Contourline Based Modeling of Vague Regions

A. Hallez, R. De Caluwe, J. Verstraete, G. De Tré
Ghent University, Dept. of Telecommunications and Information Processing
Sint-Pietersnieuwstraat 41, B-9000 Ghent, Belgium
Axel.Hallez@rug.ac.be

Abstract

This paper introduces a model based on contour lines to represent a vague region modeled as a fuzzy set. The model allows the user to adjust the accuracy of the membership function to the needs of the application and can enforce the continuity of this function whenever desirable.

keywords: Vague Regions, Spatial Data.

1 Introduction

Recently there has been a growing interest in the modeling of vagueness in spatial regions. Several authors have proposed a representation of spatial regions with non-crisp boundaries by introducing broad boundaries ([1],[2],[4]). In those papers a vague region is characterized by an inner and an outer border as shown in figure 1. The *outer border* delimits the set of points that belong to the region to a certain extent, the *inner border* delimits the set of points that belong to the region for certain. The subregion of points lying between these borders is called the *broad boundary*.

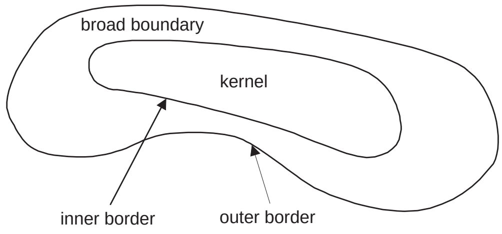


Figure 1: Sample of a vague region.

The papers cited above mainly discuss the topological properties of regions that are represented by an inner and an outer border, ignoring the available information about points in the broad boundary.

In [6] a method to assign a membership degree to the points lying in the broad boundary has been presented. In this method the membership degree of a point is based on the distances from the point to the inner and to the outer border. This is a very straightforward method, but the membership function is completely determined by both borders only, so there is no possibility to refine the membership function in the broad boundary if desired.

This paper presents an extension of this method, that allows for a more accurate modeling of the membership function in the broad boundary by introducing contourlines. In fact, these contourlines provide a very flexible means of defining a membership function with arbitrary accuracy.

Section 2 explains how contourlines are defined. Section 3 and section 4 explain how the contourlines are used to calculate the membership degrees of individual points of a vague region.

2 Contourlines in the broad boundary

A region is a subset of a two dimensional Euclidian space U. Vague regions are a generalization of regions. In the model explained below, vague regions will be represented by fuzzy sets. The outer border of a vague region X is the boundary[1] of the support of X ($supp(X) = \underline{X}_0$) [2], the inner border is the boundary of the kernel of X ($ker(X) = X_1$).

The concept of using a border to define α-level cuts can be applied to membership degrees other than 0 and 1, using information about points in the broad boundary. It is possible to define additional borders, each of these borders corresponding to a specific membership degree $m \in]0,1[$ such that it is the boundary of X_m.

The set of membership degrees for which a border is defined to model a vague region X, will be denoted $M_X \subseteq [0,1]$.

To be useful in the context of the presented model (as will be explained further), M_X must be a finite set that contains at least two elements, including 0. It is not required to include 1 in M_X, which makes it possible to use non-normalized fuzzy sets to represent vague regions. Since X_0 is undefined, the membership degree 0 has to be treated as a special case.

The border corresponding to a given membership degree m does not necessarily correspond to one single line. In the case that X_m has a "hole" in it or is a disconnected set, the corresponding border will correspond to two or more closed lines.

A closed line that is part of a border will be called a *contour line*[3]. The set of all contour lines that are defined to model a vague region X will be denoted C_X. A contour line corresponding to the membership degree m will be called an m-contour line. The subset of C_X of all contour lines that correspond to the membership degree m will be denoted C_{X_m}.

In order for the presented model to be applicable on a vague region, the following restrictions must be satisfied:

- The vague region must be bound. (It is not possible to define the membership function of an unbound region in this model, as at least the 0-contour lines cannot be defined).
- The support of the vague region must be a connected set. If necessary, a vague region with a disconnected support can be modeled as the union of two or more vague regions with a connected support.
- The α-cuts for membership degrees other than zero can be disconnected sets, but they have to be the union of a limited number of connected subsets (to exclude regions with an infinite number of contour lines).

Figure 2 is an example of a vague region. This region has one contour line for the support (0-contourline) and one 0.5-contour line. The kernel consists of two disconnected parts, hence its boundary is made up of two contour lines.

Given the definition of contour lines, it is possible that several contour lines, corresponding to different membership degrees, (partially) coincide. As a most striking example, the contourlines of a crisp region defined with this

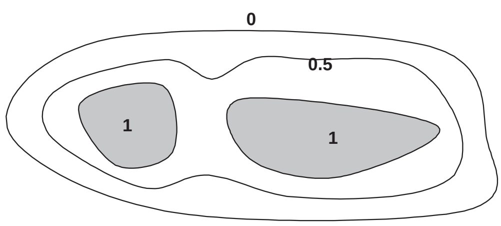


Figure 2: Sample of a fuzzy region with a disconnected kernel (marked grey).

[1] A point p is called a boundary point of a (crisp) set S iff every non-empty environment of p intersects with both S and the complement of S. The set of all boundary points of S is called the boundary of S, denoted by $Bd(S)$.

[2] $X_\alpha ::= \{x | X(x) >= \alpha\}$ for $\alpha \in]0,1]$, $\underline{X}_\alpha ::= \{x | X(x) > \alpha\}$ for $\alpha \in [0,1[$, see [3].

[3] In [5] (p.66) contour lines are defined as follows: "Contour lines are a representation of isolines for a sample set of elevation values".

model will be the same for all membership degrees, as the outer border (0-contour lines) and the inner border (1-contour lines) will completely coincide.

Given a fuzzy region defined by a set of contour lines, the computation of the membership degree of a point lying in the broad boundary consists of two parts. Firstly, the contour lines that will be used for the computation of the membership degree of the point, called the *relevant contour lines*, have to be determined. Secondly, the membership degree of the point has to be computed from the membership degrees associated with these relevant contour lines.

3 Determining the relevant contour lines

The contour lines of a region divide the broad boundary in adjoining non-overlapping *subregions*, so that each subregion is bound by a subset of the contour lines and each contour line is part of the boundary of at least one subregion.

This notion of subregions is used to determine the relevant contour lines of a point lying in the broad boundary of a vague region. If a point p is not located on a contour line, it is possible to find the unique subregion that contains p. The contour lines that border this subregion will be used to calculate the membership degree of p (section 4).

3.1 A contour line tree

A tree of contour lines will be used to determine the relevant contour lines.

A contour line divides the plane into two parts. One part is the *inside*[4] of the contour line and the other part the *outside* of the contour line. The points of the contour line itself belong to the inside. When a point p belongs to the inside of a contour line C, it is said that C *encloses* p. If all points of a second contour line C' are enclosed by C, then C *encloses* C'.

[4]the region defined by the inner normal of the contourline.

Consider a vague region X. The following rule defines a *tree of all contour lines* in C_X:

$$\begin{aligned}\forall A, B \in C_X : & A \text{ is an ancestor of } B \\ & \Leftrightarrow A \text{ encloses } B.\end{aligned}$$

If there are no contour lines that totally coincide, this rule defines the tree unambiguously.

If two contour lines completely coincide, one contour line encloses the other and vice versa, which causes ambiguity. In this case, a natural order for these contour lines has to be found. If the contour lines correspond to holes in their corresponding α-level cuts, the contour line with the higher membership degree is designated as the ancestor. In all other cases, the contour line with the lower membership degree is designated as the ancestor.

As it is assumed that the support of a vague region is connected, C_X contains exactly one 0-contour line that encloses all other contour lines. As a result, this contour line will be the root of the tree.

A contour line $A \in C_X$ will be a *child* of contour line $B \in C_X$ if B encloses A and

$$\neg \exists C \in C_X : (B \text{ encloses } C) \wedge (C \text{ encloses } A)$$

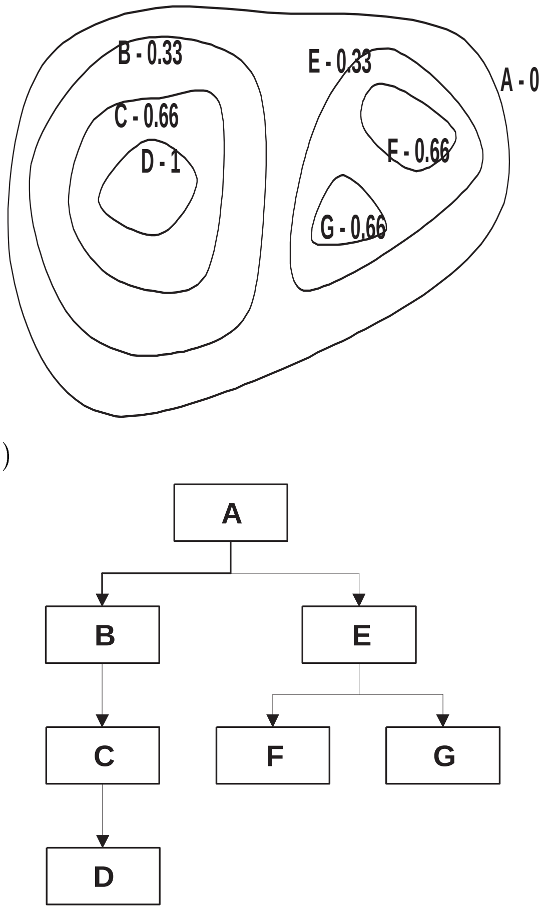
Figure 3(a) shows a more elaborate example of a vague region. Figure 3(b) shows the corresponding contour line tree.

If the contour line tree is known, the search for the relevant contourlines for a point p is carried out by applying the following algorithm:

- *Step 1:* Start with the root of the tree as the current node. If the corresponding contour line encloses p, carry out the step 2 of the algorithm, otherwise the algorithm ends as the region does not contain p (whiwh will be assigned the membership degree 0).

- *Step 2:* Repeat this step until the current node does longer change:
  For each child-node of the current node: if the corresponding contour line encloses p, make this child the current node.

a)



b)

Figure 3: (a) An example of a vague region with labeled contourlines. (b) the corresponding contourline-tree.

The relevant contour lines are the lines corresponding to the current node, obtained at the end of the algorithm, together with its childnodes if any. The process always ends because the set of specified contour lines must be finite in this model.

The next section shows how the contour lines and their associated membership degrees are used to calculate a membership degree for p.

4 Interpolation method

The distance d between a point p and a contour line C is defined as:

$$d(C,p) = d(p,C) = \min_{p' \in C} d(p',p)$$

It can be proved that this function is continuous in every point $p \in U$ for a given C.

If a point lies on a contour line, this point is assigned the membership degree corresponding to this contour line. If the point lies on more than one (coinciding) contour lines, it is assigned the highest membership degree corresponding to these contour lines.

The calculation of the membership degree of a point p not lying on a contour line in a vague region X is based on the interpolation of the membership degrees of the relevant contour lines. Based on the number of relevant contour lines for p, three general cases can be distinguished.

4.1 One relevant contour line

In this case the only relevant contour line does not enclose any other contour line (e.g. contour line F in figure 3) p is assigned the membership degree associated with the relevant contour line.

As a result the membership function of X will be continuous for all points inside the relevant contour line.

4.2 Two relevant contour lines

In this case (e.g. points lying between B and C in figure 3) the membership degree of p is defined as the weighted average of the membership degrees associated with both relevant contour lines. The weights are calculated from the distances from p to the contour lines:

$$X(p) = \frac{\frac{1}{d(p,C_1)}X(C_1) + \frac{1}{d(p,C_2)}X(C_2)}{\frac{1}{d(p,C_1)} + \frac{1}{d(p,C_2)}} \quad (1)$$

$$= \frac{d(p,C_2)X(C_1) + d(p,C_1)X(C_2)}{d(p,C_1) + d(p,C_2)} \quad (2)$$

where C_1 and C_2 are the relevant contour lines, $X(C_1)$ and $X(C_2)$ the associated membership degrees.

Taking into account that $d(C_1,p)$ and $d(C_2,p)$ are continuous functions in all points $p \in U$, it can be proved that (1) is a continuous function for all points $p \in U \backslash (C_1 \bigcup C_2)$. Moreover, the limit of the function in points of C_1 is $X(C_1)$ and the limit of the function in points of C_2 is $X(C_2)$.

4.3 More than two relevant contour lines

This is the most complicated case. (e.g. the case which holds for points lying between the contour lines A, B and E in figure 3.) There are several options for calculating $X(p)$. Depending on the option that is used, the resulting membership function will have different properties.

One option is to choose two elements of $\{C_i | i = 1, ..., n\}$, thus reducing this case to the previous case. It is obvious that two contour lines will be chosen based on the distance of the contour lines to the p. However, simply selecting the two contour lines that lie closest to p will not result in a good membership function. First of all the membership degree of points that are equidistant to two or more contour lines is not always well defined in this case. Of course, this could be solved by choosing one of both equidistant contour lines (for instance the contourline that results in the highest membership degree), but this way of defining the membership function would conflict with the provided contour lines! For instance, if there are three relevant contour lines C_1, C_2 and C_3 with $X(C_1) = X(C_2) = m_1$, $X(C_3) = m_2$ and $m_1 > m_2$, the points for which C_1 and C_2 are used to calculate the membership degree, will be assigned membership degree m_1. However, the lower membership degree of C_3 indicates that the subregion between these three contour lines does not belong to X_{m_1}. By definition, points that do not belong to X_{m_1} have a membership degree lower than m_1. To get rid of this problem the restriction is introduced that the two contour lines that are used to calculate the membership degree of p, have to correspond to two different membership degrees whenever possible (e.g. if not all contour lines correspond to the same membership degree). With this restriction, the resulting membership function will not only be consistent with the definition of the contour lines, but it will also be continuous and the limit of the function in points lying on a (relevant) contour line will be the membership degree that corresponds to this contour line. These last two properties make the choosen approach useful.

Another option is to generalize (1), so that $X(p)$ becomes the weighted average of all membership degrees associated with all relevant contour lines:

$$X(p) = \frac{\sum_{i=1}^{n} \frac{1}{d(p,C_i)} X(C_i)}{\sum_{i=1}^{n} \frac{1}{d(p,C_i)}} \qquad (3)$$

with $\{C_i | i = 1, ..., n\}$ the set of relevant contour lines for p.

It can be proved that the function (3) is continuous in all points $p \in U \backslash \bigcup_{1 \le i \le n} C_i$. The limit of function (3) taken in a point $p \in C_i$ is $X(C_i)$.

4.4 Continuity of the membership function

In the previous subsections, it has become clear that with the presented method the membership function of a vague region is made up of many different functions. In fact, every subregion as mentioned in section 3 has its own membership function. It has been mentioned that all of these functions are continuous in the part of the region where they are used as the membership function.

As the points lying on a contour line are assigned the corresponding membership degree and the limit of the membership functions, used in the adjacent subregions, in a point lying on the contour line equals to the same membership degree, the membership function will also be continuous on the contour lines itself as long as the contour lines are not coincident. Where two or more contour lines coincide, the membership function will be discontinuous.

As a result, the membership function being continuous or not is not a side effect of the calculations, but is determined by the relative position of the contourlines only. This means that the user of this model has full control over the continuity of the membership function, which can be highly desirable.

5 Conclusion

A very general method for modeling vague regions has been introduced. With this method it is possible to model a very wide range of vague regions and to modify the precision of the membership function by increasing or decreasing the number of membership degrees for which contour lines are defined. It is also possible to control the continuity of the membershipfunction by defining the contour lines to be coincident or not.

References

[1] E. Clementini, P. Di Felice; *An algebraic model for spatial objects with undetermined boundaries*; 1994; GISDATA Specialist Meeting - revised version.

[2] A.G. Cohn, N.M. Gotts; *Spatial regions with undetermined boundaries*; 1994; Proceedings of the Second ACM Workshop on Advances in Geographic Information Systems; pp. 52-59.

[3] D. Dubois, H. Prade; *Fundamentals of Fuzzy Sets*; 2000; Kluwer Academic Publishers.

[4] N.M. Gotts, A.G. Cohn; *A mereological approach to representing spatial vagueness*; 1995; Working Papers, Ninth International Workshop on Qualitative Reasoning; pp. 246-255.

[5] P. Rigaux, M. Scholl, A. Voisard; *Spatial Databases with Application to GIS*; 2002; Academic Press.

[6] J. Verstraete, B. Van Der Cruysen, R. De Caluwe; *Assigning Membership Degrees to Points of Fuzzy Boundaries*; 2000; Proceedings of the 19th International Conference of the North American Fuzzy Information Processing Sociaty; pp. 444-447.