Spatial Relations for Semantic Similarity Measurement

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

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

Abstract. Measuring semantic similarity among concepts is the core method for assessing the degree of semantic interoperability within and between ontologies. In this paper, we propose to extend current semantic similarity measures by accounting for the spatial relations between different geospatial concepts. Such integration of spatial relations, in particular topologic and metric relations, leads to an enhanced accuracy of semantic similarity measurements. For the formal treatment of similarity the theory of conceptual vector spaces—sets of quality dimensions with a geometric or topologic structure for one or more domains—is utilized. These spaces allow for the measurement of semantic distances between concepts. A case study from the geospatial domain using Ordnance Survey's MasterMap is used to demonstrate the usefulness and plausibility of the approach.

1 Introduction

Successful communication of concepts depends on a common understanding between human beings and computer systems exchanging such information. In order to achieve a sufficient degree of semantic interoperability it is necessary to determine the semantic similarity between these concepts. Various approaches to measure semantic similarity between concepts exist and often such calculations of semantic distances are based on taxonomic and partonomic relations. When determining semantic similarity between geospatial concepts it is important to account for their spatial relations in the calculation process. All geospatial objects have a position in space with regard to some spatial reference system and therefore a spatial relation to each other. Spatial relations are also central characteristics on the conceptual level. In this paper, we present an approach of integrating spatial relations into semantic similarity measurements between different geospatial concepts. Such integration improves the quality of the measurements by enhancing the accuracy of their results.

For the formal representation of concepts and the calculation of their semantic similarities we utilize Gärdenfors' idea of a *conceptual space*—a set of quality dimensions within a geometric structure [1]. Such a representation rests on the foundation of cognitive semantics [2], asserting that meanings are mental entities, i.e. mappings from expressions to conceptual structures, which themselves refer to the real world. They therefore allow us to account for the fact that different people have different conceptualizations of the world.

A case study, in which a customer specifies query concepts based on a shared vocabulary and wants to extract similar concepts from the database of a mapping agency, is used to demonstrate the importance of accounting for spatial relations in a real scenario. It is shown that without the inclusion of spatial relations, the customer is presented with answers which do not fully match her requirements.

2 Related Work

Most knowledge representations use a definitional structure of concepts to describe their semantics: Concepts are specified by necessary and sufficient conditions for something to be its extension. The nature of these conditions is distinct: Properties (features, dimensions) describe the characteristics of concepts, while semantic relations describe concepts through their relationships to other concepts.

The following section describes the formalization of natural-language spatial relations used for the description of geo-concepts. In section 2.2 we give an overview of semantic similarity measures and evaluate how they include relations. The final section describes conceptual spaces, the representational model used in this paper.

2.1 Formalization of Natural-Language Spatial Relations

Describing a concept with relations closely resembles the human way of structuring knowledge: According to the associationist theory humans memorize knowledge by building relations between concepts. The importance of spatial relations arises from the geographic reference of most of our data. All geo-objects have a position in a spatial reference system and each pair of geo-objects is spatially related. The same goes for the conceptual level: due to their functional dependence the geo-concept 'floodplain' is always situated near a water body. We consider spatial relations to be fundamental parts of the semantic description of geo-data.

While formal spatial relations—topologic [3], distance [4] and direction relations [5]—have well defined semantics, natural-language spatial relations have more complex semantics and often imply more than one type of formal spatial relation. People are more familiar with using spatial terms in their natural languages, but systems use definitions based on a computational model for spatial relations. To bridge this gap Shariff et al. developed a model defining the geometry of spatial natural-language relations following the premise *topology matters, metric refines* [6].

The *computational model for spatial relations* [7, 8] consists of two layers: first it captures the topology between lines and regions based on the 9-Intersection model. The second layer analyzes the topologic configuration according to a set of metric properties: splitting, closeness and approximate alongness.

Splitting determines the way a region is divided by a line and vice versa. The intersection of the interior, exterior or boundary of a line and a region is one- or two-dimensional. In the 1-D case the length of the intersection is measured, in the 2-D case the size of the area. To normalize length and area, they are divided either by the region's area or the length of the line or the region's boundary. *Closeness* describes the distance of a region's boundary to the disjoint parts of the line. It distinguishes between Inner/Outer Closeness and Inner/Outer Nearness. *Approximate alongness* is a combination of the closeness measures and the splitting ratios: it assesses the length of the section where the line's interior runs parallel to the region's boundary.

To capture the semantics of geo-objects with spatial relations it is important to use natural-language terms, because they are plausible for humans. Equally important is to have an unambiguous, formal interpretation of these natural language terms. The computational model by Shariff et al. provides a set of natural-language spatial relations with a formalization verified by human subject tests [8]. In our investigation we use a subset of those natural-language terms, which can be applied within the case study. People's choice of spatial relations to describe two objects differs depending on the meaning of objects, their function, shape and scale. We consider only hydrological geo-objects within a large-scale topographic map for defining spatial relations and do not take into account specific semantics of relations depending on the object's meaning, function, or shape.

2.2 Semantic Similarity Measurement

Geometric representations model objects within a multidimensional space: Objects are described on dimensions spanning a vector space [9]. Dimensions are separable into attribute-value pairs: terms that can be evaluated to a value with different, mutually exclusive levels, e.g. the flow speed of a river is either slow, middle or fast. Geometric representations evolved from multidimensional scaling (MDS) [10, 11]: while MDS starts from similarity judgments and determines the underlying dimensions, geometric models represent objects on pre-known dimensions and convert their spatial distance, interpreted as a semantic distance, to a similarity value. Similarity measures in geometric models are metric, though various extensions to account for non-metric properties exist (e.g. Distance Density Model [12], Relative Prominence Model [13]). Geometric representations use only properties for semantic description. It is not possible to describe relationships between objects or concepts.

Feature representations model concepts as sets of features. Features are unary predicates, e.g. a concept 'water body' has the feature 'flowing' or ¬'flowing'. The Feature Matching Model proposed by Tversky [14] is a nonmetric similarity measure comparing two concepts as two sets of features: common features increase and distinct features decrease similarity. It was also applied in other similarity measures such as the Matching-Distance Similarity Measure [15, 16]. In feature representations the description of concepts is limited to atomic features. Relations between objects cannot be represented in a structured way: some approaches construct compound features, but compound features do not allow for structured comparison, e.g. no similarity would be detected between 'nearRiver' and 'veryNearRiver'.

*Network representation*s describe concepts by their relation to other concepts in semantic nets. Relations are n-ary predicates with concepts as arguments. Shortest path algorithms such as Distance [17] are used as a similarity measure. The representation of relations is the strength of network models, but most similarity measures restrict the type of relations to taxonomic and partonomic relations.

Alignment models such as Goldstone's SIAM model [18] describe concepts by features and relations. For similarity measurement they additionally take into account whether features/relations describe corresponding parts: aligned matches increase the similarity more than non-aligned ones. SIAM measures similarity between spatial scenes. Applying SIAM for concept similarity is difficult: due to the different granularity of concept descriptions, an alignment of elements is often not possible. Therefore this model does not return good results when comparing geo-concepts.

This paper extends geometric models to represent spatial relations on dimensions. Geometric models are chosen, because the dimensional structure allows for modelling the degree of relations. Conceptual vector spaces provide a solid mathematical basis for representing information at the conceptual level.

2.3 Conceptual Vector Spaces as Geometric Model

The notion of a *conceptual space* was introduced by Peter Gärdenfors as a framework for representing information at the conceptual level [1]. He argued that cognitive science needs this intermediate level in addition to the symbolic and the subconceptual level. Conceptual spaces can be utilized for knowledge representation and sharing and support the paradigm that concepts are dynamical systems. A conceptual space is a set of quality dimensions with a geometric or topologic structure for one or more domains. A domain is represented through a set of integral dimensions, which are distinguishable from all other dimensions. For example, the colour domain is formed through the dimensions hue, saturation and brightness. Concepts are modelled as n-dimensional regions and every object is represented as a point in a conceptual space. This allows for expressing the similarity between two objects as their spatial distance.

In [19], a methodology to formalize conceptual spaces as vector spaces was presented. Formally, a conceptual vector space is defined as $\mathbf{C}^n = \{(c_1, c_2, \ldots, c_n) \mid c_i \in \mathbf{C}\}$ where the c_i are the quality dimensions. A quality dimension can also represent a whole domain, then $c_j = \mathbf{D}^n = \{(d_1, d_2, \ldots, d_n) \mid d_k \in \mathbf{D}\}$. The fact that vector spaces have a metric allows for the calculation of distances between points in the space. In order to calculate these so-called *semantic distances* between instances and concepts it is required that all quality dimensions are represented in the same relative unit of measurement. This is ensured by calculating the percent ranks for these values [20].

3 Case Study

A customer of the British national mapping agency Ordnance Survey, such as the Environment Agency of England and Wales, wants to set up a flood warning system [21]. An overview of existing flooding areas is needed to analyze the current flood defence situation in Great Britain.

OS MasterMap contains geographic and topographic information on every landscape feature—buildings, roads, plants, fields and water bodies. It also contains information on areas used for flooding, but these are not explicitly designated as such [22]. While labels such as 'floodplain' allude to something used for flooding, geo-objects named 'watermeadow', 'carse' or 'haugh' are identified as flooding areas by their semantic description only. The semantics of all geo-objects within OS MasterMap are described as concepts in an ontology using an accurately defined shared vocabulary. The shared vocabulary does not contain concept labels such as 'flooding area', but only terms to describe properties, e.g. 'waterlogged' and relations between concepts, e.g. 'flooding area is next to river'.

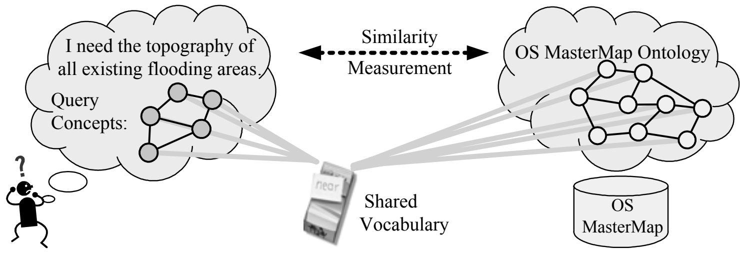

Fig. 1. Visualization of the case study.

The customer searches for topographic information about rivers and flooding areas (figure 1). The semantics of the required information is defined within query concepts using the same shared vocabulary as in the OS MasterMap ontology[1]. To retrieve data according to their relevance, a semantic similarity measure is used to match the query concepts with OS MasterMap concepts.

Table 1. Spatial relations in the shared vocabulary.

spatial relation	examples
along	flooding areas lie along a river bank
connected to	rivers are connected to a river, a lake or the sea
in	rivers lay in a river basin
end at	rivers end at river mouths
end in	rivers end in the sea
end just inside	ship ramps end just inside rivers
end near	port feeders end near the sea
near / very near	flooding areas are near / very near a river

The shared vocabulary was developed for this case study and contains only expressions necessary for this particular similarity measurement. It does not raise the claim of completeness nor of being a representative set of spatial relations for a geo-ontology. From the set of natural-language spatial relations formalized by Shariff et al. we identified a subset of those relations being relevant for the case study (table 1).

Many geo-concepts such as 'flooding area' and 'river' can be well described by their relation to other geo-concepts. The customer uses the spatial relations listed above and a set of dimensions to specify the query concept. The complete shared vocabulary and measurements can be found at http://ifgi.uni-muenster.de/~eidueidu/er05.zip.

[1] The customer can use natural-language spatial relations, while semantics in the OS MasterMap ontology is based on the formal definition of such relations.

4 Formalization of Measurement

For the described case study the similarity of a set of OS MasterMap concepts to the query concepts 'flooding area' and 'river' are measured. The occurring properties and spatial relations are formalized as dimensions of a conceptual vector space.

Concepts can be described by their properties or relations to other concepts. In a conceptual space, properties are represented by dimensions or domains. A property can be formalized by a dimension with a value— *dimension(concept) = value*. The value is defined within a specific range [1]. Figure 2 gives an example for the dimension 'waterlogged' with values for two concepts.

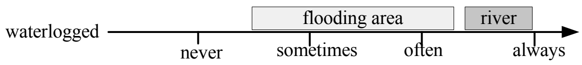

Fig. 2. Representation of dimension 'waterlogged' for concept 'flooding area' and 'river'.

To model relations[2] as dimensions, the dimension need not only represent one concept and its values, but one concept and its values with regard to a second concept. We propose to represent relations between two concepts by introducing dimensions depending on the first argument of the relation. The second argument is represented with its value on the dimension in the conceptual space.

Table 2. Relations are represented on a Boolean or ordinal dimension with numerical values.

relation	scale	original values	numerical values
along river	Boolean	yes	1
		not specified	0
nearness	Boolean	yes	1
		not specified	0
	ordinal	low nearness	0
		near	1
		very near	2
		not specified	-

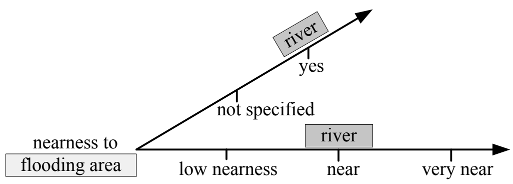

Fig. 3. Modelling relations as dimensions on Boolean and ordinal scale.

We distinguish two types of relations (table 2): Boolean relations do not have any degree of existence, e.g. the relation 'along' is either applicable to two concepts or not. They are represented by one Boolean dimension. Other relations have different degrees: the 'nearness' relation can state that two objects are very near, near, or

[2] In this case study we limit our investigation to binary relations.

somewhere around (low nearness). These relations require a domain consisting of two dimensions—a Boolean and an ordinal: If the relation holds, the Boolean dimension has the value 'yes' and the ordinal dimension is assigned a value specifying the degree (figure 3). If the relation is not applicable the Boolean relation has the value 'not specified' and the ordinal dimension has no value.

According to the rank-order rule ordinal values for the degree of relations are transformed into ordered numerical values[3] [23]. Boolean dimensions are represented by the values {1,0}. The value 0 does not state that the relation does not hold, but that it was not specified by the user or in the system[4].

For the similarity measurement we use the Euclidian and the city block metric in order to calculate distances [24]. The results of the case study demonstrate that our approach is robust and provides good results independent of the metric applied.

Here, we focus on the measurement of semantic distances on the conceptual level. Concepts are convex regions in the conceptual space. Since the Euclidian and the city block distances are between two points rather than between two regions, all concepts are approximated by their prototypes [25], i.e. representing the average value for each interval on each dimension.

5 Results of the Case Study

For the similarity calculation we compare each relation of the query concept separately to the relations of the concepts in the data source. Tables 3 and 4 show the results for the similarity measurements to the query concepts 'flooding area' and 'river' with and without spatial relations. The semantic distance values are calculated based on the differences of the standardized values for each dimension[5]. The final values are normalized by the number of dimensions used in the calculation. An Ordnance Survey expert divided the OS MasterMap concepts into three classes according to their similarities to the query concept: matching, similar (concepts must be modified to match) and non-matching.

Independent of the spatial relations the water bodies 'river', 'stream', 'channel' and 'canal' are considered as very different from the query concept 'flooding area', i.e. their semantic distances are large (table 3). The similarity measurement without spatial relations ranks 'lowland', 'meadow' and 'land' more similar to the query concept 'flooding area' than 'haugh'. Distances measured with spatial relations provide correct results. Since 'meadow' and 'land' do not necessarily lie near rivers such as flooding areas do, they are not typically used for flooding. 'Lowland' though, does not lie

---

[3] The different degrees of nearness result from different distances. The degrees of relation 'end' depend on the prepositions implying different distances, e.g. the distance between two concepts related via 'ends near' is greater than 'ends in'. The numerical values are applied according to the values for inner/outer closeness from the human subject test in [9].

[4] Boolean dimensions representing properties such as 'flowing' yes/no have also the values {1,0}, but here the value 0 explicitly negates the property. If a property is not applicable, this dimension of the conceptual space is not specified.

[5] Another possibility is using the z-transformation (as done in [19]), but this requires that the values for each dimension are normal [26].

explicitly along water bodies, but due to the fact that it is low and rivers typically flow through lowland, it is described by being near rivers and the sea. Therefore the semantic distance does not increase much when including spatial relations.

Table 3. Standardized semantic distances to query concept 'flooding area'.

		Euclidian metric		city block metric	
	OS Expert	without spat. rel.[6]	with spat. rel.	without spat. rel.	with spat. rel.
OSMM floodplain	match	59	40	38	20
OSMM river	no match	98	94	100	93
OSMM stream	no match	94	91	92	88
OSMM watermeadow	match	62	44	53	30
OSMM channel	no match	100	100	98	100
OSMM haugh	match	64	44	52	29
OSMM land	no match	62	79	49	69
OSMM meadow	no match	43	71	31	59
OSMM paddock	no match	66	82	54	74
OSMM lowland	similar	53	59	36	42
OSMM canal	no match	82	88	59	75
OSMM carse	match	53	38	38	22

Table 4. Standardized semantic distances to query concept 'river'.

		Euclidian metric		city block metric	
	OS Expert	without spat. rel.	with spat. rel.	without spat. rel.	with spat. rel.
OSMM floodplain	no match	39	44	64	67
OSMM river	match	9	6	20	17
OSMM stream	match	7	6	21	19
OSMM watermeadow	no match	43	46	68	70
OSMM channel	similar	16	17	27	25
OSMM haugh	no match	95	85	88	87
OSMM land	no match	29	46	56	59
OSMM meadow	no match	95	96	88	88
OSMM paddock	no match	100	100	100	99
OSMM lowland	no match	95	98	97	100
OSMM canal	similar	7	11	20	19
OSMM carse	no match	32	35	52	54

[6] The distance values are based on different numbers of dimensions. Adding new dimensions to the similarity measurement either increases the distance or it stays the same. To make the distances comparable, they are calculated relative to the number of dimensions and then scaled on a range of [0;100].

The similarity measures to the query concept 'river' are shown in table 4: The similarity measure ranks with and without relations the concepts 'river', 'stream', 'channel' and 'canal' correctly as the most similar concepts to query concept 'river'. But the inclusion of spatial relations changes their order: 'river' and 'stream' flow towards the sea and lay in a river valley, while 'canal' and 'channel' being artificial, man-made geo-objects, do not necessarily have these relations. With spatial relations 'river' and 'stream' are classified as more similar to the query concept than 'canal' and 'channel'.

A comparison of all distance values shows that the inclusion of spatial relations leads to more sensible values for every compared concept. The results of both metrics are good, though the city block metric shows the differences between matching and non-matching concepts more explicitly. This goes along with findings that the city-block metric is more adequate with separable dimensions (e.g. [24]). To reiterate, the results of the case study demonstrate that similarity measurements are more accurate and realistic when spatial relations are included for the calculation of semantic distances between geo-concepts.

6 Discussion

In the following, the assumptions and further requirements for this similarity measurement are evaluated and discussed.

Reducing Concepts to Prototypes. To measure distances between concepts, they are represented by their prototypes in the conceptual space [27]. For some concepts this may lead to a substantial information loss, e.g. generic concepts such as 'land' with broad intervals on each dimension are semantically narrowed down to single points.

Semantic Post-Processing. All concepts are described based on a common shared vocabulary. This vocabulary does not contain concept labels, but to specify the relations other concepts such as 'river valley' are needed. Since the shared vocabulary does not define the semantics of these, they are adjusted manually, e.g. query concept 'river' is described as 'contained within river valley'. This is aligned to 'contained within river basin' of the 'OSMM river'. The concepts 'river valley' and 'river basin' are considered the same for the semantic similarity measurement. Such manual alignment could be automated by using ontologies or thesauri.

Directed Similarity. The purpose of this similarity measurement is to find the most similar concepts to the query concept. We aim at measuring directed similarity from the point of view of the customer. Therefore the similarity values are calculated based on the dimensions used to describe the query concept. Other dimensions of OS MasterMap concepts do not have any effect on the similarity.

7 Conclusions and Future Work

This paper develops a way to include spatial relations between concepts for semantic similarity measurement within conceptual spaces. We model spatial relations as dimensions and show how they can be used in similarity measurement. A case study demonstrates how conceptual spaces extended by spatial relations lead to more

accurate retrieval results. Based on a shared vocabulary, a customer defines her data requirements by query concepts. Through the similarity measure the system identifies matching concepts within OS MasterMap, Britain's national topographic database.
The paper leads to different directions for future research:

1. Here we make the simplifying assumption that quality dimensions of a conceptual space are independent. This is often not true: In the case study several dimensions are used to describe the amount and time period when a concept is covered with water, e.g. 'fullOfWater' and 'waterlogged'. It will be necessary to investigate the covariances between dimensions and to account for these in the conceptual space representations. Human subject tests are a way to identify the quality dimensions for a concept and to infer their dependencies—see, for example, [24]—which would lead to non-orthogonal axes in the representation.
2. Concepts are typically convex regions in a conceptual space. As mentioned in the discussion, they are approximated by points to calculate the distances which entails information loss. To measure similarity between concepts a distance measure between regions must be developed. This can be done by calculating distances from each point of a concept to the reference concept. The resulting distance is an n-dimensional surface that can be transformed to a similarity value through its integral [28].
3. In the case study we focus on spatial relations formalized within the computational model by Shariff et al. This model is currently restricted to line-region relations. It seems possible to extend it for region-region relations and model these relations as dimensions in the same way as for the line-region relations.

Acknowledgements

We like to thank Tony Cohn and Hayley Mizen for taking part in our experiment. Moreover we thank the members of MUSIL and the 3 anonymous reviewers for their valuable comments.

References

1. Gärdenfors, P., Conceptual Spaces: The Geometry of Thought. 2000, Cambridge, MA: MIT Press. 317.
2. Lakoff, G., Cognitive Semantics, in Meaning and Mental Representations (Advances in Semiotics), U. Eco, M. Santambrogio, and P. Violi, Editors. 1988, Indiana University Press: Bloomington, Indianapolis, USA. p. 119-154.
3. Egenhofer, M.J. and R.D. Franzosa, Point-Set Topological Spatial Relations. International Journal of Geographical Information Systems, 1991. 5(2): p. 161-174.
4. Hernández, D., Relative Representation of Spatial Knowledge: The 2-C Case, in Cognitive and Linguistic Aspects of Geographic Space, D.M. Mark and A.U. Frank, Editors. 1991, Kluwer Academic Publishers: Dordrecht, The Netherlands. p. 373-385.
5. Freksa, C., Using Information Orientation for Qualitative Spatial Reasoning, in Theories and Methods of Spatio-Temporal Reasoning in Geographic Space, A.U. Frank, I. Campari, and U. Formentini, Editors. 1992, Springer: Berlin-Heidelberg-New York. p. 162-178.

6. Egenhofer, M.J. and D.M. Mark, Naive Geography, in Spatial Information Theory - A Theoretical Basis for GIS, International Conference COSIT, A.U. Frank and W. Kuhn, Editors. 1995, Springer: Berlin-Heidelberg-New York. p. 1-15.
7. Egenhofer, M.J. and A.R. Shariff, Metric Details for Natural-Language Spatial Relations. ACM Transactions on Information Systems, 1998. 16(4): p. 295-321.
8. Shariff, A.R., M.J. Egenhofer, and D.M. Mark, Natural-Language Spatial Relations Between Linear and Areal Objects: The Topology and Metric of English Language Terms. International Journal of Geographical Information Science, 1998. 12(3): p. 215-246.
9. Suppes, P., et al., Foundations of Measurement - Geometrical, Threshold, and Probabilistic Representations. Vol. 2. 1989, San Diego, California, USA: Academic Press, Inc. 493.
10. Shepard, R.N., The Analysis of Proximities: Multidimensional Scaling with an Unknown Distance Function. I. Psychometrika, 1962. 27(2): p. 125-140.
11. Nosofsky, R.M., Similarity scaling and cognitive process models. Annual Review of Psychology, 1992. 43: p. 25-53.
12. Krumhansl, C.L., Concerning the Applicability of Geometric Models to Similarity Data: The Interrelationship Between Similarity and Spatial Density. Psychological Review, 1978. 85(5): p. 445-463.
13. Johannesson, M., Modelling Asymmetric Similarity with Prominence. British Journal of Mathematical and Statistical Psychology, 2000. 53: p. 121-139.
14. Tversky, A., Features of Similarity. Psychological Review, 1977. 84(4): p. 327-352.
15. Rodríguez, A. and M. Egenhofer, Determining Semantic Similarity Among Entity Classes from Different Ontologies. IEEE Transactions on Knowledge and Data Engineering, 2003. 15(2): p. 442-456.
16. Rodríguez, A. and M.J. Egenhofer, Comparing Geospatial Entity Classes: An Asymmetric and Context-Dependent Similarity Measure. International Journal of Geographical Information Science, 2004. 18(3): p. 229-256.
17. Rada, R., et al., Development and application of a metric on semantic nets. IEEE Transactions on systems, man, and cybernetics, 1989. 19(1): p. 17-30.
18. Goldstone, R.L., Similarity, Interactive Activation, and Mapping. Journal of Experimental Psychology: Learning, Memory, and Cognition, 1994. 20(1): p. 3-28.
19. Raubal, M., Formalizing Conceptual Spaces, in Formal Ontology in Information Systems, Proceedings of the Third International Conference (FOIS 2004), A. Varzi and L. Vieu, Editors. 2004, IOS Press: Amsterdam, NL. p. 153-164.
20. Bortz, J., Statistik für Sozialwissenschaftler. Vol. 5th. 1999, Berlin, Heidelberg: Springer.
21. Schwering, A. Semantic Neighbourhoods for Spatial Relations (Extended Abstract). in Third International Conference on Geographic Information Science (GIScience). 2004. Maryland, USA: Regents of the University of California.
22. OrdnanceSurvey, OS MasterMap™ real-world object catalogue. 2001, Ordnance Survey.
23. Malczewski, J., GIS and Multicriteria Decision Analysis. 1999, New York: John Wiley. 392.
24. Johannesson, M., Geometric Models of Similarity. Lund University Cognitive Studies. Vol. 90. 2002, Lund, Sweden: Lund University. 171.
25. Rosch, E., Principles of Categorization, in Cognition and Categorization, E. Rosch and B. Lloyd, Editors. 1978, Lawrence Erlbaum Associates: Hillsdale, New Jersey. p. 27-48.
26. Devore, J. and R. Peck, Statistics - The Exploration and Analysis of Data. 4th ed. 2001, Pacific Grove, CA: Duxbury. 713.
27. Goldstone, R.L. and A. Kersten, Concepts and Categorization, in Comprehensive Handbook of Psychology, A.F. Healy and R.W. Proctor, Editors. 2003, Wiley: New Jersey. p. 599-621.
28. Schwering, A. and M. Raubal. Measuring Semantic Similarity between Geospatial Conceptual Regions. in 1st International Conference on GeoSpatial Semantics. 2005 forthcoming. Mexico City, Mexico: Springer.