Specification for a Shared Conceptual Layer in GIS

Enguerran Grandchamp
Département Mathématiques et Informatique
LAMIA, Université des Antilles et de la Guyane
Pointe-à-Pitre, Guadeloupe France
e-mail: egrandch@univ-ag.fr

***Abstract* - During the last decade, GIS have suffered of the lack of structured information and semantic definition of the data. The numerous origins of the data and the spatial localization of the object inferred many inconsistencies among them. Indeed, the way to collect data and the level of sampling of the boundary differ from one application to another and from one expert to another. This leads to different representations of the same reality and different models of semantically identical objects or boundaries. To allow the matching of the objects, or parts of the objects, when combining different information layers we introduced a conceptual layer which described a scene with concepts and spatial relationships based on the definition of an ontology. This layer allows detecting and solving conflicts between the common boundaries of two objects. We present in this paper the specification and the way to build and use the ontology.**

Keywords - GIS; Information Fusion; Ontology.

I. INTRODUCTION

Nowadays, geographic information invades our life with more and more applications (navigation, localization, survey, security, etc for tourism, energy, environment, urban deployment, etc.). This growing of applications requires more and more data collected by different ways and different peoples leading to a great heterogeneity.

Geographic database are not well structured and consistent because of the numerous and non experts producers of information. Indeed, there are so many information layers and terminologies as people that produce them.

Moreover, the way to collect the data could be very different from high technologic tools (GPS receptors) to visual lecture on a map. Depending on the application, the expert needs few descriptions of the objects geometry or on contrary detailed ones.

All these considerations lead to different representations of the same reality with different accuracy (localization, description, etc.). This is particularly the case for natural or semi-natural limits such as forest, agriculture, etc. which are not well defined or localized (in contrast to buildings for example). The main problem is that most of the representations do not coincide; but we could not say that one is worst or better than the others. This is just an arbitrary choice of the vertices to sample the objects.

If the impact of such differences is not a problem for a thematic exploitation, there is a consequently propagation of errors when dealing with multi thematic problematic.

Indeed, the main drawback of this heterogeneity is that the power of the spatial analysis is reached when combining different thematic layers (agriculture and environment, roads and emergency, rainfall and population, etc.) and due to localization mistakes this combination leads to inconsistent information. Applications like the ones presented in [13] and [14] suffer from these localizations errors.

In order to avoid mistakes, we decide to refer to a conceptual view of the scene which describes the objects in terms of spatial relationships, geometry and semantic.

This view is an ideal representation of the scene and is based on the definition of an ontology and spatial relations.

An ontology is considered as being a set of concepts and relation between them. It gives semantic information to the scene.

Section 2 presents related works on spatial ontologies, ontology integration in GIS and localization errors correction. Section 3 explains the way to build the ontology. Section 4 deals with the conceptual layer. Section 5 explains the way to use the conceptual layer to solve conflicts and then Section 6 gives the conclusions and perspectives of this study.

 ISBN: 978-1-61208-118-2

II. RELATED WORKS

We present here related works concerning ontology integration in GIS and spatial object localization correction.

As Fonseca et al. [7][8] and later Cruz et al. in 2005 [5], defined an ontology driven GIS, there isn't any operational framework for an ontology integration in GIS.

The association between spatial data and ontology is done with external tools such as SPIRIT [18] or GeoSVM [7].

The first step in integrating ontologies in GIS is to define ontologies on spatial data. The main objective of these ontologies is to represent concepts linked to a specific expert domain (biology, geology, tourism, urgency, etc.) and also spatial relationships between them [25]. Many authors ([7], [8], [16], [23]) propose specific ontologies which differs from the used terms and features to describe the concepts. Indeed, depending on the application and domain some concepts could be described by different ways and with different levels of detail as if they are semantically identical.

The choice of the most adapted ontology for an application is not easy.

The second step is to introduce ontologies in GIS. Many authors introduce them in different ways.

Viegas et al. [24] and Baglioni et al. [2] create an intermediate semantic layer between the user and the geodatabase in order to facilitate the user's queries.

In [6], [10] and [11] the spatial ontology is used to facilitate object classification.

Concerning the correction of localization errors of spatial objects, to the best of our knowledge, the only approach to solve the problem of corresponding points between different information layers is expressed in [26] and uses buffer area to fuse vertices. This approach has two main drawbacks: the process is applied over all vertices *(i)* without distinction between semantically different ones *(ii)* without integrating spatial relationships.

III. ONTOLOGY BUILDING

There are many ways to build a geographic ontology. Many different ontologies had been defined in the literature, each one done for a specific application.

We have to choose among them the most appropriate ones. The first thing to do is to extract the *Intersection Knowledge* between these ontologies in order to have a common and recognized basis. After that, we select the *Augmenting knowledge* linked to a specific domain or expert [12]. To compare these ontologies and detect similarities we have to define metrics ([9], [19], [21]).

Figure 1. extracted from [12] summarizes the notion of *Intersection and Augmenting Knowledge*.

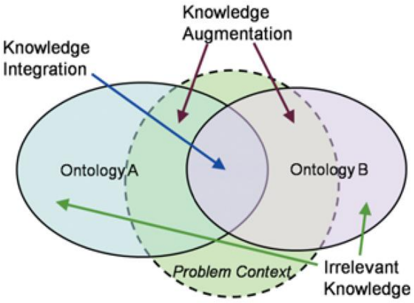

Figure 1. Ontologies intersection ([12])

We built the ontology by joining different independent ontologies compared to more general ontologies using similar or synonym terms ([17], [20], [24]) with the following criterion: *(i)* recognized ontology or commonly used one, *(ii)* built with a complete hierarchical object model, *(iii)* written with OWL (Ontology Web Language). In particular, we use SWEET (Semantic Web for Earth and Environmental Terminologies) ontology.

The ontology is introduced in the GIS using annotations in the data and in the metadata [22].

There is also automatic process to produce the ontology from the geographic data themselves [1], but the restricted area of interest and the numerous experts domains limits the advantage and efficiency of such a method and we prefer a manual building based on metrics evaluation.

IV. THE CONCEPTUAL LAYER

The process presented in this article is summarized in figure 2. Objects A and B come from different layers and represent different kind of objects or not. Depending on the application we decide to trust the boundary of object A (the case illustrated) or B. Decision criterions could be *(i)* the scale sampling of the objects *(ii)* the distance in the semantic space between the belonging class of A and B and the application (i.e. we prefer to trust river boundaries in an application concerning water). In the illustrated example, the boundary between A and B must be the same in the conceptual layer (example: A and B are to related agricultural fields).

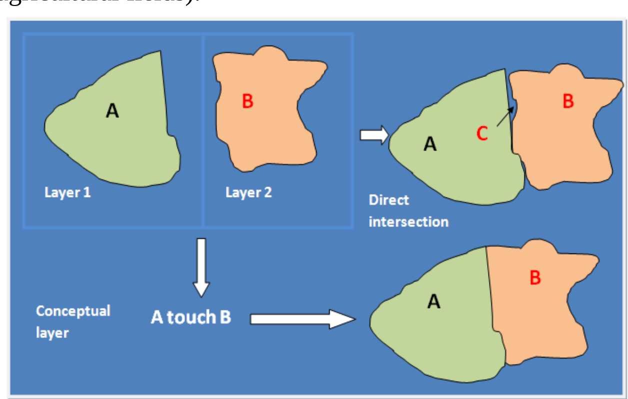

Figure 2. Process

 ISBN: 978-1-61208-118-2

But the boundaries are not the same in the two layers (Area C appears between objects A and B). We decide to conserve the boundary of object A because it is the most reliable one in the context of the application.

Because they will always have different ways to model the same object, we do not have to adopt a unique representation. As an example, depending on the observation scale, a building could be seen as a point, a square, or a more detailed polygon. Depending on the expert, the focus will be done on environment objects or urban ones. So we prefer adding semantic information concerning the data and the instances in order to have a common reference to compare the different view of the same data.

But nowadays, there isn't any way to define a vertex or a boundary as being an external resource or reference. As an example, if in a soil occupation layer we want to refer to a river defined in a hydrology layer we couldn't do it in a simple and normalized way. We can only do it during the layer creation step if we define an object as being the intersection or union of others object taken from sources layers. But as soon as the layer is created there is no link kept between layers and objects. So if we update the sources layers, changes are not propagated to the built layer.

To solve this problem we propose a common conceptual layer which regroups an ontology, representing the concepts of the scene, and the spatial relationships (topology) between concepts and instances.

A. *Ontology*

The ontology regroups the concepts present in the scene. We have to link these concepts with each object in the information layer in order to specify that an object is an instance of one or several concepts. To do so, we add a descriptive field in the table of the information layer which refer to concepts in the ontology and give information such as *is instance of*.

B. *Spatial Relation*

The proposed conceptual layer is a kind of extension in the GIS environment of the spatial relation hierarchy proposed in [4].

The authors of [4] define three levels for the representation of spatial relationships: the geometric level, the computer level and the user level. The geometric level is an abstract representation of the relations that we can consider to be exact (we use this level for the conceptual layer). The computer level takes into account the object reality and is a spatial relationships and geometric view of the information layer. If object geometry or relationships are altered in the information layer, this level will also be altered. We will compute this level directly from the data by analyzing the spatial relation between objects in order to qualify and quantify the relations. The user level is a restricted view of the information linked to the domain of the user and won't be used in the scope of this framework.

At the geometric level, we characterize: *(i)* the relations between the concepts of the ontology (more a qualification than a quantification). For example, the concept *private house* could be considered as being *linked to* the concept *road*. *(ii)* the relations between the instances of the concept (more a quantification than a qualification). For example, *house n°123, touch, road n°12*.

To characterize the spatial relations, Egenhofer [27] defines topological relations with intersection and overlapping (based on 4 or 9 intersection). These relations take into account:

1. The relative position which are strict (touch, intersects, etc.) or vagueness (close to, etc.). They could be binary (on the left, far from, etc.) or ternary (between [3]), and rarely quaternary (neighborhood).
2. The proportion (partially, completely, etc.) which are also vagueness.
3. Uncertainty: perhaps, certainly.

Some of these relations imply dependences between objects [15]. As an example *Objects A and B are around C* doesn't implies constraints between A and B but in the formulation *Objects A and B are on both sides of C* implies a dependences between the location of A and B.

Among all these characterizations of the spatial relationships, the strict relative position such as *touch, intersect, within or disjoined* are useful to solve conflicts. Other relations such as *close to* or *perhaps* don't give enough information to solve conflicts.

V. SOLVING CONFLICTS

Now let us look at the way to use the conceptual layer to solve conflicts between different representations of the same topological elements (objects or boundaries).

Firstly, it's easy to compute the computer level. We only have to describe each spatial relation between objects of the scene. Secondly, it's easy to compare the computer level with the geometric level and to localize the conflicts (as an example Object A intersects Object B in the computer level and Object A touches Object B in the geometric level). This could be done with the concepts or instances. Because most of the objects are localized with a relative good accuracy, conflicts between the two levels will not concerned relations such as *far from* instead of *close to*. Most of the conflicts will be generated by close objects having touching or merged boundaries in the geometry level and intersected one in the computer level.

Thirdly, we have to solve conflicts leading to unwanted intersections when combining layers. The aim is not to change the representation of the objects and to modify the layers but only to intervene during the combination (union or intersection) of the layers. So we built the results of the combination by considering the geometry level and by choosing a unique representation for the objects concerned by the conflict. The choice is done by considering the domain and accuracy of each layer. The priority is given to the layer close to the domain of the expert using the GIS or to the more accurate one. Indeed, as explained earlier in this paper, the same object or boundary could be described at different scales leading to matching problems.

VI. Conclusion and future work

This paper present the specifications for a conceptual layer shared between information layers in a GIS. The main objective of this layer is to propose an intermediate step to solve the matching problem of two corresponding objects or boundaries during the combination of different information layers. Indeed, if two vertices representing the same topological or semantic objects don't match, the union or intersection of the layers will lead to inexistent objects.
By referring to the conceptual layer, which regroups both concepts and spatial relationships, we will fuse the two different representations of the same object and choose a unique representation according to the application.

At present, the proposed framework is only at the specification step, but every technological ways to implement it have been studied and the implementation has already started and will be presented in a future paper.

References

[1] F. Fonseca, M. A. Rodríguez, and S. Levashkin, "Building geospatial ontologies from geographical databases", In Proceedings of the International Conference on GeoSpatial Semantics (GeoS 2007), Springer Lecture Notes in Computer Science, vol. 4853, pp. 195–209, Berlin, 2007

[2] M. Baglioni et al., "Ontology-supported Querying of Geographical Databases", In Transactions in GIS, vol. 12, issue s1, pp. 31–44, december 2008

[3] I. Bloch et al., "Modeling the spatial relation between for objects with very different spatial extensions", In Proceedings of Reconnaissance des Formes et Intelligence Artificielle (RFIA'06), Tours, France, Janvier 2006

[4] E. Clementini and R. Laurini, "Un cadre conceptuel pour modéliser les relations spatiales", In Revue des Nouvelles Technologies de l'Information (RNTI, vol. 14, pp. 1-18, novembre 2008

[5] I. Cruz, W. Sunna, and K. Ayloo, "Concept-level matching of geospatial ontologies", In Proceedings of GISPlanet, Lisbon, Portugal, 2005

[6] S. Derivaux, G. Forestier, C. Wemmert et S.Lefèvre, "Extraction de détecteurs d'objets urbains à partir d'une ontologie", In proceedings of Extraction et Gestion de Connaissances (EGC), vol. 2, pp. 71-81, Sophia, 2008

[7] F. Fonseca, M. Egenhofer, P. Agouris, and G. Câmara, "Using ontologies for integrated geographic information systems", In Transactions in GIS, vol. 6, pp. 231–57, 2002

[8] F. Fonseca, J. Martin, and A. Rodríguez, "From geo to eco-ontologies", In Proceedings of the International Conference on Geographic Information Science, pp. 93–107, Boulder, Colorado, 2002

[9] F. Fonseca, G. Camara, and A. M. Monteiro, "A framework for measuring the interoperability of geo-ontologies", In Spatial Cognition and Computation, vol. 6, pp. 309–31, 2006

[10] G. Forestier, S. Derivaux, C. Wemmert et P. Gançarski, "Interprétation d'images basée sur une approche évolutive guidée par une ontologie", In proceedings of Extraction et Gestion de Connaissances (EGC), vol. 2, pp. 469-474, Sophia, 2008

[11] G. Forestier, S. Derivaux, C. Wemmert et P. Gançarski, "An Evolutionary Approach for Ontology Driven Image Interpretation", In Proceedings of the International Conference on Applications of Evolutionary Computation (Evo'08), Springer, Lecture Notes in Computer Sciences, vol. 4974, pp 295--304, Italy, 2008

[12] M. Gahegan et al., "A Platform for Visualizing and Experimenting with Measures of Semantic Similarity", In Ontologies and Concept Maps, Transactions in GIS, vol. 12 issue 6, pp. 713-732, 2008

[13] E. Grandchamp, "GIS information layer selection directed by remote sensing for ecological unit delineation", In IGARSS, 2009

[14] E. Grandchamp, "Raster-vector cooperation algorithm for GIS Application to ecological units delineation", In GeoProcessing, 2010

[15] H. W. Guesgen and J. Albrecht, "Imprecise reasoning in geographic information systems", In Fuzzy Sets and Systems, vol. 113, pp. 121-131, 2000

[16] L. Gutierrez, "WP 9: Case study eGovernment D9.9 GIS ontology, "http://dip.semanticweb.org/documents/D9-3-improved-eGovernment.pdf ", last access : 17/01/2011, 2006

[17] C. Hudelot, J. Atif, and I. Bloch, "Ontologie de relations spatiales floues pour le raisonnement spatial dans les images", In proceedings of Représentation et Raisonnement sur le Temps et l'Image (RTE), 2006

[18] Jones C. B. et al., "Spatial Information Retrieval and Geographical Ontologies: An Overview of the SPIRIT project", In Special Interest Group on Information Retrieval (SIGIR), ACM Press, pp.387 – 388, Finland, 2002

[19] M. Kavouras, M. Kokla, and E. Tomai, "Comparing categories among geographic ontologies", In Computers and Geosciences, vol 31, pp 145–54, 2005

[20] N. E. Maillot and M. Thonnat, "Ontology based complex object recognition", In Image and Vision Computing, vol 26, pp 102–113, 2008

[21] M. A. Rodriguez and M. Egenhofer, "Determining semantic similarity among entity classes from different ontologies", In IEEE Transactions on Knowledge and Data Engineering, vol. 15, pp. 442–56, 2003

[22] N. Schuurman et al.,"Ontology-Based Metadata", In Transactions in GIS, vol. 10, pp. 709–726, 2006

[23] B. Smith, "Ontology and information systems", http://www.loa-cnr.it/Papers/FOIS98.pdf, last access : 17/01/2011, 2001

[24] R. Viegas and V. Soares, "Querying a Geographic Database using an Ontology-Based Methodology", In Brazilian Symposium on GeoInformatics (GEOINFO 2006), pp. 165-170, Brazil 2006

[25] N. Wiegand et al., "A Task-Based Ontology Approach to Automate Geospatial Data Retrieval", In Transactions in GIS, vol. 11, issue 3, pp. 355–376, 2007

[26] P. A. Zandbergen, "Positional Accuracy of Spatial Data: Non-Normal Distributions and a Critique of the National standard for Spatial Data Accuracy", In Transactions in GIS, vol. 12, issue 1, pp. 103–130, 2008

[27] M. Egenhofer and R. D. Franzosa, "Point-Set Topological Spatial Relations", In International Journal of Geographical Information Systems, vol. 5(2), pp. 161-174, 1991