Real estate and the ontology of multidisciplinary, e.g. Cadastral studies

Stubkjær, Erik

Published in:
Geographical Domain and Geographical Information Systems

Publication date:
2000

Document Version
Publisher's PDF, also known as Version of record

Citation for published version (APA):
Stubkjær, E. (2000). Real estate and the ontology of multidisciplinary, e.g. Cadastral studies. In Winter, S. (ed.) (Ed.), *Geographical Domain and Geographical Information Systems* (pp. 97-109). Vienna University of Technology. GeoInfo Series No. Vol. 19

Geo.info

19

Geographical Domain and Geographical Information Systems

Stephan Winter (Editor)

Institute for Geoinformation
Vienna University of Technology
Gusshausstr. 27-29/127
1040 Vienna, Austria

Series Editor

Andrew U. Frank
Institute for Geoinformation
Vienna University of Technology
Gusshausstr. 27-29/127
A-1040 Vienna, Austria
frank@geoinfo.tuwien.ac.at

ISBN 3-901716-20-3 GeoInfo Series Vienna

Printed in Austria

Typesetting: Camera ready by author
Printing and binding: Hochschülerschaft TU Wien
Wirtschaftsbetriebe GmbH

Foreword

This book contains the abstracts from the *EuroConference on Ontology and Epistemology for Spatial Data Standards* in La Londe-les-Maures (France), 22-27 September 2000.

The conference focused on bridging the gap between research in ontology and epistemology of spatially extended objects on one hand, and research in representational models of spatial phenomena on the other hand. This interrelation between what exists, knowledge and formalisms is still an open question for research, especially in case of uncertainty and vagueness. In the absence of a formal theory, we are in a situation where system designers may decide how to treat uncertainty. Standardisation is necessary, but premature. Ontology of spatially extended objects can provide a basis for the development of formal models of spatial uncertainty. Thus the conference dealt with ontology of space, spatial cognition, spatial approximation, spatial hierarchies, generalization, and related topics.

EuroConferences consist typically of talks from invited speakers, and presentations from (mostly young) researchers who were selected by peer review on their submitted abstracts. This volume contains the revised abstracts, and additionally some abstracts from the invited speakers.

The final program and actual abstracts are available online[1].

[1] http://www.geoinfo.tuwien.ac.at/events/Euresco2000/gdgis.htm

Acknowledgements

I would like to express my thanks to the scientific committee that has provided valuable help in all stages preparing and carrying through the conference: Max Egenhofer (USA), Peter Fisher (United Kingdom), Andrew U. Frank (Austria), and Michael Gould (Spain). Many efforts of the local organizers, Sally Lewis at the European Science Foundation and Gwendolyn Raubal at La-Londe, made the organization of the event smooth. This conference was funded by the European Science Foundation and the European Commission, with additional funding from MobileGIS Ltd. (Ireland), SICAD Geomatics (Germany), and the Institute for Geoinformation, Vienna University of Technology (Austria), which is gratefully acknowledged.

Stephan Winter

Contents

Seraphim Alvanides*

Modelling contiguity in a zone design context

Despite the wide availability and use of zonal data in geographical applications, problems still occur in representing and analysing such datasets in a GIS environment. One of the most frequently encountered problems with zonal data is the definition of spatial contiguity based on topological relationships between areal units. Contiguity is expressed in the form of a matrix and plays an instrumental role both in representing zonal data and in modelling spatial processes using areal units. For example, when smaller areal units are aggregated into larger zones a contiguity matrix is first constructed and then used throughout the aggregation process in order to secure that the zones are spatially contiguous.

A number of alternative methods have been suggested for describing topological relations and topological representations between areal units. Such methods focus on the spatial configuration of the areal units in question and usually proceed by providing a contiguity matrix that describes these relations in a mathematical form. Two contiguity representation methods are reviewed and their relevance to zone design is examined. Zone design is simply defined here as the aggregation of smaller areal units into larger zones, subject to contiguity constraints. First, the switching point method suggested by Macmillan and Pierce (1994) is examined as a contiguity constraint. The switching

**Department of Geography, Daysh Building, University of Newcastle, Newcastle upon Tyne, NE1 7RU, U.K., s.alvanides@ncl.ac.uk, http://www.staff.ncl.ac.uk/s.alvanides/.*

point method concentrates on the boundaries of the areal units that are to be aggregated. The method is proved to be highly efficient with simple topological structures that fully exhaust space. However, the method is shown to be inefficient with more complicated structures where areal units form islands or gaps in their contiguity matrix.

Subsequently, a generic approach to the contiguity problem is presented based on the traditional contiguity concept. A contiguity matrix is first created based on the spatial contiguity of areal units. The matrix is then transformed into a network that consists of straight lines connecting the contiguous areal units. The contiguity network can then be easily extended to cater for the more complicated structures mentioned above. It is shown that the network is a very robust method for modelling and checking contiguity in a zone design context. In addidion, it presents a more realistic way for measuring straight-line distances between areal units in the absence of a real netwrork. The method is demonstrated using administrative boundaries such as Local Authority Districts and Electoral Wards in the United Kingdom. Finally, the method can then be extended to handle flow or interaction data where the contiguity constraint plays a different role in zone design as demonstrated in Alvanides et al. (2000).

Bibliography

Alvanides, S., Openshaw, S., Duke-Williams, O., 2000. Designing zoning systems for flow data. In: Atkinson, P., Martin, D. (Eds.), Geocomputation. Vol. 7 of Innovations in GIS. Taylor & Francis, London.

Macmillan, W. D., Pierce, T., 1994. Optimisation modelling in a GIS framework: the problem of political redistricting. In: Fotheringham, S., Rogerson, P. (Eds.), Spatial Analysis and GIS. Taylor & Francis, London.

Dimitris Ballas[*]

GIS and spatial microsimulation for urban systems modelling: a new conceptual framework for the representation and analysis of local labour markets

This paper explores the potential of GIS and object-oriented spatial microsimulation frameworks for the representation of urban systems and in particular for modelling local labour market objects, such as economically active households and firms. First, it gives an overview of fundamental topics related to computer simulation, GIS, spatial microsimulation, epistemology and methodology. Various modelling approaches to urban system representation are discussed and their advantages and drawbacks are outlined. Further, the paper puts the case for an object-oriented spatial microsimulation approach to urban systems modelling. Examples of applied spatial microsimulation of households in a local labour market context are presented and the merits and disadvantages of the methodology are briefly discussed. It is shown how problems such as the ecological fallacy can be tackled in spatial microsimulation contexts and outputs from a spatial microsimulation model for the Leeds urban system are presented. In addition, it is argued that spatial mi-

[*] *School of Geography, University of Leeds, LS2 9JT, UK.*

crosimulation frameworks can enhance our understanding of the behaviour of different agents (objects) in the labour market (e.g. households and firms) and that it can provide new insights into the workings of urban and regional economic systems and help in the formulation of new labour market theories. The ontological characteristics of the microsimulated urban objects are also discussed. The paper discusses and evaluates the suitability and applicability of spatial microsimulation frameworks as a means of representing urban system entities. It is also shown how GIS and spatial microsimulation techniques can be used to perform social and economic policy analysis at different geographical scales and the epistemological properties of this analysis are discussed. Further, the paper discusses the role of geographical boundaries in labour market analysis and how spatial microsimulation can be combined with remote sensed data to estimate the precise residential location of household objects. Finally, the potential of microsimulation for spatio-temporal modelling of urban systems is addressed and a research agenda for dynamic spatial microsimulation modelling is outlined.

J. A. Bañares, F. J. Zarazaga, J. Nogueras, J. Gutiérrez and P. R. Muro-Medrano*

Construction and use of concept hierarchies from word taxonomies for searching geospatial data

The OpenGIS Consortium[1] uses the term Catalog to describe the set of service interfaces which support organization, discovery, and access of geospatial information. Catalog services help users or application software to find information that exists anywhere in a distributed computing environment. A Catalog can be thought of as a specialized database of information about geospatial resources available to a group or community of users. We are developing a Catalog in compliance with OpenGIS Catalog Services using Java as programming language.

The Catalog contains metadata that describe the capabilities and contents of the geospatial data. The use of standard representations of the metadata, such as the American Federal Geographic Data Committee (FGDC) standard or the ISO TC 211 standard (currently in draft version), makes the interoperability labours easier. The recommendations on metadata proposed by the Center for Earth Observation Programme (CEO programme) of the European Commission are being used in or-

**Department of Computer Science and System Engineering, Centro Politécnico Superior, University of Zaragoza, María de Luna 3, 50015 Zaragoza (Spain), Phone: 34 976 761928, Fax: 34 976 761914, {banares, javy , prmuro}@posta.unizar.es, http://iaaa.cps.unizar.es.*

[1] http://www.opengis.org

der to manage the controlled keywords. These lists of controlled keywords are been extended to support different kinds of thematic such as the Spanish political organization.

In order to complete the catalog capacity for searching information, some utilities able to derive new knowledge from metadata stored in the database are being built out of the OpenGIS Consortium Catalog specification. These utilities are based on artificial intelligence and information retrieval (IR) techniques and allow us to work with semantic relations among geographical coordinates and controlled keywords, and controlled keywords among themselves. This kind of techniques facilitates not only horizontal integration (same type of data for different areas) but also vertical integration (different types of data for the same area), which is a strong need in a distributed geolibrary.

Concept queries are inherently hierarchical in nature. We may expand a concept or keyword in the query to a large number of synonyms and narrower and general concepts related by *part-whole* and *is-a* relationships. For example, in order to find spatial references of a political unit, such as a village, the query may be expanded with the knowledge of political subdivisions such as the subdivision of a country in regions, a region in villages, etc. Data at the level of village may be not available, but data of superior levels which include the village may be sufficient to answer the question.

Concept extraction depends heavily on a good quality lexical knowledge base. Ontologies or lexical knowledge bases such as WordNet[2] and EuroWordNet[3], have in common that they are large-scale lexical databases and they were developed from a global point of view. WordNets are structured in terms of synsets (sets of synonymous words) with basic semantic relations between them (meronyms, hyponymy, ...). The semantic content and the large coverage of WordNet allows to use it to conceptual retrieval as opposed to match exact keywords.

The problem of the use of Anglo-Saxon concepts which are

[2]Princeton WordNet, Miller et al. 1990

[3]`http://www.hum.uva.nl/ ewn/`

difficult to translate to users from different cultural or linguistic background must be considered. This problem has been considered in EuroWordNet, where different language WordNets of lexicalized concepts are interlinked with a separate ontology of formalized semantic distinctions. This will allows relating concepts in different languages.

Our approach proposes to import, as Oracle Thesauruses, different geospatial keywords from Thesaurus, dictionaries and in special from the English and Spanish WordNets and some lists of controlled Spanish keywords using different sources as the codification of the Spanish Institute of Statistic[4]. These thesauruses would be completed with specific software to connect low levels with instances stored into relational tables as instances of the concepts represented by those nodes. Looking for the generic development of this software we are using the reflective capabilities of programming languages as CLOS or Java.

Medata contains keywords, and textual descriptions which allows retrieval by concepts. The best results in using WordNet for information retrieval are obtained with short texts, while standard IR techniques does not perform as well. With very short documents the opportunity to find the original keywords are lower than for average-sized documents.

We are developing a set of utilities for maintaining the controlled keyword lists and adding new ones. In order to exploit the hierarchical nature of these sources of knowledge information the InterMedia Text retrieval engine integrated with the Oracle8i database supports the use thesaurus operators in any SQL query.

[4]INE: `http://www.ine.es`

Lucy Bastin and Peter Fisher*

Modelling vegetation as fuzzy spatial entities: patches, stands and ecotones in Greek semi-natural phrygana

Semi-natural landscapes are those where natural vegetation succession processes are modified by management such as grazing or burning. Such landscapes can be of great conservation value, and spatial variation within them varies along a continuum from abrupt boundaries (between distinct landcovers, easily modelled as polygons), to gentler transitions (usually represented by raster pixels). Furthermore, the apparent grain of a landscape depends on the scale at which it is sampled. In the context of GIS, a surface must often be interpolated from point samples. It is therefore important to characterise the objects in the real landscape (discrete, homogeneous polygons or fields with blurred boundaries) and the landcovers which constitute those objects (mutually exclusive classes or overlapping clusters). This paper discusses two approaches used to model semi-natural vegetation in Northern Greece. Firstly, fuzzy clustering is used to classify vegetation samples, in an attempt to recognise both spatial mixing (transitions and interdigitations between vegetation types) and semantic mixing (overlap between defined plant species communities, such as 'heath' and 'wet heath'). Secondly, two interpolation techniques are evalu-

* *University of Nottingham, Dept. of Geography, Nottingham NG7 2RD, UK, lucy.bastin@nottingham.ac.uk.*

ated: co-kriging (using raster satellite data to inform the interpolation) and the mixed model of spatial interpolation (which uses ancillary vector segmentation). It is hoped that the second method in particular can make best use of both the vector model of reality (landcovers as discrete polygons) and the raster model of reality (landcovers as continuous fields).

Alain Becam, Maryvonne Miquel and Robert Laurini*

A distributed environment using ontology for the interoperability of urban data and models

Urban models simulate specific phenomena of the city and are a reliable decision-support for urban planners and decision makers. They are in general finished software products and are not designed in a way we can use some models together, using the results of one for input of another. They are heterogeneous, by the system they are running on, by their data formats, ... , and are used by experts, with special objectives. It may be interesting to interconnect different models, to profit their specificities. For instance for estimating traffic noise, we need to connect several urban models such as a home-to-work and home-to-service travel model, a traffic model and a model for generating noise levels.

In our project, we propose to encapsulate the models to transform them into standard models allowing their interconnection into a distributed environment. An encapsulation can be defined as a meta-model, a description of the model, specifying the method to use it, the data at inputs, outputs, parameters and finally the "semantic position" of this model. This meta-model is built using a standard language, following the XML

* *INSA of Lyon – LISI, 20, avenue Albert Einstein, F-69621 Villeurbanne Cedex, Fax : +33 4 72 43 87 13, abecam@lisi.insa-lyon.fr, miquel@if.insa-lyon.fr, laurini@if.insa-lyon.fr.*

syntax and must be "ontology compliant". This compliance is ensured by the procedure of description, which uses the elements of several ontology bases. A meta-model may include (1) a real model, (2) a specific model, which has a utility role into our system, (3) a virtual model, constituted only by this description, or finally (4) a super-model, a schema of interconnection viewed like a simple model.

Our system uses a Core Ontology, which covered a top-level ontology, from entity to top-level categories, and at least two distinct domain ontologies, for the semantic elements, such as roads, rivers, towns and for the typological elements, such as integers, characters, ... This distinction allows the user to adapt the data independently of the semantics implications. Indeed, we think one expert may have a better view of the domain than our system. Also, some specific domain ontologies may be added, relevant or not to an upper ontology. So, a very specific ontology may be added, which is independent from another ontologies, assuming we do not need any reference to the top-level ontology to use the models covered by this specific ontology. Of course, this possibility is practical to easily describe some models, taking into account only the actual elements of these models, but this ontology is absolutely not shareable. Thus, we must offer an easy way to describe a model using existing ontologies and to properly extend existing ontologies.

Furthermore, the system distinguishes the Core Ontology, Extension Ontologies, User Ontologies and the Session Ontology. The Core Ontology is the consistent base, immutable and sufficient for simple typological adaptation and semantic descriptions. The Extension Ontologies are domain specific, referencing the Core Ontology, adding strong specifications relevant to one domain. The User Ontologies are relevant or not to the Core Ontology and/or to one or more Extension Ontology, adding some missing elements for the user. Finally, the Session Ontology is a temporary ontology, linked or not with any other upper ontology, designed by the description of the meta-models. This last one allows a proper extension of the domain knowledge, adding no permanent weak information. If the user finally has confidence in his Session Ontology, he may process this Ontology to turn it into an User Ontology.

These ontologies may be written or graphically constructed. For the integration of one standard, such as SDTS or FGDC, we supply two distinct possibilities: (1) to describe the way the system may use one foreign description of this standard and (2) to construct a tool to translate this description into one ontology for our system. Allowing the adding of some new ontologies is necessary, because of the impossibility to describe all the universe, but dangerous. The hybrid ontology composed with the linked ontologies may include redundancy of definitions, useless informations and semantic conflicts. So, some tools must search for possible problems and try to resolve them. Furthermore, each extension of the core ontology may exist in different versions, considered at parallel extensions and benefits of one degree of confidence. This information about confidence is used to restrict the extensions according to the level. One extension constructed on the base of another extension inherits of this last confidence.

Specified in this way, the meta-models can be used in our environment of interconnection, which uses a three tiered architecture to interconnect standardised models. The upper layer, the schema for our user, exploits the medium layer, which interconnects the meta-models, to respect this schema and, finally, the medium layer uses the bottom layer to execute the models covered by the meta-models involved in the current session.

In the bottom layer, the models are heterogeneous, accompanied by their descriptions, on a computer allowing the use of Java RMI. The encapsulation is the mapping from this physical layer to the second layer, the medium layer. Using the description of one model, the system is able to raise an appropriate instance of encapsulation (Fig. 5.1), using the model just like a specific device.

The medium layer is a powerful environment of interconnection, using an object approach and an ”all model” view. At

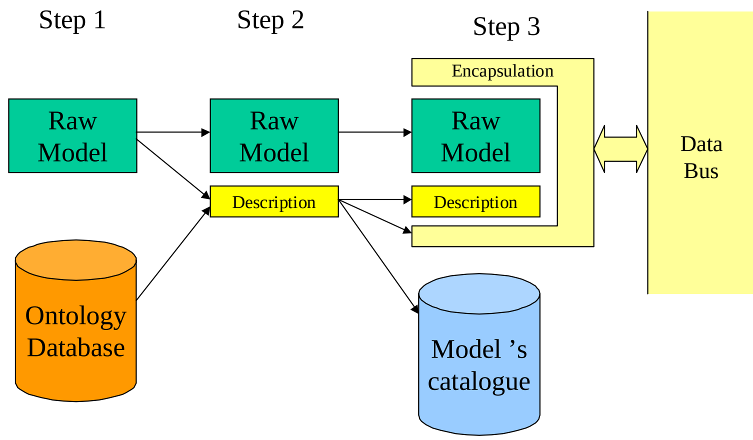

Figure 5.1: One model, from its raw form (1) to its encapsulated form (3). Using the Ontology Database and a language of description, the description of the model is added, the model with its description being a meta-model (2). With this description, the model server can register the model into a catalogue of models and the supervisor can query an encapsulation (3).

this level, we also have facility models, for adaptation of data, human-machine interface, ... The different elements of this level are all used in the same way like standard models. Some servers allow the execution of a session, at least one global server, which registers the models and prepares the interconnection; one model server on each computer with a model, which manages its local models, and one supervisor for one session. This last one may request some instances of encapsulation to follow a schema of interconnection, the third layer, the logical layer. Also, the supervisor is the link between the user and the system, and, following the user choice, it may try typological adaptation of data and semantic adaptation and verification. The supervisor accesses the ontological bases and, knowing its session's models, checks data typology and semantics. It invokes instance of special models for adaptation of data, if needed, in the form of mediators and wrappers, and, if possible, executes the simulation, using the medium layer to follow the logical schema.

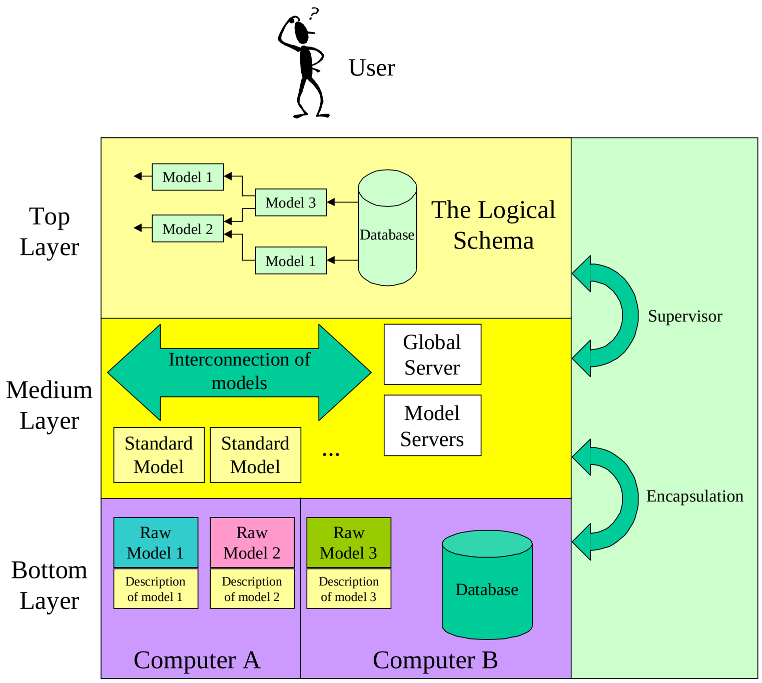

Figure 5.2: The three layers of our environment. The final user only interacts with the top level. The encapsulation process transforms the set of heterogeneous models into standard models. Some facility models may be added, such as database access models. These models are interconnected and managed by some servers, which can be located on remote computers through Internet. Finally, the supervisor utilises these models to follow the logical schema constructed by the final user.

Stefania Bertazzon*

(Re)-defining the concept of spatial contiguity: the metaspace of applied analysis

The need for analyzing the concept of spatial contiguity emerges from applied work in spatial regression analysis. Unlike standard regression, spatial regression applies to spatial units, which are irregular in shape and distribution. A contiguity matrix is the structure used to assign weights to each unit, in order to calibrate such spatial irregularities. Indeed, the entire model estimation relies on the definition of a contiguity matrix. Two or more regions are defined contiguous if they share a border. Such concept of contiguity is often inappropriate for spatial interaction, and particularly spatial regression models: shared borders can hardly represent nearness; borders may not be defined; and contiguity often has to be established across sets of different regions. In spatial regression analysis the dependent variable and the set of explanatory variables often are measured at different locations, as they are attributes of different spatial features. In socio-economic applications (e.g. shopping patterns, recreation decisions) population data (origin) are measured on spatial units such as census subdivisions, while destinations (shops, tourist attractions) are points or other features, however distinct from census subdivisions. The same problem may occur in environmental applications (e.g. bioaccumulation model of marine organisms): in this instance, the locations at which organisms

* *University of Venice, Ca' Foscari, sbertazz@unive.it.*

are sampled do not coincide with those of sediment samples. Moreover, samples are taken at points, which are assumed to be representative of an area, or region, but no borders are defined for such regions.

In an attempt to overcome these problems, contiguity is often re-defined based on alternative measures of distance, such as travel time in socio-economic applications (which still requires requires an ad hoc specification for each model). An alternative approach is the definition of contiguity, or nearness, derived from spatial autocorrelation (or semivariance) analysis. In both cases, nearness is no longer a relationship among entities in space, but an attribute of them, depending on other attributes and on (implicit) measurements of distance. It might be defined by the researcher (who decides weather contiguous regions lie within 1 or 10 hour traveling time) or it is an explicit function of attribute values and distance, as in spatial autocorrelation.

A re-definition of spatial contiguity can be provided based on an ontological view of the model space. The sampling process (in environmental models), or the identification of origins and destinations (in socioeconomic models), produce an ontological transformation of space. The space thus obtained exists only because and only where attributes and relationships (e. g. organism and sediment samples, origins and destinations) exist. It is a metaspace defined by attributes of entities in the "real" space. The metaspace is discrete, as it is formed by the objects (attributes and relationships); attributes of space become properties of the metaspace. Thus new distances, metrics, and topologies can and must be defined on the discrete metaspace of objects. Contiguity can thus be re-defined on new conceptual bases: application-specific parameters may still be required, but having defined consistently a metaspace, standard criteria can be esuggested for an operational definition of contiguity within the ontological framework of metaspace.

Steffen Bittner*

Levels of reality in a cadaster

All over the world great effort is invested in the maintenance of cadastral systems. There is a major demand to construct efficient cadastral systems (Dale and McLaughlin, 1989). The construction of efficient cadastral systems is founded in a deep understanding of the reality, which the system should correctly represent. In the field of cadaster institutional concepts, e.g. property and ownership, play a major role. They determine the structure of reality in a cadaster. A theory of the institutional structure is fundamental to represent the relevant aspects of reality in a cadaster.

In the past formal models where not very successful at modelling institutional reality. The goal of the work of the author is to create a formal and executable model of reality in a cadaster. Searle's theory of institutional reality (Searle, 1995) introduces an approach, which is very promising for this project.

This abstract applies Searle's theory to the field of cadaster. It introduces the distinction between the physical and institutional level of reality in the analysis of cadaster and argues that both are essential for a model of a cadaster.

For the formalization on the foundation of Searle's theory it is a crucial point to represent human intentions and behaviour. Agent theory (Ferber, 1999; Wooldridge and Jennings, 1995) gives the necessary tools for this task. The author hopes that Searle's theory applied in an agent- based model is the foundation to successfully represent the relevant elements of reality in

* *Department of Geoinformation, TU Vienna, Austria, stebi@geoinfo.tuwien.ac.at.*

a cadaster so that it helps to deepen our understanding of the issues involved in the domain of cadaster. The work presented here is a step on the way to reach this goal, it is the necessary precondition for the formalization.

In his book *The Construction of Social Reality* Searle distinguishes *physical* from *institutional reality* and within these parts *brute physical facts* from *institutional facts* (1995). Brute physical facts exist in external reality independent of human observers and human intentions. In opposition institutional facts exist only by human agreement, they are created by the assignment of status functions to physical facts. For instance collective intentionality assigns the status function 'money' to a rectangular piece of paper. A physical fact is always the foundation for an institutional fact.

According to Searle's theory the *institutional* and *physical level* of reality must be distinguished in the analysis of a cadaster. Why is it important? The example of an ownership transfer of a parcel between two persons will be used in the following to explain the ideas: On the institutional level, one legal person, the owner of a parcel, transfers ownership of the parcel to another legal person and they sign a deed as proof for this transaction. On the physical level there are two human beings writing on a piece of paper. But only in a special context the event 'two people writing on a piece of paper' counts as signing a contract, which causes the transfer of ownership. What happens if people make mistakes, for instance, if they mean different parcels? On the physical level the situation is the same: two persons writing on a piece of paper. But on the institutional level the preconditions for the transfer of ownership are not satisfied and therefore no transfer happens. Activities on the physical level are different to activities on the institutional level. A model of reality in cadaster has to deal with both aspects.

A crucial aspect in Searle's theory is that the rules creating the possibility of institutional reality, the *constitutive rules* (Searle, 1995, p. 80), can be codified (Searle, 1995, p. 87) in laws. Cadastral systems are highly determined by laws, especially by the cadastral law. By the analysis of the cadastral law it is possible to extract the rules constituting institutional reality. For example it is possible to extract the rules for the creation of ownership. According to Searle the rules for creating ownership (the constitutive rule) must be distinguished from the fact ownership itself. This distinction of facts and rules on the institutional level does hold on the physical level as well. For instance the rules defining the capability of a human being to write on a paper is different from the event of writing itself. Rules that create the possibility of activities are called *powers* (e.g. the power to write on a piece of paper) on the physical level and *rights* (e.g. the right to transfer ownership of a parcel) on the institutional level (Zaibert, 1998).

For an analysis of reality in a cadaster it is necessary to distinguish facts and rules on the institutional level and on the physical level. Status functions can be assigned to different ontological categories of phenomena: subjects, objects and events (Searle, 1995, p. 97). Within these categories there exist facts on both levels of reality.

The following overview shows the ideas according to the example of an ownership transfer.

- The physical level
  - Facts:
    - Subjects: human beings
    - Objects: paper
    - Events: writing on a piece of paper
  - Rules:
    - Powers: the physical capability of a person to write on a piece of paper.
- The institutional level
  - Facts:
    - Subjects: owner, legal person
    - Objects: deed

    - Events: transfer of ownership
  - Rules:
    - Constitutive rules: Signing a contract counts as transfer of ownership.
    - Rights: The owner of a parcel has the right to transfer ownership.

The example shows the structure of the issues involved in the analysis of a cadaster founded on the basic distinction of institutional and physical reality based on the example of an ownership transfer. Different countries have distinct cadastral systems but they are different solutions to a similar problem (Frank, 1996). The structure of reality in a cadaster determining the system is comparable. Therefore the analysis is not limited to this example or a specific cadastral system in a specific country and can be generalized.

Bibliography

Dale, P. F., McLaughlin, J. D., 1989. Land Information Management, An introduction with special reference to cadastral systems in Third World countries. Oxford University Press, Oxford.

Ferber, J., 1999. Multi-Agent Systems. An introduction to Distributed Artificial Intelligence. Addison-Wesley.

Frank, A., 1996. An object-oriented, formal approach to the design of cadastral systems. In: Kraak, M. J., Molenaar, M. (Eds.), Advances in GIS Research II – Proceedings of 7th International Symposium on Spatial Data Handling. Vol. 1. IGU, Delft, The Netherlands.

Searle, J., 1995. The construction of social reality. The Free Press, New York.

Wooldridge, M., Jennings, N. R., 1995. Intelligent agents: Theory and practice. Knowledge Engineering Review 10 (2).

Zaibert, L. A., 1998. Real estate as institutional fact. Tech. rep., http://wings.buffalo.edu/philosophy/farber/zaibert.htm.

Thomas Blaschke[*]

Operationalisation of the patch-concept in landscape ecology

Landscape mosaics are described by the landscape components of patches, corridors, and the surrounding matrix. Patches, corridors and matrix directly influence the spatial patterning and flows in a landscape. Spatial scale also greatly affects landscape structure, heterogeneity, and connectivity. Landscape structure is determined by the flow of materials, animals, energy, and water through the landscape elements of patches, corridors, and matrix. Factors such as patch size and shape, corridor characteristics, and connectivity work together to determine the pattern and process of the landscape. The correlation between pattern and process results in an interdependency between landscape structure and function. Landscape patterns influence process, which in turn affect the patterns. The feedback between structure and function is evident in the landscape in the world around us.

The basic element for many spatial disciplines is a patch or a landscape element. In landscape ecology, patches are spatial units at the landscape scale. Patches are areas surrounded by matrix, and may be connected by corridors. The geomorphology of the land interacting with climate factors, along with the other factors such as the establishment of flora and fauna,

[*] *University of Salzburg, Department of Geography & Geoinformation, thomas.blaschke@sbg.ac.at.*

soil development, natural disturbances, and human influences work to determine patch size, shape, location, and orientation. The size, shape, and nature of the edge are particularly important patch characteristics. Patch size can affect species habitat, resource availability, competition, and recolonization. Spatial scale is especially important when dealing with patches because an area large enough to be a patch to one species, may be a barrier or insignificant to another species.

Patches and corridors are imbedded in the matrix, which is usually the most extensive and connected landscape element present. However, the matrix may play a dominant role in the functioning of the landscape without being the most extensive landscape element. Determining what is the matrix in a landscape depends on either connectivity, dominance, or function. Landscapes vary greatly in their degree of heterogeneity. Factors which influence heterogeneity are aggregation, contrast and porosity. Aggregation is the degree to which patches are clumped together. Contrast refers to the diversity of patches, patch richness, number of patches, evenness (lack of dominance of one patch type). Measures of heterogeneity must include the vertical and horizontal structure of landscapes. The degree of heterogeneity in a landscape plays a crucial role in determining the distribution and the habitat use by organism, as well as the abiotic functioning of the landscape.

As Bittner and Winter (1999) point out, for a better understanding of the relationship between objects of the real world and their representations, a better understanding of the underlying ontological and epistomological foundations is necessary. As described above, the author has approached this research field from an application's point of view and is still searching for answers. Current research tries to extract 'meaningful' objects from remotely sensed data through multiresolution image segmentation and object-oriented classification (Blaschke et al., (in press). So far, image segmentation seems to be the only operational solution to provide context-information to pixels aiming to overcome limitations of the 'pixel-paradigm'. The definition is usually crisp generating hard boundaries, although Cheng (1999) presented an ambitioned 'fuzzy spatial objects' approach. In current research, a group at the University of Salzburg is aiming for the definition and description of landscape elements incorporating internal heterogeneity. Heterogeneity or the differences and diversity within a landscape is a basic concept in landscape ecology. In fact, there are at least two levels of heterogeneity: The degree of complexity of landscape pattern (composition) and diversity within landscapes elements or stands. A natural forest landscape, for example, normally includes a variety of species of trees, shrubs, herbs, animals, and microorganisms, as well as a diversity of ecological stand types, varying according to moisture, slope, elevation, aspect, soil, and so forth. This kind of natural diversity is important to ensure that all the parts are available for forests to function. But since vegetation scientists have elaborated methods to calculate species diversity, there are no standard techniques to estimate a landscape ecological within-patch diversity. The concept of diversity can work on many different scales. Each species, for example, operates on its own scale. What appears to be a uniform patch of habitat to a large species, such as bear or Douglas fir, may comprise a very diverse, patchy environment to a small species, such as a bark beetle or a mushroom. Hence, a multiscalar, multi-hierarchical approach is developed to describe a landscape through its patch-matrix and the texture of the patches as well as the texture of the landscape.

Bibliography

Bittner, T., Winter, S., 1999. On ontology in image analysis. In: Agouris, P., Stefanidis, A. (Eds.), Integrated Spatial Databases. Vol. 1737 of Lecture Notes in Computer Science. Springer, Berlin, pp. 168–191.

Blaschke, T., Lang, S., Lorup, E., Strobl, J., Zeil, P., (in press). Object-oriented image processing in an integrated gis/remote sensing environment and perspectives for environmental applications. In: Cremers, A., Greve, K. (Eds.), Umweltinformation für Planung, Politik und Öffentlichkeit. Metropolis Verlag, Marburg.

Cheng, T., 1999. A process-oriented data model for fuzzy spatial objects. ITC publication series 68, ITC, Enschede.

Bénédicte Bucher*

Users access to geographic information resources: a model of tasks and roles to specify intentional uses regarding available resources

The exploitation of geographic information relies on the interoperation of several components: human (like cartographers, GIS experts, end users), or implemented components (data and processes). Numerous researches in the field of geographic information actually aim at enhancing the exploitation of geographic information by improving interfaces between those components that are to work together (Abel et al., 1999; Buehler and McKee, 1998). End users are crucial. They indeed can express needs for intentional use of geographic information that should lead to the development of new applications. But their ability to do so, their access to geographic information resources, i.e. data and processes, and to their utility, is still problematic. To improve this, we think that a necessary step is to help the users understand what they can do with these available resources. We aim at facilitating the following process: a end user s expressing a need for geographic information, and specifying it by confronting it to the domain of geographic information, i.e. available data and processes, corresponding handling ex-

**COGIT – IGN , 2 av Pasteur , 94 165 St Mande Cedex , France, benedicte.bucher@ign.fr.*

pertness. It is not intended to build a system that processes geographic data to produce knowledge on demand, but rather a knowledge-based system that identifies the necessary resources, i.e. data and processes, which are inputs, i.e. useful knowledge, to build an application that would produce the knowledge required by an end user.

An analysis of what supports the transfer of knowledge from a component to another, more specifically from information resources to users, and more specifically in the geographic case, has led us to formalise two main specificities of our intended system.

1. It is functionally determined as to help users, different in their domain and in their expertness, access different information resources. This has been formalised as follows. The system should have a language to describe available heterogeneous geographic information resources as bricks of procedural reasoning. And it should hold terms for a user to perform a formal reasoning, co-operatively with the system, combining those bricks to conceive an application that should reach his working objective.

2. The body of knowledge about the existence of geographic information resources, their handling and their possible uses, is huge. We do not put the focus on enumerating it but rather on building a flexible system to further integrate this relevant knowledge. At the moment, our research objective is to evaluate the feasibility of such an approach by finalising a prototype of this system. The key element of the prototype is the underlying knowledge model that should fulfill both above requirements.

Our issue can be put back into the context of knowledge sharing and reuse which is currently seen as a major challenge in Artificial Intelligence (AI) researches (Chandrasekaran et al., 1998; Gomez and Benjamins, 1999). We apply lessons learned in AI to design our system. Those lessons are to integrate ontologies and problem-solving knowledge (Gomez and Benjamins, 1999).

Building ontologies about geographic information is a complex undergoing process (Egenhofer and Mark, 1995; Frank, 1997). We thus focus on : representing problem-solving knowledge in our system, and further integrating existing ontologies. We use an adequate modelling technique, that settled by the ESPRIT project CommonKADS which encompasses the notions of task, inference, and role (Schreiber et al., 1999).

Geographic tasks are objectives that can be reached through the exploitation of geographic information. As recalled in Cauchard (1999), three levels of abstraction are usually identified in tasks :

- Intentional tasks answering what-to-do- questions, e.g. to draw a map, to detect entities close to a given entity.
- Functional tasks answering how-to-do-questions, e.g. to determine a location, to measure a proximity.
- Operational tasks answering with-what-questions, e.g. to change co-ordinate systems, to overlay representations.

Depending on his expertness in GIS and geographic databases, a user may express his working objective by specifying an intentional, functional or operational task. Geographic inferences are what produce geographic knowledge: production rules and GIS processes. The static information resources, data, are described thanks to the language of the domain and to that of roles, which are the functional specification of inputs and outputs of the inferences. Eventually, the answer of the system to a user, who has specified a task, is the description of a method to reach the goal of the task and the corresponding information resources: GIS processes to drive the inferences, and data to fulfil the roles.

Bibliography

Abel, D., Gaede, V., Taylor, K., Zhou, X., 1999. SMART: Towards spatial internet marketplaces. Geoinformatica.

Buehler, K., McKee, L., 1998. The OpenGIS (r) guide: Introduction to interoperable geoprocessing and the OpenGIS specification. Tech. rep., Open GIS Consortium Technical Committee.

Cauchard, V. F., 1999. Realisation d un systeme d aide a la conception d applications de traitements d images : une approche basee sur le raisonnement a partir de cas. these en informatique, Universite de Caen.

Chandrasekaran, B., Josephson, J. R., Benjamins, V. R., 1998. The ontology of tasks and methods. In: KAW 98.

Egenhofer, M. J., Mark, D. M., 1995. Naive geography. In: Frank, A. U., Kuhn, W. (Eds.), Spatial Information Theory. Vol. 988 of Lecture Notes in Computer Science. Springer, Berlin, pp. 1–15.

Frank, A. U., 1997. Spatial ontology: a geographical information point of view. In: Stock, O. (Ed.), Spatial and Temporal Reasoning. Kluwer Academic Publishers, pp. 135–153.

Gomez, A., Benjamins, R., 1999. Overview of knwoledge sharing and reuse components: Ontologies and problem-solving methods. In: International Joint Conferences on Artificial Intelligence.

Schreiber, A. T., Akkermans, J. M., Anjewierden, A. A., de Hoog, R., Shadbolt, N. R., van de Velde, W., Wielinga, B. J., 1999. Knowledge Engineering and Management, The CommonkADS Methodology. MIT Press.

Carola Eschenbach[*]

Ontology, predicates, and identity

The ontology of a theory or of a formal model is tightly connected to the use of predicates (concepts, attributes, features, relations) applying to the entities or relating the entities, and to criteria of identity and distinction that allow to recognize an entity.

In the study of spatial ontology in geographic domains several levels of entities can be considered. There is a purely spatial, geometric level of entities that are identified on the basis of geographic coordinates. Concrete, material spatial entities, such as rocks or portions of water occupy space and can change place, i.e. they can move. Natural spatial units such as mountains, lakes or woods are formed by material entities based on spatial properties. Furthermore, political and social units (cities) have a location or territory, they are also tightly connected to material entities (streets, houses).

Given that a geometric model of space is provided by a Geographical Information System, the modeling of the spatial predicates and the other ontological levels is critical. In addition, there is a strong interdependence between spatial predicates and spatial entities. For example, the information about wooded areas can be represented with different ontological backgrounds. A spatial predicate `wooded-area` can be attributed to entities (r) of the geometrical level as in (10.1). A more complex rep-

[*] *University of Hamburg and Institute for Advanced Study Berlin, eschenbach@informatik.uni-hamburg.de.*

resentation is provided in (10.2). Woods are taken as natural spatial units that are represented in addition to the regions they occupy. The expression $\texttt{wood}(u)$ expresses that the natural unit u is a wood, and $r = \texttt{loc}(u)$ expresses that u occupies the spatial region r.

$$\text{wooded-area}(r) \tag{10.1}$$

$$\text{wood}(u)\ \&\ r = \text{loc}(u) \tag{10.2}$$

If the second representation is chosen as the basic form, the predicate `wooded-area` could be defined as in (10.3): A region is a wooded area, iff it is part of the location of a natural spatial unit that is a wood.

$$\begin{aligned}&\text{wooded-area}(r) \equiv \\ &\quad \text{exists } u\ [\text{wood}(u)\ \&\ \text{part-of}(r, \text{loc}(u))]\end{aligned} \tag{10.3}$$

On the other hand, if the predicate `wooded-area` is chosen as the basic attribute, an attribute that applies to maximal connected wooded areas can be defined based on the topological structure of regions as in (10.4) (it says that a maximal connected wooded area r is a wooded area, self-connected and maximal in the sense that every wooded self-connected area that is connected to r is also part of r). However, maximal connected wooded areas as defined in (10.4) and the natural spatial units in the sense of (10.2) need to be distinguished. The first ones are wooded areas, the second ones occupy wooded areas.

$$\begin{aligned}&\text{max-con-wooded-area}(r) \equiv \\ &\quad \text{wooded-area}(r)\ \&\ \text{self-connected}(r)\ \& \\ &\quad \text{forall } r'\ [\text{wooded-area}(r')\ \&\ \text{self-connected}(r')\ \& \\ &\quad \text{connected}(r', r)\ \rightarrow \text{part-of}(r', r)]\end{aligned} \tag{10.4}$$

The representation of woods in addition to spatial regions (as in 10.2) is ontologically more complex than the simple form (10.1). Reasons to use the complex version can derive from two sources. If the temporal dimension has to be modeled in addition to the spatial dimensions, then the representation of natural spatial units is essential for assuming their development to be independent of the development of the region they occupy at some moment. To represent woods as natural units that move, grow, or shrink, makes it possible to assume that the conditions of diachronic identity of woods are not purely geometrical. (Obviously, in this context the representations of (10.1)-(10.4) have to be elaborated regarding the temporal dimension.)

Further reasons to represent individual woods derive if two woods can exist at the same time that cannot be distinguished or delimited by only geometric features. Thus, the ontologically more complex form of representation is mainly justified if criteria of identity and distinction are substantial for further reasons.

This talk will discuss the formal background of the two ways of representing spatial information and their interaction. It is meant to inspire the discussion of the following topics from formal and applicational points of view.

- The role of identity for the ontology of a theory, formal model or GIS
- The interaction between identity and predicates for ontological decisions
- Different levels of spatial entities based on different criteria of identity
- Options to express (diachronic) identity and (synchronic) distinction of spatial objects based on geometry and spatial predicates
- Independence of identity of spatial entities from geometry and spatial predicates
- Ontologies of vague spatial objects and the role of vague spatial predicates
- Dependencies between synchronic and diachronic aspects of identity
- Natural language ontology, predication and identity
- Cognitive ontology, predication and identity

- Geographic / scientific ontology, predication and identity
- Do natural language, cognition and science conflict or supplement each other regarding identity?
- The role of identity for Geographical Information Systems
- Consequences of conflicting criteria of identity

Sara Irina Fabrikant [*]

The ontology of semantic information spaces

Information visualization has emerged as a subfield within the human-computer interaction (HCI) community to facilitate access to large, complex databases. Graphic depictions of large non-spatial databases are not only increasingly based on the spatial metaphor, but many examples are also explicitly geographic. These representations are known as graphic spatializations, or information spaces. Abundant spatialized depictions exist in the literature that document the rapid developments within this relatively young research area (see for example Card et al., 1999). However, a structured approach to formalize the employed spatial and graphic transformations seems to be lacking.

Two major concerns can be identified: the use of space as a data generalization strategy, and the use of spatial representations or maps to depict these data abstractions. First, the majority of the spatialization examples are only concerned with the use of formal properties of space, such as location and distance between items in the information space, specifically of the metric, Euclidean kind. This is particular problematic when information spaces resemble geographic space. Geographic space is more than just Euclidean geometry. Entities and their relationships in space carry experiential and socially constructed meanings. As argued in this paper, usable information spaces need also to be based on a sound semantic abstraction frame-

[*] *Department of Geography, University of California Santa Barbara, Santa Barbara, CA 93106-4060, USA, sara@geog.ucsb.edu, www.geog.ucsb.edu/s̃ara.*

work, including cognitive aspects of space. Second, information space depictions come in an infinite variety of ways. Often these depictions are called maps, but lack a coherent representational framework as provided for example by cartographic design principles. Unfortunately, spatialization research seems mostly taking place without the participation of GIScientists, who, with their geospatial domain and mapping knowledge would be perfectly suited for the task. A few exceptions within the GI-Sciences are the works of Couclelis (1998); Fabrikant (2000); Fabrikant and Buttenfield (1997, (submitted); Kuhn and Blumenthal (1996); Skupin and Buttenfield (1996, 1997); Tilton and Andrews (1994).

This paper aims at creating an ontology of semantic information spaces, using empirical evidence on its usability as a starting point. Results of usability evaluations not only allow deriving explicit design guidelines, but also enable to construct of a solid theoretical and representational framework. A theory of semantic information spaces, as argued in this paper, should be based on ontological, semiotic and semantic considerations. An ontological approach helps to conceptualize the entities populating an information space (e.g. documents in an archive). Explicit formal knowledge of entities and their relationships must be determined prior to applying the metaphorical mapping. The semantic strategy deals with identifying sound metaphorical mappings to encapsulate the meaning that needs to be preserved within the information space. Finally, semiotic considerations assure that information spaces are depicted following a sound representational strategy.

Two research areas can benefit from this approach. First, semantic information space design is the basis for effective and unambiguous communication between information providers and information seekers. Generalization through semantic abstraction of data archive content will increase in importance if data archives are expected to grow exponentially. A sound theoretical spatialization framework enables information designers to construct conceptually robust and usable information spaces and allows information seekers to more efficiently extract knowledge buried in large digital data archives. On the other hand, a representational framework of space grounded on ontological and semantic principles can be transferred to the explicit geographic domain as a basis to reduce current limitations of how geographic space is represented within GISystems.

Bibliography

Card, S. K., Mackinlay, J. D., Shneiderman, B., 1999. Readings in Information Visualization: Using Vision to Think. Morgan Kaufmann, San Francisco, CA.

Couclelis, H., 1998. Worlds of information: The geographic metaphor in the visualization of complex information. Cartography and Geographic Information Systems 25 (4), 209–220.

Fabrikant, S. I., 2000. Spatialized browsing in large data archives. Transactions in GIS 4 (1), 65–78.

Fabrikant, S. I., Buttenfield, B. P., 1997. Envisioning user access to a large data archive. In: Proceedings, GIS/LIS '97. Cincinnati, OH.

Fabrikant, S. I., Buttenfield, B. P., (submitted). Exploring semantic spaces: Geographic metaphors for information access. Annals of the Association of American Geographers.

Kuhn, W., Blumenthal, B., 1996. Spatialization: Spatial Metaphors for User Interfaces. Vol. 8 of GeoInfo Series. Department of Geoinformation, Technical University of Vienna, Vienna, Austria.

Skupin, A., Buttenfield, B. P., 1996. Spatial metaphors for visualizing very large data archives. In: Proceedings GIS/LIS '96. Denver, CO.

Skupin, A., Buttenfield, B. P., 1997. Spatial metaphors for display of information spaces. In: Proceedings AUTO-CARTO 13. Seattle, WA.

Tilton, D. W., Andrews, S. K., 1994. Space, place and interface. Cartographica 30 (4), 61–72.

Sébastien Gadal and Georges Nicolas[*]

Locus: fourth geographical dimension of the geomap

Abstract. In the field of geography, we have been examining the contribution of the morpho-genetic representation to the comprehension of the cycle: reality, representation, model, reality. A distinction is made between map and geomap based on the notion of *locus-object*.

1. Foundations and limits of cartographic representations

1.1. The limits of the interface "man / information" in cartography

The maps allow to reproduce, take decisions or act on the surface of the Earth. The user has "to recognise" represented information and give a geographical "concrete" sense to the signs used. On ancient maps these (the signs) were directly inspired by the objects. These signs have become abstract and by the use, are fossilised. Furthermore, a symbolic sense is superposed to the semiologic meaning of forms, generated by the use of signs. The symbolic of signs employed and their permanence have therefore a constraining role in the perception of results. Maps have

[*] *CAMS/LATES (UMR 8557), EHESS, 54, Bd Raspail F-75270 Paris cedex 06, gadal@ehess.cnrs-mrs.fr. Georges Nicolas, University of Lausanne, Switzerland, 15, rue Alfred de Musset, CH-25300 Pontarlier, nicorad@fc-net.fr.*

become ”maps-images”.

1.2. Supplementary constraints imposed by the numerical

The ” digital map images” are submitted to all constraints of the ”classic map images” (on paper or other materials). However, the passage of the analogical information on the ”classic” map to the digital information on the ”digital” map has displaced constraints. The immediate symbolic sense is nearly absent but can remerge at any given time. The multiplication of placement form possibilities, of relations and representation in real time, allows to highlight some geographical objects’ properties not immediately visible. The ”digital map image” generates new computing and mathematics problems that are linked with the characteristics of geographical objects.

1.3. The rappport object / form / structure in digital cartography

The image of remote sensing requires the fabrication of very elaborate intermediate layers between the information received and its utilisation. It requires digital information which has no direct report with the usual form of objects. Consequently, we are brought to examine the sense of the discontinuities that appear in their numerical imagery restoration. Does it concern differentiation in an object or differentiation between different objects? Therefore, it is necessary to understand their geographical meaning of limit or observable frontiers. And at this point appear the epistemological, or even, ontological problems.

2. Location and map, locus and geomap

2.1. The locus-object in geographies

All objects represented on a digital or classic map have a locus because on Earth there cannot exist object without locus. We can however conceive a locus without any object: the *empty locus* or *empty object locations.*

Four cases can appear for the couple: locus = λ, object = o:

1. $\lambda_1 \neq \lambda_2$ and $o_1 \neq o_2$: strong differentiation between the locus and the object,
2. $\lambda_1 \neq \lambda_2$ and $o_1 = o_2$: weak differentiation between the locus,
3. $\lambda_1 = \lambda_2$ and $o_1 \neq o_2$: weak differentiation between the object,
4. $\lambda_1 = \lambda_2$ and $o_1 = o_2$: lack of differentiation (undifferentiating).

A same location can contain several objects, and even several objects in a same object. The *location* of an object is therefore the location of its *locus*: the information given in each location has to be therefore semantically and spatially differentiate.

2.2. The map and the geomap

The "locus-object" and the location are therefore distinct but can be employed simultaneously in geographical representations. If the locus is differentiate, but not the object, we say that the couple is *weakly differentiate* by the locus. The information generates an analytic map as a *morpho-genetic initial map.* If the object is only differentiate, we say that the couple is *weakly differentiate* by the object. We make a synthetic map with multiple information on the object in each location like the *morpho-genetic intermediary map.* Finally, if locus and objects are both differentiate, we say that the couple is *strongly differentiate* by locus and objects, as *morpho-genetic final maps.* These are the *geomaps.* If locus and objects are undifferentiate, we say that the couple is undifferentiate. Representations that do not differentiate neither the place nor the object, generate images of the undifferentiation: for example the axis traced on a white leaf.

3. The morpho-genetic method of geographical representation

3.1. What is represented?

A current problem in classic cartography consists in trying to represent in the same location several different objects which have each a locus. The inverse problem is faced when somebody uses only electromagnetic impulses as in the case of morpho-genetic representations: How can we identify the places of concrete objects from the variations of a single type of information about a "locus-object" having multiple locations? The detection of spatial discontinuities in the information about this "locus-object" calls for the following method of calculation : Karhunen-Loeve's transformed aims on the one hand at concentrating the maximum of statistical information in terms of variance of the same axis, and on the other transforming morphologically the decrypted image.

The use of this transformed allows the recognition, origin of spatial discontinuities and their indexation by means of a digital attribute. Every discontinuity becomes then a spatial entity which expresses the couple "locus-object" in each part of the discontinuity segment, with a measure and a geometry. Finally, the locations of intrinsic spatial discontinuities confer on every object a morphology and a differentiation in comparison to the other objects. Then, the differentiated object sends back (dismisses) to the location of locus. The morpho-genetic method of geographic representation generates so maps of spatial discontinuities which are representations of structures of the geographic space. The final result of the exploitation of these maps is a morpho-genetic geomap of "locus-objects".

3.2. The geographical circuit: reality, information, representation, reality

The "image maps" created by the morpho-genetic method of geographical representation give a geomap perception of the space by the introduction of "locus-object" differentiation criteria. They can be either considered as the basis for computing reality or as an action instrument. However, this reconstruction of the reality is conditioned by the physical and geometrical characteristic measures of the sensor. It derives fundamentally from the "spatial statistical unit / geographical represented information dialectic ("location / locus-object" dialectic). The data used influences considerably the representation of the space and its reality and gives as many representations of the observable realities as possible.

4. Conclusion

"Classic" or "digital" geographical graphic representations are of the same type. However, the "image maps" created by the morpho-genetic method of geographical space representation allow to integrate a succession of maps and geomaps in a unique operative circuit. What is habitually separated in classic representation can be realised by the same user. The notion of "locus-object" allows understanding the established circuit between all geographical representation types (maps and geomaps) with the employ of the morpho-genetic representation method. Consequently, the locus is present at all the stages of the cycle controlled by the observer: reality, information, representation, reality and can therefore be considered as the *fourth implicit dimension* of the map and the *fourth explicit dimension* of the geomaps.

Myke Gluck[*]

Ontologies for geographic knowledge discovery: augmented seriation and dynamic interfaces

There currently is a poor cognitive fit between tool functionality for search and exploration in geographic multimedia datasets and the user's efficient and effective completion of geographically-oriented tasks. These issues were examined at a recent US National Science Foundation Varenius Project specialist meeting entitled 'Discovering geographic knowledge in data-rich environments' held in Seattle, USA in April 1999[1] in which the author was a participant. The meeting was the first to address issues of design for the emerging field of geographic data mining and knowledge discovery. Understanding space and geographic objects' various epistemological and ontological meanings for real users underscores users' ability to analyze spatial information. In this work we assume space is a creative developmental cognitive process and not merely an object for extrinsic discovery. Standards will be useful only as long as they are aligned with users' spatial information needs.

The need for embedding more user-centered meta-ontologies as well as ontologies implemented in tool functionality is appropriate to address the meaningfulness of geographic data mining

[*] *School of Information Studies and Department of Geography, Florida State University, Tallahassee, FL, USA, mgluck@lis.fsu.edu.*

[1] http://www.ncgia.ucsb.edu/varenius/

and knowledge discovery activities. That is, although algorithms and methods are important what is the best way to use them and under what scenarios are they most useful? In some sense methods such as clustering or association data mining techniques are answers in search of questions. Using these systems requires the ability to set thresholds and establish useful benchmarks for data mining/knowledge discovery. Computer scientists and statisticians have left the parameterisation or threshold settings to the user to declare but few users know a priori what significant thresholds to set. Thus, the computer scientists reframe the data mining/knowledge discovery problem from finding a method to establishing a threshold or set of weights. Such reframing unfortunately does little for the user of the methods with expertise in the data but not in the methods or for those who have expertise in the methods but little experience in the content domain. The epistemological and certainly the ontological basis for such shift of burden to users is understandable, but improved ontologically embedded tools may better assist users to find geographic resolutions for their spatial information needs.

Hence, the traditional view of users needing to adapt to systems through training or extensive hands-on experience is made problematic by viewing user goals as processes that can and should be embedded within system designs, especially in the design of user interfaces. Most geographic data mining systems are developed without strong methodologies to collect information on user processes during the completion of tasks aimed at achieving the user goal of geographic knowledge discovery. That is, traditional systems present users with categorical menus that are merely toolkits of system functionality and lack guidance for users' sequential use in geographic search or data exploration (Raper and Bundock, 1993). Current research in geographic data mining and exploration has focussed on algorithms for efficient pattern recognition without support for interpretation of the results nor the concern for the wider process requiring a range of geographic tools (Han, 1999). However, in the geographic domain the two, three and four dimensional nature of typical datasets makes search and discovery critically dependent on the cognitive fit between the user and the system.

Research in the fields of geographic information retrieval (Gluck and Fraser, 1997), geographic knowledge discovery with multimedia (Raper and Livingstone, 1995) and geographic data visualisation by Wood et al. (1999) has suggested that enriching the ontology of the process is the key to improving the efficiency of data mining and the effectiveness of knowledge discovery. We have found that without the means to declare/uncover the cognitive model of the user with its associated ontologies, data mining activities are unproductive at a user level. This means that the system should be capable of learning user characteristics, the interface should adapt to learning and the results must be contextualised for each user. In the domain of geographic data and knowledge the two, three and four dimensional nature of multimedia datasets adds considerable complexity to the process. Focussing on the hard problems posed by such multidimensional datasets is expected to yield general solutions not just disciplinary ones.

We have built a system whose initial interface and tools allow for user-sensitive and adaptive mechanisms. The tool enables users to perform exploratory data analysis using augmented seriation (Gluck and et al., 1999) for spatial data using visualization, sonification (sound), cartographic display, and multidimensional spreadsheets to explore and build hypotheses for spatio-temporal datasets. As such the research focuses on the deep ontological and epistemological aspects of discovering geographic knowledge. We believe that discovering geographic knowledge is a process and although it may vary among individuals there are general epistemological, ontological, and cognitive sets of structures that can and must be embedded in systems to assist users. Recent work on 'naive geography' (Egenhofer and Mark, 1995) implores researchers to utilise geographic ontologies for the representation of user concepts of space in cognitive models.

We hope in the future to modify the interfaces based upon spatial intelligent agent-monitored user profiles which are then augmented by user based experiences of use (Rodrigues and Raper, 1999). For example, menu items that are activated in various sequences will replace those initially derived through the task and information elicitation and the user profiles. We expect to support users mental models and provide such adaptive structuring with ontological supportive tools to guide tasks and reduce cognitive overhead through the development of dynamic interface redesign.

In summary, epistemological and ontological concepts of geographic objects are themselves embedded in human analytic processes. This work illustrates a user-based approach to understanding the spatial information needs of spatial analysts and concurrently adds to the developing body of knowledge on the ontological and epistemological aspects of geographic objects and their meanings for people performing real work.

Bibliography

Egenhofer, M. J., Mark, D. M., 1995. Naive geography. In: Frank, A. U., Kuhn, W. (Eds.), Spatial Information Theory. Vol. 988 of Lecture Notes in Computer Science. Springer, Berlin, pp. 1–15.

Gluck, M., et al., 1999. Augmented seriation: Usability of a visual and auditory tool for geographic pattern discovery with risk perception data. In: Proceedings of GeoComp '99. Fredricksburg, VA.

Gluck, M., Fraser, B., 1997. Usability of geospatial metadata: Metalevel issues for information retrieval. In: Proceedings of 1997 GIS/LIS Annual Conference. Cincinnati, OH.

Han, J., 1999. An overview of methods for geo-spatial data mining (from a database researcher's point of view). In: Proc. Varenius Specialist Meeting 'Discovering geographic knowledge in data-rich environments'. Seattle, USA.

Raper, J. F., Bundock, M. S., 1993. Development of a generic spatial language interface for GIS. In: Mather, P. M. (Ed.), Geographical Information Handling – Research and Applications. Wiley, Chichester, pp. 113–143.

Raper, J. F., Livingstone, D., 1995. The development of a spatial data explorer for an environmental hyperdocument. Environment and Planning B 22 (6), 679–687.

Rodrigues, A., Raper, J. F., 1999. Defining spatial agents. In: Camara, A., Raper, J. F. (Eds.), Spatial multimedia and virtual reality. Taylor & Francis, London, pp. 111–129.

Wood, J., Fisher, P., Dykes, J., Unwin, D., Stynes, K., 1999. The use of the landscape metaphor in understanding population data. Environment and Planning B 26, 281–295.

David Gross*

Approximate geometry and topology

Geographical information systems handle spatial entities by multiple representations. These representations include geometry and topology, that respectively describe the spatial location of objects and their spatial relationships.

Computing the geometry and topology of a query can be expensive, especially for high-dimensional data (e.g. spatio-temporal data). One way to overcome this problem is to consider their approximation.

We propose a model of approximation for spatial objects, with the notion of approximation schemes for spatial queries. As an example, we approximate a query on implicitly represented data, using random point membership tests.

The approximation quality increases with the number of points, converging to the exact expected result. If this quality is very high, the final user can not distinguish the approximate result and the exact one.

**LRI, Batiment 490, Universite Paris-Sud 91405 Orsay, France, fax: (33)01 69 15 65 86, dgross@lri.fr, http://www.lri.fr/ dgross.*

Stephen C. Hirtle[*]

The ontology of neighborhoods

Wayfinding is often directed by the use of neighborhoods and landmarks. For example, a restaurant in Chicago may be described as being near the Sear Towers (a landmark) in the Loop (a neighborhood), or an individual may be told drive along a road until you reach a small village, then turn left at clock tower. There is an interesting relationship between neighborhoods and landmarks, which in turn has strong implications for the usability of geographic information systems. Landmarks, which refer to points in space, are in many ways complementary to neighborhoods, which consist of small regions in a space. A neighborhood can be defined as the region surrounding one or more landmarks and a landmark can be defined as the most prominent building in a neighborhood. There are even geometric tools, such as Delaunay triangulation, which allows one to establish equivalencies between points and regions.

That said, there are several important distinctions between landmarks and neighborhoods. The fuzziness of a landmark tends to be tied to the conceptual notion of "landmarkedness." That is, there are some buildings or locations that are more likely to act as a landmark for most individuals, while are others are more personal landmarks or contain fewer of the characteristics typically associated with landmarks (Sorrows & Hirtle, 1999). In contrast, the fuzziness of a neighborhood is often tied to the extent or size of the region, so that the boundaries of a neighborhood are often vague or indeterminate. Furthermore, a

[*] *School of Information Sciences, University of Pittsburgh, Pittsburgh, PA 15260 USA, hirtle+@pitt.edu.*

landmark is often identified by name, whereas a neighborhood is defined by type. Thus, a landmark may be part of the ”what” system, as described by Landau and Jackendoff (1993), whereas a neighborhood may be part of the ”where” system.

Two interesting conclusions result from the analysis. First, given that neighborhoods are more often described by type, they are more likely to open to bias in the media. A village, town, city, and suburb, while referring to a basic municipal district, induce very different connotations and can be used strategically to bias readers. Second, while neither the use of landmarks nor regions is common in automated wayfinding systems, landmarks are easier to incorporate into current systems. Finally, it is argued that wayfinding systems will be lacking in usability, unless both landmarks and neighborhoods are systematically included as part of route descriptions.

*Alexander Klippel**

Representing qualitative spatial knowledge in schematic maps

This contribution aims at defining a theory of representing qualitative spatial knowledge in maps and, especially, in maps for wayfinding. The need for such a theory arises as maps, and especially schematic maps, are regarded as efficient — in the sense of fast and sound — means for reasoning on environmental knowledge. Yet, a theory that integrates research on qualitative spatial reasoning and maps is still missing, even though uncertain spatial knowledge has to be represented by graphical means. The foundation of this theory consists of two building blocks: First, the distinction of different kinds of spatial knowledge into three main categories, i.e. topological, ordering, and metrical knowledge, used in different disciplines (cf. Egenhofer et al., 1991; Schlieder, 1995; Vieu, 1997). Second, the notion of a representation theory first described by Palmer (1978) and extended for representation theoretic problems in AI by Freksa and his coworkers (Freksa et al., 1985).

The above mentioned distinction of three kinds of spatial knowledge enables the analysis of representing qualitative spatial information in schematic maps with respect to the geometric richness, possible inferences, and the degree of spatial abstraction that is applied to the map. It turns out, that within this triple ordering information plays the essential role (cf. Eschen-

* *University of Hamburg, Department for Informatics and Doctoral Program in Cognitive Science, Vogt-Koelln-Str. 30, 22527 Hamburg, Germany, klippel@informatik.uni-hamburg.de.*

bach et al., 1999, 1998). It sets stronger constraints for the characterization of spatial relations than topology but weaker constraints than metrics. For a more detailed analysis different kinds of ordering information have to be distinguished that are not necessarily maintained in a map all at the same time. Basic kinds of ordering relations exist between point-like and linear objects that are related through incidence relations, i.e. that points 'lie' on a line. Subway stations on subway lines are a good example. Here, ordering relations, for example, the precedence of one station to another within a single line is always preserved. The same holds for ordering information based on the notion of different sides with respect to outstanding landmarks and linear objects. The fact that an object is to the left of another object is maintained if the second object is important for spatial orientation or for spatial inferences in general. For example, if cities that are, originally, on the right side of the Rhine were depicted on the left side of the Rhine, we would get a great disorientation. In comparison to these cases, ordering information between points is not as essential. The circular order, i.e. a panorama in the sense of Schlieder (1995), is not preserved for every single point-like object, but, perhaps for outstanding ones, e.g., in the case of panorama maps.

It follows from this approach that basic entities have to be defined for which the described inferences are valid. Prior to definition these entities have to be isolated by different segmentation processes. Some of them are described within this work. The combination of basic entities and the commitment to valid inferences is summarized as conceptualized spatial structure that, following the approach of a toolkit by Tversky and Lee (1999), constitute the basic building blocks for the design of schematic maps. As an alternative to map generalization approaches it is suggested to start the design of schematic maps with these underdetermined spatial structures. Default assumptions, like the toolkit, for representing underdetermined spatial structures are necessary as the representational medium sets constraints, too. Whereas, in propositional representation formats discounting possibilities is one main characteristic graphic representations require the decision for exactly one alternative.

The characterization of different kinds of spatial knowledge in addition to ideas on segmentation and conceptualization builds the first basis for a representation theory that aims at bridging the gap between qualitative spatial reasoning and schematic maps. The second step is to theoretically substantiate the descriptively recorded aspects of the topics mentioned above by integrating them into a representation theoretic framework based on Palmer (1978).

Bibliography

Egenhofer, M. J., Herring, J. R., Smith, T., Park, K. K., 1991. A framework for the definition of topological relationships and an algebraic approach to spatial reasoning within this framework. NCGIA Technical Report 91-7, National Center for Geographic Information and Analysis.

Eschenbach, C., Habel, C., Kulik, L., 1999. Representing simple trajectories as oriented curves. In: Kumar, A. N., Russell, I. (Eds.), Flairs-99. Orlando, Florida.

Eschenbach, C., Habel, C., Kulik, L., Leßmöllmann, A., 1998. Shape nouns and shape concepts: A geometry for 'corner'. In: Freksa, C., Habel, C., Wender, K. F. (Eds.), Spatial Cognition. Lecture Notes in Artificial Intelligence. Springer, Berlin, pp. 177–201.

Freksa, C., Fuhrbach, U., Dirlich, G., 1985. Cognition and representation: An overview of knowledge representation issues in cognitive science. In: Laubsch, J. (Ed.), GWAI-84. Springer, Berlin.

Palmer, S. E., 1978. Fundamental aspects of cognitive representation. In: Rosch, E., Lloyd, B. B. (Eds.), Cognition and categorization. Lawrence Erlbaum, Hillsdale, NJ, pp. 259–303.

Schlieder, C., 1995. Reasoning about ordering. In: Frank, A. U., Kuhn, W. (Eds.), Spatial information theory. Lecture Notes in Computer Science. Springer, Berlin, pp. 341–349.

Tversky, B., Lee, P. U., 1999. Pictorial and verbal tools for conveying routes. In: Freksa, C., Mark, D. M. (Eds.), Spatial information theory. Lecture Notes in Computer Science. Springer, Berlin, pp. 51–64.

Vieu, L., 1997. Spatial representation and reasoning in ai. In: Stock, O. (Ed.), Spatial and temporal reasoning. Kluwer, Dordrecht, pp. 3–41.

Margarita Kokla and Marinos Kavouras[*]

Concept lattices as a formal method for the integration of geospatial ontologies

Effective communication and smooth interaction between different sources of geodata requires a method for sharing and integrating different ontologies. Lattices of Spatial Concepts constitute a new tool for information organization and semantic integration, in order to provide reuse of data between heterogeneous geographic information systems.

The methodology is founded on Formal Concept Analysis, a theory of concept formation and conceptual classification. The integration of multiple geospatial classifications, which exhibit differences in spatial and thematic resolution, allows the creation of an ontology for the geospatial domain. Spatial Concept Lattices can be used as a formal method to compare geospatial classifications created for different levels of detail and from different contexts. Lattices, in contrast to trees, support multiple inheritance and thus, are powerful structures for the representation of multidimensional, overlapping geographic categories.

As far as original classifications are fully described, the method identifies similarities and differences, and reveals interrelations between them. Therefore, it leads to the creation of a single, integrated, unambiguous schema from different thematic classi-

[*] *National Technical University of Athens, 9, H. Polytechniou Str., 157 80 Zografos Campus, Athens - Greece, Tel: 30+1+772-2637/2731, Fax: 30+1+772-2734, mkokla, mkav@survey.ntua.gr.*

fications, which includes all concepts and relationships identified by the original classifications. The resulting geospatial ontology with its generalisation hierarchy represents information at different resolution levels. The above concepts are illustrated using a clear, real-case example.

Werner Kuhn[*]

How to produce ontologies: an approach grounded in texts

Abstract. A method is proposed to derive formal domain ontologies from natural language texts describing the domain. Apart from its textual grounding, the key characteristics of the proposal are an emphasis on the behaviors afforded by objects, and on a layering of these behaviors. Using the example of the German traffic code, the presentation shows the steps to derive formal specifications for the domain concepts occurring in a text. The need and means to make these specifications executable are addressed, opening the perspective of executable ontologies in the sense of layered, agent-based models of domain behaviors.

Motivation

Ontologies have become the subject of engineering processes. They are specifications of concepts that occur in a domain. Like any other specification, they require methods to produce and validate them. I will present a first cut at such a method by way of an extensive example. I will also argue for a pragmatic view of the role of formalisms in ontological design.

My method differs from other approaches by attempting to close what might be called the "formality gap", i.e. the diffi-

[*] *Institute for Geoinformatics, University of Muenster, kuhn@ifgi.uni-muenster.de.*

culty to go from informal ideas about a domain in the heads of knowledge engineers to formal expressions in logic or other formalisms. I claim that this gap is generally too wide and that important application semantics tends to get lost in it.

There are two main reasons for this problem: First, knowledge engineers have to use their own understanding of a domain as a basis for formalization, gathered from interviews, brainstormings, requirements documents, existing databases, etc. The communication of application semantics to them is not supported by appropriate languages or techniques. Second, the product of the knowledge engineer's work is often too abstract and difficult for domain experts to verify. Whether it contains expressions in first order logic or diagrams in something like the Unified Modeling Language (UML), it is unlikely that enough misunderstandings or omissions will be uncovered through inspection by typical domain representatives. Teaching these people logic or UML, on the other hand, is neither practical nor useful, as these languages are per se not very good for ascertaining consistency and completeness.

Thus, more suitable languages for the communication between domain experts and knowledge engineers are needed. Practical experiences in several ontological design projects have taught me that

- domain experts should already be familiar with such a language,
- some domain knowledge should already be expressed in the language,
- translating from the language to a formal ontology should be supported by tools.

These are tough requirements to fulfill. Indeed, if they are all to be met, the choice reduces to just one candidate. It is the language with which all domain experts are familiar to the greatest extent; it is the language in which most domain knowledge is already expressed; and it is the language for which the broadest range of processing tools exists: natural language.

Consequently, I will present a method that takes natural language domain descriptions and produces formal ontologies from them. It is simply called *Ontologies from Texts*[1].

Method

The presentation will demonstrate the first results of developing and experimenting with the method, using an example from the car navigation domain. While one might not expect this domain to be easily amenable to a method starting from natural language descriptions of activities, there are actually several such descriptions available or easily obtainable:

- traffic codes describe all objects and activities relevant to driving,
- driving instructions contain information for successful navigation,
- travel narratives provide an account of observations and decisions during navigation.

I have chosen the German traffic code[2] as the text to apply the method to and develop it. One reason was that such a code defines by definition all activities and constraints that are relevant to driving behavior (at least from the legal point of view). Note, however, that such codified descriptions are frequently available for domains where spatial information is being used, due to the legal or administrative regulations coming with many spatial activities. In fact, any agency using spatial information is bound to have detailed descriptions about the operations implementing its mandate.

[1] see also the abstract with this title at http://www.giscience.org/GIScience2000/-program.html

[2] http://www.fen.baynet.de/ na1723/law/stvo.html

The presentation will show how the contents of the traffic code can be translated step-by-step into a formal ontology of driving legally on roads. Without loss of generality for the method, the code has first been reduced to the rules concerning cars alone (eliminating pedestrians and bikers, for example). Further simplifications included the elimination of some special behaviors and traffic conditions.

Roughly, the method involves the following procedure:

1. Extract behaviors and affordances from the text
    (a) Verbs and verbal expressions, e.g., "drive" or "keep distance"
    (b) Gerunds and other nouns, e.g. "driving" or "speed reduction"
    (c) Affordances and behavior restrictions, e.g., "right of way", "speed limit"
2. Define classes of behaviors by
    (a) merging different grammatical forms, e.g. "reduce speed" and "speed reduction"
    (b) merging synonyms, e.g. "reduce speed" and "break"
3. Identify object classes participating in these behaviors
    (a) Subjects of clauses, e.g., "the car turns"
    (b) Direct objects, e.g., "change the lane"
    (c) Indirect objects, e.g., "yield to driver"
4. Relate these object classes to behaviors
    (a) As arguments of operations, e.g., drive on a lane
    (b) As inheritors of (multiple) behaviors, e.g., a car is a physical body that moves
5. Assign attributes
    (a) to object classes, e.g., road width

   (b) to behaviors, e.g., driving speed

6. Identify and rank layers of behavior
   (a) Identify the prototypical behavior(s) involved (e.g., driving)
   (b) Build a layered model of increasingly complex behavior (e.g., standing still, driving at constant speed, changing speed, turning, passing etc.).

Contributions

The proposed method is still under development, and meant to complement rather than replace existing approaches. In order to determine where such complementing could occur, it appears worthwhile to consider some key characteristics by which the method differs from those proposed in the literature so far. These are

- its grounding in texts
- its behavior-oriented approach
- the layering of the resulting ontologies
- the executability of the ontological specifications.

Grounding an ontology in some tangible document or artifact outside the imagination of a knowledge engineer (or even outside that of a domain specialist or of any kind of committee), seems like a good idea. A recent survey of ontology design methods[3] has shown that existing methods have at least one thing in common: they use brainstorming or undisclosed techniques to arrive at rather improvised collections of concepts in a domain. This issue of grounding ontologies is where the *Ontologies from Texts* approach may offer its most significant contribution.

[3] http://sunsite.informatik.rwth-aachen.de/Publications/CEUR-WS/Vol-18/

The approach is behavior-oriented in the sense that relevant behaviors (as opposed to objects) are identified first. Object classes (or concepts in the narrow sense often used in the literature) are then identified as participating elements in these behaviors. For example, roads and cars are objects required for driving, and intersections occur in turning. Attributes are used (in a more restricted role than usual) to further qualify both, behaviors and objects. For example, a lane object and a driving behavior both have a direction attribute (which must have matching values for a particular instantiation of driving).

The ontologies resulting from the method's application are layered in the sense that increasingly complex behaviors get defined on top of levels with simpler behaviors. For example, turning is defined in a layer above driving, which is itself defined above standing (and consists of the simpler driving at constant speed and the more complex driving at variable speeds). This specification of ontological concepts at multiple layers of increasing complexity satisfies a key requirement of ontological design. It also establishes a link to artificial intelligence approaches for modeling complexity. More particularly, it is postulated here that the increasing complexity of the environment results from (responds to) behaviors, rather than the other way round.

In the case of the traffic code, an interesting and useful constraint and guideline for deriving layers of behavior was to keep the ontology incrementally consistent: The specified behavior at each level has to obey all the rules that are applicable up to this level. For example, if cars can only stand still, they are not permitted on lanes. Note, by the way, that this still allows for accidents. For example, if cars can only drive at constant speed, a fast car can crash into the one with the bumper sticker "I'm slow, but ahead of you". Thus, driving is made safe(r) by adding more complex behaviors and corresponding rules.

Finally, the resulting formal ontology will consist of an executable specification (in a functional language) of behaviors and object classes. The key advantage of such a rich and executable model is that an ontology can demonstrate itself to its designer as well as to a domain expert. Most inconsistencies and omissions are either avoided by construction during the design of the ontology or lead to observable errors during execution Frank and Kuhn (1999).

Combining layering with executability will eventually lead beyond a typical formal ontology (in the sense of a collection of theorems and axioms). The goal of the implementation part in the case study is actually to produce an agent model of legal driving behavior in Germany. However, the current presentation will not go into details on this aspect.

Conclusions

Apart from the claimed benefits of the method, at least three important questions need to be addressed:

1. In what way is the legal code used a special case of a natural language text?
2. What happens if no natural language texts are available for a domain?
3. What about the notorious ambiguity of natural language descriptions?

Concerning the special nature of a traffic code as a natural language text, there are clearly some properties in a legal code that other texts (such as narratives) lack. Among them are completeness, consistency, and minimized ambiguity. We have yet to collect experience on the suitability of other, less-structured texts. However, it seems that the possibility to obtain even an incomplete ontology from unstructured texts would represent a significant step forward in a domain modeling process. Also, domain descriptions with regulatory purpose occur quite frequently, though not always in the polished form of legal codes.

Questions two and three appear to have encouraging answers as well. Again, for most application domains, some kinds of work regulations or workflow descriptions are available. In cases where no texts can be found, the best procedure is to have such texts written first by domain experts, explaining their activities in simple terms, but exhaustively and in detail. Often, this phase is part of a requirements analysis anyway and leads to useful clarifications in the process of an information system design.

The ambiguity issue is non-trivial, but pertains to the general problem of moving from informal, imprecise statements to formal models. Ambiguous descriptions for which domain experts feel competent could be preferable to (possibly!) unambiguous models, which coerce them into accepting something that they might not fully understand or agree with. If there is ambiguity in the texts (and there usually is), it is likely to be revealed in executing the specification. If there is ambiguity in non-executable ontologies, it will be revealed much later (and at high costs) in system implementations. On the other hand, if today's formal ontologies are unambiguous, they are likely to have lost crucial application semantics (for the reasons given above) and may thus lead to systems that are of limited use.

In conclusion, the presentation will emphasize the need for a practical procedure to derive executable domain ontologies from natural language descriptions. As this procedure is work in progress, the issues covered in this abstract as well as others raised by the participants will be discussed at the conference, hopefully leading to an improved method and a better understanding of its potentials and limitations.

Acknowledgments

In 1984, Andrew Frank mentioned a paper proposing some kind of conceptual analysis of natural language requirements documents. Since then, we have both been looking in vain for that source which seems to contain the seed for this work.

In 1997, the German Agency for Waterways (BAW, Bundesamt für Wasserstrassen) has funded a feasibility study for the design of a feature-attribute catalogue. Many ideas presented here took shape during that study.

In 1999, MobileGIS Ltd. began funding a project to derive methods for the design of mobile services and their ontologies. Steve Smyth suggested to start that work with an analysis of travel narratives, which provided a second case study to develop the method.

Since early 2000, the members of the Special Interest Group on Object-Orientation (OOSIG) and those of the Special Interest Group on Scenarios (ScenarioSIG) at our institute have been invaluable discussion partners and critics in the development of the method.

Bibliography

Frank, A. U., Kuhn, W., 1999. A specification language for interoperable GIS. In: Goodchild, M. F., Egenhofer, M., Fegeas, R., Kottman, C. (Eds.), Interoperating Geographic Information Systems. Kluwer, Norwell, MA, pp. 123–132.

Lars Kulik*

Vague spatial reasoning based on supervaluation

Abstract. This paper presents an account, originally developed in linguistics and philosophy, to deal with indeterminate geographical objects, the theory of supervaluation. In current geographical information systems fuzzy set theory and fuzzy logic are typically employed to capture vague spatial information. An example shows that there are spatial constellations where the inferences obtained by supervaluation are more adequate than those obtained by fuzzy logic. In contrast to fuzzy theories, the theory of supervaluation does not rely on numbers to model vagueness. Therefore, it is able to support spatial databases in qualitative spatial inferences.

1. Introduction

There is a controversial debate in the AI community (cf. Elkan, 1994) whether fuzzy logic should be employed to model reasoning about vagueness. Despite this fact, the GIS community (cf. for instance Burrough, 1996) considers fuzzy methods as

---

*Department for Informatics, University of Hamburg, Vogt-Klln-Str. 30, D-22527 Hamburg, Germany, kulik@informatik.uni-hamburg.de. — The research reported in this paper has been supported by the Deutsche Forschungsgemeinschaft (DFG) in the project 'Axiomatics of Spatial Concepts' (Ha 1237-7). I am in particular indebted to Carola Eschenbach, Christopher Habel, and Heike Tappe for their valuable comments.

the primary tool to deal with vagueness. There are only few alternative approaches that are not based on fuzzy set theory to reason about entities with vague boundaries. These approaches, for instance the accounts of Cohn and Gotts (1996) or Clementini and Di Felice (1996), mainly come from the spatial reasoning community. They are primarily concerned with topological relations and are extensions of theories modeling sharply bounded entities. Cohn and Gotts modify the RCC theory (Randell et al., 1992) whereas Clementini and di Felice take up the 9-intersection model of Egenhofer and Herring (1991). None of these qualitative theories has been developed to cope with specific challenges of spatial vagueness like gradual changes given by a transition of a forest to a meadow. To capture such smooth transitions I propose a theory of spatial vagueness that is based on the theory of supervaluation. Since the theory of supervaluation does not rely on numbers, it seamlessly fits into qualitative approaches.

2. Spatial Supervaluation

Vagueness is considered in this paper — in correspondence with most theories — as semantic vagueness. Therefore, it is not necessary to assume that the geographic entities themselves are vague but our language or concepts about them. There are different theories about vagueness. The epistemic view assumes that vagueness has to be regarded as ignorance about exact spatial boundaries of a geographic entity like a forest whereas the degree theory presumes that a predicate like *forest* has a certain degree of applicability to a geographic object. According to the theory of supervaluation (cf. Fine, 1975; Kamp, 1975) vagueness results from a semantic indecision: A vague predicate distinguishes entities to which it definitely applies and entities to which it does not apply. Hence, a predicate like *forest* singles out spatial regions or locations that are undeniably part of the forest from regions which are unquestionably not part of the forest. There might be still some remaining regions that cannot be clearly assigned to one of these two groups. Theories of supervaluation model this fact by assuming that there is not one single interpretation of a vague predicate but several equally good ones. Some interpretations consider these regions as part of the forest and other ones do not take them as part of the forest. All the regions which definitely belong to the forest constitute the *positive extension* of the forest, all regions that do not belong to the forest are the *negative extension*, and the remaining ones represent the *penumbra*.

Every single interpretation which assigns a meaning to a predicate like *forest* is called *admissible* if it makes the predicate true in the positive extension, false in the negative extension, and either true or false in the penumbra. Hence, in a single interpretation every region of the penumbra either count as belonging to the forest or as being not part of the forest (see Figure 19.1b). It follows that every admissible interpretation subdivides the underlying space into two regions: one region that represents the spatial extension of the forest and another region that does not belong to the forest. Every admissible interpretation is precise and accordingly called a *precisification*.

Each statement like ‘this region belongs to the forest’ is either true or false on a given interpretation. The corresponding assignment of a truth value to the statement is called a *valuation*. There is no reason to prefer any precisification to another one. Thus, all precisifications are considered. The assignment of truth values for all interpretations is called a *supervaluation*. A statement is *supertrue* (*superfalse*) if it is true (false) for all admissible interpretations. It is a remarkable feature that the technique of supervaluation maintains the law of excluded middle and the law of non-contradiction: Given a statement ‘p’ the formula $p \vee \neg p$ is supertrue whereas the formula $p \wedge \neg p$ is superfalse even if p is based on vague predicates. If there are interpretations for which the statement is true, and other ones for which the statement is false, then in the classical theory of supervaluation no truth value is assigned. Kamp (1975) shows that the idea of degrees of truth can be captured within the framework of supervaluation. The main idea is to measure the set of admissible interpretations for a predicate. However, to define such a measure in the general case of an arbitrary predicate is very difficult if not impossible. But in the case of spatial regions it is possible to associate such a measure by using spatial knowledge like ordering information. The details are omitted because I want to emphasize the possibility to reason about spatial vagueness without numerical concepts.

3. Reasoning about Spatial Vagueness

This section shows that the theory of supervaluation is able to draw inferences without a numerical assignment of truth values. Hence, it is not always necessary to assume a degree theory of truth in order to reason about vagueness. In a simple scenario we investigate whether a certain region is a possible habitat of an animal A. We know that the animal A only settles in an area if it finds at least one of two different plants Pl_1 and Pl_2. The statement 'the plant Pl_i is found at location P' is abbreviated by $p(Pl_i, P)$, and the fact that a position P is in a region R is symbolized as $P \iota R$. If P is a point of the forest (meadow) this is written as $\mathsf{f}(P)$ ($\mathsf{m}(P)$). A forest region is uniquely determined by its points: $\mathsf{F}(R) \Leftrightarrow_{\text{def}} \forall P[P \iota R \Rightarrow \mathsf{f}(P)]$. This holds in the same way for a meadow region, abbreviated by $\mathsf{M}(R)$.

We assume two rules: first, the plant Pl_1 is found everywhere in the forest region $\mathsf{F}(R)$, second, a meadow region denoted by $\mathsf{M}(R)$ is covered everywhere with the plant Pl_2 (see Figure 19.1a, 19.1c). These two rules can be summarized as $\mathsf{f}(P) \Rightarrow p(Pl_1, P)$ and $\mathsf{m}(P) \Rightarrow p(Pl_2, P)$. Moreover, for the sake of simplicity we assume that every point that does not belong to the forest belongs to the meadow: $\neg\mathsf{f}(P) \Rightarrow \mathsf{m}(P)$. Hence, there is a uniquely determined region $\mathcal{R}$ that contains the points of the forest and the points of the meadow: $\forall P[P \iota \mathcal{R} \Leftrightarrow \mathsf{f}(P) \vee \mathsf{m}(P)]$. Since neither the forest nor the meadow have sharply bounded regions, the regions associated with their positive extensions are called Core(F) and Core(M), respectively. The question is whether the animal A considers the boundary region of the forest and the meadow $\Delta(\mathsf{F},\mathsf{M}) := \mathcal{R} \setminus (\mathrm{Core}(\mathsf{F}) \cup \mathrm{Core}(\mathsf{M}))$ as a possible habitat. Therefore, we have to determine the truth value of the disjunction $p(Pl_1, P) \vee p(Pl_2, P)$ if $P \,\iota\, \Delta(\mathsf{F},\mathsf{M})$. According to the construction of this scenario we expect that the animal clearly considers the region $\Delta(\mathsf{F},\mathsf{M})$ as a possible habitat (see Figure 19.1c). It turns out that the theory of supervaluation corresponds to this intuition whereas fuzzy logic provides an affirmative answer only to a certain degree.

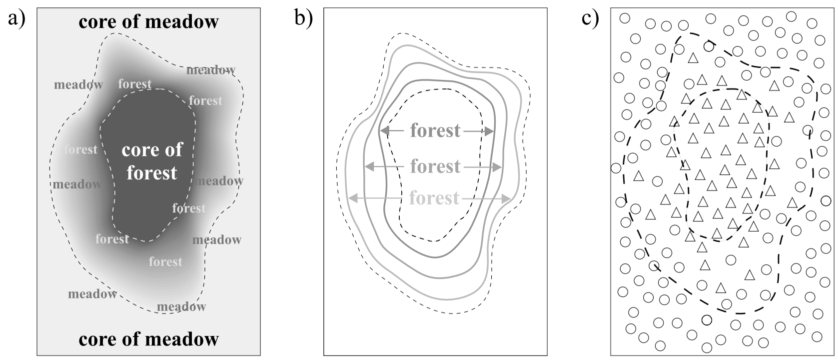

Figure 19.1: The left figure shows a forest that is surrounded by a meadow. Since the forest has a vague spatial boundary the transition of the forest to the meadow is a region. The core of the forest is surrounded by a white dashed line and the core of the meadow is the area between the black dashed line and the rectangle. The middle figure depicts three different precisifications of the vaguely bounded region of the forest. The darker the grey line, the more interpretations count the enclosed region as part of the forest. The right figure shows a distribution of two plants in the forest and in the meadow. The plants are symbolized as circles and triangles.

If we employ the theory of supervaluation, we obtain that for every precisification of the boundary $\Delta(\mathsf{F},\mathsf{M})$ the location P in question either belongs to the forest or to the meadow (see Figure 19.1b). If it belongs to the forest ($\mathsf{f}(P)$), we know that $p(Pl_1, P)$ holds, otherwise it follows from $\mathsf{m}(P)$ that $p(Pl_2, P)$ holds. Therefore, in every precisification at least one of the plants Pl_1 or Pl_2 can be found at P. This means the statement $p(Pl_1, P) \vee p(Pl_2, P)$ is true. Since this holds for every precisification, the disjunction is supertrue. As a result it turns out that the boundary of a forest is definitely a possible habitat for the animal A.

To formulate the same problem in terms of fuzzy logic, I briefly compile the required prerequisites typically assumed in fuzzy logic. Given two statements p and q the truth value (denoted by $|\cdot|$) of the disjunction is $|p \vee q| = \max\{|p|, |q|\}$, and of the negation is $|\neg p| = 1 - |p|$. To be able to reason with fuzzy logic, we use a system that ensures soundness and completeness (cf. Novak, 1989). It is based on the proof-theoretic notion that a formula p is provable to (at least) a degree of α: $\vdash_\alpha p$. Correspondingly, the modus ponens has the following form: if $\vdash_\alpha p$ and $\vdash_\beta (p \Rightarrow q)$ then $\vdash_{\max\{0,\alpha+\beta-1\}} q$. We use the soundness and completeness result stating that p is provable to a degree of α if and only if p is true in every model to a degree of α.

For the forest example this leads to the following interpretation. A location P of the boundary $\Delta(\mathsf{F}, \mathsf{M})$ belongs to the forest to a degree of $\alpha_P \in]0, 1[$ (see Figure 19.1a) that is to say $|\mathsf{f}(P)| = \alpha_P$ and therefore $|\neg(\mathsf{f}(P))| = 1 - \alpha_P$. Since $\neg(\mathsf{f}(P)) \Rightarrow \mathsf{m}(P)$ is definitely true ($|\neg(\mathsf{f}(P)) \Rightarrow \mathsf{m}(P)| = 1$) using $\vdash_{1-\alpha_P} \neg(\mathsf{f}(P))$ we obtain $\vdash_{1-\alpha_P} \mathsf{m}(P)$, and consequently $|\mathsf{m}(P)| = 1 - \alpha_P$. Since the rules $\mathsf{f}(P) \Rightarrow p(Pl_1, P)$ and $\mathsf{m}(P) \Rightarrow p(Pl_2, P)$ are definitely true, we obtain $|p(Pl_1, P)| = \alpha_P$ and $|p(Pl_2, P)| = 1 - \alpha_P$. Hence, it follows for the truth value of the disjunction $|p(Pl_1, P) \vee p(Pl_2, P)| = \max\{\alpha_P, 1 - \alpha_P\} \in [0.5, 1[$. That means depending on the location P it is true only to a degree between 0.5 and 1 to find at least one of the two plants Pl_1 and Pl_2. Therefore, it is possible to a degree of at least 0.5 that the animal A settles in $\Delta(\mathsf{F}, \mathsf{M})$, whereas the theory of supervaluation states that it is definitely possible that the animal A settles in the boundary.

4. Conclusions

The above example shows that there are spatial constellations where the inferences obtained by supervaluation are more adequate then the ones obtained by fuzzy logic. The theory of spatial supervaluation has at least two further advantages. First, it enables us to employ the methods of classical logic to reason about vagueness and second, it characterizes vague boundaries of spatial entities as sets of sharp boundaries. Therefore, it is not necessary to design a spatial database from scratch to integrate reasoning about vaguely bounded entities as it is the case for fuzzy set theory. However, the computational effort to do calculations based on supervaluation still has to be evaluated.

Bibliography

Burrough, P. A., 1996. Natural objects with indeterminate boundaries. In: Burrough, P. A., Frank, A. U. (Eds.), Geographic Objects with Indeterminate Boundaries. Vol. 2 of GISDATA. Taylor & Francis, pp. 3–28.

Clementini, E., Di Felice, P., 1996. An algebraic model for spatial objects with indeterminate boundaries. In: Burrough, P. A., Frank, A. U. (Eds.), Geographic Objects with Indeterminate Boundaries. Vol. 2 of ESF GISDATA. Taylor & Francis, pp. 155–169.

Cohn, A. G., Gotts, N. M., 1996. The 'egg-yolk' representation of regions with indeterminite boundaries. In: Burrough, P. A., Frank, A. U. (Eds.), Geographic Objects with Indeterminate Boundaries. Vol. 2. Taylor & Francis, London, pp. 171–187.

Egenhofer, M. J., Herring, J. R., 1991. Categorizing binary topological relationships between regions, lines, and points in geographic databases. Technical report, Department of Surveying Engineering, University of Maine, Orono, ME.

Elkan, C., 1994. The paradoxical success of fuzzy logic. IEEE Expert 9 (4), 3–8.

Fine, K., 1975. Vagueness, truth and logic. Synthese 30, 265–300.

Kamp, J. A. W., 1975. Two theories about adjectives. In: Keenan, E. L. (Ed.), Formal Semantics of Natural Language. Cambridge University Press, Cambridge, pp. 123–155.

Randell, D. A., Cui, Z., Cohn, A., 1992. A spatial logic based on regions and connection. In: Brachmann, R., Levesque, H., Reiter, R. (Eds.), Third International Conference on the Principles of Knowledge Representation and Reasoning. Morgan-Kaufmann, Los Altos, CA, pp. 165–176.

Tõnu Oja, Jüri Roosaare, Raivo Aunap, Jüri Jagomägi*

Spatial data models in Estonia

The Estonian Land Board initiated the project "Spatial Data Models of Estonia". Creation of the reality model, determining the feature classes to be mapped, the data model determining the modeling of mapped features in the computer, and the presentation model determining the way a map looks like on screen or on paper has been put on the Institute of Geography, University of Tartu. The goal put before the project is both to find universal agreement between the different already existing spatial databases (to minimize the needs for changes) and, to reach a reasonable set of models suitable for use at any scale. Also, the national standards should be in agreement with international standards and, forseeing joining the EU, definitely in agreement with the standards in EU. The situation in handling of spatial data in Estonia is very diverse and far from standardized Mõisja (2000). Several areas of problems in creation of the models formulated during the project can be brought up. Firstly, the boundaries of the model use - topographic maps, thematic maps, planning base, state prepared databases, national registers, commercial maps, navigation maps etc cannot have one set of models as it would be too complicated. The possible solution would be to include first of all the databases needed for state purposes on terrestrial areas (topographic maps, land cadaster and maps for planning purposes). The easiest though not the wisest way is to set a limited list of maps to what the models apply. Secondly, there is a need for a generalization rule

[*] *Institute of Geography, University of Tartu, Estonia. Jagomägi: Regio Ltd..*

for use of the reality model on different scales, and resultantly, also for data and presentation model. As a solution a seven step approach to the reality has been offered where the class is characterized by the 'number of zeros' in the scale (i.e. class 4 includes five subclasses (4A - 1:7000 - 1:14000, 4B - 1:14000 - 1:28000 etc). Thirdly, there are different approaches to classification of the reality. The width of the feature class and corresponding number of the feature classes used in the model needs to solved, also grouping of the feature classes. A possible solution is to have approximately 1000 feature classes with a hierarchic grouping based on 'discovering' the world that brings to classes like point objects, communicative lines, distinguishing lines, land cover areal objects, anthropogenic areal objects, natural 3D objects, anthropogenic 3D objects and specific objects. Alternative, more common classification follows the sectorial approach - water, roads, buildings etc. Advantages and disadvantages of different approaches are under public discussion. The data model depends on the reality model and its main complexity is associated with the ability for generalization. The basic concept for the data model would be starting from the basic geometric primitives for 0D (points), 1D (lines), 2D (areas), 3D (volumes), and 4D (spatial processes). This leads us with from one to five basic models for different feature classes depending on the scale: point objects need only point data model, areas may need 2D, 1D (e.g., rivers in small scale), or 0D (lakelet in small scale) data model etc. Theoretically, the presentation model should guarantee similar outlook of a map on the same type of media independent on the software and hardware used. However, realization of the presentation depends on process line and is dependent on the media (soft- or hardcopy), hardware and software used. Also, presentation depends much on the purpose of a certain map layer (e.g., background data versus main information). Data model should include attributes enabling for specified presentation. The first stage of the project is presented in internet for public discussion (http://www.geo.ut.ee/ruum/).

Bibliography

Mõisja, K., 2000. Handling of spatial data in Estonia in 1990-2000, unpublished Manuscript.

Martin Raubal*

Ontology and epistemology for agent-based wayfinding simulation

Wayfinding and orientation are important parts of people's daily lives. We have to find our ways through cities, through buildings, or along streets and highways. Many times people find it difficult to perform such tasks because they are not provided with adequate "knowledge in the world". Environments lack sufficient wayfinding information or their architectures are badly designed, therefore they are too complex to facilitate wayfinding. Agent-based simulation of human wayfinding before actual construction of a built environment helps to determine where people face wayfinding difficulties, why they face them, and how wayfinding information and design have to be changed to avoid such difficulties.

Ontology and epistemology of space are basic concerns during the development of an agent-based wayfinding model. By defining the ontology for a specific wayfinding domain or environment, we describe what is in this domain in a general way. Paying attention to epistemology allows us to focus on the wayfinding agent's knowledge and beliefs. In this work we look at ontology and epistemology from the viewpoint of ecological science, a multidisciplinary advance to the study of living systems, their environments, and the reciprocity between the

* *Department of Geoinformation, TU Vienna, Gusshausstr. 27-29, 1040 Vienna, Austria, raubal@geoinfo.tuwien.ac.at.*

two. In particular, we consider the sub field of ecological psychology, which proposes to study the information transactions between living systems and their environments, especially in regard to the perceived significance of environmental situations for the planning and execution of purposeful behaviors.

Our main focus lies on wayfinding in buildings and we use wayfinding in airports as a case study. The ontology of this wayfinding environment is based on the ideas of J. J. Gibson, a proponent of ecological psychology, who investigated how people visually perceive their environment. Accordingly, we subdivide the wayfinding environment into a medium, substances, and surfaces. We move in a medium (of light, sound, odor, etc.) in which there are points of observation and lines of locomotion. The substances differ in chemical and physical composition, and are structured in a hierarchy of nested units. The medium is separated from the substances of the environment by surfaces. In our case, substances are cognizing agents, such as a passenger or an employee of the airport, and non-cognizing objects, such as a door, a sign, or a check-in counter.

The epistemological question of what the wayfinding agent can know about the environment and how it can accumulate such knowledge is modeled through affordances. Gibson described the process of perception as the extraction of invariants from the stimulus flux. Surfaces absorb or reflect light and Gibson's radical hypothesis was that the composition and layout of surfaces constitute what they afford. Affordances are therefore specific combinations of the properties of substances and surfaces taken with reference to an observer. There are many affordances in the environment but the wayfinding agent perceives only the affordances relevant for the specific wayfinding task, such as a door affords opening and moving through, or a hallway affords moving along. In an airport where all the necessary wayfinding information is on yellow signs, the wayfinding agent will only utilize the affordance of perceiving a yellow sign and will ignore signs in other colors.

Agents have to be coupled with the environment in which they perceive and act. The nature of this connection is the following: the environment provides percepts (i.e., affordances) to the agent, the agent decides upon and performs actions in (and therefore on) the environment, which in turn provides new percepts (i.e., affordances), etc. The complexity of this process is influenced by the properties of the environment.

Research in spatial ontology and epistemology is an important basis for the setup of an agent-based model for wayfinding. Perception and cognition of the agent can only be modeled in a useful way if the ontological and epistemological foundations are well established. In this work we try to connect ontology of space, epistemology of space, and spatial cognition, in order to come up with a practical agent-based simulation tool for wayfinding in buildings. Such a tool will help to design buildings that facilitate wayfinding.

Christoph Rüther*

Specification of terms to share heterogeneous information

In order to provide services from heterogeneous data sources, it is necessary to build systems that know about the intended meanings of the stored information. To enable information sharing between different GI communities today it is necessary to study the whole documentation of all the databases involved. This is the only way to understand which information is stored in the databases and what the used terms mean. The intended meaning of terms used in the databases cannot be found in the databases themselves. Additionally, the terms stand in no defined relation to the terminology of other data models. Thus, a name of a road may be an official name or a road number, which can make the information of two databases incomparable.

The gap is that documentations of data models, feature- and attribute catalogues are not implemented, so it is not possible to use them for ad hoc services that meet the users' interests. It is therefore necessary to give the data more information about their meaning. As an example, we can think about two databases storing road information. Both use the term *width of road*. It is not clear if the data represent just the width of the traffic lanes or include the width of sidewalk, parking lanes and bicycle paths.

The first step to enable access to this information is to specify

* *Institute for Geoinformatics, University of Münster, Robert-Koch-Str. 26-28, D-48149 Münster, Germany, Fax + 49 (0) 251 83- 3 97 63, Tel. +49 (0) 251 83-3 300 13, ruether@ifgi.uni-muenster.de.*

the intended meaning of the used terms in a computationally accessible way. The second step is to compare these specifications with specifications of other data models and derive a common ontology. By specifying a common ontology, higher-level ontologies can be identified (e.g. the concept of named objects).

As example we may take a look at the german data model ATKIS and the international GDF standard for road data. Both models use three different kinds of road names. In ATKIS they are called *Geographischer Name* (geographic name), a *Zweitname* (second name) and a *Kurzbezeichnung* (short name), and in GDF they are called *Official Name*, *Alternate Name* and *Route Number*. Additionally these terms have a corresponding semantic. We may say that names in ATKIS and GDF provide a common conceptualisation:

```
data PrimaryName        = PrimaryName Name
type GeographischerName = PrimaryName
type OfficialName       = PrimaryName
```

Both models provide the concept of named features. Names are handled as strings. A common ontology of names may be expressed as:

```
data Name = Name String
class NamedObjects n where name :: n -> String
```

Finally, we can say that names of the same type are comparable.

```
instance Eq PrimaryName where
  (PrimaryName(Name n1)) ==
     (PrimaryName(Name n1)) = n1 == n2
```

Algebraic specifications, written in the functional language Haskell, meet the requirement of being computationally accessible. They allow for unambiguous interpretation, are testable and the ontologies may be written independently from the implementation but can also be related to them. In the presentation we will show how to use Haskell to specify the intended meaning of the stored features and attributes.

Barry Smith*

t.b.d.

This talk offers a new approach to the entire range of interlinked ways in which we relate, cognitively, to objects in the world and it shows how this new approach can be applied not only to perception and judgment but also to our veridical uses of theories, classification-schemes, databases and maps.

The theory is inspired by the supervaluationistic account of vagueness, which asserts that when we use a term like *Mont Blanc* then there are many parcels of reality to which our term refers — parcels of reality which differ just a little from each other. Our use of such a term allows us to project successfully upon reality, even in spite of the fact that we do not project uniquely. It sets a certain portion of reality into relief in a quite specific way, and it traces over those other portions of reality which fall outside our purview. Our perception, too, projects in similar fashion upon a corresponding portion of reality, and so do many of our judgments and theories. We can think of the use of a singular term as imposing a single-celled partition upon reality: it focuses upon and sets into relief a certain unified portion of reality in the manner of a telescope. Judgments, theories, classification-schemes, databases and maps are associated with many-celled partitions: they focus upon and set into relief segments of reality which may be widely scattered through time and space.

The paper presents a general theory of single- and many-

* *Department of Philosophy, 135 Park Hall, University at Buffalo, NY 14260-4150, phismith@buffalo.edu..*

celled partitions. It shows how the theory of partitions can be conceived as a generalization of the theory of sets, and it shows how the phenomena of granularity and scaling apply most properly to partitions, rather than to the objects which are located in their cells. Granularity is, it turns out, just the other side of the coin from vagueness. At the same time we shall discover that partitions are not in general additive. Thus, while there is at any given time a single interlinked totality of all our veridical perceptions, judgments, theories, classification-schemes, databases, maps, and so forth — which is called knowledge — there is, and there can be, no corresponding single total partition.

Zoran Stojanovic*

Matrix-calculus based method for qualitative spatial reasoning

Qualitative spatial reasoning has gained increasing attention in the scientific community during the last decade in many application domains, especially in Geographic Information Systems (GIS). It has been proposed as a complementary mechanism for the automatic derivation of spatial relations, which are not explicitly stored, using relatively simple inference rules and methods. Spatial reasoning can also be used to answer queries given partial spatial knowledge as well as for maintaining the consistency of the spatial database.

In its simplest form reasoning over spatial relations is based on their composition in order to answer questions of the type: Given three spatial objects A, B and C, and two spatial relations, R_1 is a spatial relation between A and B, and R_2 between B and C, what is the relation R_3 between A and C. If the complete set of mutually exclusive and pairwise disjoint spatial relations are considered, then the full set of compositions of relations can be represented in so-called composition or transitivity table. Much research work have been dedicated to the study of the computation of such tables (Allen, 1983; Egenhofer, 1994; El-Geresy and Abdelmoty, 1997; Frank, 1996; Papadias and Sellis, 1994; Randell et al., 1992; Sharma, 1996). The common point of all proposed methods is a strong requirement to calculate each

**Department of Information Systems, Faculty of Information Technology and Systems, Delft University of Technology, Zuidplantsoen 4, 2628 BZ Delft, The Netherlands, Z.Stojanovic@its.tudelft.nl.*

entry of the table separately. They provide the derivation of the table through more or less tedious and "difficult to secure" process, using either exhaustive search or theorem proving in each particular case. In such a way, the inferring of new spatial information in the run-time is very time-consuming and non-efficient process.

In this paper, a general method for automatic derivation of the composition of two spatial relations is presented. The method is originally applied and proved in homogenous (topological-topological) and heterogenous (topological-direction) reasoning in the case of spatial regions in 2-D, but it can also be applied to other types of spatial objects. The standard 9-intersection model for representing topological relations between spatial objects in the space is used (Egenhofer and Herring, 1991). Furthermore we have introduced a new model for representing direction relations, called 12-intersection. According to 12-intersection model, the direction relation between two spatial objects (primary and reference) is defined by the existence of twelve intersections of four direction sectors defined by primary object according to cone-shaped direction model (north, west, south and east) with the boundary, interior and exterior of reference object, in the form of 4×3 matrix.

Proposed reasoning method is based on the matrix multiplication of previously adapted modeling matrices (0 = empty set, 1 = non-empty set) for relations R_1 and R_2 (two 9-intersection matrices in the case of homogenous reasoning, or 9-intersection and 12-intersection matrices in the case of heterogeneous one). The necessary information for the construction of corresponding modeling matrix (9-intersection or 12-intersection) for the relation R3 is provided through the values of the resulting matrix elements, using simple 4-step algorithm. The algorithm first results in, so called, template matrix, which defines fixed elements and their values ($\emptyset$ or $\neg\emptyset$) as a matching criteria. The modeling matrices (one or many) matching the template matrix represent all the possible results of the given composition. By using this method, the transitivity table for the whole set of topological and direction relations can be efficiently derived.

The method can be successfully applied to derive a result by integrating multiple inference directions. For example, given four objects, A, B, C and D, and their topological relations $R_1(A, B)$, $R_2(B, D)$, $R_3(A, C)$ and $R_4(C, D)$, the relation $R(A, D)$ can be derived in two ways, as the composition of R_1 and R_2, or as the composition of R_3 and R_4. The procedure is to construct the matrix templates for both cases, and then construct the resulting matrix template as their union. Now, this template consists of all the conditions defined in both starting templates, which makes the selection criteria stronger. It can result in a less number of possible resulting spatial relations then in the case of single-direction inference.

One of the main advantages of presented method is that it can be easily and efficiently implemented inside the spatial query processor, as a query support and consistency checking mechanism. In that way it can serve as a base for the development of an "intelligent" GIS query and analysis tool.

Bibliography

Allen, J., 1983. Maintaining knowledge about temporal intervals. Communications of the ACM 26 (11), 832–843.

Egenhofer, M., 1994. Deriving the composition of binary topological relations. Journal of Visual Languages and Computing 5, 133–149.

Egenhofer, M., Herring, J., 1991. Categorizing topological relationships between regions, lines and points in geographic databases. Technical report, Department of Surveying Engineering, University of Maine.

El-Geresy, B. A., Abdelmoty, A., 1997. Order in space: A general formalism for spatial reasoning. International Journal of Artificial Intelligence Tools 6 (4), 423–450.

Frank, A., 1996. Qualitative spatial reasoning: Cardinal directions as an example. International Journal of Geographic Information Systems 10 (3), 269–290.

Papadias, D., Sellis, T., 1994. Qualitative representation of spatial knowledge in two-dimensional space. The VLDB Journal 3 (4), 479–516.

Randell, D., Cohn, A., Cui, Z., 1992. Computing transitivity tables: A challenge for automated theorem provers. In: CADE'92.

Sharma, J., 1996. Integrated spatial reasoning in geographic information systems. Ph.d. thesis, Department of Spatial Information Science and Engineering, University of Maine.

*Erik Stubkjær**

Real estate and the ontology of multidisciplinary, e.g. cadastral, studies

Abstract. Among spatially extended objects, units of real estate constitute a distinct category. Their complex relationships with other phenomena are of an economic, legal, political, and spatial nature; the related property rights count among the basic institutions of society. A multidisciplinary project has been proposed to establish a coherent knowledge base in this field. The project focuses on real property transactions within European countries and addresses the ontology of real estate and its boundaries. The ontology of real estate must refer to conceptualizations provided by the scientific disciplines of economics, geosciences, law, and political science, and draw upon domain knowledge of multidisciplinary studies that regards real estate, e.g. cadastral studies. The complex task of reconciling the ontologies of the diverse disciplines into one common core ontology of real estate is presented and discussed. Development criteria include that the core ontology is robust against changes in professional scopes and technology, and valid across different cultures.

**Dept. of Development and Planning, Aalborg University, Fibigerstræde 11, DK 9220 Aalborg East, Denmark, fax +45 9815 6541, est@i4.auc.dk.*

1. Introduction

Among spatially extended objects, units of real estate constitute a distinct category. An approach to the conceptualization of the spatial dimension of real estate is presented in section 2. The ontology of real estate cannot be adequately described without reference to the conceptualizations provided by the scientific disciplines of law, economics, and political science. Section 3 of the paper addresses this complexity, drawing upon the notion of *information communities*. The term *ontology* is used differently by philosophers and in the field of information processing, respectively. In linguistics mention is made of *ontological categories*. These diverse approaches to the eliciting of ontology are explored (section 4).

The preparation of an ontology of real estate is a substantial task, as it includes a reconciling of some of the conceptualizations of the above mentioned disciplines. Multidisciplinary studies in the field of real estate may provide for relevant domain knowledge, and contribute to the establishing of consistent, canonical formulations of the real estate universe of discourse. *Cadastral studies* are presented as an example of such multidisciplinary endeavor (5). The closing section proposes the eliciting of a *cadastral ontology*. It raises the question whether such a project is feasible and relates this question to other projects in ontology.

2. Approaching the ontology of real estate: The spatial dimension

The most visible boundary of real estate extends in the spatial dimension: A unit of real estate has a location in space and a boundary that has a spatial dimension. (The term 'dimension' is here used in a more general sense than the usual three dimensions of physical space.) Both location and boundary are of a rather complex nature. Because a unit of real estate is an socioeconomic unit, its boundaries cannot be exhaustively described in spatial terms, to be detailed in the following section 3.

The spatial dimension of the boundary of real estate is appropriately described by the *fiat-bona fide* dichotomy introduced in Smith (1995). To give some examples: A *bona fide* boundary like a stream may be used for the definition of a *fiat* boundary, e.g. the boundary of a jurisdiction. *Fiat* boundaries may be established as mere abstract lines that are traced on a plan for the division of land, in order to structure its settlement. However as time passes, the owners occupy and use the terrain, bringing about that the initial *fiat* lines become visible in the terrain as buildings, fences, roads and that like. The *fiat* boundaries thus become *bona fide*. The *fiat-bona fide* distinction can provide a basis for comparisons of the boundary setting practices of different countries, as well as for investigations into topological problems that extend beyond the geospatial realm in what is called mereotopology Smith and Varzi (2000).

Property boundaries are often located by means of a national, geodetic reference system. However, a cadastral locational description needs more than a specification of position with reference to a geodetic reference system. This is partly because the owners of the property typically do not understand the language of geodetic referencing, and it is also because the neighbor relations among the estate units cannot be represented simply by coordinates (Laurini and Thompson, 1992). To accommodate for the needs of the owner and other citizens, the units need to be described relative to *place names*, especially road names. The neighbor relations must be made to appear within a cadastral map that depicts parcel identifiers, or through alternative information media. Using the *scales of measurement* of Stevens (1946) in an adapted version, we arrive at a minimal list of spatial reference frames (Table 25.1, Stubkjær, 1992).

A final remark regarding the spatial extension of real estate concerns its relation to other spatial, socio-economic units. Stubkjær (in Frank et al., (to appear) suggests that the unit of ownership be categorized as a *jurisdiction*, which is distinguished from other classes of socio-economic units: *place names*,

Nominal (verbal)	Place names; Cadastral identifiers; Address codes
Ordinal (graphical, topological)	Neighbor relations; House numbering sequences
Metric (numerical)	Coordinates of boundaries, road center-lines, etc.; 'Metes and bounds'

Table 25.1: Minimal list of spatial reference frames.

districts, and *regions*, respectively.

Summing up, the conceptualizations of the mentioned disciplines and the spatial concepts presented may be applied for eliciting the rational core of an ontology of real estate that is independent of the rule sets and practices of a specific European country.

3. Reconciling the conceptualizations made by established academic disciplines

As mentioned above, the most visible boundary of real estate extends in the spatial dimension. However, it is the courts, which ultimately settle the determination of the spatial boundary in cases of dispute. This implies that the boundary of real estate has to be described also by using the conceptualization of the discipline of *law*. Moreover, law describes the non-spatial boundary of a unit of real estate. For example, the question what items belong to the estate when no specific statements are made, is a legal issue. The question is not simple, as the answer may vary according to whether the context is transfer of ownership, mortgaging, or assessment for taxation.

In European countries and elsewhere, property rights to real estate are formalized and recorded in land registries maintained by the courts. The real property rights formalized therein provide the basis for fairly transparent real property markets. Now, the field of real property markets is the object of the discipline of *economics* and its sub-fields, e.g. microeconomics, and new institutional economics. Furthermore, real estate is an object of taxation with profound political implications. Also, policy issues are involved in the recurrent change of the administrative systems, which are needed for taxation, recording of real property rights, and regulation according to spatial planning, etc. As a consequence, the conceptualization of subfields of the discipline of *political science* has to be taken into account in order to understand the changes of administrative units and information systems, which are related to real estate.

The need of an investigation of these diverse conceptualizations of the phenomena of real estate by the disciplines of *law*, *economics*, and *political science* has thus been established. This raises the issue of an appropriate approach for the eliciting of a formal ontology of real estate.

Clifford Kottman introduces the notion of *information communities* by referring to John Locke (1999, p. 46f.). A similar notion is the 'thought collectives' mentioned by Fleck (Ziman, 1992, referring to Fleck, 1935/1979), and the *scientific communities* of Thomas Kuhn, who draws upon Fleck's notion (1970). A discussion of these and similar notions of *communities* is desisted. Bishr *et al.* discuss the notion of a *geospatial information community.* They suggest the following definition: "A geospatial information community is a group of spatial data producers and users who share an ontology of real-world phenomena", and consider the ontology as "a meta-language situated above data models" (1999, p. 58). Two communities may have different ontologies, but in order to share information they must have a part of their ontologies in common.

One may conceive the scholars and practitioners of the academic disciplines of law, economics, and political science, respectively, as members of three distinct *information communities.* To arrive at an ontology of real estate one has to establish the constituent ontologies of the mentioned *information communities* that refer to real estate, as well as a rational core ontology of real estate, which they share in common.

4. The diverse notions of ontology

Addressing the ontology of spatially extended objects, one should be aware of different use of the term *ontology*. Barry Smith points to the fact that the term *ontology* is used differently by philosophers, and in the field of information processing respectively (2000). He characterizes the main concerns of the two communities, and in attempting a method for bridging the different ways of understanding this term he refers to the efforts within biology to construct ontologies that apply to the term 'gene' and similar fundamental biological units. The methods include the preparation of common vocabularies, and the formulation of appropriate translation rules between the diverse nomenclatures of the different branches of biology. Biologists cooperate with ontology engineers, as well as with philosophers in this endeavor. This approach motivates similar efforts in other scientific fields, as we shall see in the next section.

It should be noted that the term *ontology* is used by a further community, namely that of linguists. In *Semantics and Cognition* Ray Jackendoff discuss how visual information, linguistic information and other peripheral information is mapped onto mental representations (1983). Through an analysis of human perception he arrives at the following list of *ontological categories*: THING, PLACE, DIRECTION, ACTION, EVENT, MANNER, and AMOUNT. The list is not meant to be complete. He claims, however, that "the total set of ontological categories must be universal: it constitutes one basic dimension along which humans can organize their experiences." (Jackendoff, 1983; Stubkjær, 1994).

This section identified three communities, which are concerned with eliciting of ontology: philosophers, linguists, and ontology engineers. Their diverse methods of eliciting ontology may be applied to the real estate universe of discourse. In fact, it is an approximation to speak of one universe of discourse of real estate. Rather, the diverse conceptualizations of real estate that are provided by law, economics, political science, and the geosciences, respectively, in fact constitute four diversely overlapping universes. The task is thus to reconcile these different ontologies into one common ontology of real estate.

5. Real Estate, an entity within the multidisciplinary cadastral universe of discourse

Smith notes that "(e)very scientific field will .. have its own preferred ontology, defined by the field's vocabulary and by the canonical formulations of its theories" (Smith, 2000, p. 1). This agrees with Ziman, who regards the change of educational curriculum the final effect of a new (theory-based) discipline (Ziman, 1992, p. 94). However, university departments and especially branches of studies have indeed sprung up in response to societal needs, rather than to the maturing of theory-based efforts. These university fields are often multidisciplinary (i.e. drawing upon theory-based and other scientific fields), and they have yet to establish canonical formulations of their universe of discourse.

The scientific fields that want to establish their 'preferred ontology' are assumed to face the same problems as the ontology engineer, and may apply the same methods. Their specific role in the general project of eliciting of ontology may be to communicate explicit knowledge of their universe of discourse. Their 'domain knowledge' may to some extent depart from the domain knowledge of the practitioner within the same field, as well as depart from the applied theoretical domain knowledge. The author presents his own discipline: cadastral studies, as an example of a multidisciplinary study that is in the process of establishing canonical formulations.

In Europe, the cadastre has developed since 1700 in the context of centralization of administration and the issuing of tax ordinances (Sommer, 1930). Cadastral concerns at university level branched from geodesy and land surveying. The university teaching of cadastral issues: Cadastral law, property rights, and spatial planning regulations, is largely bound to a national scope. Comparing to the study and teaching of natural languages: English, German, etc., which similarly are bound to a national setting, one can note that the concept and study of a 'cadastral grammar' is missing.

However, from the 1970s a concern for an international scope manifests itself. For example, the *Fédération Internationale des Géomètre* (International Federation of Surveyors) issued among others *The FIG Statement on the Cadastre* (FIG, 1995) and the *Bathurst Declaration on Land Administration for Sustainable Development* (FIG, 1999).

Those rather normative statements from the surveyors' profession have been accompanied by research from the point of view of formal modelling (Frank, 1996), benchmarking (Steudler et al., 1998), or with a view to charting the interrelated technical, legal, and organisational aspects of this field (Zevenbergen, 1998). With a view of establishing a theoretical basis for cadastral studies, Stubkjær recently surveys research in information systems development, and research within geographical information science (1996; 1999). Subsequently a view of the cadastral universe of discourse is presented in a submitted paper. This view was drafted with reference to the Soft Systems Methodology (SSM) in its early version (Checkland, 1981). The view is graphically rendered through Figure 25.1. (The notion of 'problem domain' is borrowed from SSM, and replaced here by 'universe of discourse'.)

Stubkjær and Ferland (2000) discuss the design of research within the multidisciplinary cadastral domain. SSM suggests a 'political system analysis', but SSM does not provide the concepts and methods of political science. Knowledge of these concepts and methods are needed in order to elicit a valid ontology of the cadastral universe of discourse. At least in principle, the same applies regarding the conceptualizations used by the other theory-based disciplines. However, given our currently established research methods, it is often unfeasible to monitor and apply research results from several disciplines, especially for a

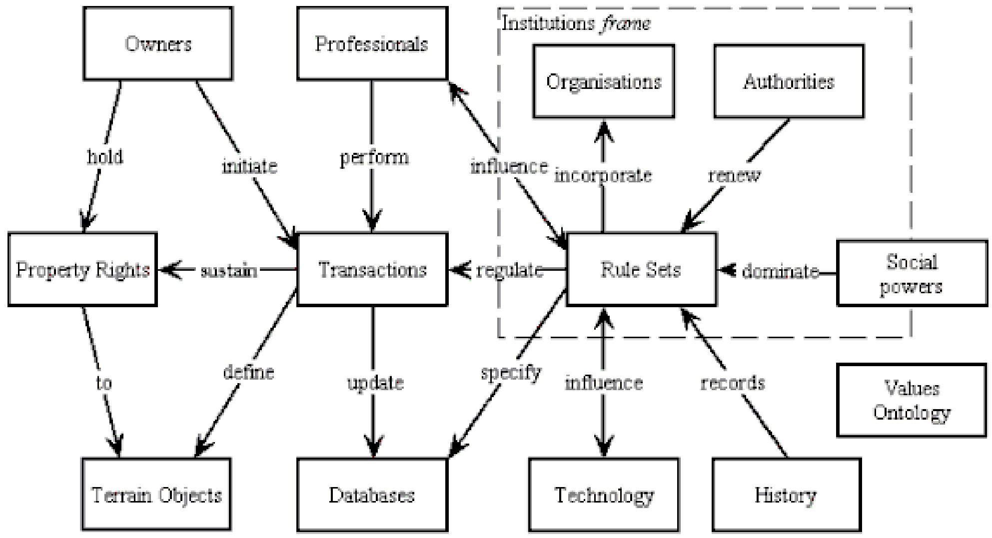

Figure 25.1: A view of the cadastral problem domain (Stubkjær, 2000).

project made in the context of a Ph.D.-study. This leads to the following *suggestion for a research strategy*: The elements of the cadastral universe (cf. Figure 1) are described with reference to the main works of the relevant disciplines through an international project. Research projects of the PhD-type can then draw upon the outcome of this effort and — within the scope of the project — investigate and apply the most recent research of the established disciplines.

A multidisciplinary European project has been proposed as a *COST action* to establish a coherent knowledge base in this field. The project focuses on the transactions of real property within various European countries, and addresses the ontology of real estate (Stubkjær, 2000).

6. The feasibility of eliciting a cadastral ontology

Multidisciplinary studies of the sort envisaged need to find ways to cope with constant changes within their target society, including changes in the scope of different professions and changes due to new technology. Furthermore, there are changes due to research outcomes from the scientific fields, which they draw upon. A proactive strategy for addressing such changes is to establish a robust ontology of the relevant universe of discourse. As a further benefit, the development of rich and rigorous knowledge structures of the universe will, potentially, allow for the accumulation of generally valid knowledge, and thus reduce the amount of context specific teaching, relating, for example, to specific national rules and practices. In the present case, a robust *cadastral ontology* is to be established.

Multidisciplinary studies are likely to be implicitly bound to a specific culture. For example, the study and teaching regarding an operating cadastre and other traits of formal property rights seem to presuppose a specific societal culture, including a fairly uncorrupt administration, manned with skilled staff (a 'Weberian' administration). In order to transgress the boundary of European/Western culture, the eliciting of the core ontology must be rooted in the philosopher's conception of phenomena like man, society, government, reality, language, and representation, in order to account for the diverse expectations related to government, for example. Without such basic concepts clearly stated, and their implications for the unit of real estate spelled out, the knowledge that is generated through cadastral studies can hardly transgress cultural boundaries.

It may be that such a project is not feasible. An exploration of this claim may lead to possible feasibility boundaries of the project of establishing formal representations of spatially extended objects.

Bibliography

Bishr, Y. A., Pundt, H., Kuhn, W., Radwan, M., 1999. Probing the concept of information communities – a first step toward semantic interoperability. In: Goodchild, M., Egenhofer, M., Fegeas, R., Kottman, C. (Eds.), Interoperating Geographic Information Systems. Kluwer, Boston, pp. 55–69.

Checkland, P., 1981. Systems Thinking, Systems Practise. John Wiley, Chichester.

FIG, 1995. The fig statement on the cadastre. FIG Publications 11, The International Federation of Surveyors (FIG), Frederiksberg, Denmark.

FIG, 1999. The bathurst declaration on land administration for sustainable development. FIG Publications 11, The International Federation of Surveyors (FIG), Frederiksberg, Denmark.

Frank, A., 1996. An object-oriented, formal approach to the design of cadastral systems. In: Kraak, M. J., Molenaar, M. (Eds.), Advances in GIS Research II – Proceedings of 7th International Symposium on Spatial Data Handling. Vol. 1. IGU, Delft, The Netherlands.

Frank, A., Raper, J., Cheylan, J.-P. (Eds.), (to appear). Life and Motion of Socio-Economic Units. ESF Series. Taylor & Francis, London.

Jackendoff, R., 1983. Semantics and Cognition. MIT Press, Cambrigde, MA.

Kottman, C. A., 1999. The Open GIS Consortium and progress toward interoperability in GIS. In: Goodchild, M. F., Egenhofer, M., Fegeas, R., Kottman, C. (Eds.), Interoperating Geographic Information Systems. Kluwer, Norwell, MA, pp. 39–54.

Kuhn, T. S., 1970. The Structure of Scientific Revolutions, 2nd Edition. Vol. 2(2) of International Encyclopedia of Unified Science. The University of Chicago Press, Chicago.

Laurini, R., Thompson, D., 1992. Fundamentals of Spatial Information Systems. Vol. 37 of The A.P.I.C. series. Academic Press.

Smith, B., 1995. On drawing lines on a map. In: Frank, A. U., Kuhn, W. (Eds.), Spatial Information Theory. Vol. 988 of Lecture Notes in Computer Science. Springer, Berlin, pp. 475–484.

Smith, B., 2000. Ontology: Philosophical and computational, draft: July 22, 2000.

Smith, B., Varzi, A. C., 2000. Fiat and bona fide boundaries. Philosophy and Phenomenological Research 60 (2), 401–420.

Sommer, L., 1930. Cameralism. In: Encyclopaedia of the Social Sciences. MacMillan, pp. 158–161.

Steudler, D., Williamson, I. P., Kaufmann, J., Grant, D., 1998. Benchmarking cadastral systems. The Australian Surveyor 42 (3), 87–106.

Stevens, S., 1946. On the theory of scales of measurement. Science 103 (2684), 677–680.

Stubkjær, E., 1992. The development of national, multi-purpose spatial information systems – danish experiences in a theoretical context. Computers, Environment and Urban Systems 16 (3), 209–217.

Stubkjær, E., 1994. Employing the linguistic paradigm for spatial information. In: Waugh, T. C., Healey, R. G. (Eds.), Advances in GIS research – Proceedings of the Spatial Data Handling Conference. Vol. 1. Edinburgh.

Stubkjær, E., 1996. A theoretical basis for cadastral development. In: Proceedings of the Third Seminar of the European Co-operation Network for Education and Research in Land Information Systems (EUROLIS). Delft University of Technology, Delft.

Stubkjær, E., 1999. Cadastral research – issues and approaches. Kart og Plan 59 (3), 267–278.

Stubkjær, E., July 2000. Modelling real property transactions, proposal for a new Cost Action.

Stubkjær, E., Ferland, Y., 2000. An adequate research design for cadastral development. In: AAG 2000 – Session on Cadastral concerns I. Pittsburgh PA.

Zevenbergen, J., 1998. The interrelated influence of the technical, legal and organisational aspects on the functioning of land registrations (cadastres). In: Proceedings FIG XXI International Congress (Commission 7). Brighton.

Ziman, J., 1992. An Introduction to Science Studies – The Philosophical and Social Aspects of Science and Technology. Cambridge University Press, Cambridge, reprint from 1984.

Heiner Stuckenschmidt*

Semantic translation of land-use classifications: a case study

Over the last few years much work has been conducted in regards to the research topic fully interoperable GIS (Vckovski, 1998). GIS's share the need to store and process large amounts of diverse data, which is often geographically distributed. Most GIS's use specific data models and databases for this purpose. This implies, that making new data available to the system requires the data to be transferred into the system's specific data format. This problem is addressed by the Bremen University Semantic Translation Project[1]. The main deliverable of the project is a knowledge-based system for the integration of (geographic) data. The system distinguishes three levels on integration: Syntactic, Structural, and Semantic integration.

In this paper we focus on the semantic aspects of intelligent information integration that tries to preserve the intended meaning of information entities in a different context. We propose to interpret this semantics-preserving context transformation as a classification task (Stuckenschmidt and Visser, 2000): translating an information item from one context to another then becomes the task of taking the properties of information item from its source context, and use these properties to re-classify the item in its target context, resulting in a re-interpretation of

* *Intelligent Systems Group, Center for Computing Technologies, University of Bremen, P.O.B.: 33 04 40, D-28334 Bremen, Germany, heiner@tzi.de, http://www.tzi.de/~heiner/.*

[1] `http://www.semantic-translation.de`

the information item in the new context. We investigated the role of ontologies for providing a shared terminology and support for the integration process. After addressing these questions in principle we summarize the results of a case study were we used the Ontology Interchange Language OIL (Fensel et al., 2000) in order to support the integration of different catalogue systems. We assess useful features of the language and point towards open problems that have to be addressed in future research.

The results of a case study showed that the integration of two terminologies for land-use classifications is possible in principle using the ontology specification language OIL in combination with the description logic reasoner FaCT (Horrocks, 1998). However, there are still some open questions that may explain the difficulties we also experienced (Stuckenschmidt, 2000). The most striking difficulties arise from the fact, that the terminologies we tried to integrate were defined in terms of natural language statements that are often affected by the use of vague expressions, inaccessible criteria and definitions that can only be understood making extensive use of commonsense knowledge. These uncertainties in the definition made it nearly impossible to provide sound and complete specifications to be used in the translation process. As a consequence, translation by subsumption reasoning often failed.

In our opinion there are two possibilities to overcome this problem:

1. We can use heuristic and approximate classification methods to get a better handle on the inherent uncertainty of the knowledge involved.

2. We have to assure that catalogue systems use a well-founded basis in terms of a formal ontology in order to define their terminology.

The evaluation of pros and cons of both approaches are subject to current research.

Bibliography

Fensel, D., Horrocks, I., van Harmelen, F., Decker, S., M., E., Klein, M., 2000. Oil in a nutshell. In: Proceedings of the 12th European Workshop on Knowledge Acquisition, Modeling and Management. Lecture Notes in Artificial Intelligence. Springer, Berlin.

Horrocks, I., 1998. Using an expressive description logic: Fact or fiction? In: Cohn, A. G., Schubert, L., Shapiro, S. C. (Eds.), Principles of Knowledge Representation and Reasoning: Proceedings of the Sixth International Conference (KR'98). Morgan Kaufmann Publishers, San Francisco, California.

Stuckenschmidt, H., 2000. Using oil for intelligent information integration. In: Proceedings of the Workshop on Application of Ontologies and Problem-Solving Methods at the European Conference on Artificial Intelligence ECAI 2000. Berlin.

Stuckenschmidt, H., Visser, U., 2000. Semantic translation based on approximate re-classification. In: Proceedings of the Workshop "Semantic Approximation, Granularity and Vagueness" at the International Conference on Principles of Knowledge Representation and Reasoning KR' 2000. Breckenridge, Colorado.

Vckovski, A., 1998. Interoperable and Distributed Processing in GIS. Taylor & Francis, London.

Nectaria Tryfona*

The role of ontologies in spatial data mining

Spatial data mining is to mine high-level spatial information and knowledge from large geographic data sets (Han et al., 1997). A typical example is find all areas with vegetation growth (5% and vegetation type "pine". While for traditional application areas such as banking systems, efficient methods and functions to mine information do exist, this is not the case for spatial applications (Tryfona, 2000). On the other hand, as the number of applications dealing with spatial data is growing rapidly and, at the same time, the amount of spatial data is increasing, the efficient querying and employment of stored information for the extraction of further knowledge is emerging.

In this work we discuss the role of ontologies in the spatial mining process. Gruber (1997) defines an ontology as an explicit specification of conceptualization. Fronseca et al. (2000) refer to spatial ontologies as dynamic, object-oriented structures that can be navigated.

It is our position that in order to successfully perform spatial data mining the role of ontologies must be investigated and comprehended. In the spatial mining framework:

- We propose a systematic categorization of spatial ontologies, such as spatial entity, map, boundary and topology participating in the geographic environments to be mined.

* *Department of Computer Science, Aalborg University, Fredrik Bajersvej 7E, 9220 Aalborg Øst, Denmark, tryfona@cs.auc.dk.*

- We discuss relations among spatial ontologies, as well as their participation in constraints and hierarchies, two fundamental aspects of the mining process.
- We show how traditional data mining methods can be combined with spatial ontologies.

The proposed outcome (i) serves as a guideline for the designer on the modeling process of a spatial environment on which data mining will be performed, (ii) helps the designer to recognize the basic ontologies, of an already existing spatial environment, and their roles in the mining process, (iii) provides the designer with a language to perform spatial data mining and, (iv) leads towards an ontology-driven architecture (Fronseca et al., 2000) for a spatial miner.

Bibliography

Fronseca, F., Egenhofer, M., Davis, C., Borges, A., 2000. Ontologies and knowledge sharing in urban GIS. Computer, Environment and Urban Systems 24 (3).

Gruber, T., 1997. A translation approach to portable ontology specifications. Technical Report TR-KSL 97-71, Knowledge Systems Labolatory, Stanford University, Stanford, CA.

Han, J., Koperski, K., Stefanovic, N., 1997. Geominer: A system prototype for spatial data mining. In: Proc. of the ACM SIGMOD International Conf. On Management of Data. Tucson, Arizona.

Tryfona, N., 2000. Constraint-based spatial data mining. Computer, Environment and Urban Systems, (in press).

Andrew Turk*

Tribal boundaries of Australian indigenous peoples

This paper explores the concept of boundary ('limit of country') held by indigenous Australians. It is not, however, a work of Anthropology. Rather, it is an interpretation of the writings of some anthropologists on this topic, aided by discussions which the author has had over a number of years with people from the Ngaluma, Injibarndi and Banjima tribes in the Pilbara region of Western Australia.

This analysis provides a partial understanding of the nature of tribal boundaries, especially variations which occur in the physical definition of boundaries and their (intentional and unintentional) indeterminacy. It suggests that the description of any specific boundary needs to incorporate the source of the boundary construct (e.g. the relevant indigenous law and customs), the people holding that construct (e.g. members of the tribe or language group who have responsibility for that 'country') and the form of representation of the boundary (e.g. singing a sequence of place names). The nature of the 'country' itself is also a key determinate of the expression of the boundary concept at any particular place. Some parallels in the concepts of boundary held by non-indigenous Australians are also discussed.

The paper goes on to draw some conclusions regarding the representation of indigenous boundaries in the property cadastres of Australian States and Territories. This has particular

* *Murdoch University, Western Australia, a_turk@central.murdoch.edu.au.*

importance in the context of the implementation of legislation concerning native title land claims. If such 'official' boundaries are to do justice to indigenous law and culture, they must reasonably reflect the ontology and epistemology of the concepts of boundary held by indigenous Australians.

Achille C. Varzi [*]

Are there vague boundaries?

Vagueness is a pervasive phenomenon of human thought and language and geography is not exempted from its grasp. It is, in fact, a characteristic feature of most ordinary geographic concepts that they involve some degree of vagueness: How small can a town be? How long must a river be? How many islands does it take to have an archipelago? More importantly, many individual names and descriptions used in geography appear to refer to entities whose boundaries are only vaguely defined: What are the borders of Mount Everest? Where does the Outback begin? What exactly is the territory occupied by the capital of Italy? by Rio de Janeiro? by Greenwich Village?

In this talk I will focus on the vagueness exhibited by individual terms such as these and I will examine two main ways of thinking about it. On one conception, the relevant vagueness is ontological (or *de re*): geographic names and descriptions may be vague because they may refer to vague objects. For instance, on this view 'Mount Everest' would refer to an object, Mount Everest, whose boundary is genuinely fuzzy: some molecules are inside it, some molecules outside, and some have an indefinite status (there is no objective, determinate fact of the matter about whether they are inside or outside). By the same pattern, valleys, deserts, dunes, rivers, bays, forests, cities, neighborhoods, states (with few exceptions such as Wyoming) are all genuinely vague denizens of reality.

On a second conception, the vagueness exhibited by geo-

[*] *Department of Philosophy, Columbia University, New York, av72@columbia.edu.*

graphic names and descriptions is exclusively semantic (or *de dicto*). It lies exclusively in the representation system, not in the represented entity, and to say that the referent of a term is not sharply demarcated is to say that the term vaguely designates an object, not that it designates a vague object. For example, on this view there is no such thing as a vague mountain. Rather, there are many things where we conceive the mountain to be, each with its precise boundary, and when we say 'Mount Everest' (or when the founder of the Indian Geodetic Office baptizes a certain piece of land, at the border between Tibet and Nepal, 'Mount Everest') we are just being vague as to which thing we are referring to: each one of a large variety of slightly distinct but perfectly determinate aggregates of molecules has an equal claim to being the referent of that name. If we wish, we can add that it is ultimately the vagueness of the relevant sortal concept (the concept mountain, in this case) that is responsible for the way in which the referent of our expression is vaguely picked out. But the stuff out there is all but vague.

The point of this talk is to compare these two conceptions of vagueness and to offer reasons for preferring the latter. There are no vague boundaries in reality but, rather, vague ways of drawing boundaries. In addition, I will also outline a way of dealing with de dicto vagueness which appears to be particularly suited for application in the geographic domain. Broadly speaking, this is based on the method of supervaluations: when dealing with vague terms (or vague representations more generally), consider the many possible ways in which those terms can be made precise and compute the pattern of agreement among them. If a statement comes out true on some "precisifications" and false on others, then there is little one can do: the relevant vagueness induces a truth-value gap. But if a statement is true under all such precisifications (or false under all precisifications), then one can naturally regard it as being true (false) in spite of the relevant vagueness; the unmade semantic stipulations do not matter. This allows us to explain why, for example, we can confidently assert that Mount Everest is in Asia and we can confidently deny that it is in Europe, though we must suspend judgment when it comes to saying whether Mount Everest is mostly in Tibet: the truth-value of such a statement depends crucially on how much land one includes in the referent of 'Mount Everest'.

If time permits, in the final part of the talk I will also deal with some general difficulties relating to this account: these include various analogues of the sorites paradox, the phenomenon of higher-order vagueness, and the relationship between vagueness and the general issue of geographic ontology.

Publications from the GEOINFO-SERIES

1. Experiential Realism and its Applications to Geographic Space
   Compiled by Irene Campari and Andrew Frank

2. Temporal Data in Geographic Information Systems (2^{nd} print)
   Compiled by Andrew Frank, Werner Kuhn and Peter Haunold

3. Geo-Information Management Systems
   A Major Public Issue and its Educational Challenges
   Compiled by Andrew Frank and Irene Orchard

4. Geographic Information Systems
   Materials for a Post-Graduate Course – Vol. 1: Spatial Information
   Andrew U. Frank, Editor

5. Geographic Information Systems
   Materials for a Post-Graduate Course – Vol. 2: GIS Technology
   Andrew U. Frank, Editor

6. Geographic Information Systems
   Materials for a Post-Graduate Course – Vol. 3: GIS Organization
   Andrew U. Frank, Editor

7. Semantics of Geographic Information
   Werner Kuhn

8. Spatialization: Spatial Metaphors for User Interfaces (2^{nd} print)
   Werner Kuhn and Brad Blumenthal

9. COSIT '95 Doctoral Consortium
   Compiled by Werner Kuhn and Sabine Timpf

10. Hierarchical Spatial Reasoning: Theoretical Consideration and its Application to Modeling Wayfinding
    Adrijana Car

11. Aufdeckung von numerischen Problemen in geodätischer Software
    Christine Goldenhuber

12. Gofer as used at GeoInfo/TU Vienna
    Andrew U. Frank, Werner Kuhn, Werner Hölbling, Hartmuth Schachinger, Peter Haunold

13. Hierarchical Structures in Map Series
    Sabine Timpf

14. Organistion des Katasters – Ziele, Grundsätze und Praxis (2^{nd} print)
    Christoph Twaroch

15. Die Modellierung eines Grundbuchsystems im Situationskalkül
Steffen Bittner

16. Multi-Agency Databases to Manage Geographic Information
Andrew U. Frank, Martin Raubal, Maurits van der Vlugt (Editors)

17. A Technical Concept for Pay-per-Use in Geomarketing Services
Peter Gustav Wenzl

18. Ego-Reorientierung und Abruf egozentrischer Targetrichtungen — Modell und Empirie
Annette von Wolff

19. Geographical Domain and Geographical Information Systems
Stephan Winter, Editor

20. Unified Behavior of Spatial Data Representations
Stephan Winter

The books are available from:

Institute for Geoinformation
Vienna University of Technology
Gusshausstrasse 27-29/127
1040 Vienna, Austria
fax: ++43-1-58801-12799
e-mail: geoinfo@geoinfo.tuwien.ac.at
http://www.geoinfo.tuwien.ac.at