

Core Geographical Concepts: Case Finnish Geo-Ontology

Riikka Henriksson
Helsinki University of Technology
P.O. Box 1200
FIN-02015
+358 40 563 45 16
riikka.henriksson@tkk.fi

Tomi Kauppinen
Helsinki University of Technology and
University of Helsinki
P.O. Box 5500
FIN-02015 TKK
+358 9 451 3355
tomi.kauppinen@tkk.fi

Eero Hyvönen
Helsinki University of Technology and
University of Helsinki
P.O. Box 5500
FIN-02015 TKK
+358 9 451 3362
eero.hyvonen@tkk.fi

ABSTRACT

In this paper we examine 1) the scope of geo-ontologies used especially for the purposes of information retrieval on the Web, 2) the core geographical concepts and their mutual relations, and 3) the properties the concepts have. Furthermore, we present the Finnish geo-ontology (Suomalainen paikkaontologia, SUO) and discuss the theories and principles that have governed the development process, as well as the limitations and requirements the use of geographical dictionaries as an instance data source have imposed to the content and the structure of SUO.

Categories and Subject Descriptors

D.3.1 [Formal Definitions and Theory]: Semantics; H.3.3 [Content Analysis and Indexing]: Dictionaries

General Terms

Theory, Experimentation

Keywords

Geo-ontology, Place, Region, Hierarchy

1. INTRODUCTION

Due to rapidly growing availability of information about the Earth, and thus the increasing use of it in many application areas, new methods and techniques are needed in geographically referenced data discovery and retrieval. As pointed out in [8] the searches made by web users specify the queries more often with geographically referenced terms, such as place names. However, these simple string-matching keyword queries which are used in conventional search engines typically cause low precision and recall; place names are not unique, places may have variant names, names may change or even disappear throughout time. This has led to efforts to develop 1) ontologies that specify the concepts and their relationships in the domain of geographical information (e.g. in [8]), and 2) special spatially-aware search engines to find places or other resources, such as documents or photographs corresponding to the places that are referred to in a query (e.g. in [7, 9]).

A geographical ontology, which is denoted in this paper as geo-ontology, defines [3] concepts that correspond to 1) things from the physical and social world having a location on the surface of the Earth and 2) semantic and spatial relations (i.e. topology and mereology) between these things. Furthermore, these concepts are associated either with discrete geographic objects that have well defined boundaries or with continuous fields over space (Table 1). The data for populating a geo-ontology may be derived from a variety of sources, including geographical dictionaries (i.e.

gazetteer). A gazetteer contains entries which are composed of information such as a place name, coordinates, and a place type assigned to a place [6].

Table 1. Basic components of geo-ontology according to [17].

	Physical World	Social World
Discrete object	e.g. lake and mountain	e.g. cemetery and nation
Continuous field	e.g. desert and drainage basin	e.g. newspaper's circulation area and farmstead

The concepts of 'place' and 'region' are widely used in geo-ontology research (e.g. in [1, 4, 8]), but their meaning and mutual relationships are likely to be weakly defined; for example in [8] 'place' is only assumed to refer to any geographical phenomenon with the condition that the phenomenon has a given name or a literal description. Furthermore, the process of building geo-ontologies for the uses of spatially-aware search engines is studied in [4]. This paper proposes that the geo-ontology building process begins with the definition of underlying conceptual model, which then serves as a base for the geo-ontology. However, this conceptual model lacks consideration of the meaning of the geographical concepts, and hence can miss some relevant information.

The point made here is to emphasize the fact that researchers working in the area of geo-ontology construction should make clear the theories, thoughts and ideas behind the geographical concepts. In this paper we examine a set of core geographical concepts, their mutual relations and properties that are then used in the Finnish geo-ontology, called Suomalainen paikkaontologia (SUO). Furthermore, we will address the middle-out approach to develop geo-ontologies, and the limitations as well as the requirements the approach places for it.

2. SCOPING THE GEO-ONTOLOGY

The main traditions of geography [14] are a spatial tradition, an area study tradition, a man-land tradition and an earth science tradition. This framework gives the starting point to determine the work areas discussed in [19] that corresponds to the contents, and also to the core concepts to be used in geo-ontologies.

2.1 Spatial tradition

The spatial tradition has emphasized the use of quantitative methods in geographic research; places are studied in terms of their spatial attributes, namely location, position and geometry [14]. Location is fundamental spatial attribute and can be

described in terms of site and situation. Site refers to the absolute location (i.e. position) of a place including the elements from natural and cultural phenomenon that occupy it, and is described as a set of coordinate values (e.g. latitude and longitude values) in a given coordinate reference system (CRS). Situation, on the other hand, refers to the relative location of place by describing its external relationships with other places, i.e. how place is connected or localized in relation to other place [15]. Finally, the basic geometry of points, lines and polygons serves as the foundation for examining the shape, unity (i.e. one's individuality, as the boundary separates it from surrounding environment), as well as the contact and separation [1, 5] of places. 'Place' is seen merely as a location of something without the consideration of the meanings it has for individually or socially.

The aspects of the situation of place are studied in geo-ontology research in terms of spatial relations, i.e. topological and mereological relations (e.g. in [2]). Topological reasoning can be obtained by assigning properties such as [12] disjoint, equals, contains, crosses, intersects, overlaps, touches and within between the geometric objects that represent places, while property is-part-of between places supports mereological reasoning [2]. Figure 1 presents some of the most fundamental geographical concepts that can be derived from the work area of spatial tradition.

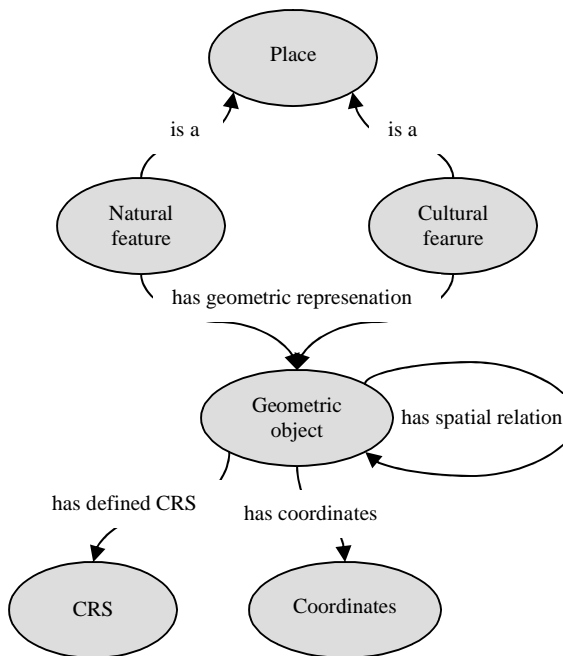


Figure 1. Some fundamental geographical concepts from the work area of spatial tradition. Real world natural and cultural phenomena are abstracted into discrete geographic features, which are represented as a geometric object. Each geometric object has a set of spatial relations to other objects and coordinates defined by some coordinate reference system.

2.2 Area studies tradition

In area studies tradition (also called as regional geography) [14] researchers have divided the world into smaller units based on their dominant features or characteristics. These units (i.e. region or area) are often subdivided into four static typologies [16], namely 1) formal (also called as uniform), 2) functional (also

called as nodal), 3) administrative, and 4) perceptual region. Formal regions are similar from some point of view, i.e. they share the same characteristics (e.g. land use patterns). Functional regions are defined by some spatial activity, which is most intense at the center of a region and decreases step by step going further away from the center (e.g. marketing areas). Administrative regions are identified usually based on their administrative or institutional boundaries (e.g. municipalities, nations, etc.), and could also be considered as a type of a formal region. Finally, perceptual regions exist only in the minds of people and do not necessarily possess any explicit feature or characteristics (e.g. the West and the Bible Belt).

The concept of 'region' is central in area study tradition. 'Region' [13] is seen as socially and culturally constituted spatial structure, a sort of socio-spatial unit which has a historical duration. Regions cannot be reduced to the history of an individual. The definition is problematic for the reason that spatial units are not always socially constructed; they can exist independently of all human cognitive acts (e.g. desert). For this reason we have adopted the concept of 'area' to refer to the non-social spatial units (figure 2). Since 'area' is not socially and culturally weighted concept, it cannot be further divided into administrative or perceptual units.

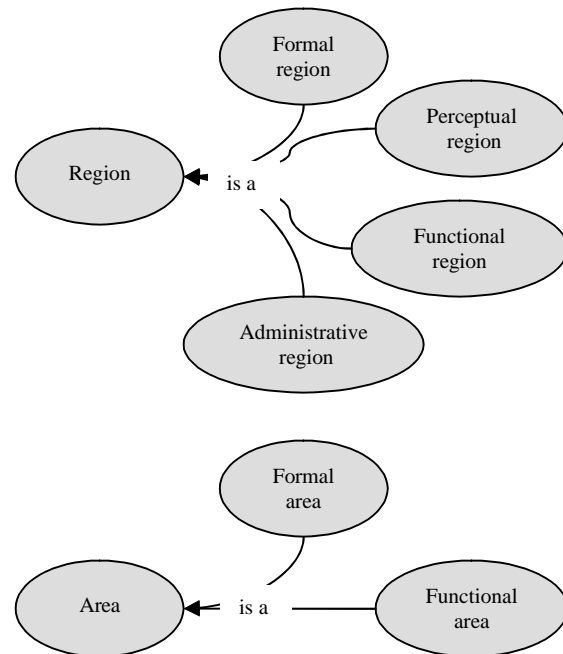


Figure 2. Some fundamental geographical concepts from the work area of area studies tradition. Spatial units are divided into two main concepts, namely region and area, which are further subdivided into more specific concepts based on their dominant characteristics.

2.3 Man-land tradition

A man-land tradition [14] entails a focus upon the relationships and interactions between societies and natural environments, i.e. how people activities are affected and controlled by the physical environment - and nowadays also vice versa. Concepts of our interest are those describing some specific land use pattern (e.g.

agricultural or residential land) - i.e. a region, which is uniform from the perspective of some specific type of a land use (figure 3).

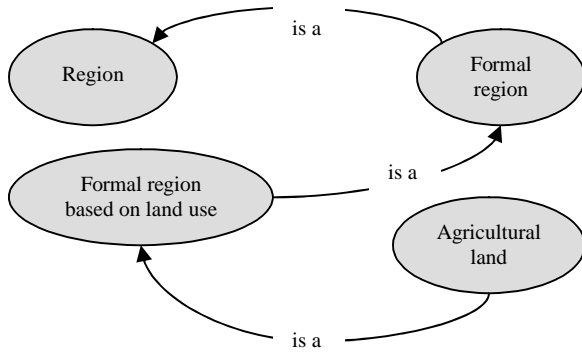


Figure 3. Work area of man-land tradition. Concept 'Formal region based on land use' is a group concept, under which different concepts representing the land use classes (e.g. agricultural land) can be gathered.

2.4 Earth science tradition

The fourth tradition, the earth science tradition [14], lays the focus fully on the physical environment around us; the waters of the earth, landforms, vegetation, soils, topography, etc. Each one of the topic categories can be positioned either under the concept of 'natural feature', 'functional area' or 'formal area' (Figure 4).

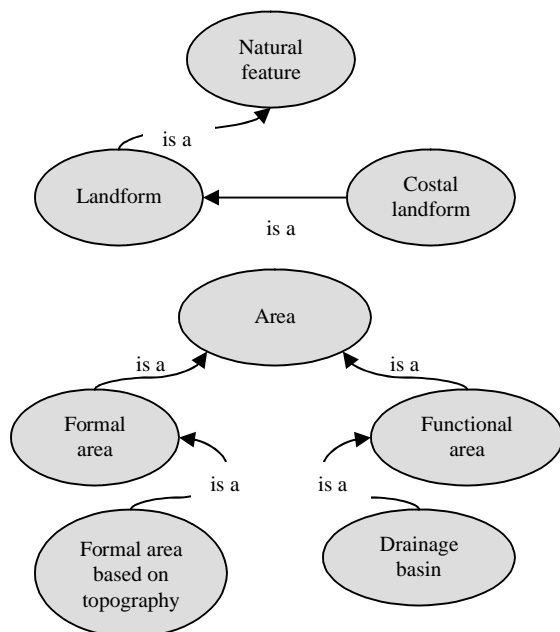


Figure 4. Concepts from the work area of earth science tradition can be placed under the concepts of 'natural feature', 'formal area' or 'functional area'.

2.5 The higher and lower level concepts

People live, act and orient in a world that is differentiated into places. The concept of 'place' has been the subject of interest and debate in geography for decades and it is perhaps the most multilayered and multipurpose words in current geographical language [13, 15].

There exist two main perspectives to study places [15]: 1) from the quantitative perspective (i.e. spatial tradition), where 'place' is seen as a location, a spatial unit within a hierarchy of units, and 2) from the phenomenological perspective, where 'place' is seen as a meaningful center of people immediate experiences of the world. The former studies places from the outside. Places are seen from the distance, as settings for physical objects and social actions. Individuals' values or feelings towards a place are not considered (called as outsidersness). On the latter perspective, on the other hand, places are experienced in one's lived-space and mind - this is called as insidersness: to be inside a place, and thus the meanings assigned to a place are seen to be constituted through individuals' attachment to particular portions of space

The distinction into insidersness/outsideensness is necessary when defining the concept of 'place' - one of the most fundamental geographical concepts. The foundation should be made clear from the beginning of geo-ontology construction, because the definition determines the bottom level concepts that are of interest. For example, a concept related to the 'insidersness domain' is 'home', while the 'outsidersness domain' concept is 'building'.

Using gazetteers as an instance data source has some crucial impacts on the 'place' definition: as centers of meaning [17], the number of places is enormous and no gazetteer can contain such amount of information. There are places such as a fireplace, workplace and the home. Neighborhood, beach, city and a lake are all places; a distinctive region is a place, and so is the continent. They all have a meaning to someone. Moreover, not all places are named. Naming a place is just to give it explicit recognition at the conscious level. This forces to look places objectively from the outside [15]; to separate individual from place in order to be able to study it selectively in terms of place's special attribute or activity. A place is seen merely as a thing having certain attributes, within some theory of location.

3. CASE: FINNISH GEO-ONTOLOGY SUO

Following the principles described in the previous section, SUO contains classes that describe spatial, regional, man-land and earth science aspects of geography. The upper and middle level concepts were derived from the scientific literature, while the bottom level concepts were derived from the Finnish Place Name Register (PNR) provided by the National Land Survey of Finland¹. PNR contains over 800 000 Finnish places (place names in Finnish, Swedish and Sami language, coordinates and the part-of hierarchy) classified in 52 different place types. Furthermore, we have also ontologized GEONet Names Server (GNS) [11] and mapped it with SUO. GNS contains 648 place types and millions of places all over the world excluding the USA.

Figure 5 illustrates the partial class hierarchy of SUO. The use of the gazetteers forced us to view the place-related concepts objectively from the outside. This set some requirements and limitations to the definition of 'place': it is described in SUO as a meaningful location of something and the properties assigned to it are only spatially weighted. Properties include the notions of mereology (e.g. is-part-of) and topology (e.g. overlaps and crosses). Classes describing the basic geometry and position are also part of SUO. The use of SUO is presented in [7].

¹ www.mml.fi



Figure 1. Partial class hierarchy of SUO. Classes that are derived form GNS have the prefix ‘GNS’. SUO was developed by using ontology editor Protégé²

4. CONCLUSIONS

This paper has summarized 1) the content of geo-ontologies and 2) the middle-out approach of geo-ontology development process. Geo-ontologies should contain classes that describe the spatial aspects of places (e.g. location), regional geography (e.g. administrative regions), patterns based on human interaction with nature (e.g. land use), and aspects related solely to the physical environment (e.g. landforms).

The bottom level concepts are usually derived from the gazetteers, which places some limitations and requirements on the geo-ontology. Gazetteers reduce the amount of instances since they do not contain places from the individual’s experiences (e.g. missing concepts such as ‘home’ and ‘workplace’). However, this gives more flexibility to define the top level concept ‘place’; it can refer both the discrete natural and man-made features, as well as the continuous natural and social phenomena.

5. ACKNOWLEDGMENTS

The research was part of the National Semantic Web Ontology Project in Finland³ (FinnONTO) 2003-2007 funded mainly by the Finnish Funding Agency for Technology and Innovation (Tekes).

² <http://protege.stanford.edu/>

³ <https://www.seco.tkk.fi/projects/finnonto/>

6. REFERENCES

- [1] Arpinar, B., Sheth, A., Ramakrishnan, C., Usery, E. L., Azami, M. and M. Kwan (2006). Geospatial Ontology Development and Semantic Analytic. *Transactions in GIS* 10 (4), pp. 551-575.
- [2] Casati, R., Smith, B., and A.C. Varzi (1998). Ontological Tools for Geographic Representation. In Guarino, N. (ed.), *Formal Ontology in Information Systems*, pp. 77-85.
- [3] Fonseca, F., Camara, G. and A. Monteiro (2006). A Framework to Measure The Interoperability of Geo-ontologies. *Spatial Cognition and Computation*, 6(4), 307-29
- [4] Fu, G., Jones, C.B., and A. Abdelmoty (2005). Building a Geographical Ontology for Intelligent Spatial Search on the Web. *Databases and Applications*, pp. 167-172.
- [5] Gatrell, J. D., Bierly, G. D. and J. J. Jensen (2005). *Research Design and Proposal Writing in Spatial Science*. Springer.
- [6] Hill, L.L., Frew J., and Q. Zheng (1999). Geographic Names: The Implementation of a Gazetteer in a Georeferenced Digital Library. *D-Lib Magazine* 5(1).
- [7] Hyvönen, E., Lindroos, R., T. Kauppinen and R. Henriksson (2007). An Ontology Service for Geographical Content. *ISCW 2007 + ASCW 2007*, 11-15 November, Busan, Korea.
- [8] Jones, C.B., Alani, H. and D. Tudhope (2001). Geographical Information Retrieval with Ontologies of Place. In *Proceedings of International Conference on Spatial Information Theory: Foundations of Geographic Information Science*, pp. 322-335. Springer-Verlag, 2001.
- [9] Jones, C.B., Purves, R., Ruas, A., Sanderson, M., Sester, M., van Kreveld M. and R. Weibel (2002). Spatial Information Retrieval and Geographic Ontologies An Overview of the Spirit Project. In *Proceedings of the 25th annual international ACM SIGIR conference on Research and development in information retrieval*, Tampere, Finland. pp. 387-388. ACM Press.
- [10] National Geospatial-Intelligence Agency (2008). *GEONet Names Server (GNS)*.
- [11] OGC (2005). *Simple Feature Access - Part 1. Common Architecture. Version 1.0.0, OpenGIS® Implementation Specification*. OpenGIS Project Document: OGC 05-126
- [12] Paasi, A. (2002). Place and region: regional words and words. *Progress in Human Geography* 26(6), pp. 802-811.
- [13] Pattison, W.D. (1964). Four Traditions of Geography. *Journal of Geography* 63 (5), pp. 211-216.
- [14] Relph, E. (1976). *Place and placeness*. London, UK: Pilon.
- [15] Salter, C.L. and J.J. Hobbs (2000). *Essentials of World Regional Geography*. Pacific Grove, CA: Brooks/Cole-Thomson.
- [16] Tuan, Y. (1975). Place: An Experiential Perspective. *Geographical Review* 65, pp. 151-165.
- [17] Ushold M. and M. Gruninger (1996). *Ontologies: Principles, Methods and Applications*. *Knowledge Engineering Review* 11(2).

