Proceedings of the 5th WSEAS Int. Conf. on Artificial Intelligence, Knowledge Engineering and Data Bases, Madrid, Spain, February 15-17, 2006 (pp366-371)

MSRIC: A Model for Spatial Relations and Integrity Constraints in Topographic Databases

KALUM PRIYANATH UDAGEPOLA, LI XIANG, YANG XIAOZONG, A.W. WIJERATNE
School Of Computer Science and Technology, Harbin Institute of Technology, Harbin 150001, CHINA

Abstract: Spatial relations and Integrity constraints are key components of topographical databases which have been used for the development of efficient models. Efforts have been made on formal definitions of spatial relations with the 4-intersection model (4IM) and 9 – intersection model (9IM), which are significant components in designing any model in topographical databases. Earlier definitions were also used for this model with required better modifications. This paper introduces new integrity constraint rules and semantic constraint rules which are used to design the model. The main contribution of this paper is to implement Conceptual frameworks for constraint rules for efficient and effective topographical database management.

Key-words: Spatial Relations, Integrity constraint rules, Semantic constraint rules, Constraint repository, Topographic database, Group relation, Spatial quires

1 Introduction

The number of applications using spatial or geographic data has been ever-increasing over the last decade. A topographic database has to be accurate, valid and consistent in order to be used as base data for any application. Inconsistencies in topographic database arise from the violation of constraint rules. This depends on the underlying data model, database management system and the spatial relationship. To define the constraint rules, the relationship among the objects should be known. Databases are intended to keep an integrated and consistent set of data that provides the information needed to support application requirements from one or several user communities. Without loss of generality, our work considers two specific facets that may induce Spatial Relations and Integrity Constraints to topographic data. This paper proposes a spatial relation model for Topographic database, called MSRIC, which offers and object based modeling of constraint rules with spatial features. MSRIC is supported by formal definitions, establishing a theoretical basis to build manipulation operations.

2 Spatial relations act in MSRIC

Spatial relations are the key issues in the design of MSRIC. A spatial database should be able to identify spatial relations to represent information about space. Spatial relationships are categorized according to different spatial concepts on which they rely. Relationships are also important to specify consistency constraints in proposed model. The model described by [6] offers content-based retrievals of scenes containing a particular object or those satisfying certain spatial constraints on them. The fundamental spatial relationships between objects can be broadly classified into three fundamental mathematical concepts, such as topological, order and metric relationships.

a. Topological relationship: eg. Neighbour, Disjoint
The topological relation between arbitrary objects based on set theory is formalized by Egenhofer [4,5].

b. Order relationship: eg. Behind. A spatial relation can also be expressed in terms of order theory based on sets [7]

c. Metric relationship: eg. Distance & directions
Spatial relations expressed in terms of distance and directions are grouped as metric relations [3].

Complex spatial data lead to the fragmental data structures and loose relationships among themselves [12]. In spatial databases, objects are represented with their geometric and attribute data. Location data are simply structured with co-ordinates and topology. The x, y co-ordinates are used to identify the location of the feature and topological data are used to identify node and polygon relationships. The focus of this MSRIC was based on topological relations. The semantic rules what we developed for the model were based on topological relations.

2.1 Topological relationships

The topological relations between arbitrary objects based on concepts of algebraic and set theory is formularized by Egnhofer [4,5]. These relations are preserved under groups of transformations, such as scaling, rotation, and translation. The model for binary topological relations have considered the possible intersection of boundary and interior of objects and called 4- intersection. Later they have extended the region which is defined as 2-dimentional point-set with three parts (interior, boundary & exterior) of non-empty and connected area. With these intersections it will be possible to formulate consistency rules as different groups of relations.

2.1.1 MSRIC connection with 4-intersection

There are two principal familiars of spatial relations which are the intersection model developed by Egenholer [4, 5], and the schemes based on the region connection calculus developed by Cohn el al. In case there are two basic schemes one of which sub schemes the others [11]. The binary relationship between any two elementary geographic objects is denoted by the boundary (O) of an elementary object and its interior (Θ). The boundary topological relationships between two region objects A & B based on the 4-intersection approach has been defined as intersection of boundaries, interiors and boundary and other interior (OO,ΘΘ,OΘ, ΘO). Each of this intersection can be empty or non-empty, thus creating 16 binary topological relationships. In this case each binary digit of 1 means for non-empty and 0 means empty intersection.

In the 16 possible relations for two regions, the following are the possible relations for two spatial regions. The relations r_2, r_5, r_8, r_9, r_{12} and r_{13} cannot occur between two spatial regions (Table 1). The relation r_{14} can occur only when one region has a hole on it. For example regions without holes, there are 8 distinct topological relations. They have been called Disjoint, Touch, Equal, Overlap, Inside, Contains, Covers, and CoveredBy.

r	$\mathrm{O}A \cap \mathrm{O}$	$A\Theta \cap B$	$\mathrm{O}A \cap B\Theta$	$A\Theta \cap \mathrm{O}B$	**Binary relation**	**Remarks**
r_o	$\emptyset$	$\emptyset$	$\emptyset$	$\emptyset$	0000	A & B are Disjoint
r_1	$\neg\emptyset$	$\neg\emptyset$	$\emptyset$	$\emptyset$	1000	A & B are Touch
r_2	$\emptyset$	$\neg\emptyset$	$\emptyset$	$\emptyset$	0100	
r_3	$\neg\emptyset$	$\neg\emptyset$	$\emptyset$	$\emptyset$	1100	A & B are equals
r_4	$\emptyset$	$\emptyset$	$\neg\emptyset$	$\emptyset$	0010	
r_5	$\neg\emptyset$	$\emptyset$	$\neg\emptyset$	$\emptyset$	1010	
r_6	$\emptyset$	$\neg\emptyset$	$\neg\emptyset$	$\neg\emptyset$	0110	A is Inside of B (B contains A or A is contained by B)
r_7	$\neg\emptyset$	$\neg\emptyset$	$\neg\emptyset$	$\emptyset$	1110	A is CoveredBy B (B Covers A)
r_8	$\emptyset$	$\emptyset$	$\emptyset$	$\neg\emptyset$	0001	
r_9	$\neg\emptyset$	$\emptyset$	$\emptyset$	$\neg\emptyset$	1001	
r_{10}	$\emptyset$	$\neg\emptyset$	$\emptyset$	$\neg\emptyset$	0101	A contains B (B is Inside of A or B is contained by A)
r_{11}	$\neg\emptyset$	$\neg\emptyset$	$\emptyset$	$\neg\emptyset$	1101	A Covers B (B is CoveredBy A)
r_{12}	$\emptyset$	$\emptyset$	$\neg\emptyset$	$\neg\emptyset$	0011	
r_{13}	$\neg\emptyset$	$\emptyset$	$\neg\emptyset$	$\neg\emptyset$	1011	
r_{14}	$\emptyset$	$\neg\emptyset$	$\neg\emptyset$	$\neg\emptyset$	0111	
r_{15}	$\neg\emptyset$	$\neg\emptyset$	$\neg\emptyset$	$\neg\emptyset$	1111	A & B overlap

Table 1 Resulting of 16 possible combinations [3]

2.1.2 MSRIC connection 9-intersection

The 9-intersection approach was proposed in Egenhofer [4,5] for formalizing binary topologic relationship between arbitrary objects. In this approach, binary topological relations between two objects A and B are defined in terms of the nine intersections of A's boundary (O), A's interior (Θ), A's exterior (⁻) with the boundary B, interior B and exterior B of B. Each object A and B can be a point, a line or a region. The 9-intersection would be able to distinguish more details than the 4-intersectin. The topological relations between regions A and B are concisely represented as 3x3 matrix, called the 9 – intersection.

$$\begin{bmatrix} A\Theta \cap B\Theta & A\Theta \cap \mathrm{O}B & A\Theta \cap B^- \\ \mathrm{O}A \cap B\Theta & \mathrm{O}A \cap \mathrm{O}B & \mathrm{O}A \cap B^- \\ A^- \cap B\Theta & A^- \cap \mathrm{O}B & A^- \cap B^- \end{bmatrix}$$

With each of these nine intersections being empty (0) or non empty (1), the model distinguished 512 different relations actually exist in $\Re^2$ between regions, lines and points. Exactly one of these topological relations holds true between any two regions, because the nine empty/non empty intersections describe a set of relations that are mutually exclusive and provide a complete coverage. The actual number of reliable

Proceedings of the 5th WSEAS Int. Conf. on Artificial Intelligence, Knowledge Engineering and Data Bases, Madrid, Spain, February 15-17, 2006 (pp366-371)

relations depends on the dimensions of the objects and the dimensions of their embedding space. From the 512 possible relations, only eight can be realized between regions in the two dimensional space. We call these eight relations Disjoint, meet, equal, Inside, contains, Covers, CoverdBy, and overlap. This is similar to the 4-intersection approach.

2.2 Group relations

Recent works underline that the integration of topological relationships into the Object Constraint Language (OCL) is an important field of investigation. The final goal is to provide an expressive language adapted to precisely model alphanumerical and topological constraints. [2] focuses on the integration of the 9IM into OCL. It shows that this OCL+9IM language is especially suitable for the specification of topological constraints implying composite spatial objects. 9IM can be applied to all kinds of geographical objects. The extension of region features with line and point features results in six major groups of binary relationships: point/point, point/line, point/region, line/line, line/region, and region/region. In these 6 major groups of binary relationships, there are large numbers of different relationship available and each of this has different name. Egenhofer[4,5] gave the list of relations (Table 2) that are embedded in 2-D space. In this model two converse relations are available, such as:

A is Inside B → B contains A
A CoveredBy B → B Covers A

Group of relations	9-intersection	4-intersection	Extended 4-intersection
Region/Region	8	8	9
Region/Line	19	11	17
Region/Point	3	3	3
Line/Line	23	16	18
Line /Point	3	3	3
Point/Point	2	2	2

Table 2 Number of relations in three intersections [1,4,5]

Checking converse relations may not show any inconsistency. Therefore these converse relations have been combined to one relation. The resulting relations between the regions will be Disjoint, Touch, overlap, Inside, CoverdBy and Equal. Relations between region and line have number of relations. It may be hard to remember all of the relations and the users might become confused. Therefore these relations are needed to be generalised according to the user interactions. Clementini [1] extend the number of relations to five topological relationships. Such as Touch, In, Cross, Overlap, and Disjoint. They have proved that these relationships are close to human use of concepts and powerful enough to handle a wide variety of cases. The *In* relations from Clementini & *Inside* relations between regions give the same result for all the other groups. Therefore *Inside* relation has been used for all the other groups. The summary of relations has been prepared from Clementini[1][9], & SDE [10] to illustrate the possible relations between two spatial objects.

3 Integrity constraints and constraint rules For the MSRIC

Integrity constraints described the conditions that must be satisfied by every legal instance of relation. An integrity constraint is similar to a condition or a predicate, which has to be verified. Consistency tests can be carried out by constraint rules. In spatial databases the geometry and the topology are linked with the geometric representation. The spatial queries can be easily performed if the geometric relations between objects are explicitly stored. However it is very expensive to store all the relations with the geometry in spatial databases because it requires large amount of disk space. Therefore relationship has to be verified in order to maintain the consistency of the database. The spatial integrity constraints are generally classified into three categories: topological, semantic, and user defined integrity contrarians.

3.1 Topological integrity constraints

Topology is the branch of mathematics that deals with properties of spaces that remain invariant under certain transformations. The data model must provide means to define and enforce topological consistency in a database.

Topological consistency relations are listed below:

- Everything must be bounded by two nodes (start node and end node).
- For every arc, there exist two polygons (left polygon and right polygon).
- Every polygon has a closed boundary consisting of an alternating sequence of nodes and arcs.
- Around every node, there exists an alternating closed sequence of arcs and polygons.
- Arcs do not intersect except at nodes.

In ArcInfo data model the arc-node data structure supports three major topological concepts:

1. Connectivity: Arcs are connected to each other at nodes. Arc is defined by two end points, the from-node and to-node (Arc-node topology).

2. Area definition: Arcs that connect to surround an area define a polygon (Polygon-Arc topology).

3. Contiguity: arcs have direction and left and right sides (left-right topology).

3.2 Semantic integrity constraints

Topological relations are based on the shape of objects, but semantics of objects have to be taken into account to decide whether a topological scene is an error or not. It is agreed that topological relations are of great importance regarding to GIS data consistency. Semantic errors, such as a road within a lake, a building represented by a line, are coming from the real world description. Such error can not be found without using the semantics of real world entities. Topological relations are based only on shape of objects, but semantics of objects have to be taken into account to decide whether a topological scene is an error or not. It is difficult to make a generic algorithm for resolving semantic constraints. But a monitoring procedure can be formulated between topologically corrected objects.

3.3 User defined integrity constraints

User defined constraints allow database consistency to be maintained according to user defined constraints, which are similar to business rules in non spatial DBMS. For example, a firework factory cannot be located within a distance from the centre of the town.

4 Semantic constraint rules for the MSRIC

4.1 User defined integrity constraints

In the topographic database designed in MSRIC, all the layers have been classified into sub classes of features and each feature is identified by a feature code. Semantic integrity constraints will be defined between two geographical objects (possibly between two layers). The topological relation between two objects is the main part of the constraint. In the spatial domain, they are mostly a group of forbidden relationships between pairs of objects [8]. It is easier to define a case that should not happen than to define a case that must exists [13]. The formula for the constraint rule will be as follows:

CONSTRAINT=Object1,Spatail-Operator,Object2, Specification

This constraint is defined as the association of two geographical objects, a topological relation (spatial operator) between them and a specification.
The specification can be one of the following
(a) Forbidden (b) Unless: condition (c) Allowed
(d) At least once (e) At most once

"Unless" specification will be followed by a "**Condition"**, where the feature code of the object will differ from the features. For example if a buildup area intersects with the road, the intersected feature will have the feature code with the last digit of 2 or 9 (last digit 2 = object is below the other object and 9 = object is above the other object).
This model spatial operator can be one of the following spatial relations

(a) Disjoint (b) CoverdBy (c) Inside
(d) Touch (e) Equal (f) Overlap (g) Cross

The objects associated with the constraint could be further combined with the sub classes of the feature.

CONSTRAINT= Object1(Subclass1), Spatial Operator,
Object2(Subclass2), Specification

To provide a more usable interface, topological relations sharing common attributes have been grouped into subsets. Such subsets have been built in each group of relations. For example, buildings cannot be ***Inside*** the roads. In this case we have to test three constraints, such as a Overlap, CoveredBy and Equal. All the possible constraint rules have been listed as follows.

a. Semantic constraint between polygon/polygon objects

- High raised buildings cannot **intersect** the roads
- Buildings cannot be **Inside** the water bodies unless it is a special building
- Any two land use parcels cannot **overlap**.
- Parking place must have a **access to** road.
- All built-up areas must have a **access to** road.
- Roads should be **adjoining** to other road types(within the theme)
- Road can **intersect** the built-up area, if the intersection is part of the underneath object code.
- Dockyard must be **adjoining to** the water body.
- Landing stage must be **adjoining to** the water body.

b.Semantic constraint between line/polygon objects

- Ditches cannot **cross** the buildings or built-up area
- Ditches cannot **cross** the road unless the intersect portion has the underneath object code.
- Wall cannot **cross** the buildings or built-up area
- Railroad cannot **cross** the buildings or built-up area unless the intersection is underneath/above object code
- Footpath and street cannot **cross** the highway unless it is below/above the other object
- Bridge must be **part of** road or water-bodies
- Shoreline or riverbank cannot **cross** the road
- Shoreline or riverbank cannot **cross** the

buildings/built-up area

- Lock door for ships is **part of** water body

c. **Semantic constraint between point/polygon objects**
- Police office, post office, municipality office, religious building, railway station(buildings symbol) should be **within** the buildings or built-up area
- Milestone pole should be **within** 6 meters of the roads
- Signpost should be **within** 6 meters of the roads
- Culvert must be **part of** road/railway/river
- Dam must be **adjoin to/part of** the water bodies
- A monument should not **Inside** road
- Buildings symbol cannot be **Inside** the road

d. **Semantic constraint between line/line objects**
- Railway line should not **cross** the ditch unless the intersection is an underneath object code
- Contour line cannot **cross** another contour line

e. **Semantic constraint between point/line objects**
- Railway milepost should be **within** 6 meters of the railway line
- High tension pillars are **part of** high tension line should intersect on common point

f. **Semantic constraint between point/ point objects**
- Tree cannot **equal to** buildings symbol
- Any buildings symbol cannot **intersect** with itself

Intersect access to, overlap, and cross are the constraint operators. These operators should be unique for all the groups of relations. Rather than selecting different names for constraint operators, it is easy to define the relations using spatial relations called spatial operators.

Constraint rule:

a. Buildings cannot intersect the roads

To translate this constraint rule into topologic relationships, we have to test three relations. Then the system must warn the user about the inconsistency.

Buildings (All) Overlap Roads (All) : forbidden
Buildings (All) CoverBy Roads (All) : forbidden
Buildings (All) Equal Roads (All) : forbidden

b. Parking place must have an access to roads

Landuse(Parking place) Overlap Roads (All) : forbidden
Landuse(Parking place) CoverBy Roads (All): forbidden
Landuse(Parking place) Equal Roads (All) : forbidden
Landuse(Parking place) Touch Roads (All) :at least

c. Building or build-up areas cannot cross ditches

Buildings (Parking place) **Cross** Streams (Ditches) : forbidden
Buildings (Parking place) **Inside** Streams (Ditches) : forbidden
Buildings (Parking place) **Touch** Streams (Ditches): allowed

4.2 Problems of spatial integrity constraints

Some of the problems are encountered during the process of listing spatial integrity constraints. For example, it is difficult to obtain a complete list of spatial integrity constraints and identify the minimum constraints in order to avoid redundancy in checking. However the list will not be fulfilled to consider all the possibilities of the real world situation. User can also define the rules to check the consistency during the updating process.

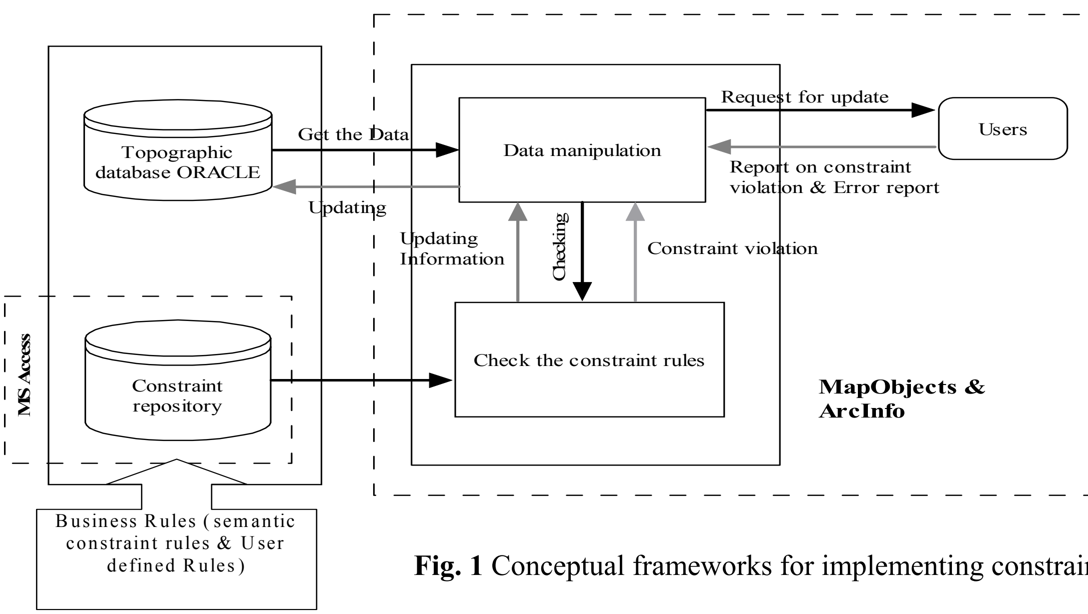

Fig. 1 Conceptual frameworks for implementing constraint rules

4.3 Implementing integrity constraints spatial database

Figure 1 shows the implementation process of the constraint rules in this model. The topographic data is in the Oracle database. The constraint repository is used to store the semantic and user defined rules. A user interface was created in MapObjects with Visual Basic for checking the constraint rule process. Data manipulation process involves getting the data from database and updating the database after checking the inconsistencies of the new data.

5 Conclusions

During the study, two major model outputs were generated; Identification of spatial relations and Identification of integrity constraints.

Identification of spatial relations for the MSRIC model: Four intersection methods and nine intersection methods were absorbed to the model. Group relations were identified as a special method for geographical databases. Identification of integrity constraints: Topological integrity constraints, symmetric integrity constraints and user defined integrity constrained were recognized.

Six symmetric constraint rules were defined and out of which three constraints rules were identified such as Building can not intersect roads, Parking place must have an access to roads and building or built-up areas can not cross ditches.

Finally, a conceptual repository model was developed for the MSRIC and spatial integrity constraint problems were identified as future works for the further development of the model.

References:

[1] E. Clementini, P.D. Felice and P.V. Oosterom, A small set of formal topological relationships suitable for end-user interaction, in 3rd International symposium on large spatial databases, SSD'93 Lecture notes in computer science Vol 692, 1993,pp. 277-295.

[2]M. Duboisset, F. Pinet, M. Kang and M. Schneider, Precise Modeling and Verification of Topological Integrity Constraints in Spatial Databases: From an Expressive Power Study to Code Generation Principles, ER 2005: 24th International Conference on Conceptual Modeling, Klagenfurt, Austria, 2005, pp.465-482.

[3] M. Egenhofer and J. Herring, AMathematical Framwork for the definition of Topological Relationships, In: 4th International Symposium on Spatial data handling, Zurich, Switzerland, 1990, pp. 803-813.

[4] M. Egenhofer and R. Franzosa, Point-set topological spatial relations, International Journal of Geographical Information Systems, Vol.5, No.2, 1991,pp.161–174.

[5] M. Egenhofer and J. Herring, Categorising binary topological relations between regions, lines and points in geographic databases, Technical Report, Department of Surveying Engineering , University of Maine, http://www.spatial.maine.edu/~max/pubs_TR.html , 1991.

[6] J.H. Hwang, J.H. Ro, K.H. Lee, J.S. Park and S. Hwang,, An XML-Based 3-Dimensional Graphic Database System, In: 1st International Conference, Human Society Internet 2001, Vol. 36, No. 1-2, 2002, pp.195-232.

[7] W. Kainz, Logical consistency, In: Guptill and J.L. Morrison, Elements of spstial data quality. Published on behalf of the Internatinal Cartogrphic Conference, Pergamon Press, 1995, pp. 109-137.

[8] O.Kufoniyi, Spatial coincidence modeling automated database updating and data consistency in vector GIS, Phd-Thesis,ITC publication number 28,1995.

[9] A.J. Morris, A.I. Abdelmoty, B.A. El-Geresy and C.B. Jones, A filter flow visual querying language and interface for spatial databases, Geoinfomatica, Vol. 8, No. 2, 2004, pp.107-141.

[10] SDE, SDE C Developer's Guide, Environmental Systems research institute, Inc, 1999.

[11] J.G.. Stell, Part and Complement: Fundamental Concepts in Spatial Relations, Annals of Mathematical and Artificial. Intelligent, Vol. 41, No.1, 2004, pp. 1-17.

[12] Z. Tie-wen, Z. Zhi-nong and J. Ning, Managing Method of Spatial data in Relational Database management Systems, Journal of Software, Vol. 13, No.1, 2002.

[13] T. Ubeda and M. Egenhofer, Correcting Topological errors in GIS, In: fifth International Symposium on large spatial databases, Berlin, July 1997.