FIFUS: A Rule-Based Fuzzy Inference Model for Fuzzy Spatial Objects in Spatial Databases and GIS

Anderson Chaves Carniel
Dept. of Computer Science
University of São Paulo
at São Carlos
São Carlos, SP 15560-970,
Brazil
accarniel@gmail.com

Markus Schneider
Dept. of Computer &
Information Science &
Engineering
University of Florida
Gainesville, FL 32611, USA
mschneid@cise.ufl.edu

Ricardo Rodrigues Ciferri
Dept. of Computer Science
Federal University of
São Carlos
São Carlos, SP 15565-905,
Brazil
ricardo@dc.ufscar.br

ABSTRACT

Decision support based on spatial (and not only alphanumerical) data has received increasing interest in geographical applications, such as geoscience, agriculture, and economics applications, and has led to *Spatial Decision Support Systems* (*SDSS*). SDSS use spatial database systems and Geographical Information Systems as their data management and analysis components in order to get and handle the needed spatial data and perform recommendations, estimations, or predictions. For instance, farmers want to know what the best areas of their farmland are to grow a specific crop. In most cases, the extent and the properties of the spatial phenomena of interest are vague and imprecise. They can be adequately represented by *fuzzy spatial objects* (e.g., fuzzy points, fuzzy lines, fuzzy regions). In this paper, we formally propose a model named *Fuzzy Inference on Fuzzy Spatial Objects* (*FIFUS*), which infers recommendations, estimations, and predictions based on fuzzy rules and knowledge of domain specialists. It incorporates fuzzy spatial objects into the components of the existing fuzzy inference methods in order to take into account the spatial imprecision found in the real world. As a main advantage, FIFUS is a general-purpose model and can thus be applied in many geoscience applications.

Categories and Subject Descriptors

H.2.8 [**Database Management**]: Spatial databases and GIS

General Terms

Design, Languages

Keywords

Spatial decision support system, fuzzy spatial objects, fuzzy inference, spatial fuzziness

SIGSPATIAL'15, November 03-06, 2015, Bellevue, WA, USA
© 2015 ACM. ISBN 978-1-4503-3967-4/15/11$15.00
DOI:http://dx.doi.org/10.1145/2820783.2820874.

1. INTRODUCTION

Decision support and decision making based on spatial (and not only alphanumerical) data have been increasingly required in geographical applications, such as geoscience, agriculture, and economics applications. This requirement has led to the development of *Spatial Decision Support Systems* (*SDSS*). The goal of a SDSS is to combine conventional analysis (e.g., by using alphanumerical data) with spatial analysis aiming to aid in strategic decision making. Commonly, SDSS use spatial database systems and Geographical Information Systems as their data management and analysis components in order to get and handle the needed spatial data to be considered. As a result, SDSS can make *recommendations*, *estimations*, or *predictions* for the user. An application example are farmers that want to know what the best areas of their farmland are to grow a specific crop.

In most cases, the extent and the properties of the spatial phenomena of interest are vague and imprecise. *Spatial fuzziness* (or *vagueness*) characterizes real-world phenomena that do not have exact locations or that have some imprecision in the definition of their boundaries and interiors. *Fuzzy spatial data types* for fuzzy points, fuzzy lines, and fuzzy regions have been defined to adequately represent spatial fuzziness by using fuzzy set theory as a basis. In general, a membership degree, a real value between 0 and 1, is assigned to each point of a fuzzy spatial object. For instance, we can represent a sandy soil by using a fuzzy region. Each point of this region has a different membership degree indicating the possibility of this point to be sandy.

Since spatial fuzziness is present in real-world phenomena, *fuzzy inference systems* have been a popular tool in the SDSS context. A fuzzy inference system infers new conclusions by using approximate reasoning according to previous knowledge, such as Mamdani's Inference. In this sense, SDSS employ fuzzy inference systems to make recommendations, estimations, and predictions. However, the available approaches that use fuzzy inference systems in SDSS face several problems, such as the consideration of only alphanumeric characteristics of fuzzy spatial objects. Another problem of the majority available approaches is that they do not combine fuzzy inference methods together with spatial fuzziness, that is, they use only one kind of analysis in order to make recommendations, estimations, and predictions.

In this paper, we fill this gap by proposing a model named *Fuzzy Inference on Fuzzy Spatial Objects* (*FIFUS*). It does not have the aforementioned drawbacks but offers the following benefits: (i) it provides a formal way to model a SDSS problem, (ii) it incorporates an adequate representation of spatial fuzziness provided by fuzzy spatial objects into concepts of fuzzy inference systems by extending the fuzzy inference engine, and (iii) it is a general-purpose model which can be applied in many spatial applications. Thus, users are able to model their application problems and obtain spatial recommendations, estimations, or predictions inferred from real-world phenomena represented by fuzzy spatial objects.

This paper is organized as follows. Section 2 discusses related work. Section 3 summarizes the needed theoretical concepts used in this paper. Section 4 formally defines FIFUS. Finally, Section 5 concludes the paper and presents future work.

2. RELATED WORK

There are many approaches that incorporate spatial fuzziness found in real-world phenomena into SDSS [1, 2, 4, 6]. We characterize these approaches with regard to (i) the range of the possible characteristics considered in spatial phenomena, (ii) the consideration of spatial fuzziness in fuzzy inference components, and (iii) the support provided for GIS and spatial database systems.

With respect to characterization (i), we can distinguish the approaches according to the quantity of characteristics a point in a spatial phenomenon can have. The approaches in [1, 2] consider only one possible characteristic to classify a point. However, it loses spatial imprecision that could be important in the decision support. Spatial imprecision give us the doubt to define if a point does or does not belong to a specific phenomenon. For instance, we want to classify if a point of a farm is *dune*, *beach*, *bush*, or *crop*. However, due to spatial fuzziness, we do not know if this point is dune or beach. Thus, we consider that this point has different membership degrees for both classes (e.g., 0.4 for beach and 0.6 for dune). On the other hand, our FIFUS model and the approaches in [4, 6] consider that a point can belong to more than one characteristic and, thus, represent spatial imprecision found in real-world phenomena adequately.

With respect to characterization (ii), the available approaches use either alphanumerical data from spatial objects that represent spatial fuzziness or the membership degrees of the points directly in the fuzzy inference components. The majority of the approaches [1, 6] extract properties of spatial objects and consider them as input in the fuzzy inference method, such as the area or diameter of a (fuzzy) region, the length of a (fuzzy) line, or some aggregation operation (e.g., the average of the membership degrees). However, only crisp characteristics are considered so that the full spectrum of the spatial fuzziness is not used. On the other hand, our FIFUS model and some other approaches [2, 4] consider the membership degrees of each point as input in the fuzzy inference method.

Finally, with respect to characterization (iii), our FIFUS model and few approaches [1, 4] consider (fuzzy) spatial data stored in spatial database systems, which provide management support for the application in order to handle real-world phenomena.

3. TECHNICAL BACKGROUND

The FIFUS model incorporates fuzzy spatial objects into fuzzy inference systems. Thus, we use fuzzy sets [7], fuzzy inference systems [3], and fuzzy spatial data types [5] as a basis. Fuzzy sets are used in applications in order to provide support for the uncertainty or vagueness expressed by *linguistic values* (LVal) in the context of a *linguistic variable* (LVar) [7]. A LVar is a variable and assumes LVals in order to give characterizations to a determined situation. Let l be a LVar and v be a LVal. Then, a LVar has a domain of LVals, that is, $dom(l) = \{v_1, ..., v_n\}$ for $n \in \mathbb{N}$.

By using LVars and LVals we are able to construct fuzzy rules. A fuzzy rule is an *IF P THEN Q* rule with *fuzzy propositions* in the antecedent P and in the consequent Q [3]. Let l be a LVal and $v \in dom(l)$. Then, a fuzzy proposition has the following format l *is* v. Fuzzy propositions can be linked by using logical operators, such as AND or OR. Further, this means that the fuzzy propositions of P implies in the fuzzy propositions of Q, that is, $P \to Q$.

Fuzzy rules, fuzzy sets, and a fuzzy inference engine compose a fuzzy inference system [3]. The fuzzy inference engine executes a combination of *implication* and *composition*, which can vary according to existing fuzzy inference methods [3]. For each fuzzy rule, it computes the membership degree of each input value of each antecedent. Thus, each antecedent has an input value. Then, these membership degrees are combined according to logical operators among the propositions of the antecedent. The resulting value is applied in the consequent of the fuzzy rule (*implication*). Finally, the resulting implications are combined (*composition*), which yields the reasoning conclusion. The kind of calculation of the logical operator, implication, and composition is defined by the fuzzy inference method [3]. For instance, *Mamdani's inference* [3] uses the minimum operator for the AND logical operator and implication and the maximum operator for the OR logical operator and composition. Finally, the result of the composition may pass by a defuzzification method in order to extract a numeric element that best represents a fuzzy set [3] and, thus, a value with meaning for the user.

Since FIFUS uses fuzzy spatial objects, here we only provide a general view of each fuzzy spatial data type. Their special properties and formal definitions can be found in [5]. A fuzzy point (*fpoint*) object is defined as a set of tuples defined as (x, y, λ) such that $(x, y) \in \mathbb{R}$ and $\lambda \in]0, 1]$. Thus, a fuzzy membership function for a simple fuzzy point p can be defined as $\mu_{\tilde{p}(x,y)}(a, b) = \lambda \in]0, 1]$ if $(x, y) = (a, b)$ holds, and $\mu_{\tilde{p}(x,y)}(a, b) = 0$ otherwise. A fuzzy line (*fline*) object is defined as a set of simple fuzzy lines. A simple fuzzy line can be defined as a sequence of simple fuzzy points defined as $\langle(x_1, y_1, \lambda_1), ..., (x_n, y_n, \lambda_n)\rangle$ for some $n \in \mathbb{N}$. Each segment is composed by two simple fuzzy points that are neighbors. For instance, $(x_{n-1}, y_{n-1}, \lambda_{n-1})$ and (x_n, y_n, λ_n) compose the last segment of a simple fuzzy line. An interpolation function can be used as fuzzy membership function to calculate membership degrees between two simple fuzzy points of a segment. Finally, a fuzzy region (*fregion*) object is defined as a set of simple fuzzy regions. A simple fuzzy region $\tilde{R}$ has the same geometric format than a crisp region but is equipped with a fuzzy membership function $\mu_{\tilde{R}} : \mathbb{R}^2 \to]0, 1]$ that assigns a membership degree to each point in $\tilde{R}$, that is, $\tilde{R} = \{((x, y), \mu_{\tilde{R}}(x, y)|(x, y) \in \mathbb{R}^2\}$. Thus, a fuzzy region is a bounded and regular bi-dimensional subset of $\mathbb{R}^2$. The distribution of membership degrees in a fuzzy simple region can be smooth, continuous, or piecewise continuous.

4. FIFUS: A FUZZY INFERENCE MODEL FOR FUZZY SPATIAL OBJECTS

In this paper, we propose a model called *Fuzzy Inference on Fuzzy Spatial Objects* (*FIFUS*), which incorporates fuzzy spatial objects into the fuzzy inference process by using two main components: (i) Data Source Component (Section 4.1) and (ii) Extraction Component (Section 4.2). In order to illustrate the execution of our proposed FIFUS model, we implemented a web agriculture application by using the Java programming language and PostGIS to store spatial objects. This application can be accessed at `http://gbd.dc.ufscar.br/fifus/`.

4.1 Data Source Component

The Data Source Component comprises the following elements: (i) fuzzy sets, (ii) fuzzy rules set, and (iii) fuzzy spatial objects. The element (i) and (ii) compose the knowledge base according to fuzzy inference systems. However, we additionally incorporate the element (iii).

We firstly begin to define how fuzzy spatial objects are represented in our FIFUS model. For this purpose, we define a *Fuzzy Spatial Antecedent* (FSA). FSAs are antecedents of a fuzzy rule and thus represent the main characteristics of a problem. Further, each FSA is a layer composed of several fuzzy spatial objects that represent real-world phenomena having some spatial imprecision. Therefore, we define a FSA as a pair $\langle LVar, V \rangle$ such that

(i) $V = \{(fso_1, lval_1), ..., (fso_k, lval_m)\}$
(ii) $k, m \in \mathbb{N} \wedge k, m \geq 1$
(iii) $1 \leq i \leq k : fso_i \in \alpha \in \{fregion, fline, fpoint\}$
(iv) $1 \leq i \leq k, 1 \leq j \leq k, i \neq j : fso_i \neq fso_j$
(v) $1 \leq i \leq m : lval_i \in dom(LVar)$

A FSA is annotated by a LVar, and it is composed of one or more pairs (conditions (i) and (ii)) of fuzzy spatial objects (*fso*) associated with LVals (conditions (iii) and (v)). The fuzzy spatial objects must be different (condition (iv)) and thus each fuzzy spatial object has only one LVal.

Since a fuzzy rule is composed of antecedents and consequents, we also need to define the format of the consequents. We define a *Consequents Set* (CS) as the expected result of the application, e.g., a recommendation, estimation, or prediction. We define a CS as a pair $\langle LVar, V \rangle$ such that

(i) $V = \{(lval_1, fs_1), ..., (lval_k, fs_k)\}$
(ii) $k \in \mathbb{N} \wedge k \geq 1$
(iii) $1 \leq i \leq k : lval_i \in dom(LVar)$
(iv) $1 \leq i \leq k : fs_i$ is a fuzzy set

A CS has a LVar associated and a set of pairs (condition (i) and (ii)). Each pair is formed by a LVal that represents qualitatively its LVar (condition (iii)) and a traditional fuzzy set (condition (iv)).

Since we have the definition of antecedents (FSAs) and consequents (CSs), users have to define the fuzzy rules. FIFUS does not automatically generate fuzzy rules. Thus, the fuzzy rules set must be define textually by specialists by using the LVars and their respective LVals provided by the FSAs and CSs. The process to define a fuzzy rule is similar to what is done in the traditional fuzzy inference systems. The main difference is that we use fuzzy spatial objects as antecedents in fuzzy rules. However, the domain specialists do not need to know how the fuzzy spatial objects are modeled. Therefore, fuzzy rules constructed from previous inference systems that have the same LVars and LVals can be also used in the FIFUS.

4.2 Extraction Component

The Extraction Component takes a target spatial object from the user as input to be inferred a new knowledge by considering the data from the Data Source Component (Section 4.1). It comprises two interacting elements: (i) an extraction engine, and (ii) a fuzzy inference engine. The extraction engine is the main element, which takes the target spatial object as input to be inferred according to fuzzy rules, FSAs, and CSs. A target spatial object t is a crisp spatial object of the well known spatial data types *point*, *line*, and *region*, that is, $t \in \alpha \in \{point, line, region\}$. However, since a crisp line and a crisp region is composed of an infinite point set, we are not able to compute computationally *all* their points. Hence, we make a *discretization process* as an intermediary step. A discretization process extracts, from a line object or a region object, a finite point set by using some strategy. For instance, for crisp lines, the discretization can be achieved by intercalating points on the line extension. For crisp regions, we can use its *Bounding Box* (BBOX) to trace lines parallel to the BBOX boundary and then extract intercalated points from these lines.

Algorithm 1 provides our extraction algorithm. Its inputs are

1. P: a crisp point as a (x, y) coordinates pair.
2. RS: the fuzzy rules set represented as a hash table with the following format: the key of a record is formed by the linguistic values of the antecedents separated by commas and the value is formed by the linguistic values of the consequents separated by commas.
3. DB: a list of FSAs. Each position (i.e., each FSA) of this list has an array that stores its fuzzy spatial objects with their linguistic values.
4. IM: a fuzzy inference method to be used in the reasoning process (e.g., Mamdani's inference).
5. DM: a defuzzification method to be used. Note that this is only required if the result of the IM is a fuzzy set.

The first step of the algorithm is to find the fuzzy spatial objects that contain the point P with some membership degree greater than 0 (lines 1 to 5). For this purpose, we make use of N which is a list of arrays and has the same structure as DB (line 1). Then, for each array l_i in DB (line 2), we traverse its elements (line 3) to verify if P has a membership degree greater than 0 in a value v_j of l_i (line 4). In the positive case, we store the element v_j in the array N_i (line 5). Thus, the array N_i will only have the fuzzy spatial objects that contain P with some membership degree greater than 0; it corresponds to the position i of DB. It is important to know the position of these elements for their use in the fuzzy rules since the position i in DB means that its linguistic variable has the position i in a fuzzy rule.

The next step performs the Cartesian product among the values of the arrays stored in N in order to find all combinations of the linguistic values of the stored fuzzy spatial objects (line 6). The result is a matrix K that has n columns, which corresponds to the same number of arrays, and hence FSAs, and m rows, which is the product operator among the sizes of all arrays of N. After this, we are able to start the inference process. In order to execute it, we need the following variables (line 7): (i) *result* which will store the inferred value, (ii) *ant* which is an array to store the linguistic values of the antecedents contained in each cell of the matrix, (iii) *input* which is an array to store the membership degrees of P in a fuzzy spatial object contained in a cell of the matrix, and (iv) *impl* which is a fuzzy set to store the result of the implication process.

Each row of the matrix indicates a fuzzy rule to be executed by the fuzzy inference method *IM* (lines 8 to 14). Each cell of a row indicates the fuzzy spatial object of each antecedent of a fuzzy rule. Then, for each cell of a row of this matrix (line 9), we add the linguistic values of the fuzzy spatial objects in the antecedent array (line 10). It is performed to know which fuzzy rule we will execute by considering the *RS*. Further, we compute the membership degree of P in the current fuzzy spatial object (cel_j), which is stored in the input array (line 11). It is performed to know the input of the antecedent j. Therefore, we are carefully considering the order of the antecedents in fuzzy rules.

The next step is the execution of the implication process according to the fuzzy inference method *IM* (line 12). In order to perform this operation, we firstly determine which fuzzy rule will be executed from the *RS* according to the antecedents (function *getFRule* of line 12). Then, $IM_{implication}$ executes the current fuzzy rule according to the inputs. It is done for each row of the matrix K since each row represents a fuzzy rule to be executed. If the antecedents do not exist in *RS*, this fuzzy rule is discarded by the inference method. The result of $IM_{implication}$ is aggregated incrementally by the composition process of $IM_{composition}$ (line 13). Finally, if *IM* is a fuzzy inference method that returns a fuzzy set as result (line 15), we use *DM* to defuzzify the resulting fuzzy set and assign it to the crisp point P (line 16). Otherwise, P is assigned with the inferred value (line 18). In both cases, P is returned with a value that indicates the reasoning conclusion.

5. CONCLUSIONS AND FUTURE WORK

This paper proposes FIFUS, a novel model that considers fuzzy spatial objects in fuzzy inference methods. This model allows that users of SDSS take into account spatial imprecisions found in real-world phenomena in their strategic decision making. In this sense, FIFUS has as main advantages: (i) an adequate modeling of application problems by using the Data Source Component, which uses fuzzy spatial objects to represents spatial fuzziness, (ii) the Extraction Component that has an algorithm in which it considers fuzzy spatial objects in the execution of the reasoning process, (iii) the use of existing fuzzy inference methods, which allows the employment of several existing fuzzy inference engines, and (iv) applicability in many spatial applications.

Future work topics mainly are in the empirical studies of applicability of FIFUS in real applications. Our goal is to evaluate FIFUS by using case studies with users as well as conduct a performance evaluation. Other future work is to remove the discretization process and consider a crisp region or a crisp line directly as input. Thus, comparisons between these approaches would be performed.

Algorithm 1: Fuzzy Inference on Fuzzy Spatial Objects

Input: P, *RS*, *DB*, *IM*, and *DM*
Output: a crisp point with an inferred value associated
1 N is an empty list of arrays with size equal to $size(DB)$;
2 **foreach** l_i **in** DB **do**
3   **foreach** v_j **in** DB_i **do**
4   **if** ***membership***$(P,\ v_j) > 0$ **then**
5   **add**$(N_i,\ v_j)$;
6 $K \leftarrow$ **cartesianProduct**(N);
7 *result* $\leftarrow$ *empty*(); *ant*[] $\leftarrow$ *empty*(); *input*[] $\leftarrow$ *empty*(); *impl* $\leftarrow$ *empty*();
8 **foreach** fr_i **in** K **do**
9   **foreach** cel_j **in** fr_i **do**
10   **add**(*ant*, **lval**(cel_j));
11   **add**(*input*, **membership**(P, cel_j));
12   *impl* $\leftarrow IM_{implication}$(**getFRule**(*RS*, *ant*), *input*);
13   *result* $\leftarrow IM_{composition}$(*result*, *impl*);
14   *ant* $\leftarrow$ *empty*(); *input* $\leftarrow$ *empty*();
15 **if** *result is a fuzzy set* **then**
16   **return** ***assign***$(P, DM(result))$;
17 **else**
18   **return** ***assign***$(P, result)$;

Acknowledgments

First and third authors have been supported by the Brazilian research agencies FAPESP, CAPES, and CNPq.

6. REFERENCES

[1] A. Calzada, J. Liu, H. Wang, and A. Kashyap. A GIS-based spatial decision support tool based on extended belief rule-based inference methodology. In *4th International Workshop on Knowledge Discovery, Knowledge Management and Decision Support*, pages 388–395, 2013.

[2] A. Jadidi, M. A. Mostafavi, Y. Bédard, and K. Shahriari. Spatial representation of coastal risk: A fuzzy approach to deal with uncertainty. *ISPRS International Journal of Geo-Information*, 3(3):1077–1100, 2014.

[3] J.-S. Jang. Fuzzy inference systems. In *Neuro-Fuzzy and Soft Computing: A Computational Approach to Learning and Machine Intelligence*, pages 73–91. Prentice-Hall, 1997.

[4] J. Jasiewicz. A new GRASS GIS fuzzy inference system for massive data analysis. *Computers and Geosciences*, 37(9):1525–1531, 2011.

[5] M. Schneider. Fuzzy spatial data types for spatial uncertainty management in databases. In J. Galindo, editor, *Handbook of Research on Fuzzy Information Processing in Databases*, pages 490–515. IGI Global, 2008.

[6] X. Tang. *Spatial Object Modeling in Fuzzy Topological Spaces with Applications to Land Cover Change*. PhD thesis, International Institute for Geo-information Science & Earth Observation, 2004.

[7] L. A. Zadeh. Outline of a new approach to the analysis of complex systems and decision processes. *IEEE Transactions on Systems, Man and Cybernetics*, SMC-3(1):28–44, Jan 1973.