Sharing feature based geographic information - a data model perspective

Joseph Greenwood and Glen Hart

Research and Innovation, Ordnance Survey,
Romsey Road, Southampton, SO16 4GU.
Telephone: +44 (0)23 8079 2125,
Email: jgreenwood@ordsvy.gov.uk, ghart@ordsvy.gov.uk

Biography

Research Scientist, Research and Innovation, Ordnance Survey, 2001 – to date. Research interests and conference publications in the areas of data modelling, data infrastructures and map generalisation.

Abstract

This paper reassesses the role of the "Feature" as the atomic unit of geographical information exchange. It proposes the desegregation of the feature into components so that information related to the geographical entity represented by the feature may be exchanged at a finer level of detail and examines the implications of this with respect to the geodata model.

Introduction

The ability to effectively share geographic information between different technical disciplines and organisations is becoming increasingly important, be this the inter-disciplinary nature of scientific research, the need for 'joined-up' governmental and commercial processes or to create mobile services for multiple consumer groups. Each domain has its own 'world view' which describes the conceptualisation of the real world objects of phenomena under consideration in the domain. From a data perspective this is encapsulated in data capture and maintenance specifications and regimes. As Burrough and Masser (1998) note it is "the fact that related sciences and technical methods do not necessarily describe or recognise the same spatial phenomena in the same way".

Setting aside the political or economic barriers, the technical task for sharing geographic information becomes that of being able to explicitly model data that represents differences in definition and interpretation of the same geographic phenomena. This must be done such that it enables the similarity and differences to be explicitly managed to enable human or computer decisions of whether the information can be exchanged.

Features

The basic unit of geographic information within most models of geographic information is the 'feature', where a feature is an abstraction of a real world phenomenon and a geographic feature is a feature associated with a position on earth (OGC, 2002). Features can include representations of a wide range of phenomena that can be located in time and space such as buildings, towns and villages or a geometric network, geo-referenced image, pixel, or thematic overlay. This means that traditionally a feature encapsulates all that a given domain considers about a single geographic phenomena in one entity.

Features can be seen at two levels *feature instances* and *feature types*. Feature instances are the individual discrete representations of geographic phenomenon in the database with geographic and temporal dimensions. The instances may then be grouped into classes with common characteristics to form feature types. The feature type is effectively a template for instances of that type and the relationship between the instance and the type is typically that of instantiation of an instance in a single type. In OGC terms however features are not fixed in their class but have application-oriented views that are classed (OGC, 2002) that is to say one view may class an instance as house while another may class it as a building. It is apparent then that the feature is not the atomic unit of geographic information as the phenomena it represents encapsulates differing human concepts resulting in multiple types.

A closer examination of the behaviour of those in involved in the collection, management, supply and exploitation of geographical information shows that different organisations utilise different collections of discrete sets of information about geographical entities. In this analysis the feature is a higher order representation of a shared understanding of a specific geographical representation and the actual atomic units of exchange are the discrete sets of information.

It therefore follows that in order to enable the exchange of feature based geographical information a model is required that allows the comparison of features not only as whole units but at a finer level that can consider the difference in what they encapsulate.

Data Model

Several authors have recognised the difficulty in reconciling differences in feature definition and have proposed, for varying reasons, modelling or defining data at a sub-feature level to allow the explicit management of varying feature definitions such as Rodriguez and Egenhofer (2003), Hart and Greenwood (2003), Usery *et al*. (2002) and Brodaric and Hastings (2002). The 'component based' approach to feature modelling of Hart and Greenwood (2003) is a meta-level model where features are modelled as a set of discrete components of *identity* and *data* and an ontological framework for description and logical combination of components. A data component represents a discrete package of data that corresponds to a single concept defined in the ontology, the feature is modelled as a collection of data components and a unique, unambiguous and persistent identity component as shown by Figure 1.

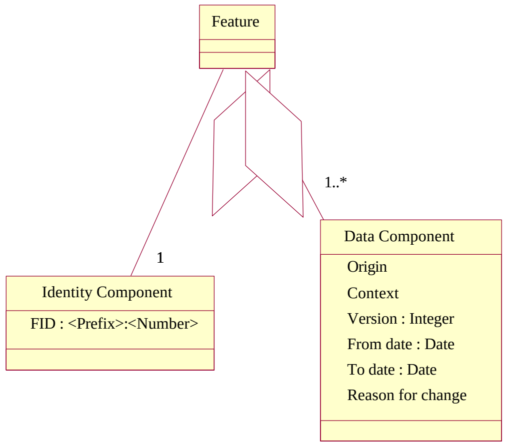

Figure 1

This approach provides the potential to:

- Enable different data suppliers to gather different components representing different world views and assign them to the same feature;
- Manage this data at the most appropriate level – that of the data component, so change metadata, ownership (including versioning if required) and quality data is held at this level, not that of the feature.
- Make it possible for the task of data sharing to be carried out through a selection by the consumer (or end chain supplier) of the data components held by one or more data suppliers that correspond to the feature of interest.
- Enable the classification of the features based on the needs of the consumer since the classification can be derived from the supplied components as these are directly related to the ontological framework. This relaxes the requirement of having to define all feature types *a priori* and also enables new world views to be constructed by a data supplier or consumer
- The differences in domain definitions of features may also be examined as the difference in the set of associated component used.

The full paper will concentrate on the logical implementation of data component and feature aspects of the component based approach in order to enable data sharing at a feature level. The relationships to an ontology and ontological modelling will not be covered in depth in the paper. Specifically this paper will address 1) how to mange components of features which gain their meaning or existence from shared identity 2) how to build features from a set of components 3) how to manage multiple feature types, in a standard object-relational environment, particularly the use of the type mechanism for multiple feature types for one instance. Further it will explore the implications on model design for sharing features across more than one database.

References

Brodaric, B. and Hastings, J. (2002) An object model for geologic map information. In: *Proceedings Spatial Data Handling 200*, July 9-12, Ottawa, Canada.

Burrough, P. and Masser, I. (1998) International aspects of spatial data exchange. In P. Burrough and I. Masser (Eds.) *GISDATA 5 European Geographic Information Infrastructures opportunities and pitfalls*. pp 1-14. Taylor and Francis, London, 1998

Hart, G. and Greenwood, J. (2003) A component based approach to geo-ontologies and geodata modelling to enable data sharing, In: *Proceedings of the 6th AGILE Confernce,* Lyon, April 24-26, In Press.

OGC (2002) *OpenGIS Reference Model*. Open GIS Consortium Inc. Wayland, MA, USA.

Rodríguez, A. M. and Egenhofer, M. J. (2003) Determining semantic similarity amount entity classes from different ontologies. *IEEE Transactions on Knowledge and Data Engineering*, Vol. 15, No. 3, pp442-456.

Usery, E. L., Timson, G. and Coletti, A. (2002) Multidimensional Geographic Feature System. In G. Zavala (Ed) *GIScience 2002 Abstracts*. *GIScience 2002 The Second International Conference on Geographic Information Science.* Boulder, Colorado, USA pp 279-282. University of California Regents