Spatial Data Description by Means of Knowledge-Based System

Víctor Montes de Oca, Miguel Torres, Serguei Levachkine, and Marco Moreno

Geoprocessing Laboratory - Centre for Computing Research - National
Polytechnic Institute, Mexico City, Mexico
`vicmoca@sagitario.cic.ipn.mx`,
`{mtorres, sergei, marcomoreno}@cic.ipn.mx`

Abstract. In this paper, we propose the use of a knowledge based system, which has been implemented in SWI-Prolog to approach the automatic description of spatial data by means of some logic rules. The process to establish the predicates is based on the topological and geometrical analysis of spatial data. These predicates are handled by a set of rules, which are used to find the relations between geospatial objects. Moreover, the rules aid the searching of several features that compose the partition of topographic maps. For instance, in the case that any road *intersects* with other, we appreciate that a *connection relation* exists between different destinies, which can be accessed by these roads. Furthermore, the rules help us to know each possible access for this case. Therefore, this description assists in the tasks of geospatial data interpretation *(map description)* in order to provide quality information for spatial decision support systems.

1 Introduction

The technique of knowledge-based systems consists of manipulating the information to support human decision-making, learning and action. Such systems are capable of cooperating with human users and so the quality of support given, as well as the knowledge representation *(the information quality)* are important issues to consider when developing such systems.

Through pattern recognition, the process of acquisition of spatial data is automated, which is generally stored in raster or vector format, being this last one the most useful to make topological and geometrical analysis from which spatial knowledge is acquired. It is necessary to count with a correct representation of this knowledge to assist to understand, manage and share information of a spatial domain region.

Expert systems are attempting to introduce human knowledge about problem solving into computer software. The general objective is to emulate the problem-solving capabilities of the human expert [1]. Traditionally, expert systems were employed to aid on specific domain tasks, and its solution set was delimited by the knowledge base that constitute them, which could not be extended to other related domain tasks without reengineering the database. Knowledge based systems (KBS) extend this principle by allowing new knowledge to be inferred from the existing one adding it to the database.

J.F. Martínez-Trinidad et al. (Eds.): CIARP 2006, LNCS 4225, pp. 502 – 510, 2006.

In this paper, we propose the use of a KBS, which has been implemented in SWI-Prolog to approach the automatic description of spatial data by means of some logic rules. The process to establish the predicates is based on a topological and geometrical analysis of the spatial data to find some basic properties. These predicates are handled by a set of rules, which are used to find the relations between geospatial objects. Moreover, the rules aid the searching of several features that compose topographic maps. The focus is on development and execution of knowledge-based "micro systems", specialized for a specific region (case study) by means of logic predicates, built in a pattern recognition process. These predicates compose the knowledge base and are used by "universal" rules that infer new knowledge for the domain of topographic maps.

For instance, in the case that any road *intersects* with other, we appreciate that a *connection relation* exists between different destinies, which can be accessed by these roads. Furthermore, the rules help us to know each possible access for this case and this knowledge is valuable for the tasks of geospatial data interpretation *(map description),* in order to provide quality information for spatial decision support systems.

Because we manage "basic data" obtained by means of a pattern recognition process and rely on that data to generate knowledge, we are convinced of its importance on the automation of task and data manipulation.

The rest of the paper is organized as follows. In section 2 we present an overview related to knowledge-based systems (KBS), their importance on AI's research and how we address the issues of knowledge representation, acquisition and processing. Also we show a brief example on how spatial knowledge can be derived through PROLOG's inference capabilities and what kind of information composes of the knowledge base. Section 3 contains the algorithm and logic that we use as the workforce toward generating the map description, and also a brief example of a description generated so far with the KBS developed. Our conclusions are outlined in section 4.

2 Why Knowledge-Based Systems Are Important?

Knowledge is of paramount importance, and AI research has shifted its focus from an inference-based paradigm to a knowledge-based paradigm. Knowledge is viewed as consisting of facts and heuristics. The *facts* constitute a body of information that is widely shared, publicly available, and generally agreed-upon by experts in a field. The *heuristics* are most private, little-discussed rules of good judgment (rules of plausible reasoning, rules of good guessing) that characterize expert-level decision making in a field [9]. "…however, this does not restrict the knowledge base to a traditional ES (Expert System) approach but it could include more indirect forms of knowledge representation" [1]. Heuristics are also embedded into the process of pattern recognition and in the results obtained by such algorithms.

Three major research issues of AI's knowledge-based paradigm are grouped as issues of knowledge representation, knowledge utilization, and knowledge acquisition [9]. In this paper, we specify how we approach these issues to obtain the automatic spatial data description by means of a knowledge-based system (KBS).

2.1 Knowledge Representation

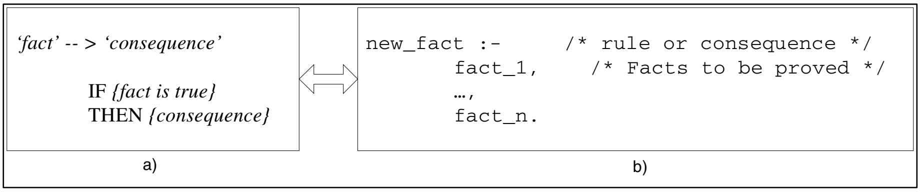
Since we use SWI-Prolog as the platform to develop our KBS, it is necessary to build the knowledge in the form of predicates or *facts* about the spatial data. We propose to use these facts to generate a description of the spatial data.

2.2 Knowledge Utilization

The knowledge is used as first order logic statements that help us to discover, by means of inference procedures, more advanced (or complex) relations that topologic and geometric analysis are not aware of because they are out of their scope at runtime.

With that, we would like to state that even when we still use inference to acquire *new knowledge* about spatial relations among the objects that compose a map (the spatial data). There is an interesting twist on the way the *basic knowledge* base is constructed; instead of a human expert being the one who inputs the knowledge, this is acquired by the automatic process of topologic and geometric analysis (a kind of pattern recognition), which is done to a map, whose in turn, uses heuristic methods to obtain this information.

How logic rules can help to describe spatial data? Logic rules are formed by two elements: facts and a consequence. The consequence is considered true if all the facts that the rule needs to prove turn out to be true (see Fig. 1).



'fact' -- > *'consequence'*

IF *{fact is true}*
THEN *{consequence}*

```
new_fact :-      /* rule or consequence */
       fact_1,    /* Facts to be proved */
       …,
       fact_n.
```

Fig. 1. Methods to express rules. a) From the traditional way. b) To the PROLOG way.

In Fig. 1 1a) we note that the traditional way is a series of *facts* (or just a single fact) that triggers a *consequence*. In PROLOG 1b) we search for all those *facts* (they must be in the knowledge base, so they can prove to be true) so if and only if every fact proves true then the *rule* is said to be true [12]. This reveals a *new fact,* we did not know, but always existed *implicitly* in the knowledge base, we could add this new knowledge *explicitly* to the knowledge base to aid on more advanced inferences. That is the very essence of a KBS.

In Fig. 2, we used first order logic predicates together with the topological relations of the 9-intersection model [6][7], because in this model two spatial objects 'A' and 'B' share a topological relation, we can group and compare them (even as a single entity) with other objects that have the same relation (and different ones) and make more complex analysis.

If **A** *contains* **B** then **B** *is inside* **A**
also
if **C** *contains* **B**
then we conclude that
C *contains* **A** or that **A** *is inside* **C**

Fig. 2. Example of a inference process, in which a topological relation between *A* and *C* is discovered, by means of a common *relation* they both share with *B*

2.3 Knowledge Acquisition

Straight from topologic and geometric analysis [10][11][13], we can automatically construct a knowledge base containing the following *basic facts* about the spatial objects that compose a map:

- Topologic relations: *disjoint, meet, overlap, coveredBy, covers, contains, inside, equal.*
- Relative directions: *north, south, east* and *west* [8].
- Geometric measures: *area, perimeter, distance, large…*
- Type attribute: *area, line, point.*

If the map is already classified into a spatial database, we can *import* the following elements:

- Theme attribute: *road, population, hydrological, land type …*
- Descriptive attribute: *name …*

Although there are many possible advanced ‘functions’ or ‘rules’, we would like to have in the inference engine of the KBS, we concentrate on a few and we think that they are the most useful to describe a map such as Fig. 3, which is very important to consider in the following sections.

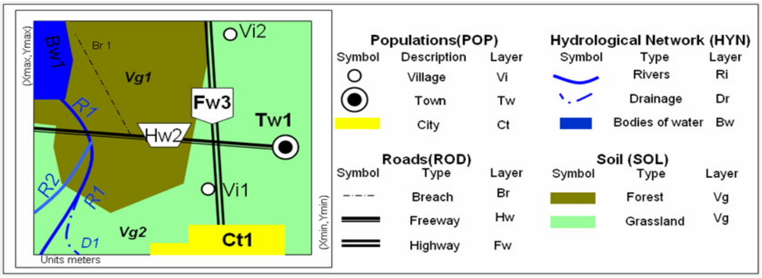


Fig. 3. Case study map for testing a KBS spatial data description. This map[1] is composed of 13 objects: 4 areal, 6 linear and 3 punctual. They are organized on 4 themes: *Population, Hydrology, Roads and Soil.*

---

[1] The legend of the map shows the geospatial objects types of each thematic, in this case they are 11 representations. However the map is composed of 13 objects, because it has 2 rivers (R1, R2) and 2 villages (Vi1, Vi2).

Interconnects() – Returns those spatial objects that constitute urban places or ‘destinies’ like villages, towns and cities, that are related *(or can be reached)* to an object that suggests some kind of communication as roads, highways and freeways.

We shall note that the logic of this relation can be extended to search for interconnections of objects of the same theme, such as body of water objects are connected by rivers, or even connections among the nodes of river networks. For example, the following PROLOG query:

```
          interconnects(fw3, Ans).
should return:
          Ans = vi1
          Ans = vi2
          Ans = ct1
```

It turns out that with this information added to the knowledge base, we could generate the following rule:

Destinations() – Returns all those ‘destinies’ that can be reached from one place by any number of ‘roads’ that *intersects* between them, which turns out to be the ‘cross product’ of all the destinies that *interconnects* each of those ‘roads’. Since the highway **Hw2** interconnects with town **Tw1** and intersects the freeway **Fw3** the query is the following:

```
          destinations(ct1, Ans).
should return:
          Ans = vi1
          Ans = tw1
          Ans = vi2
```

A more general rule, mainly used for debugging the knowledge base construction is:

What_Relation() – This rule returns *true* for all the relations that the knowledge base states between two specified objects. For this purpose, it is necessary to prove each rule seen so far, from the basic to the derived ones, so special care should be taken to include in the code every new rule generated or derived. For example the query:

```
          what_relation(r1, bw1).
should return:
          connects = yes.
          south = yes.
```

Even though the last rule was conceived for debugging purposes, we can use the information provided to construct a rule called:

Explain() – It mentions everything that we know *(that is in the knowledge base)* about a spatial object. For instance, how it is related to other objects *(topology)*, its name, classification attributes as *type (line, point, area)*, also its *theme: {[Roads(freeway, highway,…)]; [population(city, village, town,…)]; [hydrological(body of water, river, drainage,…)]; [land(grassland, forest, breach,…)]}* and its metrics, if we have such information. The query is:

```
          explain(bw1).
should return:
          type = area
          theme = hydrological
          is_a = body of water
          inside = vg1
          covered_by = vg1
          meet = vg1 ; meet = r1
          north_of = hw2 ; north_of = r2 ;
          north_of = d1 ; north_of = ct1
          west_of = br1 ; west_of = vi2 ;
          west_of = tw1 ; west_of = vi1 ;
          west_of = ct1
```

It should be obvious that, while all the basic predicates or facts in the knowledge base are generated through topological and geometrical analysis and some more advanced facts are inferred through rules, is this last rule along with the algorithm explained in the following section, the workforce of the map description generation.

3 Map Description Process

Verbal descriptions of spatial situations are frequently ambiguous and may easily lead to misinterpretations, because geographic concepts are often vague, imprecise, little understood, or not standardized [2].

Experiments in psychology and cartography showed that topology is among the most critical information people refer to when they assess spatial relationships in geographic space, while metrical changes are frequently considered to be of lesser importance. This is based on the premise *topology matters, metrical refines.* In [2] referring to [4, 5].

For this reason, we only consider topology characteristics for the “first levels” of spatial descriptions, which involve the metrics only for those spatial objects that are selected to be of relevance, as in the rule *(early in detail)*:

```
explain(object).
```

3.1 The Description Generator Algorithm

The KBS constructs sentences about the spatial state of the map and the relations between spatial objects that constitute it, in such a way that this knowledge generated as sentences can be looked out on a search by exact word match.

The description is generated considering the following algorithm:

1. Start from the ‘leftmost’ spatial object at the top and work to the right toward the bottom of the map.
2. For each spatial object:
3. Describe its type, theme and name.
4. Describe its topological relations in order of similitude *(more on this later…)*
5. IF the description level > 1 AND we have metric data, describe it.
6. Search the object to the east with which it has the following topological relation, in order of importance:

   - Overlap.
   - Meet.
   - Disjoint.
7. Set it as the new work object and repeat from the step 3.
8. If there is no object to the east, look for objects to the south considering conditions in the step 6
9. Repeat from the step 2
10. If there are no more objects to the east and south, end.

In the step 4, we use the “Conceptual neighborhood graph of the eight region-region relations” [3] shown in Fig. 4 to determine the order in which the elements of description should appear, basing on the importance and similarity of topological relations.

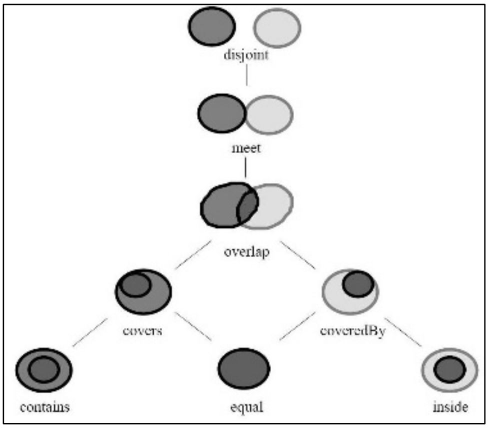


Fig. 4. Conceptual neighborhood graph of the eight region-region relations. It states the similitude between the binary topological relations.

This procedure produces the following map description:

[a body of water](bw1) *inside* [a forest](vg1) *connects* [a river](r1) {to the east} *disjoint* [a breach](br1).

[a forest](vg1) *inside* [a grassland](vg2) *contains* [a body of water](bw1) *contains* [a breach](br1) *contains* [a river](r1) *contains* [a river](r2) *covered by* [a grassland](vg2) *meets* [a freeway](fw3) *intersects* [a highway](hw2)

[a grassland](vg2) *contains* [a river](r1) *contains* [a river](r2) *contains* [a forest](vg1) *contains* [a village](vi2) *contains* [a freeway](fw3) *contains* [a highway](hw2) *contains* [a town](tw1) *contains* [a village](vi1) *contains* [a drainage](d1) *contains* [a city](ct1)

. . .

4 Conclusions

In this paper, we propose a set of logic predicates that state some basic characteristics of spatial objects, which can be generated after some pattern recognition analysis such as topology, relative direction, and geometric measures.

In addition, we propose automatic methods based on the measures above mentioned, to construct the knowledge base. These predicates or facts can be managed by rules to infer new knowledge that reveals more sophisticated relations that typical analysis are not aware of, or are out of their scope. With this information, it is possible to automatically generate richer descriptions that make sense of the map as it 'explains' more attributes about each spatial object. The rules presented are "universal" in the sense that they can be used for any given map to generate its description as long as its predicates are formed in the same way suggested in this work.

The advantages of this approach is that the analysis and inference processes are executed only once for each map, since the new knowledge is stored on the knowledge base, which is the foundation to make more complex analysis and generate richer descriptions. Also, this approach is used to *share* new data and provide them for spatial decision support systems.

Future works are related to allow us changes on the map at runtime in order to update the knowledge base, which should reflect the new state of the relations between the spatial objects of the map.

References

1. Openshaw, S. and Openshaw, C.: *Artificial Intelligence in Geography.* John Wiley & Sons (1997).
2. Egenhofer, M.: Query Processing in Spatial-Query-by-Sketch, *Journal of Visual Languages and Computing,* 8(4), (1997) 403-424.
3. Bruns, T. and Egenhofer, M.: "Similarity of Spatial Scenes", *Seventh International Symposium on Spatial Data Handling,* M.-J. Kraak and M. Molenaar, eds., Delft, The Netherlands, (1996) 31-42.
4. Mark, D.: "Counter-Intuitive Geographic "Facts:" Clues for Spatial Reasoning at Geographic Scales" *Theories and Methods of Spatio-Temporal Reasoning in Geographic Space, Pisa, Italy,* A. Frank, I. Campari, and U. Formentini, eds., *Lecture Notes in Computer Science,* 639, Springer-Verlag, Berlin, (1992) 305-317.
5. Stevens, A. and Coupe, P.: "Distortions in Judged Spatial Relations" *Cognitive Psychology,* vol. 10, (1978) 422-437.
6. Egenhofer, M. and Franzosa, R.: "Point-Set Topological Spatial Relations", *International Journal of Geographical Information Systems,* 5(2), (1991) 161-174.
7. Egenhofer, M. and Herring, J.: Categorizing Binary Topological Relationships Between Regions, Lines, and Points in Geographic Databases, *Technical Report of the Department of Surveying Engineering*, University of Maine, Orono, ME, (1991).
8. Frank, A.: "Qualitative Spatial Reasoning about Cardinal Directions", *Autocarto 10,* D. Mark and D. White, eds., Baltimore, MD, (1991) 148-167.
9. Randall, D. and Lenat, D.: *Knowledge-Based Systems in Artificial Intelligence.* Advanced Computer Science Series, Mc Graw Hill, (1982).

10. Verastegui, K.: Conceptualization of Geometrical Aspects of Geospatial Data*, Proceedings of the First International Conference on Geospatial Semantics*, Poster Session GeoS 2005. ISBN: 970-36-0304-2 (2005).
11. Rosales, M.: Semantic Topological Descriptor for Topographic Maps*, Proceedings of the First International Conference on Geospatial Semantics*, Poster Session GeoS 2005. ISBN: 970-36-0304-2 (2005).
12. Townsend, C.: *Mastering Expert Systems with Turbo Prolog*. Howard W. Sams & Company, First Edition, (1987).
13. Torres, M., Moreno, M., Quintero, R. and Fonseca, F.: Ontology-driven description of spatial data for their semantic processing. *Proceedings of the First International Conference on Geospatial Semantics*, Springer-Verlag, 3799, Mexico City, Mexico (2005) 242-249.