Spatial Relations for Semantic Similarity Measurement

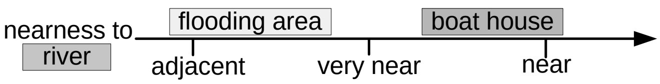
Angela Schwering[1, 2] and Martin Raubal[2]
[1] Ordnance Survey of Great Britain, UK
[2] Institute for Geoinformatics
University of Muenster, Germany
{angela.schwering; raubal}@uni-muenster.de

Measuring semantic similarity among concepts is the core method for assessing the degree of semantic interoperability. Most similarity measures utilise properties of concepts as semantic description and misjudge the importance of semantic relations for the human conceptualization and categorization process.

All geo-objects have a position in a spatial reference system and are therefore spatially related. We use the computational model by Shariff et al. for formalizing natural-language spatial relations [1, 2] and extend Gärdenfors' conceptual spaces [3]—a geometric model for representing information at a conceptual level and measuring its semantic similarity—to account for spatial relations between geospatial concepts.

A conceptual space is formed by a set of quality dimensions with a geometric or topologic structure. To represent concepts or objects in a conceptual space values are assigned to them for each dimension. 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 the distance between their points in the conceptual space.

The difficulty of modelling relations in conceptual space is that relations hold between two or more concepts, but dimensions describe a property of exactly one concept. Therefore we introduce a new kind of dimension where the dimension itself is dependent on one concept—e.g. nearness to river—and the other concepts—e.g. floodplain or boathouse—are modelled on this dimension.



We distinguish two types of *relational* dimensions: Boolean relations—e.g. river ends in sea—have only two values 'yes' and 'no' on the dimension. Ordinal relations such as in the figure above require two integral dimensions to provide correct similarity results: The first dimension is a Boolean (as for ends in) which indicates whether the relation holds or not. If the relation holds, the second and ordinal dimension is assigned a value. This value is specified according to the proposed values from the computational model.

The approach was evaluated in an information retrieval task with Ordnance Survey's MasterMap. Similarity measurement including spatial relations enhances the accuracy of the results.

1. Shariff, A., M. Egenhofer, D. Mark, Natural-Language Spatial Relations between Linear and Areal Objects. Int. Journal of Geogr. Inform. Science, 1998. 12(3) 215-246.
2. Egenhofer, M., A. Shariff, Metric Details for Natural-Language Spatial Relations. ACM Transactions on Inform. Systems, 1998. 16(4) 295-321.
3. Gärdenfors, P., Conceptual Spaces: The Geometry of Thought. 2000, Cambridge, MA: MIT Press. 317.