Uncertainty handling in spatial relationships

Abstract
It is recognized that the existing spatial relations do not cover the areas of metrics (i.e. to what degree are two objects A, B overlapping, or to what degree they meet), and uncertainty (i.e. objects A, B overlap significantly, object A is far away from B etc.). The goal of the paper is to deal with the *uncertainty* of spatial relationships on the assumption of *crisp* spatial objects. To achieve our goal we investigate the combination of spatial relationships with *metric* properties. For topological relations our interest refers to the degree of overlapping and adjacency between two spatial objects. For direction and distance relationships we would like to be able to model smooth and continuous transitions between different directions and between different concepts of distance (e.g. *far* and *very far*), respectively. All these applications embody the inherent property of uncertainty. We use fuzzy methodologies for modeling these uncertainty features. Also a query processing scheme for binary spatial relationships with uncertainty is presented.

1. Introduction

Spatial relationships are binary predicates yielding a Boolean decision whether a certain relationship holds between two spatial objects or not[Ege91,Pap95]. They are influenced by several factors like the underlying natural language, the semantics of the objects involved, the context in which objects appear, and the geometric structures of the objects and their extent. For expressing spatial relationships, people usually favor qualitative models of space, i.e. models that offer natural-language terms for these relationships. These models come closer to human thinking than quantitative models which are metric-based, involve precise distance and angular measures, for instance, and are primarily supported in current GISs. Queries illustrating the users' needs are for example: "Is object *A inside* or *adjacent* to object *B*?", "Is object *A north* or *left* of object *B*?", or "Is object *A far away* from or *near* object *B*?", or "which are the objects that are *far away* and *north* to object A?". These example queries also show that spatial relationships can imply precision or certainty (e.g. in the case of *inside* and *adjacent*) but also uncertainty or vagueness (e.g. in the case of *north*, *left*, *far away*, and *near*).

It is recognized that the existing spatial relations do not cover the areas of metrics (i.e. to what degree are two objects A, B overlapping, or to what degree they meet), and the area of uncertainty (i.e. objects A, B overlap significantly, object A is far away from B etc.) [Alt94]. Support of such concepts to a satisfactory degree, would greatly assist and broaden spatial decision making.

Spatial relations are currently designed as binary predicates yielding a Boolean and thus strict decision whether a certain relationship holds for two spatial objects or not. Well-known examples are the topological predicates like *overlap*, *meet*, *equal*, *disjoint*, and *inside* [Ege91]. Spatial databases store information about spatial objects, and most likely their extent. Thus is feasible to compute their spatial relationships (topological, directional, distance) in a Boolean manner (i.e. A overlaps B) as well as in a quantitative one, using their metrics (i.e. A is 23 m away from B).

Uncertainty in the context of the current work is a concept that represents the belief that is associated with the metrics of spatial relationships. For instance the phrase "A overlaps a lot B" bears inherent uncertainty as it represents an assessment of the degree to which the object A overlaps B. Similar nature of uncertainty is conveyed by the phrase: "vessel A is north to island B", since vessel A could be anywhere between NW and NE to island B. In a similar manner consider the phrase "City A is far from village B", again we have a qualitative characterization that gives an estimation of the distance. Of course here the context is important as the observer has some predefined perception of what is far of close.

As it is apparent the quantitative aspects of the relationships are supporting the qualitative characterizations that were described above.

The objective of this paper is to bridge the gap between i. the quantitative aspects of relationships between spatial objects and ii. their qualitative characterizations that often occur in queries. There is a great amount of knowledge that resides in the spatial and temporal relations among objects in a relevant framework. We are interested in modeling this knowledge and reason about this taking in account the uncertainty related to spatial relations' features (topology, direction, metrics).

The paper deals with the *uncertainty* of spatial relationships on the assumption of *crisp* (i.e. sharply bounded) spatial objects. The scope of this paper comprehends relationships between pairs of regions. To achieve our goal we investigate the combination of spatial relationships with *metric* properties (such as normalized length and area measures) of the participating crisp spatial objects. This procedure leads to a *refinement* and thus to an extension of existing spatial relations and to a coupling of their qualitative and quantitative aspects. For topological relations our interest refers to the degree of overlapping and adjacency between two spatial objects. If two regions *A* and *B* intersect, certain metric properties can be obtained, i.e. the area of the intersection compared to the area of one of the participating objects. We design a model that normalizes metric values for relationships between regions with respect to the length of the line, the area of a region, and the perimeter of a region. It then makes a difference whether object *A* intersects object *B* to a large or only to a small extent, and the degree of object *A* intersecting object *B* will usually be distinct from the degree of object *B* intersecting object *A*. For direction and distance relationships we would like to be able to model smooth and continuous transitions between different directions and between different concepts of distance (e.g. *far* and *very far*), respectively. All these applications embody the inherent property of uncertainty. We will use fuzzy logic methodologies for modeling these uncertainty features.

The paper organization follows: in the next section we present related work, while in the next section we briefly introduce the notions of fuzzy logic that are exploited in this paper. In Section 4. The model of spatial relations is presented. In Section 5 a query processing scheme for spatial relation queries is presented along with the evaluation of this scheme. Finally we conclude by summarizing and discussing further research and development directions.

2. Related Work

The exploration of *spatial relationships* between objects in space has turned out to be a multi-disciplinary research issue involving disciplines like linguistics, cognitive science, psychology, robotics, artificial intelligence, geography, computer science, and mathematics. Besides the definition of spatial data types, models of spatial relationships (e.g. [Cui93, , Ege91, Peu87, Ran92, Sa85]) in particular play a central role in *geographical information systems* (*GIS*) [Bu86] and *spatial database systems* [Gue94]. From a database perspective, their developments were motivated by the necessity of formally defined spatial operators (predicates) in spatial query languages, both at the query definition level and at the query processing level.

Based on different mathematical concepts, spatial relationships can be classified into three major categories. *Topological relationships* (e.g. *inside*, *adjacent*, *overlap*) [Cui93, , Ege91] use topological properties like interior, boundary, connectivity, overlapping, and inclusion for their description. They require the concept of neighborhood and are invariant under topological transformations like translation, rotation, and scaling. *Directional relationships* (e.g. *north*, *left*) [Peu87][Pap95] are relative concepts and based on the existence of a vector space. They compare the relative position between two spatial objects and are subject to change under rotation, while they are invariant under translation and scaling. *Distance relationships* (e.g. *near*, *far*) [Her95] are based on the concept of metric and therefore change under scaling but are invariant under translation and rotation. Sometimes *order relationships* [Ka90] are viewed as a fourth category of spatial relationships. They are based on the definition of strict and partial order and have inverse relationships. Examples are *behind*/*in_front_of*, *below*/*above*, *inside*/*contains* and *under*/*over*. Order relationships are applicable to any of the three categories of spatial relationships since they establish an ordering over topological, directional, and distance relationships. They are not considered in this paper.

The term *spatial reasoning* is usually used for describing the process of investigating how people perceive constellations of spatial phenomena or facts related with time, how they reason about such constellations, how they express them in languages, what kind of queries they pose, and how this gained knowledge can be represented in models. There have been many efforts in this area. Related work in topology-direction reasoning has been done on (i) the composition of topological and orientation relations taken together [Her94] with the result being pairs of topological/orientation relations; (ii) defining direction relations between extended objects in terms of interval relations facilitating the retrieval of spatial objects from a database using an R-tree indexing mechanism [Pap95]; an iii. defining direction relations between extended objects using the 4-intersection [Abd94].

Spatial reasoning currently lacks of the natural language richness. Early enough we find evidence for requirements on spatial queries involving uncertainty [All84]: *"Park it in the space next to that tree", " I wandered around downtown, trying to keep the highway entrance within sight", and "I stood at the intersection where the Lexington Avenue enters Harlem", "under the shadow of a huge tenement block".* In the above queries we clearly have evidence of the necessity for representation and handling of uncertainty related to facts of real life in the context of spatial locations and relations.

In the area of GISs there have been some efforts aiming at incorporation uncertainty in modeling and decision support related to geographic data [Alt94, Ste96, Ste99, Wan90, Wan94]. More specifically In [Ste96, Ste99] we worked on the issue of incorporation of fuzzy set methodologies into a DBMS repository for the application domain of GIS which is beneficial and will improve its level of intelligence. Focusing on this direction our work addresses both a representation and a reasoning issue. Specifically, it extends a general spatial data model to deal with the uncertainty of geographic entities, and shows how the standard data interpretation operations available in GIS packages may be extended to support the fuzzy spatial reasoning. Representative geographic operations, such as the fuzzy overlay, fuzzy distance and fuzzy select, are examined, while several real world examples are given.

At present, Geographic Information Systems (GIS) though powerful toolboxes, most with hundreds of functions, suffer from several limitations which render them inefficient tools for spatial decision-making.

3. Elements from Fuzzy Logic

Fuzzy set theory is an extension or generalization of classical Boolean set theory and aims at representing the degree to which an object belongs to a set.

In this section we describe briefly some concepts and techniques from fuzzy sets that we exploit in this research effort. In many cases we have objects partially belonging to a set O. This classification property may well be represented by the degree of belief (d.o.b.) concept introduced in the fuzzy logic domain. Then each set of objects $S=\{ s_i \}$ creates a fuzzy set $Fs = \{s_i, \mu_{si} \mid s_i \in S, \mu_{si} \in [0,1]\}$.
Notice that classical sets allow only binary membership functions (i.e., TRUE or FALSE). Fuzzy set theory [Zad68] is an extension of the classical set theory dealing mainly with the partial classification (or other ways belief) of an object s_i to a set S. The belief is represented by the degree of belief μ_{si} which is a real number in the range [0,1], where 0 indicates no-membership and 1 indicates full membership and specifies the extent to which s_i can be regarded as belonging to set S.

In order to obtain the beliefs mentioned above, real world quantitative values are mapped to the fuzzy domain values (i.e. to the domain [0,1]) through mapping functions. These functions may be very diverse and they play a vital role in the representation of uncertainty as fuzzy values. A thorough presentation of mapping functions and their properties is found in [Kli95].

The choice of the membership function, i.e., its shape and form, is crucial and strongly affects the results derived from a decision-making process. In correspondence to classical set theory, two options are available for choosing the membership functions for fuzzy sets [Bur96] a) through an imposed "expert" model; and b) by a data driven multivariate procedure. Commonly used transformation functions include [Gup88] linear (decreasing, increasing, triangular), S and P functions, trapezoidal etc.

In many application domains we are interested in evaluating S as for its objects' property belonging to O. Then we search for an overall belief that the objects of S belong to O. Essentially we are searching to what degree S is a subset of O. This measure may be quantified by the Energy Metric function defined in [Gup88] as:

$$E_S = \Sigma \mu_O{}^q (s_i)$$

where q is a positive integer. A usual choice is q=2. Assuming A={a_i}, B={ b_j } two sets and let "◊" a relationship among the sets A, B, and let $\mu_{ai◊bj}$ the d.o.b. that the pair (a_i, b_j) belongs to the set of objects related with the relationship "◊". e define the energy metric A◊B of the relationship "◊" as:

$$\forall\ (x,y),\ x\in A,\ y\in B:\ E_{A◊B} = \Sigma \mu_{A◊B}{}^q (x,y)$$

which represents the overall belief that A, B are related through the relationship "◊". We will exploit this information measure in the following section extensively.

4. The integrated spatial relation model

In this section we propose a set of modeling primitives representing the uncertainty of spatial relations and in a spatial context. We propose a way of representing uncertainty that is inherent in the classification of the metrics of spatial relations and how this is reflected to the various aspects of such relations, namely: *topology, direction* and *distance*.

Uncertainty may be related to the spatial objects as well as to their relations. As regards objects there is the uncertainty related to the belief whether a point belongs or not to the extent of the object and to what degree. In this research effort we neglect this aspect assuming that the spatial objects under concern are *crisp*, i.e. it is a Boolean decision whether a point of the 2D space belongs to an object or not. We will rather concentrate in the uncertainty related to the binary spatial relations and their various aspects.

Most of the spatial relationship schemes (see the section on related work) proposed so far only deal with the qualitative aspects of the relations avoiding the quantitative attributes that in some cases may be of vital importance. The need for metrics is apparent, this issue has not been addressed adequately. Thus we introduce the *distance* as a useful aspect of a spatial relationship (i.e. far, close), which has been under-addressed.

A comprehensive model for dealing with the uncertainty of spatial relations enriched with metrics is hereafter presented. In this model, we attempt a refinement of the spatial relations in order to support metrics and moreover the use of the metrics to classify a relationship in the appropriate lexical-qualitative categories. We present our efforts in the aspects of spatial relations, namely *topological, direction, distance.*

In this design we make a set of assumptions that simplify the design of the modeling primitives:

- Spatial objects are assumed to be two-dimensional convex polygons
- for each spatial object *A* there is a function *cent(A)* that returns a point in 2D space which is the center (*centroid*) of the object

4.1. Topological relations

In this section we introduce a minimal set of primitives that enable the complete representation of the qualitative and quantitative features of topological relations. The minimal set of primitives are namely: *overlapping* and *adjacency*.

Observation: *The topological relations are quantitatively not commutative*
Intuitively the degree of overlapping between A and B is the degree to which A *occupies* B. In Figure 1 it is clear that the validity of the statement: "A occupies B" is greater than the validity of the statement: "B occupies A". This is justified as follows: although the area of overlap is common the "importance" of this area as part of each object is different. As it is clear the overlap relation is not commutative, although this is true for the qualitative aspect of the relations [Egen91].

Similarly the adjacency relation between the boundaries of two objects needs to be enriched with metrics that also do not preserve commutativity. This is illustrated in Figure 2 where the degrees of *adjacency* between the pairs of objects (A, B) and (A, C) are different although the common boundary is of equal length.

In the sequel the uncertainty measures for overlapping and adjacency between two objects are defined.

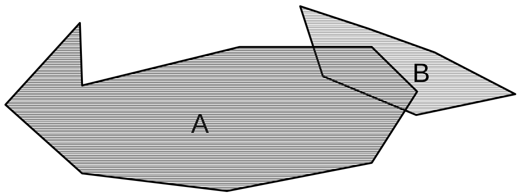


Figure 1. Two spatial objects overlapping

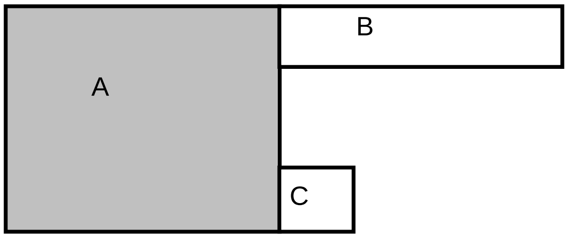


Figure 2. Degrees of adjacency between A, B and A, C are different although the common boundary is equal.

Here we have to clarify some details and assumptions:

- the assertion: “A overlaps B”, conveys that A is also above B (the visible common *area* belongs to A)
- let max(A∩B) the maximum possible overlap between A and B. This value is either A if *area*(A)< *area* (B) or B if *area*(A)> *area* (B) or some other region X= A∩B such that *area*(X)< *area*(A) if the shapes are such that A cannot be covered by B

Definition 1: Let two spatial objects *A, B*, their interior (A^o) and boundary (ϑA) [Ege91], then the degree of belief that *A overlaps* B is defined as follows:

$$\mu_{\text{intersection}}(A, B) = \frac{A^o \cap B^o}{B^o} \qquad (1)$$

For each of the topological relations we investigate the feasibility of uncertainty attachment. For instance, the relationship “overlaps” conveys uncertainty, depending on the size of the overlapping area related to the objects area. A linear increasing function is exploited to model the confidence that two objects overlap each other according to the part of object related to of overlapping. Assuming that object A “invades” object B and that the area of A is not bigger than the area of B and the objects’ shapes are such that A can totally cover B. When the part of A overlapping B grows to 100% of B’s area, this implies that either the spatially larger object (A) includes the smaller (B) or the objects are of equal area and have identical shapes. In the case that the area of B is bigger that the area of A, the degree of belief is always smaller than one (1).

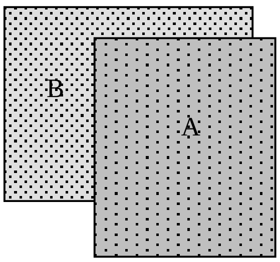


Figure 3. Topological relationship

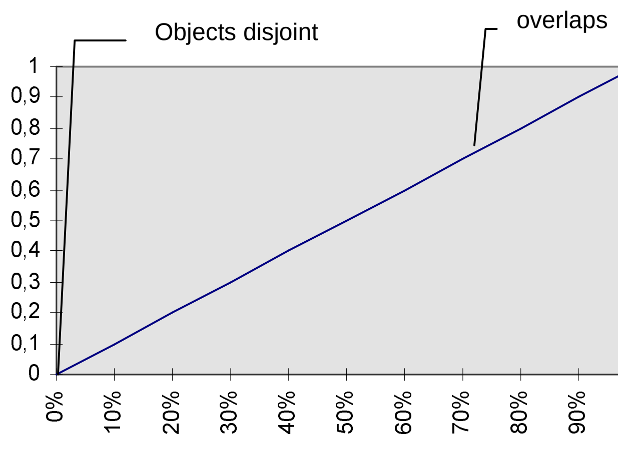


Figure 4. The mapping function regarding the overlapping features between two objects

As regards adjacency we are interested in the degree to which the boundaries of the objects A,B coincide.

Definition 2: Assuming two spatial objects *A, B* and their boundary $(\vartheta A,\ \vartheta B)$, then the degree of belief that *A* is *adjacent* to *B* is defined as follows:

$$\mu_{\text{adjacency}}(A,B) = \frac{\partial A \cap \partial B}{\partial B} \qquad (2)$$

The $\mu_{adjacency}(A,B)$ is defined as the degree to which the boundary of *A* coincides with the boundary of *B*.
In a similar way we process the adjacency measure to produce a belief related to adjacency. We exploit the uncertainty measures introduced before for overlapping and adjacency in order to represent the rest of the topological relations.

As we mentioned before the established set of topological relations [Ege91] conveys only qualitative information and thus it can be regarded as a subset of the relations framework we propose here, while the topological relations result as a combination of the above metrics. In Table 1 the reader can find the mapping of the qualitative topological relations to the degrees of overlapping and adjacency as we have defined them.

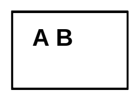
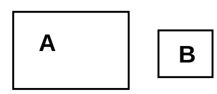
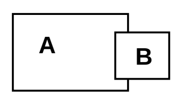
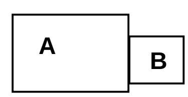
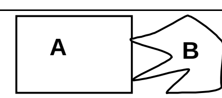
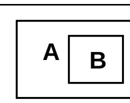
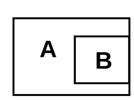
	$\mu_{overlap}(A,B)$	$\mu_{adjacency}(A,B)$	
Equal	1	1	
Disjoint	0	0	
Overlap	(0,1)	0	
Adjacent	0	(0,1)	
Meet	0	a non-infinite set of points	
Inside (Contain)	1	0	
Covers (Covered by)	1	(0,1)	

Table 1: The topological relations related to the degrees of overlap and adjacency.

4.2. Direction relations

Here we establish a representation of the uncertainty conveyed by the quantitative features of direction relations. The well-established set of direction relations is the quartet: {north, east, south, west} or alternatively {above, right, below, left}[Pap95]. Here we are interested in determining the uncertainty of a direction relationship among two spatial objects *A* and *B*. This is going to be deduced from the value of the angle φ between the line segment connecting centroids of *A* and *B* and the horizontal axis (see Figure 5). Here we make the assumption that the objects *A*, *B* are not overlapping and their distance is big compared to their extent. In the case of large objects being close to each other then the computation of the direction relationship has to take in account:

- the related objects have some extent that contributes to the direction of the relationship
- each point of the object's extent contributes to the direction relationship according to the attached weight.

The value of angle φ is mapped to the fuzzy domain for each direction relationship according the appropriate mapping functions (in this case we picked up the sinusoidal one) as it appears in Figure 6. The intuitive result is that for each pair of objects there is potentially more than one direction relations that are partially valid for that pair (i.e. in Figure 5 the object A is both "south" and also "south-west" to object B). The direction relationship between the two objects will be defined by the angle formed by their centroids. In the current approach we do not consider the contribution of the rest of the objects' extent.

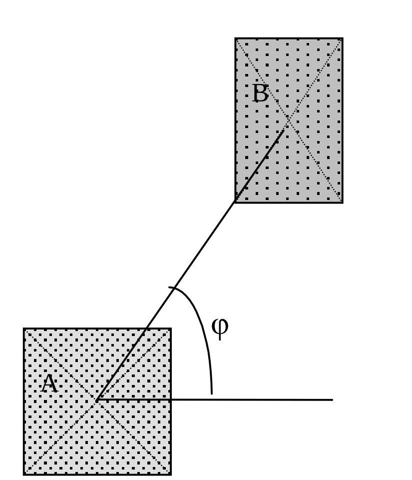


Figure 5.The angle between the two objects that determines the direction relationship

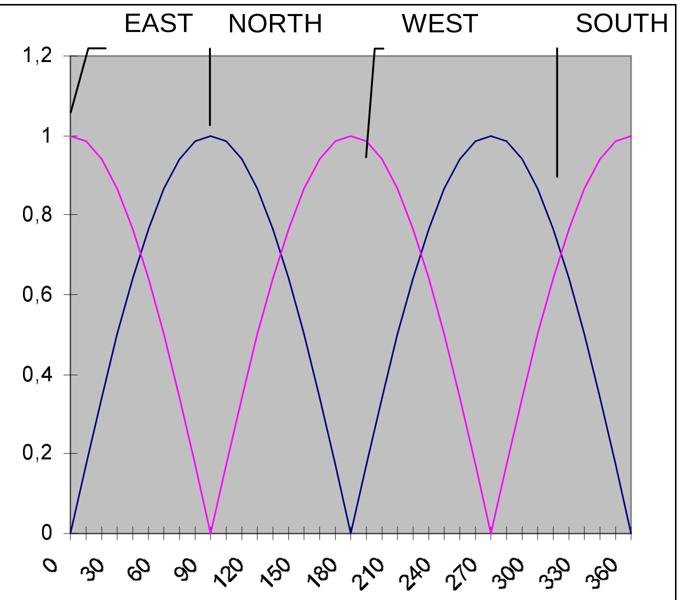


Figure 6. The mapping functions regarding the direction relationship between two objects

Definition 3: Let *A*, *B* two spatial objects and *cent(A)*, *cent(B)* their centroids. Let φ be the angle between the line segment connecting centroids of A and B and the horizontal axis. Then the belief that the direction relationship between *A* and *B* is:

$$\mu_{dir\_rel}(\varphi), \text{ where } dir\_rel \in \{east, north, south, west\}$$

and more specifically:

$$\mu_{east} = \begin{cases} \sin(\varphi + 90) & \text{for } 270 < \varphi < 360 \text{ or } 0 < \varphi < 90 \\ 0 & \text{otherwise} \end{cases} \qquad (3)$$

$$\mu_{north} = \begin{cases} \sin(\varphi) & \text{for } 0 < \varphi < 180 \\ 0 & \text{otherwise} \end{cases} \qquad (4)$$

$$\mu_{west} = \begin{cases} \sin(\varphi - 90) & \text{for } 90 < \varphi < 270 \\ 0 & \text{otherwise} \end{cases} \quad (5)$$

$$\mu_{south} = \begin{cases} \sin(\varphi - 180) & \text{for } 180 < \varphi < 360 \\ 0 & \text{otherwise} \end{cases} \quad (6)$$

The result of the above procedure is a tuple for each pair of objects (A,B) of the form: $<(A, B), \mu_{north}(\varphi), \mu_{east}(\varphi), \mu_{south}(\varphi), \mu_{west}(\varphi)>$ representing the degree to which the direction relationship among A and B is north, or east, or south or west. This model can easily be extended to cover the intermediate direction relations such as: north-east etc.

4.3. Distance relations

As regards the distance we need to take in account the context of the application, (i.e. what is considered as big or small in the application or what is considered as far or close. It is assumed that the distance is the one among the centroids of the objects, see Figure 7. This distance is mapped to the fuzzy domain taking in account the maximum context distance and we exploit linear mapping functions as depicted in Figure 8.

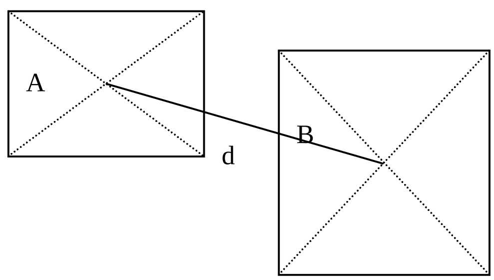


Figure 7. The distance between two spatial objects

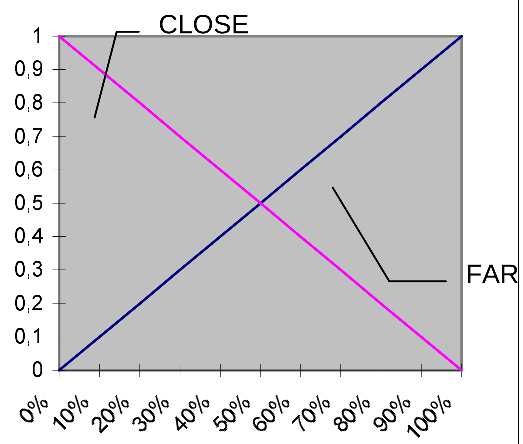


Figure 8. The mapping functions for the distance between two objects X,Y.

Again we remind the assumption that the objects A, B are not overlapping and their distance is big compared to their extent. In the case of large objects being close to each other or even overlapping then the computation of the distance is not a trivial issue and we have to take in account the special requirements of the application domain.

Definition 4: Let A, B two spatial objects and *cent(A)*, *cent(B)* their centroids. Let d_{max} be the maximum distance in the application context and d the distance of the line segment connecting centroids of A and B. Then the belief that the objects A and B are far (close) is given by the formulas:

$$\mu_{far}(A,B) = \frac{d}{d_{\max}} \quad (7)$$

$$\mu_{close}(A,B) = \frac{d_{\max} - d}{d_{\max}} \quad (8)$$

As it is clear from the above: $\mu_{far}(A,B)+\mu_{close}(A,B)=1$. Thus we can keep one of them (i.e. $\mu_{close}(A,B)$) in the model and discard the other. As for the maximum distance, it can be defined as the maximum distance between any two objects of the application context.

Another point to be stressed is that the distance relations are symmetric, i.e. $\mu_{close}(A,B)=\mu_{close}(B,A)$.

4.4. The spatial relationship model

Hereafter we define a relationship model that integrated the aforementioned aspects in one relationship and provides a mechanism for evaluating queries related to one or more of the features of the spatial relationship hereafter called Spatial Relationship Features (SRFs).

Definition 5: Let *A rel B* such an expression where *A*, *B* spatial objects and *rel* their spatial relationship including the attached uncertainty.
The relationship *rel* is a tuple of the form:

rel = (top, dir, dist)

where

$top = (\mu_{\text{intersection}}(A,B), \mu_{\text{adjacency}}(A,B))$

$dir = \{\mu_{\text{dir\_rel}}(\varphi) |$ where $dir\_rel \in \{east, north, south, west\}\}$

$dist = (\mu_{\text{close}}(A,B))$

as defined in the previous sections. The constituents of the tuple are the relations SRFs.
An interesting issue is that using the aforementioned scheme we can represent not only relations between two objects but also the spatial relationship of an object to the spatial origin of the context. Indeed, if we assume an object *A* and let *Θ* the spatial origin of the context. We assume *Θ* is an object whose interior is the empty set and its boundary is reduced to one point. Then the object *A* can be represented as it relation to the origin: Θ *(top, dir, dist) A,* where *(top, dir, dist)* defined as above.

5. Spatial relationship queries processing

In this section we propose a mechanism for evaluating spatial relationship queries submitted to a database of spatial objects. Queries in this context are classified in two categories:

- Query by example: the user provides a sample spatial relationship searching for similar relationships in the database.
- SQL-like queries: the user searches of spatial objects that have a given spatial relationship with a reference object.

In this research effort we consider only pairs of objects. The issue of object groups that form n-ary relationships will be tackled in future research work. We will simplify the problem assuming queries searching for pairs of objects similar to a query pair provided by a user.

For the spatial relationship we need to find a spatial similarity measure between the query pair and each pair of objects in the spatial context. Let a visual query Q depicted in Figure 9. From this query the system extracts the SRFs of the spatial relationship in terms of the degrees of belief introduced in the previous section:

A rel B = A (top, dir, dist) B

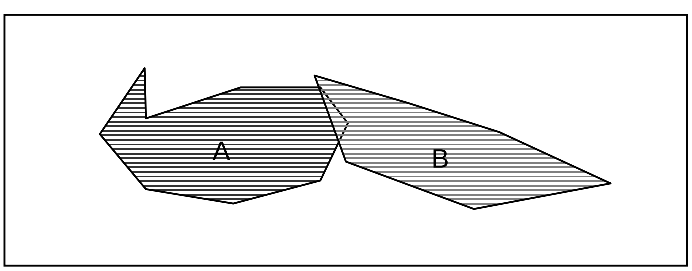


Figure 9. A visual query conveying information about the SRFs of the spatial relationship among two objects

Then for each pair of objects *(Xi, Xj)* of the context we can compute the corresponding relation tuple:

Xi rel Xj = Xi (top, dir, dist) Xj

For each of the SRFs a similarity measure (sm_i) is computed, which represents the similarity of the pair *(Xi, Xj)* to the query pair *(A,B)* as regards the *SRFi*. We exploit the Euclidean distance as a similarity measure[1]:

$$sm_i = 1 - \sqrt{(\mu_{SRFi}(A,B) - \mu_{SRFi}(Xi,Xj))^2} \quad (9)$$

where SRFi$\in \{\mu_{\text{intersection}}(A,B),\ \mu_{\text{adjacency}}(A,B,\ \{\mu_{\text{dir_rel}}(\varphi) |$ where $dir\_rel \in \{east, north, south, west\}\},\ \mu_{\text{close}}(A,B)\}$

The last step is to compute the overall similarity measure (*sm*) between the two pairs by aggregating and normalizing the partial sm_i s:

$$sm = \left(\frac{\sum_{k=1}^{7} sm_i}{7 * \max(sm_i)} \right) \quad (10)$$

The result of this procedure is a list of pairs of the form <(Xi, Xj), sm(Xi, Xj)> rating the overall similarity of the spatial relationship between the objects of the pair (Xi, Xj) to the query pair (A,B). Apparently the users may select the pairs (Xi, Xj) that have the higher values for sm.

Assume a visual query: “Find pairs of objects having a spatial relationship similar to the one appearing in Figure 10”.

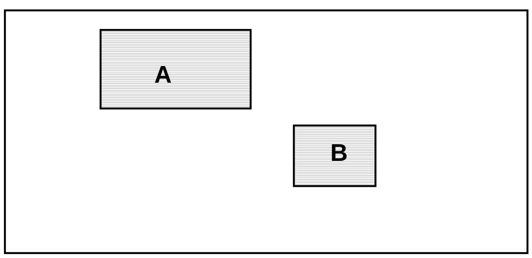


Figure 10. The spatial relationship query

From this figure there is a set of SRFs extracted which are represented by the tuple:

A rel B = A (top, dir, dist) B

Then a matching procedure starts, in order to identify the session with similar spatial relations and also temporally close to the desired time. In processing this query, it is assumed that in a session there are more than one pairs of objects so we have more than one spatial relations to examine. Then for each pair of objects we have to extract a normalized spatial similarity measure representing the overall belief that the session contains pairs similar to the ones of the query. Thus we have to normalize the similarity measure resulting from aggregating the similarity measures for each pair. Consequently the spatial similarity has to be aggregated with the temporal similarity so as to provide an overall measure of similarity on the basis of which the users will select the most relevant sessions. The algorithm follows:

[1] An altenative here could be the weighted Euclidian distance, increasing the computational cost though.

```
S={}                              // the set of session similarities
for each pair of objects (Xi,Xj)
        sm = 0                    // the spatial similarity measure
        for each SRF
                compute smi       // as in equation 9
        compute sm                // as in equation 10
        add sm to S
```

The cost of the above algorithm is approximated as follows: Let *n* the number of spatial objects in the database thus we have to compute the spatial similarity for

$$\binom{n}{2} = \frac{n!}{(n-2)!2!} \qquad (11)$$

pairs (irrelevant of order) that may be extracted from them, hence $2*\binom{n}{2}$ ordered pairs. For each of these pairs we compute the overall belief that the pair fulfills the criteria set by the SRFs as depicted in the relation A rel B. Hence the overall approximate computational cost of the above algorithm is:

$$(2*\binom{n}{2})*\mathrm{cost}(E\_sp\_similarity)$$

Thus the cost of the similarity search is $O(n^2)$ where n is the average number of objects in the database. This result can be optimized by using heuristics that result from knowledge of the specific context.
In the case of SQL like queries, a reference pair of objects that totally fulfils the criteria of the query is assumed and then the query processing scheme that was presented above is exploited. For instance consider the query: "*Find the **closest** and **most north** eastern island to Athens*" we will assume a reference pair of objects X,Y that fulfil the criteria *close* and *north* totally ($\mu_{close}(X,Y)=1$ and $\mu_{north}(X,Y)=1$). Then

For each pair (X,Y) where X= Athens and Y $\in$ Islands set the *sm* is computed and thus the desired island results. In this case the complexity of the query processing plan is O(n) where n is the number of islands, since the first member of the pair is constrained to the value "Athens".

6. Conclusions

In this paper we presented a model to represent and handle the uncertainty inherent in spatial relations. We exploited techniques from fuzzy logic as the most appropriate methodology to deal with uncertainty which is related to human perception and classification. The model represents in an integrated way the uncertainty of qualitative terms based on quantitative aspects of the various aspects of spatial relations, namely topology, direction and metrics. We also presented a related query processing scheme for binary spatial relations bearing uncertainty related to SRFs.

The proposed scheme will be extended so as to cover fully the following topics related to:

- Representation of the uncertainty of more features of spatial objects such as: *position, shape and extent*. This issue involves selection of the appropriate mapping functions to represent the degree to which a point belongs to an object. The involvement of experts in the area would be essential.
- Study of the impact of various mapping functions to the quality of the reasoning mechanism.
- Develop further the model mechanism to deal with: i. n-ary spatial relations (i.e. not only a pair of objects but a set of objects spatially related), this will make your relevance feedback mechanism much more flexible and meaningful and ii. temporal relations among facts and intervals.
- Adopt query processing and indexing techniques for high dimensional data to improve performance when large databases are present. Although this is not straight forward due to the non crisp nature of the relations and of the queries which implies that the range of a query is not a crisp object itself.

References

[Abd94] Abdelmony, A, and Williams M.H, “Approaches to the representation of qualitative spatial relationships for Geographic databases”, in M. Molenaar (Ed.) Advanced Gfeographic Data modeling (AGDM 94), Delft, Netherlands, Netherlands Geodetic Commission.

[All84] James F. Allen, “Towards a general theory of action and time”, "Artificial Intelligence Journal, vol. 23, pp123-154, 1984.

[Alt94] Altman, D., 1994, Fuzzy set theoretic approaches for handling imprecision in spatial analysis. International Journal of Geographical Information Systems, 8, pp. 271-289.

[Bu86] P.A. Burrough. *Principles of Geographic Information Systems for Land Resources Assessment.* Oxford University Press, 1986.

[Cui93] Z. Cui, A.G. Cohn & D.A. Randell. Qualitative and Topological Relationships in Spatial Databases. *3rd Int. Symp. on Advances in Spatial Databases* (*SSD'93*), Springer-Verlag, LNCS 692, pp. 296-315, 1993.

[Ege91] M.J. Egenhofer & R.D. Franzosa. Point-Set Topological Spatial Relations*. Int. Journal of Geographical Information Systems*, vol. 5, no. 2, pp. 161-174, 1991.

[Gue94] R.H. Gueting, “An Introduction to Spatial Database Systems”. VLDB Journal 3 (1994), 357-399.

[Gup88] M.M. Gupta et al., "Sinusoidal - Based cognitive mapping functions", in Fuzzy Logic in Knowledge Based systems, decision and control, North-Holland, pp. 69-92, 1988.

[Her95] D. Hernandez, E. Clementini & P. Di Felice. Qualitative Distances. 2nd Int. Conf. on Spatial Information Theory (COSIT), LNCS 988, pp. 45-57, 1995.

[Ka90] W. Kainz. Spatial Relationships - Topology Versus Order. Proc. of the 4th Int. Symp. on Spatial Data Handling, pp. 814-819, 1990.

[Kli95] G. Klir, B. Yuan, “Fuzzy Sets and Fuzzy Logic, Theory and Applications”, Prentice Hall, 1995

[Pap95] D. Papadias, Y. Theodoridis, T. Sellis, M. Egenhofer, “Topological relations in the world of minimum bounding rectangles: a study with R-trees, in the proceedings of the ACM-SIGMOD conference, 1995

[Peu87] D.J. Peuquet & Z.C.-Xiang, “An Algorithm to Determine the Directional Relationship between Arbitrarily-Shaped Polygons in the Plane”, Pattern Recognition, vol. 20, no. 1, pp. 65-74, 1987.

[Ran92] D.A. Randell, Z. Cui & A.G. Cohn. A Spatial Logic Based on Regions and Connection. Principles of Knowledge Representation and Reasoning (KR'92), pp. 165-176, 1992.

[Ste96] E. Stefanakis, M. Vazirgiannis, T. Sellis, "Incorporating Fuzzy Logic Methodologies into GIS Operations", in the proceedings of the International Conference in Urban, Environmental and Regional Planning, Greece, April 1996.

[Ste99] E. Stefanakis, M. Vazirgiannis, and T. Sellis, "Incorporating Fuzzy Set Methodologies in a DBMS Repository for the Application Domain of GIS", to appear in IGIS Journal, 1999.

[Wan90] F. Wang, B. Hall, G., Subaryono, 1990, “Fuzzy information representation and processing in conventional GIS software: database design and application”. International Journal of Geographical Information Systems, 4, pp. 261-283.

[Wan94] F. Wang, 1994, “Towards a natural language user interface: an approach of fuzzy query”, International Journal of Geographical Information Systems, 8, pp. 143-162.

[Zad65] Zadeh, L., "Fuzzy Sets," Information and Control 8:338-353, 1965.

[Zad88] Zadeh, L., "Fuzzy Logic," IEEE Computer 21 (4): 83-93, 1988.