A Statistical Semantic Approach For Analysing Change Using Lcmgb And Lcm2000.

Comber, AJ[1], Fisher, PF[1], Wadsworth, RA.[2]

[1]Department of Geography, University of Leicester,
Leicester LE1 7RH, United Kingdom;
Tel: +44 (0)116 252 3859;
Fax: +44 (0)116 252 3854;
Email: ajc36@le.ac.uk

[2] Centre for Ecology and Hydrology, Monks Wood,
Abbots Ripton, Huntingdon, United Kingdom;
Tel +44 (0) 1487 772400;
Fax +44 (0) 1487 773467

Biography

Lex Comber is a researcher on the EU REVIGIS (revising geographical information) project based at Leicester. His research interests include managing uncertainty associated with land cover data, automated approaches to land cover monitoring, and spatial data integration.

Introduction

We often want to combine data that are coincident in time or space despite it being in different formats, at different scales and with different levels of accuracy. Much work addresses how such data may be combined under various banners of "Interoperability", "Data Sharing" and "Standardisation". It focuses on the readily measurable qualities of the data (often contained in meta-data: scale, resolution, representation, format) due to the DBMS / computing science origins of these objectives, where data qualities must be numerically measurable to provide a numeric link between dataset objects, thus moving us towards greater data base federation. However, integrating activities are hindered by differences in conceptualisation as determined by different methodologies and objectives. There has been little work that considers how differences in the *meaning* of data objects (e.g. pixel or segment) may be overcome. This is because the data ontologies (i.e. meanings, semantics or conceptualisations of the data) are difficult to quantify, especially in a way that is neither specific to the dataset nor the intended integration.

In this paper we develop an approach that seeks to overcome the semantic stumbling block to data integration using UK national land cover datasets from 1990 and 2000. Because relations between different land cover semantics cannot readily be measured (they can only be interpreted), necessarily they are subjective and relative. We attempt to overcome this subjective ambiguity using expressions of Land Cover Expert opinion in order to relate the different data concepts of the 1990 LCMGB and its successor LCM2000 and thereby to identify locales of land cover change.

There are considerable methodological differences between the two datasets. The most significant of these in terms of data ontologies are changes in the way that land cover classes with a high policy value are defined. Classes such as "Bog" and "Acid

Grasslands” were not identified by their reflectance properties, rather through the application of soil masks.

Statistical Semantic Approach

The overall approach is to characterise each LCM2000 object (segment) twice: in 1990 using the intersection of LCMGB pixels and in 2000 using spectral attributes that are independent of the LCM2000 classification. The intersection with LCMGB data is interpreted via the opinion of the Land Cover Expert, and the spectral subclass attributes through published information of expected spectral confusion. For each LCM2000 polygon, for each time an Unexpected and Expected score is calculated. From these a vector showing the movement of the segment can be plotted. We hypothesised that segments whose scores had changed most were likely areas of change. Polygons identified in this way were visited and the present land cover recorded along with any contextual evidence of change.

Results

Locales of land cover change were successfully identified where both datasets were correct and where the ontological differences between the two classifications were minimised.

Change was not reliably identified for land cover classes subject to a high degree of ontological change such as “Bog” and “Acid Grasslands”. This was due to three factors: a failure by the expert to understand fully the changed ontology; LCM2000 spectral meta-data not relating to the land cover class label; the LCM2000 class not relating to the land cover on the ground.

Discussion

We discuss the meaning of the results with regard to three levels of ontological change and the different directions of the vector in the feature space of Unexpectedness and Expectedness. We discuss other aspects of the LCM2000 data highlighted by the analysis including the nature of the ontological links between the LCMGB and LCM2000 (i.e. class differences and similarities), ontological ambiguity in LCM2000 class definitions (i.e. where class definitions overlap), some of the spectral ambiguities in LCMGB.

Conclusions

We conclude that the semantic statistical approach, using Look Up Tables, provides a generic, widely applicable method for developing second level, linkage ontologies. This is a prerequisite step to being able to combine data of different ontological pedigrees, for a specific application.

We commend the LCM2000 production for including extensive meta-data, which allows a better fit to the translation between two classifications to be made than would be possible using a table of correspondence.