

Geonoemata Elicited: Concepts, Objects, and Other Uncertain Geographic Things

Marinos Kavouras

School of Rural and Surveying Engineering, National Technical University of Athens, 9, H. Polytechniou Str., 157 80 Zografos Campus, Athens, Greece, mkav@mail.ntua.gr

Abstract

Almost thirty years after the beginning of geographic information science (GIScience) as an interdisciplinary but distinct scientific field, new and deeper research questions have arisen, questions which make us return back to the fundamental issues of geographic concepts, knowledge representation, and semantically-aware approaches. The questions are very difficult to answer, yet this should not prevent us from always pursuing the very nature of geographic meaning. It is evident that meanings and understandings in the geospatial domain (thereinafter called “geonoemata”) pivot around the connection between the central representational notions of concepts and objects. The use and application of these notions can be accounted for most problems in interoperability, non-universality of approaches, misinterpretation, and semantic conflicts. In this section, an attempt is made to identify a number of open and promising research questions in great need for progress.

1 Introduction

It is always fascinating to retrospect and realize the long distance covered by Geographic Information Systems and Science in the last 30 years at least. The path has not been an easy one, the objective was neither self-evident

nor always clear, and some prominent scientists have throughout the years (and their life) significantly contributed to what has become a renowned field today. Having appreciated this long contribution, it would be interesting but also fruitful to examine what was at times the research agenda, which questions have either found a sufficient answer or present no interest anymore, and what went right or wrong in these 30 years. But, even more importantly, it would be crucial to identify which issues remain open and long-standing, which questions still have not found a sufficient answer, and which theories will provide an essential core and make GIScience not just survive but thrive in the years to come.

Indeed, being more knowledgeable now, and after having resolved or left behind us many issues especially technical ones, it is not difficult to realize that the real problem was not the initial lack of geodata or standards. Nor it was how data are formalized, structured, or organized in a database. And clearly it was not how fast algorithms perform and can be optimized. Without diminishing their role, what really kept GIScience going and not being absorbed by other technological fields, were the open (and difficult to answer) theoretical questions. The interdisciplinary but also distinct character of GIScience has attracted scientists from many other seemingly diverse scientific disciplines.

The truth is that many of these theoretical/fundamental issues are not completely new. They were addressed early enough by some envisaged scientists, but maybe too-early to be appreciated by the large geographic information community. Two reasons changed the picture.

The first reason was the expansion of “traditional” spatial science to other disciplines: Philosophy, Epistemology, Linguistics, Cognitive Science, Psychology, Knowledge Science and Engineering, etc. Parenthesis: It could be claimed that this expansion completely reversed the classic question “What is special about spatial information?” to “How much information is left that cannot be treated as spatial?”

The second reason that helped change the picture was (a) the progress of systems, applications, sensors, and communications (including the web), and (b) the enormous acquisition of geospatial¹ data. The natural next step was to associate and integrate this information, which revealed the serious issue of incompatibility, especially that of semantic conflicts. All attempts to standardise information failed at some level, simply because you cannot completely standardize the way people observe, perceive, think, and form concepts.

In other words, what proves to be essential is the way reality is carved and represented, and how the resulting differences in meaning and knowl-

¹ Throughout the Section, the terms “geographic” and “geospatial” are used interchangeably.

edge about geospace, i.e., *geonoemata*, are elicited. It is a necessity to understand this duality relation between the representational notions of *concepts* and *objects*. This duality is met in different fields under different names, but presents a strong similarity. Whether known as <concept ↔ object>, <concept ↔ symbol>, <signifier ↔ signified>, <territory ↔ map>, <word ↔ idea>, <sign ↔ thing>, (Frege 1892; Ogden and Richards 1923; Korzybski 1933) or other, it all boils down to the fact that a representation is a substitute. It is neither the concept formed when observing a real thing (geographic reality in our case), nor the thing itself. In the same context, Frank (2000) uses a multi-agent system as a model for map production, communication, and use. The model consists of two disjoint tiers: reality and beliefs the first tier represents the environment and the second the agents' understanding and knowledge of the environment.

Semantic-aware approaches became thus a critical issue to geospatial research. The different and non-complementary representations of geographic reality (lets call it "poems") are the expression of views, conceptions, and meanings of their creators ("the poets"), which at the other end are received diversely by the different users ("poetry readers") creating different views, conceptions, and meanings. With the dramatic increase of internet users, the problem escalates. Data and its intended meaning are not confined to domain experts or the community that created it, but to a wider and a-priori unknown community of users. This often prevents the development of universal approaches, while it often leads to data misinterpretation and misuse. Without resorting to Husserl's Phenomenology and deeper uses of "noema/noesis", the systematic clarification of "geonoemata" and semantic conflicts is considered to be of vital importance towards a proper elicitation of geo-knowledge. This inevitably entails significant progress in ontological research.

The deeper questions are indeed very difficult to answer, yet this should not prevent us from always pursuing the very nature of geographic meaning. The world functions for centuries (although admittedly not well) without having resolved these notions, by making assumptions and by using axioms, limited theories, vague knowledge, heuristics, rules of thumb, trial and error, etc. Probably, the methodological instruments of GIScience need to take such approaches into consideration, be aware of all the assumptions made, and proceed to pragmatic issues.

In this section, an attempt is made to identify a number of open research questions in great need for progress. Naturally, the list itself is also open - it could not be otherwise.

2 Obstacles and Challenges

Although research on geosemantics and ontologies has tremendously increased in the last ten years, its practical results are still limited. This can be accounted to a number of reasons or obstacles (Kavouras and Kokla, 2008). The same obstacles on the other hand, present open research areas and important challenges. Three main areas are the following:

1. There is no clear understanding (not to call it lack of awareness or even confusion) first about the meaning of the notions, and secondly about the objectives set by users. That is to say, semantic conflicts are not only something between concepts, but they are also a problem at a meta level, between notions very loosely used, such as *concepts*, *objects*, *ontologies*, *semantics*, *integration*, *similarity*, *context*, etc. As a result, it is not easy to design appropriate solutions to problems.
2. Besides the significant research accomplishments in GIScience, there has not been a remarkable progress towards developing a concrete and universal *corpus of geographic knowledge*. This would require a synthesis of all independent theories and methods into a lattice of meta-ontologies which after thorough analysis and reviewing would be valuable to geo-ontology engineering.
3. The third obstacle is the lack of well established *formal instruments*, to deal with the highly demanding knowledge representation/engineering needs. Traditional methodologies, algorithms and data structures cannot play this role anymore. Given a task, the user needs to know which conceptual structure is more expressive or appropriate, how conceptual structure are converted or integrated, etc. Also, the user needs to know what options exist for the elicitation of semantic knowledge from different sources including textual descriptions.

The first two obstacles, and especially the second one, appear to be the most persistent. They involve and depend on domain (geospatial) knowledge, entailing thus research topics of long term objective but also of large value. The third obstacle does not solely refer to geospatial knowledge, and could generally advance independently. Nevertheless, geospatial knowledge may present additional requirements which general purpose conceptual instruments are unable to deal with. Finally, explication and implementation issues have been purposely left out; for they are easier to be dealt with once theoretical-conceptual issues have been clarified.

Given the above framework, in the rest of the section a number of important and promising research questions are put forward. All of them, theoretical approaches as well as formal instruments, could fall in or contribute to a main objective, that is, geo-ontologies, concepts, and semantics

as well as the development of a *corpus of geographic knowledge*. The list is only indicative and could not be closed. Some questions appear to be difficult to deal with rigorously in the foreseeable future. In these cases, there is no other way than adopting a workable solution, and working constantly towards its improvement.

3 Corpus of Geographic Knowledge

Ontological research in GIScience, as in many other disciplines, has grown substantially (see review by Agarwal 2005), yet there is still an open field of challenge. Especially, as the emphasis is put on deeper (semantic) information (Kavouras and Kokla 2008), it becomes necessary to move from the explication level higher to the ontological level. In order to do so, it is essential to build an ontological corpus of geographic knowledge.

3.1 Essence in Geographic Concepts

Concept formation is a central issue in all disciplines, GIScience not excluded. Metaphysics, Cognitive Science, and Epistemology have the potential to contribute to this objective. It is highly unlikely however that an overall encompassing theory will ever be in the position to describe different, partial, yet complimentary perspectives of geographic reality. It is nevertheless important that the essence of geographic concepts is approached at a pragmatic level.

Since concepts are the basic conceptual units of an ontology, a set of principal semantic dimensions can be set to which each concept can be projected. This set of semantic dimensions is similar to what has become known as a *semantic reference system* (Kuhn 2003). The properties, relations, and other signified elements (Kavouras and Kokla 2008) in describing a concept can prove very useful, and can be based on the universal distinctions of *top-level ontologies* (Sowa 2000). The determination of such a universal set of essential properties/relations for geoconcepts would make a great research accomplishment of immediate practical use. Several prominent initiatives for the development of a top-level ontology exist, such as IEEE Standard Upper Ontology, Upper Cyc Ontology, Basic Formal Ontology, etc. However, there is not yet an established top-level ontology which would be used as a common basis for the evolvement and association of more specialized domain-dependent ontologies. Furthermore, Agarwal (2005) argues that it is debatable whether a unified approach to geospatial ontology may exist, taking into account the interdisciplinary

nature of geographic information research, and the different conceptualizations and terminology used for the same geospatial concepts.

3.2 Principal Ontologies

This research direction focuses on the development of a set of interoperable ontologies covering the most basic/common (*principal*) concepts and relations of the geospatial domain, also called *the core*. Despite the different views on the general issue of core ontologies (Borgo and Gangemi 2004), they constitute a vital area of active research. The core should probably rely on the general structure of *top-level ontologies* and include different *views* of the domain. The existing amount of work, which otherwise could be considered significant, has not reached this stage yet. Exemplary principal ontologies, without any particular order, cover the following: *views and context, location, boundaries, spatiotemporal relations, essential properties, affordances, spatial operators, change, uncertainty, etc.* The benefit of establishing such a corpus, even an imperfect one, is manifold:

- First, it would help geospatial research move forwards, without running constantly in circles.
- Secondly, it would provide a solid basis/reference for comparing different views and resolve conflicts.
- Thirdly, it could systematise, compare and harmonise GIScience education and curricula.

3.3 Light vs. Specialised Ontologies

Traditionally, the analysis and processing of geographic information were of concern to the experts. Experts, appreciating the intricacy of the world phenomena, naturally attempt to deal with the problems as they are taught to, that is, in a complex scientific way. On the other hand, nowadays in the internet era more and more non-specialists attempt to deal with geospace and its information in a much simpler way based on “common sense”. The ontologies employed by this type of users are much lighter than those used by specialists. Research-wise, it is important to semantically strengthen these approaches. In this area, worth mentioning is the work of *Semantic Geospatial Web* (Egenhofer 2002).

The problem of supporting cognitive, common-sense, naïve approaches may be quite difficult (and thus challenging), because “common-sense” may prove to be not that “common”.

3.4 Conceptual Instruments

Besides the number of advanced conceptual structures, formalisms, and instruments developed, and their application to various fields, including GIScience, neither their suitability, nor their applicability have been sufficiently studied so far (Kavouras and Kokla 2008). Major instruments have proven to be *Formal Concept Analysis - FCA* (Ganter and Wille 1999), *Conceptual Graphs – CG* (Sowa 2000), *Chanel Theory and Information Flow - IF* (Barwise and Seligman 1997), and others. Another important challenge here is the effective synthesis of such advanced structures in frameworks. Along these lines, the Information Flow Framework (IFF) (Kent 2004) is an attempt to unify IF with FCA, while, Wille (1997) has attempted to join FCA and CGs.

3.5 Geonoemata from Natural Language

Though research on linguistic aspects of geographic space has been around for sometime (Mark and Frank 1991), traditional geospatial research has paid only limited attention to geographic knowledge and meaning elicited from natural language. Geonoemata can be systematically extracted from linguistic descriptions and formalised. This is a very exciting research direction, not only because it can explore a wealth of knowledge such as textual descriptions, but also because it has the potential to establish a “natural” and semantically rich communication between humans and machines in the future.

Advances in natural language processing/understanding (NLP/NLU), as far as geospatial concepts are involved, are absolutely necessary (Kokla 2008). An analogy with more traditional geospatial “languages” such as cartography may also prove to be useful. Some other promising NLP techniques are (Kavouras and Kokla 2008): (a) text summarization (Mani and Maybury 1999), and (b) controlled “natural” languages (Sowa 2004; Kavouras and Kontaxaki 2005).

4 Conclusions

GIScience needs not just to endure but also to establish a core identity as a real science. Developing a corpus of geographic knowledge is of utmost importance. Towards this objective, probably the most critical questions are (a) how concepts are formed, (b) how representations fulfil their role, and (c) how geonoemata are elicited. In this context, a short indicative list of challenging research areas has been introduced.

The relation their progress may have to other research areas is obvious. Such progress affects work on semantic similarity and interoperability. It also affects the way we deal with vague or imperfect geographic knowledge. The development of universal methods and tools, almost regardless to when dealing with spatial, thematic or temporal information, or when dealing with visual, cartographic, or textual descriptions of space, heavily depends on the successful tackling of the issues of the previous paragraph.

The fact that some problems are not likely to find a final solution in the near future, for they involve amongst others some deep yet unanswered philosophical issues, should not inhibit us from pursuing at least a pragmatic approach. After all, the endless effort to develop a multi-faceted yet coherent representation of geographic reality shall determine conflicts between the different geonoemata we form, and significantly improve our communication and our (admittedly partial) understanding of this world.

References

- Agarwal P (2005) Ontological considerations in GIScience. *International Journal of Geographical Information Science (IJGIS)* 19(5): 501–535
- Barwise KJ, Seligman J (1997) *Information Flow: the Logic of Distributed Systems*, Cambridge Transactions in Theoretical Computer Science, Cambridge University Press, Cambridge (UK)
- Borgo S, Gangemi A (eds) (2004) Preface. In: *Core Ontologies in Ontology Engineering 2004 - (Un)Successful cases and best practices for ontology engineering: reusing well-founded ontologies for domain content specification*, Proceedings of the EKAW*04 Workshop on Core Ontologies in Ontology Engineering, Northamptonshire (UK)
- Egenhofer JM (2002) Toward the Semantic Geospatial Web. In: *Proceedings of the Tenth ACM International Symposium on Advances in Geographic Information Systems*, McLean, Virginia
- Frank AU (2000) Spatial Communication with Maps: Defining the Correctness of Maps Using a Multi-Agent Simulation. In: Freksa C, Brauer W, Habel C, Wender KF (eds) *Spatial Cognition II (International Workshop on Maps and Diagrammatical Representations of the Environment*, Hamburg, August 1999, *Lecture Notes in Artificial Intelligence*, Vol. 1849, Springer, Berlin Heidelberg New York, pp 80–99
- Frege G (1892) On Concept and Object, originally published as *Über Begriff und Gegenstand*. *Vierteljahresschrift für wissenschaftliche Philosophie* 16: 192–205, translated in Geach, Black (1952), pp 42–55
- Ganter B, Wille R (1999) *Formal Concept Analysis - Mathematical Foundations*. Springer, Berlin Heidelberg New York
- Kavouras M, Kokla M (2008) *Theories of Geographic Concepts: Ontological Approaches to Semantic Integration*, CRC Press, Taylor & Francis Group, Boca Raton, FL, USA

-
- Kavouras M, Kontaxaki S (2005) Spatial knowledge extraction from geographical databases: an approach based on the controlled English query language Geo-Q and conceptual graphs. In: Proceedings of GIS Planet 2005, Estoril, Portugal
- Kent RE (2004) The IFF Foundation for ontological knowledge organization. In: Williamson NJ, Beghtol C (eds) Knowledge Organization and Classification in International Information Retrieval, Cataloging and Classification Quarterly, The Haworth Press Inc., Binghamton, New York
- Kokla M (2008) GeoNLP: A tool for the extraction of semantic information from definitions. In: The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. XXXVII, Part B2, Beijing 2008, pp 691–696
- Korzybski A (1933) Science and sanity: An Introduction to Non-Aristotelian Systems and General Semantics.
- Kuhn W (2003) Semantic Reference Systems. *International Journal of Geographical Information Science* 17(5): 405–409
- Mani I, Maybury MT (1999) *Advances in Automatic Text Summarization*, MIT Press
- Mark DM, Frank AU (eds) (1991) *Cognitive and Linguistic Aspects of Geographic Space*, Proceedings of the NATO Advanced Study Institute, Series: NATO Science Series D, Vol. 63, Kluwer, Dordrecht
- Ogden CK, Richards IA (1923) *The Meaning of Meaning*, Routledge & Kegan Paul, London
- Sowa JF (2000) *Knowledge Representation: Logical, Philosophical, and Computational Foundations*, Brooks Cole Publishing Co., Pacific Grove, CA
- Sowa JF (2004) Common Logic Controlled English, <http://www.jfsowa.com/clce/specs.htm>, last date accessed 01.2009
- Wille R (1997) Conceptual graphs and formal concept analysis. In: Lukose D, Delugach HS, Keeler M, Searle L, Sowa JF (eds) *Proceedings of International Conference on Conceptual Structures ICCS'97*, Lecture Notes in Computer Science, vol. 1257, Springer, Berlin Heidelberg New York, pp 290–303



<http://www.springer.com/978-3-540-88243-5>

Research Trends in Geographic Information Science

Navratil, G. (Ed.)

2009, XII, 283 p., Hardcover

ISBN: 978-3-540-88243-5