

Les chorèmes pour une visualisation synthétique des connaissances géographiques

Ibtissem Cherni^{1,2}, Karla Lopez¹, Robert Laurini¹, Sami Faiz²

¹ LIRIS – INSA de Lyon
69621 – Villeurbanne Cedex – France
{Ibtissem.Cherni, Karla.Lopez, Robert.Laurini}@insa-lyon.fr

² Faculté des Sciences Juridique, Economique et de Gestion de Jendouba
Avenue de l'UMA - 8189 Jendouba– Tunisie
Sami.faiz@insat.rnu.tn

Résumé

Les chorèmes sont des représentations schématisées du territoire. Ces chorèmes permettent de représenter visuellement des connaissances géographiques alors que jusqu'à présent les chorèmes avaient été conçus manuellement par les géographes en utilisant leurs propres connaissances mentales du territoire. C'est dans ce contexte qu'un projet ait été lancé afin de découvrir automatiquement les caractéristiques spatiales et de les visualiser en se fondant sur une base de données spatiales et les chorèmes. Cet article présente le projet international en mettant l'accent sur le sous-système de génération du langage de description des chorèmes appelé ChorML et qui sera par la suite une entrée pour le sous-système de visualisation des chorèmes.

Abstract

Chorems are schematic representations of territory and they can represent visually geographical knowledge. Until now chorems were made manually by geographers using their own mental knowledge of the territory. So, a new project has been launched in order to automatically discover spatial patterns and layout chorems by relying on a spatial database. This article presents this project and stresses the sub-system generation for ChorML which can be seen as an input for the subsystem chorem visualization.

Mots clés : chorème, base des données spatiales, système de génération, système de visualisation des chorèmes, ChorML.

Keywords : chorem, spatial database, subsystem generation, subsystem chorem visualization, ChorML

1. Introduction

Alors que les données géospatiales sont des informations portant sur des objets et événements situés sur la surface terrestre, les chorèmes [BRU 86] sont une représentation schématique des territoires et ils représentent la structure et l'organisation de ces territoires. Pour ces raisons, ils semblent constituer une solution intéressante pour l'aide à la décision spatiale. Dans ce cadre, un projet international entre la France, le Mexique et l'Italie vise à définir des solutions cartographiques capables de représenter les informations issues de fouille des données géographiques selon les chorèmes.

Cet article a pour objectif la présentation de notre projet en mettant l'accent sur le langage de description des chorèmes que nous avons baptisé ChorML, ainsi que son système de génération que nous viendrons de développer. Dans la première section, nous définissons les chorèmes à travers un exemple, ensuite dans la deuxième section nous présentons le projet ChEVIS. Dans la troisième section nous décrivons le langage ChorML et par la suite dans la dernière section est dédiée au système de génération du code ChorML.

2. Les chorèmes

Un chorème [BRU 86] est une représentation schématique d'un espace géographique avec des caractéristiques importantes que nous souhaitons représenter sur une carte et mettre en évidence pour une étude ou une meilleure compréhension.

Un pattern géo graphique [DEL 09] ou pattern est une régularité intéressante d'un certain phénomène découvert dans une base des données géographiques. Les patterns peuvent être utilisés comme point de départ pour la localisation des phénomènes spatio-temporels et des relations entre eux, comme le montre la figure 1.

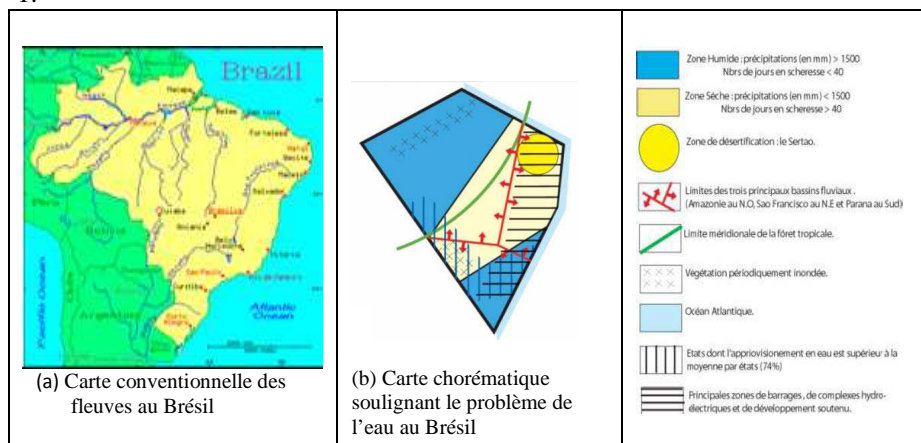


Figure 1. Carte conventionnelle des fleuves au Brésil (a) et carte chorématique soulignant le problème de l'eau dans ce pays(b) [ANONYMOS 06]

L'intérêt de cette carte chorématique est de montrer d'une manière claire la grande problématique de l'eau dans ce pays.

3.

4. Présentation générale du projet ChEVIS

Le projet ChEVIS de recherche entre trois instituts de recherche : L'Institut National des Sciences Appliquées (INSA) de Lyon (France), L'Université di Salerno (Italie) et le Tecnológico de Monterrey, Campus de Puebla (Mexique) afin d'extraire les chorèmes à partir de la fouille des données géographique. En d'autres termes, les connaissances extraites ne seront pas données en logique descriptive, mais par des représentations visuelles.

La figure 2 montre l'architecture du système ChEVIS (Chorem Extraction and Visualisation System) qui est composé de deux sous systèmes : Le système d'extraction des chorèmes, appelé Chorem Extraction System et le système de visualisation des chorèmes, appelé chorem Visualization system.

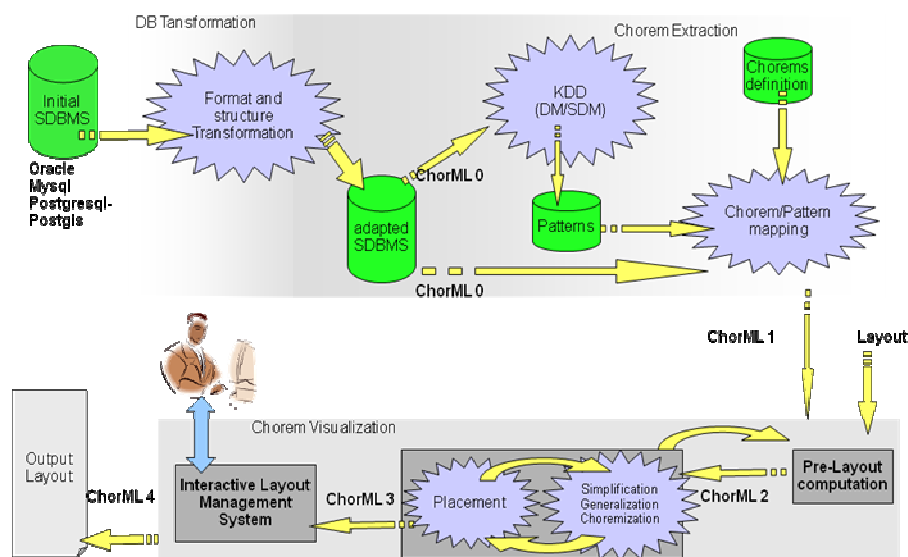


Figure 2. Architecture du système [ANONYMOS 08]

Le système d'extraction des chorèmes

Le système d'extraction des chorèmes [DEL 09] permet de transformer une liste de chorèmes, venant du système d'extraction en termes de texte, dans une représentation visuelle, ou bien dans une carte chorématique. Dans une première phase, un dictionnaire des chorèmes est importé dans la base de données pour être employé dans la phase suivante du processus d'extraction des chorèmes. Par la suite, le système exécute un processus d'association entre les patterns et les chorèmes,

ainsi qu'un processus de réduction, afin d'obtenir un ensemble restreint de chorèmes. En effet, du processus d'extraction, une grande quantité de patterns peuvent être obtenus, en risquant ainsi de surcharger d'un nombre excessif de chorèmes de la carte chorématique en cours de construction. Les deux phases permettent la participation des utilisateurs afin de sélectionner les patterns et les chorèmes les plus significatifs. Finalement, le système exécute le calcul des relations spatiales existantes parmi les éléments de chorème. Dans ce but, des fonctions d'Oracle Spatial sont utilisées par le système et les résultats sont codés en ChorML, qui sera ensuite envoyé au système de visualisation [ANONYMOS 08].

3.1. Le système de visualisation des chorèmes

Le système de visualisation des chorèmes [DEL 09] permet de transformer une liste de chorèmes dans une représentation visuelle, ou bien dans une carte chorématique ; deux phases différentes sont développées par ce système, la création des chorèmes et leur modification.

La phase de création des chorèmes a été développée à travers un système multiagents qui exécute trois phases : chorem Drawing (dessin des chorèmes), le Coordinate Translation (traduire les coordonnées) et le Best Placement (meilleure position), cette dernière phase est effectuée à travers deux opérations :

- L'opération de simplification qui détermine une version simplifiée des données géométriques, en réduisant le nombre de sommets de la forme originale. Cette technique est basée sur des fonctions spatiales, qui incorporent l'algorithme de Ramer-Douglas-Peucker (RDP) et ses variantes [RDP 73].
- L'opération de chorémisation qui conditionne la simplification en associant à la composante spatiale du chorème déjà simplifié, la forme d'un polygone régulier avec un faible nombre de côtes.

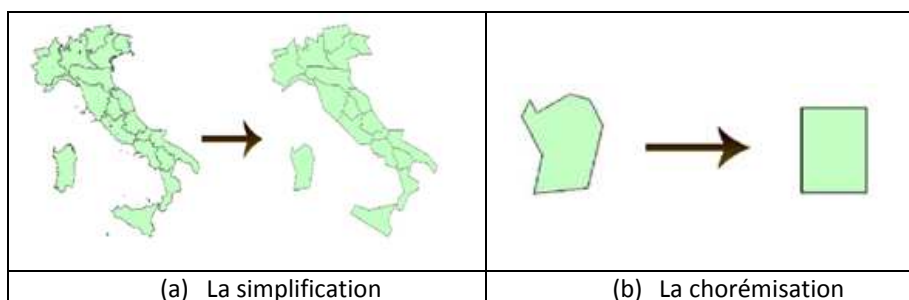


Figure 3. Les deux opérations de création des chorèmes [DEL 09]

Finalement, la phase de modification des chorèmes fournit aux utilisateurs la possibilité de modifier la carte chorématique.

4. ChorML : ChoreM Markup Language

Dans la section précédente nous avons exposé le système d'extraction des chorèmes, ces derniers sont codés à travers un langage appelé ChorML.

Le ChorML [COI 08] est un langage de type XML, multi-niveaux, avec la fonctionnalité de mémoriser l'information des chorèmes et de permettre la communication de telles informations entre les différents modules du système, ces niveaux sont décrits comme suit :

- *Le niveau 0* est composé de XML et GML [http1]. Il a été pensé pour stocker les informations sur les proto-chorèmes, l'origine des données et les fonctions appliquées afin d'obtenir les chorèmes.
- *Le niveau 1* de ChorML est également une combinaison de XML et de GML. Il spécifie les résultats des algorithmes de fouille des données spatiales. Les éléments du langage sont en particulier :
 - Les informations de caractère général, contenant l'identificateur de la carte, le nom de la carte, le nom de l'auteur, la date de création, le système de référence, le nom de la base des données originale, la dernière date de mise à jour, etc.
 - La liste des chorèmes dans lequel les données géographiques sont codées en GML.
 - Une pré-légende, qui contient une description en format texte de chaque chorème.
 - La liste des relations topologiques et non-topologiques parmi les chorèmes.
- *Le niveau 2* de ChorML est une combinaison de XML et de SVG [ANT 06]. Les éléments du langage sont en particulier :
 - Les informations de caractère général, contenant l'identificateur de la carte, le nom de la carte, le nom de l'auteur, la date de création, le système de référence, le nom de la base de données originale, la dernière date de mise à jour, etc.
 - La liste des chorèmes simplifiés résultant de la phrase de ChoreM Création.
 - La liste des chorèmes résultant de l'opération de ChoreM Editing.
 - Une légende qui contient une description en format texte et visuelle de chaque chorème.
 - La liste des relations topologiques et non-topologiques parmi les chorèmes.

La figure 4 montre la structure générale d'un document ChorML.

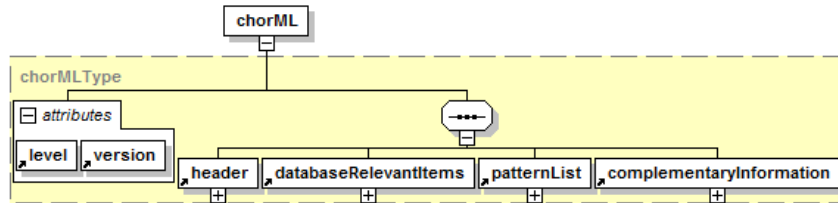


Figure 4. La structure de ChorML [COI 08]

Dans le tableau 1, nous donnons un exemple d'évolution selon les niveaux de ChorML.

Niveau	Représentation
Niveau 0	Point de longitude / latitude, représentant une ville avec sa population.
Niveau 1	Point de longitude / latitude et son importance s'il est choisi comme patterns
Niveau 2	Point avec coordonnées en pixel, représenté par un cercle de rayon et de couleur.

Tableau 1. Présentation d'un point à travers les différents niveaux de ChorML

5. Génération automatique de ChorML

Nous allons illustrer les différents composants de notre application. En fait, cette application a pour finalité la génération d'un document XML selon la spécification ChorML, et ce, en partant du résultat d'une requête SQL et PL/SQL exécuté sur une base de données Oracle.

Le système de génération des documents ChorML est composé de quatre modèles comme le montre la figure 4 : DB Connexion, SQL Generator, Data Manipulator, ChorML Generator.

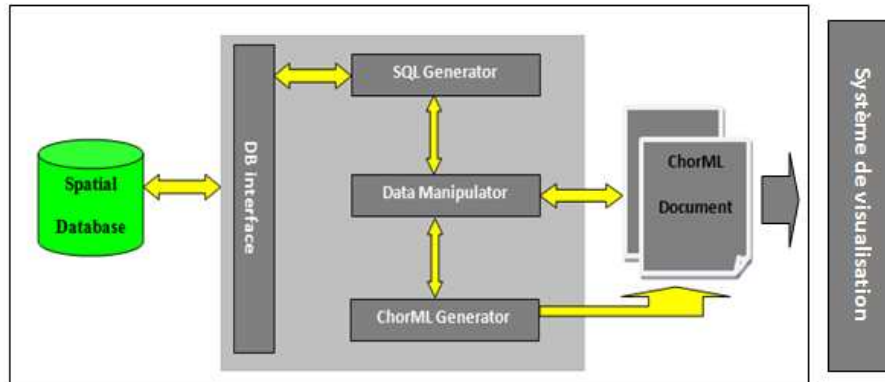


Figure 4. Architecture du ChorML generator

Dans ce qui suit nous allons présenter rapidement les quatre modèles de notre système :

- **DB Connexion**

Ce module a pour finalité d'assurer la communication entre notre application et le SGBD utilisé (Oracle dans notre cas), et ce du moment où on lance la requête jusqu'au moment où l'on récupère le résultat. Ce module assure également le bon déroulement de l'exécution de notre connexion.

- **SQL Generator**

C'est à travers ce module que les requêtes seront lancées et exécutées. En fait SQL Generator assure :

- La préparation et l'exécution de la requête SQL appropriée
- La récupération du résultat et stockage des données dans une structure bien déterminée.

- **Data Manipulator**

La tâche principale de ce module est de récupérer le résultat du SQL Generator et d'injecter les données dans le module ChorML Generator ; il permet aussi de faire des modifications sur le ChorML déjà généré.

Il est composé de trois modules :

- Une méthode qui permet de récupérer un résultat d'une requête d'un type donné (sql/plsql) afin d'appeler les méthodes nécessaires au module ChorMLGenerator, elle prend comme paramètres une requête, un nom du fichier pour stocker le résultat de génération et un type de requête.

- Une méthode qui permet de récupérer le résultat d'une requête cluster afin d'appeler les méthodes de génération de résultats correspondantes au type de requête.
- Une méthode de génération de contraintes qui seront intégrées dans le fichier final ChorML.

- **ChorML Generator**

Ce module a pour objectif la construction d'un document XML selon la spécification ChorML. Ce module est un composant essentiel dans l'application développée car il permet la génération de document XML selon la spécification ChorML, il contient:

- Un module qui permet de générer un document XML selon la spécification ChorML d'un élément simple ou d'un cluster ou région.
- Un module qui soit le plus important, c'est celui qui fait la transformation d'un élément de type SDO_GEOMETRY en ChorML en fonction de type de ce champ récupéré de la base de données.
- Un module qui permet de générer un document XML qui représente les contraintes ChorML.
- Une méthode qui permet la conversion de résultat des coordonnées sous forme d'une chaîne de caractères.
- Un module qui permet de stocker des requêtes SQL sous forme d'un document XML.
- Une méthode qui permet de récupérer le contenu d'un fichier donné sous forme d'une chaîne de caractères.
- Une méthode qui permet la génération complète d'un document ChorML, elle permet de fusionner les documents générés « l'entête » et « les informations complémentaires ».

5.1. Etat de l'implémentation

Dans cette section, nous proposons quelques étapes de la génération d'un document ChorML par notre système ChorML Generator : La génération des frontières, la génération des relations topologiques et la génération du document ChorML final.

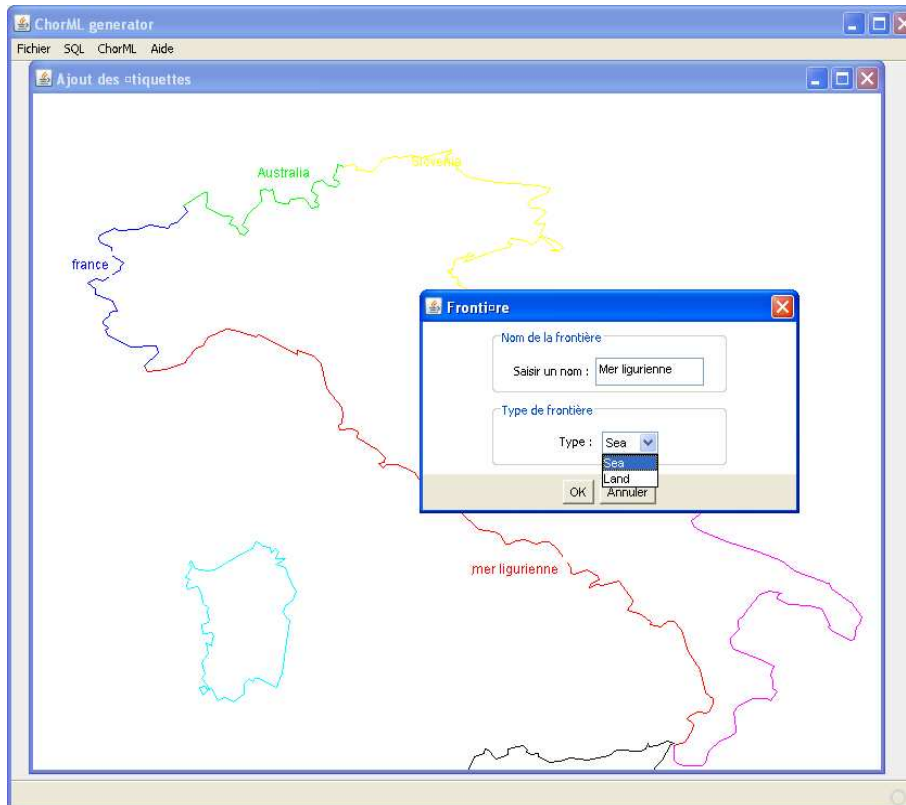


Figure 5. ChorML generator : génération des frontières

Rares sont les bases de données géographiques qui décrivent rapidement l'extérieur de leur territoire de juridiction. Afin d'assurer une plus grande intelligibilité des chorèmes, il a été décidé d'ajouter des informations extérieures comme les noms des mers, des pays voisins, etc.

La figure 5 montre l'interface de génération des frontières d'un pays à partir d'une interface qui permet de valider le nom de frontière et son type (mer, terre) aussi l'utilisateur doit avoir le choix d'ajouter des frontières régionales de la même manière que pour les frontières. Les frontières régionales sont les frontières entre un pays et une région d'un autre pays.

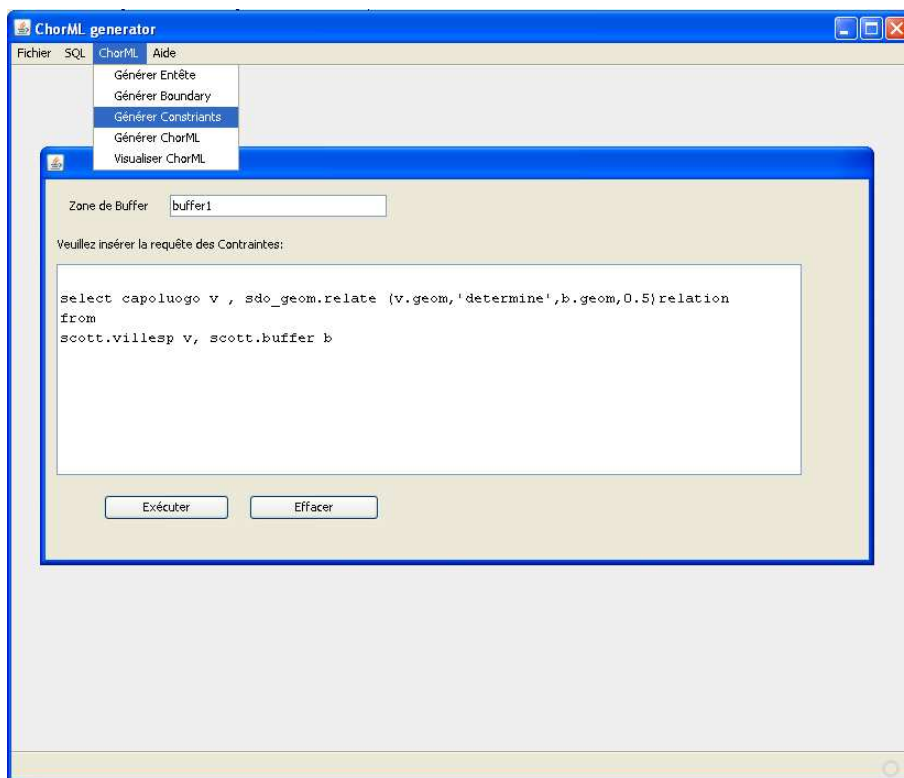


Figure 6. ChorML generator : génération des relations topologiques

La génération des relations topologiques se fait par l'exécution de deux requêtes spatiales : SDO-GEOM_RELATE et SDO-BUFFER, ces deux requêtes vont donner les relations topologiques entre les différents types géométriques. L'interface de génération des relations est montrée par la figure 6.

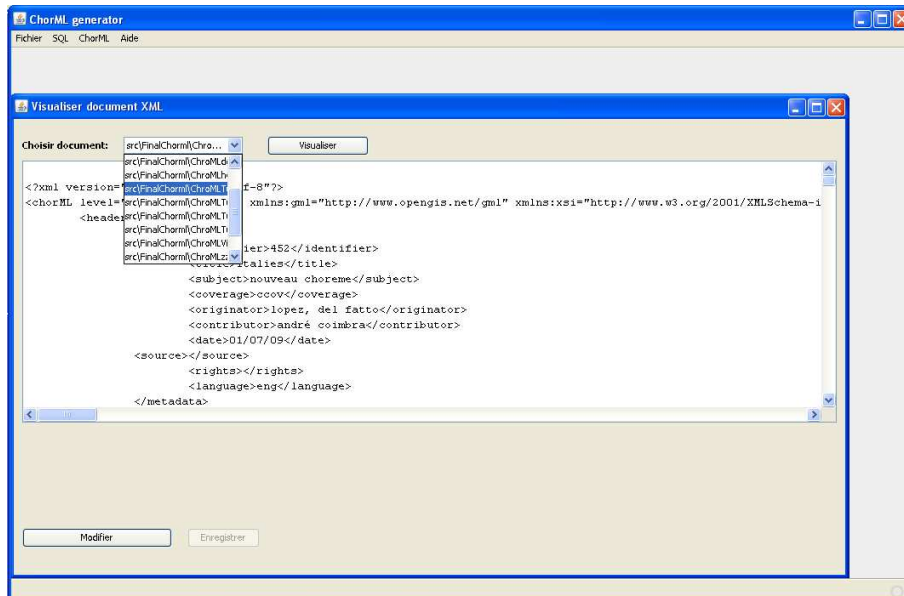


Figure 7. ChorML generator : génération et visualisation des documents ChorML

La figure 7 montre la génération d'un document ChorML qui est le résultat de fusion de trois documents XML : l'entête, la requête et les informations complémentaires.

3.1. Validation des résultats

Après avoir généré les documents ChorML, il faudrait disposer d'un outil pour valider les résultats obtenus par notre système; c'est pourquoi nous avons décidé de transformer ces résultats en KML (Keyhole Markup Language) [http2] qui est un langage basé sur le formalisme XML et destiné à la gestion de l'affichage de données géospatiales dans les logiciels Google Earth, Google Maps, Google Mobile et World Wind.

Dans la figure 8 nous présentons la visualisation d'une requête de pattern et de frontière ayant été transformée en KML et puis visualiser avec Google Earth.



Figure 8. Visualisation avec Google Earth

Conclusion

Cet article présente notre projet ChEVIS qui permet de définir des solutions cartographiques capables de représenter convenablement les informations extraites d'une base des données géographique. La solution proposée se fonde sur le concept de chorèmes et sur sa capacité de synthétiser des scènes qui comprennent des objets géographiques et des phénomènes spatio-temporels en leur associant des notations visuels schématisées.

Références

- [ANT 06] **Antoniou, B , Tsoulos,L.** "The Potential of XML encoding in geomatics converting raster images to XML and SVG", 2006, Computer and Geosciences 32.
- [BRU 86] **Brunet R ,** "La carte-modèle et les chorèmes", 1986 , Mappemonde 86/4 pp. 4-6.
- [COI 08] **André Rocha Coimbra** "ChorML: XML Extension For Modeling Visual Summaries of Geographic Databases Based on Chorems", Master project, 2008.
- [DEL 09] **VinCenzo Del Fatto,** "Visual Summaries of Geographic Databases by Chorems", Ph.D. Thesis, 2009, INSA-Lyon, University of Salerno.
- [Anonymos 06] **Laurini R., Milleret-Raffort F., Lopez K.** "A Primer of Geographic Databases Based on Chorems", 2006, In proceedings of the SebGIS Conference, Montpellier, Published by Springer Verlag LNCS 4278, pp. 1693-1702.

[Anonymos 08] **Laurini R.**, “*Visual Summaries of Geographic Databases*”, 2008 , 14th Int'l conf. On Distributed Multimedia System, Invited speaker.

[RDP 73] **Douglas D., Peucker T.** (1973) "*Algorithms for the reduction of the number of points required to represent a digitized line or its caricature*", The Canadian Cartographer 10(2), 112-122.

[http1] The Open Geospatial Consortium: Geography Markup Language (GML), <http://schemas.opengis.net/gml>

[http2] Keyhole Markup Language (KML); <http://code.google.com/apis/kml/documentation/>