RULE BASED INTERPRETATION OF UNSTRUCTURED VECTOR DATA

Manuel Weindorf
Institute for Photogrammetry and Remote Sensing
University of Karlsruhe, Germany
weindorf@ipf.uni-karlsruhe.de

KEY WORDS: GIS, rule based interpretation, unstructured vector data

ABSTRACT

The basic motivation for this work is the fact that interoperability of GIS in general and data exchange in particular is still a challenging task in industry and research. Nowadays data exchange is still done on a graphical representation level, in the way that for every combination of data source and data sink (GIS) a specific mapping (at the symbolic level) has to be defined to translate the several data formats into other ones. The idea proposed in this work is the definition of the target-systems data model in terms of a rule base, which is used later on by a map interpreter to support many data sources applying this rule base. This leads to the need of only one set of transformation rules per target system instead of many mappings (one per source system) which is the traditional way.

Therefore the presented work deals with the problems arising while interpreting digital unstructured vector data provided by low-level graphics formats (e.g. map data digitized within CAD environments) for object recognition and extraction. In the context of automated map interpretation, the problems of grouping of geometric entities and validating of possible configurations play a major role and lead usually to combinatorical problems.

At the IPF a system for automated map interpretation, which is based on the concepts of rule based production systems, is under development. This system is able to classify geometric entities out of unstructured vector data and compose objects out of these primitives. These classification and composition are achieved by applying inference rules and grammars (e.g. for text classification) in combination with spatial predicates like close_to, within, contains etc. On the one hand, these rules have to be general in the way that they have to be applicable for many data sets from different sources, similarly they have to be specific enough to discriminate between several object types in a reliable way. One problem with using a set of rules to fulfill the needs described above, is that on the local level of interpretation exist usually more than one hypothesis for one specific instance of an object. Hence a decision has to be taken without knowledge about the overall result of the interpretation (which is just available in the end). The approach proposed in this work first collects all possible object hypothesis and searches for the optimal solution by minimizing object specific cost functions. This leads to an optimization problem which is tackled by a branch and bound based search technique in combination with a constraint logic programming. The system has been tested on several datasets of medium and large scale.