The geo-attribute space: a general space-time-property representation

O. Ahlqvist[1]
[1]The Ohio State University
Telephone: +1 614 247 7997
Fax: (international codes)
Email: Ahlqvist.1@osu.edu

1. Introduction

This paper describes the tight integration of a general theory for geographic representation (Goodchild et al., 2007) with a general representation for uncertain concept information (Ahlqvist, 2004). Research on geographic data representations and categorical measurement semantics suggest a potential for a close integration of these two geographic information components (Gahegan et al., 2003). This integration is only possible if representational approaches can be harmonized. In the following I will argue that the geo-atom concept suggested by Goodchild et al. (2007) has a parallel in the Uncertain Conceptual Space representation proposed by Ahlqvist (2004). By extending the geo-dipole concept also described by Goodchild et al. (2007), I enable an integration of the spatio-temporal and the conceptual dimension that forms a comprehensive geo-attribute space for geographic information representation.

2. Geo-atoms and uncertain conceptual spaces

Geo-atoms are defined as tuples $\langle \mathbf{x}, Z, z(\mathbf{x}) \rangle$ where **x** defines a point in 4-dimensional space-time, Z identifies a measured property, and $z(\mathbf{x})$ is the particular value of that property at the specified location (Goodchild et al., 2007). This representational primitive has the capability to serve as the foundation for both object based views through aggregation as well as field based views through a convolution operation using some type of discretization. In many cases **x** is only measured in two spatial dimensions at a fixed temporal location. Z is then added on as an extra measurement dimension in which only one value is possible at each **x** location. While the geo-atom, geo-object, and geo-field representation is mostly focused on capturing the dimensionality and discretization of space-time, the Uncertain Conceptual Spaces approach (Ahlqvist, 2004) captures the multi-dimensional character of many categorical geographic attributes.

Following a similar notation as before, a category or concept is also defined as a tuple $\langle \mathbf{s}, \mathbf{p}, \mathbf{p}(\mathbf{s}), \mathbf{w} \rangle$ where **s** defines a theoretically infinite collection of approximation spaces each relating to a property that can define a concept. In this conceptual space a collection of properties **p** are identified as salient to the definition of the category of interest. $\mathbf{p}(\mathbf{s})$ is the particular values of those properties given as rough fuzzy set definitions (Ahlqvist, Keukelaar, & Oukbir, 2003; Ahlqvist, 2004) and it constitutes the *uncertain conceptual space*. The inclusion of each property to the concept definition is determined by the salience weights **w**, which can either use 0 or 1 for exclusion/inclusion, or any value in between for graded importance to the concept definition. Further details on the cognitive foundations for this concept representation approach can also be found in Gärdenfors (2000). Similar to **x**, the conceptual space **s** has infinitely many possible locations, but the dimensionality of **s** is not restricted to 4 + 1 measurement dimension. From the infinite number of dimensions in **s** only a few, **p** as determined by **w**, are measured. Thus, there is nothing in principle that makes a 2-dimensional **x** space different from a 2-dimensional **p** space. Furthermore, Goodchild et al. (2007) point out that a discretization is a required and inherent property of any 1-4 dimensional space-time field. In the uncertain conceptual space this discretization is also inherent, and explicitly represented by the approximation space (U,θ) where a granularity is imposed on the underlying universe U by an equivalence relation θ (Ahlqvist et al., 2003).

Another key difference between the two representational approaches is that the conceptual space property values are defined as fuzzy number regions or *n*-dimensional volumes in the uncertain conceptual space. This concept representation format is similar to the bona fide geo-object theory, also proposed by Goodchild et al. (2007), but by using a vector $\mathbf{p}(\mathbf{s})$ of attributes $\{p(\mathbf{s})_1, p(\mathbf{s})_2, \ldots, p(\mathbf{s})_n\}$ defined as rough fuzzy numbers, we form an *n*-dimensional uncertain conceptual space that allow for explicit representation of uncertainty caused by measurement granularity and potential concept vagueness. The concepts of interest can be thought of as ***C*** irregularly shaped hyper-volumes each occupying a specific location specified by the set of rough fuzzy numbers on the *n* attributes with the important extension that attributes can be anything from a Stevens (1946) Ratio to a Nominal variable. Note that even though the representation can take multi-variate rough-fuzzy numbers it easily generalizes down to simple numerical measurements as well, since these are just a special case where membership is restricted to 0 or 1 and the upper and lower approximations are equal.

3. Integration to form a geo-attribute space

Of specific interest now is the geo-dipole concept (Goodchild et al., 2007) that accommodates for interactions between spatio-temporal phenomena. The geo-dipole is also defined as a tuple, again connecting a property with a value, but not just for one location in space–time as in the case of the geo-atom but for two locations; $<\mathbf{x}_1,\mathbf{x}_2,Z,z(\mathbf{x}_1,\mathbf{x}_2)>$. The property Z describes a relation between the two points based on their values. Distance, directions, and flows are maybe the most common types of interaction properties in spatial analysis, but there are many other types of interactions of potential interest to analysts. Also, interaction need not be confined to spatio-temporal relations but relations in any dimensions (Upton & Fingleton, 1985). There are for example numerous examples of research in information theory and cognitive semantics that use similarity metrics in aspatial information specifications such as ontologies (Resnik, 1999; Rodriguez & Egenhofer, 2003; Schwering & Raubal, 2005; Tversky, 1977). For example, Ahlqvist (2004) identified the Euclidean distance in an uncertain conceptual space and the fuzzy number overlap as informative metrics for an evaluation of the semantic relationship between land cover categories.

Because of the structurally close resemblance between the geo-object and the conceptual space representations we can see that it does not make any formal difference to generalize the geo-dipole concept to accommodate for three different relations;

a) between two locations in physical space $<\mathbf{x}_1,\mathbf{x}_2,Z,z(\mathbf{x}_1,\mathbf{x}_2)>$

b) between two locations in conceptual space <$\mathbf{s}_1,\mathbf{s}_2,Z,\mathbf{p}(\mathbf{s}_1,\mathbf{s}_2)$>, and
c) between locations in physical and conceptual space <$\mathbf{x}_1,\mathbf{s}_2,Z,\mathbf{p}(\mathbf{x}_1,\mathbf{s}_2)$>.

Case a) are the standard relations dealt with in spatial analysis, and case b) deal with the aspatial relations mentioned above. The c) relations describe mappings from properties in physical space to properties in attribute space as when measured spectral reflectance in a satellite image is compared with the spectral signature of land cover classes defined in a conceptual space with property dimensions of reflectance. Obviously, the type of relation metric used in physical space-time and the metric used in conceptual space need to be determined based on the question at hand to produce relevant results and much research is still needed, especially for aspatial relation metrics. This extension of the geo-dipole concept also allows for analysis of space-time-theme dependence. For example, spatial autocorrelation metrics that assess the degree of spatial dependence of a property are essentially formulated as a cross product $\sum\sum w_{ij}Y_{ij}$ that summarizes the product of a spatial relation metric w_{ij} and a property relation metric Y_{ij} for a set of paired observations (Getis, 1991). This would translate into the use of a) and b) dipoles such as in the construction of an empirical semantic semi-variogram (Ahlqvist & Shortridge, 2006).

These examples demonstrate that a tight integration between the geo-atom, and uncertain conceptual space representations is possible by extending the geo-dipole concept to link between space-time and conceptual dimensions. I call this comprehensive representation a *geo-attribute space*, and through this construct we can perform many of the existing methods of spatial analysis as well as analysis of ontological structure and also combinations thereof. Further studies are required to evaluate the usefulness of the geo-attribute space construct to develop for example interactive software that provides the same tight integration of space-time and concept dimensions, a path only tentatively pursued so far, providing exciting research and application opportunities.

4. References

Ahlqvist, O. (2004). A parameterized representation of uncertain conceptual spaces. *Transactions in GIS, 8*(4), 493-514.

Ahlqvist, O., Keukelaar, J., & Oukbir, K. (2003). Rough and fuzzy geographical data integration. *International Journal of Geographical Information Science, 17*(3), 223-234.

Ahlqvist, O., & Shortridge, A. (2006). Characterizing land cover structure with semantic variograms. *Progress in Spatial Data Handling -12th International Symposium on Spatial Data Handling,* Vienna, Austria. 401-415.

Gahegan, M., Dai, X., Macgill, J., & Oswal, S. (2003). From concepts to data and back again: Connecting mental spaces with data and analysis methods. Paper presented at the *Proceedings of the 7th International Conference on GeoComputation,* University of Southampton, United Kingdom. from GeoComputation CD-ROM database. from http://www.geocomputation.org/2003/Papers/Gahegan_Paper.pdf

Gärdenfors, P. (2000). *Conceptual spaces: The geometry of thought*. Cambrige, MA: MIT Press.

Getis, A. (1991). Spatial interaction and spatial autocorrelation: A cross-product approach. *Environment and Planning A, 23*(9), 1269-1277.

Goodchild, M. F., Yuan, M., & Cova, T. J. (2007). Towards a general theory of geographic representation in GIS. *International Journal of Geographical Information Science, 21*(3), 239-260.

Resnik, P. (1999). Semantic similarity in a taxonomy: An information-based measure and its application to problems of ambiguity in natural language. *JAIR, 11*, 95-130.

Rodriguez, M., & Egenhofer, M. (2003). Determining semantic similarity among entity classes from different ontologies. *Knowledge and Data Engineering, IEEE Transactions on, 15*(2), 442-456.
Schwering, A., & Raubal, M. (2005). Measuring semantic similarity between geospatial conceptual regions. *Lecture notes in computer science, 3799*, 90-106.
Stevens, S. S. (1946). On the theory of scales of measurement. *Science, 103*(2684), 677-680.
Tversky, A. (1977). Features of similarity. *Psychological review, 84*(4), 327-352.
Upton, G. J. G., & Fingleton, B. (1985). *Spatial data analysis by example. vol. 1: Point pattern and quantitative data*. Chichester: John Wiley & Sons.