Fuzzy Spatial Data Types and Predicates: Their Definition and Integration into Query Languages

Markus Schneider

University of Florida
Department of Computer & Information Science & Engineering
E450 CSE Building
PO Box 116120
Gainesville, FL 32611-6120
mschneid@cise.ufl.edu

1 Introduction

Representing, storing, quering, and manipulating spatial information is important for many non-standard database applications. Specialized systems like geographical information systems (GIS) and spatial database systems to a certain extent provide the needed technology to support these applications. So far, spatial data modeling has implicitly assumed that the extent and hence the borders of spatial phenomena are precisely determined, homogeneous, and universally recognized. From this perspective, spatial phenomena are typically represented by sharply described *points* (with exactly known coordinates), *lines* (linking a series of exactly known points), and *regions* (bounded by exactly defined lines which are called *boundaries*). Special data types called *spatial data types* (see [23] for a survey) have been designed for modeling these spatial data. We speak of *spatial objects* as instances of these data types. The properties of the space at the points, along the lines, or within the regions are given by attributes whose values are assumed to be constant over the total extent of the objects. Well known examples are especially manmade spatial objects representing engineered artifacts like highways, houses, or bridges and some predominantly immaterial spatial objects exerting social control like countries, districts, and land parcels with their political, administrative, and cadastral boundaries. We will denote this kind of entities as *crisp* or *determinate* spatial objects.

Increasingly, researchers are beginning to realize that the current mapping of spatial phenomena of the real world to exclusively crisp spatial objects is an insufficient abstraction process for many spatial applications and that the feature of *spatial vagueness* or *spatial indeterminacy* is inherent to many

geographic data [4]. Moreover, there is a general consensus that applications based on this kind of indeterminate spatial data are not covered by current GIS and spatial database systems. In this chapter we focus on a special kind of spatial vagueness called *spatial fuzziness*. Fuzziness captures the property of many spatial objects in reality which do not have sharp boundaries or whose boundaries cannot be precisely determined. Examples are natural, social, or cultural phenomena like land features with continuously changing properties (such as population density, soil quality, vegetation, pollution, temperature, air pressure), oceans, deserts, English speaking areas, or mountains and valleys. The transition between a valley and a mountain usually cannot be exactly ascertained so that the two spatial objects "valley" and "mountain" cannot be precisely separated and defined in a crisp way. We will designate this kind of entities as *fuzzy* spatial objects.

Topological relationships characterize the relative locations of two spatial objects, for example, whether they overlap, meet, or are disjoint. In spatial databases and GIS they are important for formulating spatial selections and spatial joins. The definition of topological predicates on fuzzy spatial objects in general and fuzzy regions in particular is currently an open problem. For two fuzzy regions A and B we would like to be able to pose and answer queries like

- Do regions A and B overlap *a little bit*?
- Determine all pairs of regions that *nearly completely* overlap.
- Does region A *somewhat* contain region B?
- Which regions lie *quite* inside B?

In this chapter we present a formal object model for *fuzzy points*, *fuzzy lines*, and *fuzzy regions* in the two-dimensional Euclidean space, an effort which is to lead to a *fuzzy spatial algebra*. We propose *fuzzy set theory* and *fuzzy topology* as appropriate conceptual tools for modeling indeterminate spatial data. Fuzzy set theory is an extension and generalization of classical set theory; the approach of fuzzy sets replaces the crisp boundary of a classical set by a gradual transition zone and permits partial and multiple set membership. For fuzzy regions, different views give a better understanding of their nature and also demonstrate how these objects can be represented as (collections of) crisp regions. Consequently, the current exact object models for crisp spatial objects can be considered as simplified special cases of a richer class of models for general spatial objects.

Based on the data type definitions, we consider topological predicates on fuzzy spatial objects and show their integration into a *fuzzy spatial query language*. *Fuzzy modifiers* may be used in the query language to emphasize nuances of topological predicates and thus to modify their semantics.

Section 2 explains different aspects of spatial vagueness and presents related work. Section 3 introduces some basic definitions of fuzzy set theory and fuzzy topology as far as they are needed in this paper. Sections 4, 5 and 6 formally define fuzzy points, fuzzy lines, and fuzzy regions, respectively. Since

the definition for fuzzy regions does not expose their geometric structure, Section 7 provides several structured views of fuzzy regions based on collections of crisp complex regions. Section 8 introduces topological predicates on complex crisp regions as they are needed in Section 9 for defining topological predicates on fuzzy complex regions. Section 10 demonstrates how querying with fuzzy topological predicates can be done. Section 11 draws some conclusions and gives a prospect of future research.

2 Related Work

In the past, there has been a strong tendency to force vague spatial reality into crisp objects. In practice, however, there is no apparent reason for the whole boundary of a region to be determined. There are a lot of geographical application examples illustrating that the boundaries of spatial objects can be partially or totally indeterminate or blurred. For instance, boundaries of geological, soil, and vegetation units [1, 3, 19, 20] are often sharp in some places and vague in others; many human concepts like "the Indian Ocean" are implicitly vague.

In the real world, there are essentially two categories of indeterminate boundaries: sharp boundaries whose position and shape are unknown or cannot be measured precisely, and boundaries which are not well-defined or which are useless (e.g., between a mountain and a valley) and where essentially the topological relationship between spatial objects is of interest. According to these two categories, mainly two kinds of spatial vagueness can be identified: uncertainty and fuzziness. *Spatial uncertainty* is traditionally equated with randomness and chance occurrence and relates either to a lack of knowledge about the position and shape of an object with an existing, real boundary (*positional* uncertainty) or to the inability of measuring such an object precisely (*measurement* uncertainty). *Spatial fuzziness* is an intrinsic feature of an object itself and describes the vagueness of an object which certainly has an extent but which inherently cannot or does not have a precisely definable boundary.

The subject of modeling spatial vagueness has so far been predominantly treated by geographers but rather neglected by computer scientists. At least three alternatives are proposed as general design methods: (1) *exact models* [7, 8, 15, 22] which transfer type systems and concepts for spatial objects with sharp boundaries to objects with unclear boundaries and which model both uncertainty and fuzziness but in a restricted way, (2) *probabilistic models* [2, 3, 16, 24] which are based on probability theory and predominantly model positional and measurement uncertainty, and (3) *fuzzy models* [1, 3, 10, 11, 19, 20, 26, 27, 28, 31] which are all based on fuzzy set theory and predominantly model fuzziness.

The exact object model approach profits from existing definitions, techniques, data structures, algorithms, etc. which need not be redeveloped but

only modified and extended. Except for [15], the approaches are based on some kind of *zone* concept. Vague boundaries of a region are modeled as zones expressing the minimal and maximal possible extent of a region. *Vague regions* [15] are a generalization of these models. A vague region is defined as a pair of disjoint, crisp regions. The first region called the *kernel* describes the area which definitely and always belongs to the vague region. The second region called the *boundary* describes the *area* for which we cannot say with any certainty whether it or parts of it belong to the vague region or not. *Maybe* it is the case, *maybe* it is not. Or we could say that this is *unknown*. Vague regions are based on a *three-valued logic*, and boundaries need not necessarily be one-dimensional structures but can be regions.

Probability theory is able to represent uncertainty and defines the probability of an entity in a set by a statistically defined probability function. It deals with the *expectation* of a future event, based on something known now. Examples are the uncertainty about the spatial extent of regions defined by some property such as temperature, or the water level of a lake.

Fuzzy set theory deals only with fuzziness, that is, it describes the *admission of the possibility* (given by a so-called *membership function*) that an individual is a member of a set or that a given statement is true. Hence, the vagueness represented by fuzziness is not the uncertainty of expectation. It is the vagueness resulting from the imprecision of meaning of a concept. Examples of fuzzy spatial objects include mountains, valleys, biotopes, oceans, and many other geographic features which cannot be rigorously bounded by a sharp line.

The potential importance of fuzzy sets for geographical applications is demonstrated in [3, 20, 26] where also examples of application-specific membership functions are given. The benefits of fuzzy set theory for approximate spatial reasoning and fuzzy query languages is shown in [10, 11, 19, 27]. [28] models fuzzy objects by means of the relational data model.

Our definitions of fuzzy topological relationships are based on the so-called 9-*intersection model* [12] from which a complete collection of mutually exclusive topological relationships can be derived for each combination of spatial types. The model is based on the nine possible intersections of *boundary* (∂A), *interior* (A°), and *exterior* (A^-) of a spatial object A with the corresponding components of another object. Each intersection is tested for the topologically invariant criteria of emptiness and non-emptiness. $2^9 = 512$ different configurations are possible from which only a limited subset makes sense depending on the combination of spatial objects just considered.

A restriction of the 9-intersection model with respect to regions is that regions must be homeomorphic to the closed disc, that is, they must be connected and are not allowed to have holes. These regions are usually called *simple regions*. For two simple regions, eight meaningful configurations have been identified which lead to the eight predicates of the set $T_{sr} = \{$*disjoint*, *meet*, *overlap*, *equal*, *inside*, *contains*, *covers*, and *coveredBy*$\}$. Each predicate is uniquely determined so that all predicates are mutually exclusive and

complete with regard to the topologically invariant criteria of emptiness and non-emptiness.

Topological predicates on crisp *complex regions* have so far not been defined. In [6] the so-called TRCR (Topological Relationships for Composite Regions) model only allows sets of disjoint simple regions without holes. In [13] only topological relationships of simple regions with holes are considered; multi-part regions are not permitted. Topological predicates on fuzzy spatial objects, let them be simple or complex, have so far not been defined.

3 Fuzzy Sets and Fuzzy Topology

Crisp regions have been formally defined on the basis of point sets and point set topology (e.g., [15, 17, 23]) which mainly rest on the set operations of union, intersection, and difference. In a straightforward way we will now describe extensions of these two concepts to fuzzy set theory and fuzzy topology.

Fuzzy set theory [30] is an extension and generalization of Boolean set theory. Let X be a classical (crisp) set of objects, called the *universe* (*of discourse*). Membership in a classical subset A of X can then be described by the *characteristic function* $\chi_A : X \to \{0, 1\}$ such that for all $x \in X$ holds:

$$\chi_A(x) = \begin{cases} 1 & \text{if and only if } x \in A \\ 0 & \text{if and only if } x \notin A \end{cases}$$

This function, which discriminates sharply between members and non-members of a set, can be generalized such that all elements of X are mapped to the real interval [0,1] indicating the *degree of membership* of these elements in the set in question. Hence, fuzzy set theory permits an element to have partial and multiple membership. Larger values designate higher grades of set membership. Let X again be the universe. Then

$$\mu_{\tilde{A}} : X \to [0, 1]$$

is called the *membership function* of $\tilde{A}$, and the set

$$\tilde{A} = \{(x, \mu_{\tilde{A}}(x)) \mid x \in X\}$$

is called a *fuzzy set* in X. All elements of X receive a valuation with respect to their membership in $\tilde{A}$. Those elements $x \in X$ that in the classical sense do not belong to $\tilde{A}$ get the membership value $\mu_{\tilde{A}}(x) = 0$; elements $x \in X$ that completely belong to $\tilde{A}$ get the membership value $\mu_{\tilde{A}}(x) = 1$.

There are many ways of extending the set inclusion as well as the basic crisp set operations to fuzzy sets. We will comply with the definitions in [30]. Let $\tilde{A}$ and $\tilde{B}$ be fuzzy sets in X. Then

(i) $\neg \tilde{A} = \{(x, \mu_{\neg \tilde{A}}(x)) \mid x \in X, \mu_{\neg \tilde{A}}(x) = 1 - \mu_{\tilde{A}}(x)\}$

(ii) $\tilde{A} \subseteq \tilde{B} \Leftrightarrow \forall\, x \in X : \mu_{\tilde{A}}(x) \leq \mu_{\tilde{B}}(x)$

(iii) $\tilde{A} \cap \tilde{B} = \{(x, \mu_{\tilde{A} \cap \tilde{B}}(x)) \mid x \in X \;\wedge\; \mu_{\tilde{A} \cap \tilde{B}}(x) = \min(\mu_{\tilde{A}}(x), \mu_{\tilde{B}}(x))\}$

(iv) $\tilde{A} \cup \tilde{B} = \{(x, \mu_{\tilde{A} \cup \tilde{B}}(x)) \mid x \in X \;\wedge\; \mu_{\tilde{A} \cup \tilde{B}}(x) = \max(\mu_{\tilde{A}}(x), \mu_{\tilde{B}}(x))\}$

(v) $\tilde{A} - \tilde{B} = \tilde{A} \cap \neg \tilde{B}$

A [*strict*] α-*cut* or [*strict*] α-*level set* of a fuzzy set $\tilde{A}$ for a specified value α is the crisp set

$$A_\alpha\,[A_\alpha^*] = \{x \in X \mid \mu_{\tilde{A}}(x) \geq [>]\, \alpha \;\wedge\; 0 \leq \alpha \leq [<]\, 1\}$$

The strict α-cut for $\alpha = 0$ is called *support* of $\tilde{A}$, i.e., $supp(\tilde{A}) = A_0^*$. For a fuzzy set $\tilde{A}$ and $\alpha, \beta \in [0,1]$ holds

(i) $X = A_0$

(ii) $\alpha < \beta \;\Rightarrow\; A_\alpha \supseteq A_\beta$

The set of all levels $\alpha \in [0,1]$ that represent distinct α-cuts of a given fuzzy set $\tilde{A}$ is called the *level set* $\Lambda_{\tilde{A}}$ of $\tilde{A}$:

$$\Lambda_{\tilde{A}} = \{\alpha \in [0,1] \mid \exists\, x \in X : \mu_{\tilde{A}}(x) = \alpha\}$$

Fuzzy (point set) topology [5] is a straightforward extension and generalization of ordinary point set topology and allows one to distinguish specific topological structures of a fuzzy set like its closure or interior.

A *fuzzy topology* on a universe X is a family $\tilde{T}$ of fuzzy sets in X satisfying the following conditions:

(i) $X \in \tilde{T}, \emptyset \in \tilde{T}$

(ii) $\tilde{A} \in \tilde{T}, \tilde{B} \in \tilde{T} \;\Rightarrow\; \tilde{A} \cap \tilde{B} \in \tilde{T}$

(iii) $\tilde{S} \subseteq \tilde{T} \;\Rightarrow\; \bigcup_{\tilde{A} \in \tilde{S}} \tilde{A} \in \tilde{T}$

The pair $(X, \tilde{T})$ is said to be a *fuzzy topological space*. The elements of $\tilde{T}$ are called *open fuzzy sets*. Note that X and $\emptyset$ are crisp sets and simultaneously special fuzzy sets. X corresponds to the fuzzy set $\tilde{X} = \{(x, \mu_X(x)) \mid x \in X \;\wedge\; \mu_X(x) = 1\}$. The empty set $\emptyset$ corresponds to the empty fuzzy set $\tilde{\emptyset} = \{(x, \mu_X(x)) \mid x \in X \;\wedge\; \mu_X(x) = 0\}$. We will identify X and $\tilde{X}$ as well as $\emptyset$ and $\tilde{\emptyset}$ and use the crisp notations for these two sets.

The family $\tilde{T}'$ of all *closed fuzzy sets* in a fuzzy topological space $(X, \tilde{T})$ is given by

$$\tilde{T}' = \{\neg \tilde{A} \mid \tilde{A} \in \tilde{T}\}$$

The *closure* [*interior*] of a fuzzy set $\tilde{A}$ in a fuzzy topological space $(X, \tilde{T})$ is the smallest closed [largest open] fuzzy set containing $\tilde{A}$ [contained in $\tilde{A}$], i.e.,

$$cl_{\tilde{T}}(\tilde{A}) = \bigcap\, \{\tilde{S} \mid \tilde{S} \in \tilde{T}' \;\wedge\; \tilde{A} \subseteq \tilde{S}\}$$

$$[int_{\tilde{T}}(\tilde{A}) = \bigcup\, \{\tilde{S} \mid \tilde{S} \in \tilde{T} \;\wedge\; \tilde{S} \subseteq \tilde{A}\}]$$

4 Fuzzy Points

Due to our assumption of the point set paradigm, an understanding of the nature of a point, or more precisely a *fuzzy point*, is necessary. There are at least two meaningful definitions for a fuzzy point.

The first definition views a fuzzy point as a point in two-dimensional Euclidean space with a membership value greater than 0, since 0 documents the non-existence of a point. A *fuzzy point* $\tilde{p}$ at (a, b) in $\mathbb{R}^2$, written $\tilde{p}(a, b)$, is a fuzzy singleton in $\mathbb{R}^2$ defined by

$$\mu_{\tilde{p}(a,b)}(x, y) = \begin{cases} m & \text{if } (x, y) = (a, b) \\ 0 & \text{otherwise} \end{cases}$$

with $0 < m \leq 1$. Point $\tilde{p}$ is said to have support (a, b) and value m. Let P_f be the set of all fuzzy points. P_f is, of course, a proper superset of P_c, the set of all crisp points in $\mathbb{R}^2$. For $\tilde{p} = p = (a, b) \in P_c$, we obtain $\mu_{\tilde{p}(a,b)}(x, y) = \chi_p(x, y) = 1$, if $(x, y) = (a, b)$, and 0 otherwise.

The second definition uses a membership function that returns the *degree of proximity* of a point to a reference point $\tilde{p}$. That is, we consider that the point (x, y) is "approximately (a, b)" or "about (a, b)" to the degree $\mu_{\tilde{p}(a,b)}(x, y)$. A fuzzy point $\tilde{p}(a, b)$ is then generally defined by

(i) $\mu_{\tilde{p}(a,b)}$ is upper semicontinuous[1]
(ii) $\mu_{\tilde{p}(a,b)}(x, y) = 1$ if and only if $(x, y) = (a, b)$
(iii) $\forall\, 0 \leq \alpha \leq 1 : \tilde{p}_\alpha$ is a convex[2] subset of $\mathbb{R}^2$

The concrete "distance-based" membership function

$$\mu_{\tilde{p}(a,b)}(x, y) = c^{-\lambda\,((x-a)^2 + (y-b)^2)}$$

with $c \in \mathbb{R}^+, c > 1$, and $\lambda > 0$ illustrates this definition. The degree of proximity decreases as (x, y) moves further away from (a, b). It reaches 1 if $(x, y) = (a, b)$.

Unfortunately, this membership function with unbounded support is difficult to represent. Alternatively, we can employ the following, restricted but more practical function which defines a circle around (a, b) with radius $r \in \mathbb{R}^+$:

$$\mu_{\tilde{p}(a,b)}(x, y) = \begin{cases} 1 - \frac{\sqrt{(x-a)^2 + (y-b)^2}}{r} & \text{if } (x - a)^2 + (y - b)^2 \leq r^2 \\ 0 & \text{otherwise} \end{cases}$$

Next, we define three *geometric primitives* on fuzzy points which are valid for both definitions of fuzzy points. Let $\tilde{p}(a, b), \tilde{q}(c, d) \in P_f$ with $a, b, c, d \in \mathbb{R}$. Then

---

[1] A function $f : X \to \mathbb{R}$ is *upper semicontinuous* $:\Leftrightarrow \forall\, r \in \mathbb{R} : \{x \mid f(x) < r\}$ is open. (The notation ":$\Leftrightarrow$" means "is defined as being equivalent".)

[2] A set $X \subseteq \mathbb{R}^2$ is called *convex* $:\Leftrightarrow \forall\, p, q \in \mathbb{R}^2 \; \forall\, \lambda \in \mathbb{R}^+$ with $0 < \lambda < 1 : r = \lambda\, p + (1 - \lambda)\, q \in X$ (p, q, and r are here regarded as vectors)

(i) $\tilde{p}(a,b) = \tilde{q}(c,d)$ $:\Leftrightarrow$ $a = c$ $\wedge$ $b = d$ $\wedge$ $\mu_{\tilde{p}(a,b)} = \mu_{\tilde{q}(c,d)}$

(ii) $\tilde{p}(a,b) \neq \tilde{q}(c,d)$ $:\Leftrightarrow$ $\neg(\tilde{p}(a,b) = \tilde{q}(c,d))$

(iii) $\tilde{p}(a,b)$ and $\tilde{q}(c,d)$ are *disjoint* $:\Leftrightarrow$ $supp(\tilde{p}(a,b)) \cap supp(\tilde{q}(c,d)) = \emptyset$

In contrast to crisp points, for fuzzy points we also have a predicate for disjointedness. We are now able to define an object of the fuzzy spatial data type *fpoint* as a set of disjoint fuzzy points:

$$fpoint = \{Q \subseteq P_f \,|\, \forall\, \tilde{p}, \tilde{q} \in Q : \tilde{p}(a,b) \text{ and } \tilde{q}(c,d) \text{ are } disjoint \,\wedge\, Q \text{ is finite}\}$$

5 Fuzzy Lines

In this section we define a data type for *fuzzy lines*. For that, we first introduce a *simple* fuzzy line as a continuous curve with smooth transitions of membership grades between neighboring points of the line. We assume a total order on $\mathbb{R}^2$ which is given by the lexicographic order "$<$" on the coordinates (first x, then y). The membership function of a *simple fuzzy line* $\tilde{l}$ is then defined by

$$\mu_{\tilde{l}} : f_{\tilde{l}} \to [0,1] \text{ with } f_{\tilde{l}} : [0,1] \to \mathbb{R}^2 \text{ such that}$$

(i) $\mu_{\tilde{l}}$ is continuous

(ii) $f_{\tilde{l}}$ is continuous

(iii) $\forall\, a,b \in]0,1[: a \neq b \Rightarrow f_{\tilde{l}}(a) \neq f_{\tilde{l}}(b)$

(iv) $\forall\, a \in \{0,1\} \,\forall\, b \in]0,1[: f_{\tilde{l}}(a) \neq f_{\tilde{l}}(b)$

(v) $f_{\tilde{l}}(0) < f_{\tilde{l}}(1) \,\vee\, (f_{\tilde{l}}(0) = f_{\tilde{l}}(1) \,\wedge\, \forall a \in]0,1[) : f_{\tilde{l}}(0) < f_{\tilde{l}}(a))$

Function $f_{\tilde{l}}$ on its own models a continuous, simple *crisp* line (a *simple curve*). The points $f_{\tilde{l}}(0)$ and $f_{\tilde{l}}(1)$ are called the *end points* of f. The definition allows loops $(f_{\tilde{l}}(0) = f_{\tilde{l}}(1))$ but prohibits equality of interior points and thus self-intersections (condition (iii)) and equality of an interior with an end point (condition (iv)). The last condition ensures uniqueness of representation, i.e., in a *closed simple line* $f_{\tilde{l}}(0)$ must be the leftmost point.

Let S be the set of fuzzy simple lines, and let $T \subseteq S$. An *S-complex C over T* is a finite set $C = \{\tilde{l}_1, \dots, \tilde{l}_n\} \subseteq T$ such that[3]

(i) $\forall\, 1 \leq i < j \leq n : f_{\tilde{l}_i}(]0,1[) \cap f_{\tilde{l}_j}(]0,1[) = \emptyset$

(ii) $\forall\, 1 \leq i < j \leq n : \{f_{\tilde{l}_i}(0), f_{\tilde{l}_i}(1)\} \cap f_{\tilde{l}_j}(]0,1[) = \emptyset$

(iii) $\forall\, 1 \leq i \leq n \,\exists\, 1 \leq j \leq n, j \neq i : \{f_{\tilde{l}_i}(0), f_{\tilde{l}_i}(1)\} \cap \{f_{\tilde{l}_j}(0), f_{\tilde{l}_j}(1)\} \neq \emptyset$

(iv) For all $1 \leq i,j \leq n$ and for all $a,k \in \{0,1\}$ let $V_{\tilde{l}_i}^a = \{(j,k) \mid f_{\tilde{l}_i}(a) = f_{\tilde{l}_j}(k)\}$. Then we require: $\forall\, 1 \leq i \leq n \,\forall\, a \in \{0,1\} : (|V_{\tilde{l}_i}^a| = 1) \vee (|V_{\tilde{l}_i}^a| > 2)$

(v) $\forall\, 1 \leq i \leq n \,\forall\, a \in \{0,1\} \,\forall\, (j,k) \in V_{\tilde{l}_i}^a : \mu_{\tilde{l}_i}(f_{\tilde{l}_i}(a)) = \mu_{\tilde{l}_j}(f_{\tilde{l}_j}(k))$

---

[3] The application of a function f to a set X of values is defined as $f(X) = \{f(x) \mid x \in X\}$.

Condition (i) requires that the elements of an S-complex do not intersect or overlap within their interior. Moreover, they may not be touched within their interior by an endpoint of another element (condition (ii)). Condition (iii) ensures the property of connectivity of an S-complex; isolated fuzzy simple lines are disallowed. Condition (iv) expresses that each endpoint of an element of C must belong to exactly one or more than two incident elements of C (note that always $(i,a) \in V_{\tilde{l}_i}^a$). This condition supports the requirement of maximal elements and hence achieves minimality of representation. Condition (v) requires that the membership values of more than two elements of C with a common end point must have the same membership value; otherwise we get a contradiction saying that a point of an S-complex has more than one different membership value.

All conditions together define an S-complex over T as a *connected planar fuzzy graph* with a unique representation. The corresponding point set of C is $points(C) = \bigcup_{\tilde{l} \in C} f_{\tilde{l}}([0,1])$. The set of all S-complexes over T is denoted by $SC(T)$. The disjointedness of any two S-complexes $C_1, C_2 \in SC(T)$ is defined as follows:

$$C_1 \text{ and } C_2 \text{ are } disjoint :\Leftrightarrow points(C_1) \cap points(C_2) = \emptyset$$

A fuzzy spatial data type for fuzzy lines called *fline* can now be defined in two equivalent ways. The "structured view" is based on S-complexes:

$$fline = \{D \subseteq SC(S) \mid \forall\, C_1, C_2 \in D : C_1 \text{ and } C_2 \text{ are } disjoint \,\wedge\, D \text{ is finite}\}$$

Let $1_r = \mathbb{R}^2 \times \{1\}$. The "flat view" emphasizing the point set paradigm is:

$$fline = \{\tilde{Q} \subseteq 1_r \mid \exists\, D \subseteq SC(S) : \bigcup_{C \in D} points(C) = supp(\tilde{Q})\}$$

6 Fuzzy Regions

The aim of this section is to develop and formalize the concept of a *fuzzy region*. Section 6.1 informally discusses some intrinsic features of fuzzy regions and compares them to classical crisp regions. Section 6.2 provides their formal definition. Finally, Section 6.3 gives examples of possible membership functions for them.

6.1 What are Fuzzy Regions?

The question what a *crisp region* is has been treated in many publications. A very general definition defines a crisp region as a set of disjoint, connected areal components, called *faces*, possibly with disjoint *holes* [15, 18, 23] in the Euclidean space $\mathbb{R}^2$. This model has the nice property that it is closed under (appropriately defined) geometric union, intersection, and difference

operations. It allows crisp regions to contain holes and islands within holes to any finite level.

By analogy with the generalization of crisp sets to fuzzy sets, we strive for a generalization of crisp regions to fuzzy regions on the basis of the point set paradigm and fuzzy concepts. At the same time we would like to transfer the structural definition of crisp regions (i.e., the component view) to fuzzy regions. Thus, the structure of a fuzzy region is supposed to be the same as for a crisp region but with the exception and generalization which amounts to a relaxation and hence greater flexibility of the strict belonging or non-belonging principle of a point in space to a specific region and which enables a partial membership of a point in a region. This is just what the term "fuzzy" means here.

There are at least three possible, related interpretations for a point in a fuzzy region. First, this situation may be interpreted as the *degree of belonging* to which that point is *inside* or *part of* some areal feature. Consider the transition between a mountain and a valley and the problem to decide which points have to be assigned to the valley and which points to the mountain. Obviously, there is no strict boundary between them, and it seems to be more appropriate to model the transition by partial and multiple membership. Second, this situation may indicate the *degree of compatibility* of the individual point with the attribute or concept represented by the fuzzy region. An example are "warm areas" where we must decide for each point whether and to which grade it corresponds to the concept "warm". Third, this situation may be viewed as the *degree of concentration* of some attribute associated with the fuzzy region at the particular point. An example is air pollution where we can assume the highest concentration at power stations, for instance, and lower concentrations with increasing distance from them. All these related interpretations give evidence of *fuzziness*.

When dealing with crisp regions, the user usually does not employ point sets as a method to conceptualize space. The user rather thinks in terms of sharply determined *boundaries* enclosing and grouping areas with *equal* properties or attributes and separating different regions with different properties from each other; he or she has purely *qualitative* concepts in mind. This view changes when fuzzy regions come into play. Besides the qualitative aspect, in particular the *quantitative* aspect becomes important, and boundaries in most cases disappear (between a valley and a mountain there is no strict boundary!). The distribution of attribute values within a region and transitions between different regions may be *smooth* or *continuous*. This feature just characterizes fuzzy regions. There are a lot of spatial phenomena showing a smooth behavior. Application examples are air pollution (Figure 1), temperature zones, magnetic fields, storm intensity, and sun insolation.
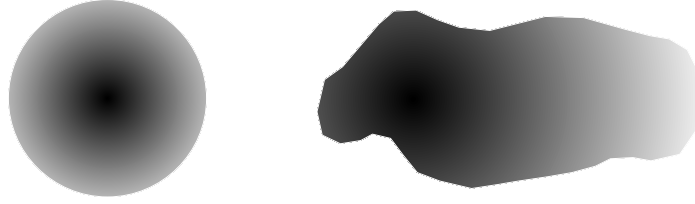
Fig. 1. This figure demonstrates a possible visualization of a fuzzy region which could model the expansion of air pollution caused by a power station. The left image shows a radial expansion where the degree of pollution concentrates in the center (darker locations) and decreases with increasing distance from the power station (brighter locations). The right image has the same theme but this time we imagine that the power station is surrounded by high mountains to the north, the south, and the west. Hence, the pollution cannot escape in these directions and finds its way out of the valley in eastern direction. In both cases we can recognize the smooth transitions to the exterior.

6.2 Formal Definition of Fuzzy Regions

Since our objective is to model two-dimensional fuzzy areal objects for spatial applications, we consider a fuzzy topology $\tilde{T}$ on the Euclidean space (plane) $\mathbb{R}^2$. In this spatial context we denote the elements of $\tilde{T}$ as *fuzzy point sets*. The membership function for a fuzzy point set $\tilde{A}$ in the plane is then described by $\mu_{\tilde{A}} : \mathbb{R}^2 \to [0, 1]$.

From an application point of view, there are two observations that prevent a definition of a fuzzy region simply as a fuzzy point set. We will discuss them now in more detail and at the same time elaborate properties of fuzzy regions.

Avoiding Geometric Anomalies: Regularization. The first observation refers to a necessary *regularization* of fuzzy point sets. The first reason for this measure is that fuzzy (as well as crisp) regions that actually appear in spatial applications in most cases cannot be just modeled as arbitrary point sets but have to be represented as point sets that do not have "geometric anomalies" and that are in a certain sense *regular*. Geometric anomalies relate to isolated or dangling line or point features and missing lines and points in the form of cuts and punctures. Spatial phenomena with such degeneracies never appear as entities in reality. The second reason is that, from a data type point of view, we are interested in fuzzy spatial data types that satisfy closure properties for (appropriately defined) geometric union, intersection, and difference.

We are, of course, confronted with the same problem in the crisp case where the problem can be avoided by the concept of *regularity* [15, 23, 25]. It turns out to be useful to appropriately transfer this concept to the fuzzy case. Let $\tilde{A}$ be a fuzzy set of a fuzzy topological space $(\mathbb{R}^2, \tilde{T})$. Then

$$\tilde{A} \text{ is called a } regular\ open\ fuzzy\ set \text{ if } \tilde{A} = int_{\tilde{T}}(cl_{\tilde{T}}(\tilde{A}))$$

Whereas crisp regions are usually modeled as *regular closed crisp sets*, we will use *regular open fuzzy sets* due to their vagueness and their usual lack of boundaries. Regular open fuzzy sets avoid the aforementioned geometric anomalies, too. Since application examples show that fuzzy regions can also be partially bounded, we admit *partial boundaries* with a crisp or fuzzy character. For that purpose we define the following fuzzy set (The notation ":=" means "is defined as".):

$$frontier_{\tilde{T}}(\tilde{A}) := \{((x,y), \mu_{\tilde{A}}(x,y)) \mid (x,y) \in supp(\tilde{A}) - supp(int_{\tilde{T}}(\tilde{A}))\}$$

A fuzzy set $\tilde{A}$ is now called a *spatially regular fuzzy set* iff

(i) $int_{\tilde{T}}(\tilde{A})$ is a regular open fuzzy set
(ii) $frontier_{\tilde{T}}(\tilde{A}) \subseteq frontier_{\tilde{T}}(cl_{\tilde{T}}(int_{\tilde{T}}(\tilde{A})))$
(iii) $frontier_{\tilde{T}}(\tilde{A})$ is a partition of n connected boundary parts (fuzzy sets)

We can conclude that $frontier_{\tilde{T}}(\tilde{A}) = \emptyset$ if $\tilde{A}$ is regular open. We will base our definition of fuzzy regions on spatially regular fuzzy sets and define a *regularization* function reg_f which associates the interior of a fuzzy set $\tilde{A}$ with its corresponding regular open fuzzy set and which restricts the partial boundary of $\tilde{A}$ (if it exists at all) to a part of the boundary of the corresponding regular closed fuzzy set of $\tilde{A}$:

$$reg_f(\tilde{A}) := int_{\tilde{T}}(cl_{\tilde{T}}(\tilde{A})) \cup (frontier_{\tilde{T}}(\tilde{A}) \cap frontier_{\tilde{T}}(cl_{\tilde{T}}(int_{\tilde{T}}(\tilde{A}))))$$

The different components of the regularization process work as follows: the *interior* operator $int_{\tilde{T}}$ eliminates dangling point and line features since their interior is empty. The *closure* operator $cl_{\tilde{T}}$ removes cuts and punctures by appropriately adding points. Furthermore, the *closure* operator introduces a fuzzy boundary (similar to a crisp boundary in the ordinary point-set topological sense) separating the points of a closed set from its exterior. The operator $frontier_{\tilde{T}}$ supports the restriction of the boundary.

The following statements about set operations on regular open fuzzy sets are given informally and without proof. The intersection of two regular open fuzzy sets is regular open. The union, difference, and complement of two regular open fuzzy sets are not necessarily regular open since they can produce anomalies. Correspondingly, this also holds for spatially regular fuzzy sets. Hence, we introduce *regularized set operations* on spatially regular fuzzy sets that preserve regularity. Let $\tilde{A}$, $\tilde{B}$ be spatially regular fuzzy sets of a fuzzy topological space $(\mathbb{R}^2, \tilde{T})$, and let $a \dot{-} b = a - b$ for $a \geq b$ and $a \dot{-} b = 0$ otherwise $(a, b \in \mathbb{R}_0^+)$. Then

(i) $\tilde{A} \cup_r \tilde{B} := reg_f(\tilde{A} \cup \tilde{B})$
(ii) $\tilde{A} \cap_r \tilde{B} := reg_f(\tilde{A} \cap \tilde{B})$
(iii) $\tilde{A} -_r \tilde{B} := reg_f(\{((x,y), \mu_{\tilde{A}-_r\tilde{B}}(x,y) \mid (x,y) \in \tilde{A} \ \wedge$
$\mu_{\tilde{A}-_r\tilde{B}}(x,y) = \mu_{\tilde{A}}(x,y) \dot{-} \mu_{\tilde{B}}(x,y)\})$

(iv) $\neg_r \tilde{A} := reg_f(\neg \tilde{A})$

Note that we have changed the meaning of difference (i.e., $\tilde{A} -_r \tilde{B} \neq \tilde{A} \cap_r \neg \tilde{B}$) since the right side of the inequality does not seem to make great sense in the spatial context. Regular open fuzzy sets, spatially regular fuzzy sets, and regularized set operations express a natural formalization of the desired dimension-preserving property of set operations. In the crisp case this is taken for granted but mostly never fulfilled by spatial type systems, geometric algorithms, spatial database systems, and GIS.

Whereas the subspace $RCCS$ of *r*egular *c*losed *c*risp *s*ets together with the crisp regular set operations "∪" and "∩" and the set-theoretic order relation "⊆" forms a Boolean lattice [15], this is not the case for $SRFS$ denoting the subspace of *s*patially *r*egular *f*uzzy *s*ets. Here we obtain the (unproven but obvious) statement that $SRFS$ together with the regularized set operations "$\cup_r$" and "$\cap_r$" and the fuzzy set-theoretic order relation "⊆" is a pseudo-complemented distributive lattice.

This implies that (i) $(SRFS, \subseteq)$ is a partially ordered set (reflexivity, anti-symmetry, transitivity), (ii) every pair $\tilde{A}$, $\tilde{B}$ of elements of $SRFS$ has a least upper bound $\tilde{A} \cup_r \tilde{B}$ and a greatest lower bound $\tilde{A} \cap_r \tilde{B}$, (iii) $(SRFS, \subseteq)$ has a maximal element $1_r := \{((x,y), \mu(x,y)) \mid (x,y) \in \mathbb{R}^2 \quad \wedge \quad \mu(x,y) = 1\}$ (identity of "$\cap_r$") and a minimal element $0_r := \{((x,y), \mu(x,y)) \mid (x,y) \in \mathbb{R}^2 \quad \wedge \quad \mu(x,y) = 0\}$ (identity of "$\cup_r$"), and (iv) algebraic laws like idempotence, commutativity, associativity, absorption, and distributivity hold for "$\cup_r$" and "$\cap_r$".

$(SRFS, \subseteq)$ is not a complementary lattice. Although the algebraic laws of involution and dualization hold, this is not true for the laws of complementarity. If we take the standard fuzzy set operations presented in Section 3 as a basis, the law of excluded middle $\tilde{A} \cup_r \neg \tilde{A} = 1_r$ and the law of contradiction $\tilde{A} \cap_r \neg \tilde{A} = 0_r$ do *not* hold in general. This fact explains the term "pseudo-complemented" from above and is no weakness of the model but only an indication of fuzziness.

Modeling Smooth Attribute Changes: Predominantly Continuous Membership Functions. The second observation is that according to the application cases shown in Section 6.1 the mapping $\mu_{\tilde{A}}$ itself may not be arbitrary but must take into account the intrinsic smoothness of fuzzy regions. This property can be modeled by the well known mathematical concept of *continuity* and results in special continuous membership functions for fuzzy regions. We say that a function f contains a *continuity gap* at a point x_0 of its domain if f is semicontinuous but not continuous at x_0. Function f is called *predominantly continuous* if f is continuous and has at most a *finite* number of continuity gaps.

Defining Fuzzy Regions. The type *fregion* for fuzzy regions can now be defined in the following way:

$$fregion = \{\tilde{R} \in SRFS \mid \mu_{\tilde{R}} \text{ is predominantly continuous}\}$$

6.3 Examples of Membership Functions for Fuzzy Regions

In this section we give some simple examples of membership functions which fulfil the properties required in Section 6.2. The determination of suitable membership functions is the difficulty in using the fuzzy set approach. Frequently, expert and empirical knowledge is necessary and used to design appropriate functions. We start with an example for a smooth fuzzy region. By taking a crisp region A with boundary B_A as a reference object, we can construct a fuzzy region on the basis of the following distance-based membership function:

$$\mu_{\tilde{A}} = \begin{cases} 1 & \text{if } (x,y) \in A \\ a^{-\lambda\, d((x,y),B_A)} & \text{if } (x,y) \notin A \end{cases}$$

where $a \in \mathbb{R}^+$ and $a > 1$, $\lambda \in \mathbb{R}^+$ is a constant, and $d((x,y), B_A)$ computes the distance between point (x,y) and boundary B_A in the following way:

$$d((x,y), B_A) = \min\{dist((x,y),(x',y')) \mid (x',y') \in B_A\}$$

where $dist(p,q)$ is the usual Euclidean distance between two points $p,q \in \mathbb{R}^2$. Unfortunately, this membership function leads to an unbounded spatially regular fuzzy set (regular open fuzzy set) which is impractical for implementation. We can also give a similar definition of a membership function with bounded support:

$$\mu_{\tilde{A}} = \begin{cases} 1 & \text{if } (x,y) \in A \\ a^{1-\frac{1}{\lambda}d((x,y),B_A)} & \text{if } (x,y) \notin A,\ d((x,y),B_A) \leq \lambda \\ 0 & \text{otherwise} \end{cases}$$

In the same way as the distance from a point outside of A to B_A increases to λ, the degree of membership of this point to $\tilde{A}$ decreases to zero.

[26] also presents membership functions for smooth fuzzy regions. The applications considered are air pollution defined as a fuzzy region with membership values based on the distance from a city center and a hill with elevation as the controlling value for the membership function. [20] models the transition of two smooth regions for soil units with symmetric membership functions.

A method to design a membership function for a finite-valued region with n possible membership values (truth values) is to code the n values by rational numbers in the unit interval $[0,1]$. For that purpose, the unit interval is evenly divided into $n-1$ subintervals and takes their endpoints as membership values. We obtain the set $T_n = \{\frac{i}{n-1} \mid n \in \mathbb{N}, 0 \leq i \leq n-1\}$ of truth values. Assuming that we intend to model air pollution caused by a power station located at point $p \in \mathbb{R}^2$, we can define the following (simplified) membership function for $n = 5$ degrees of truth representing, for instance, areas of extreme, high, average, low, and no pollution ($a,b,c,d \in \mathbb{R}^+$ denote distances):

$$\mu_{\tilde{A}}(x,y) = \begin{cases} 1 & \text{if } dist(p,(x,y)) \leq a \\ \frac{3}{4} & \text{if } a < dist(p,(x,y)) \leq b \\ \frac{1}{2} & \text{if } b < dist(p,(x,y)) \leq c \\ \frac{1}{4} & \text{if } c < dist(p,(x,y)) \leq d \\ 0 & \text{if } d < dist(p,(x,y)) \end{cases}$$

7 Structured Views of Fuzzy Regions

The formal definition of a fuzzy region given in Section 6.2 is conceptually somehow "structureless" in the sense that only "flat" point sets are considered and no structural information is revealed. In the following four subsections some "semantically richer" characterizations of fuzzy regions are presented which enable a better understanding of fuzzy regions. On the one hand they subdivide fuzzy regions into fuzzy components and on the other hand they describe them as collections of crisp regions. Moreover, they give hints for a possible implementation.

7.1 Fuzzy Regions as Multi-Component Objects

The first structured view considers a fuzzy region as a set of *fuzzy components*. For a definition we need a notion of *connectedness* for fuzzy regions. A *separation* of a fuzzy region $\tilde{Y}$ is a pair $\tilde{A}, \tilde{B}$ of fuzzy subregions satisfying the following four conditions:

(i) $\tilde{A} \neq \emptyset, \tilde{B} \neq \emptyset$
(ii) $\tilde{Y} = \tilde{A} \cup_r \tilde{B}$
(iii) $\tilde{A} \cap int_{\tilde{T}}(\tilde{B}) = \emptyset \ \wedge \ int_{\tilde{T}}(\tilde{A}) \cap \tilde{B} = \emptyset$
(iv) $|\tilde{A} \cap_r \tilde{B}|$ is finite

If a separation of $\tilde{Y}$ into $\tilde{A}$ and $\tilde{B}$ exists, then $\tilde{Y}$ is said to be *separated*, and we call $\tilde{A}$ and $\tilde{B}$ to be *disjoint*. Otherwise $\tilde{Y}$ is said to be *connected*. Note that condition (iii) of the definition uses the usual fuzzy intersection operation and not the one on spatially regular fuzzy sets since the latter requires two fuzzy regions as operands. The property of disjointedness (condition (iv)) requires that the two fuzzy subregions $\tilde{A}$ and $\tilde{B}$ may at most share a finite number of boundary points; this makes sense since otherwise they could be simply merged into one fuzzy subregion. We now continue this separation process and decompose a fuzzy region $\tilde{Y}$ into its maximal set of pairwise disjoint fuzzy components $\tilde{Y} = \{\tilde{A}_1, \ldots, \tilde{A}_n\}$ (in the spatial context this decomposition is always finite) so that we obtain with $I = \{1, \ldots, n\}$:

(i) $\forall i \in I : \tilde{A}_i \neq \emptyset$
(ii) $\tilde{Y} = \bigcup_{r \, i \in I} \tilde{A}_i$
(iii) $\forall i, j \in I, i \neq j : \tilde{A}_i \cap int_{\tilde{T}}(\tilde{A}_j) = \emptyset \ \wedge \ int_{\tilde{T}}(\tilde{A}_i) \cap \tilde{A}_j = \emptyset$

(iv) $\forall \, i,j \in I, i \neq j : |\tilde{A}_i \cap_r \tilde{A}_j|$ is finite
(v) $\forall \, i \in I : (\tilde{A}_i$ is connected $\wedge \;\; \not\exists \, \tilde{B} \supset \tilde{A}_i : \tilde{B}$ is connected)

We call each fuzzy component $\tilde{A}_i$ a *fuzzy face*. Hence, we obtain:

A *fuzzy region* is a set of pairwise disjoint *fuzzy faces*.

A question arises whether also *fuzzy holes* can be identified from the point set view of a fuzzy region. This question has to be negated. Let us briefly consider the crisp case. If A is a crisp region, its faces can have holes which belong to the complement (exterior) of A, i.e., to $\mathbb{R}^2 - A$, and are "enclosed" by A. Unfortunately, ordinary point set topology offers no method to extract holes from a (regular closed) point set as separate components; they are simply part of the complement. Note that this does not mean that regions with holes cannot be modeled. Some research work in [13, 23, 29], for example, shows that this is possible by selecting a constructive approach. Roughly speaking, the idea is to assume that the holes of A are already given as regions and to subtract these holes from a "generalized region A*" being isomorphic to a closed disc and being the union of A and the holes. But since this a pure set operation, afterwards A "forgets" how it was produced and cannot reconstruct its past. Similarly to the crisp case, holes cannot be identified from a (spatially regular) fuzzy point set, since fuzzy topology also offers no concept of holes.

Moreover, we are here faced with the problem of the nature of a fuzzy hole. By analogy with the crisp case, we could say that the fuzzy holes of a fuzzy region $\tilde{A}$ exclusively contain all points that are enclosed by any fuzzy face of $\tilde{A}$ and that have membership grade 0 in $\tilde{A}$. But then, a fuzzy hole is crisp and a subset of the set

$$H = \{((x,y),1) \mid (x,y) \in supp(\neg \tilde{A})\}$$

This model of a fuzzy hole is unsatisfactory in the sense that it only deals with those points enclosed by $\tilde{A}$ that definitely do not belong to $\tilde{A}$. It does not take into account the complement of those points of $\tilde{A}$ belonging only partially to $\tilde{A}$, i.e., the model does not consider the set

$$\underline{\tilde{A}} = \{((x,y),m) \mid (x,y) \in supp(\tilde{A}) \wedge m = 1 - \mu_{\tilde{A}}(x,y)\}$$

called the *anti-fuzzy region* of $\tilde{A}$.

One could argue that the points of $\underline{\tilde{A}}$ also belong to the fuzzy holes. And indeed, we will take this view. The consequence is that for a fuzzy face there exists exactly one fuzzy hole.

7.2 Fuzzy Regions as Three-Part Crisp Regions

The second structured view leads to a simplification of an originally smooth fuzzy region to a core-boundary region and thus to a change from a quantitative to a qualitative perspective. It distinguishes between the kernel, the

boundary, and the exterior as the three parts of a fuzzy region. For a fuzzy region $\tilde{A}$, these parts are defined as crisp regions (regular closed sets)[4]:

$$kernel(\tilde{A}) = reg_c(\{(x, y) \in \mathbb{R}^2 \mid \mu_{\tilde{A}}(x, y) = 1\})$$
$$boundary(\tilde{A}) = reg_c(\{(x, y) \in \mathbb{R}^2 \mid 0 < \mu_{\tilde{A}}(x, y) < 1\})$$
$$exterior(\tilde{A}) = reg_c(\{(x, y) \in \mathbb{R}^2 \mid \mu_{\tilde{A}}(x, y) = 0\})$$

The *kernel* identifies the part that definitely belongs to $\tilde{A}$. The *exterior* determines the part that definitely does not belong to $\tilde{A}$. The indeterminate character of $\tilde{A}$ is summarized in the *boundary* of $\tilde{A}$ in a unified and simplified manner. Kernel and boundary can be adjacent with a common border, and kernel and/or boundary can be empty. This view corresponds exactly to the already described concept of *vague regions* with its three-valued logic [15].

All in all, this view presents only a very coarse and restricted description of fuzzy regions since it differentiates only between three parts. The original gradation in the membership values of the points of the boundary gets lost. The benefit of this view lies in the implementation since efficient representation methods and algorithms for crisp regions can be used.

7.3 Fuzzy Regions as Collections of Crisp α-Level Regions

The third structured view attempts to diminish the drawbacks of the three-part view of fuzzy regions and to avoid the great information loss in this representation. It describes a fuzzy region in terms of nested α-level sets. Let $\tilde{A}$ be a fuzzy region. Then we represent a region A_a for an $\alpha \in [0, 1]$ as

$$A_a = reg_c(\{(x, y) \in \mathbb{R}^2 \mid \mu_{\tilde{A}}(x, y) \geq \alpha\})$$

We call A_a an α-*level region*. Clearly, A_a is a crisp region whose boundary is defined by all points with membership value greater than or equal to α. Note that A_a can have holes. The kernel of $\tilde{A}$, as it has been defined in Section 7.2, is then equal to $A_{1.0}$. A property of the α-level regions of a fuzzy region is that they are nested, i.e., if we select membership values $1 = \alpha_1 > \alpha_2 > \cdots > \alpha_n > \alpha_{n+1} = 0$ for some $n \in \mathbb{N}$, then

$$A_{\alpha_1} \subseteq A_{\alpha_2} \subseteq \cdots \subseteq A_{\alpha_n} \subseteq A_{\alpha_{n+1}}$$

We here describe the finite, discrete case that enables us to model and implement finite-valued and interval-based regions. If $\Lambda_{\tilde{A}}$ is infinite, then there are obviously infinitely many α-level regions which can only be finitely represented within this view if we make a finite selection of values. In the discrete case, if $|\Lambda_{\tilde{A}}| = n + 1$ and we take all these occurring membership values of a

---

[4] Correspondingly, for a crisp set A, the regularization function reg_c is defined as $reg_c(A) = cl_T(int_T(A))$ where T is a topology for a universe X and cl_T and int_T are the *closure* and *interior* operators on a topological space (X, T).

fuzzy region, we can replace "$\subseteq$" by "$\subset$" in the inclusion relationships above. This follows from the fact that for any $p \in A_{\alpha_i} - A_{\alpha_{i-1}}$ with $i \in \{2, \ldots, n+1\}$, $\mu_{\tilde{A}}(p) = \alpha_i$. For the continuous case, we get $\mu_{\tilde{A}}(p) \in [\alpha_i, \alpha_{i-1})$ which leads to interval-based regions. As a result, we obtain:

> A *fuzzy region* is a (possibly infinite) set of α-level regions, i.e., $\tilde{A} = \{A_{\alpha_i} \mid 1 \leq i \leq |\Lambda_{\tilde{A}}|\}$ with $\alpha_i > \alpha_{i+1} \Rightarrow A_{\alpha_i} \subseteq A_{\alpha_{i+1}}$ for $1 \leq i \leq |\Lambda_{\tilde{A}}| - 1$.

From the implementation perspective, one of the advantages of using (a finite collection of) α-level sets to describe fuzzy regions is that existing geometric data structures and geometric algorithms known from Computational Geometry [21] can be applied.

7.4 Fuzzy Regions as α-Partitions

The fourth structured view is partially motivated by the previous one and describes a fuzzy region as a partition. A partition in the spatial context, called a *spatial partition* [14], is a subdivision of the plane into pairwise disjoint (crisp) regions (called *blocks*) where each block is associated with an attribute and where adjacent blocks are not allowed to be labeled with the same attribute. It differs from the set-theoretic notion of a partition in the sense that it, of course, relates to space and that it incorporates a treatment of common boundary points which at the same time may belong to two adjacent blocks.

From an application point of view, different blocks of a spatial partition are often marked differently, i.e., different *labels* of some set L are assigned to different blocks. Thus, in a certain way, L determines the type of a partition. This leads to spatial partitions of type L that are functions $\pi : \mathbb{R}^2 \rightarrow L$. In most cases, partitions are defined only partially, i.e., there are blocks (frequently called the *exterior* of a partition) which have no explicitly assigned labels. To complete π to a total function, we assume a label $\perp_L$ (called *undefined*) for each label type L and require that the exterior of a partition is labeled by $\perp_L$.

Like for crisp regions, we also desire regularity for the blocks of a spatial partition. We require the interiors of blocks to be regular open sets. Since points on the boundary cannot be uniquely assigned to either adjacent block, we cannot simply map them to single L-values. Instead, boundary points are mapped to the set of values given by the labels of all adjacent blocks. This leads to the definition of a *spatial mapping* of type L as a total mapping $\pi : \mathbb{R}^2 \rightarrow L \cup 2^L$. The *range* of a spatial mapping π yields the set of labels actually used in π and is denoted by $range(\pi)$. The blocks of a spatial mapping π are point sets that are mapped to the same labels. The block for a single label l (or a set S of labels) is given[5] by $f^{-1}(l)$ $(f^{-1}(S))$. The common label

---

[5] We use the following definition of function inverse: for $f : X \rightarrow Y$ and $\forall\, y \in Y$: $f^{-1}(y) := \{x \in X \mid f(x) = y\}$. Note that f^{-1} applied to a set yields a set of sets.

of a block b of π is denoted by $\pi[b]$, i.e., $\pi(b) = \{l\} \Rightarrow \pi[b] = l$. Obviously, the cardinality of block labels identifies different parts of a partition. A *region* of π is any block of π that is mapped to a single element of L, and a *border* of π is given by a block that is mapped to a set of L-values, or formally for a spatial mapping π of type L:

(i) $\rho(\pi) = \pi^{-1}(range(\pi) \cap L))$ (*regions*)
(ii) $\beta(\pi) = \pi^{-1}(range(\pi) \cap 2^L))$ (*borders*)

Now we can finally define a spatial partition by topologically constraining regions to regular open sets and by semantically constraining boundary labels to those of adjacent regions.

A *spatial partition* of type L is a spatial mapping π of type L with

(i) $\forall\, r \in \rho(\pi) : r$ is a regular open set (i.e., $r = int_T(cl_T(r))$)
(ii) $\forall\, b \in \beta(\pi) : \pi[b] = \{\pi[r] \mid r \in \rho(\pi) \ \wedge \ b \subseteq cl_T(r)\}$

The set of all spatial partitions of type L is denoted by $[L]$, i.e., $[L] \subseteq \mathbb{R}^2 \to L \cup 2^L$.

Using the representation based on α-level regions defined in the preceding subsection, we are now able to define a fuzzy region as a spatial partition. In our case $L = \Lambda_{\tilde{A}}$, i.e., the labels are formed by all possible membership values α. We have now to determine the different blocks for regions and borders. The regions of $\tilde{A}$ are given[6] by the set $\{int_T(A_{\alpha_i} -_c A_{\alpha_{i-1}}) \mid i \in \{2, \dots, n+1\}\}$, and the borders of $\tilde{A}$ are represented by the set $\{bound_T(A_{\alpha_i} -_c A_{\alpha_{i-1}}) \mid i \in \{2, \dots, n+1\}\}$. The object $A_{\alpha_i} -_c A_{\alpha_{i-1}}$ is a region possibly with holes. Each region is uniquely associated with an $\alpha \in \Lambda_{\tilde{A}}$, and each border has all α-labels of adjacent regions.

A *fuzzy region* $\tilde{A}$ is a spatial partition of type $\Lambda_{\tilde{A}}$ (i.e., $\tilde{A} \in [\Lambda_{\tilde{A}}]$), called an *$\alpha$-partition*.

If $\Lambda_{\tilde{A}}$ is infinite, we get an infinite spatial partition.

8 Topological Predicates on Complex Crisp Regions

In this section we extend the eight topological predicates on simple crisp regions to complex crisp regions. The extended predicates serve as a basis for topological predicates on complex fuzzy regions modeled as collections of crisp α-level regions (Section 9).

In the following we use as a syntactical simplification the notation $(P_1|P_2|\dots|P_n)(F, G)$ for the term $P_1(F, G) \ \vee \ P_2(F, G) \ \vee \ \dots \ \vee \ P_n(F, G)$ where $P_i : region \times region \to \mathbb{B}$ is a topological predicate for each $1 \le i \le n$.

---

[6] In the following, the operation "$-_c$" denotes the regular difference operation on regular closed sets. The operation $bound_T$ applied to a regular closed set yields its point-set topological boundary.

8.1 Topological Predicates on Simple Regions with Holes

As a first step to a general definition of topological predicates for complex crisp regions we consider such predicates for simple regions with holes and base their definition on the topological predicates for simple regions as they have been derived from the 9-intersection model (Section 2).

Let F and G be two simple regions with holes, that is, $F = F_0 - \bigcup_{i=1}^{n} F_i$ and $G = G_0 - \bigcup_{j=1}^{m} G_j$. We consider F and G to be disjoint if they have nothing in common, that is, either F_0 and G_0 are disjoint, and thus implicitly also their corresponding holes due to the definition of F and G, or F_0 (or G_0, respectively) and implicitly its holes are completely inside a hole G_j of G (F_i of F, respectively) (Figure 2a). Formally, we can then define the predicate $disjoint_{srh}$ as

$$disjoint_{srh}(F, G) \quad := \quad \begin{aligned} &disjoint(F_0, G_0) \; \vee \\ &(\exists\, 1 \leq i \leq n : inside(G_0, F_i)) \; \vee \\ &(\exists\, 1 \leq j \leq m : inside(F_0, G_j)) \end{aligned}$$

<center>(a)   (b)</center>
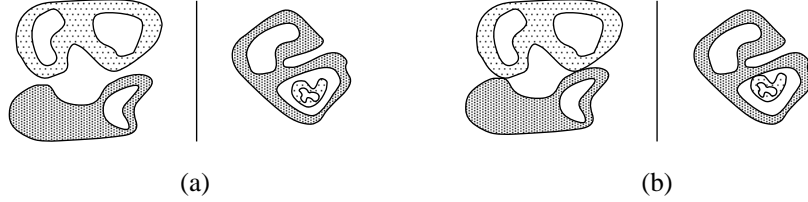
**Fig. 2.** Examples for the predicates $disjoint_{srh}(F, G)$ (a) and $meet_{srh}(F, G)$ (b).

The predicate $meet_{srh}$ is defined as follows (Figure 2b):

$$meet_{srh}(F, G) \quad := \quad \begin{aligned} &meet(F_0, G_0) \; \vee \\ &(\exists\, 1 \leq i \leq n : coveredBy(G_0, F_i)) \; \vee \\ &(\exists\, 1 \leq j \leq m : coveredBy(F_0, G_j)) \end{aligned}$$

We consider F to be inside G if F_0 is inside G_0 and if each hole G_j of G is either disjoint from F_0 or inside a hole F_i of F. (Figure 3a). The definition for the predicate $inside_{srh}$ is:

$$inside_{srh}(F, G) \quad := \quad \begin{aligned} &inside(F_0, G_0) \; \wedge \\ &(\forall\, 1 \leq j \leq m : disjoint(F_0, G_j) \; \vee \\ &\quad (inside(G_j, F_0) \; \wedge \; \exists\, 1 \leq i \leq n : inside(G_j, F_i))) \end{aligned}$$

We do not have to take into account the topological relationships between the F_i's and G_0 in our definition, because $inside(F_0, G_0) \Rightarrow inside(F_i, G_0)$ due to $F_i \subset F_0$ for $1 \leq i \leq n$. The predicate $contains_{srh}$ is symmetric to the predicate $inside_{srh}$, that is, $contains_{srh}(F, G) := inside_{srh}(G, F)$.

(a)   (b)
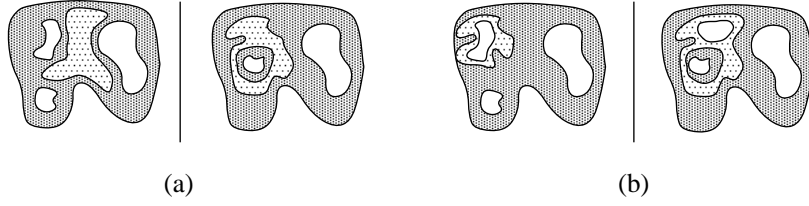
Fig. 3. Examples for the predicates $inside_{srh}(F, G)$ (a) and $coveredBy_{srh}(F, G)$ (b).

We consider F and G to be equal if F_0 and G_0 are equal, if F and G have the same number of holes, and if each hole F_i of F coincides with a hole G_j of G and vice-versa, that is,

$$
\begin{aligned}
equal_{srh}(F, G) \quad &:= \quad equal(F_0, G_0) \ \wedge \ n = m \ \wedge \\
&\quad \exists\, \pi : \{1, \ldots, n\} \to \{1, \ldots, n\}, \pi \text{ bijective,} \\
&\quad \forall\, 1 \le i \le n : equal(F_i, G_{\pi(i)})
\end{aligned}
$$

F is considered to be covered by G if F is a proper subset of G and if F's boundary touches G's boundary (Figure 3b).

$$
\begin{aligned}
coveredBy_{srh}(F, G) := \ &\neg((inside_{srh}|equal_{srh})(F, G)) \ \wedge \\
&(inside|coveredBy|equal)(F_0, G_0) \ \wedge \\
&(\forall\, 1 \le j \le m : ((disjoint|meet)(F_0, G_j) \ \vee \\
&(\exists\, 1 \le i \le n : (inside|coveredBy|equal)(G_j, F_i))))
\end{aligned}
$$

The predicate $covers_{srh}$ is symmetric to the predicate $coveredBy_{srh}$, that is, $covers_{srh}(F, G) := coveredBy_{srh}(G, F)$.

Finally, the predicate $overlap_{srh}$ (Figure 4) covers all remaining topological situations. This predicate can, of course, be defined directly in order to give an exact characterization of the remaining topological situations. But this makes the definition unnecessarily complicated and longish. We define instead:

$$
\begin{aligned}
overlap_{srh}(F, G) \quad &:= \quad \neg((disjoint_{srh}|meet_{srh}|coveredBy_{srh}|covers_{srh}| \\
&\qquad inside_{srh}|contains_{srh}|equal_{srh})(F, G))
\end{aligned}
$$

The set $T_{srh} = \{disjoint_{srh}, \ meet_{srh}, \ overlap_{srh}, \ coveredBy_{srh}, \ covers_{srh}, \ inside_{srh}, \ contains_{srh}, \ equal_{srh}\}$ provides a complete coverage of topological relationships for two simple regions with holes, and its elements are mutually
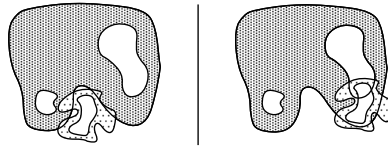
Fig. 4. Examples for the predicate $overlap_{srh}(F, G)$.

exclusive. Completeness of T_{srh} follows immediately from the complementary character of the definition of $overlap_{srh}$. Hence, at least one predicate must hold for any pair F, G of simple regions with holes. Mutual exclusion of each pair of different topological predicates P_1 and P_2 can be proved by showing that $\neg(P_1(F, G) \wedge P_2(F, G))$ holds for any pair F, G. Since we have $k = 8$ predicates, we have to check the diversity of $\frac{1}{2}(k^2 - k) = 28$ predicate pairs. We will not show the validity for the predicate pairs here in detail but only tell the strategy. First, we can use the mutual exclusion of the topological relationships for simple regions employed in the definition of some predicates. For instance, $disjoint_{srh}$ and $meet_{srh}$ exclude each other since $disjoint$ and $meet$ as well as $inside$ and $coveredBy$ are mutually exclusive in the 9-intersection model. Second, several predicates use the negation of other predicates on simple regions with holes for their definition. For instance, $overlap_{srh}$ excludes all other predicates. Similarly, $coveredBy_{srh}$ excludes both $inside_{srh}$ and $equal_{srh}$. Overall, at most one predicate is valid for any pair F, G.

The set T_{srh} of topological predicates on simple regions with holes is in two ways compatible with the set T of topological predicates on simple regions obtained by the 9-intersection model. First, if both F and G do not have holes, then T_{srh} and T coincide. Second, each of the eight topological predicates on simple regions with holes has the same boolean results for the nine intersections as the corresponding predicate on simple regions (see Section 2).

8.2 Topological Predicates on Complex Regions

With the aid of the topological predicates on simple regions with holes we are now able to define the corresponding predicates on complex regions. Let $F = \bigcup_{i=1}^{n} F_i$ and $G = \bigcup_{j=1}^{m} G_j$ be complex regions where the F_i and G_j are simple regions possibly with holes. We define the following predicates:

$$
\begin{aligned}
disjoint_{cr}(F, G) \;&:=\; \forall\, 1 \le i \le n \; \forall\, 1 \le j \le m : disjoint_{srh}(F_i, G_j) \\
meet_{cr}(F, G) \;&:=\; \neg disjoint_{cr}(F, G) \;\wedge\; (\forall\, 1 \le i \le n \; \forall\, 1 \le j \le m : \\
&\qquad (disjoint_{srh} | meet_{srh})(F_i, G_j)) \\
inside_{cr}(F, G) \;&:=\; \forall\, 1 \le i \le n \; \exists\, 1 \le j \le m : inside_{srh}(F_i, G_j) \\
contains_{cr}(F, G) \;&:=\; inside_{cr}(G, F) \\
equal_{cr}(F, G) \;&:=\; n = m \;\wedge\; (\exists\, \pi : \{1, \ldots, n\} \to \{1, \ldots, n\}, \pi \text{ bijective}, \\
&\qquad \forall\, 1 \le i \le n : equal_{srh}(F_i, G_{\pi(i)})) \\
coveredBy_{cr}(F, G) \;&:=\; \neg((inside_{cr} | equal_{cr})(F, G)) \;\wedge \\
&\qquad (\forall\, 1 \le i \le n \; \exists\, 1 \le j \le m : \\
&\qquad\qquad (inside_{srh} | coveredBy_{srh} | equal_{srh})(F_i, G_j)) \\
covers_{cr}(F, G) \;&:=\; coveredBy_{cr}(G, F) \\
overlap_{cr}(F, G) \;&:=\; \neg((disjoint_{cr} | meet_{cr} | coveredBy_{cr} | covers_{cr} | \\
&\qquad inside_{cr} | contains_{cr} | equal_{cr})(F, G))
\end{aligned}
$$

With similar arguments as in the last section we can recognize that two complex regions satisfy exactly one of these topological predicates. In other words, the topological predicates of the set $T_{cr} = \{disjoint_{cr}, meet_{cr}, inside_{cr}, contains_{cr}, equal_{cr}, coveredBy_{cr}, covers_{cr}, overlap_{cr}\}$ are mutually exclusive and complete. Note that the predicates $disjoint_{cr}$, $meet_{cr}$, $equal_{cr}$, and $overlap_{cr}$ are symmetric whereas the others are not.

9 Topological Predicates on Complex Fuzzy Regions

In this section we introduce a concept of topological predicates for fuzzy regions. In a similar way as we can generalize the characteristic function $\chi_A : X \to \{0,1\}$ to the membership function $\mu_{\tilde{A}} : X \to [0,1]$ (Section 3)[7], we can generalize a (binary) predicate $p_c : X \times Y \to \{0,1\}$ to a (binary) *fuzzy predicate* $p_f : \tilde{X} \times \tilde{Y} \to [0,1]$. Hence, the value of a fuzzy predicate can be interpreted as the degree to which the predicate holds for its operand objects. In our case of topological predicates, $X = Y = region$, $\{0,1\} = bool$, and $\tilde{X} = \tilde{Y} = fregion$ hold. For the set $[0,1]$ we have to introduce a new type *fbool* for *fuzzy booleansfuzzy boolean*.

For the definition of fuzzy topological predicates, we take the view of a fuzzy region as a set of α-level regions (Section 7.3). We know that an α-level region is a crisp complex region, and in the last section we have defined topological predicates on complex regions. This preparatory work now enables us to reduce topological predicates on fuzzy regions to topological predicates on collections of crisp regions[8].

The approach presented in this section is generic in the sense that any meaningful collection of topological predicates on complex crisp regions can be the basis for our definition of a collection of topological predicates on complex fuzzy regions. If the former collection additionally fulfils the properties of completeness and mutual exclusion (which is the case for T_{cr}), the latter collection automatically inherits these properties.

The open question now is how to compute the topological relationships of two collections of α-level regions, each collection describing a fuzzy region. We use the concept of basic probability assignment [9] for this purpose. A *basic probability assignment* $m(F_{\alpha_i})$ can be associated with each α-level region F_{α_i} and can be interpreted as the probability that F_{α_i} is the "true" representative of F. It is defined as

$$m(F_{\alpha_i}) = \alpha_i - \alpha_{i+1}$$

for $1 \leq i \leq n$ for some $n \in \mathbb{N}$ with $\alpha_1 = 1$ and $\alpha_{n+1} = 0$. That is, m is built from the differences of successive α_i's. It is easy to see that the telescoping sum $\sum_{i=1}^{n} m(F_{\alpha_i}) = \alpha_1 - \alpha_{n+1} = 1 - 0 = 1$.

---

[7] Note that χ_A is a unary crisp predicate and that $\mu_{\tilde{A}}$ is a unary fuzzy predicate.

[8] Another great benefit of this approach is its easy implementability through well known concepts for crisp spatial objects and for crisp topological predicates.

Let $\pi_f(F,G)$ be the value that represents a (binary) property π_f between two fuzzy regions F and G. Based on the work in [9] property π_f of F and G can be determined as the summation of weighted predicates by[9]

$$\pi_f(F,G) = \sum_{i=1}^{n}\sum_{j=1}^{n} m(F_{\alpha_i}) \cdot m(G_{\alpha_j}) \cdot \pi_{cr}(F_{\alpha_i}, G_{\alpha_j})$$

where $\pi_{cr}(F_{\alpha_i}, G_{\alpha_j})$ yields the value of the corresponding property π_{cr} for two crisp α-level regions F_{α_i} and G_{α_j}. This formula is equivalent to

$$\pi_f(F,G) = \sum_{i=1}^{n}\sum_{j=1}^{n} (\alpha_i - \alpha_{i+1}) \cdot (\alpha_j - \alpha_{j+1}) \cdot \pi_{cr}(F_{\alpha_i}, G_{\alpha_j})$$

If π_f is a topological predicate of $T_f = \{disjoint_f,\ meet_f,\ overlap_f,\ equal_f,\ inside_f,\ contains_f,\ covers_f,\ coveredBy_f\}$ between two fuzzy regions, we can compute the degree of the corresponding relationship with the aid of the pertaining crisp topological predicate $\pi_{cr} \in T_{cr}$. The value of $\pi_{cr}(F_{\alpha_i}, G_{\alpha_j})$ is either 1 (*true*) or 0 (*false*). Once this value has been determined for all combinations of α-level regions from F and G, the aggregated value of the topological predicate $\pi_f(F,G)$ can be computed as shown above. The more fine-grained the level set Λ for the fuzzy regions F and G is, the more precisely the fuzziness of topological predicates can be determined.

It remains to show that $0 \le \pi_f(F,G) \le 1$ holds, that is, π_f is really a fuzzy predicate. Since $\alpha_i - \alpha_{i+1} > 0$ for all $1 \le i \le n$ and since $\pi_{cr}(F_{\alpha_i}, G_{\alpha_j}) \ge 0$ for all $1 \le i,j \le n$, $\pi_f(F,G) \ge 0$ holds. We can show the other inequality by determining an upper bound for $\pi_f(F,G)$:

$$\begin{aligned}
\pi_f(F,G) &= \sum_{i=1}^{n}\sum_{j=1}^{n} (\alpha_i - \alpha_{i+1}) \cdot (\alpha_j - \alpha_{j+1}) \cdot \pi_{cr}(F_{\alpha_i}, G_{\alpha_j})\\
&\le \sum_{i=1}^{n}\sum_{j=1}^{n} (\alpha_i - \alpha_{i+1}) \cdot (\alpha_j - \alpha_{j+1}) \quad (\text{since } \pi_{cr}(F_{\alpha_i}, G_{\alpha_j}) \le 1)\\
&= (\alpha_1 - \alpha_2)(\alpha_1 - \alpha_2) + \cdots + (\alpha_1 - \alpha_2)(\alpha_n - \alpha_{n+1}) + \cdots +\\
&\quad (\alpha_n - \alpha_{n+1})(\alpha_1 - \alpha_2) + \cdots + (\alpha_n - \alpha_{n+1})(\alpha_n - \alpha_{n+1})\\
&= (\alpha_1 - \alpha_2)\big((\alpha_1 - \alpha_2) + \cdots + (\alpha_n - \alpha_{n+1})\big) + \cdots +\\
&\quad (\alpha_n - \alpha_{n+1})\big((\alpha_1 - \alpha_2) + \cdots + (\alpha_n - \alpha_{n+1})\big)\\
&= (\alpha_1 - \alpha_2) + \cdots + (\alpha_n - \alpha_{n+1}) \quad (\text{since } \sum_{i=1}^{n}(\alpha_i - \alpha_{i+1}) = 1)\\
&= 1
\end{aligned}$$

Hence, $\pi_f(F,G) \le 1$ holds.

An alternative definition of fuzzy topological predicates, which pursues a similar strategy like the one discussed so far, is based on the topological

---

[9] For reasons of simplicity, we assume that $\Lambda_{\tilde{F}} = \Lambda_{\tilde{G}} =: \Lambda$. Otherwise, it is not difficult to "synchronize" $\Lambda_{\tilde{F}}$ and $\Lambda_{\tilde{G}}$ by forming their union and by reordering and renumbering all levels.

predicates on simple regions possibly with holes, that is, on predicates $\pi_{srh} \in T_{srh}$. If F_{α_i} is an α-level region, let us denote its faces by $F_{\alpha_{i1}}, \ldots, F_{\alpha_{if_i}}$. Similarly, we denote the faces of an α-level region G_{α_j} by $G_{\alpha_{j1}}, \ldots, G_{\alpha_{jg_j}}$. We can then define a topological predicate π'_f as

$$\pi'_f(F,G) = \sum_{i=1}^{n} \sum_{k=1}^{f_i} \sum_{j=1}^{n} \sum_{l=1}^{g_j} \frac{(\alpha_i - \alpha_{i+1}) \cdot (\alpha_j - \alpha_{j+1}) \cdot \pi_{srh}(F_{\alpha_{ik}}, G_{\alpha_{jl}})}{f_i \cdot g_j}$$

It is obvious that $\pi'_f(F,G) \geq 0$ holds since all factors have a value greater than or equal to 0. We can also show that $\pi'_f(F,G) \leq 1$ by the following transformations:

$$
\begin{aligned}
\pi'_f(F,G) &\leq \sum_{i=1}^{n} \sum_{k=1}^{f_i} \sum_{j=1}^{n} \sum_{l=1}^{g_j} \frac{(\alpha_i - \alpha_{i+1}) \cdot (\alpha_j - \alpha_{j+1})}{f_i \cdot g_j} \\
&\quad (\text{since } \pi_{srh}(F_{\alpha_{ik}}, G_{\alpha_{jl}}) \leq 1) \\
&= \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{(\alpha_i - \alpha_{i+1}) \cdot (\alpha_j - \alpha_{j+1})}{f_i \cdot g_j} \cdot f_i \cdot g_j \\
&= \sum_{i=1}^{n} \sum_{j=1}^{n} (\alpha_i - \alpha_{i+1}) \cdot (\alpha_j - \alpha_{j+1}) \\
&= 1
\end{aligned}
$$

Hence, $\pi'_f(F,G) \leq 1$ holds. As a rule the predicates π_f and π'_f do not yield the same results. Assume that F_{α_i} and G_{α_j} fulfil a predicate $\pi_{cr} \in T_{cr}$. This fact contributes once to the summation process for π_f. But it does not take into account that possibly several faces $F_{\alpha_{ik}}$ (at least one) of F_{α_i} satisfy the corresponding predicate $\pi_{srh} \in T_{srh}$ with several faces $G_{\alpha_{jl}}$ (at least one) of G_{α_j}. This fact contributes several times (at most $f_i \cdot g_j$) to the summation process for π'_f. Hence, the evaluation process for π'_f is more fine-grained than for π_f.

Both generic predicate definitions reveal their quantitative character. If the predicate $\pi_{cr}(F_{\alpha_i}, G_{\alpha_j})$ and the predicate $\pi_{srh}(F_{\alpha_{ik}}, G_{\alpha_{jl}})$, respectively, is never fulfilled, the predicate $\pi_f(F,G)$ and $\pi'_f(F,G)$, respectively, yields *false*. The more α-level regions of F and G (simple regions with holes of F_{α_i} and G_{α_j}) fulfil the predicate $\pi_{cr}(F_{\alpha_i}, G_{\alpha_j})$ ($\pi_{srh}(F_{\alpha_{ik}}, G_{\alpha_{jl}})$), the more the validity of the predicate π_f (π'_f) increases. The optimum is reached if all topological predicates are satisfied.

10 Querying with Fuzzy Topological Predicates

In this section we demonstrate how fuzzy topological predicates can be integrated into an SQL-like spatial query language. The fact that the membership degree yielded by a fuzzy topological predicate is a computationally determined quantification between 0 and 1, that is, a fuzzy boolean, impedes

a direct integration. First, it is not very comfortable and user-friendly to use such a numeric value in a query. Second, spatial selections and spatial joins expect crisp predicates as filter conditions and are not able to cope with fuzzy predicates.

As a solution, we propose to embed adequate qualitative linguistic descriptions of nuances of topological relationships as appropriate interpretations of the membership values into a spatial query language. For instance, depending on the membership value yielded by the predicate $inside_f$, we could distinguish between *not* inside, *a little bit* inside, *somewhat* inside, *slightly* inside, *quite* inside, *mostly* inside, *nearly completely* inside, and *completely* inside. These *fuzzy linguistic terms* can then be incorporated into spatial queries together with the fuzzy predicates they modify. We call these terms *fuzzy modifiers* since their meaning is that of intensifying or relaxing the constraint expressed by the primary term to which they are applied. For instance, *somewhat inside* is a relaxation of the constraint *inside* since we can expect that it is better satisfied than *inside* even if some portions are outside. It is conceivable that a fuzzy modifier is either predefined and anchored in the query language, or user-defined.

We know that a fuzzy topological predicate π_f is defined as $\pi_f : fregion \times fregion \to [0,1]$. The idea is now to represent each fuzzy modifier $\gamma \in \Gamma = \{not,\ a\ little\ bit,\ somewhat,\ slightly,\ quite,\ mostly,\ nearly\ completely,\ completely\}$ by an appropriate fuzzy set with a membership function $\mu_\gamma : [0,1] \to [0,1]$. Let $F, G \in fregion$, and let $\gamma\,\pi_f$ be a quantified fuzzy predicate (like *somewhat inside* with $\gamma = somewhat$ and $\pi_f = inside_f$). Then we can define:

$$\gamma\,\pi_f(F,G) = true \quad :\Leftrightarrow \quad (\mu_\gamma \circ \pi_f)(F,G) = 1$$

That is, only for those values of $\pi_f(F,G)$ for which μ_γ yields 1, the predicate $\gamma\,\pi_f$ is true. A membership function that fulfils this quite strict condition is, for instance, the crisp partition of $[0,1]$ into $|\Gamma|$ disjoint or adjacent intervals completely covering $[0,1]$ and the assignment of each interval to a fuzzy modifier. If an interval $[a,b]$ is assigned to a fuzzy modifier γ, the intended meaning is that $\mu_\gamma(\pi_f(F,G)) = 1$, if $a \leq \pi_f(F,G) \leq b$, and 0 otherwise. For example, we could select the intervals $[0.0, 0.02]$ for *not*, $[0.02, 0.05]$ for *a little bit*, $[0.05, 0.2]$ for *somewhat*, $[0.2, 0.5]$ for *slightly*, $[0.5, 0.8]$ for *quite*, $[0.8, 0.95]$ for *mostly*, $[0.95, 0.98]$ for *nearly completely*, and $[0.98, 1.00]$ for *completely*.

Alternative membership functions are shown by the fuzzy sets in Figure 5. While we can always find a fitting fuzzy modifier for the partition due to the complete coverage of the interval $[0,1]$, this is not necessarily the case here. Each fuzzy modifier is associated with a *fuzzy number* having a trapezoidal-shaped membership function. The transition between two consecutive fuzzy modifiers is smooth and here modeled by linear functions. Within a fuzzy transition area μ_γ yields a value less than 1 which makes the predicate $\gamma\,\pi_f$ false. Examples in Figure 5 can be found at 0.2, 0.5, or 0.8. Each fuzzy number associated with a fuzzy modifier can be represented as a quadruple (a, b, c, d) where the membership function starts at $(a, 0)$, linearly
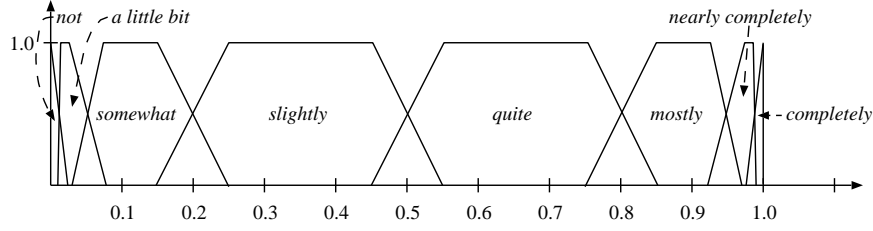
Fig. 5. Membership functions for fuzzy modifiers.

increases up to $(b, 1)$, remains constant up to $(c, 1)$, and linearly decreases up to $(d, 0)$. Figure 5 assigns $(0.0, 0.0, 0.0, 0.02)$ to *not*, $(0.01, 0.02, 0, 03, 0.08)$ to *a little bit*, $(0.03, 0.08, 0.15, 0.25)$ to *somewhat*, $(0.15, 0.25, 0.45, 0.55)$ to *slightly*, $(0.45, 0.55, 0.75, 0.85)$ to *quite*, $(0.75, 0.85, 0.92, 0.96)$ to *mostly*, $(0.92, 0.96, 0.97, 0.99)$ to *nearly completely*, and $(0.97, 1.0, 1.0, 1.0)$ to *completely*.

So far, the predicate $\gamma\,\pi_f$ is only true if μ_γ yields 1. We can relax this strict condition by defining:

$$\gamma\,\pi_f(F, G) = true \quad :\Leftrightarrow \quad (\mu_\gamma \circ \pi_f)(F, G) > 0$$

In a crisp spatial database system this gives us the chance also to take the transition zones into account and to let them make the predicate $\gamma\,\pi_f$ true. When evaluating a fuzzy spatial selection or join in a fuzzy spatial database system, we can even set up a weighted ranking of database objects satisfying the predicate $\gamma\,\pi_f$ at all and being ordered by descending membership degree $1 \geq \mu_\gamma > 0$.

A special, optional fuzzy modifier, denoted by *at all*, represents the existential modifier and checks whether a predicate π_f can be fulfilled to any extent. An example query is: "Do regions A and B (*at all*) overlap?" With this modifier we can determine whether $\mu_\gamma(x) > 0$ for some value $x \in [0, 1]$.

The following few example queries demonstrate how fuzzy spatial data types and quantified fuzzy topological predicates can be integrated into an SQL-like spatial query language. It is not our objective to give a full description of a specific language. We assume a relational data model where tables may contain fuzzy regions as attribute values.

What we need first is a mechanism to declare user-defined fuzzy modifiers and to activate predefined or user-defined fuzzy modifiers. This mechanism should allow to specify trapezoidal-shaped and triangular-shaped membership functions as well as crisp partitions. In general, this means to define a *classification*, which could be expressed in the following way:

create classification fq

$$
\begin{array}{ll}
(not & (0.00, 0.00, 0.00, 0.02), \\
a\ little\ bit & (0.01, 0.02, 0, 03, 0.08), \\
somewhat & (0.03, 0.08, 0.15, 0.25), \\
slightly & (0.15, 0.25, 0.45, 0.55), \\
quite & (0.45, 0.55, 0.75, 0.85), \\
mostly & (0.75, 0.85, 0.92, 0.96), \\
\\
nearly\ completely & (0.92, 0.96, 0.97, 0.99), \\
completely & (0.97, 1.0, 1.0, 1.0))
\end{array}
$$

Such a classification could then be activated by

set classification   fq

Assuming that we have a relation *pollution*, which stores among other things the blurred geometry of polluted zones as fuzzy regions, and a relation *areas*, which keeps information about the use of land areas and which stores their vague spatial extent as fuzzy regions. A query could be to find out all inhabited areas where people are rather endangered by pollution. This can be formulated in an SQL-like style as (we here use infix notation for the predicates):

> **select** areas.name
> **from**   pollution, areas
> **where** area.use = inhabited **and**
>   pollution.region *quite overlaps* areas.region

This query and the following two ones represent *fuzzy spatial joins*.

Another query could ask for those inhabited areas lying almost entirely in polluted areas:

> **select** areas.name
> **from**   pollution, areas
> **where** areas.use = inhabited **and**
>   areas.region *nearly completely inside*
>   pollution.region

Assume that we are given living spaces of different animal species in a relation *animals* and that their vague extent is represented as a fuzzy region. Then we can search for pairs of species which share a common living space to some degree:

> **select** A.name, B.name
> **from**   animals A, animals B
> **where** A.region *at all overlaps* B.region

As a last example, we can ask for animals that usually live on land and seldom enter the water or for species that never leave their land area (the built-in aggregation function **sum** is applied to a set of fuzzy regions and aggregates this set by repeated application of fuzzy geometric union):

```
select  name
from    animals
where  (select sum(region) from areas)
           nearly completely covers or
           completely covers region
```

11 Conclusions

In this chapter we have introduced a fuzzy spatial algebra that introduces spatial data types for fuzzy points, fuzzy lines, and fuzzy regions, and that deals with fuzzy topological predicates operating on these data types. Structure and semantics of types and predicates are formally defined. The characteristic feature of the design is the modeling of smoothness and continuity which is inherent to the objects themselves and to the transitions between different fuzzy objects.

Future work will have to deal with the design of fuzzy spatial operations and the implementation of the whole algebra.

References

1. D. Altman. Fuzzy Set Theoretic Approaches for Handling Imprecision in Spatial Analysis. *Int. Journal of Geographical Information Systems*, 8(3):271–289, 1994.
2. M. Blakemore. Generalization and Error in Spatial Databases. *Cartographica*, 21, 1984.
3. P. A. Burrough. *Natural Objects with Indeterminate Boundaries*, pages 3–28. In Burrough and Frank [4], 1996.
4. P. A. Burrough and A. U. Frank, editors. *Geographic Objects with Indeterminate Boundaries*. GISDATA Series, vol. 2. Taylor & Francis, 1996.
5. C. L. Chang. Fuzzy Topological Spaces. *Journal of Mathematical Analysis and Applications*, 24:182–190, 1968.
6. E. Clementini, P. Di Felice, and G. Califano. Composite Regions in Topological Queries. *Information Systems*, 20(7):579–594, 1995.
7. E. Clementini and P. Di Felice. *An Algebraic Model for Spatial Objects with Indeterminate Boundaries*, pages 153–169. In Burrough and Frank [4], 1996.
8. A. G. Cohn and N. M. Gotts. *The 'Egg-Yolk' Representation of Regions with Indeterminate Boundaries*, pages 171–187. In Burrough and Frank [4], 1996.
9. D. Dubois and M.-C. Jaulent. A General Approach to Parameter Evaluation in Fuzzy Digital Pictures. *Pattern Recognition Letters*, pages 251–259, 1987.
10. S. Dutta. Qualitative Spatial Reasoning: A Semi-Quantitative Approach Using Fuzzy Logic. In *1st Int. Symp. on the Design and Implementation of Large Spatial Databases*, LNCS 409, pages 345–364. Springer-Verlag, 1989.
11. S. Dutta. Topological Constraints: A Representational Framework for Approximate Spatial and Temporal Reasoning. In *2nd Int. Symp. on Advances in Spatial Databases*, LNCS 525, pages 161–180. Springer-Verlag, 1991.

12. M. J. Egenhofer. A Formal Definition of Binary Topological Relationships. In *3rd Int. Conf. on Foundations of Data Organization and Algorithms*, LNCS 367, pages 457–472. Springer-Verlag, 1989.

13. M.J. Egenhofer, E. Clementini, and P. Di Felice. Topological Relations between Regions with Holes. *Int. Journal of Geographical Information Systems*, 8(2):128–142, 1994.

14. M. Erwig and M. Schneider. Partition and Conquer. In *3rd Int. Conf. on Spatial Information Theory*, LNCS 1329, pages 389–408. Springer-Verlag, 1997.

15. M. Erwig and M. Schneider. Vague Regions. In *5th Int. Symp. on Advances in Spatial Databases*, LNCS 1262, pages 298–320. Springer-Verlag, 1997.

16. J. T. Finn. Use of the Average Mutual Information Index in Evaluating Classification Error and Consistency. *Int. Journal of Geographical Information Systems*, 7(4):349–366, 1993.

17. S. Gaal. *Point Set Topology*. Academic Press, 1964.

18. R. H. Güting and M. Schneider. Realm-Based Spatial Data Types: The ROSE Algebra. *VLDB Journal*, 4:100–143, 1995.

19. V. J. Kollias and A. Voliotis. Fuzzy Reasoning in the Development of Geographical Information Systems. *Int. Journal of Geographical Information Systems*, 5(2):209–223, 1991.

20. P. Lagacherie, P. Andrieux, and R. Bouzigues. *Fuzziness and Uncertainty of Soil Boundaries: From Reality to Coding in GIS*, pages 275–286. In Burrough and Frank [4], 1996.

21. F.P. Preparata and M.I. Shamos. *Computational Geometry*. Springer Verlag, 1985.

22. M. Schneider. *Modelling Spatial Objects with Undetermined Boundaries Using the Realm/ROSE Approach*, pages 141–152. In Burrough and Frank [4], 1996.

23. M. Schneider. *Spatial Data Types for Database Systems - Finite Resolution Geometry for Geographic Information Systems*, volume LNCS 1288. Springer-Verlag, Berlin Heidelberg, 1997.

24. R. Shibasaki. A Framework for Handling Geometric Data with Positional Uncertainty in a GIS Environment. *GIS: Technology and Applications*, pages 21–35, World Scientific, 1993.

25. R. B. Tilove. Set Membership Classification: A Unified Approach to Geometric Intersection Problems. *IEEE Trans. on Computers*, C-29:874–883, 1980.

26. E. L. Usery. *A Conceptual Framework and Fuzzy Set Implementation for Geographic Features*, pages 71–85. In Burrough and Frank [4], 1996.

27. F. Wang. Towards a Natural Language User Interface: An Approach of Fuzzy Query. *Int. Journal of Geographical Information Systems*, 8(2):143–162, 1994.

28. F. Wang, G. B. Hall, and Subaryono. Fuzzy Information Representation and Processing in Conventional GIS Software: Database Design and Application. *Int. Journal of Geographical Information Systems*, 4(3):261–283, 1990.

29. M.F. Worboys and P. Bofakos. A Canonical Model for a Class of Areal Spatial Objects. In *3rd Int. Symp. on Advances in Spatial Databases*, LNCS 692, pages 36–52. Springer-Verlag, 1993.

30. L. A. Zadeh. Fuzzy Sets. *Information and Control*, 8:338–353, 1965.

31. F. B. Zhan. Topological Relations between Fuzzy Regions. In *ACM Symp. for Applied Computing*, pages 192–196, 1997.

Index