondent's mean rating from each individual rating, and dividing the result by the standard deviation of the respondent's ratings. An analysis of variance (AOV) procedure was then used to evaluate how these normalized rating compared to the *H* values, number of categories, number of dissimilar cells, and *EC* coefficients used to construct the map pairs. When it was found that some of these parameters were related to normalized ratings via highly nonlinear functions, exploratory data analysis (EDA) techniques were used to estimate the shape of these functions.

Finally, QBAS indexes were generated for each of the 288 map pairs used in the survey, and mean survey ratings (across all valid ratings for each pair of maps) were computed. Simple correlation analyses were used to determine how well the QBAS indexes related to these mean normalized ratings.

4 Results and Discussion

After all questionable evaluations were eliminated, 2,124 remained for analysis. This represents 73.75% of the 2,880 map pairs evaluated. Virtually all of the eliminated evaluations were removed because the survey taker's notes indicated that they were at least in part evaluating point-to-point similarity and not solely arrangement similarity.

Table 1 shows the result of the AOV procedure relating normalized survey ratings to the parameters used to build the map pairs. These results indicate that ratings were highly correlated with the number of dissimilar cells and with the *EC* coefficient, which controls the degree to which dissimilar cells are associated with feature boundaries. Ratings were not related to the number of categories in the databases, nor were they related to the amount of fragmentation in the databases (i.e., the *H* value).

These results are very encouraging. The fact that ratings are related to the number of dissimilar cells is the only intuitively viable result; had these variables not been correlated, it would have been difficult to place any confidence in the survey procedures used here and/or the ECE method's ability to construct databases with controlled amounts of dissimilarity. The fact that survey ratings and *EC* values are correlated lends support to the psychological studies that concluded human dissimilarity perceptions are influenced by the proximity of dissimilarities to feature boundaries. The fact that survey ratings are not related to either the number of categories in the database or the degree of fragmentation (*H* values) implies that the ECE method is broadly applicable, and functions across at least the range of number of categories and degrees of fragmentation investigated here.

Table 1. AOV results relating normalized survey results to map pair building parameters

A. Overall AOV model results:

Source	Degrees of freedom	Sum of Squares	Mean Squares	Test Statistic	p Value
Model	4	129.7039	32.4260	2795.53	< 0.0001
Error	2119	24.5787	0.0116		
Total	2123	154.2826			

B. Breakdown of variance components:

Source	Degrees of freedom	Sum of Squares	Mean Squares	Test Statistic	p Value
Num. dissimilar	1	128.1622	128.1622	11049.19	<0.0001
EC coefficient	1	1.5379	1.5379	132.59	<0.0001
Num. categories	1	0.0036	0.0036	0.31	0.5777
H coefficient	1	0.0001	0.0001	0.01	0.9139

The relationship between survey ratings, numbers of dissimilar cells, and *EC* values were visually investigated using three-dimensional graphic techniques. It was clear that the relationship between these variables was highly nonlinear. Using EDA techniques, the nonlinear surface shown in Figure 7 was created. This surface explains over 85% of the variation on the survey data.

As expected, Figure 7 shows that number of dissimilar cells has a consistently negative relationship with similarity. This relationship seems somewhat nonlinear; as the number of dissimilar cells increases, the impact of additional dissimilar cells becomes more pronounced. This is in keeping with the psychological findings that humans are able to ignore dissimilarities up to a point, but beyond that point, additional dissimilarities have significant negative impacts on similarity perception.

The relationship between *EC* and similarity is more complex than that between similarity and number of dissimilar cells. At small values of *EC*, the relationship between *EC* and similarity is positive; increasing the *EC* results in increased similarity ratings. This is intuitive and conforms to the psychological findings that errors concentrated near boundaries are more detrimental to similarity ratings than errors spread about more randomly. However, at larger *EC* values, this relationship breaks down; similarity and *EC* become unrelated. This may be a result of interactions between the numbers of rows and columns in the database, the number of categories, and the *EC* value. As *EC* values are increased for a database of a given size and a given level of division into categories, at some point an *EC* value is reached that extends the region where dissimilarities are likely to be introduced all the way to the edge of the database. Once this point is reached, increasing the *EC* value any farther would not materially increase the randomness of the resulting dissimilarities.

Finally, when mean normalized survey ratings were correlated with QBAS indexes, a simple Pierson's analysis produced a correlation coefficient of 0.88. This implies that to a large degree, the QBAS index is measuring the same arrangement similarity phenomenon that the survey respondents were measuring.

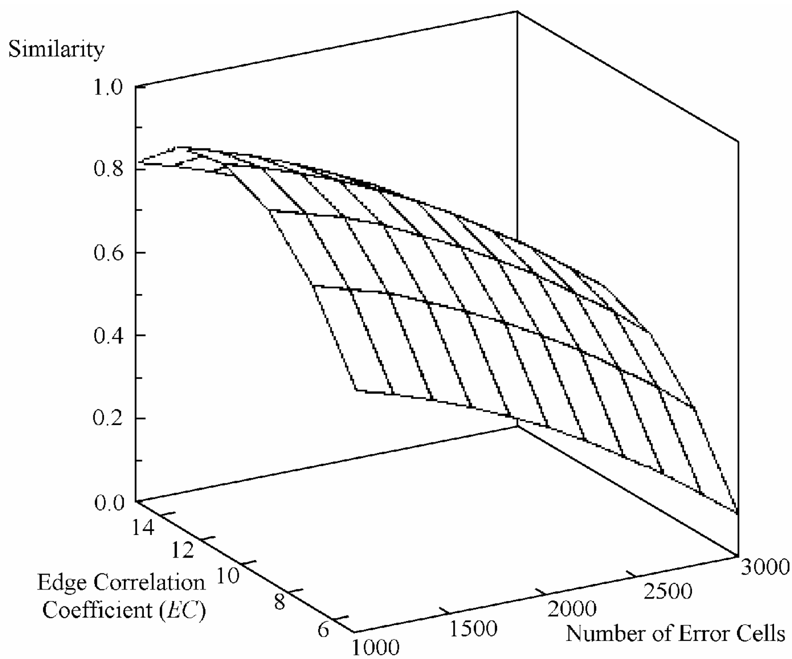


Fig. 7. Relationship between rated similarity, number of error cells, and *EC* coefficient

5 Conclusions

The results of this study were quite positive. The empirical results provided by the survey indicate that the ECE method is a viable way of producing pairs of raster databases with controlled amounts of arrangement similarity. This was certainly a useful capability within the context of this study, and as the utility of existing data generation techniques like MPDM imply, it may be a useful in the future as well.

More importantly, the survey results also indicate that the QBAS index is a viable way of measuring the degree of arrangement similarity between thematic raster databases. The QBAS index is intuitively appealing, easy to implement, inherently multiscale, and after this study, is supported by empirical results. QBAS appears to be a viable statistic, ready for use in future studies.

One area where the results of this study were disappointing is that they do not imply any obvious way to expand the QBAS approach to raster databases containing continuous data, nor how to assess arrangement similarity in vector databases. These topics will remain subjects for future research.

References

1. Jensen, J.R. *Introductory Digital Image Processing*. Prentice Hall, Upper Saddle River, New Jersey, 1996.
2. Seaborn, M., Hepplewhite, L., Stonham, J. Fuzzy Colour Category Map for the Measurement of Colour Similarity and Dissimilarity. *Pattern Recognition*, 38 (2): 165-177, 2005.
3. Tse, P. U. Complete Mergeability and Amodal Completion. *Acta Psychologica*, 102 (2-3): 165-201, 1999.
4. Lohse, G., Walker, N., Biolsi, K., Rueter, H. Classifying Graphical Information. *Behaviour & Information Technology*, 10 (5): 419-436, 1991.
5. Halpern, D.F., Fishbein, H.D., Warm, J.S. Similarity Judgments of Patterns and Maps. *Bulletin of the Psychonomic Society*, 13 (1) 23-26, 1979.
6. Liu, Z.L., Jacobs, D.W., Basri, R. The Role of Convexity in Perceptual Completion: Beyond Good Continuation. *Vision Research*, 39 (25) 4244-4257, 1999.
7. Pontius, G., Malanson, J. Comparison of the Structure and Accuracy of Two Land Change Models. *Int. J. of Geographical Information Science*, 19 (2): 243-265, 2005.
8. Pontius, R. G., Huffaker, D., Denman, K. Useful Techniques of Validation for Spatially Explicit Land-Change Models. *Ecological Modelling*, 179 (4): 445-461, 2004.
9. Fewster, R.M., Buckland, S.T. Similarity Indices for Spatial Ecological Data. *Biometrics*, 57 (2): 495-501, 2001.
10. Patil, G.P., Taillie, C. Multiscale Frequency Table Analysis of Landscape Fragmentation in Thematic Raster Maps. *Sankhya: The Indian Journal of Statistics*, 64 (A2): 344-363, 2002.
11. Belongie, S., Malik, J., Puzicha, J. Shape Matching and Object Recognition Using Shape Contexts. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24 (4): 509-522, 2002.
12. Carstensen, L.W. A Measure of Similarity for Cellular Maps. *American Cartographer*, 14 (4): 345-357, 1987.
13. Hagen, A. Fuzzy Set Approach to Assessing Similarity of Categorical Maps. *Int. J. of Geographical Information Science*, 17 (3): 235-249, 2003.

14. Power, C., Simms, A., White, R. Hierarchical Fuzzy Pattern Matching for the Regional Comparison of Land Use Maps. *Int. J. of Geographical Information Science*, 15 (1): 77-100, 2001.
15. Townsend, P. A. A Quantitative Fuzzy Approach to Assess Mapped Vegetation Classifications for Ecological Applications. *Remote Sensing of Environment*, 72 (3): 253-267, 2000.
16. Carmel, Y., Dean, D.J., Flather, C.H. Combining Location and Classification Error Sources for Estimating Multi-Temporal Database Accuracy. *Photogrammetric Engineering and Remote Sensing*, 67 (7): 865-872, 2001.
17. Carmel, Y., Dean, D.J. Performance of a Spatio-Temporal Error Model for Raster Datasets Under Complex Error Patterns. *Int. J. of Remote Sensing*, 25 (23): 5283-5296, 2004.
18. Saupe, D. Algorithms for Random Fractals. In: Peitgen, H-O., Saupe, D. (eds.): *The Science of Fractal Images*. Springer-Verlag, Berlin Heidelberg New York, pp. 71 – 113, 1988.
19. Burrough, P.A., McDonnell, R.A. *Principles of Geographical Information Systems*. Oxford University Press, Oxford, England, 1998.
20. Samet, H. *Applications of Spatial Data Structures*. Addison-Wesley, Reading, Massachusetts, 1990.
21. Samet, H. *The Design and Analysis of Spatial Data Structures*. Addison-Wesley, Reading, Massachusetts, 1990.
22. Ross, R.T. Optimum Orders for the Presentation of Pairs in the Method of Paired Comparisons. *J. of Educational Psychology*, 25 (4): 370 – 382, 1934.

Extending Semantic Similarity Measurement with Thematic Roles

Krzysztof Janowicz

Institute for Geoinformatics,
University of Muenster, Germany
janowicz@uni-muenster.de

Abstract. Semantic similarity measurement plays a significant role in semantic interoperability and in information retrieval within the geo domain as it supports the detection of conceptually close but not identical entities. In feature-based models, the similarity measurement is done by comparing common and different features such as parts, attributes and functions. This paper suggests adding thematic roles as an additional type of features to be compared, and shows why and how the usage of thematic roles may prevent wrong function matches.

1 Introduction

Ontologies specify a conceptualization of entities represented in geographic information systems (and services), and therefore allow the users to interpret the meaning of the used terms. What makes information retrieval and usage difficult is that users often have no clear class (concept) definition in their mind that could be compared to the specification of the geographic information system or both definitions do not match. Semantic similarity measurement offers the possibility to define an area of interest and to calculate the distance between the classes within this area. In contrast to rigid logic-based reasoning, the result should be more flexible and adaptable, and therefore close the gap between user-expected and system-retrieved meanings.

The Matching-Distance Similarity Measure (MDSM) [1] is such a (feature-based) measurement theory introduced for the geo domain. The intension of this paper is to present an extension to MDSM that is able to measure similarity based on the idea that entity classes whose members share a certain behavior are similar. Thematic Roles are used to model this behavioral aspect, because they offer an abstract theory (that is grounded in Sowa's [2] formal ontology) of roles an entity plays within a certain function.

The goal of this extension is to avoid wrong matches within the functional feature (FF) similarity calculation of MDSM and to improve the robustness of the model by aligning the entity classes to roles described within formal ontology.

M.A. Rodríguez et al. (Eds.): GeoS 2005, LNCS 3799, pp. 137 – 152, 2005.

2 Related Work

This section introduces Thematic Roles (TR), Matching-Distance Similarity Measure, Role-Governed and Transformational Categories as foundation for the semantic similarity measurement extension presented in this paper.

2.1 Thematic Roles

Influenced by the work of Moravcsik [3], Dick [4] and Pustejovsky [5], John Sowa [6] related Somers [7] case grid to Aristotle's idea of four causes (efficient cause, material cause, final cause and formal cause) called *aitiai*. The result is a matrix of six rows representing verb categories (or to be more precise the type of nexus [6]) and four columns representing different kinds of participants. Each of the twenty-four cells represents at least one thematic role such as Agent or Location. These thematic roles are arranged within a hierarchy of participants depending on their position in the matrix. At the top of this hierarchy Source and Product participants are distinguished. At the next level Source is further distinguished into the Initiator and Resource participants and Product subsumes the Goal and Essence participants. Location for example is a special kind of Essence (Location<Essence<Product<Participant) (see Figure 1). In contrast to roles in description logics thematic roles are not binary relations but concepts (unary predicates) [8].

With respect to entities this means that, depending on the context, each entity plays a specific thematic role. For example, a Person who arrives at a sport arena is

	Source		Product	
	Initiator	Resource	Goal	Essence
Action	Actor, Effector	Instrument	Result, Recipient	Patient, Theme
Process	Agent, Origin	Matter	Result, Recipient	Patient, Theme
Transfer	Agent, Origin	Instrument, Medium	Experiencer, Recipient	Theme
Spatial	Origin	Path	Destination	Location
Temporal	Start	Duration	Completion	Point in Time
Ambient	Origin	Instrument, Matter	Result	Theme

Fig. 1. Thematic roles matrix [2][6]

regarded as Actor whereas the sport arena is regarded as Location. The corresponding conceptual graph representation looks as follows:

[Person: Bob]←(Agnt)←[Arrive]→(Loc)→[Sport Arena]

Due to the hierarchical structure of participants the conceptual graph can be shifted on a more abstract level. This is very useful in case of ambiguity, i.e. when it is not clear what role is played by a certain entity [2][6], or for comparison between different cases as discussed later in this paper.

[Person: Sue]←(Agnt)←[Go]→(Dest)→[City: Mexico City]

In both cases the Person plays the role of an Agent and therefore no shift to Initiator is necessary whereas Location (Location<Essence<Product) and Destination (Destination<Goal<Product) have to be replaced by Product which is their immediate common superclass. Entities are not restricted to occupy the same thematic role in different cases and therefore Bob becomes the Recipient (Goal) in "*Sue sent the gift to Bob by Federal Express*" [2, p. 506].

Sowa [2][6] places the thematic roles in an intermediate level of his formal ontology and suggests creating subtypes for each kind that is of interest for a certain domain or context (e.g. TaxiDriver<Driver<Doer<Agent<Initiator<…). Sowa argues that Driver only represents persons who are actively driving a vehicle and that therefore a LicensedDriver (e.g. Chauffeur) can not be a subtype of Driver e.g. because licensed drivers are legally authorized to drive a vehicle whether they are driving it right now or not.

2.2 Matching-Distance Similarity Measure

MDSM is the asymmetric and context sensitive semantic similarity measurement approach for entity classes developed by Rodriguez and Egenhofer [1]. It can be regarded as an extension of Tverskys [9] ratio model and therefore is classified as a feature-based approach to similarity (in contrast to geometric and alignment models for example [10][11][12] which calculates the similarity using the number of common and different features. Three kinds of features can be distinguished: parts, which are structural components of a class such as wall for building; functions which describe "what is done to or with a class" [1, p. 232} such as the function educate is offered by college (the idea of functions in MDSM is close to Gibson's [13] affordances) and attributes which are additional characteristics that can not be regarded as parts or functions such as name or owner type for building.

$$S(c_1,c_2)=\omega_p \cdot S_p(c_1,c_2)+\omega_f \cdot S_f(c_1,c_2)+\omega_a \cdot S_a(c_1,c_2) \tag{1}$$

Equation 1 displays the overall semantic similarity measurement, which is regarded as the sum of the weighted similarities of the three kinds of features (parts, attributes and functions) of the compared entity classes c1 and c2.

$$S_t(c_1,c_2)=\frac{|C_1 \cap C_2|}{|C_1 \cap C_2|+\alpha(c_1,c_2)\cdot|C_1 \setminus C_2|+(1-\alpha(c_1,c_2))\cdot|C_2 \setminus C_1|} \tag{2}$$

Equation 2 describes the no-symmetric similarity function for each of the feature types. S_t (c1, c2) is defined as the similarity for the feature type t between the entity classes c1 and c2 where C_1 and C_2 are the sets of features of type t for c1 and c2, $|C_1 \cap C_2|$ is the cardinality of the set intersection and $|C_1 \setminus C_2|$ is the cardinality of the set difference.

The relative importance α (equation 3) of the different features of type t is defined in terms of the distance d between c1 and c2 within a hierarchy that takes taxonomic and partonomic relations into account. Lub denotes the least upper bound, i.e. the immediate common superclass of c1 and c2 [1].

$$\alpha(c1,c2) = \begin{cases} \frac{d(c1,\text{lub})}{d(c1,c2)}, d(c1,\text{lub}) \leq d(c2,\text{lub}) \\ 1-\frac{d(c1,\text{lub})}{d(c1,c2)}, d(c1,\text{lub}) > d(c2,\text{lub}) \end{cases} \quad where: d(c1,c2) = d(c1,\text{lub}) + d(c2,\text{lub}) \tag{3}$$

MDSM takes context into account and therefore the weighting in the overall similarity function (equation 1) is calculated depending on the domain of application using variability or commonality within the features (of each type). "Contextual information (C) is specified as a set of tuples over operations (op_i) associated with their respective noun arguments (e_j) (Equation 4). The nouns correspond to entity classes in MDSM, while the operations refer to verbs that are associated with methods of these classes." [1, p. 239] A contextual specification such as C = <(play, {})> for example expresses a domain of application that contains all entity classes which share the functional feature play.

$$C = \left\langle (op_1, \{e_1,...,e_m\}),...,(op_n,(e_1,...,e_l\}) \right\rangle \tag{4}$$

Within such a context the relevance (ω_t in equation 1) of each feature type is defined either by the variability P_t^v (equation 5) or commonality P_t^c function (equation 6) and then normalized with respect to the remaining feature types so that $\omega_p+\omega_f+\omega_a$ is always 1.

$$P_t^v = 1 - \sum_{i=1}^{l} \frac{o_i}{n \cdot l} \tag{5}$$

The variability describes how diagnostic [9][14] a feature type t is within a certain domain of application by assuming that the more characteristic each feature is for a given class the more diagnostic it is. A certain feature of type t has low relevance if it appears in many classes and high relevance if it is not common to the classes within the domain. P_t^v is the sum of the diagnosticity of all features of the type t in the domain and therefore 0 when all features a shared by all entity classes (P_t^v=1-1=0) and close to 1 if each feature is unique (where o_i is the number of occurrences of the feature within the domain) and the number of features l and classes n in the domain is high.

$$P_t^c = \sum_{i=1}^{l} \frac{o_i}{n \cdot l} = 1 - P_t^v \tag{6}$$

Commonality is defined as the opposite of variability ($P_t^c = 1 - P_t^v$) and assumes that by defining a domain of application the user implicitly states what features are relevant [1].

2.3 Role-Governed and Transformational Categories

Depending on the classification and the level of granularity several kinds of categories can be distinguished. Beside common also called feature-based categories, role-governed and transformational categories are of special interest for this paper.

Role-Governed Categories

In contrast to common categories, members of role-governed categories are not grouped together because they share a set of necessary (and sufficient) features, but due to a certain role they play within a domain or context [15][16]. Wittgenstein [17] argued, that it is difficult to find a feature-based representation of Game, but as described by Markman and Stilwell [15] Game may be regarded as a role governed category that is specified as being the second argument of the relation play(Player, Game) where Player is also defined in a role-governed way. In other words, games are the entities played by players.

The coherence between the category members is merely based on few (or even one) significant core roles and therefore the overall similarity between the members is low in general [16][18]. Nevertheless graded structure also exists for non feature-based categories and therefore similarity measurement is possible in principle [19].

Moreover role-governed categories cannot be arranged within feature hierarchies as this is possible for feature-based categories. On the one hand they do not inherit properties and on the other hand - besides a very abstract functional theory - they do not necessarily share a common role [15].

The importance of social roles for concepts such as money or president is discussed by Masolo et al. [8]. The importance of roles in the geo domain is elucidated by Kuhn [20].

Transformational Categories

Markman and Stilwell [15] claim that there is an additional kind of categories called transformational categories that specify a change in a certain selection restriction for a relation. For example, according to the specification of Markman and Stilwell a player in the relation play (Player, Game) has to be a sentient being, but a team can also play a game. Thus, team is a transformational category that transforms a group to an individual. Metonymy can be regarded as a linguistic and cognitive device for the creation of transformational categories [15].

3 Why Play and Play Do Not Match

This section discusses the relation between the functions as defined in the lightweight ontology used within MDSM on the one hand and the thematic roles on the other hand.

3.1 Two Shortcomings of Feature-Based Similarity Models

Beside others [11], there are two main shortcomings that (more or less) affect all kinds of feature-based models. All features are unary, which means that an entity which is green for example, is described by the feature green and not by a feature-value pair such as color = green. On the level of entity classes an adult (in Germany) would have to be defined by the feature over17 and not age > 17. This simplification may lead to difficulties [11] for example if it is not clear whether the feature Height of the entity class Theater [1] corresponds to the height of the building or the height of the stage. The use of two separate features such as BuildingHeight and StageHeight is impractical because it will decrease similarity to all other buildings. In the case of functional features the play function of Sport arena and Game are regarded as a common feature of both classes and therefore match. This is possible because the relation between function and entity class is very loosely defined in (the lightweight ontology used in) MDSM. A function of a class can be anything that is afforded by this class independently of which role it plays within this function (and due to polysemy of verbs one could imagine many different play functions). In the upper case the entity class is either the location where one can play or the thing that is played. The KIF like code fragments below shows two simplified specifications of play, whereas the first function might also be named played-at.

```
1. (DEFRELATION PLAY (?X ?Y)
   :=> (AND (GAME ?X) (SPORTARENA ?Y)))

2. (DEFRELATION PLAY (?X ?Y)
  :=> (AND (PLAYER ?X) (GAME ?Y)))
```

MDSM is able to deal with polysemy of entity class names using taxonomic and partonomic relations but not with polysemous feature names [1].

A second weakness of feature-based similarity models (and also geometric approaches) is that they regard classes as bags of unsorted features, which means that there is no structure connecting the features within a class or even to other classes. The topological relation above(Circle, Triangle) [11] does not describe the same fact as above(Triangle, Circle). In a similarity assessment subjects may judge above(Triangle, Circle) to be more similar to above(Rectangle, Circle) than to above(Circle, Triangle) because of the same role (being under something else) that the circle plays within the first two examples (see also [21]). A first step to solve these problems is to structure the features into types as done by MDSM. Moreover Rodriguez and Egenhofer [1] propose to investigate the semantic comparison of distinguishing features (in contrast to the comparison of their labels as done so far).

As argued by Goldstone and Son [11] in extreme cases the combination of both shortcomings may lead to a spurious match in the feature-based similarity model because "a car with a green wheel and a truck with a green hood both share the feature green" [11, p. 15]. In fact the wheel and the hood share the feature green but not the compared car and truck and therefore this kind of match should not increase similarity.

3.2 Functional Features and Thematic Roles

Functions in the feature-based approach are just synsets such as {play} or {recreate, play} [1] whereas thematic roles can be regarded as entity classes which are subclasses of Participant [2]. The relation between both is that the entity class to which the functional feature belongs is a special kind of participant and the kind of function determinates its possible role. The same synset {play} may represent functions that involve different participants. These participants such as player, game or sport arena are subtypes of thematic roles. In play(Player, Game) Player is a subtype of the thematic role Agent and Game of Theme whereas in a case such as play(Player, Sport arena) the second parameter is not a subrole of Theme but of Location (see Figure 3). If both play functions are features of two different entity classes they should not be regarded as match (a feature that is common to both classes) and thus increase similarity between the compared entity classes as done in the MDSM model so far. Functional features should only count as common if the compared entity classes both occupy the same functional roles within these functions.

Sometimes it may not be clear which thematic role has to be taken out of the twenty-four cell matrix to describe the role of an entity class within a functional feature. In many cases thematic roles can be directly excluded considering their conceptual relation as described by Sowa [2]. For example, sport arena can not be a Recipient because the corresponding conceptual relation Rcpt(Act, Animate) restricts the usage of Recipient to Act and Animate. In the case of a valid allocation such as play(Agent, Location) for play(Player, Sport arena) as functional feature of the entity class Player, the resulting conceptual graph is:

```
[Player]←(Agnt)←[Play]→(Loc)→[Sport Arena]
```

In other cases of uncertainty it is possible to use the immediate superclass instead of a concrete functional role such as Essence for Location.

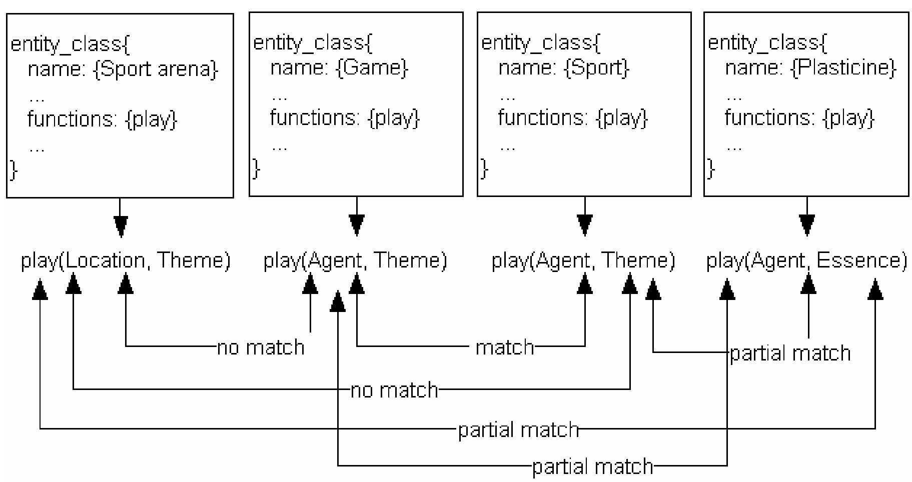


Fig. 2. Full and partial matches between the functional feature play

As depicted in Figure 2, an extended semantic similarity measurement approach should count the resulting matches as partial match. Game and Sport both occupy the same thematic role (Theme) in the play function (full match), whereas Plasticine is defined as subrole of Essence in Figure 2 and therefore the match between Sport and Plasticine is only partial.

3.3 Thematic Roles as Feature Type

While the former section discused the relation between functional features and thematic roles, this section illuminates the question whether thematic roles can be regarded as feature type additionally to parts, functions and attributes as used in MDSM.

The question what can be done to or with something or what this thing affords to its environment seems to be a suitable way to model and categorize the world and especially artifacts [13][15][22][23]. As function is defined as the "role that an entity plays in serving the goal of an agent, or its role in the operation of a larger system such as a geology, ecology, or religion" [22, p. 2] and entities may have more than one function, it follows that an entity can play different roles within different contexts. Bob is the Agent of giving but the Recipient of receiving for example.

For entity classes this means that they can be subtypes of several thematic roles such as Agent and Recipient for Person (even at the same time: hurting oneself). This seams to contradict Sowa's idea of the placement of the thematic roles within formal ontology [2][6].

A possible solution is to regard thematic roles played by an entity class as directly connected to its functions. Person, in this sense it not a subclass of Agent and Recipient but an entity class with two functions (give and receive) that impose a certain role to the class. In other words Person is only an Agent in the situation of giving and not before or after a Theme is given to a Recipient. Sowa would argue that Person is no kind of Participant at all, but Giver and Receiver are (Giver is always an Agent for example). From this point of view thematic roles cannot be regarded as an additional kind of feature but have to be directly assigned to functions. One can even argue that in a valid model each subtype of a thematic role can only have one function namely the one that makes it an Agent for example.

Nevertheless another argumentation is possible that regards thematic roles as a way to describe the potential (potential ability) of class members. In this case thematic roles can be regarded as feature type that describes the tendency of how entities of a certain class behave. Stadium and Sport arena for example share the thematic role feature Location. Thus, all their members tend to behave as locations in associated functions, which make both entity classes seem to be similar. In cases where an entity class is described by more than one role feature the same conclusion is possible. Persons for example are entities that can behave either as Actors or Recipients but not as points in time or paths. Entity classes that have thematic role features in common can be therefore regarded as more similar than classes that differ in their role features.

3.4 Thematic Roles and Transformational Categories

As example for functions, Rodriguez and Egenhofer [1] argue that the class College has the functional feature educate. The function educate(x,y) can be interpreted either in a way that the college educates students or that the college is the location where students are educated. The corresponding conceptual graph for the former interpretation is:

```
[Building: Collage]←(Agnt)←[Education]→(Rcpt)→[Student:{*}]
```

The corresponding conceptual relation of Agent Agnt(Act, Animate) "restricts the usage to an active animate entity that voluntarily initiates an action" [2, p. 508). On the one hand this would mean that College is only the location where education takes place. On the other hand this is a classical example for metonymy [24] and reflects a human way of thinking and categorizing. Other examples for Metonymy in class definitions are the functions perform and present that are defined as features of Theater by Rodriguez and Egenhofer [1]. Again two interpretations are possible: a non-metonymic interpretation where Theater is the location where a group of actors present (or perform) a play and a metonymic interpretation where the Theater stands for the actors that present a play to the visitors. In the former case Theater can be regarded as subtype of Location (if we accept roles as feature types) and in the second case as of Agent. Theater specifies a change in the selection restriction of present(x, y) in a way that first Theater stands metonymical for a group of actors and than in a second step the group is regarded as a single Agent [15], which means that when Theater is defined as a transformational concept there is no contradiction in Agnt(Perform, Theater). By using thematic roles, an ontology engineer can restrict possible interpretations to the intended one.

3.5 Requirement for a Thematic Role Sensitive Similarity Measurement

An extended (feature-based) semantic similarity measurement theory that is able to deal with polysemy of functional features, metonymy within entity class names, potential behavior of class members (entities) and classes that are mostly defined by their role (role-governed) should support both views on thematic roles (as part of functional features and as feature type) and offer weightings for full and partial matches. Moreover, it should be able to integrate thematic roles in its context definitions.

4 Extending MDSM with Thematic Roles

In this section it is shown why and how an extended matching distance similarity measure (MDSM+TR) is able to fulfill the above requirements.

4.1 Extending the Entity Class Definition

MDSM requires a special class definition format, which can be regarded as a lightweight ontology [1]. To extend MDSM this class definition needs to be changed. As shown in Table 1, functions are defined as synstes, each synset containing different words that represent the synonym symbols for a certain function. It is not possible to add the thematic roles as synset here, because they are not synonyms for functions. Instead, functions have to be referenced by pointers as this is done already for the entity classes in is_a for example. As minimum assumption for MDSM+TR, functions are defined by their name and the thematic role the possessing class (role_of_class) plays within this function. This definition is able to capture the relation between functions and thematic roles. To regard roles as an additional feature type another extension is necessary that defines the thematic role type as a list of roles as shown in Table 1.

Table 1. Differences between MDSM and MDSM+TR notation

<table>
<tr><th>BNF Notation: MDSM</th><th>BNF Notation: MDSM+TR</th></tr>
<tr><td><entity_class>::= **entity_class** {
name: {syn_set}
description: <description>
is_a: <is-a>
part_of: <part_of>
whole_of: <whole_of>
parts: <parts>
functions: <functions>
attributes: <attributes>}

<is_a>::= {}|{<pts_entity_classes>}
<part_of>::={}|{<pts_entity_classes>}
<whole_of>::={}|{<pts_entity_classes>}
<parts>::= {}|{<syn_sets>}
<functions>::= {}|{<syn_sets>}
<attributes>::= {}|{<syn_sets>}
<syn_sets>::={<syn_set>}|
<syn_sets>,{syn_set}
<syn_set>::= <word>|<syn_set>,<word>
<description>::= <word>|
<description><word>
<pts_entity_classes>::=<pointer>|

<pts_entity_classes>,
<pointer></td><td><entity_class>::= **entity_class** {
…
functions: <functions>
thematic_roles: <functional_roles>
…
< functions>::= {}|{<pts_functions>}
<pts_functions>}::=<pointer>|
<pts_functions>,
<pointer>
<functional_roles>::={}|
{<functional_role>}|
<functional_roles>,
{<functional_role>}
<functional_role>::=<{x ∈ TR}>
…

<function>::= **function** {
name: {syn_set}
role_of_class: <functional_role>}
…</td></tr>
</table>

4.2 Similarity Between Functional Features in MDSM+TR

In MDSM (Equation 2) $|C_1 \cap C_2|$ is defined by comparing the synset of each (functional) feature of c_1 to c_2. In other words, the (implicit) equal function used in MDSM examines function names (or sets of function names) and returns 1 for a match and 0 if the names do not match. For MDSM+TR this is not enough, because partial matches should be allowed too, and hence the thematic roles (role_of_class, see Table 1) have to be taken into account. Therefore the strength of a match has to be calculated (Equation 7) and the average of all matches has to be defined as weighting for $S_f(c_1,c_2)$.

$$\text{match}(tr_1, tr_2) = \frac{1}{1 + \text{arc\_distance}(tr_1, tr_2)} \tag{7}$$

The match function returns 1 for a full match (tr1=t2; e.g. match(Agent, Agent)) and $^1/_2$, $^1/_3$, $^1/_5$ or $^1/_7$ for the four different kinds of partial matches that are possible within the hierarchy of thematic roles (each is-a relation is regarded as one arc) [2]. In Equation 8 the weighting function ω_{ffp} is defined, that sums the strength of all matches within $C_1 \cap C_2$ and calculates the average ($c_1.ff_i.tr$ is the thematic role c_1 plays within the functional feature ff_i). The values for ω_{ffp} range between 1, if all matches are full matches and $^1/_7$ if only the function names match, but the roles are entirely different.

$$\omega_{\text{ffp}} = \frac{\sum_{i}^{|c_1 \cap c_2|} match(c_1.ff_i.tr, c_2.ff_i.tr)}{|c_1 \cap c_2|} \tag{8}$$

Some cells of the thematic roles matrix contain two roles, in this case arc_distance is defined as 2 (e.g. Agent-Initiator-Effector).Equation 9 shows the MDSM+TR version of $S_f(c_1,c_2)$ whereas S_p and S_a remain as there are.

$$S_f(c_1, c_2) = \omega_{\text{ffp}} \cdot \frac{|C1 \cap C2|}{|C_1 \cap C_2| + \alpha \cdot (c_1, c_2) \cdot |C_1 \setminus C_2| + (1 - \alpha \cdot (c_1, c_2)) \cdot |C_2 \setminus C_1|} \tag{9}$$

4.3 Similarity Between Thematic Role Features in MDSM+TR

In order to take thematic roles as additional feature type into account it is necessary to extend the overall similarity measure $S(c_1,c_2)$ by a weighting ω_{tr} and the similarity measurement for roles $S_{tr}(c_1,c_2)$ as described in Equation 10. The similarity function $S_t(c_1,c_2)$ is the same as in MDSM and each role can appear only one time per entity class.

$$S(c_1, c_2) = \omega_p \cdot S_p(c_1, c_2) + \omega_f \cdot S_f(c_1, c_2) + \omega_a \cdot S_a(c_1, c_2) + \omega_{tr} \cdot S_{tr}(c_1, c_2) \tag{10}$$

4.4 Thematic Roles and Context in MDSM+TR

In MDSM the weighting function ω_t is defined by variability or commonality and then normalized, so that the sum of the weightings is always 1. For P_f^v and P_f^c one has to decide whether the number of occurrences (o_i) of a certain functional feature within the domain of application is determined by its name or the combination of name and role. Partial matches can not be taken into account here, because this would violate the model of variability and commonality within MDSM. The author prefers the latter method because it reduces the effect of polysemous function names, increases variability (decreases commonality) and therefore strengthen the importance of functions within overall similarity. This is especially important for entity classes that are mostly defined by their functions (role-governed) and artifact classes (such as buildings or devices) in general.

$$\omega_{tr} = \frac{P_{tr}^{\ v}}{P_p^{\ v} + P_f^{\ v} + P_a^{\ v} + P_{tr}^{\ v}} \tag{11}$$

In the case where thematic roles are regarded as additional feature type, P_t^v and P_t^c do not need to be changed, but the weighting functions (6a, 6b, 6c and 8a, 8b, 8c in [1]) have to be extended by P_{tr}^v or P_{tr}^c as this is demonstrated for variability in ω_{tr} (Equation 11).

5 Theater, Sport Arena and Guitar

This section presents some measurement examples from a test-ontology and discusses the different results between MDSM and MDSM+TR.

5.1 Experiment

To prove the idea of the thematic roles extension semantic similarity between the entity classes Theater, Sport arena (both taken from Table 2 of [1]) and Guitar is measured using MDSM and MDSM+TR. Theater is defined in to ways: one that regards Theater as Actor of the functional features perform and present and another where Theater plays the role of a Location (see Table 2).

Table 2. Feature description for Theater, Sport arena and Guitar

Entity Class	Parts	Functions	Attributes	Roles
Theater_1	Dressing room Entrance hall Foundation Orchestra Roof Spectator stands Stage Ticket office Wall	Perform(L) Present(L) Recreate(L)	Architectural properties Ext. material construction Height Location Name Owner type Structure type User type	Location
Theater_2	As above	Perform(A) Present(A) Recreate(L)	As above	Agent Location
Sport arena	Court Dressing room Foundation Roof Spectator stands Wall	Play(L) Practice(L) Recreate(L)	Architectural properties Ext. material construction Height Location Name Owner type Structure type User type	Location
Guitar	Body Strings	Play(I) Practice(I) Recreate(I)	Type Material Color	Instrument

The context is defined as C=<{recreate, {}> which means that the domain of application contains the four entity classes displayed in Table 2. It has to be emphasized that Theater_1 and Theater_2 are both taken into account for the calculation of weightings which decreases variability within the domain. Moreover Guitar (which is used here as a kind of false-positive for the similarity calculation within functional features in MDSM and therefore contains the same functions as Sport arena) is specified by few features only which additionally decrease variability.

The aim of this similarity measurement experiment is to show how MDSM+TR behaves in certain situations in comparison to MDSM. Theater_1 and Theater_2 will never be part of the same ontology and same context in real world measurements for example.

Table 3. Some relevant values from the similarity measurement with MDSM and MDSM+TR

Model	c_1 versus c_2	P_f^v	P_{tr}^v	$S_f(c_1,c_2)$	$S_{tr}(c_1,c_2)$	$S(c_1,c_2)$
MDSM	Theater_1 vs. Theater_2	0.4	--	1.0	--	**1.0**
MDSM+TR	Theater_1 vs. Theater_2	0.7	0.58	0.71	0.66	**0.82**
MDSM	Theater_1 vs. Sport arena	0.4	--	0.33	--	**0.66**
MDSM+TR	Theater_1 vs. Sport arena	0.7	0.58	0.33	1.0	**0.7**
MDSM	Theater_2 vs. Sport arena	0.4	--	0.33	--	**0.66**
MDSM+TR	Theater_2 vs. Sport arena	0.7	0.58	0.33	0.66	**0.61**
MDSM	Guitar vs. Sport arena	0.4	--	1.0	--	**0.32**
MDSM+TR	Guitar vs. Sport arena	0.7	0.58	0.43	0.0	**0.14**

5.2 Discussion of the Results

The results presented in Table 3 show some relevant results from the similarity measurement using MDSM and MDSM+TR, where S is the overall similarity, S_f and S_{tr} are the similarities for the functional features and thematic roles and P_t^v and P_{tr}^v are the results for variability of functional features and thematic roles. The functional feature extension of MDSM+TR tends to decrease similarity because it introduces more information about functions. If name and role_of_class are equal for the compared functional features the results between MDSM and MDSM+TR do not differ (Theater vs. Sport arena), but are decreased the more different the roles of the entity classes within the compared functions are. Therefore S_f(Theater_1, Theater_2) is not 1.0 but 0.71 in the MDSM+TR approach and 0.43 instead of 1.0 for S_f(Guitar, Sport arena). The functional features of Guitar and Sport arena have nothing more than their names in common (polysemous function names).

The thematic role feature type offers an additional possibility to compare entity classes and is therefore able to increase or decrease similarity. On the one hand in S(Theater_2, Sport arena) the overall similarity is decreased because Theater_2 does not only play the role of a Location but can be regarded as an Agent too. On the other hand S(Theater_1, Sport arena) is increased by S_{tr}(Theater_1, Sport arena) because the compared classes both play the role of a Location.

In border cases such as S(Guitar, Sport arena) the differences between MDSM and MDSM+TR may be very heigh (MDSM: 0.32; MDSM+TR: 0.14) but in general the results should not vary more than between 5-20%. The thematic role feature type similarity $S_{tr}(c_1,c_2)$ has more impact on the model than the role-based partial matches for $S_f(c_1,c_2)$. Therefore the latter one can be regarded more as a refinement than an extension to the MDSM theory.

6 Conclusions and Future Work

Thematic roles can be easily integrated into MDSM and improve the theory to fulfill the requirements defined in this paper. The resulting MDSM+TR is able to handle polysemous functional feature names and metonymy within entity class names. By taking thematic roles as an additional feature type into account, MDSM+TR is able to measure similarity based on the idea that entity classes whose members behave in a common way (play a certain role) are similar. Thematic roles are more than just another feature type such as parts, functions and attributes, because they come with a very generic theory of participation that adds more structure to the entity class description (and the functional features). While the names (symbols) and the meaning of other features may differ from ontology to ontology, thematic roles are fixed within Sowa's formal ontology and therefore are able to restrict possible interpretations. The ontology design process has fundamental influence on the similarity measurement and as argued in Goldstone and Son [11] all entity classes can be made similar to each other by adding features such as lessthan5000pound or colored for example. Moreover we do not measure similarity between concepts (in our mind) or real world entities but between representations (models); what sounds trivial first, is a fundamental restriction to all assumptions made by using computational theories of similarity. Even within a single ontology granularity can vary between the concept specifications, which directly influence the resulting similarity. All we can state from this kind of measurement is that according to the examined ontology c1 and c2 are similar to a certain degree represented by a numerical value. It is up to the user to decide what similarity value is sufficient for a certain task. MDSM uses a lightweight ontology that primarily consists of meaningless labels without any relation to each other or axioms, which additionally increases the influence of the ontology engineer and makes the measurement very design and granularity dependant. Nevertheless similarity is an important theory for information retrieval and discovery within ontologies, because it is not only able to return classes suitable for a certain task but offers also a ranking. The extension presented in this paper is a first step to a more semantic comparison of distinguishing features (functional and thematic role features) as proposed by Rodriguez and Egenhofer.

A lot of work remains to be done such as human subject testing. Moreover the theory presented here only takes the participant hierarchy into account to express partial matches leaving the verb categories beside. Future work is necessary to analyze how this aspect can be added to the model. The six verb categories are not a final set and on a very abstract level, Sowa [6] argued that they can be divided into more categories if necessary. For the geo domain it would be of special interest to analyze the temporal and spatial categories and create additional sub roles if necessary.

References

1. Rodríguez, A. and Egenhofer M. J. Comparing Geospatial Entity Classes: An Asymmetric and Context-Dependent Similarity Measure. *International Journal of Geographical Information Science*, 18(3): 229-256, 2004.
2. Sowa, J. F. *Knowledge Representation: Logical, Philosophical, and Computational Foundations*. Brooks Cole Publishing Co., Pacific Grove, CA., 2000.
3. Moravcsik, J. M. What Makes Reality Intelligible? Reflections on Aristotle's Theory of Aitia. *Aristotle's Physics: A Collection of Essays*. Ed. Lindsay Judson. New York: Clarendon, pp. 31-48, 1991.
4. Dick, J. A conceptual, case-relation representation of text for intelligent retrieval, PhD thesis, Department of Computer Science, University of Toronto. Published as technical report CSRI-265, 1991.
5. Pustejovsky, J. *The Generative Lexicon*. Cambridge/London: MIT Press, 1995.
6. Sowa, J. F. Processes and Participants. In Peter Eklund, Gerard Ellis, and Graham Mann, editors, *Conceptual Structures: Knowledge Representation as Interlingua, number* 1115 in Lecture Notes in Artificial Intelligence. Springer-Verlag, 1996.
7. Somers, H. L. *Valency and case in computational linguistics.* Edinburgh: Edinburgh University Press, 1987.
8. Masolo, C., Vieu, L., Bottazzi, E., Catenacci, C., Ferrario, R., Gangemi, A., Guarino, N. Social Roles and their Descriptions. In: *Proceedings of the Ninth International Conference on the Principles of Knowledge Representation and Reasoning*, 2004.
9. Tversky, A. Features of similarity. *Psychological Review*, 84(4): 327-352, 1977
10. Goldstone, R. L. The role of similarity in categorization: providing a groundwork. *Cognition*, 52: 125-157, 1994.
11. Goldstone, R. L. and Son J. Y. Similarity. Cambridge *Handbook of Thinking and Reasoning*. K. Holyoak and R. Morrison. Cambridge, Cambridge University Press, 2004.
12. Gärdenfors, P. *Conceptual Spaces - The Geometry of Thought.* Cambridge, MA, Bradford Books, MIT Press, 2000.
13. Gibson, J. *The Ecological Approach to Visual Perception*. Boston, Houghton Mifflin Company, 1979.
14. Goldstone, R. L., Medin, D. L. and Halberstadt, J. Similarity in context. *Memory and Cognition*, 25: 237-255, 1997.
15. Markman, A. B. and Stilwell, C. H. Role-governed categories. *Journal of Experimental and Theoretical Artificial Intelligence*, 13: 329-358, 2001.
16. Gentner, D., and Kurtz, K. Learning and using relational categories. In W. K. Ahn, R. L. Goldstone, B. C. Love, A. B. Markman & P. W. Wolff (Eds.), *Categorization inside and outside the lab*. Washington, DC: APA (in press)
17. Wittgenstein, L. *Philosophical investigations, trans. G.E.M. Anscombe*. New York: MacMillan, 1968.
18. Jones, D. M. and Love B. C. Beyond common features: The role of roles in determining similarity. *CogSci 2004 - 26th Annual Meeting of the Cognitive Science Society*, Chicago, US, 2004.
19. Barsalou, L. W. Ad hoc categories. Memory & Cognition 11(3) (1983) 211-227
20. Kuhn, W. Modeling the Semantics of Geographic Categories through Conceptual Integration. GIScience 2002, Boulder, CO, USA, D. Mark, Editor. Springer: Berlin, pp. 108-118, 2002.
21. Markman, A. and Gentner D. Structural Alignment during Similarity Comparisons. *Cognitive Psychology*, 25: 431-467, 1993.

22. Barsalou, L.W., Sloman, S.A, and Chaigneau, S.E. The HIPE theory of function. In L. Carlson & E. van der Zee (Eds.), *Representing functional features for language and space: Insights from perception, categorization and development*. Oxford: Oxford University Press, (in press).
23. Khoshafian, S. and Abnous, R. *Object Orientation: Concepts, Languages, Databases, and User Interfaces*. New York, John Wiley & Sons, 1990.
24. Radden, g. and Kövecses, Z. Towards a Theory of Metonymy. In *Panther*, Klaus-Uwe; Radden, Günter (eds.), 1999.

Exploiting Geospatial Markers to Explore and Resocialize Localized Documents

Christophe Marquesuzaà [1], Patrick Etcheverry[1], and Julien Lesbegueries[2]

IDEE Research group (Interaction, Electronic Document and Education)
http://idee.iutbayonne.univ-pau.fr
Computer Sciences Laboratory of the "Université de Pau et des Pays de l'Adour"
http://liuppa.univ-pau.fr
[1] IUT de Bayonne – Pays Basque,
Château-Neuf, Place Paul Bert,
64100 Bayonne – France
Christophe.Marquesuzaa@iutbayonne.univ-pau.fr,
Patrick.Etchevrry@iutbayonne.univ-pau.fr
[2] UFR de Sciences et Techniques de Pau, Département Informatique,
Avenue de l'Université,
64000 Pau Cedex BP 576 – France
Julien.Lesbegueries@univ-pau.fr

Abstract. Corpus resocialization can be achieved through digital revitalizing, i.e., making a corpus available for new users with new uses. We propose a semi-automatic identification of spatial and temporal markers that make up metadata for localized documents. Metadata is then exploited by concrete applications in environments for tourism and teaching.

1 Introduction

In all Western European countries, national institutions take initiatives to emphasize their cultural heritage by relying on digital and web technologies. However, the valorization of resource collections with less notoriety is much more difficult because it applies to corpuses of more restricted interest or more geographically localized. Nevertheless, this field requires to be developed because the heritage available in these spaces is as rich as varied and potential for its valorization remains undeniable.

The wealth of localized documents we focus on suffers not to be sufficiently shared because of its inaccessibility. These document collections generally rest in the depth of archives, museums and libraries and live only for a small number of specialists who know of their existence.

Our research work focuses on the specificities of localized documents and the ways to re-socialize them for an increased number of people. This does not "simply" consists in digitizing and redistributing our document collections. Resocializing consists not only of putting these inaccessible documents at the users' disposal but also of thinking the way to facilitate their content appropriation for specific users, with new usages.

Problems tackled in this paper are of two kinds. From one hand, we focus on the investigation and the classification of a massive and varied document set, on the other hand, the methods of transmission of their contents towards new users.

M.A. Rodríguez et al. (Eds.): GeoS 2005, LNCS 3799, pp. 153 – 165, 2005.

We will first present the characteristics of localized documents (paragraph 2), and will focus on the methods to make them available. Because the re-socialization of these documents is achieved through their suiting to new public and new uses [1], we discuss the means of exploiting a set of spatial and temporal metadata to express the documentary contents in a relevant way (paragraph 3). After presenting the principles for locating spatial segments (paragraph 3.1), we explain the way to exploit them. A first research work enabled us to identify three publics (scientists, tourists, teachers/learners) to which we associate particular uses of the corpus. Then, we present a specific way of using our spatial and temporal metadata to use and revitalize localized documents in tourist (paragraph 3.2) and academic contexts (paragraph 3.3).

Then, in order to illustrate this work, we describe the practical aspects and the technical tools used to validate each one of our proposals: a tool for generating spatial and temporal metadata (paragraph 4.1), a possible use of the metadata to revitalize localized documents thanks to applications for tourism (paragraph 4.2) and teaching (paragraph 4.3).

We conclude this contribution by presenting the limits of our proposals and the further research envisaged.

2 Localized Documents

Our interest is related to localized documents left out at the bottom of archives, museums and libraries. Our work aims at defining tools allowing to re-socialize these lost treasures to make them accessible to a larger public again [1].

We think that “rough and massive” digitalization and broadcasting of these document collections do not constitute in itself a satisfactory solution to facilitate access to their contents. It must be coupled with some thinking on the means of transmitting the contents of the documents to the potentially interested users [2]. This “adapted” transmission of the contents presupposes a structured organization of the document collections in order to be able to locate within the bulk of documents (or extracts of documents) those considered to be relevant for the user. However the variety and the quantity of the considered document collections stresses the issue of document exploration and classification.

Therefore, we first focus on the design of tools allowing to improve the study, the characterization and classification of the document collections to revitalize. Our localized documents are characterized by contents strongly attached to a territory and its history. This property manifests itself in the omnipresence of place names and of historical events defining the concerned territory. Whether they are newspapers, popular or erudite literature and so on, documents abound of notations and spatial and temporal references.

This assumption leads us to consider the criteria of time and space as significant entrance points to explore localized document collections [3]. Therefore, the constant presence of space and time markers leads us to imagine tools able to consider a corpus and to automatically or semi-automatically outline all the places and historical events mentioned within each document [4].

3 Document Revitalization: Defining and Using Spatial and Temporal Markers

As specified in the preceding paragraph, we focus on spatial and temporal references within the concerned documents. We then define tools allowing a non-expert user (teacher/learner, tourist, general public) to access to digitized corpus. Moreover, these tools must be not only easy to use and available in familiar digital environments (web sites, interactive terminals, virtual museums, etc.) but also associated with the usual activities of these publics. We will expose hereafter the method employed to detect spatial segments and then define the corresponding markers. The way of exploiting these spatial markers is then explained within two scopes related, on the one hand, to general public (tourist-oriented) and, on the other hand, to the educational environment (teaching-oriented).

3.1 Semantic Treatment of the Document Contents

Considering a restricted corpus allows one to implement more sensitive scans that take into account the document contents. Our contribution is a non exclusive alternative to traditional search methods used in libraries and based on descriptive forms (metadata). We aim at considering in a more accurately way the user request and at increasing the results relevance, returning for example a document extract instead of a whole document.

The document contents are taken into account by a specific semantic process which exploits the localized property of the corpus and focuses on spatial and temporal entities. The data processing sequence used to highlight spatial and temporal markers is composed of four main steps [5]:

- the lemmatization carries out a segmentation of the words,
- the lexical and morphological analysis proceeds to a word recognition,
- the syntactic analysis, based on grammars, allows to find the bonds between words,
- the “semantic” analysis carries out a more specific analysis allowing the extracted syntagms to be interpreted.

More precisely, the data processing sequence used to detect spatial entities is implemented as follows:

1. currently, the text “tokenisation” corresponds to a simple lemmatization phase. It would be nevertheless interesting, in a further work, to retrieve the document structure, i.e. paragraphs, titles, etc.
2. the detection of spatial entities that are potentially georeferencables (words with capital letters, words belonging to a lexicon) is carried out at the same time than a morphosyntactical analysis which allows to retrieve words type.
3. a grammar-based analysis, allowing the interpretation of the extracted syntagms (adjacency, distance to another spatial entity, etc.), is then carried out at the same time of a georeferencing process applied on the spatial entities previously detected. This georeferencing process is based on services close to the gazetteers (services available on the web and allowing named entities to be georeferenced).

This data processing sequence allows each spatial entity, previously extracted, to be interpreted. Figure 1 highlights an indirect entity “la lisière du bois de Zouhoure” (“the edge of the Zouhoure wood”) defined by a direct entity (Zouhoure) and a topological relation of “adjacency” type (“the edge of ”).

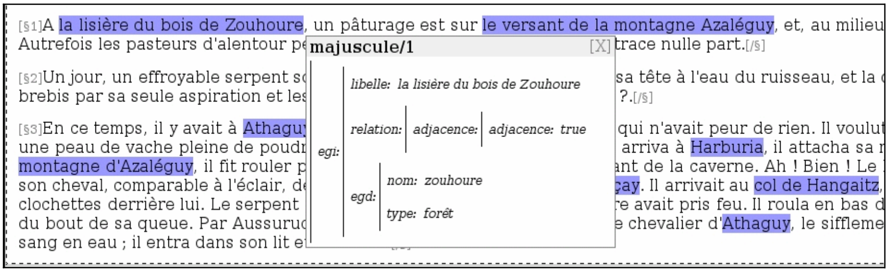


Fig. 1. Interpretation of the spatial segment « la lisière du bois de Zouhoure » (“the edge of the Zouhoure wood”)

From a technical point of view, the marking tool architecture relies on web services which re-use and adapt modules of the Linguastream platform[1] developed to solve similar problems in the field of the text analysis. Using the XML technology, we produce an intra-documentary marking: the initial document is enriched with markers highlighting spatial and temporal text segments.

Developed by the GREYC research group[2] since 2001, LinguaStream is a generic platform used for automatic natural language processing. LinguaStream is based on the concept of data processing sequence. In a data processing sequence the results produced by an upstream chain link are exploited by a downstream link.

LinguaStream allows the design and the evaluation of complex data processing sequences, by assembling various analysis modules: morphological, syntactic, semantic, discursive or statistics. Thus, each stage of the data processing sequence results in the discovery and the marking of new information, which subsequent analyzers will be able to base upon. At the end of the chain, various tools allow to display the analyzed documents and their associated marks. The advantage of this platform lies in the fact that all the modules of the data processing sequence accept an XML flow as entry and produce an output XML flow.

Thanks to a simple and powerful GUI (Figure 2), the user builds his/her specific processing link while linking the different modules that he/she deems necessary for his/her analysis. It is also possible to shunt the XML flow into a file in order to trace the output of a given module.

In order to detect geographical entities in our document collections with the Linguastream platform, we have built a specific language processing link which is made up of several macro phases [6]. This detection system is not based on a fixed model or

[1] www.linguastream.org
[2] www.greyc.unicaen.fr

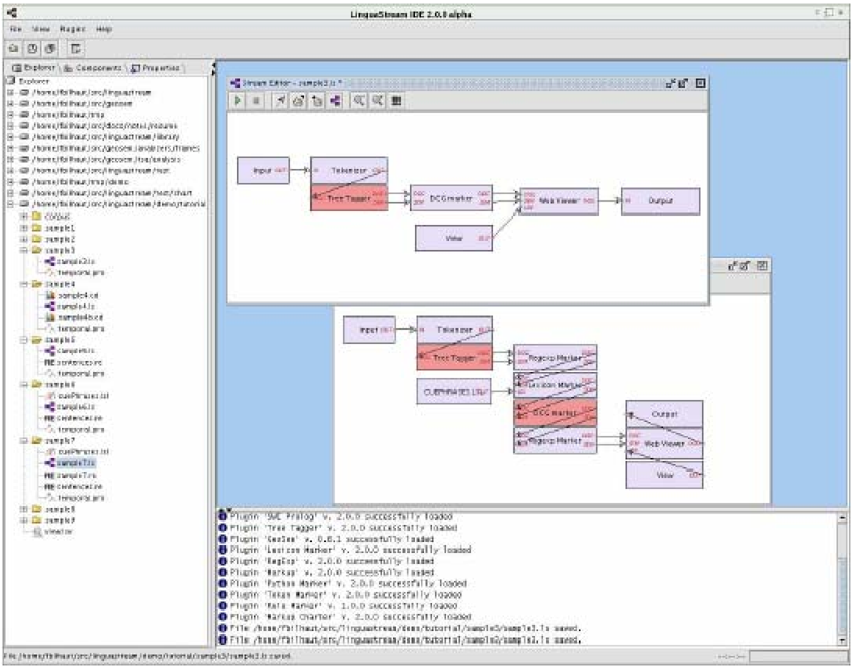


Fig. 2. The Linguastream platform

an ontology but on a set of heuristic rules. Here is an example of Prolog rules integrated in our prototype, allowing to detect adjacent relations between geographical entities.

```
root (X) --> ega (X).

ega ( entite_geo : X ) -->eg (X).

eg ( egi : X ) --> egi (X).

eg ( egd : X) --> egd (X).

egi ( libelle : libelle .. relation : X .. egi : EGI ) -
-> relation (X) , egi (EGI).

egi ( libelle : libelle .. relation : X .. egd : EGD) --
> relation (X) , egd (EGD).

relation ( adjacence : X ) --> adjacence (X).

adjacence ( type_adj : proche ) --> ls_token ( 'proche' ).

adjacence ( type_adj : proche ) --> ls_token ( 'près' ).

adjacence ( type_adj : proche ) --> ls_token ( 'à' ) ,
ls_token ( 'de' ).
```

```
egd (X) --> prep , lexique (X) .
%Preposition
prep --> ls_token ( 'de' ) .
prep --> ls_token ( 'd\'' ).
% Candidate : Direct Entity
lexique ( libelle : N .. type : commune ) --> N @ eg :
egd .. egn : oui.
```

More precise details concerning this marking technique are available in [6] and [7].

3.2 Revitalizing Documents in a Tourist Context : Principles

In a tourist context, the user mainly needs to explore, discover and be oriented in a territory. Our contribution consists in proposing the discovery of a territory via a navigation based on spatial (and temporal) markers identified as metadata within the corpus. The territory visit is built through an itinerary among all the available marked documents.

Spatial and temporal markers available in the corpus will then be exploited by the navigation system in a transparent way for the tourist. For example, the selection of the documents relevant to himself is realized according to the route defined on a map. Markers (place names, house names, position, date, times, etc.) defining the route become the selection criteria of significant documents in our corpus.

Thus, the territory discovery is performed according to a spatio-documentary navigation which consists in defining the possible interactions and dependences between two interdependent spaces:

- the navigation in a geographical space is achieved by (re)defining the route on the map and thus updating the corresponding consultable documents that deal with the route,
- the navigation in a document space is not only carried out by consulting the proposed documents but also by following the intra-documentary hyperlinks which can possibly refer to new places. This may entail the wish of a change of itinerary.

The success of tools based on these principles depends more on their interactional qualities, on their ease of appropriation and on their adequacy to a target public expectations than on the exhaustiveness of the markers located inside the corpus. These tools can then be described like scenarized environments exploring the marked resources. “A scenario describes and organizes the communicational structures that take part in an activity, in accordance with an intention. The scenario describes the activities that support interactions.” [8]

Within this scope, we are developing, in partnership with a local inter-district media library (MIDR in Pau - Médiathèque Intercommunale à Dimension Régionale) a project called “*Virtual Routes in the Pyrenean heritage*” (« *Itinéraires virtuels dans le patrimoine Pyrénéen* »). This project focuses the needs from both a cultural and tourist point of view. It appears as a cartographic visualization of a route (determined by the user) allowing to structurally display the result of the information retrieval. Our prototype uses a loosely coupled architecture based on web services [7]. It notably supports the execution of geographical requests on georeferenced data, information retrieval, result representation; some web services are already available on the web. The next paragraph presents the principles related to the utilization by and for the educational environment.

3.3 Revitalizing Documents in a Teaching/Learning Context: Principles

In the current technological and economic context, the issue of educational resources acquires an ever increasing importance. This remains true with regard to their development, their indexing and standardization. [9] defines some guidelines of relevant concern in his research in Human Learning Interactive Environments. Documents should be considered as very significant elements in Problem-Based Learning Situation (PBLS) because the learner faces problems in which the documents can be seen as suggestions for potential solutions and thus constituting *a cognitive resource, a vector or a proof of learning* [10].

However, according to [11], the current state of Information and Communication Technologies does not seem to be able to provide simple computing tools for processing, indexing and capturing the information contained in the documents. Therefore, to overcome these difficulties, we have chosen to limit our research work to documents including spatial and temporal information. For such documents, we wish to provide the pedagogues-designers data-processing tools allowing them to semantically mark the contents of documents used for modeling PBLS.

From spatial and temporal markers arranged in a semi-automatic way in documents in relation to a given PBLS, our research work [12][3] consists in developing cognitive tools for various users. Therefore, the pedagogues-designers need to critically examine and specialize existing tools dedicated for production, coordination and co-operation. Actually, in the utilization phase, these tools will be used by the learners during their activities for which they handle extracts of the documents provided by the designers. The interaction processes between the available documents and the learning activities (and conversely) will lead the pedagogues-designers to both prototype and finalize the documents to didactize, and the activities suggested by the PBLS.

SMASH [13] is a co-operative PBLS designed [14] to check and to implement our research projects. This case is designed for 10-12 years children and the underlying learning objectives relate to road safety. We have chosen this specific co-operative PBLS because it implies documents including both spatial and temporal information. In texts resulting from witnesses' accounts, we used spatial and temporal information for teaching purposes. In the next paragraph, you will find screenshots relating to the work in progress.

4 First Results

We hereafter present the technique and the result of spatial and temporal documents marking. We then detail the use of the marked documents within two contexts for two different uses.

---

[3] This work is developed in our IDEE research group – http://idee.iutbayonne.univ-pau.fr.

4.1 Marking of Spatial References

The text below is an extract of a legend described in [15] and is used as support in an article published in Pays Basque Magazine, a tourist magazine intended to promote the Basque Country [16]. Figure 3 highlights a manual spatial marking:

– we bolded place's names, house's names and people's names belonging to a real territory,
– we underlined the spatial references in the considered document.

> En ce temps il y avait à **Athaguy** un chevalier, cadet de cette maison, qui n'avait peur. Il voulut savoir s'il serait maître du dragon. Il met une peau de vache pleine de poudre sur sa monture et il va. Quand il arriva à **Harburia**, il attache sa monture à une aubépine. De la crête de la montagne d'**Azaléguy**, il fait rouler par bonds et par sauts la peau au-devant de la caverne. Ah ! bien ! ! Le bon Dieu lui avait donné l'agilité. Il monte son cheval, comparable à l'éclair, descend le vallon, et se tourne vers **Alçay**. Il arrivait au col de **Hangaitz**, lorsqu'il entend comme un bruit de cent clochettes derrière lui. Le dragon ayant avalé la peau de vache, la poudre avait pris feu. Il roule en bas du bois d'**Ithe** fracassant les jeunes hêtres du bout de sa queue. Par **Aussurucq** il arriva à la mer et s'y noya. Pour **le chevalier d'Athaguy**, le sifflement du dragon et le bruit convertirent son sang en eau ; il entra dans son lit et mourut.

Fig. 3. Manual spatial marking

To compare, Figure 4 presents an automatic spatial marking on the preceding text (almost equivalent) carried out by our tool [7].

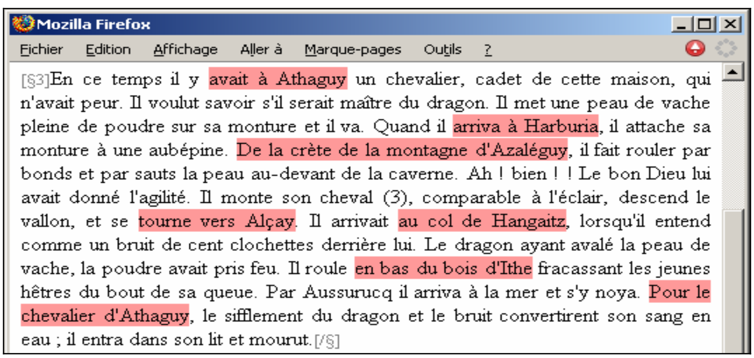


Fig. 4. Automatic spatial marking

The automatic text marking techniques for segments related to space are not of the same complexity level. A first set of tools allows to automatically locate the terms in bold representing toponyms and house names, such as "Athaguy" or "Azaléguy". Spotting them is based on typographical scanning (here all the words with a capital letter) and morphological (here, all the names). This locating phase is only based on knowledge rules. Results could be improved by using for example a geographical and linguistic atlas [17].

This first phase of recognition is then supplemented by the use of tools able to identify, by extension, the terms or expressions such as “coming to the pass of” (“*au col de*”) or “down the wood of” (“*en bas du bois de*”), or to interpret more complex segments such as “the peak of the mountain of Azaléguy” (“*la crète de la montagne Azaléguy*”).

However we observe, a difference with manual marking if we consider spatial references (like “in front of”, “nowhere”, etc). Work remains to be done in order to integrate rules able to process this type of spatial references.

Temporal analysis is carried out according to the same principles. The “stopover points” are then dates, operators defining intervals (“from X to Y”, “between X and Y”, “years X”) and new classes of operators, such as (“the beginning of X”, “around X”, etc).

Spatial and temporal documents marking constitutes a first step aiming at facilitating the following phase because these marks are exploited to re-socialize the document collections by conveying relevant extracts to the users. The two following paragraphs present two projects using the tool for spatial and temporal marking on document collections related to the Basque Country.

4.2 Revitalizing Documents in a Tourist Context: Application

We have imagined a first case study that consists of proposing a tourist to discover the Basque Country through a virtual visit. The user defines a route on a map and the system then proposes a set of documents in which the places on the itinerary are mentioned. Figure 5 illustrates the discovery of the area where the Cerquand tale presented in paragraph 4.1. is located.

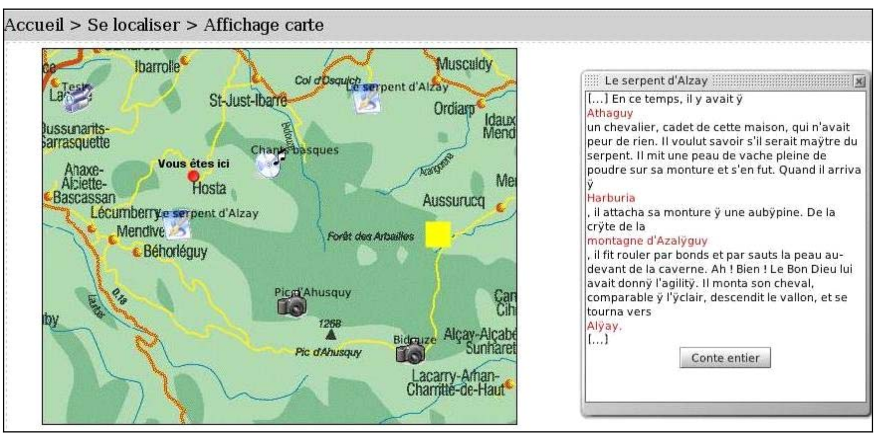


Fig. 5. Re-socialization of localized documents for a tourist use

In the current prototype, the route is simply defined by a starting point and a point of arrival (or a displacement range) that the user selects by pointing on the map. Using a graph representing the possible ways to go from a place to another, the system calculates a possible route. During this process, it indexes the names of the places defining the stages on the route and queries the pre-marked documentation base in order to extract the documents likely to interest the tourist. The map displays, next to each stage place, icons that refer to the documents in which the place name appears. The possibilities of spatio-document navigation are then implemented according to principles described in paragraph 3.2.

4.3 Revitalizing Documents in a Learning/Teaching Context: Application

This paragraph stresses the importance of the concepts of space and time in the scope of teaching activities. We can easily highlight the interest of spatially and temporally marked documents to teach disciplines such as geography or history. However, we remain convinced that the exploitation of spatial and temporal markers may extend to other subject areas.

The first example we will mention relates to the Smash case presented in paragraph 3.3. This example does not envisage the problems of localized documents revitalization but illustrates a way of exploiting the spatial and temporal references in a teaching activity which is not related to the teaching of history or geography. In the Smash situation, learners have to an accident scene from witnesses' accounts in order to understand what occurred (and thus detect/learn the associated road safety requirements). The accident reconstitution phase consists in positioning the various protagonists of the history (cyclist, witnesses, and traffic signs) on the map representing the place of the accident. Learners carry out this work while reading and interpreting the various accounts available. To help them in this activity, we used our spatial marking tool to provide them (if they face learning obstacles), a second version of accounts where the spatial references were highlighted. Figure 6 presents an extract of spatial marking for a testimony ("*I had just left the store and was ready to cross the pedestrian walkway situated on my way to go to the Brasier road*.").

"Je venais juste de sortir du magasin et je me tenais prêt à traverser au passage piéton qui se situe sur mon chemin pour aller route du Brasier."

<spatial_entity id= "E8G8"> au <name_places id= "E8F16" type = "element_de" localisation= "non" position= "internal" > passage piéton </name_places> </spatial_entity>
....
<spatial_entity id= "E8G14"> <name_places id= "E8F26" type = "place" localisation= "non"> route </name_places> du <name_places type = "nom_propre" localisation= "oui" position= "internal"> Brasier </name_places> </spatial_entity>

Fig. 6. Spatial marking in the Smash application

Our second example focuses on the revitalization of localized documents in educational applications. This project, still in progress, aims at exploiting our local document collections (*cf.* paragraph 4.2) in order to familiarize learners with the signs allowing to estimate the potential evolution of a village from different point of views (demographic, economic, spatial). This PBL teaching application, will consist in studying shared documents (population registers, land property record, demographic evolution tables, etc.) describing two villages that have evolved differently throughout time (expansion, disappearance). From this study, learners will have to deduce some indicators (population, trading activity, etc.) allowing to estimate if a given village is about to extend or, on the contrary, to decay and for which reasons.

From a technical point of view, the application will also be based on the Smash architecture in order to establish the tools underlying a PBL: generic whiteboard implemented on Lazlo application server[4] coupled to an Open Distance Learning platform such as OpenUss[5], capable of using the information system. Spatial and temporal marking tools will be available for:

– the pedagogue, to extract from the documentary collections a set of documents considered as relevant for learning,
– the learners, to search for document extracts, mentioning the relevant indicators at various times and/or on various places, relying on spatial and temporal criteria.

This project is only at a design stage and aims to validate the assumption that the concepts of space and time are omnipresent in many learning/teaching situations.

5 Conclusions and Perspectives

In this paper, we presented a research work focused on the revitalization of localized documents. In a first stage, we proposed to use the specificity of these documents in order to develop tools allowing a set of spatial and temporal metadata to be generated. This metadata can be considered as criteria which can be used to propose a first classification of all the documents.

In a second step we stated the principles exploiting this metadata in order to socialize these documents according to various uses: socialization for a tourist purpose and socialization for a teaching purpose, each one being illustrated by concrete projects.

These research tasks are currently in progress and thus require further investigations and many more experiments. Indeed, the results obtained at this stage entail us to support the assumption that spatial and temporal markers represent a significant entrance point for the exploration of localized documents. However, we keep in mind the underlying difficulties of the marking stage and therefore the effort that remain to be done to approach the quality of manual marking. Subsequent work on this aspect will consist to enrich the current marking rules on the one hand and, on the other hand, to find new ways to exploit these marks in order to favor the broadcasting of the document contents for new users.

Currently, our tools concern only text resources. However, localized document collections consist of extremely varied documents: postcards, photographs, video and sound extracts, etc. These other kinds of documents must also be taken into account, in a long-term project whose aim is to revitalize a cultural heritage.

In the same way, the proposed tools were only tested on space/time markers. In the longer term, research tasks will consist in developing a reflection on the feasibility to find and exploit new kinds of markers.

---

[4] www.openlaszlo.org
[5] http://openuss.sourceforge.net/openuss/index.html

Acknowledgments

"*Virtual Routes in the Pyrenean heritage*" project is led in partnership with the community of agglomeration of Pau and the multimedia library of Pau (MIDR). We want to thank them for their help (by the provision of the digitized corpus, notably) and their support.

Moreover this project is shared in the whole research team IDEE. That's why we want to thank its members for works performed together.

References

1. Casenave, J., Marquesuzaà, C., Dagorret, P., Gaio, M. La revitalisation numérique du patrimoine littéraire territorialisé, *Colloque ENSSIB-EBSI*, 13 - 15 octobre, Montréal, Québec, Canada, to be published, 2004.
2. Sédes, F., Calabretto, S., Lallich-Boidin, G. Les temps du document numérique, *Colloque ENSSIB-EBSI*, 13 - 15 octobre, Montréal, Québec, Canada, to be published, 2004.
3. Etcheverry, P., Marquesuzaà, C., Lesbegueries, J. Revitalisation de documents territorialisés : Principes, outils et premiers résultats, Atelier MetSI, *Congrès INFORSID*, 24 - 27 avril 2005, Grenoble, France, to be published, 2005.
4. Egenhofer, M.J. Towards the semantic geospatial web, GIS'02, *Proceedings of the 10th ACM international symposium on Advances in geographic information systems*, ACM Press, pp. 1-4, 2002.
5. Abolhassani, M., Fuhr, N., Gövert, N. Information Extraction and Automatic Markup for XML documents. *Intelligent Search on XML Data - Applications, Languages, Models, Implementations*, and Benchmarks (Lecture Notes in Computer Science Vol.2818), Ed. by Henk M. Blanken, T. Grabs, Hans-Jörg Schek *et al.*, Published Springler, Berlin, pp. 159-174, 2003.
6. Loustau, P. Traitement sémantique de documents dans leur composante spatiale, Application au Patrimoine Pyrénéens, Rapport de Master Recherche TI, Laboratoire Informatique de l'Université de Pau et des Pays de l'Adour, LIUPPA, Equipe Idée, 2005.
7. Lesbegueries, J. Des services web destinés à l'indexation de la recherche spatio-temporelle dans un corpus territorialisé, Rapport interne, Laboratoire Informatique de l'Université de Pau et des Pays de l'Adour, LIUPPA, Equipe Idée, 2005.
8. Latapy, M., Lopistéguy, P., Dagorret, P. Genre Based Design as a Response to User Expectations, Designer, User, Meaning, Maker: Rethinking Relationships for a More Creative *HCI - International Workshop at HCI 2004*, 18th British HCI Group Annual Conference, ISBN : 0-904354-90-3, 2004.
9. Tchounikine, P., Baker, M., Balacheff, N., Baron, M., Derycke, A., Guin, D., Nicaud, J.-F., Rabardel, P. Platon-1 : quelques dimensions pour l'analyse des travaux de recherche en conception d'EIAH, Synthèse des travaux menés par un groupe pluridisciplinaire dans le cadre de l'action spécifique "Fondements théoriques et méthodologiques de la conception d'EIAH du département STIC du CNRS, AS 51 du RTP 39, 2004.
10. Laflaquière, J. Cognition située et application aux espaces documentaires, Sciences Cognitives, Mémoire de Maîtrise, Université Victor Segalen, Bordeaux II, pp. 28, 2002.
11. Azouaou, F.l, Weiqin, C., Desmoulins, C. Semantic annotation Tools for Learning Material, SW-EL'04: Semantic Web for E-Learning. *Applications of Semantic Web Technologies for Educational Adaptive Hypermedia*, Eindhoven, The Netherlands, 2004.

12. Nodenot, T., Gaio, M., Bessagnet, M.-N., Sallaberry, C. Creating rich collaborative learning scenarios: a model driven approach for contextualizing software components and electronic documents. *8th World IFIP Conference on Computers in Education, WCCE*, Univ. of Stellenbosch, Cape Town, South Africa, to be published, 2005.
13. ACTIS, Smash Scenario, Project boxes, www.projectboxes.co.uk/catalogue/scenarios.html#smash, 2001.
14. Oudot, C. Modélisation UML d'une situation d'apprentissage cooperative, Rapport de stage du DESS « Sciences Cognitives Appliquées », Université de Bordeaux 2, 2003.
15. Cerquand, J.-F. Le chevalier de Çaro et le dragon, "*Légendes et récits populaires du Pays Basque*", Récité par Marianne Etchebarne, transcrit par M. Basterreix, dialecte souletin, orthographe de la langue basque conforme à l'édition originale des traductions parue en 1872 dans le *Bulletin des Sciences, Lettres et Arts de Pau*, pp. 211-213, 1872.
16. Laxalt, T. Forêt des Arbailles - le souffle de la légénde, Pays Basque Magazine, numéro 32, 4° trimestre, 40-45, 2003.
17. Videgain, C. Atlas linguistique de la langue basque, sous l'égide l'Académie de la Langue Basque (to be published)

Ontology Matching for Spatial Data Retrieval from Internet Portals

Hartwig H. Hochmair

St. Cloud State University, Department of Geography,
720 Fourth Avenue South,
St. Cloud, MN 56301, USA
hhhochmair@stcloudstate.edu

Abstract. This paper analyses reported difficulties in spatial data retrieval from Internet portals, based on results from an empirical survey. The analysis reveals the problems which cause user dissatisfaction and failure in the search process for spatial data sets. These problems are addressed in a new data search architecture. Within the architecture, the paper focuses on query expansion, which helps to overcome mismatches between the user's taxonomy of geographic features and the taxonomies in the knowledge base of the metadata search engine. The proposed data search architecture is formalized in an algebraic specification language, followed by the simulation of a search scenario for spatial data. The simulation demonstrates the advantage of query expansion techniques over simple keyword matching.

1 Introduction

An Internet portal is a single entry point for accessing information on the World Wide Web. A GIS clearinghouse, such as the FGDC National Geospatial Data Clearinghouse[1], or the Alexandria Digital Library[2], is an Internet portal that is able to search a network of (meta)-data servers that host digital geographic data primarily for geographic information systems. Geodata warehouses hold data from various data providers at one server, and offer filter functionality over various metadata attributes, such as boundary coordinates. For successful data retrieval, metadata need to be provided for published datasets. Metadata are data about data, describing characteristics of the data sets and the data provider. For geographic data, the two most important metadata standards are the Content Standard for Digital Geospatial Metadata (CSDGM) version 2 (FGDC-STD-001-1998), which is a US standard, and the ISO/TC 211 19115-2003 standard.

Although data search engines allow to search after numerous criteria, they lack intelligent reasoners. A lack of understanding of the semantics of the user's entered search terms, as well as a lack of spatial inference mechanisms, causes the search results to often be too narrow or too large. Recent developments, such as the Semantic

[1] http://clearinghouse3.fgdc.gov/
[2] http://webclient.alexandria.ucsb.edu/mw/index.jsp

M.A. Rodríguez et al. (Eds.): GeoS 2005, LNCS 3799, pp. 166 – 182, 2005.

Web (Berners-Lee et al. 2001) or place-footprints (Brandt et al. 1999; Jones et al. 2003) use (geo)ontologies in their platform, which provides a basis for automated reasoning during the Web crawling for spatial data. Various domain specific data warehouses offer powerful thesauri as basis for their keyword based search.

The goal of this work is to present a conceptual framework which addresses various problems in geodata retrieval from internet data sources. There is an extensive literature regarding Web search on non-geographic information (Hölscher and Strube 2000; Pirolli 2002). We extend this work with a survey about user satisfaction with searching and downloading GIS data from geodata warehouses on the Internet. Based on the study, we classify observed shortcomings, describe their impact on data search failure, and provide a set of guidelines to overcome the reported problems with a more intelligent user interface. We focus, however, on the large group of reported keyword matching problems. The assumption of this paper is that a server sided knowledge base (KB) with a well-designed domain ontology, combined with a set of query expansion rules, can increase the number of useable datasets retrieved from geodata servers as compared to simple keyword matching.

The mapping between the ISO-19115 and FGDC-STD-001-1998 metadata standards is necessary for automated crawling of several metadata servers. The Digital Geographic Information Working Group[3] provides mappings in Microsoft Excel format, i.e., not in machine-readable format. The OWL metadata ontologies for FGDC-STD-001-1998 and ISO-19115 are available online[4]. Bermudez and Piasecki (2004) propose the use of the Ontology Web Language (OWL) (W3C 2004) for translating the equivalence of attributes between two metadata standards in machine-readable form. The equivalence of attributes can be expressed with *owl:equivalentProperty*, as shown below for retrieving the cost of acquiring data sets.

```
<rdf:Description rdf:about="&iso:fees">
  <owl:equivalentProperty rdf:resource="&fgdc:Fees">
</rdf:Description>
```

The remainder of the paper is structured as follows: Section 2 summarizes findings of the empirical survey on geodata search in the Internet. Section 3 describes the role of ontologies and taxonomies in a geo-data search. Section 4 presents the new conceptual framework for spatial data search. The formalization of the framework is given in section 5, which is followed by a simulation in section 6. Section 7 summarizes the results and presents directions for future work.

2 Assessing User Satisfaction: Results of a Survey

This survey examined experiences that participants made when searching for GIS datasets in the Internet. The 34 participants were undergraduates from Saint Cloud State University in an introductory GIS course who participated in the study to receive partial course credit. Participants noted their search experiences in questionnaires during the data search (part I) and after the data search (part II).

---

[3] http://metadata.dgiwg.org/

[4] http://loki.cae.drexel.edu/~wbs/ontology/fgdc-csdgm.htm

2.1 Experimental Design – Task and Observations

The task given to the participants consisted of a scenario that required a data search on four different geodata sources, namely three geodata warehouses[5] and one geodata metaserver[6]. These sources were chosen as they provide different query methods (e.g., pulldown menus, graphical selection, browsing). Thus, they should reveal a great range of user interface problems concerning geo-data search.

In the scenario, each participant was asked to imagine being an employee at a federal planning institution who is responsible for locating potential spots for a new bird reservation on fresh-water water bodies. He or she would therefore be in charge of finding a dataset that helps making a decision according to the scenario. No time limit was set for the search process.

For each of the four Internet portals visited, each participant filled in some statistics on the questionnaire concerning the search process, including familiarity with the Web site, success or failure of the search attempt, frequency of misdirection during the search, and search time. Further observed problems could be written down in free text. This information was collected during the data search (part I). Further the questionnaire asked for participants' subjective importance ratings of seven pre-defined problems (q1-q7) after completion of all four search attempts (part II). An additional rating was asked on how frustrating the data search was on average for all four Web sites (q8). Problems to be rated (q1-q7) referred to the user interface, and were kept general and not specifically tied to spatial data search (slightly modified):

1. Too many filter parameters asked at a time
2. Meaning of provided keywords to select from is unclear
3. Filter criteria are too static and cannot be refined during the query
4. Relevant filter attributes are missing
5. Too little information is provided on how to fill in interactive forms
6. Difficult to estimate how many filter criteria to be utilized
7. Hard to predict how the change or selection of a filter affects the result

2.2 Results

Each of the given Web sites hosted the data sets required in the scenario (i.e., water bodies with fresh-water such as lakes or rivers). However, the data sets were not always identifiable as such in the user interface, which caused problems in the data retrieval. According to the completed questionnaires, in 43.5 % of the trials the requested data could be found and downloaded, in 26.7% the data could be found or visualized but not downloaded, and in 29.8% of the trials the search failed completely. Most users were unfamiliar with the Web sites (mean = 1.49 on a scale from 0 to 10, 0…not familiar at all, 10…very familiar). For 74.4% of the successful attempts, participants stated that they “got lost” at least once.

---

[5] http://gislab.fiu.edu/dynamic/searchmeta.cfm?tid=7
http://www.nationalatlas.gov/atlasftp.html
http://data.geocomm.com/catalog/

[6] http://www.geographynetwork.com/

Average importance ratings of problems q1-q7 ranged from 4.29 (q1: too many filters) to 6.47 (q5: problems with filling in interactive forms) on the importance scale from 0 to 10 (**Fig. 1**a). The average value for frustration (q8) was found to be 6.44.

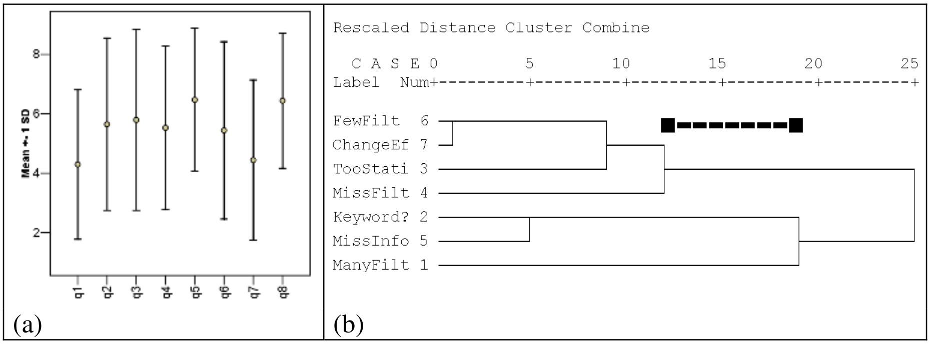


Fig. 1. Rating the problems with data retrieval from the Internet (a); dendrogram showing clustering of seven posted questions using Average Linkage, Within Group method (b)

We grouped the problems q1-q7 based on correlations of individual ratings. A high correlation between two problems means that these problems may be caused by an underlying, more general problem. We applied a hierarchical cluster analysis on the ratings for q1-q7. The gap in the dendrogram (**Fig. 1**b) between four and three clusters (bold bar) indicates that a classification with three clusters is a good solution. The first cluster, C1 (the upper most in the dendrogram), comprises problems related to limited filter functionality, i.e., questions 3, 4, 6, 7. The next cluster (C2) contains problems 2 and 5 and refers to semantic problems, namely unclear keywords, and lack of information concerning the filling in of interactive forms. The third cluster (C3) consists of one problem only, namely of too many filter parameters being requested at a time (q1). Taking the importance ratings from **Fig. 1**a, the average importance of the three clusters are 5.30 for C1, 6.06 for C2, and 4.29 for C3, which indicates that semantic problems (C2) cause most problems. In the free text section, 22.9% of the participants mentioned missing functionality as part of their problem statements. An example is “dropdown menus are too restrictive, type-in fields would be useful”.

Computation of Pearson-correlations between the number of failed search attempts, and participants’ importance ratings of problems q1-q7 revealed that correlations were significant between failed searches and “q1-too many filters” ($r=0.460$, $p<0.01$), and failed searches and “q2-unclear keywords” ($r=0.367$, $p<0.05$).

3 Ontologies and Taxonomies

The problems identified in the previous section are heterogeneous. Each group of them requires specific attention and consideration on the way to more intuitive GIS user interfaces. In this paper we restrict ourselves to semantic problems in data retrieval, which can be ascribed to mismatches between used taxonomies. A key to successful data retrieval is the construction of a powerful KB in the application. The underlying taxonomy defines the system's "organization of the world". Simple keyword based searches often yield zero hits. Contrary, a well-designed KB-taxonomy enables the system to find similar data sets using query expansion.

Ontologies are content theories about sorts of objects and their relations for a given domain of knowledge. An ontology structures knowledge and forms the core of any system of knowledge representation for that domain. The AI communities use ontology in the sense of an explicit specification of a shared conceptualization (Gruber 1993). Ontology provides potential terms for describing knowledge about a domain.

3.1 Shared Features in Taxonomies of Real World Domains

A taxonomy is part of an ontology and sorts concepts (classes) of the ontology in a generalization hierarchy. Although differences exist between ontologies, even for the same domain, general agreement exists between ontologies on many issues (Chandrasekaran et al. 1999). Some commonalities are motivated by the Aristotelian ontology, which is a theory of substances (things or bodies), and instances (qualities, events, processes): There are objects in the world that have attributes which can take values; objects can have parts; there can exist relations between objects; and there are events that occur at different time instants and processes that occur over time. Classes are the focus of most ontologies, as they describe concepts in the domain. Subclasses represent concepts that are more specific than the superclass. Different ontologies exist because of the many ways to subdivide a class into subclasses. Ontologies generally appear as a taxonomic tree of conceptualizations, from very general at the top to increasingly domain-specific further down in the hierarchy. What aspects of reality are adopted for encoding an ontology depends on the task. Such adopted ontology is often called an application ontology (Guarino 1997). Due to different domain tasks, an ontology is unlikely to cover all potential uses. An ontology together with a set of individual instances of classes constitutes a knowledge base. We use the term knowledge base in a slightly different way, namely for a system that contains the taxonomy of feature classes, a set of synonym names, and a set of query expansion rules used for the metadata crawler (no instances are required).

When it comes to building ontologies for spatial data retrieval, Aristotle's idea of *essence* comes into play. For each substance there are one or more qualities, which are inseparable from it. They are called its *essences* or *essential* qualities. One essence of a river is that it is an inland water and that is does not flow within an ocean. The taxonomy underlying a knowledge base should consider the essences of geographic feature classes, as this provides an intuitive classification of entities. Substances may carry other inessential qualities, which are called *accidental*. For example, that a river flows into the Atlantic Ocean is an accidental quality. Accidental qualities may vary between instances of the same feature class (i.e., rivers in this example). They are not helpful in the a-priori categorization of objects in data search: As long as the knowledge base does not "know" whether a spatial data set in the Web carries specific accidental qualities, the query cannot be expanded to these qualities. Accidentals are potentially specified in the metadata sets. Once metadata sets are retrieved, the accidental qualities could be dynamically added to the taxonomy of the KB later on.

Shirky (2005) proposes an alternative idea of organizing information more organically than current categorization schemes, namely using links, which can point to anything. The problem of hierarchization in a data search becomes apparent if the user wants to *browse* a taxonomy, where the ontologist doing the categorization has the responsibility to organize the world. If the user is looking for something that has not been categorized in the way she thinks, this will result in failure. The *search* paradigm (i.e., typing in keywords) is different, as no a-priori classification is offered to the user. Thus a good user interface for geospatial data search should support both searching and browsing, and provide links to related terms within the categorization.

3.2 Sample Taxonomy for Showcase

Entities in the real world can be categorized after numerous criteria, which can be seen when comparing online catalogs of geo-data and thesauri. For example, water bodies may be classified after the salinity of the contained water (i.e., freshwater vs. saline water), or after the geographic location (i.e., inland vs. coastal). The depth of categorization layers may vary. One may, for example, not consider the salinity as relevant classification criterion, and rather list basic categories (Rosch 1978), such as lake, river, or sea as members of the next subclass level.

Although some ontologies and thesauri with their hierarchical structures have been fully implemented in machine readable form, it still needs a manual step to build an application ontology, which is mainly caused by inconsistencies between various taxonomies, a different terminology, and difference data formats (e.g., OWL, RDF, KIF). We use taxonomies from three Web sources for building the sample taxonomy in the knowledge base. Using several taxonomies helps to cover a wider range of relevant feature classes and relations. Any other sources could be used for this task, as long as they deal with geographic concepts and state their taxonomy explicitly.

The first source used is WordNet (Miller 1995), an online lexical reference system for English nouns, verbs, adjectives and adverbs which are organized into synonym sets (called *synsets*). Besides hierarchical relations, WordNet offers also part-of relations, called *meronyms*. WordNet states for example that water is a substance meronym of "body of water" (and not a superclass). The online application[7] gives following subsets (called *hyponyms*) for “bodies of water” (excerpt):

=> backwater -- (a body of water that was created by a flood or tide...)
=> bay -- (an indentation of a shoreline larger than a cove but smaller than a gulf)
=> channel -- (a deep and relatively narrow body of water…
…
=> lake -- (a body of (usually fresh) water surrounded by land)
  => reservoir, artificial lake -- (lake used to store water for community use)
  => pond, pool -- (a small lake; "the pond was too small for sailing")

CERES (2005) is an online information system for environmental data in California. The user interface allows the user to enter a combination of geographic area, theme, and data type for data search. The theme-based thesaurus produced for the keyword “Bodies of water” the following results (excerpt):

[7] http://wordnet.princeton.edu/

NT (narrower terms): bays, lakes and ponds, channels, seas, streams, waterfalls,…
BT (broader terms): Water

Whereas the narrower terms are similar to the categories found by WordNet, the "Broader Term" result reveals that the CERES taxonomy does not distinguish between substances and meronyms: "Water" is being considered as a broader term of "Bodies of water", which differs from the WordNet taxonomy. In WordNet, "thing" is considered as direct superclass (*hypernym*) of "bodies of water". Both taxonomies have in common that they do not explicitly distinguish between saltwater and freshwater water bodies (e.g., lakes). This needs eventually to be manually done for a domain ontology. The top level categories of the CERES taxonomy are natural resources, natural environment, and human environment, the second one including hydrosphere.

The Florida International University Geo-Spatial Database[8] offers the following hydrology-related keywords: hydrographic features, hydrography, lakes, reservoir, rivers, streams, surface water, water bodies, and water boundaries. Although the user interface does not suggest a taxonomy, it provides relevant concepts of this domain.

Among the common elements found in taxonomies (section 3.1), our sample taxonomy for the KB includes objects, essential qualities (i.e., attribute values), and relations. **Fig. 2**a shows the taxonomic tree. Essential qualities appear twice: Once they are shown as a design template attached to the class "water bodies", indicated with dashed lines and underlined text. In the subclasses, the essences with their values assigned to qualities are allusively indicated within brackets. The complete set of essential qualities can be found in **Table 1**. Essences are inherited to subclasses: As a stream, for example, flows inlands, this must also be true for sub classes, such as branches, creeks, and rivers. The reasoner for intelligent query expansion can use the values for essentials. Some feature classes may have several attribute values for essential qualities. A lake, for example, can contain salt water or freshwater. Thus, searches both for salt and freshwater bodies would include search for lakes. For a more general thematic search on hydrographic objects, one may add the essence of "related to hydrography" for the feature class hydrosphere. The list of essentials in Table 1 is not comprehensive but meant to demonstrate the idea behind it.

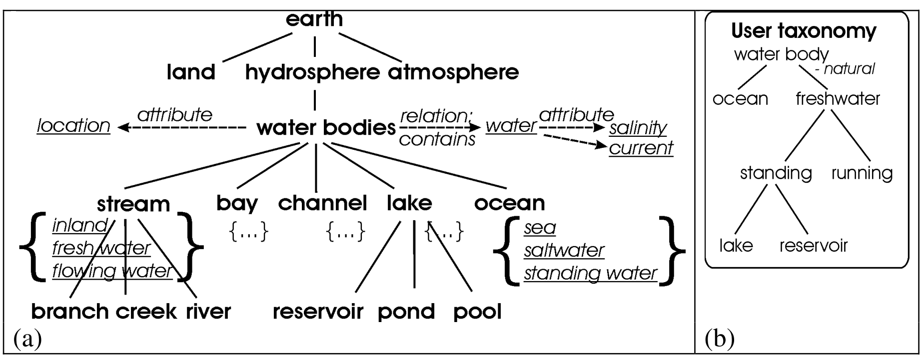


Fig. 2. Geospatial feature taxonomy of the knowledge base (a) and user taxonomy (b)

[8] http://gislab.fiu.edu/

Table 1. Essential qualities of objects in the features class "water bodies"

	Location	Salinity	Current
stream	inland	freshwater	flowing
branch	inland	freshwater	flowing
creek	inland	freshwater	flowing
river	inland	freshwater	flowing
bay	sea	saltwater	standing
channel	inland, sea	freshwater, saltwater	flowing, standing
lake	inland	freshwater, saltwater	standing
reservoir	inland	freshwater, saltwater	standing
pond	inland	freshwater, saltwater	standing
pool	inland	freshwater, saltwater	standing
ocean	sea	saltwater	standing

The user taxonomy enclosing the characteristics of searched data differs from the KB-taxonomy (**Fig. 2**b), namely in the used concept names (e.g., "water body" vs. "water bodies") and in the a-priori subset relations. As opposed to the KB, the fictive user distinguishes between freshwater water bodies and ocean, thus omitting other salt water bodies, such as saltwater lakes.

The user may consider accidental qualities for feature classes, such as *natural* water bodies that are located *inside Florida*. Query expansion for the second request can use knowledge from geographic thesauri, such as TGN (Getty 2005). Accidental qualities can also refer to the data file itself, such as the cost to purchase the dataset. The ideal dataset the user has in mind can be quite complex and refer to various qualities and relations. The desired data set may therefore be described within more than one single taxonomy (Hochmair and Frank 2001).

4 Searching for Geographic Features: Conceptual Framework

This section takes a broader look at data search problems identified in the survey (section 2). We give five guidelines of user interface design for geodata search, which are partly adopted from participants' comments in the survey. The presented conceptual framework of human-computer interaction for spatial data retrieval from Internet portals will take the following guidelines into consideration.

a. Provide immediate feedback on modified search parameters.
b. Provide alternative search options or similar results.
c. Provide an extra window that shows intermediate results.
d. Provide both type-in search and browsing functionality.
e. Avoid shenanigan functions that have no effect or lead to zero hits.

The conceptual framework for (meta)-data retrieval refers to interaction between the user, the knowledge base (KB), and the metadata servers hosting metadata of distributed data sets (DS) (**Fig. 3**). The user formulates a data search request, e.g. by entering a keyword. The server sided KB forwards the request to a metadata crawler, telling which metadata fields should be searched for different metadata standards. If no metadata sets are found in the initial search, the KB utilizes its taxonomies and synsets to expand the query. Once the search is successful, the harvested metadata are stored in the client's memory, and the results are presented to the user. The synsets refer to feature classes ("ocean" ⇔ "sea"), geographic names ("Florida" ⇔ "FL"), and data file types (".shp" ⇔ "ESRI Shapefile"). The query expansion may be combined with a spell checker, similar to Web search engines, such as Google.

An important feature of the proposed concept is that, by storing retrieved metadata sets locally, the search system can immediately hide or deactivate further search options that would give zero hits if applied. Such a system would immediately present relevant statistics about the retrieved metadata sets, such as showing the number of shapefiles among the harvested datasets. It also supports *direct manipulation* (Shneiderman 1983), as the user could predict the consequence of a parameter change, even without restarting a new search. The initial search itself, however, requires intelligent expansion on searched terms to avoid zero hits. This is especially important if the user formulates a free-text query.

The taxonomies in the KB can be stored in any machine-readable format that allows hierarchical reasoning and defining relations between feature classes, such as OWL. See, for example, the geospatial ontologies on the Geosciences Network (GEON) project Web page[9]. A common format for storing metadata is in XML representations of various standards, such as FGDC and ISO. Metadata crawlers can access metadata documents on a metadata server using a client that communicates using standardized protocols, such as ArcXML or Z39.50.

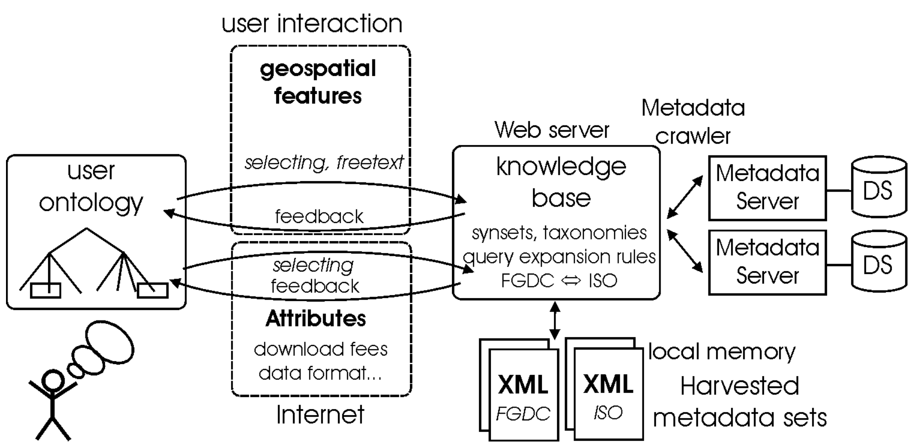


Fig. 3. Conceptual framework for metadata retrieval

4.1 Query Expansion for Keyword Based Data Search

The core of the proposed system is a knowledge base that provides intelligent query expansion. We propose rules of query expansion for geographic feature classes, leaving aside other hierarchical domains, such as data formats or geographic places. We show the steps necessary to retrieve some initial matching data for typed-in terms. The flow diagram in **Fig. 4** visualizes the processing steps within the knowledge base. Terms on the right indicate the functions used in the formalization (section 5).

[9] http://www.geongrid.org/

The user starts by entering one or more keywords that are sent to the knowledge base via the internet. If the search terms cannot be matched with any concept in the KB-taxonomy, the submitted keyword is searched in the set of synonyms (synset) of the knowledge base. If the matching succeeds, the query is expanded with the matched synonym term. Otherwise an error message tells the user that the keyword is unknown. When it comes to query expansion using the KB-taxonomy, we treat four different cases. Each of them is numerated as a separate path in **Fig. 4**.

(1) If the searched concept is a branchnode in the taxonomy graph, query expansion adds all concepts on the subtree of the taxonomy to the query.
(2) If the searched concept is not on a branchnode, and if it is the first search attempt, no changes are made to the searched terms.

If the searched concept is located on a leaf node in the graph (i.e., not on a branchnode), and if it is not the initial search, query expansion must widen the search to the next higher branchnode. Simply looking at broader classes could bring non-relevant results given the broader scope of some classes. Therefore the expansion rule keeps essences when expanding the search. We use two methods for that:

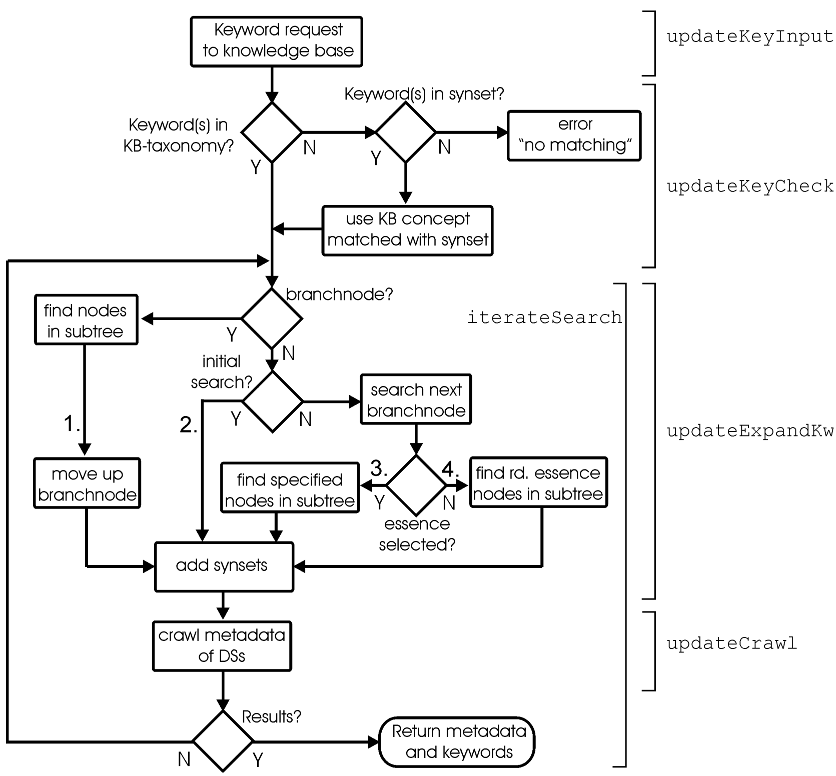


Fig. 4. Schematized processing steps of a free-text query for concepts and essential qualities

(3) If the query phrase itself contains an essence, only features in the sub tree of the shifted branchnode are searched that contain that specified essence.
(4) If no essence has been defined, query expansion selects randomly an essence of the leaf concept, and searches for sister terms containing this essence.

As an example for case (4) let us consider a search for the concept "bay", which is a leaf in the KB-taxonomy (see **Fig. 2**a). Simple broadening of the search to the next higher level would give 11 feature classes, namely all water bodies in the taxonomy. However, if the expansion algorithm randomly selects, for example, the essence "location", which carries the value "sea" for "bay", query expansion finds only the additional concept "ocean", which is more similar to "bay" than for example "river".

Finally, to increase the chance of useable hits in the crawling process, all synonyms of the selected feature(s) are added to the set of searched concepts. If the subsequent metadata crawling process is successful, the metadata are presented to the user. Otherwise the search is iterated. If query expansion was applied, the system might indicate the semantic distance between the matched expanded terms and the original keyword, based on the underlying KB-taxonomy (Rodriguez et al. 1999).

5 Formalization

The conceptual model of query expansion is formalized through algebraic descriptions using classes with their instantiated functions. The result is an executable agent-based computational model. We thereby use *agent*—a concept from Artificial Intelligence—as a conceptual paradigm to represent the human data searcher that interacts with the system components. We implemented the algebraic specifications through the functional programming language Haskell (Thompson 1999). The formal model contains abstractions of the data searching person, the knowledge base, the metadata sets dispersed in the Web, and the processes of interaction. In this paper we show only a few lines of simplified code that refer to keyword based searches, and describe the functionality of the operations and the structure of objects.

5.1 Objects

The complete system (*System*) consists of the simulated agent, the knowledge base, and the metadata sets. The system is abstracted as a product type (using the keyword “data”) that consists of three components. It is created with the type constructor “S”. The simulated person (*HumanAgent*) consists of perceptions (i.e., the expanded keywords and retrieved datasets), the user's taxonomy of geographical feature classes (**Fig. 2**b), a set of synonyms, and a search goal. The knowledge data base consists of the search term entered by the user, the list of expanded search terms, the current branchnode in the taxonomy hierarchy, the taxonomy of the KB (**Fig. 2**a), a set of synonyms, and the retrieved metadata sets. The taxonomies are formalized as undirected graphs, each node consisting of a numerical identifier, the concept term, and quality slots with their values. The metadata sets (*DS*) are formalized as composite type, namely as a list of the datatype *Metadata*. Different metadata standards (i.e., FGDC and ISO-19115) are formalized as a union type, where the union is expressed with the vertical bar ('|') symbol. This data type can be read as disjoint union, which means that each metadata set is either a FGDC or an ISO-19115 type, depending on the constructor function.

```
data System = S HumanAgent KB DS
data HumanAgent = HA Perc UTax Synsets Goal
data KB = KB KeyInp KeyExt BrNode KBTax Synsets Results
type DS = [Metadata]

data Metadata = FGDC IdInfo SpatDataOrg DistrInfo
     | ISO MD_Id MD_SpatRepr MD_Distr
```

Pattern matching provides a tool to define case expressions, i.e., different semantics for different patterns. This method allows the instance of a single function to observe the corresponding components (e.g., theme keywords) from two different metadata types. In the example below, the first line shows the semantics (i.e., the instantiation) of the *themeKW* function for reading the theme keyword from an FGDC metadata set, the second line shows the same for an ISO metadata set. Pattern matching therefore corresponds to the *equivalentProperty* in OWL which was mentioned in section 1.

```
themeKW (FGDC (IdInfo de (KeyW theme place)) s di) = theme
themeKW (ISO (MD_Id a sr (MD_KeyW place theme)) s di) = theme
```

5.2 Operations

A class is used to model the behavior of a data type or a parameterized family of data types. Classes can be used to express *polymorphic* functions (which are equally defined for all data types), and to *overload* functions (which use different definitions at different types). The class header introduces the name of the class and the parameters. The signatures of operations consist of a name, arguments and the result of each operation. An instance describes how to apply operations of a class to a particular data type. Operations in an instance are given in the form of executable equations.

As an example we describe one class used in the formalization. The operations on the highest hierarchical level change the complete system (and not one component only). They correspond to the operations shown in the flow chart in **Fig. 4** (right side), and are members of the class *SystemClass* that takes one parameter. The semantics of each operation is indicated with a square bracket in the flow chart.

```
class SystemClass system where
  updateKeyInput, updateKeyCheck, updateExpandKw,
  updateCrawl, iterateSearch, kwSearch,...:: system -> system
```

We show the instances for two of these functions (underlined) for the data type *System*. Function *iterateSearch* is recursively defined and iterates expanding the keywords and crawling the databases (function *kwExpCrawl*) until the agent perceives a list of retrieved datasets with a length > 0 (i.e., more than one hit). The *kwSearch* function uses function composition, indicated by the dot (.) operator, to combine the five sub functions in **Fig. 4**. The first function (*updateKeyInput*) is located at the end.

```
instance SystemClass System where
iterateSearch s@(S (HA (Perc db kw) utax g ssa ksub) k d)
  |length db > 0 = []
  |otherwise = kwExpCrawl s : iterateSearch (kwExpCrawl s)
kwSearch = iterateSearch.updateCrawl.updateExpandKw
  .updateKeyCheck.updateKeyInput
```

6 Testing the Formal Model

This section uses a small data search scenario, which serves as a demonstration of the executable implementation of the model, and provides the reader with a better understanding of the core functions used in the conceptual model. To keep the showcases simple, the search criterion is restricted to feature class types, excluding geographic area or data type filters. The testing is not empirical. To evaluate the conceptual model empirically, another step would be necessary, namely the implementation into a graphical application environment. Such empirical investigation is, however, considered as part of the future work.

6.1 Creating and Initializing Objects of the System

We simulate a simple case scenario where the agent tries to find metadata of data sets about water bodies using a keyword-based search. We start with the construction of the virtual agent *tom*, and initialize its four components. The agent's perceptions (*perc*) are initialized as two empty lists, its taxonomy is adapted from **Fig. 2**b (*uTax*), the set of synonyms is a list of pre-defined feature class synonym sets (*synFeatAg*), and its goal consists of one concept only (i.e., "water body").

```
tom = HA perc ut syn goal key where
  perc = Perc [] []
  ut = uTax
  syn = synFeatAg
  goal = ["water body"]

synFeatAg= [["water body", "waters", "water bodies",...],...]
```

The initial knowledge base consists of empty features or null values except for the taxonomy (*kbTax*), formalized as graph from **Fig. 2**a, and another list of synsets (*synFeatKB*).

```
knbase = KB kin kext bn tax syn res where
  kin=""
  kext= []
  bn=(0,"")
  tax=kbTax
  syn=synFeatKB
  res=[]
```

A metadata set is formalized as a nested structure of data types, which is adapted from the metadata standard definitions. The use of a union type for a metadata set allows one to assign a different data type structure to each standard. We adapted ten existing metadata sets of Florida[10] as data sources, and implement eight of these sets in FGDC and two in ISO standard. We demonstrate the initialization of one metadata set (*dsF4*) in FGDC standard. Each metadata set contains abstract, theme and place keywords, spatial reference format, object information, file format, and fees. All metadata sets in the system are summarized as a list (*dsList*), which comprises 10 datasets. Finally, the system *initSys* is built from all initialized components (agent, KB, data sets).

[10] The metadata files can be found at http://gislab.fiu.edu/dynamic/searchmeta.cfm?tid=7.

```
dsF4 = FGDC (IdInfo descr keyw)
    (SpatDataOrg dirspatref spaobjinfo) (DistrInfo st) where
  descr ="Polygon coverage of Miami-Dade County water bodies"
  keyw = KeyW themekw placekw
  themekw = ["Water Bodies"]
  placekw = ["Dade County","Florida"]
  dirspatref = "Vector"
  spaobjinfo = PoiVecInfo [(SDTS_Term "G-polygon" 2899)]
  st = StOrder digi fees
  digi = DigiForm (DigiTrans "shp")
  fees = "None"
dsList=[dsF1,dsF2,dsF3,dsF4,dsF5,dsF6,dsF7,dsF8,dsI1,dsI2]
initSys = S tom knbase dsList  -- creating the initial system
```

6.2 Simulation

We demonstrate stepwise the behavior of system operations (specified in sections 4.1 and 5.2) using the initialized system *initSys*. We show how the relevant values of the agent and the knowledge base change by applying system functions, where changed values are underlined. The *updateKeyInput* function copies the requested search term from the agent to the knowledge base, which gives an updated system *sys1*:

```
test input> updateKeyInput initSys  -- copy search term to KB

>> sys1: SYSTEM:
HUMAN AGENT:
PERC:     RETURNED DATASETS:     RETURNED KEYWORDS:
GOAL: “water body”
KNOWLEDGE BASE:
KEYINP: “water body” KEYEXT: BN: KbNode (0,””) CRAWL RESULTS:
```

The *updateKeyCheck* function tries to match the search term with the concepts in the KB-taxonomy in *sys1*. As the used taxonomy (**Fig. 2**a) does not contain a feature class “water body”, the synonym “water bodies”, which is in fact part of the taxonomy, is used instead as search term in the updated KB of system *sys2*.

```
test input> updateKeyCheck sys1

>> sys2: SYSTEM:
...
KNOWLEDGE BASE:
KEYINP: “water bodies” KEYEXT: BN: KbNode (0,””)...
```

The *updateExpandKw* function, applied on *sys2*, finds that the concept “water bodies” is a branchnode. According to **Fig. 4** the function adds all concepts of the subtree of “water bodies”, as well as the synonyms for these concepts, to the search terms. This gives an updated list of expanded terms in the knowledge base used for data crawling.

```
test input> updateExpandKw sys2

>> sys3: SYSTEM:
...
KNOWLEDGE BASE:
KEYINP: water bodies
```

```
KEYEXT: "water body", "waters", "water bodies", "stream",
"watercourse", "creek", "pool", "pond", "lake", "ocean",
"lakes", "sea", "reservoir", "branch", "bay",...
BN: "water bodies"
...
```

The *updateCrawl* function returns all metadata sets for which any of the expanded theme keywords matches any of the theme keywords in the data sources in *sys3*. The function further updates the *result* slot of the knowledge base. The search yields three metadata sets, two in FGDC and one in ISO standard. The matched theme keywords are “Water Bodies” and “Lakes”.

```
test input> updateCrawl sys3

>> sys4: SYSTEM:
...
KNOWLEDGE BASE:
...
CRAWL RESULTS:
FGDC: ThemeKW: Water Bodies,
PlaceKW: Dade County,Florida, Vec/Raster: Vector;
Format: shp, Fees: None,
FGDC: ThemeKW: Water Bodies,
PlaceKW: Miami-Dade County,Florida; Vec/Raster: Vector;
Format: TGRLN; Fees: None
ISO: ThemeKW: Lakes;
PlaceKW: Miami-Dade County,Florida; Vec/Grid: Vector;
Format: ArcInfo Coverage; Fees: None
```

The *iterateSearch* function completes the search process. As the initial search yields some results, there is no need to broaden the search. The harvested metadata sets and the utilized keywords are copied to the agent’s perceptions.

```
test input> iterateSearch sys4

>> sys5: SYSTEM:

HUMAN AGENT:
PERC: RETURNED DATASETS:  FGDC:..., FGDC:..., ISO:...
RETURNED KEYWORDS: "Water Bodies", "Lakes"
...
```

Substituting the demonstrated sequence of operations with the *kwSearch* function on the initial system *initSys* would give the same result. However, simple keyword matching with the submitted search term “water body” in the crawling process would yield zero hits, as no metadata set contains such a keyword.

7 Conclusions and Future Work

An empirical study revealed that keyword matching problems and too many selection options are one of the main reasons for user frustration when searching GIS data in existing Internet portals. We proposed a conceptual framework to overcome some of the reported problems. The core of the architecture is a server sided knowledge base that is able to expand keyword-based searches through a domain ontology and a set of rules for query expansion. To allow direct manipulation, harvested metadata sets should be stored in the client's local memory. A command-line based simulation demonstrated the higher success rate of spatial data harvesting with intelligent keyword expansion over simple keyword matching for a small scenario.

In the given case scenario, the merging of taxonomies was done by hand. However, numerous ontologies have been implemented in machine-readable form. For example, SUMO (Niles and Pease 2001) is a shared upper ontology in KIF. It provides various formalized domain ontologies, also one for geography, which could serve as a basis for a KB-taxonomy. Also WordNet has been completely mapped to SUMO[11]. When using existing taxonomies as is for automated query expansion, the large number of ambiguous terms (e.g., "pool") may cause problems. Part of the future work will be to analyze how machine-readable taxonomies can be used for query expansion, and which techniques of automated taxonomy merging are appropriate. Although the presented simulation gives a proof of concept, another part of future work will be the empirical testing of the proposed conceptual search framework with a graphical interface on an implemented client - server architecture. Empirical tests will reveal if the proposed guidelines actually create a more intelligent user interface.

Acknowledgments

I want to thank Andrew Frank and Jennifer Fu for their helpful comments on an earlier version of this paper. David Wall helped to improve the style of this work.

References

1. Bermudez, L. and Piasecki, M. *Achieving Semantic Interoperability with Hydrologic Ontologies for the Web*. AWRA Special Spring Meeting, Nashville, TN, 2004.
2. Berners-Lee, T., Hendler, J., and Lassila, O. The Semantic Web. *Scientific American*, 284 (5): 35-43, 2001.
3. Brandt, L., Hill, L. L., and Goodchild, M. F. *Digital Gazetteer Information Exchange (dgie) - final report*. Digital Gazetteer Information Exchange Workshop, 1999.
4. CERES. Online information system for California. Retrieved 05/22/2005 from *http://ceres.ca.gov/catalog/* , 2005.
5. Chandrasekaran, B., Josephson, J., and Benjamins, R. What Are Ontologies, and Why Do We Need Them? *IEEE Intelligent Systems*, 1: 20-26, 1999.
6. Getty. Getty Thesaurus of Geographic Names Online. Retrieved 05/20/2005 from *http://www.getty.edu/research/conducting_research/vocabularies/tgn/*, 2005.
7. Gruber, T. R. A Translation Approach to Portable Ontology Specifications. *Knowledge Acquisition*, 5: 199-220, 1993.
8. Guarino, N. Semantic Matching: Formal Ontological Distinctions for Information Organization, Extraction, and Integration. In M. T. Pazienza (Ed.), *International Summer School on Information Extraction: A Multidisciplinary Approach to an Emerging Information Technology* (LNCS 1299, pp. 139-170). London: Springer, 1997.
9. Hochmair, H. H. and Frank, A. U. A Semantic Map as Basis for the Decision Process in the WWW Navigation. In D. R. Montello (Ed.), *Conference on Spatial Information Theory (COSIT'01)* (LNCS 2205, pp. 173-188). Berlin: Springer, 2001.

[11] http://cvs.sourceforge.net/viewcvs.py/sigmakee/KBs/WordNetMappings

10. Hölscher, C. and Strube, G. *Web Search Behavior of Internet Experts and Newbies*. 9th International World Wide Web Conference, Amsterdam, 2000.
11. Jones, C. B., Abdelmoty, A. I., and Fu, G. Maintaining Ontologies for Geographical Information Retrieval on the Web. In *Proceedings of OTM Confederated International Conferences CoopIS, DOA, and OOBASE* (pp. 934-951), 2003.
12. Miller, G. A. WordNet: A Lexical Database for English. *Communications of the ACM*, 38 (11): 39-41, 1995.
13. Niles, I. and Pease, A. Towards a Standard Upper Ontology. In C. Welty and B. Smith (Eds.), *Proceedings of the 2nd International Conference on Formal Ontology in Information Systems (FOIS-2001)* (pp. 17-19), 2001.
14. Pirolli, P., Fu, W.-T., Reeder, R. and Card, S. K. A User-Tracing Architecture for Modeling Interaction with the World Wide Web. *Advanced Visual Interfaces*, Trento, Italy, 2002.
15. Rodriguez, M. A., Egenhofer, M. J., and Rugg, R. D. Assessing Semantic Similarities Among Geospatial Feature Class Definitions. In A. Vckovski, K. Brassel and H.-J. Schek (Eds.), *Interoperating Geographic Information Systems '99* (LNCS 1580, pp. 189-202). Berlin: Springer, 1999.
16. Rosch, E. Principles of categorization. In E. Rosch and B. B. Lloyd (Eds.), *Cognition and categorization* (pp. 27-48). Hillsdale, NJ: Erlbaum, 1978.
17. Shirky, C. Ontology is Overrated: Categories, Links, and Tags. Retrieved 05/30/2005 from *http://shirky.com/writings/ontology_overrated.html*, 2005.
18. Shneiderman, B. Direct manipulation: a step beyond programming languages. *IEEE Computer*, 16 (8): 57-69, 1983.
19. Thompson, S. *Haskell: The Craft of Functional Programming*. Harlow, England: Addison-Wesley, 1999.
20. W3C. OWL Web Ontology Language. Retrieved 05/22/2005 from *http://www.w3.org/TR/2004/REC-owl-features-20040210/*, 2004.

Geospatial Semantic Web: Architecture of Ontologies

Dave Kolas, John Hebeler, and Mike Dean

BBN Technologies, Arlington, VA USA
{dkolas, jhebeler, mdean}@bbn.com

Abstract. An effective ontology architecture enables the development of a geospatial semantic system that forges multiple geospatial data sources into a powerful cross-discipline knowledge source. This paper suggests types of ontologies that could support a geospatial semantic system. Motivations of each of the ontology types will be expounded, as well as potential areas for standardization by the geospatial community. Finally, the use of this approach within the OGC GSW Interoperability Experiment will be discussed.

1 Introduction

The assimilation of semantic web technologies into the world of geospatial information presents several interesting challenges, of which perhaps one of the most obvious is how to incorporate and leverage existing knowledge representations such as Geography Markup Language [GML], the ISO 19100 series [ISO19100], descriptions of geospatial services, and descriptions of geospatial queries into a coherent set of ontologies. A number of different organizations have begun to convert geospatial knowledge representations, but a more focused approach must be developed and led by the Open Geospatial Consortium[1] if the effort is to succeed. Standardized geospatial ontologies form the foundation upon which many more specific geospatial ontologies should be built. Indeed, this area is one of the primary focuses of the Geospatial Semantic Web Interoperability Experiment [GSW IE] within the OGC, of which BBN is an initiator.

This paper outlines five distinct ontology types in OWL [OWL] that contribute to forming a geospatial semantic system. Each ontology type represents a key role in establishing a rich, dynamic, and flexible geospatial knowledgebase. The five ontology types are:

1. Base geospatial ontology – Provides the core geospatial knowledge vocabulary and knowledge structure.
2. Feature data source ontology – Provides an ontological view of WFS data. Allows WFS and GML data sources to fully participate with knowledgebases and ontologies.
3. Geospatial service ontology – Enables knowledgebase discovery and execution of all registered geospatial services.
4. Geospatial filter ontology – Enables the integration of geospatial relationships into the queries.
5. Domain ontology – Provides a knowledge representation that is organized, customized, and aligned with a specific domain and/or user.

[1] http://www.opengeospatial.org

M.A. Rodríguez et al. (Eds.): GeoS 2005, LNCS 3799, pp. 183 – 194, 2005.

We explore each in turn by using a reference example to highlight each ontology's contributions. In addition, the paper highlights areas that require formal standardization.

2 Example Scenario and Architecture

We construct a simple scenario to illustrate the role and benefits of each ontology type.

2.1 Scenario: An Aircraft Emergency

Imagine a scenario in which a Boeing 747 aircraft is flying over a major metropolitan area. Without warning, the plane begins experiencing engine trouble. Emergency crews scramble to determine the best course of action for the plane. The scramble includes querying many data and knowledge sources. Where are the nearest airports? Are these airports capable of supporting a 747? Would the weather conditions at the airports allow an emergency landing? At which airport will a runway be most easily cleared? What other activities are already underway at each location? Which airport has access to the best medical facilities and mechanical facilities?

Current approaches to this situation require the consultation of many different sources of information, some geospatial and some not. Each would likely be in a proprietary format and require special, independent access methods. Much of the integration of the knowledge would occur manually. The success or failure of the actions would depend upon the emergency crew's inherent knowledge of where to find appropriate data and the timeliness of their decisions. All of this takes time and skill that the flight crew can ill afford.

In the semantic web vision, however, much of this data retrieval and integration could be performed automatically. Ideally, instead of seeking out disparate sources of information, an emergency worker could formulate one query to one system which would break the query down into components, access the appropriate data sources, and return an answer or answers. The gained knowledge is then easily incorporated into the knowledgebase to assist in future, similar situations.

2.2 The Geospatial Ontology Architecture

Fig. 1 shows the proposed ontology architecture in supporting a knowledgebase:

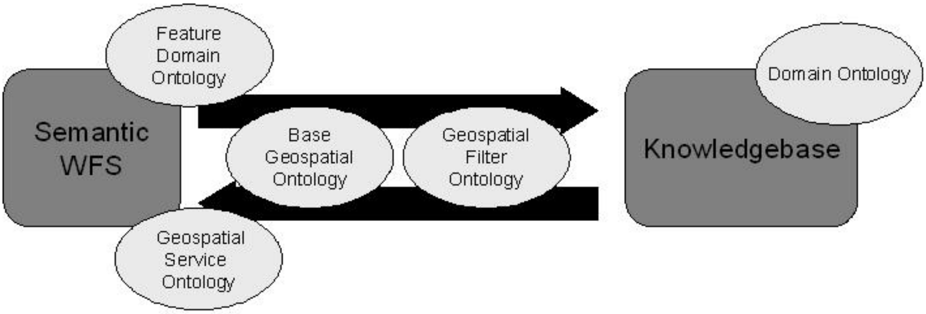


Fig. 1. Geospatial Ontology Architecture

The base geospatial ontology is the core vocabulary that all of the other ontologies must reference. It is a common language and knowledge structure used by both the knowledgebase and the WFS to represent geospatial data, and is analogous to the GML specification [GML]. The domain ontology represents the data from the perspective of a specific user or group. Data is mapped into the domain ontology from the base geospatial ontology and the feature data source. The knowledgebase and the domain ontology are shown on the far right of the diagram to imply that the WFS [WFS] does not require any knowledge of their existence. The filter ontology is a representation of WFS filters as well as geospatial relationships within the query and within the knowledgebase, and is inherently built on base geospatial ontology concepts. It, like the base geospatial ontology, is shared by the WFS and the knowledgebase. The feature data source ontology defines the data from the perspective of the WFS, and is built on base geospatial concepts. While the feature data source ontology is primarily associated with the WFS, the knowledgebase must also have some understanding of it to function as desired. The geospatial service ontology links the base geospatial ontology and the filter ontology to OWL-S [OWL-S] to create semantic web feature service descriptions. Each of these ontologies will now be discussed in detail, with focus given to the base geospatial ontology, the geospatial service ontology, and the geospatial filter ontology which could be standardized by the community.

All five ontologies serve in forming an effective response to the above situation. A query built with the domain ontology initiates a request on the worker's behalf and from the worker's perspective. In order to express the geospatial relationships and filter the data geospatially, the query uses the geospatial filter ontology and the base geospatial ontology. Parts of the query can then be translated for the WFS, which is located via the geospatial service ontology. The service ontology describes the types of services offered by the WFS. The WFS data is returned using its associated ontology, the feature data source ontology. The feature data source ontology translates and maps the underlying formats of WFS into an ontology. This bridges the technology gap between the WFS and ontologies. The WFS responds to the request by forming a response in the feature data source ontology. This data is then mapped to the domain ontology and formatted for the user.

3 Base Geospatial Ontology

The base geospatial ontology forms the ontological foundation of geospatial information. It provides a common base to which geospatial knowledge representations can be linked.

3.1 Motivations for a Common Geospatial Ontology

The motivations for creating a base geospatial ontology start with the same goals that led to the creation of GML:

- A standard way to communicate geospatial data between applications
- An expandable definition of types such that applications can extend the types with their own data
- The capacity to link geospatial application schemas

The current XML representation of GML can realize all of these goals, but a base geospatial ontology extends its power with the significantly greater expressiveness of OWL and the ability to link this data to knowledge outside the geospatial realm. This expands the overall usefulness of the geospatial data while enriching it with complementary information.

3.2 Relationship to GML

The base geospatial ontology should leverage the work done in creating GML by including the same geometry types that are in GML. By using the same concepts that already exist, we enable easier translation from GML to the ontology. In fact, an initial version of this ontology might simply be a conversion of the schemas into OWL. While this would not add any expressiveness per se, it would allow the other ontologies to link to the geospatial content in a widely used manner. As development on the ontology continues, further refinement of the relationships between the types could be added to fully utilize the descriptive power of OWL.

3.3 Linking to the Base Geospatial Ontology

Each of the other ontologies to be discussed has a direct relationship to the base geospatial ontology. Both the domain ontology and the feature data source ontology could extend the types within the base geospatial ontology through OWL just as a WFS can extend the base GML types through XML schema. The other two ontologies, the filter ontology and the service ontology, will necessarily contain properties which use base geospatial types.

3.4 Opportunities for Development and Standardization

Because the base geospatial ontology is the foundation upon which the other ontologies must be built, it should be the primary focus of standardization among the geospatial community. The substantial work that has gone into the creation of the GML schemas as well as the ISO 19100 series of standards will be of enormous benefit both in scoping such an ontology and defining its content.

4 Geospatial Service Ontology

The geospatial service ontology conforms to the OWL-S specification enabling full knowledgebase queries against service offerings. This enhances and extends the current offerings via web services.

4.1 Motivations for Semantic Service Descriptions

While some of the data services that the knowledgebase employs may be known a priori, an effective solution requires the ability to discover and utilize these services automatically. In this way, the goals for the interaction are closely aligned with those of OWL-S [OWL-S]: automatic discovery, invocation, and composition of semantic web services.

In order for a service to be discovered automatically, it must advertise in an abstract sense what it is and what it can do. This corresponds to the concept in OWL-S of a *profile*. The profile of an OWL-S description describes who provides the service, what the service does, and other properties characterizing the service. This description allows the knowledgebase to decide whether or not a particular service is appropriate to its needs, particularly, whether or not the service has data relevant to the query being asked of the knowledgebase.

Automatic service invocation requires both an abstract concept of how the individual services work and the concrete way in which the knowledgebase must interact with them. OWL-S provides these details through the *process model* and *grounding*. For each individual service provided, the process model states the *inputs, outputs, preconditions, and effects.* This means for each service, our knowledgebase knows what it needs to send to the service, what it will get back, what the state of the world must be for the service to succeed, and what the state of the world will be when it finishes. This allows our knowledgebase to ensure that the service meets its needs, as well as knowing what data it must provide in order to achieve the desired result. The grounding brings the interaction down to the most concrete level, telling the knowledgebase which ports, protocols, and encodings to use for invocation.

Finally, the semantic services description enables our knowledgebase to create composite services. While this may initially seem irrelevant to fetching data from a Web Feature Service, there are situations in which our knowledgebase could certainly profit from its usage. Consider a Web Feature Service with a semantic description that does not have the ability to return its feature data in OWL. The knowledgebase could then chain this service to one that can translate GML data of this WFS's particular type schema to OWL of a given ontology.

4.2 Linking OWL-S Service Descriptions and the WFS Specification

It is easy to draw parallels from the parts of an OWL-S service description to the various parts of the Web Feature Service specification [WFS] and a Web Feature Service's typical advertisement, the capabilities document. However, when drawing the parallels in the opposite direction, there are concepts that are unique to the interaction with a Web Feature Service and have not yet been addressed.

In order for the knowledgebase to make use of the WFS, it needs to not only know what the WFS does and how to interact with it, but also what types of content the WFS can return. This content is defined both in terms of feature types and the extent of the data available within those types. Not only does a WFS tell you that it can find airports, for example, but also that it can find airports between a certain set of latitudes.

The scope of OWL-S, however, does not by default have a place for such statements. To add this content to the service description, an ontological definition of the remaining concepts is necessary. This leads to the definition of the Geospatial Service Ontology.

4.3 Usage of the Geospatial Service Ontology

Because the Geospatial Service Ontology would allow the service description to indicate the type of content that can be found on the server and thus allow the knowledgebase to decide if the service is appropriate, it seems most likely to fit into the realm of the profile. Along with the listing of the types of features that the server can return, the service description would probably need to provide a reference to the ontology that defines these types. This function would be analogous to the *describeFeatureType* operation presented by current Web Feature Services, allowing the knowledgebase to find the ontology from which the types are derived.

It is likely that the service descriptions of various Web Feature Services would be quite similar. In fact, much of the process model and grounding could be nearly identical for different Web Feature Services. This would allow significant reuse from WFS to WFS, and thus quicker development, as well as providing potential grounds for standardization.

4.4 The Relationship Between Semantic Service Descriptions and Registries

It is important to note that semantic web service descriptions are complementary to, and not a replacement for, service registries. By providing information about what the WFS offers, the registry then only needs to catalog these descriptions and provide an interface to search them. Exactly how this interface might work is a nontrivial topic that is outside the scope of this paper.

Useful for discussion, however, is how a registry or catalog might fit into the architecture. While the initial description of the architecture was kept as simple as possible, a likely scenario, especially long-term, would add interaction with a service registry before the interaction with the Web Feature Service. In this situation, before the knowledgebase made any contact with a WFS, it could contact the registry with the type of data it requires and get back descriptions of available appropriate WFS's. Another advantage of the semantic service description for a WFS is that it could participate in both specialized WFS registries as well as more general semantic web service registries. This could allow geospatially inclined semantic applications to make the most efficient use of the available servers, while still allowing more general semantic applications to interact with the WFS.

4.5 Opportunities for Development and Standardization

Since the specification for Web Feature Services falls under the purview of the Open Geospatial Consortium, it seems like a logical progression for the definition of the Geospatial Service Ontology and potentially a standard piece of an OWL-S service description for a WFS to be standardized by the OGC as well. Some of the knowledge of which components standardization would benefit will hopefully be derived from the OGC's Geospatial Semantic Web Interoperability Experiment, which is currently in progress.

5 Geospatial Filter Ontology

The Open Geospatial Consortium currently defines a filter encoding [Filter Spec] for use with Web Feature Services. This encoding is used to represent some geospatial and logical relationships that can be used to filter results of a *getFeature* request. In our scenario, we could construct a filter that would represent all features within a 100 mile radius of the plane's current location:

```
<Filter>
    <DWithin>
        <PropertyName>Geometry</PropertyName>
        <gml:Point>
            <gml:coordinates>(plane's coordinates)</gml:coordinates>
        </gml:Point>
        <Distance units='miles'>100</Distance>
    </DWithin>
</Filter>
```

5.1 A Common Language for Geospatial Relationships

While the filter encoding is very effective for interaction with a Web Feature Service, it could be rendered significantly more powerful if built into an ontology. The combination of the concepts of this filter ontology and the concepts the base geospatial ontology would result in a standardizable language that could be used to describe all spatial relationships in a way that is widely accepted and understood, thus useful for the expression of geospatial knowledge outside the realm of the Web Feature Service.

As geospatial semantic web technology becomes more commonplace, there will be an increasing need for a common language for geospatial relationships. For example, the word “within” could conjure many meanings to many people. However, if the concept of “within” is defined in a spatial relationship ontology, then its meaning can be interpreted without ambiguity through an end to end semantic system such as the one described here. There are four areas within the system where the benefits of this filter ontology are immediately apparent: the semantic description of the Web Feature Service; the interaction with the Web Feature Service’s *getFeature* operation, the creation of the client-side query and its representation, and the creation of semantic rules, rule functions, and their implementations within the knowledgebase. Each of these uses will be addressed individually, starting with those related to the Web Feature Service directly.

5.2 The Filter Ontology and the Semantic Web Feature Service

The use of the filter ontology as related to the semantic services description of the WFS is reasonably straightforward. The filter ontology gives the service description a way to reference which types of filters it does and does not support. This could be given by referencing the filter types from the filter ontology that are supported.

The next use of the filter ontology is in performing the actual *getFeature* operation. In a truly semantic system, this request could be formatted in OWL, allowing for easy integration with semantic clients like our knowledgebase. The existence of the filter ontology enables the building of a filter for a feature request as OWL instance data.

```
<filter:DWithin rdf:ID=featuresWithinRadius">
    <filter:propertyName rdf:resource="&gml;#location" />
    <filter:measuredFromGeometry
            rdf:resource="#CurrentAircraftCoordinates" />
    <filter:distance>
```

```
            <units:Miles>
                <units:value>100</units:value>
            </units:Miles>
        </filter:distance>
    </filter:DWithin>
    <gml:Point rdf:ID="CurrentAircraftCoordinates">
        …
    </gml:Point>
```

5.3 The Filter Ontology and the Knowledgebase

Note that to this point no additional power of the filter ontology has been demonstrated, but existing capabilities have been mapped to its use. The additional power derived from using the filter ontology in the semantic web system comes from its use in the client side of the architecture. The first such use is utilizing the ontology to formulate geospatial semantic queries.

Semantic queries with geospatial relationships initially present a problem for RDF query processes; some of the relationships, DWithin and Beyond, are n-ary. This clearly prevents them from being expressed as one simple RDF property. One potential solution to this problem involves user defined functions within the query. This approach is supported by the emerging query language SPARQL [SPARQL]. While this method will certainly enable the geospatial relationships to be placed in the query, it has a few significant disadvantages to using the filter ontology to represent them. First, it requires whatever front-end client that is being used to be able to specially format its semantic queries in this way. While this may not be a significant problem for systems which are focused specifically on geospatial information processing, it makes it less likely that a system designed to build queries from existing ontological concepts will be able to build specialized geospatial queries. Moreover, the addition of more spatial relationships then would require enhancing both the clients and the knowledgebase instead of just the knowledgebase. Secondly, using the special functions within the query could place the burden of calculation on the query processor instead of the knowledgebase. Since the trend now is for integrated rule processing at the storage layer, it seems beneficial to use the mechanisms that will likely already be in place rather than add an additional dependency.

The final area in which the filter ontology could be used is in rule processing within the knowledgebase. While queries that are handled entirely by a WFS could be filtered entirely by the WFS, those that incorporate existing data or require the processing of combinations of data will require the spatial calculations to be performed within the knowledgebase. This ability could be manifested in a set of SWRL built-ins [SWRL], implementations of these built-ins, and SWRL rules making use of the built-ins.

By referencing the same spatial relationships here as are used in the query, there is no requirement for any special knowledge of geospatial relationships in the client front-end; that is, no knowledge about how to process them is required above and beyond how to process any semantic relationship.

5.4 Opportunities for Development and Standardization

Just as the filter encoding specification is standardized by the OGC, it seems logical for such a standards body to develop and approve a specification for an ontology based thereupon. Aside from the standardization of this ontology, it could be beneficial for the OGC to participate in the development of the geospatial SWRL built-ins and potentially a standard set of SWRL rules to imply the geospatial relations (e.g. propagation of containment relations).

6 Domain Ontology

The domain ontology is not a specific ontology, but rather a class of ontologies representing the perspective of the user community that requires geospatial information. As such, it could be made up of or derived from many other ontologies. The domain ontology will very likely be built upon many public ontologies, and could very well be public itself; however, it will be considered private in this scenario because only the client side of the architecture requires any knowledge thereof.

The domain ontology represents the data in the knowledgebase from the perspective of subject matter experts in the relevant domain. This ontology relates the geospatial features and all other data into terms that the user understands. For example, in our emergency scenario, an airplane is something that flies between airports; from the domain of airplane manufacturing, however, the same airplane could be viewed as a product.

The primary purpose of the domain ontology is to represent the concepts over which the user will query. In fact, this is the standard upon which a good domain ontology should be measured; if the user has the vocabulary to ask the questions that they want to ask, then the domain ontology is successful.

In order for the knowledgebase to successfully interact with the Web Feature Service, the domain ontology must have a connection to the base geospatial ontology. This link could be achieved in a number of ways, including subclassing of base geospatial ontology concepts or mapping from base concepts to domain concepts through rules. Either way, this link provides the means to use the previously mentioned geospatial relationships of the filter ontology within queries on the domain ontology.

Due to the domain-specific nature of these ontologies, it is unlikely that they would be general enough with respect to geospatial concepts to be targeted for standardization by the geospatial semantics community.

7 Feature Data Source Ontology

Like the domain ontology, the feature data source ontology is a class of ontologies. However, this class of ontologies is used to represent the domain of the Web Feature Service as opposed to that of the client knowledgebase. Perhaps even more than the domain ontology, the feature data source ontology will almost certainly be built upon or derived from ontologies or data models in common use.

The feature data source ontology is quite analogous to the feature type schemas returned by the WFS *describeFeatureType* operation. It serves exactly the same purpose: to describe the type of data that will be returned beyond the base GML, or in this case base geospatial ontology, types.

The separation of feature data source and domain ontologies also accommodates semantic integration among heterogeneous data sources. For example, WFSs containing aeronautical features corresponding to the US DAFIF [DAFIF] and European AIXM [AIXM] standards could be mapped into a common aeronautical domain ontology.

Clearly, for this additional data to be of use to the client system there must exist a mapping from the feature data source ontology to the client's domain ontology. However, it is not necessary for a client which understands the base geospatial ontology to understand everything or anything from this ontology; the data is simply there if it is desired.

In the long term, it is expected that some ontologies that represent current data schemas as well as new types of data will be developed and become publicly available. This is necessary for semantic clients to make full use of such semantically enabled Web Feature Services. These ontologies are more likely to be standardized by particular communities of interest than by the geospatial community as a whole.

8 The Ontologies in Action: GSW Interoperability Experiment

This ontology approach is currently being evaluated as part of the OGC GSW Interoperability Experiment. Rather than attempting to create full versions of any of these ontologies, we will attempt to validate the approach by creating minimal parts of each ontology to be used as a thread through the system. The current working scenario is very similar to the one described earlier; a specific type of aircraft needs to find a suitable place to land near a city. We have created a knowledgebase which holds OWL data and receives queries in the domain ontology. These queries are then decomposed and translated into appropriate WFS queries. Data returned from the WFS is translated from GML to OWL, and SWRL business rules are applied. The result is then returned to the user.

8.1 Domain Ontology

The domain ontology in this example is that of airports. The ontology defines airports, runways, and other physical features of airports, as well as what it means for an airport to be able to support the plane. It currently also contains the concept of a City, within a different sub-ontology of what would be considered the domain ontology. The domain ontology links to the base geospatial ontology in that a city has Point geometry data.

8.2 Feature Data Source Ontology

The feature data source ontology currently being used is based on the DAFIF schema. This allows for representation of data served by a WFS in DAFIF format. Just as the DAFIF schema links to GML constructs, the DAFIF ontology links to the base geospatial ontology.

8.3 Geospatial Filter Ontology

The filter ontology is designed to mimic the OGC Filter Encoding specification, for easy translation into filters that can be placed in WFS queries. Currently the filter DWithin is being used in the query to define the area in which the airports are being requested. Naturally, this ontology references the base geospatial ontology where the filters require a geometry.

8.4 Base Geospatial Ontology

Currently the base geospatial ontology has only one type of geometry in it: a Point. The point is clearly the simplest form of geometry that we could use, but it nonetheless demonstrates how each of the other ontologies link to base geospatial ontology constructs.

8.5 Geospatial Service Ontology

This is the only ontology concept that has not yet been created or put to use. Using semantic constructs for geospatial service discovery is a current area of research for the Interoperability Experiment.

8.6 Current Usage

A query is formulated for the system using primarily the domain ontology and using the geospatial filter ontology to define the spatial relationships. This query is processed by the knowledgebase, and an appropriate query to what is now a known WFS is formed. The WFS call is made, and the resulting GML is processed into OWL in the form of the DAFIF feature data source ontology. This data is then mapped into the domain ontology through SWRL rules, and the query results are constructed from the product.

9 Conclusions

Developing a rich, standardized set of geospatial ontologies will significantly advance the usefulness and effectiveness of geospatial data regardless of format. This requires the involvement of the whole community to both accurately incorporate the significant progress-to-date and enable the full benefits of the semantic web. Our approach suggests the use of five distinct ontology types to produce a flexible, powerful semantic system – three of which require some form of standardization. Developing a base geospatial ontology, a geospatial service ontology, and a filter ontology to represent geospatial relationships should be goals of the community to further semantic interoperability. The standardization of such ontologies by the OGC would give significant momentum to efforts attempting to create end-to-end semantic geospatial systems, as well as those related to integration of geospatial and non-geospatial data sources.

References

[AIXM] Aeronautical Information Exchange Model. http://www.eurocontrol.int/ais/aixm/

[GML] Cox, S., Daisey, P., Lake, R., Portele, C., Whiteside, A., eds. OpenGIS Geography Markup Language (GML) Implementation Specification, Version 3.0. *Open Geospatial Consortium*, 01/29/2003. https://portal.opengeospatial.org/files/?artifact_id=7174,2003.

[DAFIF] Digital Aeronautical Flight Information File. https://164.214.2.62/products/digitalaero/index.cfm.

[SWRL] Horrocks, I., Patel-Schneider, P.F., Boley, H., Tabet, S., Grosof, B., Dean, M. SWRL: A Semantic Web Rule Language Combining OWL and RuleML. *W3C Member Submission,* http://www.w3.org/Submission/2004/SUBM-SWRL-20040521/,05/21/2004.

[ISO 19100] ISO/TC 211 Geographic Information/Geomatics. http://www.isotc211.org/scope.htm.

[GSW IE] Lieberman, J., et al. Geospatial Semantic Web Interoperability Experiment. Open Geospatial Consortium. http://www.opengeospatial.org/initiatives/?iid=168.

[OWL-S] Martin, D., et al. OWL-S: Semantic Markup for Web Services. *W3C Member Submission*. http://www.w3.org/Submission/2004/SUBM-OWL-S-20041122/, 11/22/2004.

[OWL] McGuinness, D.L., van Harmelen, F., eds. OWL Web Ontology Language Overview. *W3C Recommendation*. http://www.w3.org/TR/2004/REC-owl-features-20040210/, 02/10/2004.

[SPARQL] Prud'hommeaux, E., Seaborne, A., eds. SPARQL Query Language for RDF. *W3C Working Draft*. http://www.w3.org/TR/2005/WD-rdf-sparql-query-20050419/, 04/19/2005.

[Filter Encoding] Vretanos, P., ed. Filter Encoding Implementation Specification, *Version 1.1. Open Geospatial Consortium.* http://portal.opengeospatial.org/files/?artifact_id=8340, 05/03/2005.

[WFS] Vretanos, P., ed. Web Feature Service Implementation Specification, *Version 1.1. Open Geospatial Consortium.* https://portal.opengeospatial.org/files/?artifact_id=8339, 05/03/2005

Formal Approach to Reconciliation of Individual Ontologies for Personalisation of Geospatial Semantic Web

Pragya Agarwal*, Yongjian Huang, and Vania Dimitrova

School of Computing, University of Leeds, UK
{pragya, scs4, vania}@comp.leeds.ac.uk

Abstract. Geospatial domain is characterised by vagueness, especially in the semantic disambiguation of the concepts in the domain, which makes defining universally accepted geo- ontology an onerous task. This is compounded by the lack of appropriate methods and techniques where the individual semantic conceptualisations can be captured and compared to each other. With multiple user conceptualisations, efforts towards a reliable Geospatial Semantic Web, therefore, require personalisation where user diversity can be incorporated. In this paper, a formal approach for detecting mismatches between a user's and an expert's conceptual model is outlined. The formalisation is used as the basis to develop algorithms to compare models defined in OWL. The algorithms are illustrated in a geographical domain using concepts from the SPACE ontology, and are evaluated by comparing test cases of possible user misconceptions. The work presented in this paper is part of our ongoing research on applying commonsense reasoning to elicit and maintain models that represent users' conceptualisations. Such user models will enable taking into account the users' perspective of the real world and will empower personalisation algorithms for the Semantic Web.

1 Introduction

The recent developments in the Semantic Web have great potential for the geospatial community, in specific, because the focus on the incorporation of data semantics will lead to better retrieval and more reliable integration methods by tapping into the semantics during the search process on the web. However, the basic semantic web and the technological developments are not targeted to the specific needs of the geospatial community. The idea of a more focussed 'Geospatial Semantic Web' has been recognised as a research priority within UCGIS initiatives (Fonseca and Sheth 2002). There is a distinct move away from structure and syntax in the geospatial community accompanied by an increased awareness that semantics is the backbone for a successful ontology to enable translation of data from different resources and users. Agarwal (2005) discuss in detail the problems associated with ontology development in the geo-spatial domain primarily due to semantic ambiguities. Egenhofer (2002) identified the need to support queries based on meanings and better definition of spatial terms across a number of disciplines, and the need to integrate multiple

* Supported by Leverhulme Early Career Fellowship.

M.A. Rodríguez et al. (Eds.): GeoS 2005, LNCS 3799, pp. 195 – 210, 2005.

terminological ontologies as a backbone for an effective Geospatial Semantic Web (GSW). Success of a standardised geo-ontology for the semantic web will be determined by the level of acceptance by the users of the services- both experts and naïve, and the level to which the basic geo-ontology is semantically compatible with the users' conceptualisations. Users' preferences, expectations, goals and tasks differ while using the web for information resources. Moreover, people form different conceptual models of the world and these models dynamically change over time. The knowledge-enhanced web services are normally driven by some description of the world which is encoded in the system in the form of an ontology defined by knowledge engineers. The users' conceptualisation of the world may differ, sometimes significantly, from the conceptualisation encoded in the system. If not taken into account, the discrepancies between a user's and a system's conceptualisation may lead to the user's confusion and frustration when utilising Semantic Web services, which, in turn, can make these services less popular.

With the technological developments in search engines and with the vast amount of spatial and earth sciences data and resources now located and available over the web, the number of people who use web-based services is expanding, and hence, dealing with user diversity and providing personalisation functionality becomes paramount (Dolog *et al.* 2003). The one-size-fits-all-users approach to developing web applications is becoming outdated. Personalized information systems aim at giving the individual user optimal support in accessing, retrieving, and storing information. Many different research disciplines have contributed to explore personalization techniques and to evaluate their usefulness within various application areas: e.g. hypertext research has studied personalization in the area of adaptive hypertext systems, artificial intelligence techniques have been widely used to cluster web data, usage data, and user data, reasoning and uncertainty management has been adopted to draw conclusions on appropriate system behaviour. Previous work, such as GLUE, has also attempted to apply machine learning approaches to ontology mapping on the semantic web using heuristics and multi-strategy learning approaches (Doan *et al.* 2002). However, in most systems, there are no mechanisms to capture the interaction and context of the user. There is an urgent need to include the people as an axis in the design, development, and deployment of semantically enriched services. Computational models are needed that can process the different terminological and semantic ontologies and process the semantic incompatibilities between users and the expert's geo-ontology. Semantic personalisation of web-based services is required to exploit the user intentions and perspectives. Development of automated reasoning tools to detect mismatches and discrepancies between the user and the expert ontology forming the backbone for the web-based resources will be a step forward.

This paper outlines the development of a formal approach to user modelling and to aligning and co-ordinating different conceptualisations for achieving personalised services for a GSW. Algorithms are developed from the formal approach to compare different user models based in Ontology Web Language (OWL). Patterns of discrepancies between a user and a system's conceptualisation are analysed and a formal approach is proposed, based on Description Logic (DL), to define these patterns. OWL-based rules are then derived and implemented in a demonstration prototype that compares an expert ontology and a user's conceptualisation, both represented in OWL. The discrepancies are identified and registered as misconceptions between the different models and allows the identification of mismatches between individual conceptualisations in the domain. The innovativeness of this paper lies in proposing personalisation approach for a Geospatial Semantic Web by formalising semantic mismatches, developing algorithms based on these formalisations, combining knowledge elicitation methods and user models with ontology mapping and integration approach, and developing test-bed for evaluation of core geo-ontologies against multiple conceptualisations to allow integration of different perspectives in information systems.

The rest of the paper is structured as follows. Section 2 proposes an argument for the semantic personalisation of web-based geo-spatial services. A brief review of previous methods and initiatives, relevant to the purposes of this paper, in user modelling and personalisation, and in ontology mapping is presented in section 3. The overall methodological framework for the work presented in this paper is outlined in section 4, with details on the formalisation approach and the basic assumptions and notations. The demonstration of the algorithms based on the DL formalisations is presented using the domain ontology for semantic web and test cases of user conceptualisations in section 5, and the results summarised. The conclusions and future research directions are outlined in section 6.

2 Semantic Personalisation of the Geospatial Web

Indeterminacy and ambiguity in meanings are key issues in the development of ontologies in the geographic domain (Agarwal 2005). Empirical results show that individual conceptualisations are characterised by semantic heterogeneity (Hameed *et al.* 2001, Agarwal 2004). To deal with a diverse user population having different preferences, goals, understanding of tasks and conceptual models, existing design paradigms in geo-spatial services will have to be redefined. Furthermore, new diagnostic techniques and models are needed to capture the long-term development of users' capabilities, the dynamics of user's goals and conceptual understanding, the uncertainty and inconsistency of naive users' conceptualizations, and so on. The ambitious target is to offer manageable, extendible and standardized infrastructure for complementing and collaborating applications tailored to the needs of individual users. Without the benefit of deeper semantic or ontological knowledge about the underlying domain, personalization systems cannot handle heterogeneous and complex objects based on their properties and relationships. Nor can these systems possess the ability to automatically explain or reason about the user models or user recommendations. This realization points to an important research focus that combines the strengths of web mining with semantic or ontological knowledge. Traditional personalization and adaptation architectures were suited to deal with closed-world assumption, where user modelling methods, such as overlay, bug library, constraint-based modelling and other marked discrepancies in a user and expert's semantics as erroneous, and often called them misconceptions. New approaches for open-world user modelling that facilitate elicitation of extended models of users are needed to deal with the dynamics of a user's conceptualization. Similarly, methods that acknowledge semantic discrepancies and heterogeneity are required for effectively personalising the web-based services for the geo-spatial community.

The availability of geo-spatial knowledge resources on the web enables members of the public to take advantage of trusted knowledge built by domain experts, e.g. for planning travel routes and for accessing weather information. Geospatial services and systems are also unique in the way that they use data, which are related to locations in space and time, and that the processing of the data with respect to these spatial locations is possible. People's questions to geospatial tools have a spatio-temporal context. One can ask "where is a certain object" or "where are all objects with certain properties" at a given time when trying to find the nearest health services for the elderly; or one can ask "what are the properties of a certain area in space (as well as time)" when trying to ascertain the suitability of an environment (for example, crime rates) while renting or buying a property. Users access the web services with different goals, often; these services require integration of the various different resources to provide a comprehensive result for the user search for their specific requirements. For example, in a 'what is in my backyard' service provided by the Environment Agency (EA) in the United Kingdom, members of the public can see what pollutants may be scattered across their neighbourhood. End-users will have their own contexts of use: property evaluation, ecology, etc. and for a member of the public, a general interest (based on a topographic view of different areas in the city). Each could potentially view the data provided by the others but form their own conceptual understanding of the location-based information.

Resolving the discrepancy between psychological user variables and physical system variables in the area of Geospatial Services goes beyond the user-interface level. Rather than a closed view of the world, the personalisation efforts for geo-spatial services design will ensure that the different perspectives and semantic conceptualisations of the real world are maintained as 'open'. The underlying principle for the methodology adopted in this paper is that an ontology, whatever the scale or granularity, maps the tacit knowledge from the real world (Smith 2003), and makes this knowledge explicit by specification of relations and rules (Agarwal 2005). In this paper, we focus on user modelling and alignment of different semantic models for personalisation. The approach defined in this paper is an effort to allow the system to reconcile the user conceptual model with the core ontology and therefore identify the discrepancies and similarities, and thereby allowing the system to identify the differences in the user conceptualisations with the so-called expert ontology. This will allow, first, for the development of systems that allow personalisation by incorporating user models and diversity and second, as a means to test any core ontologies that are developed as the basis for a geo-spatial services against user conceptualisations for discrepancies and thereby evaluate its reliability as a standard, re-usable ontology. Moreover, the personalisation approach allows flexibility and the possibility of using the user models to enrich the available information resources with shared semantics instead of relying on fixed ontologies available to the developers at the design stage. Using this approach, the research towards specification of well-defined standards and ontologies for inter-operability in the geo-spatial domain can be enhanced and personalised to provide extendibility, flexibility, interoperability, and re-usability. Automating the mapping of multiple conceptualisations and personalisation of web-based services will also facilitate pervasive computing in mobile services and enable more effective use of mobile GIS services.

3 Overview of Existing Approaches for Ontology Mapping

A number of methods are proposed for ontology integration and there is not always an agreement in this community on the meanings of terms that are used (Klein 2001). In the context of the approach adopted in this paper, '***alignment***' and '***mapping***' are most relevant, as these are the least reductive of all methods maintaining the semantic consistency and coherence of the original ontology while comparing and mapping across the different ontologies. Alignment is focussed on a concept-level approach and on finding corresponding semantic properties in the two ontologies (Wache *et al.* 2001). Merging and integration, as mapping processes, are not relevant because the aim is not to develop an integrated resource, but to compare and identify similarities and differences in ontologies for meanings of different concepts from diverse user models. The mapping process during alignment is goal-oriented and since alignment is only carried out for parts of the ontology with corresponding concepts and semantics, the demands on resources, in time and system, are not as high as for merging.

3.1 Existing Methods and Tools - Problems and Limitations

A major problem faced in conceptual alignment is caused by variation in knowledge granularity and discourse domains for the concept across the different ontologies. Visser *et al.* (1998) classify this as a problem in '***conceptualisation***' where domain is interpreted in different ways. Such problems are solved by human interpretation and different conceptualisations can be used together by aligning the overlapping parts of the ontology. However, Noy and Musen (2000) state that finding terms that need to be (and can be) aligned is difficult. Klein (2001) summarises the main issues involved in combining ontologies, such as conflicts in domain coverage, concept scope, synonyms, homonyms, concept description, paradigm, and encoding. In addition, naming conflicts, as proposed by Bishr (1997) and Goh (1997), arising from semantic ambiguities in the use of homonyms and synonyms for concept description, can cause problems in alignment and ontology comparison (Visser *et al.* 1998). Use of natural language specifications and the difference in detail in the ontology can cause problems in the extent to which instances, properties and relations are explicated for a concept, causing 'conceptualisation mismatch' (Klein 2001).

Besides this general framework for mapping information from ontologies, specific methods include a range of top-level to bottom-up approaches for finding semantic associations between ontologies. Silva and Rocha (2003) propose '***semantic bridges'*** where certain mediator agents are used to define the mapping between ontologies. This method is employed in tools such as KRAFT (Preece *et al.* 1999) and MAFRA (Silva and Rocha 2003). However, Wache *et al.* (2001) argue that such mappings fail to maintain the semantics of the concepts as the user is allowed to propose mappings even if these conflict with the internal semantic arrangement of the ontologies. Heuristics-based methods rely on lexical relationships such as synonym, homonym and hypernym for mapping the similarity between concepts (Klein 2001). Other methods such as '***formal ontology methods'*** (Guarino and Welty 2000) rely on inheritance where the different ontologies are linked to a top-level ontology and mappings between the different ontologies are formed by inheriting a common super-class for all lower-level concepts. Such methods adhere closely to formalised frameworks but problems occur when concepts in different ontologies do not overlap or are not terminologically coincidental. Rodriguez and Egenhofer (2003) present a semantic similarity model for geographic data types where linguistic analysis and contextual variability in semantic heterogeneity is incorporated to assess semantic similarity between entity types. Noy and Musen (2002) present a set of criteria for a comparison of these tools based on usability, knowledge expressiveness and inter-operability across different representation languages. It is proposed that as part of future work from this paper, an evaluation procedure with our proposed methodology with existing tools will be carried out.

To sum up, this summary has demonstrated that while previous research has attempted to describe mismatch patterns that may occur between two conceptualisations, the descriptions provided are vague at times and there is a lack of formal descriptions of these patterns. In the following sections, we present a brief overview of the formalisation approach to capture possible mismatches between a user's and an expert's conceptual models.

4 Methodological Framework

Resolution of semantic differences is more crucial than syntactic resolution for aligning ontologies, and conflicts in terminological mismatch are of greater concern while developing re-usable and shareable models from a comparison and similarity assessment of existing ontologies. For the personalisation of geospatial services, the aim is to find points of mismatches between different conceptual models to define ways of either reconciling these differences or using the variability in semantics to find the most suitable information source from the available resources. Therefore, mapping is goal-oriented, with a definitive articulation, and the issue of finding terms on which to align the ontology is not relevant. The methodology employed in this paper is a hybrid approach that is based primarily on principles of *'semantic coordination'* (Bouquet *et al.* 2003), where instead of assuming generic abstract structures for aligning the different ontologies, an agreement on the meanings of concepts is realised by comparing how different knowledge models map onto each other. This is done by considering different levels of semantic knowledge and making the semantic relations explicit for a comparison of the meanings of concepts in different hierarchical structures. Although 'semantic coordination' is distinguished from 'meaning negotiation' in some AI literature (Magnini *et al.* 2002), these are considered as equivalent terms in the approach adopted here.

In brief, the methodological approach can be summarised as such:

1. define a formalisation that captures certain phenomena (misconceptions in our case)
2. implement a demonstration program that follows the formalisation; normally the demonstration are in a specific domain and for a specific problem
3. empirically test the demonstrator to verify the algorithms and the formalisation

4. fine-tune both the demonstrator (i.e. some problems might be due to implementation decisions rather than flaws in the formalisation) and the formalisation (i.e. there may be aspects of the phenomena that may have been missed or defined inappropriately).

Section 4.1 details out the formalisation approach for defining the mismatches, along with the basic assumptions and notations. The prototype developed to reason with OWL-based ontologies is introduced in section 4.2. Section 4.3 outlines the domain, the specific ontology considered for the test scenario and the process adopted for capturing individual user conceptualisations and models. The demonstration of the tool developed in JAVA, on the test ontologies captured from the user study, and based on the formalisation approach presented in section 4.1, is summarised in section 4.4.

4.1 Formalisation Approach

Formal approaches allow the design of algorithms at levels higher than the specific applications, and therefore, bring considerable insights into the design of intelligent system. We use Description Logic (DL) to formally define discrepancies between a user's and a system's conceptualisations. The formal descriptions can be followed in algorithms for user modelling in a variety of domains. In our formalisation approach, we define concept as ‘having meaning’ is distinguished from a ‘term’ that is a referent for the concept to the real world and therefore does not necessarily has semantic content. We hold that a concept is associated with four parts: term, definitions, instance and property (role). If two concepts match all of the four parts, then we consider that there is no misconception between them. Some of the notations used are as follows:

Cu: the concept from the user’s perspective.
Ce: the concept from the expert’s perspective.
Iu: the individual of a concept from the user’s perspective.
Ie: the individual of a concept from the expert’s perspective.
Pu: the property from the user’s perspective.
Pe: the property from the expert’s perspective.
Term(C): the name of the concept
$\forall$R.E(C): value restriction for the concept C.
=: equality, as owl:sameAs
≡: equivalence, as owl:equivalentClass

At this stage of development, we have assumed that all the intentional meaning of a concept is reflected by and only reflected by its term, definitions and the properties. It is accepted that the semantic of the concept and the intentional meanings will also lie in the relations to other neighbouring concepts and this is expected to be taken into account in further development of the reasoning algorithms. One of possible solutions to that exceptional mismatch is to use ‘owl:sameAs’ to explicitly indicate that the two concepts are equal. The definition (or definiens) is expressed by the language of DL, that is, we treat a concept as a set of individuals. Our approach for misconception identification is to first determine the relationship between two concepts by reasoning with their definitions and then check the term and properties for misconception.

We define five types of relationship between two concepts, namely, **equality**, **equivalence**, **subsumption**, **partial overlap** and **disjointness**. Equality, which indicates two concepts have exactly the same intentional meaning, is a special case of equivalence, which merely indicates two concepts have the same set of individuals. Equivalence, again, is a special case of subsumption, which shows one concept is a sub-class of another. Partial overlap refers to two concepts sharing part of instances yet not equivalent. Disjointness defines the relationship between two concepts without any common instance. The complete set of misconceptions and their formalisations have been presented in a related work (Huang *et al.* 2005) and can be referred to for further details. Here, we outline few example misconceptions and their definitions that were tested and identified in the user study to follow.

1. Mismatches based on equivalence

Two concept are considered equivalent if they have the same set of individuals, i.e. $C_u \equiv_{def} D$, C_e
$\equiv_{def} D \vdash C_u \equiv C_e$, where D can be either atomic concept or combination of other concepts.

Term Mismatch $Term(C_u) \neq Term(C_e) \rightarrow$ *Term Mismatch*

Two concepts have the same sets of individuals; however, the concepts may have different intentional meaning. There are many examples in the space ontology. For instance, *edge* is equivalent to *boundary*, yet these two concepts have different intentional meanings in their own rights.

Attribute Mismatch $C_u \equiv_{def} D \cap \forall R \bullet E$, $C_e \equiv_{def} F \cap G \vdash R_u \equiv R_e$, where D, F, G can be any concepts.

$F \subseteq \forall R \bullet E$, $Term(C_u) = Term(C_e) \rightarrow$ *Attribute Mismatch*

This is so-called attribute assignment mismatch (Visser *et al.* 1997), which is a property misconception occurring when two properties are the same except the domains, with one being a subset of another. For instance, a user assigns to *AdministrativeRegion* the attribute of *haslocation*, which could be assigned to *Region* in the expert ontology.

Abstraction Mismatch $C_u \equiv_{def} (D_1 \cup D_2 \cup \ldots \cup D_m)$, $C_e \equiv_{def} (D_1 \cup D_2 \cup \ldots \cup D_m) \vdash C_u \equiv C_e$, where D_i can be either atomic concept or combination of other concepts

C_u does not exists $\rightarrow$ *Abstraction Mismatch*

This mismatch occurs when user has a concept whose abstraction does not exist in expert ontology. For example, *Coordinate* usually include *HorizontalCoordinate* and *VerticalCoordinate*, but *Coordinate* could be missing in the user's conceptualisation, that is, the user is not aware that both of coordinates form the whole coordinate for a location.

2. Misconceptions based on subsumption
Subsumption shows that one concept is a sub-class of another.

Structure Mismatch $C_u \equiv_{def} (D_1 \cap D_2 \cap \ldots \cap D_m \cap \forall R_1 \bullet F_1 \cap \forall R_2 \bullet © \cap \ldots \cap \forall R_k \bullet F_k)$, $C_e \equiv_{def} (E_2 \cap \ldots \cap E_n \cap \forall S_1 \bullet G_1 \cap \forall S_2 \bullet G_1 \cap \ldots \cap \forall S_1 \bullet G_l) \vdash C_u \subseteq C_e$, where D_i, E_i can be either atomic concept or union of other concepts

$\forall$ i, $1 \leq i \leq m$, $\exists$ j, $1 \leq j \leq$q n, $D_i \rightarrow E_j$, and
$\forall$ i, $1 \leq i \leq k$, $\exists$ j, $1 \leq j \leq l$, $R_j = S_j$, $F_j \rightarrow G_j$

$Term(C_u) = Term(C_e) \rightarrow$ *Structure Mismatch*

The description is similar to a subsumption problem. The only difference is the last condition, which indicates concept subsumption with structure mismatch or Definiens Mismatch. For instance, the user may define *Top* as *maximalheight* (*Top* $\equiv_{def} \forall$ *hasHeight* $\bullet$ *Maximum*) whereas expert ontology defines *Top* as *Maximum* with *updirection* (*Top* $\equiv_{def}$ *Maximum* $\cap \forall$ *hasDirection* $\bullet$ *Up*).

4.2 A Prototype for Discovering OWL-Based Mismatch Patterns

Based on the formal descriptions of mismatches, we have implemented algorithms to capture a user's misconceptions defined as the discrepancies between the user's and the expert's perspective of the world. Because the misconception patterns were defined in DL, they could easily be applied to conceptualisations defined in OWL. Although OWL is limited in providing direct support for representing the semantics of the procedures for processing geospatial data, it is a commonly used standard for ontology creation and development for the semantic web and therefore we use it here as a standardised reasoning mechanism for developing our algorithms. Semantic restraints are imposed in OWL on terms and concepts. This means that the meanings are included, along with relationships and objects, and a richer set of specifications such as disjointness, cardinality, equality, symmetry and properties can be stated in an OWL specification (McGuinness and van Harmelen 2003). In OWL, 'classes' define concepts from the domain, 'individuals' represent specific instances and 'properties' define the values that each individual can take.

The five types of relationship are defied in OWL, as follows:

owl:same	**equality**
owl:eq	**equivalence**
owl:subsume	**subsumption**
owl:po	**partial overlap**
owl:dj	**disjointness**

For the implementation of the algorithms, we have used rule-based OWL inference engine Jena2.. Following the triple-based nature of Jena, we have defined a set of rules to capture the mismatches defined in the formalisation. For example, to check for *Structure Mismatch* for domain based on *subsumption* relation between two concepts, the following rule will be passed to Jena:

[rule3: (?C rdfs:domain ?D), (?D rdfs:subClassOf ?E), (?E rdfs:subClassOf ?F)-> (?C rdfs:domain ?F)]");

The rule states that if an object property C has a Class D as its domain and D is a subclass of E, which is a subclass of F, then C takes class F as another domain. For instance, in the SPACE ontology (see section 4.3), the object property *hasCapital* takes *City* as its domain and *City* is a subclass of *AdministrativeRegion*, which is a subclass of *Region*. Thus, we can deduce that *hasCapital* also takes *Region* as its domain, which corresponds to the user's perspective.

4.3 Case Ontology and Domain

There is a lack of comprehensive geo-ontologies for the Semantic Web. This is partly because of the vagueness and ambiguity in the geographic domain and because many of geographic concepts and terminology are anchored in human cognition. The SWEET (Semantic Web for Earth and Environmental Technology) suite of ontologies is constructed by NASA to provide an upper level ontology as a basis for a common semantic framework for the GeoSciences. The web interface created for SWEET (http://sweet.jpl.nasa.gov) is aimed at supporting user intervention and is based on the assertion that a comprehensive ontology should include collaborative capabilities and community participation, thus allowing the users to update the terms and concepts and include their own conceptualisations in the existing knowledge base. These are hierarchical ontologies, for example, 'hydrosphere' is a parent concept for 'surface water' which is a parent for 'river' which is a parent for 'Mississippi River'. The Global Change Master Directory (GCMD) was used along with keywords from the Earth Science Modelling Framework (ESMF) to populate the ontologies. The SPACE ontology contains the maximum relevant concepts for spatial divisions and locations in the geographic context, and includes terminology specific to the spatial domain, focussing on spatial extents, such as country, equator, boundary, and relations such as has capital, has location, top of, north of etc. Table 1 shows an example of how a concept such as '*Region'* is conceptualised in the SPACE ontology.

Table 1. Conceptualisation of 'Region' in SPACE ontology

space#region	physical, material
< supertype space#NumericGeometricObject_2D; space#SpatialObject	
>subtype space#AdministrativeRegion	
attributes (direct and inherited) space#inside	
disjoint space#NumericGeometricObject_3D	

4.4 Eliciting User Semantics

The overall purpose of the experimental work is to conduct more appropriate and comprehensive evaluation of the algorithms and ontology mapping procedures based on real data of multiple user conceptualisations.

4.4.1 Experimental Framework

Previous related work (Agarwal 2004) has shown ways in which user perspectives, especially for semantic content of the domain, can be extracted through mapping individual conceptualisations. For the work presented in this paper, the primary aim was to extract user models and semantic conceptualisations for the concepts present in the 'expert' ontology. SWOOP (http://www.mindswap.org/2004/SWOOP/) was used as an exploration tool to identify the concepts and relations in the SPACE ontolog. The concepts to be used in the experiment were delineated based on previous related work (Galton 2001, Agarwal 2004a) that have shown the inherent ambiguity in several of these concepts, such as *region*, *boundary*, and *location*. These concepts were selected because of their links to the real world as well as to the human commonsense reasoning. Also, these concepts were identified to be commonly used to extract web-based geographic information, and that require multiple levels of transformation, for example, ‘show all the hospitals in my region’ or ‘show the boundary of the most economically developed area’, or the ‘boundary for the flood prone area in my region’, ‘show information about pollution level near the location of a specific industry’.

The questionnaire included a list of the relevant concepts without making any inherent hierarchy or relations apparent. The spatial concepts and the respective sub and super classes included in the questionnaire are ***Region, Zone, Spatial Object, Administrative Region, Geometrical Object, Boundary, Edge, Administrative Boundary, Country, State, City, Political Division, Position, Location, distance.*** The design of the questionnaire was enabled capturing of partial conceptual models of the users, and was focussed solely on the concepts that were delineated for the purpose of the study. Along with this, detailed definitions along with examples were provided for subclass, superclass, property and synonym to minimise any individual biases in interpretation of these terms. **Synonym** of a term ‘t’ was defined as ‘Similar in meaning to ‘t’, such as *table is synonym with desk*’. **Subclass** of a term ‘t’ is defined to hold ‘when a term is a child term of ‘t’, such as a *coffee table is a subclass of a table*’. **Superclass** of term ‘t’ exist ‘when a term is a parent of term ‘t’, such as *furniture is a superclass of table*’, and **Property** of term ‘t’ is stated to be ‘when a term is a characteristic of term ‘t’, such as *‘has legs’ is a property of table*’. The questionnaire were sent to a wide range of end-users, from a cross section of disciplinary backgrounds, including Geography, Information management, Linguistics and Computer Science. Although the user responses were treated as anonymous, some personal information was also requested on previous experience in using web-based services for geographic information, the websites that were used and examples of problems that were faced in using web services for geographic resources and information. Most of the users admitted familiarity with a range of web-based weather and travel services as well as for environmental information concerning their neighbourhoods and localities. As this was a test scenario, implemented primarily to test the effectiveness of the matching algorithms, detailed control experimental settings were not practical. In addition, the primary focus in this study was on testing the methodology for alignment of diverse models, and therefore email communication was used. The respondents were asked to work independently and it was expected that the simple, self-explanatory design of the questionnaire enabled the respondents to express their internal semantic conceptualisation and understanding of the different concepts.

5 Results from the User Studies

The user ontology constructed from the aligned model in OWL, expert ontology (in this case, SPACE ontology) and a user concept is taken as an input in the automated tool 'Conceptualisation Comparison' developed for this purpose. The program first collects all the information related to the user's central concept (*Region* in our experiment). These include the subclasses, super classes, synonyms and properties (with range). Secondly, it searches for the corresponding concepts and properties in the expert ontology, based on the rules of pattern matching and formalisations of misconceptions, examples of which are provided in section 4.1. Finally, the user's perspectives on *Region* are mapped to the corresponding concepts and properties in expert ontology. The perspectives that have no correspondents are reported as misconceptions or mismatch. Figure 1 shows the interface of the automated tool, showing pull-down menus for concepts in user and expert ontology, an initialisation button, and a window that shows the misconceptions after aligning, mapping and comparing the two input ontologies.

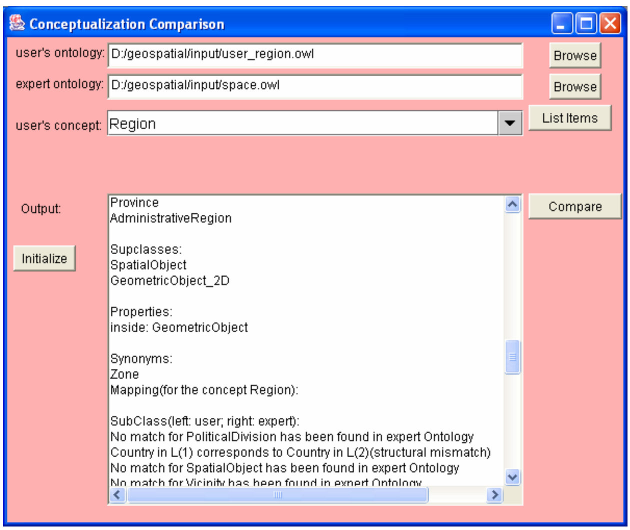


Fig. 1. The *Conceptualisation Comparison* Interface allowing the visualization of mismatches between the input ontologies based on DL algorithms and OWL-based reasoning

The approach adopted in this test case to demonstrate the effectiveness of the automated prototype for comparing individual semantic conceptualisations is based on a comparison of the central concept from the user ontology with all the concepts in the expert ontology. This differs from the approach that we adopted in a related work (Huang *et al*. 2005) where artificially constructed ontologies were used to test the tool and therefore, concepts, both from the user as well as the expert ontology, were specified for comparison to limit the computational complexity. Unlike other real-world ontologies, which usually consist of a large number of inter-related concepts, this user ontology, as a demonstration, was limited to one concept and the relevant semantic relations and properties, as extracted from user questionnaires. Therefore, in this case, the program compared the complete user ontology with the expert ontology, and looked for all relevant concepts. However, as the scale of the user ontology increases with further concepts, relations and semantic properties, the computing complexity will increase significantly, because the program will compare every concept and property related to each concept in the user ontology against a the total number of concepts and properties in the expert ontology.

A few of the representative results from the automated mapping are as follows:

SubClass(left: user; right: expert):
Country in L (1) corresponds to Country in L(2)(*structural mismatch*)
City in L(1) corresponds to City in L(2)(*structural mismatch*)
State in L(1) corresponds to State in L(2)(*structural mismatch*)
No match for Vicinity has been found in expert Ontology
No match for Suburb has been found in expert Ontology

SuperClass(left: user; right: expert):
SpatialObject match SpatialObject
No match for Continent has been found in expert Ontology
No match for State has been found in expert Ontology
No match for City has been found in expert Ontology

Synonyms(left: user; right: expert):
Zone match Zone
No match for District has been found in expert Ontology
No match for Area has been found in expert Ontology

Properties:
isPartOf(City) corresponds to City in L(2) in expert Ontology
hasCapital corresponds to hasCaptial(City) in expert Ontology
isA(SpatialObject) corresponds to Spatial Object as Superclass in expert Ontology
No match for hasPosition has been found in expert Ontology
No match for hasContext has been found in expert Ontology
No match for hasEdge has been found in expert Ontology
No match for hasBoundary has been found in expert Ontology
No match for hasCommonCharacteristics has been found in expert Ontology
No match for hasLocation has been found in expert Ontology
No match for hasClimate has been found in expert Ontology

The primary misconceptions that have been identified by the program consist of **Abstraction Mismatch, Structural Mismatch** and **Attribute Mismatch** (see section 4.1 for formalizations). However, some limitations were also noted in the demonstration tool. The program fails to report all misconceptions which were apparent on manual inspection. For example, in the user's OWL file, 'AdministrativeRegion' is both subclass and super class of 'Region', which can be interpreted in two ways: (a) 'AdministrativeRegion' is equivalent to 'Region' and; (b) there exist conflicting relationships on the concept. The program, which has discovered that 'AdministrativeRegion' is a subclass of 'Region', however, has no means to solve confusion and ambiguity from user's perspective. For the latter reason, the user's ontology has some properties such as *isPartOf(City)*, which can be arguably translated as a subclass of *City*, which is actually a subclass of *'Region'* in expert ontology. This is partly because of the inherent organization of the user's OWL file and partly because the program at this stage lacks robust mechanisms to also handle the semantic meaning of properties along with its capacity to reason with semantic meanings of concepts. So, although the innovativeness of the formalization approach allows us to make the semantic misconceptions between concepts apparent, the tool itself needs further development to enable identification of semantic mismatches also at the property level. The systematic methodological approach has facilitated the evaluation of the demonstration tool and identify areas where it needs fine-tuning to make the formalizations more effective in identification of semantic mismatches between individual user conceptualizations in the domain.

6 Conclusions and Future Work

The work presented in this paper is part of our ongoing research on applying commonsense reasoning to elicit and maintain models that represent users' conceptualisations of the real world. Such user models will enable taking into account the users' perspective of the world and will empower personalisation algorithms for the Semantic Web. A formal approach for detecting mismatches between a user's and an expert's conceptual model is outlined. The formalisation is used as the basis to develop algorithms to compare two conceptualisations defined in OWL. The algorithms are illustrated in a geographical domain using a geo-ontology in OWL developed as part of the SWEET initiative for the Semantic Web by NASA, and have been tested by using test cases of possible user misconceptions.

A number of possible benefits that the above approach can afford to the development of personalised geospatial services are foreseen. The approach defined in this paper is an effort to allow the system to reconcile the user conceptual model with the core ontology and therefore identify the discrepancies and similarities, and thereby allowing the system to identify the differences in the user conceptualisations with the so-called expert ontology. This will allow, first, for the development of systems that allow personalisation by incorporating user models and diversity; second, this approach can be used to test core ontologies developed as the basis for a geo-spatial system/service against user conceptualisations for discrepancies. This will be useful in evaluating the reliability of ontologies for standardisation and re-usability. Moreover, the personalisation approach allows the possibility of using the user models to enrich the available information resources with shared semantics instead of relying on *'fixed'* ontologies available to the developers at the design stage.

The Semantic Web paradigm requires the deployment of appropriate user modelling approaches that capture and maintain different user perspectives. At this stage, the identification of suitable concepts from the core ontology and the capturing of the user conceptualisations, as well as development of user ontologies from the results are manual and require human intervention. It is proposed that this process be deployed on a web-based service and be largely automated. Future work will develop on this work to develop automated web-based interfaces that can use the different semantics, detect the semantic mismatches and process the information available to integrate the knowledge resources based on individual conceptualisations of the domain. The long-term goal of our research is to apply commonsense reasoning approaches to capture and maintain users' conceptual models and to use these models for personalised, semantically-enhanced search on the web. For this, we consider that the domain expertise is encoded in some ontology (or several ontologies) pre-defined by domain experts and knowledge engineers. This expertise is used to guide the intelligent behaviour of the system and is combined with some model of the user that corresponds to the user's conceptualisation of the domain. Work is undergoing in developing more robust knowledge elicitation methods to capture individual conceptualisations in the geographic domain. These methods will be used along with the algorithms and formal approaches described in this paper to test and fine-tune the algorithms. Future work will also include incorporation of uncertainties in user models and semantic conceptualisations, and target more complex mappings and mismatches. We are also looking into the possibilities of using the mismatch detection algorithms in combination with additional reasoning to deal with vagueness and heterogeneity problems. The aim is also to explore the possibility of including other ontology language, standards and reasoning methods (for example, SWRL, RDF and XML) within these algorithms. For this purpose, evaluation of transferability between different web languages for geographic concepts is currently being carried out.

Acknowledgments

The authors wish to thank the three anonymous reviewers whose comments helped improve the quality of the paper. Pragya Agarwal's research is supported by the Leverhulme Early Career Fellowship.

References

1. Agarwal, P. Contested nature of 'place': knowledge mapping for resolving ontological distinctions between geographical concepts, In Egenhofer, M., Freksa, C., Harvey, M. (Eds.) *Third International Conference, GIScience* 2004, Lecture Notes in Computer Science, Springer-Verlag, Berlin, LNCS 3234, pp. 1-21, 2004.
2. Agarwal, P. Ontological considerations in GIScience. *International Journal of Geographical Information Science* (IJGIS), Vol. 19, No. 5, pp. 501 – 535, 2005.
3. Bishr, Y. A. Semantic Aspects of Interoperable GIS, ITC Publication No. 56, Enschede, NL, 1997.
4. Bouquet, P., Serafini, L. and Zanobini, S. Semantic coordination: a new approach and an application, *Proceedings of 2nd International Semantic Web Conference (ISWC2003*), Sanibel Island, Florida, 20-23 October, 2003.
5. Doan, A., Madhavan, J., Domingos, P., Halvey, A. *Learning to map between ontologies on the semantic web*. In Proceedings of WWW'02, pp. 662—673, 2002.
6. Dolog, P., Henze, N., Nejdl, W., and Sintek, M. Towards the Adaptive Semantic Web. *1st Workshop on Principles and Practice of Semantic Web Reasoning (PPSWR'03*), co-located with International Conference on Logic Programming (ICLP'03), Dec. 8, 2003, Mumbai, India, 2003.

7. Egenhofer, M.J. Toward the Semantic Geospatial Web. In *Proceedings of the Tenth ACM International Symposium on Advances in Geographic Information Systems,* McLean, Virginia, 2002.
8. Fonseca F., and Sheth A., The Geospatial Semantic Web, UCGIS White Paper. http://www.ucgis4.org/priorities/research/2002researchagenda.htm, 2002.
9. Goh, C. H. *Representing and Reasoning about Semantic Conflicts in Heterogeneous Information Sources*. Doctoral Dissertation, MIT, MA, 1997.
10. Galton, A. On the Ontological Status of Geographical Boundaries, *Meeting on Funadamental Questions in GIScience*, Manchester, 2001.
11. Guarino, N. and Welty, C. A Formal Ontology of Properties, *Proceedings of the ECAI-00, Workshop on Applications of Ontologies and Problem Solving Methods*, Berlin, Germany, pp. 121-128, 2000.
12. Hameed, A., Sleeman, D. and Preece, A. Detecting Mismatches among Experts' Ontologies Acquired through Knowledge Elicitation, R&D in Intelligent Systems XVIII, *Proceedings of ES2001:the 21st SGES International Conference on Knowledge Based Systems and Applied Artificial Intelligence*. Springer-Verlag, London, Cambridge, U.K., pp. 9–24, 2001.
13. Huang, Y, Dimitrova, V. Agarwal, P. Detecting mismatches between a user's and a system's conceptualizations to empower user modelling for the Semantic Web, Proceedings of the Workshop on Personalisation of Semantic Web, PersWeb 05, *10th International Conference on User Modelling*, 24-29 July 2005, Edinburgh, pp.53-63, 2005.
14. Klein, M. Combining and relating ontologies: an analysis of problems and solutions, IJCAI-2001 *Workshop on Ontologies and Information Sharing*, Seattle, WA, pp. 53–62, 2001.
15. Magnini, B., Serafini, L. and Speranza, M. Using NLP Techniques for Meaning Negotiation, *Proceedings of VIII Convegno AI*IA*, Siena, Italy, 11-13 September 2002.
16. McGuinness, D. L. and vanHarmelen, F. OWL Web Ontology Language Overview: W3C Candidate Recommendation 18 August 2003, http://www.w3.org/TR/owl-features/./ , 2003.
17. Noy, N. F. and Musen, M. A. Evaluating Ontology-Mapping Tools: Requirements and Experience, *Proceedings of the Workshop on Evaluation of Ontology Tools at EKAW'02* (EON2002). Available as SMI technical report SMI-2002-0936, Siguenza, Spain, October 2002.
18. Preece, A. D., Hui, K.-Y., Gray, W. A., Marti, P., Bench-Capon, T. J. M., Jones, D. M. and Cui, Z. The KRAFT Architecture for Knowledge Fusion and Transformation, *19th SGES International Conference on Knowledge-based Systesm and Applied Artificial Intelligence* (ES'99). Springer, Berlin, 1999.
19. Rodríguez, A. and Egenhofer, M. Determining Semantic Similarity Among Entity Classes from Different Ontologies. *IEEE Transactions on Knowledge and Data Engineering*, 12(2): 442-456, 2003.
20. Silva, N. and Rocha, J. MAFRA – An Ontology Mapping FrAmework for the Semantic Web, *Proceedings of the 6th International Conference on Business Information Systems*, UCCS, Colorado Springs (CO), USA, May 2003.
21. Smith, B. Ontology: Philosophical and Computational. In: L. Floridi (Editor), *The Blackwell Guide to the Philosophy of Computing and Information*. Blackwell, Oxford, 2003.
22. Visser, P. R. S., Jones, D. M., Bench-Capon, T. J. M. and Shave, M. J. R. Assessing Heterogeneity by Classifying Ontology Mismatches. In: N. Guarino (Editor), *Proceedings of Formal Ontology in Information Systems, FOIS' 98*. IOS Press, Amsterdam, The Netherlands, Trento, Italy, pp. 148-162, 1998.
23. Wache, H., Vögele, T., Visser, U., Stuckenschmidt, H., Schuster, G., Neumann, H. and Hübner, S. Ontology-based Integration of Information - A Survey of Existing Approaches, *Proceedings of IJCAI-01 Workshop: Ontologies and Information Sharing*, Seattle, WA, pp. 108-117, 2001.

Incorporating Update Semantics Within Geographical Ontologies

Xuan Gu and Richard T. Pascoe

Department of Computer Science and Software Engineering (CSSE),
University of Canterbury, Canterbury, NZ
mark.guxuan@gmail.com,
richard.pascoe@canterbury.ac.nz
http://www.cosc.canterbury.ac.nz/richard.pascoe

Abstract. In this paper is described a systematic technique by which geographical ontologies, descriptions of concepts and relationships that exist for geographical domains of interest, may incorporate *update policies*, knowledge that governs the updating of data described by these ontologies. Of particular interest are those ontologies describing distributed geographical data where different components are maintained by separate organizations. As provider organizations change their individual contributions to the distributed data set, the efficacy of a local copy of this distributed data will decline. The incorporated update policy of the associated ontology for this local copy will be used to determine when an accumulation of changes, described by *update notifications*, justifies updating the local copy. Update policies and update notifications are assumed to have a common ontological basis. Ontologies are described using the Unified Modelling Language (UML) [4] with the semantics of an update policy being expressed using a UML profile described in this paper. The intent is to implement software agents that will execute the update policy and when justified will generate a plan by which the local copy can be updated to reflect the distributed data currently available.

1 Introduction

Geographic Information Systems (GISs) are used by many organizations, governments, research institutes and other bodies for tasks such as gathering, transforming, manipulating, analyzing, and producing information related to spatial data. For example, police and fire departments may use GISs to locate landmarks and hazards, plot destinations, and design emergency routes. The tasks are often further complicated by organizations using shared or distributed data sets in their analysis to reduce costs and improve consistency across related data sets.

The environment being considered in this research is analogous to the notion of a Geospatial Information Community (GIC) initially described by McKee and Buehler [1] and refined by Bishr [2] where organizations share data sets within the context of a common ontology. In such an environment. individual

M.A. Rodríguez et al. (Eds.): GeoS 2005, LNCS 3799, pp. 211–226, 2005.

organizations using this data may also be responsible for maintaining elements of this common data to reflect changes in the real world phenomena represented. Thus, an organization may fulfill either or both of the following roles:

a data provider P_n where this organization provides others with access to at least one data set λ. For a given system, there is a collection of data sets $\Theta = \{\lambda_g\},\ 1 \leq g \leq G$ provided by N providers ($G \geq N$). In the particular environment being considered, Θ will typically be large and transmitting large parts of this data will be resource intensive and potentially slow.

a data consumer C_m where this organization at some point in time t creates a local copy $\Lambda_{p,m,t}$ of a distributed data set $\Lambda_p \subseteq \Theta_t$. The local copy $\Lambda_{p,m,t}$ is needed to satisfy accessibility and performance requirements.

The focal point of this research is the ease with which changes to Θ over time can be selectively propagated to $\Lambda_{p,m,t}$. That is, over a period of time $t \ldots t',\ t < t'$, any provider P_n may apply changes $\Delta\lambda_{g,t'}$ to their data set $\lambda_{g,t}$ to form an updated data set $\lambda_{g,t'} = \lambda_{g,t} + \Delta\lambda_{g,t'}$. In doing so, any consumer C_m of the distributed data set $\Lambda_p : \lambda_g \in \Lambda_p$ may update $\Lambda_{p,m,t}$ to form $\Lambda_{p,m,t'}$. The conditions under which C_m updates $\Lambda_{p,m,t}$ is determined by comparing a description of $\Delta\lambda_{g,t'}$, provided by P_n, with the update policy for $\Lambda_{p,m}$. The notation being used throughout this paper is summarized in Table 1.

Table 1. Summary of notation used to refer to elements of the proposed system

Symbol	Description	
P_n	Provider n	$1 \leq n \leq N$
C_m	Consumer m	$1 \leq m \leq M$
$\lambda_{g,t}$	Data set g provided by P_n at time t	
Θ_t	All data provided within the system at time t.	$\Theta_t = \{\lambda_{g,t}\}\ 1 \leq g \leq G,\ G \geq N$
$\Delta\lambda_{g,t'}$	Changes to $\lambda_{g,t}$ occurring between t and t'	$t < t'$
Λ_p	Distributed data set p	$\Lambda_p \subseteq \Theta$
$\Lambda_{p,m,t}$	The local copy of a distributed data set $\Lambda_p \subseteq \Theta$ created at some point in time t by C_m	
$\Delta\Lambda_{p,t'}$	Changes to $\Lambda_{p,t}$ occurring between t and t'	

Notes:

1. For any given point in time t:
   (a) Any data set λ contributed by P_n is uniquely identified by g;
   (b) Any distributed data set Λ is uniquely identified by p.
2. For any given two points in time $t,\ t' :\ t < t'$:
   (a) $\lambda_{g,t'} = \lambda_{g,t} + \Delta\lambda_{g,t'}$
   (b) $\Lambda_{p,t'} = \Lambda_{p,t} + \Delta\Lambda_{p,t'}$ where $\Delta\Lambda_{p,t'} = \{\Delta\lambda_{i,t'} : \lambda_i \in \Lambda_p\}$

Although the problem of synchronizing a local copy of a distributed data set could be solved by having consumers access the distributed data set directly rather than indirectly through a local copy $\Lambda_{p,m,t}$, such an ideal situation is unlikely, because of the typically large quantities of data being shared, security and reliability concerns, cost and flexibility, and the different performance and business requirements of individual consumers.

The proposed approach uses software agents to synchronize $\Lambda_{p,m,t}$ according to the associated update policy embedded in the associated ontology. While these agents may act independently of each other, the expectation is that they will collaborate where the agents have common goals within their update policies thereby optimizing the process by which $\Lambda_{p,m,t}$ can be synchronized. The remainder of this paper contains descriptions of our progress towards achieving the following research objectives:

1. To determine what characteristics of a change in data value or structure need to be described in an update notification $\Delta\lambda_g$ to facilitate research objective 3, described below.
   To be effective, these update notifications must contain detailed descriptions of the changes so that each software agent can better evaluate the significance of $\Delta\lambda_g$ in the context of the update policy for $\Lambda_{p,m}$.
2. To provide a means by which an organization's update policy can be easily incorporated into the ontology for $\Lambda_{p,m}$. The approach adopted here involves expressing the ontology using UML that has been extended by a UML Profile for clearly expressing the semantics of an update policy.
3. To implement
   (a) agents that can individually or collectively
      i. determine the importance of a particular update notification in the context of each agent's update policy. In essence, addressing issues such as:
         A. computing $\Delta\Lambda_{p,m,t'}$ by accumulating and merging update notifications $\Delta\lambda_g$ from the providers of $\lambda_g \in \Lambda_p$.
         B. evaluating $\Delta\Lambda_{p,m,t'}$ within the context of the associated update policy to determine whether the difference is significant thereby justifying the update of $\Lambda_{p,m,t}$.
      ii. create an *update execution plan* that details the steps by which to synchronize $\Lambda_{p,m,t}$ with $\Lambda_{p,m,t'}$.
   (b) a system that will execute the update execution plan

Objectives 1 and 2 are discussed in Sections 3 and 4 respectively while some preliminary ideas for objective 3 are briefly discussed in Section **??**. The general approach adopted for this research is described next.

2 Overall Approach

Geographical ontologies are expressed as UML models [6] encoded as XML Metadata Interchange (XMI) documents [13] in a manner compatible with the notion of application schemas that conform to ISO 19109 [7]. This standard is one of many being defined by Technical Committee 211 (TC211) of the International Organization for Standardization to facilitate the interoperability of GISs.

The intent is to draw upon the many ISO TC211 standards and the associated harmonized UML model underlying these standards as the basis for geographical ontologies that are analogous to the notion of a Platform Independent Model (PIM) within the Model Driven Architecture (MDA) defined by the Object Management Group (OMG) [8].

At least one such PIM or ontology is expected to be defined for the entire distributed data set Θ available to organizations participating by either providing components of the distributed data set λ_g, consuming components of this data set as a local copy $\Lambda_{p,m,t}$, or both. Each organization will elaborate upon the distributed PIM or ontology to form another that is also platform independent, but which is restricted to only those components of the PIM that are of interest to the organization. This restricted ontology, for the local copy of the relevant components of the distributed data set, is further elaborated to form a Platform Specific Model (PSM) that introduces platform specific semantics reflecting the particular storage techniques being employed. While there is likely to be only one PIM describing the entire distributed data set, there will be at least one pair of restricted PIM and corresponding PSM for each organization.

Use of the MDA in this way clearly distinguishes between heterogeneity arising from conceptual differences in the way each organization views the shared geographical data and heterogeneity arising from the each organization using different implementation specific technologies for managing and processing this data. This distinction facilitates the development of flexible, scalable, loosely coupled systems. Furthermore, use of the MDA and the UML for expressing ontology partially addresses the problem that F. Fonseca mentioned in his paper about the gap between ontologies and the software components [5].

Seth and Larson's notion of a five-level schema framework for distributed systems [9] is combined with the OMG's notion of PIMs and PSMs elements of the MDA to form the four-level ontology framework shown in Figure 1. Within this framework, an ontology may fulfill one of five roles:

local PSM In this role, an ontology describes an organization's geographic domain of interest and reflects the specific platform managing the data. Each organization is likely to have quite different ontologies at this level of abstraction.

local PIM In this role, an ontology also reflects an organization's geographic domain of interest but in a platform neutral manner ideally using appropriate elements of the ISO 19100 harmonized model. In this role, an ontology is analogous to what Sheth and Larson [9] refer to as a component schema.

export PIM In this role, an ontology defines those parts of the local PIM for which each organization m is contributing data (λ_g) to form the distributed data set Θ. Such an ontology will be a subset of the local PIM ontology.

import PIM In this role, an ontology will define the local copy of the distributed data set $\Lambda_{p,m}$ that is of interest to organization m. Such an ontology may be a (part of a) of an ontology generically referred to as distributed PIM.

distributed PIM In this role, an ontology will define data being shared by participating organizations.

When fulfilling any of the export, import, and distributed, PIM roles, the ontology is expressed using a vocabulary common to all organizations and consistently uses elements of the ISO TC211 harmonized model. An ontology fulfilling one of these three roles corresponds to Sheth and Larson's notion of an export, external, and federated schemas respectively. The export and import PIMs roles are regarded as being at the same level of abstraction and will be defined by an organization acting as a provider or a consumer respectively, while each the three remaining roles are regarded as being defined at a distinct levels of abstraction: therefore, there are four levels of abstraction in the proposed framework.

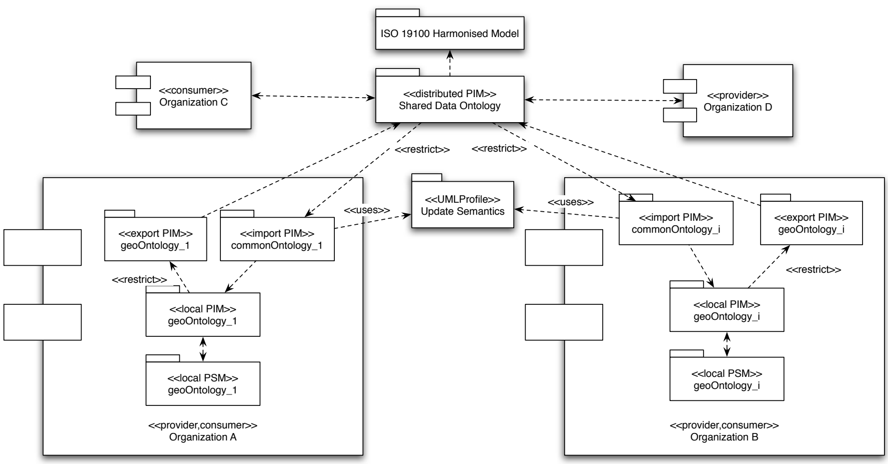


Fig. 1. Proposed four-level ontology framework where UML component symbols in the diagram represent organizations acting in the role of provider, consumer or both as indicated by the assigned stereotypes; UML package symbols are used in the diagram to depict ontologies with each being stereotyped to indicate the role the ontology fulfills in the framework

Of particular interest to this research is the import PIM ontology describing $\Lambda_{p,m}$ since this ontology is to incorporate the semantics of the organization's desired update policy. These semantics are to be incorporated by applying a UML Profile containing extensions (stereotypes, tagged values and constraints) [4–page 711] to the UML that will allow the individual characteristics of the organization's update policy to be specified consistently across all organizations.

An organization's local copy, Λ_p, of Θ conform to an ontology in the role of an import PIM. This ontology is expressed using UML that has been extended

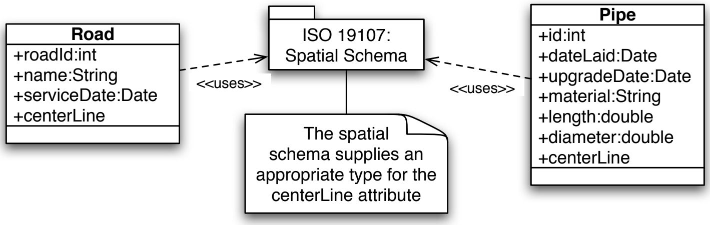


Fig. 2. The ontology for $\Lambda_{1,3,t}$. This ontology fulfills the role of an import PIM.

by the UML Profile for update semantics described in Section 4.1 thereby incorporating an update policy reflecting the organization's business priorities for data synchronization.

Once an update policy has been defined by organization m, execution of this policy involves analyzing the update notifications, $\Delta\lambda_g$, either broadcast by, or requested from, the various organizations contributing components of $\Lambda_{p,m}$ to determine when, within the context of the update policy, a significant difference exists between the local copy $\Lambda_{p,m,t}$ and $\Lambda p, m, t'$. When such a significant difference exists, an update execution plan is formulated to govern the steps by which $\Lambda_{p,m,t}$ is updated. The proposed system is viewed as an Ontology Driven Information System [3] since an ontology with the embedded update policy "plays a central role in the system's lifecycle" [*op cite*, page 16].

To illustrate the system in practice, a use case is now described to demonstrate update notifications and policies in the envisaged distributed environment.

Two local government authorities, P_1 and P_2, each provide a data set: P_1 provides λ_1 containing information about roads; and P_2 provides λ_2, containing information about underground water pipes. At time t a consumer organization, C_3, consumes both P_1 and P_2 to form a local data set $\Lambda_{1,3,t}$ which provides mappings between roads and water pipes based on their geospatial locations.

The concepts of update notifications and update policies are examined in more detail in Sections 3 and 4 respectively. In each, scenarios associated with the above use case are given to further illustrate these concepts.

3 Update Notifications

An update notification $\Delta\lambda_{g,t'}$ is a message from a data provider P_n to any consumer C_i that describes changes to data set $\lambda_{g,t'}$. In simple terms the following tasks are of interest:

- when the data is changed, create a description of this change (Sections 3.1); and
- provide consumers with access to this description (Section 3.2).

To illustrate the concept of an update notification consider the following scenarios within the context of the use case described earlier.

Scenario 1

Organization P_1 modifies λ_1 to:

- introduce two new road centre lines as a consequence of a new subdivision. This introduction also requires an existing road centre line to be altered;
- modify an existing centre line for a segment of road that has been realigned to remove a bend that was causing serious traffic accidents.

By prior arrangement, P_1 immediately sends $\Delta\lambda_1$, a description of these changes, to C_3.

Scenario 2

Organization P_2 modifies λ_2 to:

- alter the export PIM ontology defining λ_2 by deleting the attribute called 'diameter' and introducing a new attribute called 'comment';
- modifying the name of an existing attribute called 'id' to become 'pipeId'.

C_3 retrieves $\Delta\lambda_{2,t'}$ and $\Delta\lambda_{1,t'}$ from each provider's 'Blackboard' (see Section 3.2) as part of the periodic review process in place for updating $\Lambda_{1,3,t}$.

In each scenario, the description of the changes form the content of an update notification and only document characteristics that will be needed to determine whether the changes are significant in terms of the various update policies defined for the relevant distributed data sets. Inaccurate or incomplete update notifications will lead to poor decisions about when to synchronize the relevant Λ_p. Furthermore, subsequent planning and execution of the updates to $\Lambda_{p,t}$ may be inhibited because consumers lack sufficient information to determine what they need to update and how to retrieve and possibly transform the changed data to form $\Lambda_{p,t'}$.

In the case of Scenario 1, content of the update notification would include, for example: the spatial bounding box for each of the new roads; and the identifiers of existing roads that have modified values. The specific details (such as the spatial location of the changed road centre lines) are supplied when the update execution plan is being implemented. Further explanation of the content of update notifications is provided in Section 3.1.

As illustrated by the two Scenarios, update notifications are available either directly by sending the description to one or more consumers (Scenario 1), or indirectly by posting the description in a storage location that all consumers can access when they are interested (Scenario 2). Further explanation of the content of update notifications is provided in Section 3.2.

3.1 Content of Update Notifications

Update notifications describe changes to a particular data set. As shown in Figure 3, information contained within the notification are grouped into those

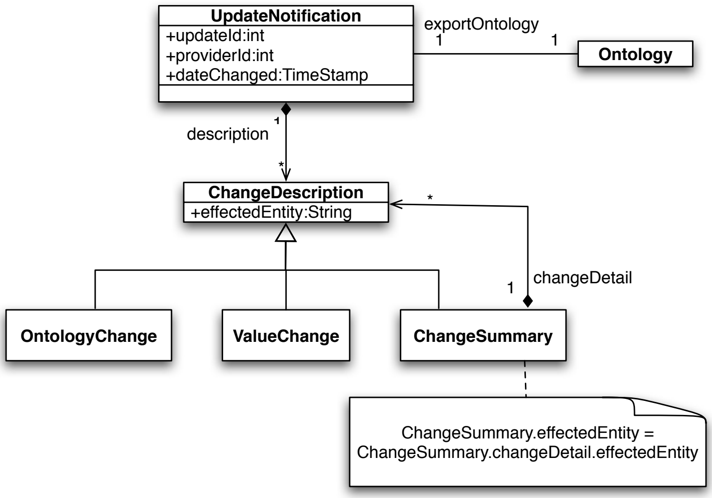


Fig. 3. UML Class diagram of Update Notifications

elements concerned with administrating the notifications themselves and those concerned with describing the changes. Administrative elements of a notification are:

Provider ID which is a unique identifier for each provider within the system so that the consumer knows the source of the notification.

Update ID is specified in sequence by data provider, and is unique to each update generated by this provider. This ID can also be treated as data set version number.

Export PIM Ontology describes provider data set in order to aid the consumer to further identify the provider and subscribed data sets.

Date Changed specifies the date and time when these changes were applied to the data.

Each change description describes one modification to an entity. This description is focussed on either

a value change, described by a statistic that summarizes the changes to values of this class of entity within $\Delta\lambda_{g,t'}$. A value change may involve the insertion, deletion, or modification of a value as illustrated by Scenario 1; or

an ontological change involving the insertion, deletion, or modification of elements (classes, attributes, associations) of the ontology as illustrated by Scenario 2.

More than one kind of change affecting the same entity is aggregated by a ChangeSummary description.

Describing Value Updates. A value update notification describes the content of $\Delta\lambda_{g,t'}$, the changed values since the last update notification at time t. This description is a list of statistics that are calculated for $\Delta\lambda_{g,t'}$ and only these statistics are included in the update notification. Examples of statistics that may be included are the number of insertions, deletions, or modifications, for different classes of entity, and such like. By transmitting only the statistics of the changed values, the transmission of a large amount of data across the network may be avoided if consumers decide to ignore such changes.

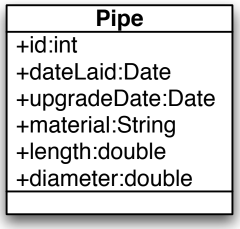
Describing Ontology Updates. Insertions, deletions and modifications to an existing ontology are expressed as an XMI document [13– page 1-31]. All differences described by XMI document must be applied in order so that the model integrity can be maintained. Thus, when a consumer receives an ontology update notification, the consumer can either ignore and discard that notification or address all differences described by the notification and any earlier notifications that were ignored.

Ontology updates may have serious consequences and need to be considered with care. Changes to the ontology correspond to changes to the database, which in turn may impact upon existing database processing software developed by each organization. One strategy being considered involves providers intending to make an ontological change sending consumers an ontological update proposal. This proposal is analyzed by each data consumer in order to assess the desirability of the proposed ontological update. For example, adapting an software application to the attribute name change would be much easier than adapting the same application to the changes that involve attribute deletion. After analysing the update proposal, each consumer sends feedback to the provider indicating whether such a change is desirable. The provider evaluates this feedback while deciding whether to make the ontological change.

3.2 Notifying Consumers of Updates

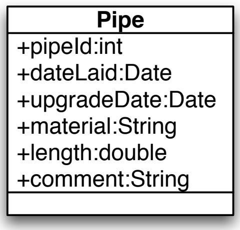
A local copy is updated when the consumer becomes aware that this data differs significantly from that which is available from the provider(s) as they change the available data to reflect modifications to the relevant real world phenomena. Consumer awareness typically occurs in one of two ways: either a consumer periodically checks with the provider(s) for updates, or provider(s) broadcasts update notifications to all relevant consumers whenever the available data has been updated. These two methods are referred to as *Pull* and *Push* respectively with both methods having advantages and disadvantages.

The Push method allows consumer to receive update notifications as soon as they are published by the providers. As the number of consumers increases, however, this method is likely to become impractical. Software such as Microsoft Windows and Norton Antivirus are used by a great many people around the world and having the providers of this software broadcast update notifications to every consumer may be impractical. The situation is exacerbated by providers




```
<xmi:XMI version="2.0">
   ...
  <UML:Class xmi.id="S.11" name="Pipe">
    <UML:Attribute xmi.id="S.19" name="id">
       ...
    </UML:Attribute>
    <UML:Attribute xmi.id="S.22" name="diameter">
       ...
    </UML:Attribute>
  </UML:Class>
   ...
</xmi:XMI>
```

(a) oldPipe.xmi




```
<xmi:XMI version="2.0" ...>
   ...
  <UML:Class xmi.id="S.11" name="Pipe" ...>
    <UML:Attribute xmi.id="S.19" name="pipeID">
       ...
    </UML:Attribute>
    <UML:Attribute xmi.id="S.25" name="comment">
       ...
    </UML:Attribute>
  </UML:Class>
   ...
</xmi:XMI>
```

(b) newPipe.xmi

```
<xmi:XMI version="2.0" xmlns:UML="org.omg/UML"
                       xmlns:xmi="http://www.omg.org/XMI">
    <!-- deletion of "diameter" attribute -->
    <difference xmi:type="xmi:Delete">
        <target href="oldPipe.xmi#S.22"/>
    </difference>

    <!-- addition of "comment" attribute -->
    <difference xmi:type="xmi:Add" addition="S.25">
        <target href="oldPipe.xmi#S.11"/>
    </difference>
    <UML:Attribute xmi.id="S.25" name="comment" ...>
        ...
    </UML:Attribute>

    <!-- change of "id" attribute -->
    <difference xmi:type="xmi:Replace" replacement="S.19">
        <target href="oldPipe.xmi#S.11"/>
    </difference>
    <UML:Attribute xmi.id="S.19" name="pipeID" ...>
        ...
    </UML:Attribute>
</xmi:XMI>
```

(c) diffPipe.xmi

Fig. 4. An example of an ontological update. The initial ontology for a pipe feature (a) is modified to become (b) by the following operations listed in (c): 1. the attribute ‘diameter’ is removed; 2. the attribute ‘comment’ is added; 3. the name of the attribute ‘id’ has been updated to ‘pipeID’. Note that for illustrative purposes only the relevant XMI is shown.

who frequently publish update notifications. The Pull method solves these problems, since this method allows each consumer to configure a strategy of checking for update notifications which reflects the organization's unique business rules for data synchronization. However, the drawback to the Pull method is that consumers may not get critical updates in a timely fashion.

In this paper, both Pull and Push methods are considered, and the proposed system architecture allows both to work cooperatively to achieve the best result. Those consumers with low to moderate demands for up to date data may use a Pull method while organizations with high demands will be prepared to pay providers to be included in a Push method of receiving notifications.

Even though update notifications may be broadcasted to all subscribers, these updates might not be executed or might even be discarded because of the update policies (see Section 4) defined at the consumer end. In order for these consumers to retrieve for non-updated updates in a later stage, all update notifications need to be stored and maintained by both data providers and consumers for a configurable period of time. The storage space is analogue to *Blackboard*, which allows interaction and information can be exchanged indirectly and asynchronously between different organizations. Figure 5 shows the whole process described above.

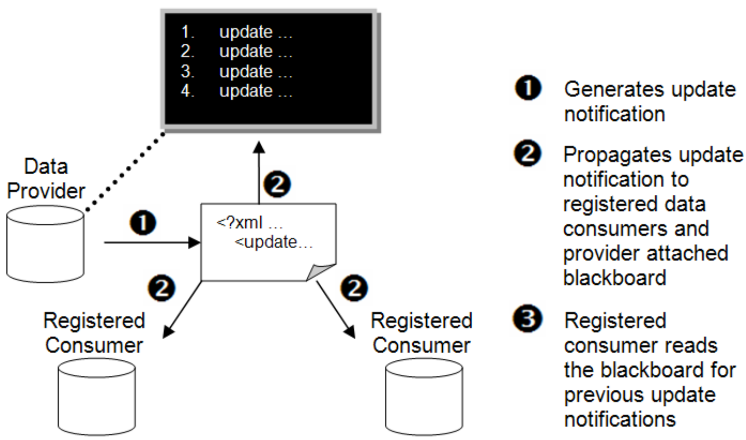


Fig. 5. Propagation of update notification

However, if more and more update notifications are displayed on the Blackboard, the storage space will eventually run out. In order to avoid this problem, a *Version control* mechanism is used to control the version of all subscribers' data sets by using the "Update ID" field in the notification message to maintain the information about which update has been successfully extracted and applied by which subscriber. Once an update has been extracted by all subscribed consumers, or has been displayed for a certain period of time, both the update notification and $\Delta\lambda_{g,t'}$, the changed data values, will be removed from the Blackboard.

If a notification has been removed and not all consumers have executed that update, then those consumers need to either compute the changes by themselves by comparing their local copies $\Lambda_{p,m,t}$ with $\Theta_{t'}$ or simply update $\Lambda_{p,m,t}$ regardless of the quantity and nature of the changes made in the period of time between t and t'.

4 Update Policy

Consumer defined update policies consist of a series of rules that determine the circumstances under which a data consumer C_m will update its local data set $\Lambda_{p,m,t}$. These rules primarily reflect the higher level business requirements and constraints unique to each consumer.

Once update policies are defined by consumers, they are incorporated into primarily import PIM ontologies using UML extensions defined by a profile for update semantics described in Section 4.1. Providers may wish to indicate immutable elements (for example identifiers) of an export PIM ontology: therefore, such a stereotype is also provided within the profile for Update Semantics.

An update policy is used to evaluate the significance of a collection of update notifications to a particular consumer. When deemed significant, notification of these updates will initiate the updating of $\Lambda_{p,m,t}$. Informally this may be described by the following expression:

$$updatePolicyFunction(\Lambda_{p,m,t}, \Delta\Lambda_{p,m,t'}) \mapsto \{true, false\}$$

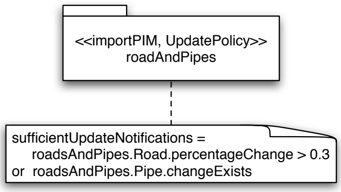
where the *updatePolicyFunction*() comprises an OCL Expression involving tagged values defined by the UML profile for update semantics.

To illustrate the concept of an update policy, consider Scenario 3 within the context of the use case described earlier in Section 2.

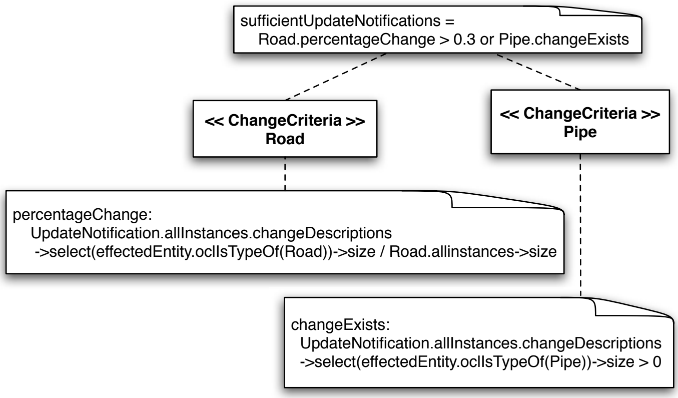
Scenario 3

> Based on the associated cost for update and the importance for each data set, C_3 only updates $\Lambda_{1,3,t}$ when either at least 30% of the data values in $\lambda_{1,t}$ have been changed or $\lambda_{2,t}$ is changed in any way. The import PIM ontology shown in Figure 2 is enhanced to incorporate this update policy as shown in Figure 6.

Sometimes, C_3 can neglect some minor updates in order to save network bandwidth and time or reduce the associated cost while the degree of inconsistency between $\Lambda_{1,3,t}$ and $\lambda_{1,t}$ or $\lambda_{2,t}$ is acceptable. However, under this situation, it needs extra care due to the accumulation effect of minor updates. For example, in this case, if 25% of data has been updated in $\lambda_{1,t+\Delta t}$ at $t + \Delta t$ and another 10% of data has been updated at t', then C_3 should update $\Lambda_{1,3,t}$ when the second update notification generated by P_1 has been received, since 35% of data in total has been updated.



(a) The package enclosing the ontology with the UpdatePolicy stereotype applied.



(b) The content of the package with stereotypes and tagged values

Fig. 6. The import PIM ontology shown earlier in Figure 2 enhanced to incorporate the update policy described in Scenario 3

4.1 UML Profile for Update Semantics

A complete description of the UML profile for Update Semantics is beyond the scope of this paper: therefore, selected stereotypes and tagged values defined by this profile are described to illustrate the intent.

«**UpatePolicy**»
This stereotype is applied once to a UML package symbol within which is defined an ontology (typically either an export or an import, PIM ontology). Associated with this stereotype are the following tagged values:

sufficientUpdateNotifications is an OCL expression which when true indicates that the available Update Notifications $\Delta\Lambda_{p,m,t'}$ justifies the updating of $\Lambda_{p,m,t}$.

updateSchedule which is assigned a value indicating the frequency with which $\Lambda_{p,m,t}$ is to be updated regardless of what has changed. Values that may be assigned to this tag are: `none`, `daily`, `weekly`, `monthly`, and `yearly`.

significantUpdate which is assigned the value of the following OCL expression: sufficientUpdateNotifications or (updateSchedule <> 'none'). When true, $\Lambda_{p,m,t}$ is to be updated.

Use of the updateSchedule tag indicates the consumer adopts a *pull* method while use of the sufficientUpdateNotification indicates the adoption of a *push* method for update notification.

«**ChangeCriteria**»

This stereotype is applied to classes for which any change may initiate the updating of the data set. This stereotype has the following tagged values to characterize the changes:

numberOfChanges an integer value indicating the number of change descriptions applicable to instances of this class or to this class itself.

percentageChange a real value between 0 and 1 indicating what percentage of the instances of this class in $\Lambda_{p,m,t}$ have been changed in $\Lambda_{p,m,t'}$.

changeExists a boolean value indicating that this class or instances of this class have been changed.

priority an integer value between 1 and 9 indicating the importance of the instance of this class. The lower the priority value, the higher the importance.

«**SpatialExtentCriteria**»

Like «ChangeCriteria», this stereotype is also applied to classes for which any change may initiate the updating of the data set: however, these changes are characterized by occurring to values within some specified spatial extent as defined by one of the following tagged values:

boundingBox the minimum bounding rectangular extent within which any change notifications or part thereof are regarded as potentially significant.

spatialBuffer For example all changes to pipes within 10 kilometers of a specified road centerLine.

«**Immutable**»

This stereotype will be applied to classes and or attributes of a class within an export PIM ontology to convey that instances of this class or values of this attribute will never be changed by the provider of λ_g.

The tagged values are to be used within OCL constraints to convey update policies unique to each consumer.

5 Conclusions and Future Research

In this paper is described ongoing research into facilitating the synchronization of a consumer's local copy of a distributed data set. The proposed solution involves each consumer incorporating into their platform independent import ontology of the local data set, defined using UML, a policy that governs when the data set is to be updated. The semantics of the update policy are documented within the ontology as OCL constraints expressed in terms of the stereotypes and tags defined by the proposed UML Profile. The platform independent import ontology is one of four in the proposed four level framework which results from merging the OMG's MDA [8] with Seth and Larson's [9] notion of a five-level schema framework for distributed systems.

The platform independent import ontology with an embedded update policy is represented within an XMI document which is processed by software, an update agent, that implements the embedded update policy. Over time, this update agent listens for relevant update notifications from providers of the distributed data and evaluates the changes described either in isolation of, or in collaboration with, other update agents to determine when the local copy of a distributed data set is to be synchronized.

As the research continues, the UML profile for documenting the update policy will be refined to address issues that arise from implementation of this system as briefly described in the next section. Preliminary results suggest that the approach is feasible: however, much more experimentation is necessary to evaluate the efficacy of the solution proposed.

The system described in the paper is currently being implemented using Agent technologies such as OPAL [10], which incorporates an implementation of the Java Agent Services (JAS) specification [11], and two implementations of the Open GIS Consortium Web Feature Service (WFS) Specification [12] (Intergraph and geoserver, an Open Source project). Different versions of various data sets have been kindly made available by the Christchurch City Council to be used to create meaningful and realistic sequences of update notifications from two implementations of the WFS specification. Using this evolving implementation as a testbed for the research described here is the next phase.

An issue yet to be fully explored is the use of spatial operators in the definition of update policies. Consider, for example, the following scenario.

> C_3 would like to update $\Lambda_{1,3,t}$ when there are at least 5 changes to pipes within 10 kilometers of a specified road center line.

Creating an update policy with such a constraint expressed using OCL is difficult because, by default, OCL does not support spatial operations, such as 'Contains' and 'Overlaps'. Expressing such operators within OCL remains an open problem.

Acknowledgments

The research described here has been funded by the Foundation for Research, Science, and Technology subcontract number UOOX0208.

References

1. L. McKee, K. Buehler. The Open GIS Guide, *Open GIS Consortium, Inc*, Wayland, MA, 1996.
2. Y.A. Bishr, H. Pundt, W. Kuhn, M. Rdwan. Probing the concepts of information communities - A first step toward semantic interoperability. *Interoperating Geographic Information Systems*, pp. 55-70, Kluwer, Norwell, MA, 1999.

3. Frederico T. Fonseca and Max J. Egenhofer. Ontology-driven geographic information systems. In *Proceedings of GIS '99: Proceedings of the 7th ACM international symposium on Advances in geographic information systems*, ACM Press, New York, NY, 1999.
4. Object Management Group (OMG). UML 2.0 Superstructure Specification. October 8, 2004 http://www.omg.org/technology/documents/modeling_spec_catalog.htmi, 2004.
5. F.T. Fonseca. Users, Ontologies and Information Sharing in Urban GIS. *ASPRS Anual Conference*, Washington D.C., 2000.
6. S. Cranefield, M. Purvis. UML as an ontology modelling language. In *Proceedings of the Workshop on Intelligent Information Integration, 16th International Joint Conference on Artificial Intelligence* (IJCAI-99), 1999.
7. International Standards Organisation / Technical Committee 211 Geographic Information - Rules for application schema, ISO, 2005
8. Object Management Group (OMG) MDA Specifications. Available from: http://www.omg.org/mda/specs.htm.
9. A. Sheth, J.A. Larson. Federated Database Systems for Managing Distributed, Heterogeneous, and Autonomous Databases. *ACM Computer Surveys*, 22(3), 1990.
10. M. Purvis, S. Cranefield, G. Bush, D. Carter, B. McKinlay, M. Nowostawski and R. Ward. The NZDIS Project: an Agent-based Distributed Information Systems Architecture, In R.H. Sprague, Jr. (ed.) *CDROM Proceedings of the Hawaii International Conference on System Sciences (HICSS-33)*, IEEE Computer Society Press, 2000.
11. Francis G. McCabe. Java Specification Request 87: Java Agent Services Available from: http://www.jcp.org/en/jsr/detail?id=087.
12. Panagiotis A. Vretanos Web Feature Service Implementation Specification OpenGIS project document number OGC 04-094, version 1.1.0 3 May 2005. Available from: https://portal.opengeospatial.org/files/?artifact_id=8339.
13. Object Management Group. XML Metadata Interchange (XMI) Specification, Version 2.0. Available from: http://www.omg.org/mda/specs.htm, Document: formal/05-05-01.

Purpose-Driven Navigation

Neeharika Adabala and Kentaro Toyama

Microsoft Research India,
Bangalore, India
{neeha, kentoy}@microsoft.com

Abstract. Navigation in the context of GIS (Geographic Information Systems) is associated with a sequence of pan and zoom operations that lead to a specific destination. Navigation, in this context, leads a user to an *a priori* desired destination. There are cases, however, when users may not have a clear idea of a single destination. In this work, we propose richer navigational schemes by augmenting the concept of navigation to be broader than the goal of arriving at a single destination. This is achieved by identifying typical patterns of map use and the purposes behind such patterns, and defining corresponding navigational schemes. The proposed technique enables what we call *purpose-driven navigation* of maps, e.g., "scan region" or "explore neighborhood". We present example scenarios that demonstrate the benefit of purpose-driven navigation.

1 Introduction

It is widely acknowledged that moving data into the semantic domain greatly enhances the value as it enables richer and more natural queries than traditional keyword-based queries [1],[2]. The main disadvantage of traditional keyword queries is that the search results are dictated by the choice of keywords rather than the concept behind them. Similarly, the ability to define meaningful, purpose-driven interactive techniques with an information system can result in more compelling user experiences.

One of the key interactions with GIS is navigating in the virtual geographic environment. Navigation in GIS is usually defined as a point-to-point process. Such navigation is implemented as a sequence of pan and zoom operations that lead from start to destination location. This can be perceived as a fundamental form of navigation. However, not all interactions with geographic environments involve navigating from point to point. In some cases navigation may be motivated by goals other than just reaching a destination – for example, the purpose of navigation may be to explore a neighborhood to discover what prominent landmarks are present. Currently, support for such rich navigational schemes in existing GIS is absent. We address this gap in existing GIS by introducing the concept of parameterized, purpose-driven navigational schemes. This approach enables us to associate semantics with navigation.

The rest of the paper is organized as follows, the next section describes motivating scenarios that are not easy to solve with existing navigational support

M.A. Rodríguez et al. (Eds.): GeoS 2005, LNCS 3799, pp. 227–233, 2005.

in GIS. We go on to discuss the issues that render the problems described difficult for the current systems. Our solution is then proposed in section 4 where purpose-driven navigational schemes are identified and parameterized. Section 5 gives details on how the proposed technique enables easy implementations of the scenarios described in section 2. We present conclusions in section 6.

2 Motivating Scenarios

In this section, we describe some scenarios that prompted us to develop a scheme for navigation based on intent. A GIS that includes road maps and satellite images of geographic regions is the focus of the scenarios presented. We present scenarios where people repeatedly create similar navigation patterns because of the purpose behind their interaction with a GIS.

- **Scenario 1:** Alice is considering relocating to a new neighborhood and would like to explore the region to know the locations of grocery stores, shopping malls, schools, etc. In this case, she would like a navigation scheme that automatically guides her through the neighborhood amenities in the context of her relocating to the neighborhood. She most probably will check more that one location to help her decide which location best suits her requirements.
- **Scenario 2:** Ben accompanies a friend to a location and later wants to find the address of where they had visited. All Ben can remember, however, is the appearance of the location (some trees and the kind of buildings in the region) and some general idea of where the location may have been geographically - like the southern region of Washington State. In this case he would like to be able to mark out a particular region on the map, that does not necessarily correspond to regions as defined in the data set, and scan this region till he finds the place that most resembles the place visited.
- **Scenario 3:** Charlie would like to view details in a wide region in the virtual geographic space and try to identify places of interest to him to tour when he visits the region. A brute-force scanning of the whole region is time consuming. Also, he may have certain regions that he knows he wants to visit.
- **Scenario 4:** Dana is planning a long drive and wants to choose some "interesting" roads to follow, she would like to know what is present on the road sides - is it scenic/mostly commercial/mountaineous, etc. She would like to view the road in the highest zoom level possible and follow roads, rather than directly "flying" to target destinations.

In all of these scenarios, navigation is complicated because there is no single start and destination location. The issues in the context of each scenario are described in the next section and possible solutions are outlined.

3 Issues and Solutions

User interaction with GIS is compelling if the navigation scheme is aware of the context of the navigation, and can automate the process of interaction using the knowledge of the user intent.

In the case of scenario 1, Alice has to find all the grocery stores in the neighborhood, in the current system she would have to enter possible names of stores and see if they exist in the neighborhood. In a semantic database, the task is simplified as all grocery stores can be identified by the labels associated with the data [4]. However, when Alice is exploring a neighborhood for relocating, she will have to make a list of all the amenities of interest to her, just to be able to query for them. This is cumbersome. The issue here is that in the current systems there is no way to point to a geographic region and indicate that our interest is to explore the neighborhood.

The solution is to define a purpose-driven navigational scheme that incorporates the concept of "explore neighborhood". This generic scheme is then parameterized to support the requirements of the user by associating with it a list of amenities to be explored in the neighborhood.

In the above scenario, associating the scheme of navigation with the user's intent enables identification of destinations. Once destinations are identified, existing point-to-point navigation with the additional information that the location specified by the user is the start point for all navigation provides the solution for "explore neighborhood".

The second scenario of Ben finding the place he had visited is more complex. The key issue is the inability to identify the destination of the navigation till one actually navigates through the relevant geo-spatial data. This is unsuitable for the existing concept of navigation that requires a well-defined destination.

The solution to the problem is to enable the Ben to select an area on the map and indicating that a scan of the region is to be performed. The system performs a zoomed-in scan of the region and the Ben can stop the scan when he recognizes the location.

In the third scenario, since the area to be toured is large, a scan-based solution will be cumbersome for Charlie. Also, he has additional information about locations in the area that he definitely wants to visit. The navigation process should therefore incorporate these points in the path it follows. One approach to implement a navigation for this scenario is to enable Charlie to mark his locations of interest as "perch locations". The navigation process then makes forays from these locations and enables identification of other locations of interest close-by.

In all of the above scenarios, there is no restriction on the path followed during the navigation. However, in scenario 4, Dana would like to follow roads; the typical navigation that flies or jumps directly to destination is not applicable here. We address this limitation in navigation by associating a property that enables a choice of type of path during navigation.

4 Generic Navigation Schemes

In this section we generalize the ideas presented in the previous section and describe generic navigation schemes; depending on the purpose of navigation we create a suitable parametrization of the navigation to enable user customization of the generic navigation scheme.

The concept of associating type with processes as opposed to only objects can explained in the context of the Object Process Methodology (OPM) [3] developed to enable challenging systems in which structure and behavior are intertwined. OPM is a holistic approach, to study and develop a system, which integrates the object-oriented and process-oriented approaches into a single frame of reference. We adopt this holistic view when defining generic navigation schemes.

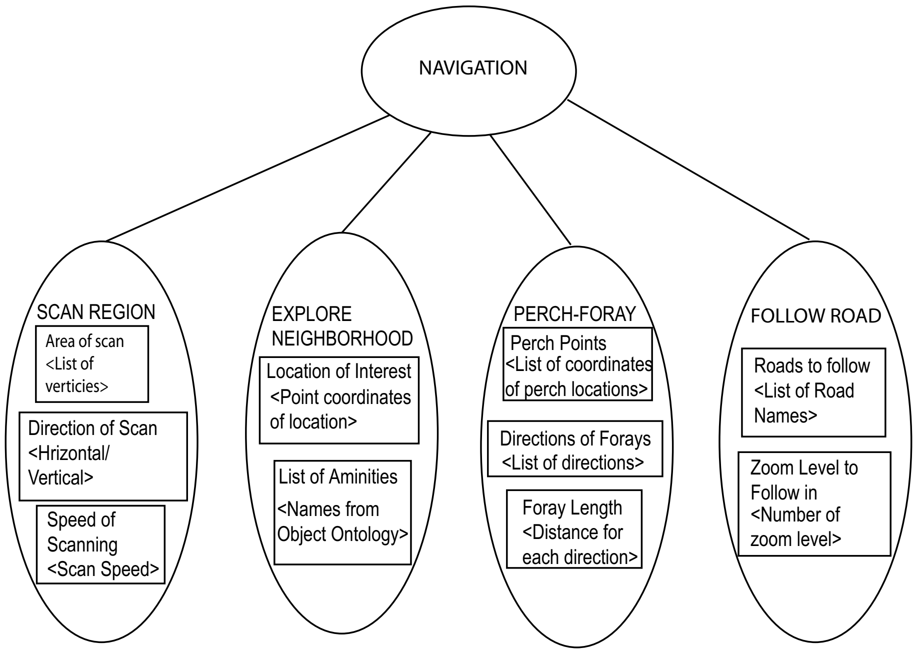


Fig. 1. Generic parameterized schemes for navigation “explore neighborhood”, “scan region”, “perch-foray” and “follow road”

Generic navigation schemes provide the user with a framework to incorporate additional knowledge in the system that automates navigation. The user no longer has to manually pan and zoom to find the required information in the GIS; the navigation system can be further tuned to the user interaction requirements using the additional parametric values defined by the user. Figure 1 outlines generic navigation schemes for “scan region”, “explore neighborhood”, “perch-foray” and “follow road”.

Figure 1 also presents the parametrization of the generic navigational schemes. In “scan region” the list of vertices that define the polygon around the area of interest and the choice of direction and speed of scan are present. In “explore neighborhood” the point of interest around which to explore and the list of amenities that are of interest to the user are given. For “perch-foray” the list of perch locations, the directions of forays and length of forays for each direction constitute the information required to automate the navigation scheme. The list of roads and the level of zoom at which the navigation is to be performed is the parameters in the "follow road" scenario.

5 Example Information Flow During Navigation

In this section we describe the navigation process when it is associated with a generic scheme with parameters described in the previous section.

Figure 2 is an illustration of a GIS with generic navigational schemes is presented. We use ellipses to represent processes and rectangles to represent data. During interaction with the GIS a generic navigation scheme may be represented as the parameter values that the user provides to customize the navigation. Therefore it is represented as data associated with the process of navigation.

When users interacts with this GIS they select a mode of interaction that is either in the form of a text query or a navigation of the data. Users who choose to navigate are given the option of selecting either the manual pan/zoom approach or the richer navigational schema, namely "scan region", "explore neighborhood", "perch-foray" and "follow road". When one of these generic navigation schemes is selected, the user is prompted to give details to set parameters related to the particular scheme. For example, if "scan region" is chosen, they will be asked to provide the vertices of the polygon enclosing the scan region and also indicate the speed and direction in which they would like the area to be scanned, vertical, or horizontal. The navigation system then carries out the task of scanning and displaying the map to the user. Users interact similarly for other navigational schemes by providing the parameters to complete the data required for a particular generic navigation scheme. The system takes over once the parameters are defined and displays a suitable traversal through the geographical environment. (The precise UI for how users answer questions about parameters could be further refined.)

Figure 3 shows a highlight of an example path followed by the navigation system when it explores a neighborhood for places where food can be purchased.

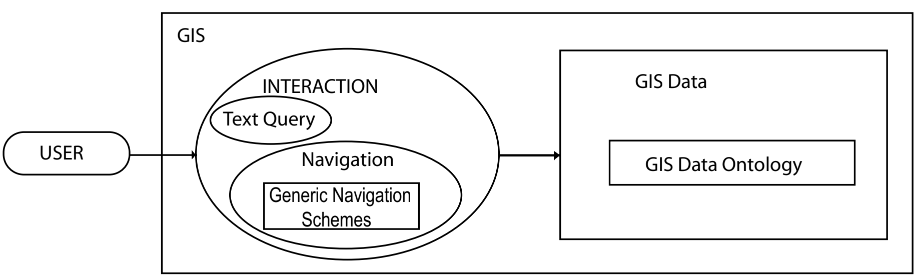


Fig. 2. Schematic representation of location of the generic parameterized navigation in GIS

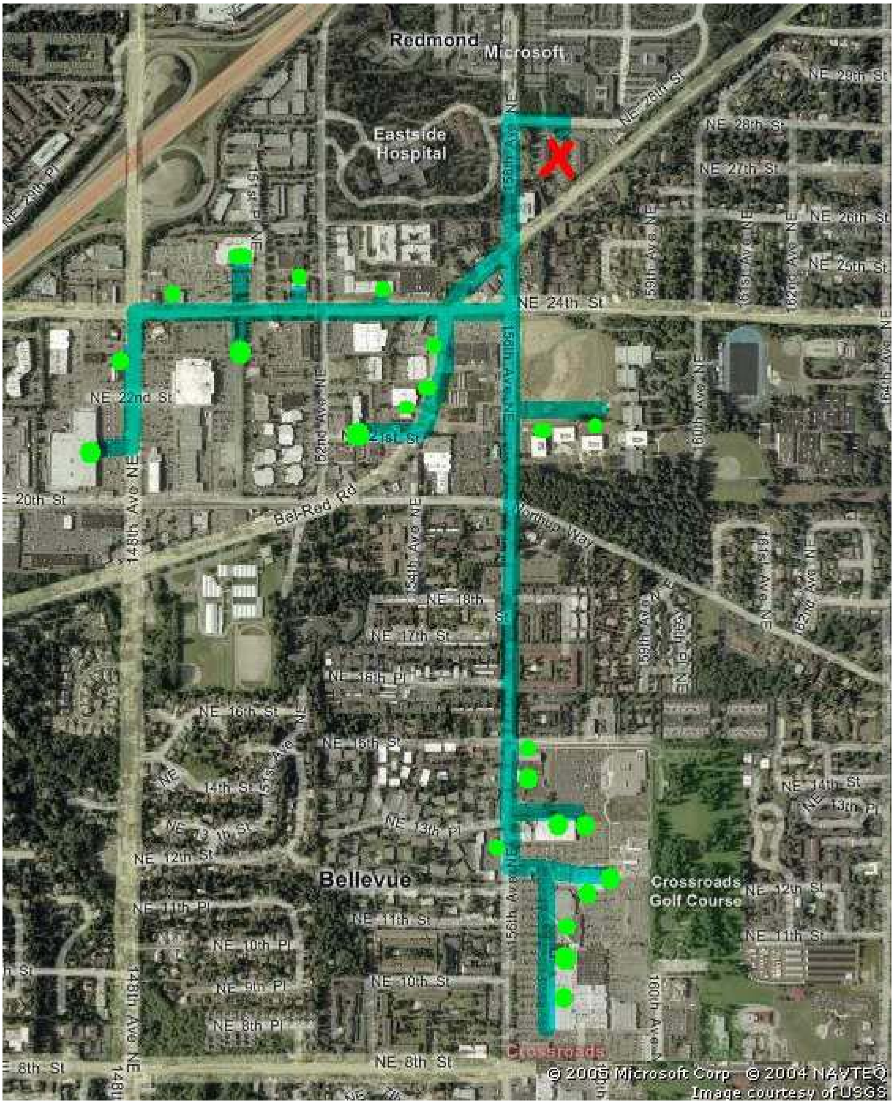


Fig. 3. Example navigation path, of the path traced by the system when the goal of the navigation is to explore all the places where food can be purchased (shown as green spots on the map) in the neighborhood of the location marked by a red x

We show a zoomed out version of the map with the path highlighted. In the actual system, the navigation tool will automatically follow the path highlighted on a zoomed-in version of the region and the user can see more details of the locations being visited by the automated navigation process.

6 Conclusions

We have presented a technique for creating complex navigation schemes in the context of GIS. Scenarios that warrant the use of purpose-driven navigation were described, and we develop suitable parameterized navigation schemes to support rich interaction. Knowledge of navigation purpose was used to automatically identify multiple destinations, and also a scan-based solution was proposed for the case when the navigation process itself leads to the identification of the destination. A "perch" and "foray" based generic navigation scheme was defined for the case of a tourist with partial information on locations of interest. A technique for enabling control of path of navigation was also proposed.

The development of generic navigational schemes described in this paper have general-purpose applications beyond the given scenarios. "Scan region", for example, can be applied in crime fighting where the law keepers may receive descriptions of locations, and navigation allows scanning of the region with reduced manual effort. "Perch-foray" can be used to gather information around a scattered set of locations in a wide region. "Follow road" can be applied by ecologists to follow the course of a river and examine various changes in environment along it. It should also be noted that we can combine these schemes like "follow road" with "explore neighborhood' to obtain richer information from GIS, like the distance in miles for a particular tour of the neighborhood, from which we can infer whether the gas in the car is sufficient to undertake a tour. Thus by defining purpose-driven navigation we are able to add compelling interactions with GIS beyond those of purely manual navigation.

References

1. Tim Berners-Lee, James Hendler, Ora Lassila. The Semantic Web. *Scientific American*, May 2001.
2. Max J Egenhofer. Towards the Semantic Geospatial Web. *ACM-GIS 2002*. McLean, VI A. Voisard and S.-C. Chen (eds.), 2002.
3. Nathan R. Soderborg, Edward C. Crawley, and Dov Dori. OPM-Based System Function and Architecture: Definitions and Operational Templates. *Communications of the ACM*, 46(10): 67-72, 2003.
4. Sebastian Hübner, Rainer Spittel, Ubbo Visser and Thomas J Vögele. Ontology-Based Search for Interactive Digital Maps. *IEEE Intelligent Systems*, 19(3): 80-86, 2004.

Mobile GIS: Attribute Data Presentation Under Time and Space Constraints

Lasse Møller-Jensen

Institute of Geography, University of Copenhagen, Øster Voldgade 10,
Copenhagen, Denmark
Lmj@geogr.ku.dk

Abstract. An ontology-based generalization scheme is presented with specific reference to object attribute data subjected to space and time constrained extraction and presentation. The method is expected to be of value in mobile GIS for providing travelers with additional spatial information while moving. The space constraint is given by the boundary of an object-specific area of information relevance. The time constraint is given by the number of time units available to present the information as a function of the speed of travel and the spatial distribution of objects. An algorithm for changing the duration of the attribute data presentation is presented. A geodata-ontology is used to specify meaningful transitions betweeen levels of generalization.

1 Introduction

The increasingly popular navigation systems based on small mobile computers (PDAs) equipped with a GPS may be viewed as mobile GIS implementations since they contain both road map data with attributes and network algorithms for findings best routes. Viewed as a GIS, however, the functionality is very specialized since the ability to provide other spatial information than what is relevant for the way-finding is rather limited. Several studies have examined the potentials of map reading from mobile devices and how this is different in nature from use of traditional paper maps (see f.ex. [16]). To date, however, few papers have described efforts to further utilize the potential ability of the mobile GIS to provide elaborate attribute-type information about any object that a user interacts with while moving around. Frank et al see the process as a transformation of the traditional GIS into a Location Based Service by providing mobility, distributiveness and egocentric awareness [5]. They further describe a system that enables a user carrying a PDA equipped with a GPS and an orientation sensor to get access to the attribute data stored for a feature of interest by standing and facing it. The general problem of selecting the physical object for which to receive information in a given spatial context (equal to the human action of pointing) has been examined by Egenhofer [1]. Current solutions are still mostly of an experimental nature and assume availability of special equipment such as directions sensors or advanced pointing devices.

It is reasonable to conclude that the potentials of the GPS-PDA to act as an advanced location-based information system are acknowledged but still not examined in-depth. Nevertheless, some systems restricted to confined locations such as cultural

M.A. Rodríguez et al. (Eds.): GeoS 2005, LNCS 3799, pp. 234 – 241, 2005.

or entertainment sites have been implemented in practice. IST describes how GPS-enabled city guides are already used in the travel industry and argues that especially the "culture and history travel market" will be able to benefit from a strategy towards location-sensitive presentation of object information [7].

2 Presenting Attribute Data while Moving

The focus of the current paper is the extraction and presentation of GIS-attribute data while moving. Furthermore, focus is put on the situation where a) the user cannot interact directly with the device or point to any location (e.g. because he or she is engaged in driving) and b) the user wants relevant attribute information to be presented more or less continuously while moving. This obviously creates a need for rethinking the traditional view of the attribute database as a container of numbers or small bits of text that may be read of the screen with no time limit given for the reading process. A relevant type of information to be considered here is sound recordings of a given duration attributed to each object. Furthermore, it is necessary to address the issue of generalization with focus on the attribute data.

The movement of the GIS-PDA during the process subjects the extraction and presentation of each bit of information to space and time constraints. The primary objective is to present the traveler with timely and complete information within the time frame given by the travel speed. The time constraint can be viewed as the number of time units available to present the information as a function of the nearness of neighboring features, the speed traveled by the audience and possibly some stated preferences concerning the desired level of information. It is important that the degree of data completeness is known. At constant speed, all information related to a specific generalization level at a given location should be presented – not a random selection. It is therefore necessary to reduce the duration of the information message if the speed and object density do not allow presentation of all information.

The space constraint is given by the boundary of the area within which the information about a specific object is relevant to a specific audience. The attribute presentation should therefore take place within this zone. In some cases, depending on the travel speed, the time required for presenting the information could take up a travel distance that far exceeds the zone of relevance. A reduced duration of the information message – if possible – is required in this situation also.

Three approaches to data generalization aiming at reducing the amount of attribute information for presentation are described below.

Selecting a more generalized data layer. Using this strategy, the ability to present attribute information on different levels of generalization is based on the existence of several data-layers within the mobile GIS. These must be of varying levels of generalization and all cover the area of movement. Traditionally, the map generalization process involves a number of techniques for changing the appearance of graphic objects in the map, e.g. simplification, enlargement, merging and selection [7]. The aim is to produce a map that is targeted to a specific presentation scale and possibly also to a specific purpose and user group. A number of methods for this have been described in literature but mainly targeted towards traditional paper maps. In the digital domain, Kulik et al [8] describe an algorithm for the generalization of line segments for use with mobile devices. This approach is based on – or driven by – a formalized ontology. According to Weibel, a main difference between the conventional and the digital context is that, "in digital systems, generalization can affect directly the map data and not the map graphics alone" [14]. Also with reference to digital systems, Weibel & Dutton argues that generalization has assumed a wider meaning as "a process which realizes transitions between different models representing a portion of the real world at decreasing detail, while maximizing information content with respect to a given application" [15]. The main objective of this model generalization is controlled data reduction for various purposes and serves the purpose of deriving datasets of reduced accuracy and/or resolution. It is also stated by Weibel & Dutton that the generalization process could be replaced by a strategy for producing multi-scale databases that "integrate single representations at fixed scales into a consistent multiple representation" [15]. In this way, the different generalization levels are constructed initially in the database.

Selecting a subset of objects. A suitable generalization may also be obtained by selecting a set of objects from the detailed map. This strategy is actually a part of the traditional methods for map generalization. It is treated separately in this context because it requires only one data layer within the mobile GIS. Furthermore, it requires the existence of attribute data that characterizes the individual object in terms of importance within a given thematic domain. Examples could include buildings that are interesting in the context of a specific historical period. Also objects that share a certain property or functionality could be selected. The result is a dynamically created data layer with a reduced amount of attribute data. This generalized data layer is created entirely based on attribute values and the location of the objects is therefore irrelevant for the selection process.

Selecting object class information. The above described generalization strategies are basically seeking to identify a suitable set of existing geo-objects within the mobile GIS. This is in accordance with the traditional role of the GIS as a provider of information concerning specific objects linked to specific locations. An approach with a different angle will be discussed briefly here. In order to achieve a required reduction in data a shift is made from object attribute data to object *class* attribute data. In other words a shift is made from location-specific information to 'encyclopedia-type' information. The object class attribute data are not linked to a specific geographic location. It consists of information that is common to a subset of objects within a GIS-layer no matter their individual location.

The described generalization strategy is seen as particularly useful in a situation where a subset of objects with similar properties within a given domain is located in a spatial cluster. Take as an example a set of trees of a specific species located close to each other. This will result in a situation where the information relevance areas of the objects are partly overlapping indicating that they may be perceived as a group when traveling through the area. In this case a pre-established object class containing general information about the tree species may be applied. This approach goes somewhat beyond the traditional definition of a GIS and it would require additional data structures to implement an object model that includes generalized class information. This is further discussed in the next section on geodata ontologies.

2.1 A Geodata Generalization Ontology

A mobile GIS to be used on the move could apply one or more of these generalization strategies for adjusting the amount of attribute information to given time constraints. In many cases, however, geo-data of different generalization levels do not constitute meaningful substitutions for each other even though they cover the same area. The adjustment process requires that knowledge about meaningful transitions is present within the system. The set of "legal" vertical moves between data layers of different generalization levels is regarded as an ontology of the system. The term ontology is used here to denote a formal specification of the concepts and relationships that can exist within a certain domain and is able to capture aspects of the semantics of this domain.

The use of ontology in software systems has been proposed in several studies, primarily within the field of information processing. Raper states generally that the new generation of digital geo-representations makes it possible for each geographic information scientist to design their own ontologies for the task at hand [11]. Rodriques & Egenhofer describe the use of ontology comparison in the process of retrieving and combining information that resides in different repositories to identify any differences in data definitions [13]. Fonseca *et.al.* argue that ontologies support the creation of conceptual models and help with information integration and propose a formal framework that explains a mapping between a spatial ontology and a geographic conceptual schema [3]. Moreover, Fonseca et al. describe a comprehensive ontology for geographic information aimed at improving data integration at different levels of details – a process that also involves object generalization and specialization [4]. Another example of ontology supported data extraction is provided by Møller-Jensen who uses ontology information represented in semantic networks to predict the textural pattern of urban objects in satellite images for semi-automated classification purposes [10].

In the present context of mobile GIS, the ontology is used as an active system component that provides guidance for the generalizing process regarding the attribute data. The process of continuously selecting and presenting relevant data layers in order to comply with the given time and space constraints is based on an examination of the ontology properties. These ontology properties are conveniently defined using semantic networks defined as knowledge representation schemes involving nodes and links - the nodes constituting objects and the links constituting relations between objects [6],[12]. In the current context the ontology objects represent data sets included in the system – either existing, static GIS layers or data sets that may be created dynamically by an attribute selection process. The object relations of the generalization ontology must include the following types: a) *is_a_spatio-thematic_generalisation_of* and b) *is_a_conceptual_generalisation_of*. Type a) relations exist between two ontology objects that represent data sets and indicate that these may be substituted by each other during the generalization process. Type b) relations exist between a sub-object that either represents a data set or a class definition and a super-object that represents a class definition. They indicate that a generalization can be achieved by applying the information associated with the super object.

The set of vertical moves that is meaningful and relevant to a specific user may be seen as an ontology that exists in parallel with the system ontology and reflects the specific thematic preferences of the user. Successful application of a mobile information system requires a high degree of similarity between the system ontology and the user ontology. Stated more informally, the information presented on various generalization levels should be relevant and meaningful to the specific user. This could be achieved by allowing the user to define properties of the system ontology given a specification of the data that is included in the system, or, alternatively, by allowing the user to choose between different pre-defined and theme-specific ontologies.

3 Algorithm and Prototype

This chapter describes the proposed algorithm for the process of adapting the attribute data presentation to a given speed. While it may be possible to identify strategies for a purely graphical 'on the fly' simplification process based on a set of rigid rules, e.g. for removing close lying vertex points, a similar approach is not possible with attribute data. Any process aiming at automatically shortening text messages would be of a completely different and complex nature and subject to a number of problems especially if the attribute information is taking the form of sound recordings.

In the proposed algorithm, therefore, the dynamic generalization process is equal to the process of choosing between data sets of different levels of generalization. Such datasets – generalized in accordance with the guide lines discussed above and providing a data hierarchy that is compliant with the user ontology - are expected to be provided in a separate work flow. An alternative development path, discussed e.g. by Kwan & Shi [9] is the application of wireless systems that streams the necessary data to the PDA in real time, responding to specific data requisitions from the PDA.

The discussion in chapter 1 and 2 is formalized by making the following definitions:

The *R-space* (relevance space) is defined as a static buffer zone surrounding an object within which attribute information for the object is of relevance to the moving audience. Visibility analysis based on digital elevation models and line of sight analysis may be applied in a more advanced stage to derive the zone. In some special cases it could be relevant also to focus on other properties than visibility, such as the spatial extension of sound or smell from an object. For the current system, however, standard fixed-width buffer zones have been created around each object to provide an indication of the R-space.

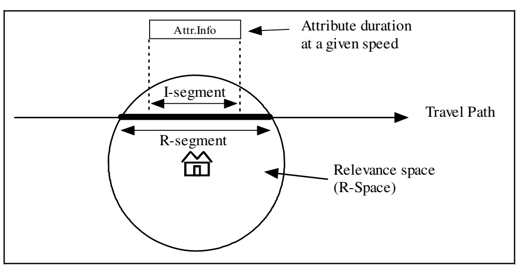


Fig. 1. I-segment indicates the duration of the message while R-segment indicates the section of the travel path where the information is relevant

The *R-segment* (relevance segment) for an object is defined as the projection of the R-space onto the path that the traveler is currently following, (see figure 1).

The *I-segment* (information segment) for an object is defined correspondingly as a segment of the current path with the following properties:

- the halfway point is equal to the halfway point of the R-segment of the object
- the length is equal to the distance covered during the time it takes to present the attribute information given the current level of speed.

The I-segment must therefore be computed dynamically based on the current speed of travel, see also figure 1.

Overlapping I-segments at a given speed indicate that there is not enough time to present all attribute information at this level of generalization. In some situations, the I-segment could be considerably longer that its corresponding R-segment. This would indicate a situation where the traveler would have "passed the object long ago" when the potential attribute presentation ends. It seems reasonable to either suppress the presentation in this situation even if there is sufficient time for the presentation or alternatively to select a more generalized level if available. The latter strategy is implemented in the current prototype.

Following the discussion above, three strategies for attribute data generalization are considered in the prototype. Strategy 1 and 2 handle implementation of the spatio-thematic type of generalization while strategy 3 handles the implementation of the conceptual generalization:

1. Generalize by selecting a subset of all objects. The selected objects should be characterized as important within a specific thematic domain.
2. Generalize by moving to a previously established less detailed GIS-layer with thematically coherent attribute information as defined by the system ontology.
3. Generalize by applying information contained in a thematically coherent object class, or – in other words – provide the user with general encyclopedia-type information. This information is geocoded on-the-fly by linking it to the smallest polygon that covers the R-spaces of a selected set of GIS-layer objects. The polygon constitutes the R-space of the class attribute information and it is used by the algorithm to position the I-segment and hence to decide whether there is time enough to present this information. The objects that are used to define the polygon must be characterized by having identical values for an attribute value that specifies the object class. Moreover, the objects must be clustered together in a way that makes their R-spaces overlap as discussed above. The coding of the algorithm is relatively straightforward and pseudo code is provided below. Note that the word "level" means "level of generalization" in the pseudo code:

1. Retrieve the current speed and location of the traveler
2. Choose the most detailed level and all objects within this level
3. Generate R-segments by projecting the R-layer buffers of the current level to the expected travel path corresponding to a selected time period (see figure 1).
4. Generate I-segments centered on the R-segments by computing the length necessary to present the attribute information.

5a. If no I-segments overlap: (there is time to present all information at this level)
    compare I and R lengths:

if (I-segment >> R-segment):
try selecting higher level data (based on system ontology), goto step 3
otherwise
present attribute information from the current level
5b. If I-segments overlap: (there is not enough time for all information)
try selecting higher level data, (based on system ontology), goto step 3
- If no higher level data can be found in step 5: pass by the objects without providing any information
6. If tour not ended: goto step 1

4 Discussion and Conclusions

The prototype software is developed for experimentation purposes and is currently in a 'laboratory' stage. The real-time nature of the attribute presentation process creates some problems for documenting its behavior. The software reports the selected speed, the computed start and ending points of attribute information, as well as the information itself. The expected future travel path (EFTP) is not given much attention in this study, although the ability to estimate this for some limited distance at any time during the travel is, indeed, important for the proper selection of attribute data. The required length of the EFTP depends on the speed of travel and the duration of the potential up-coming attribute messages. Computing the EFTP becomes easier if the mobile GIS is used also as a navigation device and providing the user with directions. The use of probabilities for route selection based on current direction and type of road etc. will be necessary if no route is given a priori.

The above discussed issue of selecting the specific generalization strategy is important, if the ontology and data availability allow for a choice between several options. It may, for example, be possible – at a given speed – to select both a subset of objects within the same layer and a more generalized new layer to reduce the duration of the information. This functionality is not handled by the prototype at this stage, since all test runs are made with only one suitable generalized data set. It would be reasonable to assume that user preferences in this case would be related to the specific type of data. As an example, the user may prefer general class information about the vegetation species as a generalization of tree objects, while preferring the spatio-thematic generalization of building objects into buildings objects of historical interest. If this is the case, it will be necessary to expand the ontology definitions to include these preferences. Following a similar line, it could be argued that a user may sometimes prefer a more generalized level of information than what is potentially possible at the current speed. This is not an ontology issue but rather a question of allowing the user to control the system behavior by having all potential attribute information compared to a certain user defined threshold that excludes the detailed information.

To conclude more generally, it would seem as if the well known concept of geographical information systems is easily depicted in a mobile information system context. The close association between attribute data residing in a database and graphical objects on the map is what defines the GIS and provides its potentials. It should therefore be clear that presenting these attribute data in a mobile context could also be highly beneficial. The high amount of existing spatial data that would be of interest to different segments of travelers is also an argument for further work towards making these data available in mobile systems.

References

1. Egenhofer, M. A family of smart maps In Timpf S: (Ed). *Euresco Conference on Geographical Domain and Geographical Information Systems*. ESF. Bad Herrenalb, 2003.
2. ESRI: Avenue – Customization and Application Development for ArcView. *Environmental Systems Research Institute*. Redlands, 1996.
3. Fonseca, F., Davis C, Camara G. Bridging Ontologies and Conceptual Schemas in Geogarphic Applications Development. *Geoinformatica*, 7(4), 2003.
4. Fonseca, F., Egenhofer, M., Clodoveu, D., Câmara, G. Semantic Granularity in Ontology-Driven Geographic Information Systems. *AMAI Annals of Mathematics and Artificial Intelligence - Special Issue on Spatial and Temporal Granularity* 36(1-2): 121-151, 2002.
5. Frank, C., Caduff, D., Wuersch, M. From GIS to LBS – An Intelligent Mobile GIS. *IfGI prints* Vol. 22 Münster, 2004.
6. Hendrix, G. Encoding knowledge in partitioned networks. In N.V.Findler (ed.): *Associative Networks – The Representation and Use of Knowledge in Computers*. New York. Associated Press, 1979.
7. IST. Location-Based Services in the Culture and History Travel Market. *Intelligent Spatial Technologies*. White Paper, 2004.
8. Kulik, L., Duckham, M., Egenhofer, M. Ontology-driven map generalization. *Journal of Visual Languages and Computing* 14 (5): 469-492, 2003.
9. Kwan, K., Shi, W. A study of dynamic database in Mobile GIS. *Symposium on Geospatial Theory, Processing and Applications*, Ottawa, 2002.
10. Møller-Jensen, L. Classification of urban land cover based on expert systems, objects models and texture. *Comput., Environ. And Urban Systems*, Vol 21, No ¾ , 1997.
11. Raper, J. *Multidimensional Geographic Information Science*. Taylor & Francis, London, 2000.
12. Ringland, G.A., Duce, D.A. (Eds). Approaches *to Knowledge Representation*. Research Studies Press. 1988.
13. Rodríguez, M., Egenhofer, M. Determining Semantic Similarity Among Entity Classes from Different Ontologies. *IEEE Transactions on Knowledge and Data Engineering* 15(2), 2003.
14. Weibel, R. Map generalization in the Context of Digital Systems. *Cartographic and geographic Information Systems*, 22(4), 1995.
15. Weibel, R., Dutton, G. generalizing spatial data and dealing with multiple representations, In; Longley PA, Goodchild MF, Maguire DJ, Rhind DW: *Geographical Information Systems – Principles and Technical Issues* Vol I Chap 10, 1999.
16. Winter, S. & Tomko, M. Shifting the focus in Mobile maps. *Proc. UPIMap2004*. 7-9 September. Tokio, 2004.

Ontology-Driven Description of Spatial Data for Their Semantic Processing

Miguel Torres[1], Rolando Quintero[1], Marco Moreno[1], and Frederico Fonseca[2]

[1] Geoprocessing Laboratory, Centre for Computing Research, IPN, Mexico City, Mexico
{mtorres, quintero, marcomoreno}@cic.ipn.mx

[2] School of Information Sciences and Technology, Pennsylvania State University, USA
fredfonseca@ist.psu.edu

Abstract. We use ontologies in this paper to search for alternative representations of geographic objects thus providing a description of these objects in cartographic vector maps. We define ontologies based on two types of concepts ("terminal" and "non-terminal") and two kinds of relations ("has" and "is-a"). These are the basic elements used to describe a map. We also present a case study in which an ontology for topographic maps is created. Our approach is oriented towards solving heterogeneity and interoperability issues in GIS.

1 Introduction

Recently, the notion and concept of ontology has gained increased attention among researchers in geographic information science [5] to address the many problems of geographic data dealing with spatial data exchange and representation. According to Smith *et al.* [8], ontologies are essential to the creation and use of data exchange standards, and to the design of human-computer interfaces. Ontologies are also important in the solution of problems of heterogeneity and interoperability of geographic data.

Moreover, there are other problems regarding geographic objects representation. For instance, spatial analysis often requires very precise description of objects to provide good solutions. For instance, in Mexico City the population outgrows the official boundaries of the city, thus invading ecological and protected areas, and creating zones without services. Mexico City is thereby a dynamic entity, in which the actual boundaries cannot be defined exactly (just with inconsistencies). In this case, the boundary of Mexico City is an imprecise geographic object, and its analysis will be imprecise as well.

There are other sources of imprecision, for example, data in different scales, different resolution levels, or attributes that are implicit in the composition of the geographic objects. Although the consequences are diverse, the data that are used in GIS-applications are often imprecise; thereby it is important to consider alternative object representations, which are independent of the imprecise nature of the data.

In this paper, the features related to representation of spatial data are described in the proposed ontology. The objective is to generate descriptions of the spatial data. Furthermore, we present an example to show how such ontology can be built.

M.A. Rodríguez et al. (Eds.): GeoS 2005, LNCS 3799, pp. 242 – 249, 2005.

At the same time, we define the concepts that compose the ontology. They are related to the properties of the set of spatial data, which are represented in the ontology. In our approach, the descriptions are generated using tuples of concepts related among them. In order to build the description, we use the relations *"has"* (Γ) and *"is-a"* (Φ) which are defined in the ontology. In addition, the set of relations is defined by the pairs, which are associated to Γ and Φ (Γ and Φ are non-reflexive, non-symmetric and transitive relations).

The rest of the paper is organized as follows. Section 2 describes related work. In section 3, we present an ontology for geographic data and an approach to generate descriptions. Section 4 shows a case study of creating a description of the topographic data by means of the ontology. Our conclusions and future work are outlined in section 5.

2 Related Works

Guarino [2] coined the term "ontology-driven information systems" and provided a broad discussion on their place in the computer and information sciences. Gruber [4] defines ontology as a formal explicit specification of a shared conceptualization.

Fonseca *et al.* [3] propose a way to link the formal representation of semantics (i.e., ontologies) to conceptual schemas describing information stored in databases. The main result is a formal framework that explains the mapping between a spatial ontology and a geographic conceptual schema. The mapping of ontologies to conceptual schemas is made using three different levels of abstraction: formal, domain and application levels.

According to Worboys [10], geographic data models explicitly represent a set of basic objects, their geometry and their properties. But much of the semantics appears in the relations linking objects. Nevertheless, although some relations are represented in data models others are not. Usually, these non-represented relations appear implicitly when one is looking at a display of a geographic database [7]. In addition, the use of descriptions allows us to explicitly represent the relations that link objects, whereby we can say that descriptions contain a high semantic content [10] and these can be used as an alternative representation method for spatial data. Furthermore, the main purpose of the obtained descriptions is to partially solve problems of heterogeneity and interoperability of the spatial data.

In Mark *et al.* [6], the authors assume that cognitive categories in the geographic realm manifest certain special features as contrasted with categories for objects at surveyable scales. They argue that these features reflect specific ontological characteristics of geographic objects.

3 Ontologies for Spatial Data Description

We use ontology that describes spatial data according to a context. Our approach is oriented to describe cartographic vector maps. We define a map as a spatial partition Ω inside of a universe of geographic objects α_i, which consists of a set of representation primitives [9] (Equation 1).

$$\Omega = \alpha_i \cup \{Rp_l, Rp_p, Rp_a\}, \; i = 1,...,n, \tag{1}$$

where Rp_l is the primitive of representation "linear"; Rp_p is the primitive of representation "punctual"; Rp_a is the primitive of representation "areal"; i represents the thematic number that involves the spatial partition.

We define two types of concepts (C) in the ontology: *terminal* (C_T) and *non-terminal* (C_N). Terminal concepts do not use other concepts for defining their meaning (they are defined by "simple values"). Non-terminal concepts use other concepts (terminal or non-terminal) in their definitions (Equation 2).

$$C = C_N \bigcup C_T \tag{2}$$

Each concept has a set of aspects, because geographic objects have aspects. Aspects are properties and relations between geographic entities. In the following, we shall use the term "relation" to denote unary relations/properties as used in Berendt *et al.* [1]. From this point of view, all aspects of a terminal concept are simple, e.g. the type of all aspects that belongs to the set of primitive types (T_P), as shown in Equation 3.

$$\begin{aligned} &T_P = \{number, character, string, enumeration, struct\}, \\ &A = \{a_i \mid type(a_i) \in T_P\}, \end{aligned} \tag{3}$$

where T_P is the set of primitive types; A is the set of aspects.

Then, the set of *terminal concepts* is defined by Equation 4.

$$C_T = \{c(a_1, a_2, ..., a_n) \ni a_i \in A, \; i = 1,..,n\} \tag{4}$$

In the same way, the *non-terminal concepts* have at least one aspect that does not belong to T_P. It is denoted by Equation 5.

$$C_N = \{c(a_1, a_2, ..., a_n) \ni \exists a_i \notin A\} \tag{5}$$

where c is a concept.

Finally, the set of relations R is defined by the pairs that are associated to Γ and Φ, where Γ and Φ are non-reflexive, non-symmetric, and transitive relations (Equation 6).

$$R = R_\Gamma \cup R_\Phi = \{(a,b) \mid a\Gamma b, \; a \in C_N, \; b \in C\} \cup \{(a,b) \mid a\Phi b, \; a \in C_N, \; b \in C\} \tag{6}$$

This description maps spatial data into the ontology. Once the concepts are defined in the ontology, we can start defining the non-terminal concepts (this means to select the aspect to be described). This process continues until we find a terminal concept. Once, the terminal concept is found, it is necessary to select a pair of geographic objects, verifying whether a relation between them exists, otherwise a part of the description needs to be generated. The terminal concepts are defined by the type of relation between two objects. Fig. 1 shows an ontology fragment for hydrological maps of the linear type: all objects that compose the hydrological network are represented by lines (drainage and rivers). The ontology consists of two types of concepts (non-terminal and terminal) and a set of relations. The relations are the following: "*has*" and "*is-a*"; there are three relations in this ontology.

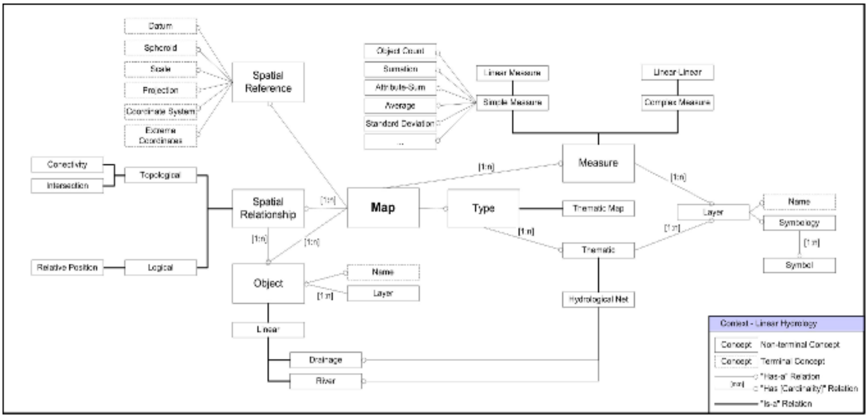


Fig. 1. Ontology fragment for hydrological maps of linear type

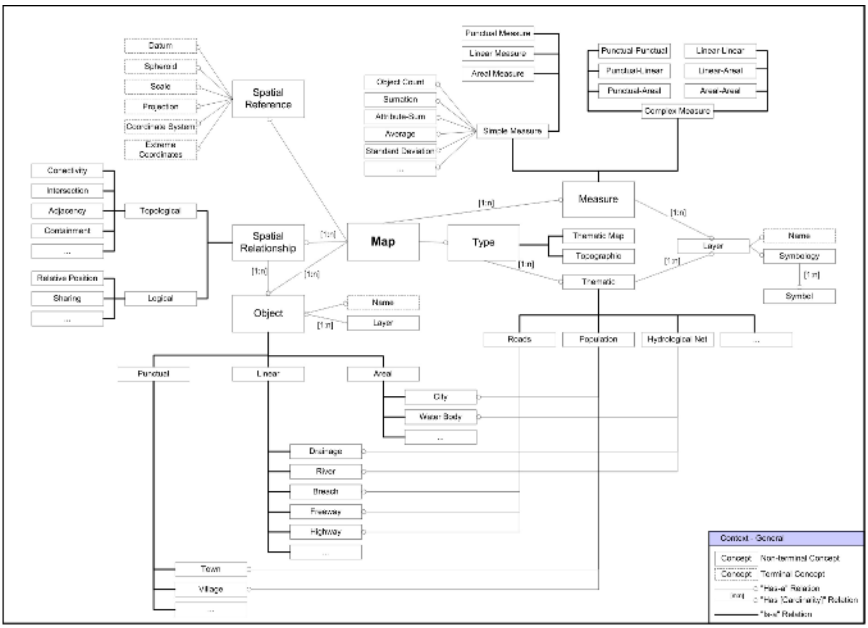


Fig. 2. Part of the ontology for topographic maps

The concepts in Fig. 1 are represented by "boxes with three points". For instance, looking at "punctual" concepts (*town* and *village*), we note that many others can exist, such as archeological sites, monuments, wells, or buildings. The proposed ontology defines all concepts required for the spatial data description, according to the standards of INEGI (the National Institute of Statistics, Geography and Informatics, who produces the official cartographic maps in Mexico). On the other hand, it is important to note that different ontologies may exist (one for each thematic layer) and they need to be joined in a "general" ontology. For example, a topographic map is composed of rivers, contours, highways, and cities, and it can be merged in one ontology. Fig. 2 depicts a part of this general ontology for a topographic map. One can see that hydrological maps are a subset of the ontology shown in Fig. 1.

4 Case Study

In this section we present a case study using the proposed ontology to describe the aspects of spatial data in cartographic vector maps.

The map of Fig. 3 depicts different thematic layers. Each layer contains geographic objects represented by spatial primitives. This map contains *Populations* (POP), *Hydrologic Network* (HYN), *Roads* (ROD) and *Soils* (SOL). In addition, each thematic layer is denoted in the map legend, and is described by specific symbols. The map is composed of 3 punctual objects, 6 linear objects, and 5 areal objects.

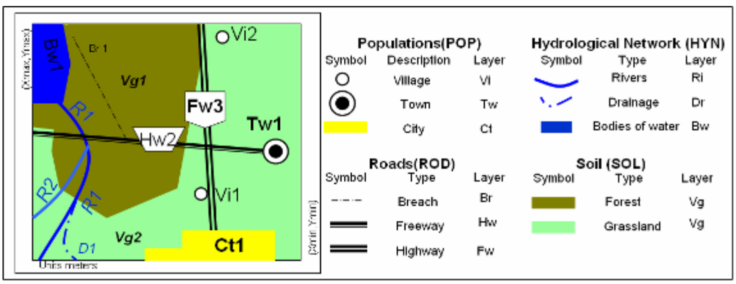


Fig. 3. Thematic map used in the case study

Thereinafter, we use the ontology to describe this map. Fig. 4 depicts the description. The description starts at the non-terminal concept called "***Map***". The non-terminal concepts are denoted by means of *rectangles* and the values of the terminal concepts are represented by *ellipses*.

According to the aspect of each non-terminal node, we establish a relation, which defines another non-terminal or terminal concept (depending on the objective). This leads to the complete description of the geographic objects that compose the partition (Fig. 4).

On the other hand, the properties (*aspects*) that belong to each terminal node contain quantitative values. Moreover, the ontology includes the topological and logical relations, symbol sets, and measurements of the map content.

In our approach, it is important to characterize beforehand the topological relations in order not to consider all of them in the semantic process because of their excessive number in some cases. Additionally, the description depends directly on the context; therefore it is not possible to count on a general context, since some semantic ambiguities may occur.

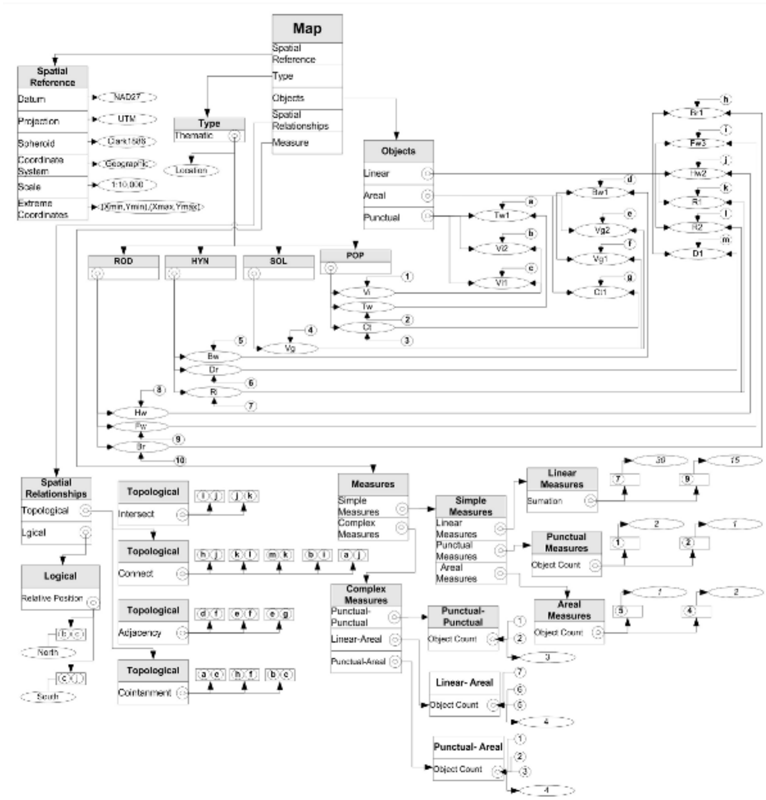


Fig. 4. Semantic description of the thematic map shown in Fig. 3

The method is focused on describing the *semantic content of a cartographic vector map*. However, this description depends on a number of spatial relations, properties and cartographic *measurements*[1] that needs to be considered. Whereby, it is possible to increase the semantic resolution in the description[2].

The description is made using *tuples* of non-terminal and terminal concepts related among themselves (they are denoted by *Concept* ***relation*** *Concept*).

For instance, Fig. 3 is composed of several spatial objects. The objects in the layer reflect the relation **"is-a"** (i.e., *HWY* **is-a** *Linear Object*). Moreover, the topological relation ***"Intersects"*** is related *Hw2* and *Fw3*, which are both linear objects. According to the ontology (Fig. 4), we see that the ***"Intersects"*** relation is a *topological relation* and at the same time, it is a *spatial relation*, whereby this relation is congruent with the description *Fw3* ***Intersects*** *Hw2*. In Table 1 all spatial relations are depicted according to the description of Fig. 4.

[1] A measure is a procedure for computing values, which are the basis to evaluate characteristics of cartographic phenomena and assess the need for and the success of a map description.

[2] This assumption is only considered for the case study, because the map description contains all the relations of the map.

Table 1. Spatial relations between geographic objects

Objects	Vi2	Vg2	Vg1	Ct1	Fw3	Hw2	R1	R2
Tw1		♠				♦		
Vi2		♠			♦			
Vi1	↑					↓		
Bw1			♥					
Vg2			♥	♥				
Vg1		♥						
Br1			♠			♦		
Fw3						♣		
Hw2							♣	
R1								♦
D1							♦	

♣=Intersect, ♦=Connect, ♥=Adjacency, ♠=Containment, ↑=North, ↓=South

5 Conclusions and Future Work

We used ontologies to describe spatial data as an alternative representation. This representation allows us to handle imprecise data, since we translate them into a conceptual schema. We also showed how to generate the ontological (semantic) map description based on the two types of concepts: "terminal" and "non-terminal" as well as two kinds of relations: "*has*" and "*is-a*". By using these elements, we described a cartographic map. As a result, maps with the same theme could be described using the same ontology (thematic ontology). As an example, we presented a fragment of the ontology and the result of its use in the description of a map. On the other hand, it is not possible to describe data that are not considered in the ontology. For instance, we can not describe the freeways of a region, if the ontology is about hydrology. From this point of view, the ontology implicitly defines the context of the map theme. A great challenge is to define formally what the context is, but at first approximation, we believe that the context could be described by means of a "context" ontology. Our ontologies are intended to be used in map production. Semantic descriptions would allow us to explicitly represent the relations that link objects. Whereby we argued that descriptions will contain a higher semantic content if we use ontologies for their construction. They can be also seen as an alternative representation for spatial data.

From this point of view, we have implicitly obtained the semantics of the spatial data by means of an ontology. That is, the description based on the ontology O of the spatial dataset D, provided the semantics S of D related to the context C, which is also defined by O. As a future research, we will attempt to provide measures of the ambiguities and inconsistencies, which can be involved into the content. We will also search for mechanisms to minimize the degree of imprecision in the content.

Acknowledgments

The authors of this paper wish to thank the CIC, CGPI, IPN and CONACYT for their support and Sergei Levashkin, who inspired this research.

References

1. Berendt, B., Barkowsky, T., Freksa, C. and Kelter, S. I. E. Spatial Cognition - An Interdisciplinary Approach to Representing and Processing Spatial Knowledge. *Spatial Representation with Aspect Maps*, Springer-Verlag, pp. 157-175, 1998.
2. Guarino, N. Formal Ontology and Information Systems. *Proceedings of the International Conference on Formal Ontology in Information Systems*, Kluwer Academic Publishers, IOS Press, Trento, Italy, pp. 3-15, 1998
3. Fonseca, F., Davis, C. and Câmara, G. Bridging Ontologies and Conceptual Schemas in Geographic Information Integration. *GeoInformatica*. 7(4): 355-378, 2003.
4. Gruber, T. R. Toward Principles for the Design of Ontologies Used for Knowledge Sharing. *International Journal of Human and Computer Studies*. 43(5/6): 907-928, 1995.
5. Mark, D., Egenhofer, M., Hirtle, S., and Smith, B. UCGIS Emerging Research Theme: Ontological Foundations for Geographic Information Science. *UCGIS Emerging Research*, 2000.
6. Mark, D., Smith, B. and Tversky, B. Ontology and Geographic Objects: An Empirical Study of Cognitive Categorization. *Proceedings of the Spatial Information Theory: A Theoretical Basis for GIS*, Stade, Germany, pp. 283-298, 1999.
7. Mustière, S. and Moulin, B. What is Spatial Context in Cartographic Generalisation? *Proceedings of the Symposiun on Geospatial Theory, Process and Applications*, Ottawa, Canada, 2002.
8. Smith, B. and Mark, D. Ontology and Geographic Kinds. *Proceedings of the 8th International Sysmposium on Spatail Data Handling*, Vancouver, Canada, pp. 308-320, 1998.
9. Torres, M. and Levachkine, S. Generating Spatial Ontologies based on Spatial Semantics. *Proceedings of the International Workshop on Semantic Processing of Spatial Data*, Mexico City, Mexico, pp. 163-178, 2003.
10. Worboys, M. F. Metrics and Topologies for Geographic Space. *Proceedings of the International Symposium on Spatial Data Handling*, Delft, Netherlands, 7A.1-7A.11, 1996.

An Interstage Change Model for Sandbox Geography

Florian A. Twaroch

Institute for Geoinformation and Cartography, Technical University Vienna,
Gusshausstrasse 27 - 29, 1040 Vienna, Austria
twaroch@geoinfo.tuwien.ac.at
http://www.geoinfo.tuwien.ac.at

Abstract. Formal models for incremental ontologies are needed in geographic information systems. While there are methods to describe ontologies for a certain purpose, it is still an open question, how to link different geographic ontologies. Observing children in a sandbox can motivate a new way of designing dynamic spatio-temporal ontologies. Contemporary developmental psychology provides evidence that knowledge about the world is acquired in piecemeal fashion. Infants form theory like concepts of the world that are revised in the light of new evidence [12]. We take these findings to build multi tiered ontologies grounded in children's spatial experience. Questions of how to structure and connect the ontologies will be addressed. The formalization of spatial concepts are investigated in an agent based approach using an algebraic framework.

1 Introduction

Children's concepts of space are different from those of adults. Infants are more general in their learning than adults, because their perceptual system is more abstract and less specific [2]. It can be assumed that children start out with certain knowledge into this world and that there are mechanisms of development that make a human proceed through life.

The formalization of these concepts seems to be a promising approach to reach a better understanding of how to build incremental ontologies for geographic space. Conceptualizations of the world change under new evidence. The paper discusses first steps in how to formalize changing concepts in small scale space.

Though objects in large scale space are not directly perceivable or manipulable, the learning process is assumed to be similar. All adults have once been children and have been going through developmental processes. It is comparable to the bootstrapping mechanism on a computer. Two directions of research are motivated. First to study the structure of spaces, built up by operations and perceptions of children in small scale space. This includes the formal description of such spaces. Second to connect the different conceptualizations to a learning theory for a cognizing agent, e.g., based on a re-weighting mechanism.

This paper will concentrate on the first part of the question, namely how to structure the ontologies, in order to build a proper base for investigating the change mechanism between ontologies. The research assumes stages of knowledge defined in an agent. Each stage corresponds to a set of algebras. This is done in order to investigate structural similarities between the stages in the future.

M.A. Rodríguez et al. (Eds.): GeoS 2005, LNCS 3799, pp. 250–257, 2005.

Stage wise acquisition of spatial knowledge has been already proposed very early [24] e.g., children go through six stages when learning about the permanence of objects. Some of these findings could be worked out in greater detail since the investigations of Piaget under the aspect of large scale space [13,25]. In a recent view conceptualizations of the world are explained as theories [12]. These theories change in the light of new evidence. To distinguish them from fully fledged theories, philosophers like Roberto Casati would rather call them "theoritas" (little theories) [4].

For the present paper a theory stands for a conceptualization of the world at a certain point of time in an infants development. In early infants behavior spatial relations seem to be missing. Concepts like containment, support or occlusion are gradually learned by making sensorimotor experiences in space. A formalization of this learning process seems promising to reach sound formal models of incremental spatial ontologies. We present a formal model for object support using an algebraic framework. The aim of building ontologies based on children's concepts of space is a contribution to the area of naive geography [6] and spatial reasoning. The goal is not to model an infant or a mental model nor to provide a tool to be used in developmental psychology.

Complex systems can be build from simple parts, this applies also for geoinformation systems [9]. The simple parts have been identified in a series of studies about humans conceptualization of space [26,12,1]. The crucial question however is how these parts are linked and interact with each other. By searching for a learning theory we hope to find mechanisms that describe the structure of and the connections between these parts for small scale space (spatial) as well as large scale space (geo-spatial) problems.

The remainder of the paper is organized as follows: Section 2 reviews related work in conceptualization of space and models proposed for it. Section 3 proposes an overall framework for an object ontology that is modular, hierarchical, dynamic and action-driven. The connection to algebra is discussed. Section 4 gives deeper insights in the modules and the modelling process. Section 5 concludes the paper and gives an outlook to future research questions.

2 Human Conceptions of Space

This section reviews what is known about the human conceptualization of space and how far it has been considered in formal models. The existence and structure of space has been of continuous research interest in geoinformation science. Conceptualizations of space have been worked out [20,16,5,22,11,21], mainly distinguishing between small and large scale space resulting in various taxonomies of spaces.

Freundschuh and Egenhofer study the links between different spaces. They base their study on a review of 15 different models of (geographic) space. The attributes: size of space, manipulability, and locomotion are used to propose a model that distinguishes six types of spaces and the connection between them. One of the open research questions is to explore what kind of spatial knowledge people acquire in each kind of space [11].

Couclelis and Gale discuss the formal difference between perceptual and cognitive spaces. Utilizing an algebraic approach they point out the difficulty to find a universal definition for the concept of space. Different spaces have different algebraic structures. Physical space is a group while e.g., sensorimotor space is a monoid [5].

The present work concentrates on small scale space and the various conceptions in a child's development of it. Recent formal models of tabletop space have been based on image schemata [8,7] in toy spaces. The proposed model differs in so far as it considers infants activities in space. Empirical experiments carried out in developmental psychology, serve to model incremental ontologies for tabletop objects.

The formalization of cognitive development has been exploited as a tool for psychologists to build sound theories. A sound theory is the result of very specific definitions, as used in computer models. Early research in autonomous agents and robots like models using production systems [18] concentrated on describing the stages of development, without explaining the shift between different levels of competence. More recent studies investigate qualitative changes in development using connectionist approaches like neural networks [14].

There is a multitude of concepts for space and spatial objects. It seems that humans are endowed with various mechanisms to encode them [23]. The learning process of infants can be a motivation to formalize object ontologies that are close to human thinking. Different levels of competence can be identified in infants development. A sound mechanism to describe the transition between these levels is still an open research question. A transition mechanism would enable the construction of a framework that allows to switch between different levels of competence and between different kind of spaces. Computational models can help to define sound theories and provide feedback for necessary investigations.

3 Structure of the Ontology

This section describes an object ontology for tabletop space. The ontology is based on simple theories that are expressed as algebras. These algebras are embedded in an agent's knowledge base. Through acting in the environment the agent elicits changes in the algebraic framework. The section will introduce examples for changes in the algebras.

3.1 A World of Theories

The state of the world is described by a set of theories. There are theories for different kind of objects. Objects that move, objects that fall, objects that are supported by other objects. We distinguish core theories that describe basic object behavior like movement and identity and more advanced theories like occlusion, containment, friction and gravity. Some of the advanced theories are derived from core theories.

The theories build the knowledge base of an agent. The agent observes objects and manipulations on objects in a sandbox world. The agent can test theories about the world through observation. The agent can predict occurrences and compare them with observations. When necessary the agent exchanges a theory. The current model assumes that observations are free of any error or uncertainty. Each theory is defined by an initial concept. In order to model beliefs of the agent and facts in the world, a multi layered ontology is applied [10].

The successive refinement of the theories demands that old theories are retested. Whenever a new theory is taken into the framework the consistency of the old theories has to be guaranteed [15]. Such a mechanism to retest will be foreseen in future work. At the time contradictions or violation of the monotony of theories did not occur in the modelling process.

The only action considered in the model at the time is observation, in a simplified error free form. The present model can benefit in future from the consideration of further activities. The consideration of activities, movement and the intention of agents leads to action-driven ontologies.

3.2 Theories as Algebras

Theories about the world are described with algebraic specifications. A change in the conceptualizations of the world is reflected in an exchange of axioms. We use the algebra in its simplest definition as a set of sorts, operations, and axioms [17]. The functional programming paradigm with algebraic specifications is used to carry out a prototypic implementation. It focuses on operations, thus activities carried out in space. The approach provides an object oriented view to the world, where algebras group operations based on the same data type. The advantage of using algebra is its mathematical sound- and compactness. It allows the reuse of code by defining sub algebras and combining different algebras [9].

The basic elements of the proposed object ontology are core objects. They consider the identification and the comparison of objects. We give an example how developmental psychology connects with algebraic specifications. There is some evidence that newborns distinguish objects based on their spatio-temporal properties [2,12]. Empirical tests show that infants at the age of a few months can identify objects via spatio-temporal attributes and later move on to a feature based object identification. This has already been expressed formally by exchangeable rule sets [18,2]:

- Rule 1: An object is a bounded volume of space in a particular place or on a particular path of movement.
- Rule 2: An object is a bounded volume of space of a certain size, shape, and color that can move from place to place along trajectories.

These rules are translated in algebraic definitions. We give the example of an algebra that defines the equality of objects. While for the first rule the definition of equality is just based on an axiom comparing the location of two objects, the follow up algebras needs to define an axiom that considers additionally a comparison of object attributes. We give an example in pseudo code (Example 1), where *loc (object)* is an operation to define the location of an object and *att(object)* an operation to derive certain attributes of an object.

Example 1. Qualitative change of an algebra for the equality of objects

```
Algebra Equality 1
isEqual (object A, object B) = loc (object A) == loc (object B)
Algebra Equality 2
isEqual (object A, object B) = loc (object A) == loc (object B) &&
                               att (objectA) == att (object B)
```

4 An Example for Theory Change - The Support of Objects

In this section we provide a more advanced theory - the theory of object support. Knowledge about the support of objects has been tested in interpreting the time children look at novel events. Infants tend to look longer at new than at unknown situations. Psychological experiments are manipulated in a way that they produce unexpected outcomes. Objects that should fall, are supported and vice versa. The infants reaction on this experiments is interpreted as knowledge. Again we use these findings to carry out a formalization of object support.

Figure 1 shows two cases of a scene were an object A supports an object B. On the left side of the figure the two objects have contact via their side surfaces. Though adults would not conceptualize this situation as object support, infants younger than 3 month do not show any sign of surprise when object B does not fall [19]. The right side of the figure shows two objects that touch via their top/bottom surface. Some cases will lead to object support while others will not. For a first model we define two axioms:

Fig. 1. Two objects in a scene touch each other. In situation (a) the objects touch on their sides no object support should be expected. Situation (b) will lead in some case to object support.

- Axiom 1: An object A supports an object B when the objects have contact with each other.
- Axiom 2: An object A supports an object B when the contact of B is via the Top-Surface of A.

This simple example can be indeed verified in empirical experiments with children in the age of 3 to 6 month. An initial concept for support seems to consider just the contact between two objects. We give an example in pseudo code (Example 2), where *contact (object1, object2)* is an operation to define the type of contact between two objects. Possible values are "Top" (objects touch via top/bottom surface) and "Side" (objects touch on their side surfaces) contact.

Example 2. Qualitative change of an algebra for the support of object

Algebra Support 1
isSupported (object A, object B) = contact (object A, object B) == Top || Side
Algebra Support 2
isSupported (object A, object B) = contact (object A, object B) == Top

Developmental psychology gives evidence for a more comprehensive theory of object support. After considering that an object that supports another object must have contact

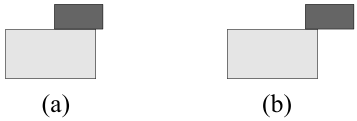


Fig. 2. The amount of contact will lead to a prediction about when an object is to fall an when it is supported. In situation (a) the dark grey object B should fall, while in situation (b) the dark grey object B should be supported.

via the top surface of the supporting object a variable like the amount of contact is recognized [1]. Figure 2 visualizes the situation.

An extension to the Algebra Support 2 of example 2 is necessary. An example in pseudo code (Example 3) is given were *amount(object1,object2)* is an operation that determines the amount of contact between two objects. It is compared to a certain threshold t.

Example 3. Extension of the Algebra Support

Algebra Support 3
isSupported (object A, object B) = contact (object A, object B) == Top &&
amount (object A, object B) == t

These algebras can be embedded in the knowledge base of an agent. In a functional prototype the agent's expectations of an event (object A falls) are compared with the actual facts (object A does not fall, it is still supported). Continuous observation of differences between expectations and facts will raise "doubt" in the agent and cause the exchange of a theory. Here a stochastic model could be useful for grading observations and predictions and is topic of future research.

5 Conclusions

We have shown how computer science and developmental psychology motivate a new way to design ontologies for the spatial realm. A computational model of an agent in a "sandbox" has been recently developed. The functional programming paradigm was chosen, that allows rapid prototyping with algebraic specifications.

Simple parts have been identified that contribute to an object ontology that is close to human thinking. In this paper we concentrated on small examples like object identity and object support. Like physicists searching for the smallest parts of the universe, geoscience is on the way to identify the smallest elements of processes in space. The next step is to define a sound mechanism for linking the parts and to extend the functional model.

Further theories about object motion, gravity, inertia, occlusion relations and many others have to be developed. On the way of defining small theories about the world, children undergo changes, that have to be considered in the agent model. To give an example of how body experience influences the development of we look at conceptions of gravity and support. Children have an early notion of gravity when they see things falling down. When infants learn to sit (body experience) they also seem to get an advanced understanding for support relations. Another example is the consideration of self-movement that provides a link to large scale spaces in the proposed theory of an developing agent.

A futher aspect is the acquisition of language, that has influences on the conceptualization of the world, e.g., the development of counting [3]. Formal models for these concepts to extend the agent model are a potential topic for future research. People's different beliefs about the world could be modeled in an multi agent system.

Additional investigations of common-sense geographic models and the transition between these models using stochastic methods is planned. We are confident that there are several possibilities to link the ontologies and that this mechanism carries a certain redundancy. The consistent mathematical framework of algebra should be further exploited.

Acknowledgments

I like to thank Andrew Frank for joining me along the way of my PhD. I appreciate the help of George Luger, who raised my interest in artificial intelligence and gave so useful pointers to delve deeper into the topic. Thanks to the valuable comments of three anonymous reviewers and to Claudia Achatschitz for correcting the final text. This work has been funded by the E-Ten project C 517343 GEOCOMPASS.

References

1. Renee Baillargeon. Infants' reasoning about hidden objects: evidence for event-general and event-specific expectations. *Developmental Science*, 7(4):391 – 424, 2004.
2. T.G.R. Bower. *The Rational Infant - Learning In Infancy*. W. H. Freeman and Company, New York, 1989.
3. Susan Carey. Bootstrap knowledge. *Deadalus, Winter*, pages 59 – 68, 2004.
4. Roberto Casati. *Time and Motion of Socio-Economic Units*, chapter The structure of shadows, pages 99–109. London: Taylor and Francis, 2000.
5. Helen Couclelis and Nathan Gale. Space and Spaces. *Geografiska Annaler*, 68B(1):1–12, 1986.
6. Max J. Egenhofer and David M. Mark. Naive Geography. In Andrew U. Frank and Werner Kuhn, editors, *Lecture notes in computer science (cosit '95, semmering, austria)*, volume 988, pages 1–15. Springer Verlag, 1995.
7. Max J. Egenhofer and Andrea M. Rodriguez. Relation Algebras over Containers and Surfaces: An Ontological Study of a Room Space. *Journal of Spatial Cognition and Computation*, 1(2):155–180, 1999.
8. Andrew U. Frank. Specifications for interoperability: formalizing spatial relations 'in', 'auf' and 'an' and the corresponding image schemata 'container', 'surface' and 'link'. Internal report, Dept. of Geoinformation, Technical University Vienna, 1998.
9. Andrew U. Frank. One Step Up the Abstraction Ladder: Combining Algebras - From Functional Pieces to a Whole. In Christian Freksa and David M. Mark, editors, *Spatial information theory - cognitive and computational foundations of geographic information science (int. conference cosit'99, stade, germany)*, volume 1661 of *Lecture Notes in Computer Science*, pages 95–107. Springer-Verlag, Berlin, 1999.

10. Andrew U. Frank. Tiers of ontology and consistency constraints in geographic information systems. *International Journal of Geographical Information Science*, 15(7 (Special Issue on Ontology of Geographic Information)):667–678, 2001.
11. Scott M. Freundschuh and Max J. Egenhofer. Human Conceptions of Spaces: Implications for Geographic Information Systems. *Transactions in GIS*, 2(4):361–375, 1997.
12. Alison Gopnik and Andrew N. Meltzoff. *Words, thoughts, and theories*. Learning, Development, and Conceptual Change. MIT Press, Cambridge, Massachusetts, 1997.
13. Roger A. Hart and Gary T. Moore. The development of spatial cognition of large scale environments: a review. Technical report, Clark University, 1973.
14. Kazuo Hiraki, Akio Sashima, and Steven Phillips. From Egocentric to Allocentric Spatial Behavior: A Computational Model of Spatial Development. *Adaptive Behaviour*, 6(3/4):371–391, 1998.
15. Thomas S. Kuhn. *Die Struktur wissenschaftlicher Revolutionen*. suhrkamp taschenbuch wissenschaft, Frankfurt, 1976.
16. Benjamin Kuipers. Modeling spatial knowledge. *Cognitive Science*, 2(2):129–153, 1978.
17. Jacques Loeckx, Hans-Dieter Ehrich, and Wolf Markus. *Specification of abstract data types*. John Wiley and B.G. Teubner, Chichester, UK and Stuttgart, 1996.
18. George F. Luger, Jennifer G. Wishart, and T.G.R. Bower. Modelling the stages of the identity theory of object-concept development in infancy. *Perception*, 13:97–115, 1984.
19. Yuyan Luo and Renee Baillargeon. When the ordinary seems unexpected: evidence for rule-based physical reasoning in young infants. *Cognition*, 95(3):297 – 328, 2005.
20. Kevin Lynch. *The image of the city*. MIT Press, 1960.
21. D. Mark, C. Freksa, S. Hirtle, R. Lloyd, and B. Tversky. Cognitive models of geographical space. *IJGIS*, 13(8):747–774, 1999.
22. Daniel R. Montello. Scale and multiple psychologies of space. In *Cosit*, pages 312–321, 1993.
23. Nora S. Newcombe and Janellen Huttenlocher. *Making space: the development of spatial representation and reasoning*. Learning, Development, and Conceptual Change. MIT Press, Cambridge, Massachusetts, 2003.
24. Jean P. Piaget and Bärbel Inhelder. *Die Entwicklung des räumlichen Denkens beim Kinde*, volume 6. Klett-Cotta, Stuttgart, 3 edition, 1999. translated by Heipcke, R.
25. Alexander W. Siegel and Sheldon H. White. *The Development of Spatial Representations of Large-Scale Environments*, volume 10 of *Advances in Child Development and Behavior*, chapter 2, pages 9–55. Academic Press, Inc. (London) Ltd., 1975.
26. Elizabeth S. Spelke. Principles of object perception. *Cognitive Science*, 14:29–56, 1990.

Author Index