A Primer of Geographic Databases Based on Chorems

Robert Laurini, Françoise Milleret-Raffort, and Karla Lopez

LIRIS, INSA de Lyon, F-69621 Villeurbanne
Robert.Laurini@insa-lyon.fr, Francoise.Raffort@insa-lyon.fr,
Karla.Lopez@insa-lyon.fr

Abstract. The goal of this paper is not to present outcomes of research, but rather present a new research plan in the use of chorems in geographic information systems. Created by R. Brunet, chorems are a schematic representation of a territory. Presently, geographic decision-makers are not totally satisfied by conventional cartography, essentially because they want to know where and what are the problems. And so chorems appear as an interesting approach to unveil geographic problems, and so to help decision makers understand their territory, their structure and their evolution. After having given the definition and presented some applications of chorems, we show how chorems can be discovered by spatial data mining, can help decision making, and also how chorem maps can be a novel approach to visually entry a geographic database or datawarehouse. Comparing with the Ben Shneiderman's approach, chorems can give an overview of the territory; then by zooming and filtering, and sub-chorem maps can be generated for smaller territories. Finally a list of barriers to overcome is given as main landmarks for a new research program in order to design new kind of geographic information systems or spatial decision support systems.

Keywords: chorems, spatial data mining, geographic databases, geographic datawarehouses, geographic knowledge, GIS, visual entry system, research plan.

1 Objectives and Organisation of the Paper

Chorems were created in 1980 by Pr. Roger Brunet, a French geographer. They are a schematic representation of a territory. This word comes from the Greek chôra which means space, territory. It is not a raw simplification of the reality, but rather aims at representing the whole complexity with simple geometric shapes. Even if it looks a simplification, the chorem tries to represent the structure and the evolution of a territory with a rigorous manner.

The basis of a chorem is in general a geometric shape in which some other shapes symbolize the past and current mechanisms. Brunet has proposed a table of 28 elementary chorems, each of them representing an elementary spatial configuration, and so allowing them to represent various spatial phenomena at different scales. According to Brunet, chorems are a tool among other to model the reality, but it is a very precious tool not only as a visual system, but also as a spatial analysis tool.

R. Meersman, Z. Tari, P. Herrero et al. (Eds.): OTM Workshops 2006, LNCS 4278, pp. 1693 – 1702, 2006.

As an introductive example, let us mention the water problem in Brazil[1] as depicted in Figure 1. In Figure 1a, there is a conventional map of rivers in Brazil; as usual, at this scale only the main rivers are mapped, and this map will not be very useful for decision making. However, nothing indicates what and where the problems are. However, Figure 1b gives a chorem map of the situation according to the caption given Figure 1c. We can see more accurately where the humid and dry zones are, where water is missing, where dykes are located, where water is more demanded and so on. As a conclusion, this kind of drawing is much more informational to any decision-maker than the conventional river map.

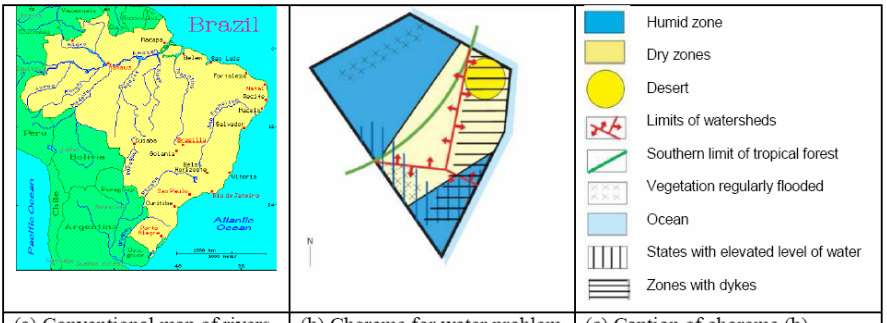

Fig. 1. Comparing a conventional map of rivers in Brazil and a chorem map emphasizing the water problem in this country

The objective of this paper is to give the first outlook of a research program on using s for geographic decision-making. After a more accurate definition of chorems, we will describe how chorem map can assist decision-makers, and more important, we will describe the links with spatial data mining.

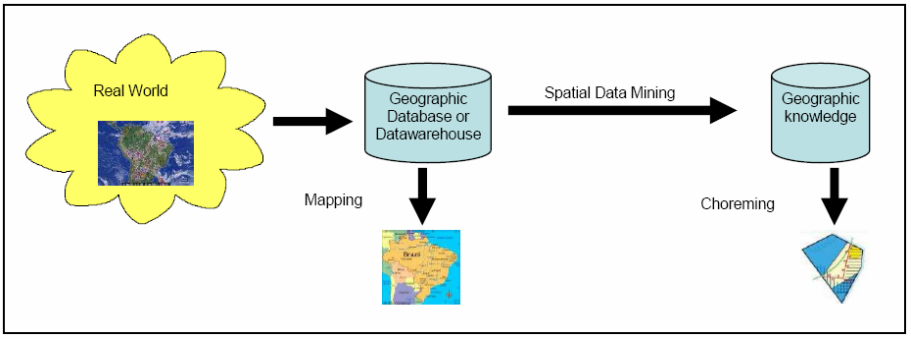

Fig. 2. Main steps of the proposed methodology

[1] This example is drawn from the web site http://histoire-geographie.ac-bordeaux.fr/espaceeleve/bresil/eau/eau.htm

Then, we will present how chorem map can help create a novel progressive entry system for geographic databases following Ben Shneiderman's mantra. We will finish this paper by emphasizing the barriers to be overcome to reach these objectives. The ultimate goal will be to create the main concepts in order to implement a decision support system which can be schematized as follows (Figure 2): starting from the real world, a geographic database or datawarehouse will be constructed and populated. Them by applying appropriate spatial data mining and filtering techniques, geographic knowledge will be discovered and then visualized as chorems (Lopez, 2006).

2 What Are Chorems? How Can They Help Decision-Making?

2.1 Elementary Chorems

As initially told, a table of 28 chorems were created by Roger Brunet (1986), Figure 3 gives this table (in French) representing the basic chorem vocabulary.

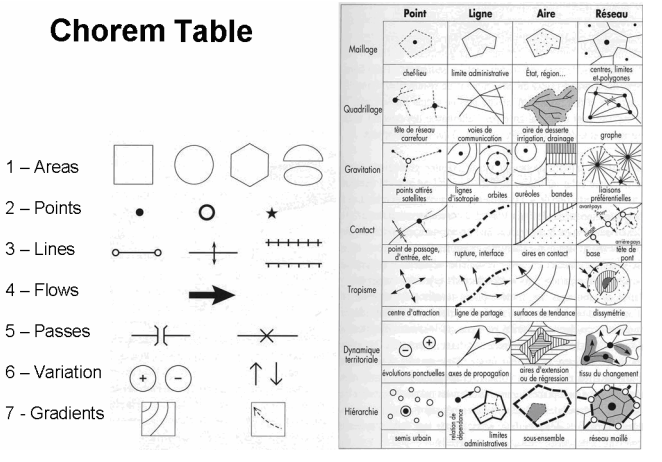

Fig. 3. Table of Brunet's chorems (in French)

Finally chorems can be defined as a visual vocabulary allowing the description of salient characteristics and problems of a territory. From our point of view, chorems are a nice basis for decision making, because they emphasize the more significant aspects, leaving apart secondary problems. Let us say that when it is necessary to understand the structure of a territory, a one-to-one map is not useful, whereas a small schema can be more useful. So chorem maps are a key-tool to schematize a territory, and then after, decision-makers or politicians can get a clearer view of the situation. Among applications, let us mention the schematic visualization for:

- salient political, economic and demographic problems,
- salient features in environment and climatology,
- main evolution in epidemiology,

- natural and technological risks or disasters,
- etc.

Regarding other applications, let us mention a very interesting approach to use chorems for way finding (Klippel et al. 2005).

2.2 Approaches, Manual Versus Automatic?

In the state of the art review, it appears that there are only manual approaches of discovering chorems. By analyzing the examples of Bolivia (Arreghini, 1995) given Figure 4, and Zaire (Bruneau-Simon, 1991), for a whole country, the typical manual approach for a global diagnostic can be as follows:

- Relief and climatology,
- Ecosystems, environment,
- History, population and demography,
- Rural and urban dynamics,
- Communication networks,
- Economy and international relationships.

However, when research must be done on some more specific problems, such as the water problem for Brazil, or politics in Switzerland, some other steps must be defined.

As far as we know, no automatic approaches seem to exist. The objective of this project is to set up a methodology based on spatial data mining. Doing that, chorems appear as geographic knowledge which must be visualized. In other words, we can define now chorems as a kind of visual geographic knowledge, whereas knowledge is information useful for problem solving.

To be more precise, at the discovery level of chorems, geographic knowledge is chased. Once it is discovered, it must be not exactly mapped, but rather visualized. That is to say that some layout procedure must be implemented in order to arrange the appropriate elementary chorems. During this step, perhaps some spatial knowledge must be used to perform the layout, for instance by using non-overlaps, spatial organization of elementary diagrams, etc.

Spatial data mining could be an interesting approach to discover chorems. Data mining can be defined as a systematic way of analyzing data stored in databases or in datawarehouses and of extracting patterns, i.e. useful information for decision makers. And spatial data mining (Ester et al. 1997) is the extension to the study of spatial data.

According to Pech-Palacio (2005), there are several kinds of spatial data mining or knowledge discovery techniques, such as aggregation, clustering, classification, trend detection, etc.; and he has proposed a new way based on graphs. Since geographic knowledge is hidden by coordinates, a way is to transform the geographic database content (perhaps limited to the territory under consideration) into a huge graph which is mined by the SUBDUE algorithm (http://cygnus.uta.edu/subdue/) developed by the University of Texas at Arlington (Holder-Cook, 2005).

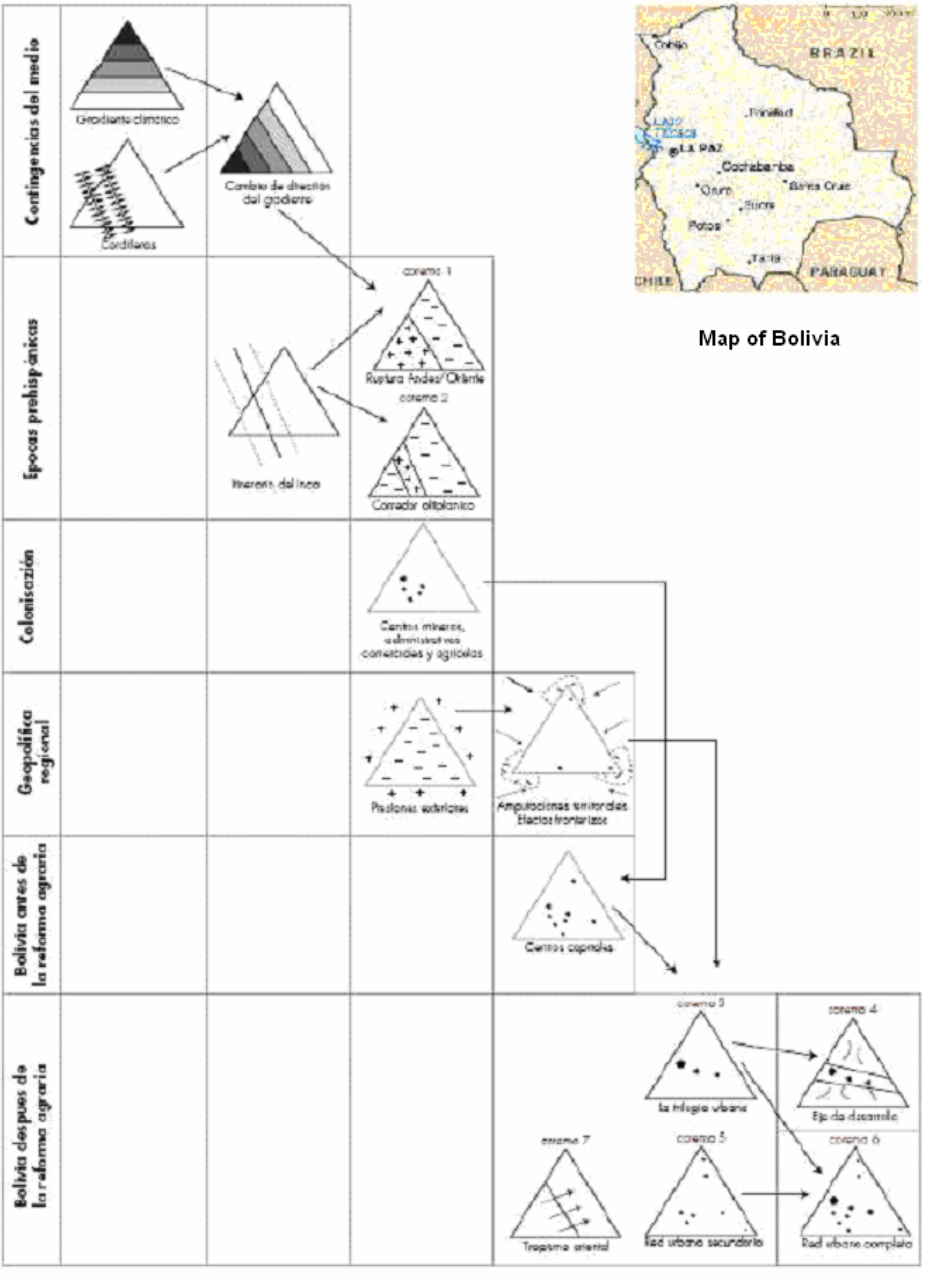

Fig. 4. Example of approach for designing chorems for Bolivia (Arreghini, 1995)

The big advantage of the Pech-Palacio's approach is to be a global approach, i.e., it considers not only one class of geographic objects as done by several spatial data mining techniques, but rather all kinds of geographic objects. By using this method, geographic knowledge will be discovered.

The next step will be to select the more salient knowledge. For that, a list of criteria must be set up to filter all that knowledge until reaching to a dozen of geographic knowledge which can be represented as chorems.

3 Chorems as a New Way to Access Geographic Databases

Another point of view of chorems lays on Ben Shneiderman's mantra for designing human interfaces "*Overview, zoom and filter, details on demand*" (Shneiderman 1997a, 1997b), i.e. macroscopic versus microscopic approach. Indeed, we can state that chorems can be an excellent candidate at "overview" level when studying a territory.

Indeed, for conventional databases, approaches such as starfield or space filling treemaps were created for relational or object-oriented databases. The starfield system is targeted to layout instances of a database object or a relation into a screen: a procedure is given for selecting the two axes from attributes, and then a third axis is selected for colours; the result is called a starfield. The more known example is the starfield system made for Hollywood movies (Ahlberg, Shneiderman 1994). For databases with different objects, another metaphor is used based on so-called space filling treemaps; personally, we would prefer to name this approach the "bookshelf" metaphor.

Back on geographic databases and datawarehouses, the chorems approach can have a similar target. In this case the territory chorem gives an overview of the situation, whereas the "details on demand" step can be represented by a detailed mapping. And by "zooming and filtering", we can gracefully and gradually reduce the search space. Here zooming will mean using different geographic scales or thematic disaggregation, whereas filtering reflects conditions and criteria (geographic and semantic zooming). By zooming and filtering, a sort of sub-chorem can be defined. By sub-chorem, we mean a chorem made for a smaller territory. For instance, the first step can be a chorem for a whole country, then chorems for regions and so on.

In other words, chorems can be seen as a new way to enter geographic databases. Table 1 gives a comparison between the use for conventional databases, and the approach to geographic databases and datawarehouses. And Figure 5 schematized the comparison of various styles of database entry systems.

So, a new way of entering a geographic database can be sketched:

1 – at the opening, a global chorem map can be displayed

2 – then by semantic and geographic filtering some sub-chorem maps can be visualized

3 – finally, the final query answer (map or table) can be displayed.

Table 1. Comparing accesses to conventional and geographic databases

Ben Shneiderman's mantra	Conventional databases	Chorem-based approach
Starting point	Relational or object-oriented database of an organization	Any kind of data which can be useful
1 – Overview	Generally the "overview" is visually presented by means of starfield or space filling treemaps; they are both structure- and content-oriented.	The territory-level chorem can give an overview, perhaps more linked to problems than to data contents.
2 – Zoom and filter	Criteria can be used to reduce the search space.	The territory can perhaps be split in different zones, each of them with a sub-chorem (geographic zoom). A second way can be to reduce the number of topics (semantic zoom)
3 – Details on demand	The final step delivers what could be necessary for the user, usually as a table.	Here both tables and maps can be the final steps, depending on the user's needs.

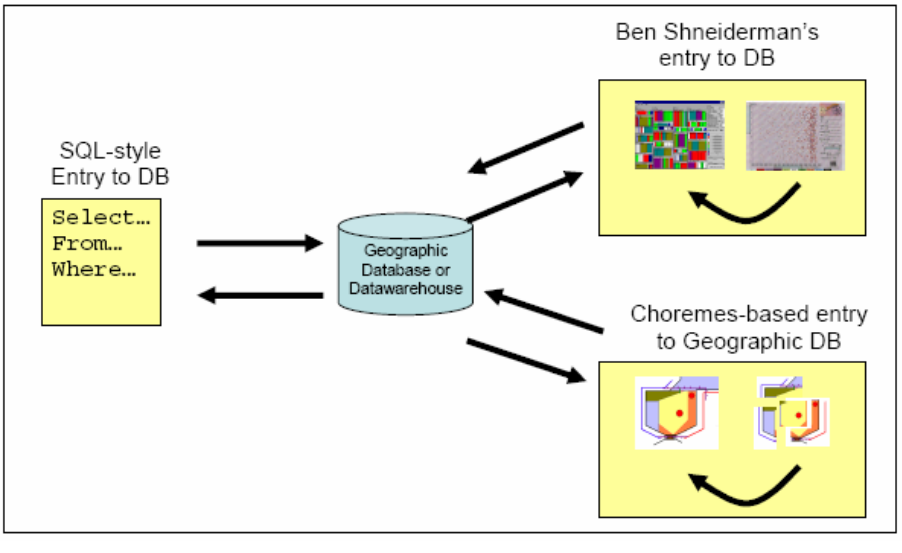

Fig. 5. Comparing various styles of database entry systems

4 Chorems as Visual Representation of Geographic Knowledge

It seems useful to distinguish spatial knowledge from geographic knowledge. "Spatial Knowledge" corresponds to relations which are common in any kind of space

(whether 2D, 3D, etc.) such as Egenhofer topological relationships and their mathematical consequence; for us spatial knowledge is quite similar to geometric knowledge. Whereas by "Geographic Knowledge" we mean knowledge located in the earth; for instance a pattern such as "when there is a lake and a road leading to the lake, there is a restaurant" can be consider as geographic knowledge.

Finally, we can now define chorem as a visual representation of geographic knowledge. Until now, geographic and spatial knowledge were essentially represented verbally or by using some mathematic tools such as descriptive logics; now a new possibility is emerging based on chorems.

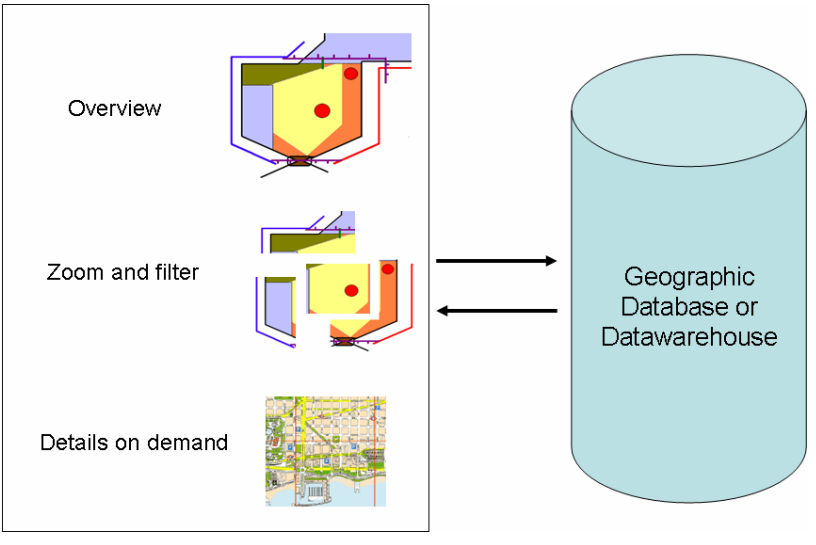

Fig. 6. Chorem-style entry to geographic databases

5 Landmarks for a Research Plan: Barriers to Be Overcome

In order to reach those objectives, the following barriers have already been identified, and must be overcome, each of them corresponding more or less to a future PhD.

1 – Chorem modelling
The modelling of chorems means several things. First as a graphic object, we must define precisely its graphic attributes, where to locate the chorem, etc. Another very important aspect is the modelling of all conditions which govern the chorem. For that, not only conditions must be modelled, perhaps by using descriptive logics, but also a definition language must be defined.

2 – Discovery of salient features
Let us consider the first elementary chorem corresponding to the more important points. What are important points? Maybe there are major cities from a political or an

economic point of view. Do we have to select one, two, five points? What are the limits?

In other words, among all features, we have to discover the more important, that is to say we have to define "importance". Suppose we have got 20 important features. Is it too much or too less? A second problem is the limitation. Too much will lead to some impossibility to understand, too less, the result can be biased by criteria weights.

In other words, once the geographic knowledge is discovered by spatial data mining, a methodology must be created to select the more important geographic knowledge.

3 – Spatial arrangement

Once the more salient features are discovered, the next step is the chorem map. By that, we must organize them spatially not only without overlaps, but also a nice spatial distribution to cover the whole territory, and color selection. Remember that geographic shapes must be generalized (Buttenfield et al. 1991), so implying to store those simplified shapes into the database, or to reconstitute them when necessary. This barrier appears to be overcome by using different agents, each of them using different spatial knowledge.

4 – From chorem map to sub-chorem maps

In connection with the zoom-and-filter phase, the generation of sub-chorem maps must be studied. Two solutions are possible, either to re-begin the chorem phase for a smaller territory, or to reuse the already discovered geographic knowledge to generate the sub-chorem map. Since it is well known that data mining is very consuming, the second approach looks more interesting.

5 – From chorem discovery to chorem-based access

Whereas chorem discovery in a territory can be seen as one task, the use of chorems to access a database is a little bit different. To accelerate the access of the geographic databases or the datawarehouses, chorems and sub-chorems can be stored and used when necessary as a visual entry system. For that an appropriate GUI must be created by following Ben Shneiderman's mantra.

6 – Interoperability

Once the system is made, it must work for any geographic database. Two solutions can be envisioned: the first one is to transform manually the whole content of the existing geographic database to match the structure and the names used in the chorem system; an alternative can be the definition of views as in relational databases. The second solution is to provide an interface to allow interoperability between any geographic database and the chorem system. Here an approach such as a geographic ontology can be a good candidate.

7 – Cognitive aspects

Once a first prototype is made, it will be important to study several kinds of users to test their reactions, their behaviours with such a novel approach, their understanding of the system, their performance, their acceptability, and so on.

6 Final Remarks

The goal of this paper was not to present any result, but rather present some landmarks for a research plan in geographic decision-making based on chorems. Two main aspects were emphasized:

- chorems can be a way to represent geographic knowledge, as a visual outcome of spatial data mining,
- chorems can be a new way to enter geographic databases as a global vision.

Anyhow, this research plan must be detailed, concepts must be clarified, and experimentations must be made through several prototypes.

References

AHLBERG, C., SHNEIDERMAN B. (1994) "*Visual Information Seeking: Tight coupling of dynamic query filters with starfield displays*", Proc. of ACM CHI94 Conference (April 1994), 313-317.

ARREGHINI L. (1995) "*La modélisation graphique dans la réalisation des atlas pour le développement*". Journées Géographiques de l'ORSTOM, 11-12/09/1995, Maison de la Géographie, Montpellier, France.

BERTIN J. (1973). "*Sémiologie graphique*", Mouton/Gauthier-Villars, 2ème édition, 1973.

BRUNEAU J.-C., SIMON T. (1991) *"Zaïre, l'espace écartelé*". Mappemonde 4/91.

BRUNET R. (1986) "*La carte-modèle et les chorèmes*", Mappemonde 86/4 pp. 4-6.

BRUNET R. (1993), "*Les fondements scientifiques de la chorématique*", in "La démarche chorématique", Centre d'Études Géographiques de l'Université de Picardie Jules Verne.

BUTTENFIELD B., McMASTER R., (1991) "*Map Generalization: Making Rules for Knowledge Representation*", Longman, London, 1991.

ESTER M., KRIEGEL H.-P., SANDER J. (1997) "*Spatial Data Mining: A Database Approach*". Proc. of the Fifth Int. Symposium on Large Spatial Databases (SSD '97), Berlin, Germany, Lecture Notes in Computer Science Vol. 1262, Springer, 1997, pp 47-66.

HOLDER L.B., COOK D. J., (2005) "*Graph-based Data Mining*", J. Wang (ed.), Encyclopedia of Data Warehousing and Mining, Idea Group Publishing, 2005.

KLIPPEL A., RICHTER K.-F., HANSEN S., (2005) "*Wayfinding Choreme Maps*". In Proceedings of 8th International Conference, VISUAL 2005, June 2005 Edited by S. Bres and R. Laurini "*Visual Information and Information Systems*". Springer, Berlin. Lecture Notes in Computer Sciences 3736, pp. 94-108.

LOPEZ K. (2006) *"Génération Automatique de Cartes Chorématiques",* Master Thesis in Computing, INSA of Lyon, June 2006.

PECH PALACIO M. (2005) *"Spatial Data Modeling and Mining using a Graph-based Representation*". PhD in Computing, jointly awarded December 12, 2005 by INSA de Lyon, France, and Universidad de las Américas de Puebla, Mexico.

SHNEIDERMAN B. (1997a) "*Information visualization: Dynamic queries, Starfield Dis-plays, and LifeLines*" White paper: http://www.cs.umd.edu/hcil/members/bshneiderman/ivwp.html

SHNEIDERMAN B. (1997b) "*Designing the User Interface*." Third edition. Addison-Wesley Publishing Company. 600pp.

ZEITOUNI K. YEH L., (2000) *"Le data mining spatial et les bases de données spatiales*", Revue internationale de Géomatique, Editions Hermès, Vol. 9, 4 (99), pp 389-423, Avril 2000.