

Semantic Description of Topological Relations in Spatial Databases

Miguel Martinez

Geoprocessing Laboratory-CIC- National Polytechnic Institute, Mexico City, Mexico
miguelrosales@sagitario.cic.ipn.mx,

Abstract. An approach focused on incorporating semantic content into Spatial Databases is proposed. Our methodology is based on a conceptualization of a geospatial domain restricted to recover the meaning of topological relations between geographic objects by means of concepts. Indeed, in spatial databases only a small set of topological relations is explicitly represented. While, a semantically enriched set of such relations may be required, but this sometimes can only be identified at the time when the geospatial data are displayed or analyzed by the user. Thus, we define six relations, which are obtained considering the behavior of diverse themes such as Hydrology, Land Use, Transportation Networks, and Settlements. Geospatial objects are analyzed to identify the topological relationships. We consider two analysis levels: intrinsic and extrinsic. Then, the descriptions are automatically generated in form of tuples $\{O_i, R, O_j\}$, where O_i and O_j represent a pair of geospatial objects, and R represents the concept (relation). Each tuple represents the meaning of a topological relation. For example, a highway (O_1) crosses (R) a roadway (O_2). The conceptual representation has some advantages with respect to the traditional approaches: the conceptualization does not depend on the data scale, geo-reference system, dimension, etc.

1 Introduction

Nowadays, the spatial databases commercially used have a little or null semantic content, great part of this content is represented implicitly in the data and requires being extracted analyzing geographic data. In general, the geospatial data have different properties that cover diverse aspect of geographical data; within these are the topological, geometrical, thematic and logic properties.

Additionally, the topological relationships between geographical objects are not explicitly represented in the spatial databases. Frequently these relationships are identified when the data are displayed or analyzed [12].

Therefore, actually it is necessary that the GIS lead the efforts on investigation to describe the spatial relationships explicitly, by means of objects conceptualization and the relationships that maintain whit other entities in some area of interest. In addition to make use of the semantic to solve problems that traditionally are dealt with numerical or classic processing, for example the interoperability into heterogenic database.

When realize a topological description of spatial database in explicit way, based in concepts that represent the topological relationships, these can be stored explicitly in a spatial database. It's possible to automatically identify topological relationships analyzing different themes hat compose to a database.

With respect to analysis of relations, exits tow ways to analyzed spatial relations [15]:

- Intersection models, developed by [4] [5] [6].
- Schemas based in RCC (Region Connection Calculus), developed by [13], [14], [3], [1] and [2].

Considering these models to analyzed spatial relations, it was chosen the intersection model. Because we considerate that this model is the most adapted. This define the topological component for geographic objects, based on set point theory, which can be used to analyzed and formalized the topological relations between spatial objects. Additionally the necessities topological components are considerate to analyzed the relations between objects that are represented with different geometrical primitives of representation, same that the RCC model does not consider.

Semantic content can be appended to a spatial database by means of tuples. These tuples can be compound by a concept that represents the topological relationships. Of this way, each tuple represent the mean of a topological database.

The context of this work, we will focus to the topology and in specific on topological database among geographic database. In order to represent semantic content in spatial databases, we use a conceptual representation of the topological relationships. Fig. 1 show the methodology that we propose to integrate semantic content to databases. And Fig. 2 show the representation conceptual of a topological description.

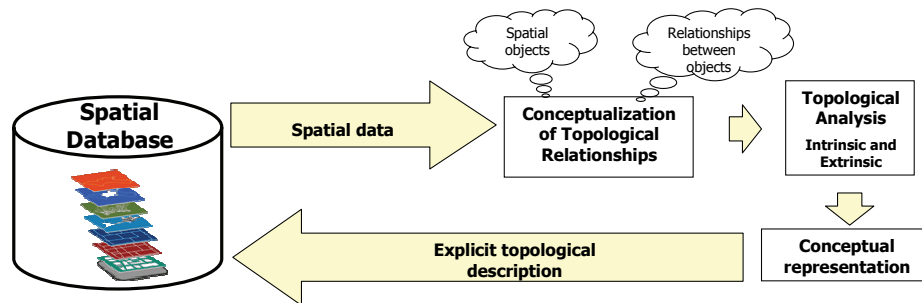


Fig. 1. Methodology proposed to integrate semantic content to spatial databases.

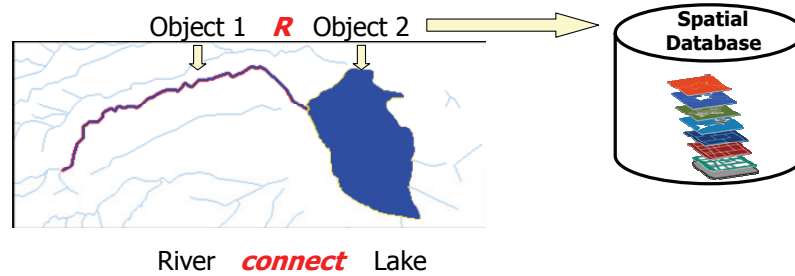


Fig. 2. Representation conceptual of a topological description.

This paper is organized as follows: Section 2 describes the schema of topological descriptor, that it is the computational system used to describe a spatial database. Section 3 describes the conceptualization of topological relationships used to describe the relations between to objects. Section 4 describes the levels of topological analysis: intrinsic and extrinsic. Sections 5 show the experimental results, and Section 6 sketches out our conclusion and future work.

2 Schema of Topological Descriptor

The design of topological descriptor consists of five principal parts. *Data source.* The system is design to allow tow source of data. These spatial data are that compose the spatial database. These data can be “shapefile” or “geofile” (property format). *Converting to geofile format.* If the source of data is “shapefile”, the system should make a transformation processes y “geofile” format. *Relations analysis.* In this stage, the system makes an analysis en two levels: *intrinsic* relationships analysis and *extrinsic* relationships analysis, which are made in independent way. *Topological descriptions.* These are the result of analysis of relationships among geo-spatial objects that compose the spatial database. The descriptions are stored in dBASE files (.dbf), with this, the topologic descriptions are represented in explicit way. *Spatial database,* this component store the topological descriptions, making one table for each topologic descriptions that is obtained with that intrinsic and extrinsic analysis. These tables contain the topological descriptions and the spatial objects are stores in “geofile” format. Additionally, is possible to export the topologic descriptions for to use these in other systems. Fig. 3 shows the topological descriptor model.

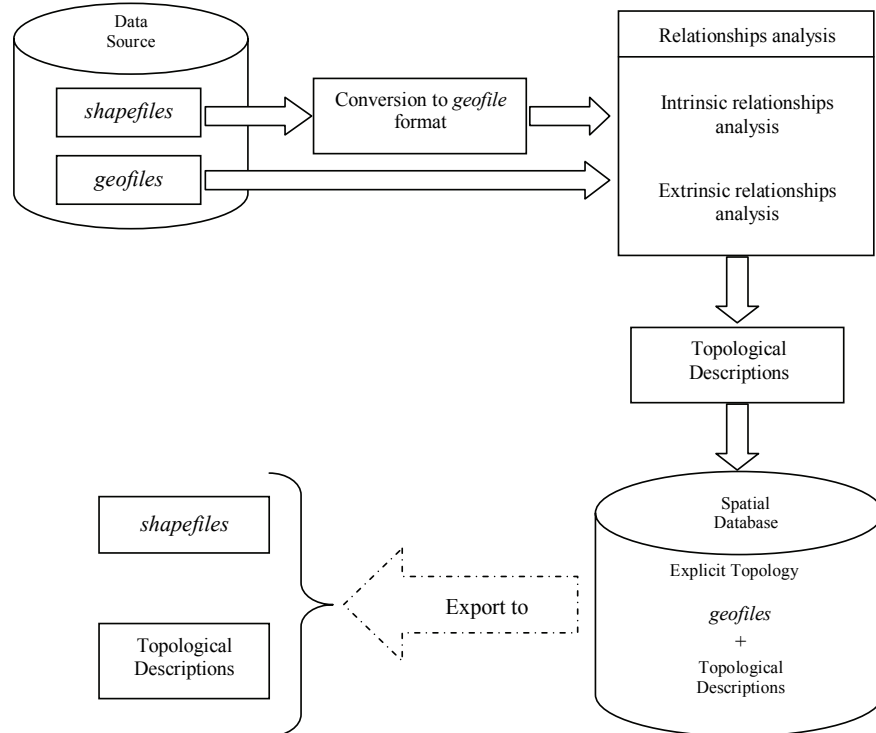


Fig. 3. Schema of Topological Descriptor.

3 Conceptualization of topological relationships

The conceptualization is performed by means of the relationships defined in 9-intersection model [6] and INEGI¹ specification [7][8][9]. We defined a set of 6 relations to describe topological descriptions between two spatial objects. These relations are based on the properties described in the topographic data dictionaries.

The description of six relations is showed in Table 1, as well as the symbol with which each relation is briefed.

¹ National Institute of Geography, Statistics and Informatics, National Mapping Agency

Table 1. Set of relations

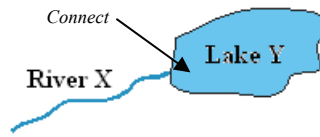
Relations	Symbol	Description
<i>Connect</i>	<i>C</i>	Valid to relate linear object to another object. The initial or final node of the linear object is connected to a limit [6] of another object. For instance: A river <i>connect</i> to a prey; ; A road <i>connect</i> to another road; A highway <i>connect</i> to a population.
<i>Share</i>	<i>S</i>	Valid to relate area objects to area or linear objects. Pairs of this objects have common elements, except the boundary. For instance, a river that is part of the boundary of a country.
<i>Share limit</i>	<i>Sl</i>	Valid to relate area objects to another area objects. The only common element is the boundary. For instance, the boundary between two states.
<i>Cross</i>	<i>X</i>	Valid to relate linear objects to areal or linear objects. (7) part of linear object is inside of a area object; (5) two linear objects are intersected, but the flow is not shared. For instance, a highway <i>cross</i> a railroad.
<i>Intersect</i>	<i>Y</i>	Valid to relate pairs of linear object. This relations describe an intersection and their flow is shared. For instance, an street <i>intersect</i> with another street.
<i>Inside</i>	<i>I</i>	Valid to relate any kind of objects, if they are inside of an area object. For instance, a town <i>inside</i> an state.

Connect relations is described to detail as example of as the other relations are described in [10]. This relation represents the connection between two objects. There are three variants of this relation: relations between Line/Area, Line/Line and Line/Point object. These possibilities depend on the spatial objects that are related. That is:

- When the relation is between one line object O_L and one area object O_A , we say that O_L *connect* with O_A ; i.e. if some node of O_L connect with the limit of some O_A . Using the topological components [11] to define this relation, we have the following:

$$\partial O_L \cap \partial O_A \neq \emptyset \text{ and } {}^\circ O_L \cap {}^\circ O_A = \emptyset \text{ and } \neg O_L \cap \neg O_A \neq \emptyset$$

For example, “The river X *Connect* with the lake Y” (see Fig. 4).

Fig. 4. The river X *Connect* with the lake Y.

- If the relation is between two line objects O_{L1} and O_{L2} , that is when a node of O_{L1} exist inside or in the limit of O_{L2} , we say that O_{L1} *Connect* with O_{L2} . Using the topological components to define this relation, we have the following:

$$(\partial O_{L1} \cap \partial O_{L2} \neq \emptyset \text{ or } \partial O_{L1} \cap {}^\circ O_{L2} \neq \emptyset) \text{ and } \neg O_{L1} \cap \neg O_{L2} \neq \emptyset$$

For instance, “The river X *Connect* with the river Y” (see Fig. 5a) and “The street X *Connect* with the street Y” (see Fig. 5b).

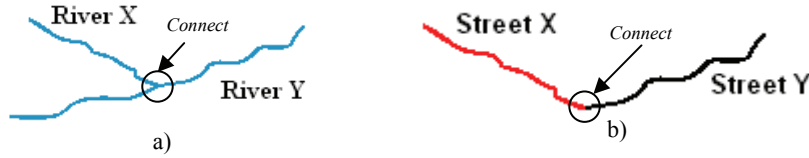


Fig. 5. a) The river X *Connect* with the river Y. b) The street X *Connect* with the street Y.

- When the relation is between some line object O_L and some point objects O_P , we say that O_L *Connect* with O_P , if O_P is in the limit of O_L . According with the Egenhofer’s definition, the limit of the point object is itself. Using the topological components to define this relation, we have the following:

$$\partial O_L \cap \partial O_P \neq \emptyset \text{ and } \neg O_L \cap \neg O_P$$

For example, “The highway X *Connect* with the population Y” (see Fig. 6).

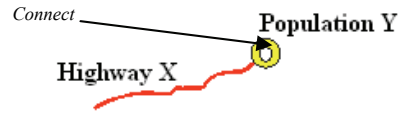


Fig. 6. The highway X *Connect* with the population Y.

The objects that they were used to compose the spatial database are classified in four thematics. The objects that were used to compose the spatial database are classified into four themes:

- Hydrology
- Land use
- Communication network
- Settlements

4 Topological analysis: intrinsic and extrinsic relationships

4.1 Intrinsic relationships

The intrinsic relationships are those that exist between objects that compose a same theme; for example, relations that exists inside Communications Networks theme. To

identify these relations, we use a diagram form each theme, like the shown in Fig. 7 to identify intrinsic relationships in Communication Networks.

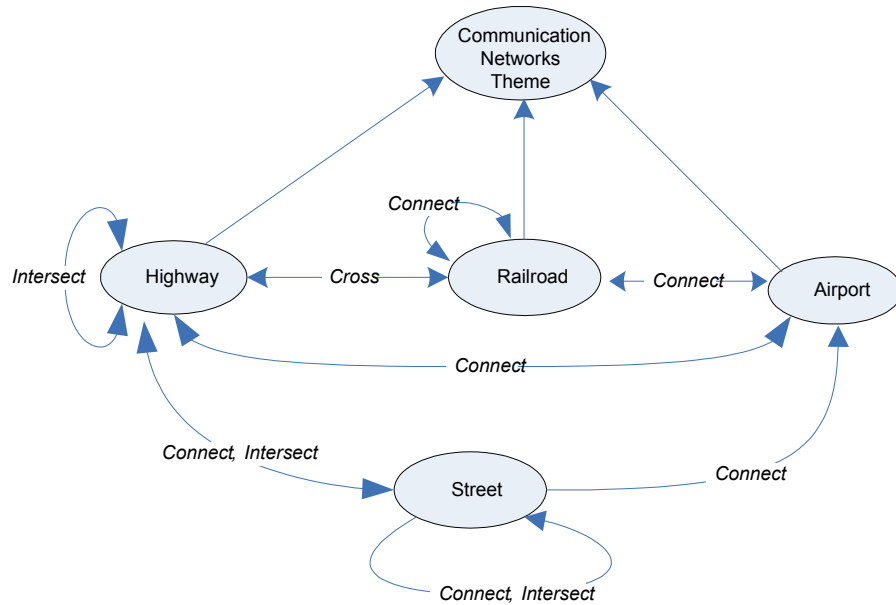


Fig. 7. Relations in Communication Networks.

4.2 Extrinsic relationships

The extrinsic relationships are those that exist between different thematics, for example, the relations between objects that belong to Hydrology and Settlement thematics. For example, in Fig. 8 is showed the relations among Hydrology and Settlement thematics.

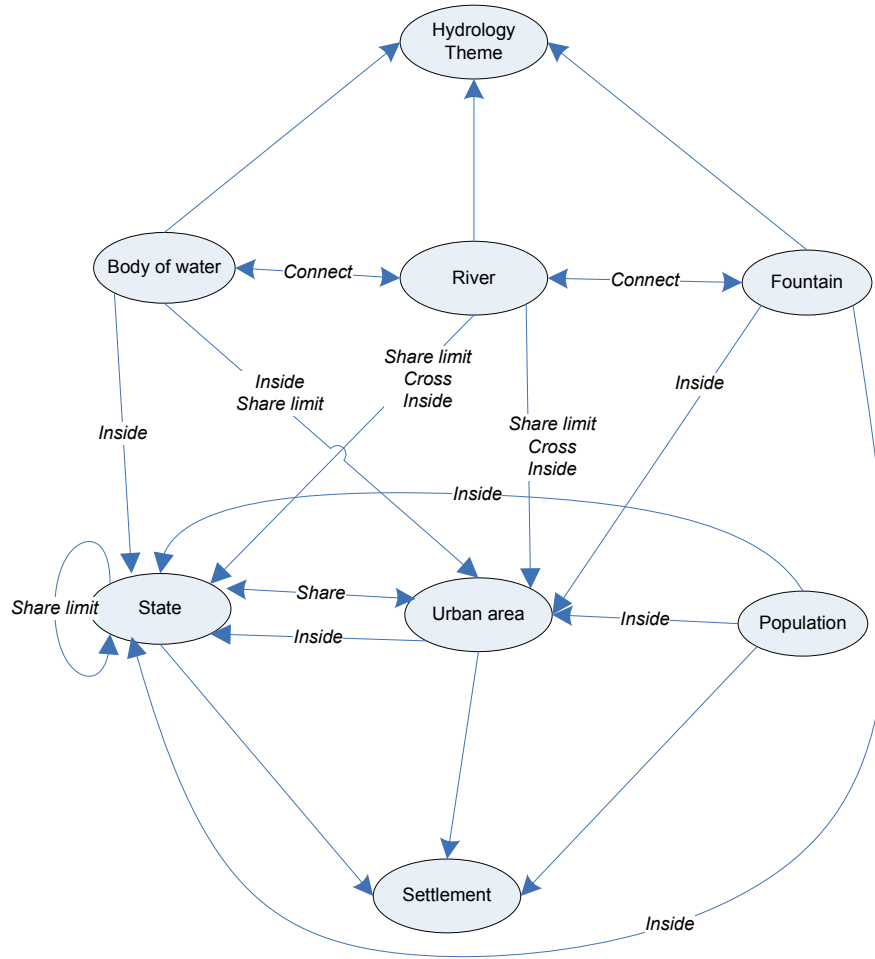


Fig. 8. Relations between Hydrology and Settlements

5 Experimental results

The results that were obtained from the topological descriptor are data stored in a table, when each table represents the descriptions of the topological relationships. These descriptions are stored in dBASE format (.dbf). Table 2 shows the attributes that compose the descriptions table.

Table 2. Attributes of descriptions table.

Attribute	Description
ID#R	Identifier existent relation between tow objects.
ID_OBJ_1	Index of first object. This Index corresponds to index of attribute table.
LAYER_BELONG_1	Data layer name of first object.
ID_OBJ_2	Index of second object. This Index corresponds to index of attribute table.
LAYER_BELONG_2	Data layer name of second object.
RELATION	Contain the symbol that identifies the relation that exists between tow objects.

To make the topological analysis was developed a library of classes that store and manage the spatial object in a proprietary format. These classes are implemented in Borland C++ Builder. The implementation of the classes is focused on working on vector data.

The functions in C++ to identify the *Inside relation* between polygons are:

```

bool Relations::Inside(Poly *_p1, Poly *_p2) {
    bool flag=false;
    if(_p2->polyInPoly(_p1))
        flag=true;
    return flag;
}

bool Poly::polyInPoly(Poly *_poly) {
    bool flag=true;
    for(int i=0;i<_poly->narcs;i++)
        for(int j=0;j<_poly->arcs[i].npoints;j++)
        {
            if(pointOutPoly(&(_poly->arcs[i].point[j])))
            {
                flag=false;
                break;
            }
        }
    return flag;
}

```

The classes implement methods to obtain the basic topology (connectivity and adjacency), the topological relationships between pairs of spatial objects provide methods for computing geometrical measures, For instance, sinuosity measure of an arc.

The topological relationships are obtained using the methods *Connect()*, *Share()*, *ShareLimit()*, *Inside()*, *Cross()*, and *Intersect()*.

Fig. 9 depicts an example in which we show *Inside* relationships between two spatial objects; “Green urban area” and “Urban area” (explicitly represented in the database).

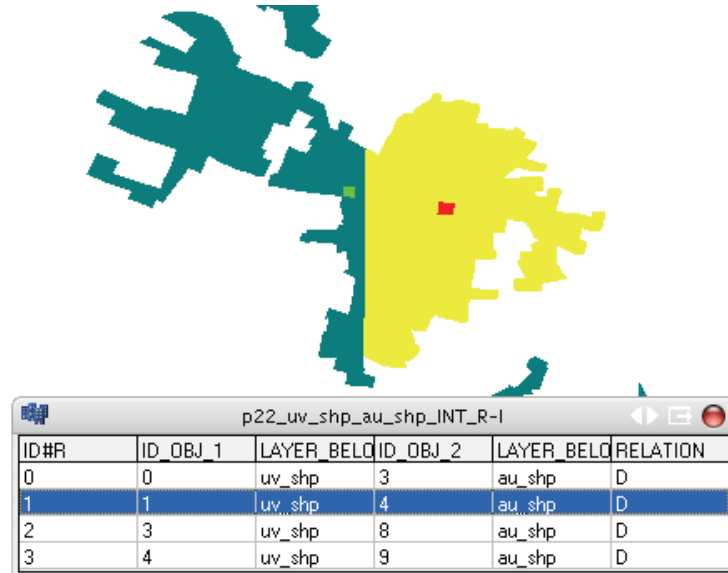


Fig. 9. Green urban area *Inside of* Urban area.

6 Conclusions and Future Work

With this classification in thematic the analysis of relations in two levels was carried out; reason why they were analyzed the existent relations between objects of each theme, to which it was denominated *analysis of intrinsic relations*. Afterwards it was analyzed the existent relations between thematic, to which it was called *analysis of extrinsic relations*.

The *content semantic* of relations between data is expressed by *concepts*. The conceptualization of topological relationships and concepts allow integrate a little of *semantic* to the GIS applications. The semantic content is obtained relating pairs of objects with one topological relationship. Reason why this work is one of the first attempts in this direction.

The concepts are generated using sets of geographic data. The concepts represent the interpretation of spatial data and the meaning of the relations between geo-spatial objects. With this work, we try to capture or determinate the proportions the semantic content that implicitly the spatial data contain and they do not depend on other factors, like scale or projection.

As a future work, the conceptualizations of these relations can be enhanced incrementing the number of thematic and topological relations, with which can be added new relations between objects. One interesting aspect that can be included is an analysis that depends of the context. In future will be very interest to analyze the changes in the relations that depends of context.

Is important to project that a domain conceptualization is useful to build ontologies, which represent (globally) the context of that domain, while the vocabulary of concepts and its relations describe the semantics (locally).

This descriptor can be applied to improve the results and possibly improve the performance in automatic generalization process.

Acknowledgements

The author of this paper wishes to thank to Marco Moreno-Ibarra and Miguel Torres for their support.

References

1. Cohn, A., Randell, D. and Cui, Z.: Taxonomies of logically defined qualitative spatial relations, *Formal Ontology in Conceptual Analysis and Knowledge Representation*, Kluwer, 1994.
2. Cohn, A., Bennett, B., Gooday, J. and Gotts, N.: Qualitative spatial representation and reasoning with the region connection calculus, *GeoInformatica*, 1:275-316, 1997.
3. Cui, Z., Cohn, A.G., and Randell, D.A.: Qualitative and Topological Relationships in Spatial Databases, *Third Symposium on Large Spatial Databases*, Lecture Notes in Computer Science N° 692, pages 296-315, Singapore, June 23-25, 1993.
4. Egenhofer, M. and Franzosa, R.: Point-Set Topological Spatial Relations, *International Journal for Geographical Information Systems*, 5(2): 161-174, 1991.
5. Egenhofer, M.: A Model for Detailed Binary Topological Relations, National Center for Geographical and Analysis and Department of Surveying Engineering, Department of Computer Science, University of Main, Orono, ME 04469-5711, U.S.A., 1993.
6. Egenhofer, M. and Herring, J.: Categorizing topological spatial relationships between point, line and area objects, The 9-intersections: formalism and its use for natural language spatial predicates, *Technical report* 94-1, National and Analysis, Santa Barbara, 1994.
7. Instituto Nacional de Estadística Geografía e Informática, "Diccionario de Datos Topográficos (vectorial), escala 1:250 000", 1995.
8. Instituto Nacional de Estadística Geografía e Informática, "Diccionario de Datos Topográficos (vectorial), escala 1:50 000", 1996.
9. Instituto Nacional de Estadística Geografía e Informática, "Diccionario de Datos Topográficos (vectorial), escala 1:1000 000", 1997.
10. Martínez, M.: Descriptor Topológico para Mapas Topográficos, Tesis de Maestría, México, Junio 2006.
11. Molenaar, M.: An Introduction to the Theory of Spatial Object Modelling for GIS, (Department of Geo-Informatics, International Institute for Aerospace Survey and Earth Science, Enschede, The Netherlands).
12. Mustière and Moulin, B.: What is spatial context cartographic generalisation?, *Symposium on Geospatial Theory, Processing and Applications*, Symposium sur la théorie, les traitements et les applications des données Géospatiales, Ottawa, 2002.
13. Randell, D. and Cohn, A.: Modelling topological and metrical properties of physical processes, *Proceedings First International Conference on the Principles of Knowledge Representation and Reasoning*, Morgan Kaufmann, Los Altos, pages 55-56, 1989.

14. Randell, D., Cui, Z. and Cohn, A.: A spatial logic based on regions and connection, Proceedings Third International Conference on Knowledge Representation and Reasoning, Morgan Kaufmann, San Mateo, pages 165-176, 1992.
15. Stell, J.: Part and Complement: Fundamental Concepts in Spatial Relations, School of Computing, University of Leeds, Leeds, LS2 9JT, U. K, 2002.