Utilising Partonomic Information in the Creation of Hierarchical Geographies

Omair Chaudhry and William A Mackaness Institute of Geography, University of Edinburgh UK Tel: 44 131 6502532 Email: <u>O.Chaudhry@sms.ed.ac.uk</u>

Abstract

Intuitive and meaningful interpretation of geographical phenomenon requires their representation at multiple levels of detail. This is due to the scale dependent nature of their properties. Considerable interest remains in capturing once geographical information at the fine scale, and from this, automatically deriving information at various levels of detail and scale via the process of generalisation. Prior to the cartographic portrayal of that information, model generalisation is required in order to derive higher order phenomenon associated with the smaller scales. This paper presents an approach for aggregation of source database objects into composite objects at higher level of abstraction based on partonomic relationships. The benefits of these relationships in terms of database transformation and spatial analysis is discussed and illustrated with results.

Keywords: partonomies, boundaries, database transformation, spatial analysis

1.0 Problem Statement

In the Geosciences we want to model different geographies of the world (both in analysis and in visualisation). It is in this sense that a map can be considered to be a model of the world at a particular abstraction – one that reflects our perception of the world at a particular 'granularity', level of detail or scale. A map represents a phenomenological view – that we do not see lines points and text – but that we see things such as mountainous regions, airports or footpaths. Some of those features can be viewed as base objects from which we can create composite objects – so a close grouping of specialist buildings, car parks, and sport facilities create 'the school', and that in turn, a collection of schools, other municipal buildings, and dwellings creates the composite thing 'city'. When we see a 'dot' with the word London next to it, we understand the functional definition of what city means – according to some prototypical view of what a city. And the cartographer takes advantage of the fact that we have a shared view of what 'city' means (or for that matter any other thing we choose to display on our map).

Depending on the task, the user will require any number of things (base or composite objects) to be displayed on a thematic map (scaled and symbolised in a way that takes account of the intended use). The vision is that we survey (capture) only the base (component) objects, and structure the information in such a way that we can deliver this type of output at any level of detail. This paper has as its focus, a summary of some techniques that can automatically create composite objects, and also structure such things in a way that supports intuitive retrieval and display of such geographic information. It should be pointed out that base objects can contribute to the creation of more than one composite thing. This paper therefore

discusses the non exclusive nature of classification schema, and how multiple representations can be created, stored and provide context to spatial queries.

There are significant advantages to the automatic creation of composite objects (airports, harbours, mountain chains, forested regions, cities, etc) from base objects. These are quality control in the definition and creation of composite objects, and automatic update (so that growth in suburban house building automatically leads to revision of the city boundary). These things are usually stored within 'multi representational databases' (MRDB) and afford more meaningful query, analysis and display of geographic phenomenon.

Section 2 provides a brief overview of model generalisation; section 3 explains the process of determining functional (or partonomic) relationships of composite objects by using container boundaries; section 5 presents different approaches to deal with multiple partonomies during aggregation process section 4 illustrates the role of these partonomies in querying spatial databases and visualisation before concluding with a few ideas on future work.

2.0 Generalisation: Model and Cartographic

Different scales represent different information useful for different applications- there being a strong relationship between scale and phenomena(Sheppard & McMaster, 2004). There is no single (ideal) scale at which to view the world. The techniques developed within automated environments to control the abstraction of geographic phenomenon at different scales come under the heading of map generalisation. Map generalisation can be broadly categorised into model and cartographic generalisation (Figure 1). The focus of this paper is on aspects of model generalisation and does not review cartographic generalisation techniques.



Figure 1: DLM, DCM, Model and Cartographic Generalisation.(Brassel & Weibel, 1988; Grunreich *et al.*, 1992)

Model generalisation focuses on database transformation and cartographic generalisation focuses on visual representations. The objective of model generalisation techniques is to

reclassify and reduce down the detail, and give emphasis to entities associated with the broader landscape – thus enabling us to convey the extent of the forests rather than see the trees, or to see the island chain along the plate margin, rather than the individual island. Typically model generalisation precedes cartographic generalisation. Model generalisation is also required in response to a non-visual query, or as a prerequisite to data analysis. For example the question 'what modes of travel exist between the cities of Edinburgh and Glasgow?' requires us first to aggregate together phenomena at the fine scale (in this case dense regions of buildings) in order to define the extent and general location of these two entities. Only then can we identify, for example, the major roads that connect these two urban centres.

2.1 Model generalisation

Model generalisation involves transformation of a database instance of a spatial data model at higher level of detail, into a database instance of a data model at higher level of abstraction. Such a transformation involves the creation of composite objects. Composite or higher order objects are formed via the process of thematic and spatial abstraction. In thematic abstraction the number of distinct attributes of objects in the database is reduced. In spatial abstraction the number objects are reduced by means of aggregation or elimination. Thematic abstraction often triggers spatial abstraction. For instance objects having similar attribute structure can be categorised into classes under the process of classification. Each object then becomes an instance of a particular class and that class defines an object's properties in terms of its attribute structure. If different classes share some attributes then a super class or parent class can be created whose attributes are the common attributes of its child classes. This creates a hierarchy in which complex classes are present at the detailed (low end of a hierarchy) and increasingly abstracted classes are present as we go up the hierarchy. This type of hierarchy is called a taxonomy or classification hierarchy (Figure 2a). The creation of these hierarchies is an important way of modelling changing levels of detail and provides a basis for creating generalised maps (Figure 2b).



Figure 2: Using taxonomy for aggregation of source objects

Another complimentary hierarchy useful in the creation of composite objects is a partonomy. Whereas a taxonomy refers to a 'is-a' relationship, a partonomy refers to 'part-of' relationships between parent and child classes – reflecting more of a functional and conceptual division of geographic space. Over large changes in scale it is necessary aggregate objects belonging to different classes in order to create composite objects. A prototypical

view of a city might be defined as a dense collection of municipal and industrial buildings, and multi modal transportation infrastructures. Once represented in partonomic form, the resulting hierarchical structure can be used as a basis for combining objects together – in this case moving from the detail of the house, land parcel and pavement, to a simple building block (Figure 3).



Figure 3: Using partonomy for aggregation of source objects

2.2 Creating Partonomic Structures

A variety of techniques exist to create partonomies; one is to create functional boundaries (in the way that we might 'group' railway sidings, stations and track to define 'the railway network'). Others involve the use of models to define the degree to which something is part of something else. For example Chaudhry & Mackaness (2006a, 2007) used a model to define city and forest 'container' boundaries (Figure 4). By creating these boundaries (Figure 5) we are able to structure phenomenon according to (hierarchical) geographies – creating a set of nested geographies that 'fit; with our perception of reality'.



Figure 4: a) Input database OS MasterMap (1:1250/1:10,000/1:25,000) and 4b) a surface representing the value of citiness for objects in Figure 3.

If we can successfully utilise boundary information to create these partonomies, then we can automatically create a hierarchy of geographies. This is invaluable in:

- automatically populating MRDB;
- simplifying/ facilitating the model generalisation process;
- creating responses to spatial queries that are meaningful;
- modelling explicitly the interdependence that we typically find among / between geographic phenomenon thus supporting the creation of semantic reference systems;
- supporting intuitive 'geographic' query rather than simply 'spatial' query;
- supporting novel ways of visualising geographic phenomenon.

3.0 Making Partonomies Explicit and Exclusive

Unlike taxonomic relationships, partonomic relationships are not exclusive. This means that an object can be part of more than one higher order object. Different approaches into database transformation have been proposed together with different ways of determining the partonomic relationship between objects (Molenaar, 1998; van Smaalen, 1997; van Smaalen, 2003). In this paper container boundaries defining the limits of higher order objects are used in order to determine the partonomic relationships of the source objects in terms of the higher order or composite object.

Boundaries of natural geographic phenomena are much less distinct and discrete than anthropogenic ones. This observation is reflected in research on the modelling of fuzzy boundaries. Nevertheless a boundary separates the entity from its environment and is one of the marks of its individuality (Casati *et al.*, 1998). Geographic boundaries are dependent upon the scale of observation; sometime they are vague and fuzzy (for instance the extent of a wetland, or mountainous region). Boundaries can be divided into two basic types: bona-fide boundaries and fiat-boundaries (Smith & Varzi, 1997; Smith & Varzi, 2000). Bona-fide boundaries are a result of either spatial discontinuity or qualitative heterogeneity whereas fiatboundaries are the result of human cognition. Bona-fide boundaries include boundaries of a building, pavements, and roads whereas fiat boundaries include the boundary of a city, district, and block. In our earlier research (Chaudhry & Mackaness, 2006a; Chaudhry & Mackaness, 2007) we presented an approach for the creation of fiat boundaries (Figure 5) of higher order objects such as a city or forest from bona fide objects such as buildings and tree patches from a large scale source database (Ordnance Survey MasterMap Figure 4a).



Figure 5: Boundaries for settlement and forest objects (Chaudhry and Mackaness 2006, 2007). The two red rectangles marked in the figure are the areas shown in detail in Figure 6 (top) and Figure 7(bottom).

3.1 Determining Partonomies

The resultant boundaries (such as those created in Figure 5) can act as a 'containers' – all objects within are classified as 'part of' the settlement or forest (Figure 6). Within the database this is achieved via a spatial join between the boundary and the source objects (Figure 4a). Using the nine intersection model (Egenhofer & Franzosa, 1991; Egenhofer & Herring, 1990) topological relationships between the source objects and resultant boundaries can now be identified. All objects in Ordnance Survey MasterMap (source database) used and resultant boundaries are area (polygon) objects. If the source objects are completely 'contained by' the resultant boundary they are deemed part of the higher order object. If 'disjoint' they may be considered part of some other higher order object. If it's an 'overlap' relationship then the percentage of area intersection can be used to govern membership. This results in the creation of multiple partonomies. This is reflected in Figure 6 in which different regions (numbered) partially overlap – reflecting shared membership most often in proximity between the two boundaries.



Figure 6: Determining membership (partonomy) of source objects in terms of higher order objects: the case of overlapping boundaries

3.2 Multiple representation of a 'palimpsest' of partonomies

If an object is completely overlapped by the resultant boundary object then the object can be viewed as definitely being part of the higher order object. In some cases objects are partially covered by overlapping boundaries (Figure 6) or the boundaries are themselves part of another boundary (Figure 7). In these instances multiple, overlapping partonomies will exist.

Case a: In Figure 6 the resultant boundaries of the higher order objects are overlapping. In these cases the source objects will be partially covered by these boundaries. A threshold can be used to determine which higher order object boundary has the greatest amount of area intersection. An alternative approach is to allow the source objects to have multiple partonomies. Instead of deciding whether an object is (or is not) part of a higher order object we can define degrees of membership. In Figure 6 for instance building object 11 is partially covered by boundaries of settlement as well as boundaries of forest objects. The amount of area intersection can be used to determine the degree of membership.

Case b: In Figure 7 the boundaries of the higher order object identified are themselves part of other higher order object boundaries. For instance a settlement boundary can be part of forest boundary (Figure 5 and Figure 7). In such cases the source database object will again have multiple partonomies. Thus a building object that is part of a settlement object which is part of a forest can be considered to be part of the forest as well. This is analogous to the example of piston being part of engine and the engine part of a car such that the piston is also part of the car. This is axiom of *Transitivity* in partonomy (Smith & Varzi, 1997; Varzi, 2003). In Figure 7 for instance Building object 17 is part of the settlement as well as part of the forest even though the area intersection between building and forest object is null.

A $Pxy \land Pyz \rightarrow Pxz$ Transitivity

(Axiom of transitivity: if x is part of y and y is part of z then x is also part of z)



Figure 7: Determining membership (partonomy) of source objects in terms of higher order object.

4.0 Multiple Partonomies for Aggregation and Visualisation

As explained in the introduction section for database transformation overlarge changes in the level of detail we need to aggregate objects belonging to different taxonomies. We need to determine the partonomic relationships for transformation of source objects into target or higher order objects. But when objects have multiple partonomic relationships which partonomic relationship will be used by aggregation. A few possible solutions are presented.

Case a: Based on percentage of partonomy. As explained above the percentage of area intersection between the source object and resultant higher order object boundary illustrates the degree of membership. This degree can then be used for aggregation process by having a threshold. All objects that have an intersection area percentage above certain threshold are selected for aggregation process (Figure 8).



Figure 8: Aggregation based on degree of membership. Here scrub is aggregated with settlement object

Case b: Based on a predefined context. A database transformation is application dependent process. Each application imposes different sets of conditions and constraints that affect the process of transformation and thus affecting the process of aggregation. For instance if the resultant database is transformed for urban planning then settlement partonomy can be given higher precedence where as if the application is vegetation analysis then forest partonomy can be given precedence. Thus all objects that are part of a settlement, no mater the degree membership is, are selected for the process of aggregation.

Case c: Based on similarity matrix. Many studies on aggregation process have proposed different ways of expressing compatibility between classes (Yaolin *et al.*, 2002). Compatibility values between the output classes of the resultant database and classes of the source database can be identified in advance and a resultant matrix can be created. During the process of aggregation the highest compatibility value between the source object and the target object's class is determined from the matrix and this is used as criteria for aggregation of objects having multiple partonomies. For instance if an scrub object is part of both settlement and forest then its class similarity value is checked from matrix since a scrub is more similar to a forest as compared to a settlement the value will be large and the object will be aggregated into a forest object as illustrated in Figure 9.





But these similarity values are difficult to generate automatically (Bregt & Bulens, 1996). So a combination of above cases might be more appropriate. For instance if a road source object is part of a settlement and a forest object and there is no similarity value between these classes. Then case 'a' could be used to determine the higher degree of partonomy for aggregation. But if the degree of membership is equal then case 'b' could be used. Thus the order in which these cases are used is case 'c' than case 'a' and then case 'b' for aggregation of objects having multiple partonomies.

A few results using the above approach are presented in Figure 10 and Figure 11. In Figure 10 the output composite objects are created from component objects in the source database

(OS MasterMap) illustrated in Figure 4a. The number of objects has been reduced from 57730 (Figure 4a) to 980 (Figure 10). Note that important objects such as major roads and railways have not been aggregated and are selected directly from the source database (Chaudhry & Mackaness, 2006b). Figure 11b shows the output composite objects for a different area (Figure 11a). The number of objects have been reduced from 25000 (Figure 11a) to 340 (Figreu11b) objects.



Figure 10: Resultant composite objects created via aggregation of components objects in Figure 6a using partonomic relationships



Figure 11: a) Test area showing component or base objects in the source database (OS MasterMap). b) Resultant composite objects and their output classes

5.0 The Value of Creating Partonomies – 'Geographical' Query

Once the partonomies relationships have been determined for source objects in terms of the target higher order objects, these relationships can be utilised for spatial analysis. For instance 'retrieving all building or coniferous trees' that are part of Edinburgh city can now be performed by a simple query that checks the partonomic relationship of all source objects in terms of 'Edinburgh city. Similarly queries can be made more complex by retrieving only those objects that are only part of Edinburgh i.e. degree membership for source objects in terms of Edinburgh city object is 100% and 0% with any other higher order object. Being able to convey and query based on partial membership, accommodate the level of detail associated with a query. It is importance to stress that such a query was not possible in the original database since there were no objects representing the particular city.

Having multiple partonomies is useful for supporting intuitive 'geographic' query rather than 'spatial' query – i.e. the query fits the 'scale' associated with the task. Thus a query concerning how to get from Easter Calder to Livingston needs to pitch the question at the right 'geography' – at the level of the city (not the region, and not the building level) (Figure 12).



Figure 12: The railway network partonomy (station and a railway line) are one of the means by which the city partonomy of Glasgow and Edinburgh are connected.

Partonomies to support richer/innovative forms of visualisation of phenomenon (Figure 13). In Figure 13 the degree of membership of objects in terms of citiness (Figure 4b) can be combined with crisp boundaries of a forest to visualise the relationship between different phenomena. This can used for identification areas where phenomena transits from one to another.



Figure 13: Automatically creating new types of features from the combination of partially overlapping partonomies.

6.0 Conclusion

Maps and spatial databases at different levels of detail present different information levels. Partonomic relationships revels the interdependence between different phenomena at different levels of detail. Unlike taxonomies an object can be part of more than one object. Multiple partonomies are not only useful for database transformation from one level of detail to lower levels of detail. But provides enriched databases that facilitate users in reasoning about space more intuitively. This paper has presented how partonomic relationships can be determined by the use of pre determined boundaries. The resultant partonomic relationships, whether multiple or not, can be used for spatial analysis, visualisation and database transformation.

7.0 References

- Brassel, K.E. & Weibel, R. (1988) A review and conceptual framework of automated map generalisation. *International Journal of Geographic Information Systems*, 2, 229-244.
- Bregt, A.K. & Bulens, J. (1996). Application-Oriented Generlisation of Area Objects. In Methods for the Generlisation of Geo-Databases (ed M. Molenaar), pp. 57-64. NCG, Delft.
- Casati, R., Smith, B., & Varzi, A. (1998). Ontological Tools for Geographic Representation. In *Formal Ontology in Information Systems* (ed N. Guarino), pp. 77-85. IOS Press, Amsterdam.
- Chaudhry, O. & Mackaness, W. (2006a) Creation of Fiat Boundaries in Higher order Phenomenon. In Workshop of the ICA Commission on Map Generalisation and Multiple Representation Vancouver - WA.
- Chaudhry, O. & Mackaness, W. (2006b) Modelling Geographic Phenomena at Multiple Levels of Detail. In AutoCarto, Vancouver- WA.
- Chaudhry, O. & Mackaness, W. (2007) Automatic Identification of Urban Settlement Boundaries in Multiple Representation Databases. *Computer Environment and Urban Systems*, (under review).
- Egenhofer, M.J. & Franzosa, R.D. (1991) Point set Topological Spatial Relations. International Journal Geographical Information Systems, 5, 161 - 174.
- Egenhofer, M.J. & Herring, J.R. (1990). A Mathematical Framework for the Definition of Topological Relationships. In *Proceedings of the 4th International Symposium on Spatial Data Handling, held in ZŸrich, July, 1990,*, pp. 803 - 813.
- Grunreich, D., Powitz, B., & Scbmidt, C. (1992) Research and Development in Computer-Assisted Generalisation of Topographic Information at the Institute of Cartography, Hanover University. In Proceedings of EGIS '92, Vol. 1, pp. 532-541. EGIS Foundation, Munich.
- Molenaar, M. (1998) An Introduction to the Theory of Spatial Object Modelling for GIS Taylor and Francis, London.
- Sheppard, E. & McMaster, R.B. (2004) *Scale and Geographic Inquiry: Nature Society and Method* Blackwell Publishing.
- Smith, B. & Varzi, A.C. (1997) The Formal Ontology of Boundaries. *The Electronic Journal* of Analytic Philosophy.
- Smith, B. & Varzi, A.C. (2000) Fiat And Bona Fide Boundaries. *Philosophy And Phenomenological Research*, 60, 401-420.
- van Smaalen, J. (1997) Spatial Abstraction Based on Hierarchical Reclassification. *Cartographica Monograph* 47, 33(1), 65-73.
- van Smaalen, J. (2003) Automated Aggregation of Geographic Objects: A New Approach to the Conceptual Generalisation of Geographic Databases. *Doctoral Dissertation, Wageningen University, The Netherlands.*
- Varzi, A.C. (2003) Mereology. In Stanford Encyclopedia of Philosophy (ed E.N. Zalta). Stanford: CSLI (on line publication).
- Yaolin, L., Molenaar, M., & Kraak, M.J. (2002) Semantic similarity evaluation model in categorical database generalisation. In Proceeding ISPRS Commission IV Symposium, Ottawa, Canada.