6-1-2006

Shape Types for Labeling Natural Polygon Features with Maplex

Charlie Frye
Environmental Systems Research Institute Inc. (Esri)

cartographic techniques

Shape Types for Labeling Natural Polygon Features with Maplex

Charlie Frye
Environmental Systems Research Institute, Inc.
cfrye@esri.com

INTRODUCTION

Automated label or text placement has made great strides in recent years, particularly with respect to labeling point and line features. However, the same cannot be said for polygon features. One of the main difficulties is determining whether the text should be inside or outside the polygon's perimeter, and whether the text should follow the general trend of the polygon or just flow horizontally within the polygon. The problem is relatively easy to solve when the polygon is substantially larger than the text, or if the text is substantially larger than the polygon. However, for many natural features such as smaller lakes, rivers, canyons, valleys, ridges, or mountain ranges, the text will occupy an area that is not substantially larger or smaller. Also, many of these kinds of features are not simple shapes, but instead have prongs, blobs, or bottlenecks; or are simply splotchy. Each of these kinds of shapes should be approached differently when it comes to cartographic text placement. This article describes a methodology for automatically describing such shapes in order to have Maplex, ESRI's cartographic label placement extension automatically place their names on a map.

OVERVIEW

This paper is about a method for describing polygons that represent natural features, so that feature label placement rules can be derived for Maplex, a cartographic label placement extension for ArcGIS. These label placement rules must take into account preferred placement style (curved, horizontal, etc.), options for type placement (for example character spacing), and coping strategies to deal with contextual circumstances that impinge on the space available for text placement.

The main issue here is that all kinds of polygon shapes can exist in polygon datasets for hydrography or physiographic features. For some of these shapes the best rule would be to curve the text, inside the polygon, along its major trending axis, provided the polygon is large enough. If that polygon is too small, then the text could be curved outside the polygon, along the edge (if the polygon is large enough. If not, then the label could be placed horizontally, and if need be, with a leader line connecting the text and the feature. For round-ish polygons, it makes sense to place the text horizontally inside the features, and if absolutely necessary overrun the feature's edges by a small amount; otherwise if the text is too large, place it outside of the feature. The task here is how to tell Maplex what the general shape of the feature is so that the correct placement rule can be applied.

METHODOLOGY

To accomplish this goal two pieces of information are needed. First is to find the shape of the polygon: is it relatively long, oblong, or round? Second, is to find out how large the polygon is within the context of the space needed to place text. To do this, the polygon's minimum bounding rectangle (MBR) is used. The ratio of the length to the width of the MBR gives a good indication of how long or round the polygon is. Then the percentage of the MBR's area that the polygon occupies indicates whether the polygon is substantial.

The term, minimum bounding rectangle has been used extensively in many computational contexts, and therefore means different things to different people. In this case, it is the smallest rectangle that can be fitted around the polygon, often requiring the rectangle be rotated away from alignment with the x and y axis of the coordinate system used to define the polygon. Figure 1 shows an example of a MBR, and Figure 2 shows, for the sake of this discussion, a minimum bounding envelope (although such has been called a MBR in many other contexts, such as Zhou, *et.al*, 1999).

MBRs have been used extensively in selection and spatial indexing in GIS, to describe shapes and text in rough but efficient terms for conflict detection, and in many other computational and analytical contexts (Examples include Abdelmoty and El-Geresy, 2004; Papadias, *et. al.*, 1995). MBRs have also been used as a rudimentary basis for automated text placement

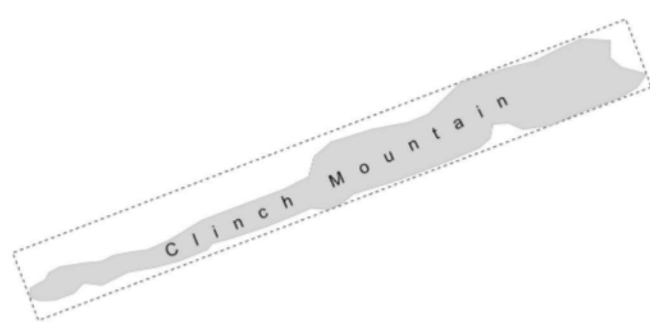

Figure 1. Example of a minimum bounding rectangle.

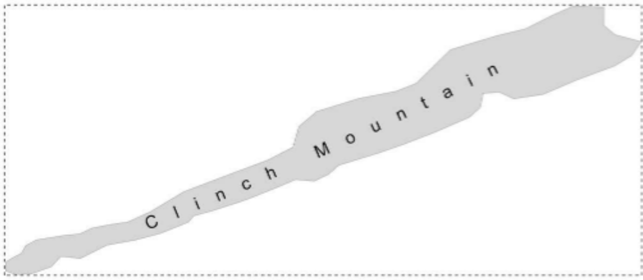

Figure 2. Example of a minimum bounding envelope.

algorithms. This paper formalizes a more refined use of MBRs for text placement.

In order to use this method on a polygon data set, four additional fields that contain information that enables label placement are needed in the attribute table. These four fields would contain information on:

- RatioL2W: double/float. This will contain the ratio of the MBRs length to its width for each feature; it is for analysis and evaluation of results
- MBRArea: double/float. This will contain the percentage of the MBR's area that the polygon occupies; it is for analysis and evaluation of results
- LabelSize: short integer. This is optional, but useful if the range between the size of the smallest polygon and the size of the largest polygon is more than two orders of magnitude.
- LabelType: short integer. This is required and contains values ranging from 1 to 7, describing the seven types of shapes will be the basis for Maplex rules. The seven shape types are:

1. Round-ish: (see figure 3)
2. Oblong (see figure 4)
3. Long (see figure 5)
4. Long and Skinny (see figure 6)
5. Splotchy (see figure 7)
6. Snaky or Pronged (see figure 8)
7. Snaky or Pronged and Skinny (see figure 9)

In order to create the information that is stored in these fields, a Python script was written and uses an ArcGIS 9.2 geoprocessing command called "GEOMETRY:HULLRECTANGLE". This command returns a string containing the eight coordinates of the MBR. With these coordinates the RatioL2W and MBRArea field values can be set. The values for the LabelType field are set based on the following pseudo-code logic:

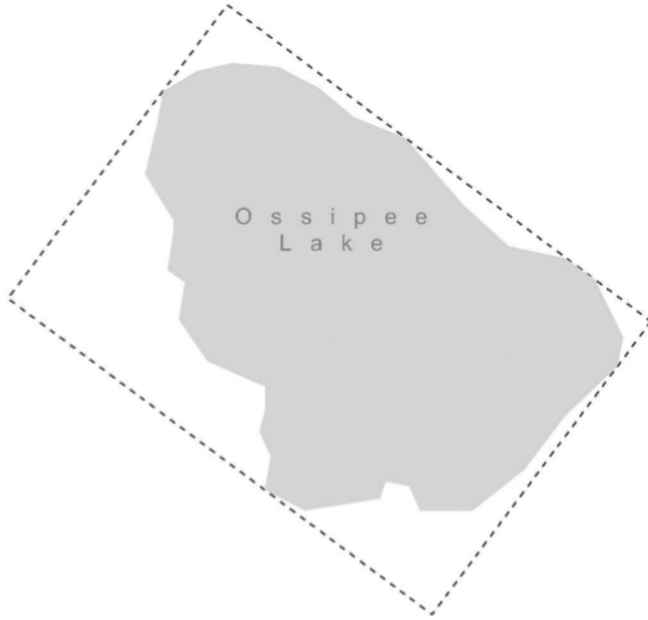

Figure 3. Example of round-ish polygon.

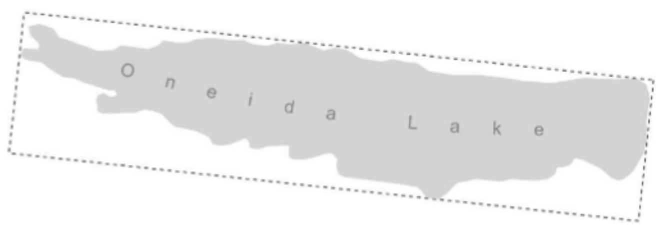

Figure 4. Example of an oblong polygon.

```
If RatioL2W < 4 and MBRArea > 60%
   Label Type = "Roundish"
Elseif RatioL2W < 8 and MBRArea > 25%
   LabelType = "Oblong"
Elseif RatioL2W >= 8 and MBRArea > 10%
   LabelType = "Long"
Elseif RatioL2W >= 8 and MBRArea <= 10%
   LabelType = "Long and Skinny"
Else
   If RatioL2W < 4 and MBRArea >= 20%
      Label type = "Splotch"
   Elseif RatioL2W < 8 and MBRArea > 12%
      Label Type = "Snaky or Pronged"
   Elseif RatioL2W < 8 and MBRArea <=12%
   Label Type = "Snaky or Pronged and Skinny"
```

This logic is essentially first determining whether the shape is round-ish, and if not, if it is oblong or long, and if not, if it is a splotch, or snaky or pronged. The specific thresholds may need to be tuned to specific cartographic requirements.

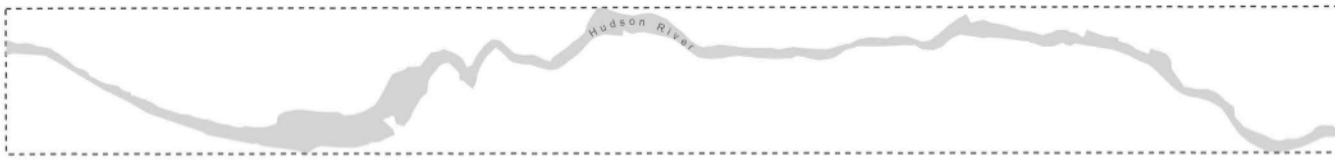

Figure 5. Example of long polygon.

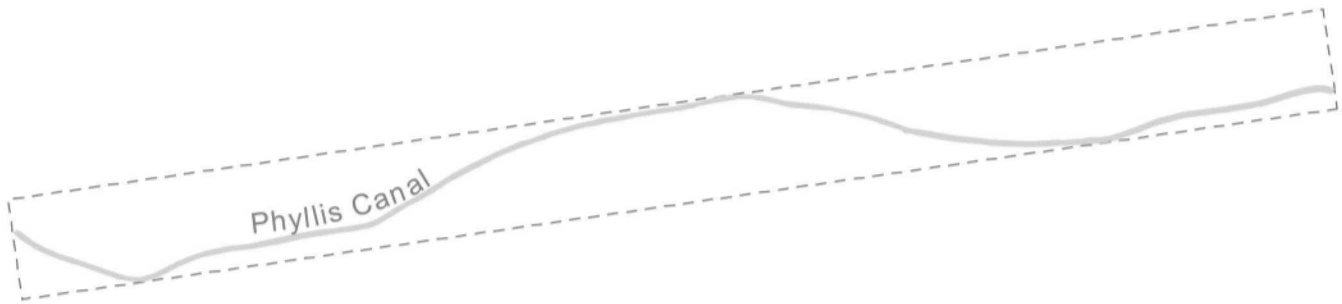

Figure 6. Example of a long and skinny polygon.

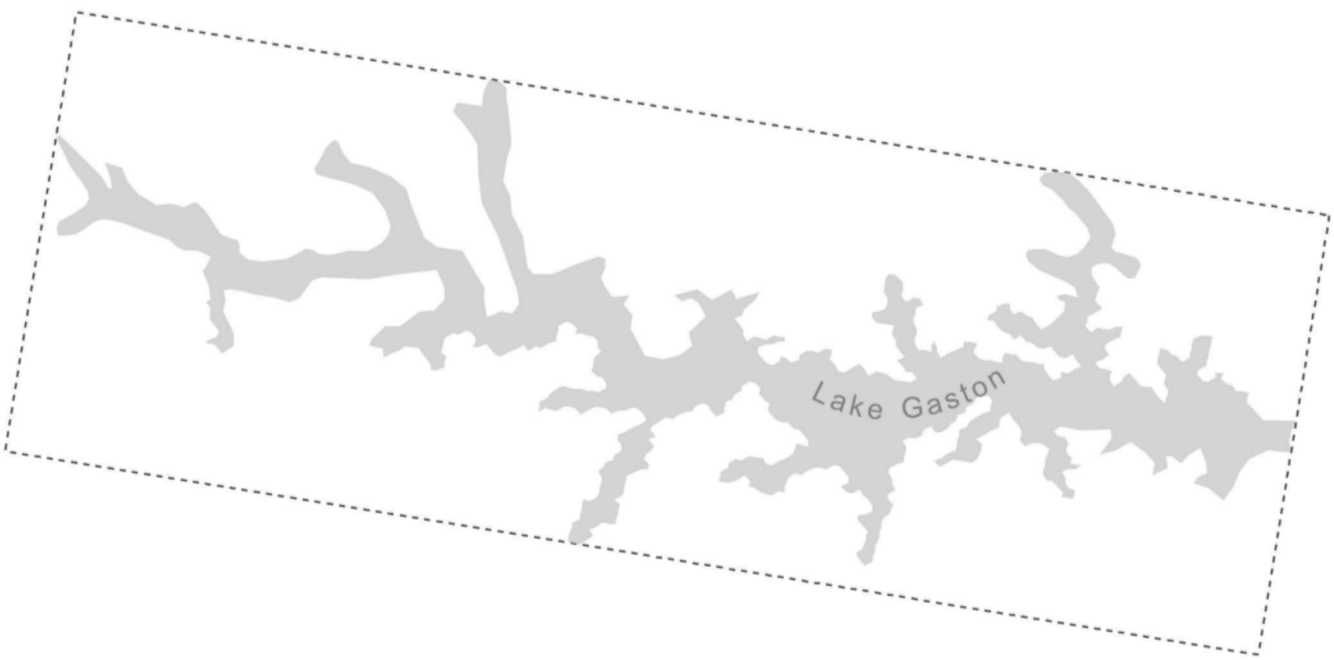

Figure 7. Example of a splotch polygon.

The final field, LabelSize, is based on a binary regression classification. That is, size classes are determined based on the range area between the smallest polygon and the largest polygon. The classes are determined by initially cutting the range in half and the upper portion becomes the first class (for the largest features). Then the lower half is cut in half and its upper portion becomes the next class. This is repeated until the desired set of size classes is codified.

Figure 8. Example of a snaky or pronged polygon.

Figure 9. Example of a snaky or pronged and skinny polygon.

To use the result in ArcMap with Maplex, the polygon data is added as a layer and the following steps are implemented:

1. In the label properties dialog's symbology tab, choose to symbolize with the data using Categories: Unique Values Many Fields option.
2. Choose the LabelType and LabelSize fields and click the "Add All Values" button; then click the OK button to close the layer properties dialog box.
3. Open the Label Manager dialog (on the labeling toolbar) and click on your layer's uppermost line in the Label Classes list.
4. In the "Add label classes from symbology categories" section click the "Add" button. Click yes in the resulting message box to overwrite any existing label classes. (Note at this point you can close the Label Manager dialog and go back and set the layer's symbology to any desired method and symbols.)
5. In the label manager the following Maplex rules can be applied. These rules are described in general for just one size class; typically the size class would dictate the size or size range of the text symbol.

   a. Round-ish
      i. Placement: Curved
      ii. Try Horizontal First = true
      iii. May Stack = true
      iv. Character Spacing = up to 200%
   b. Oblong
      i. Placement: Curved
      ii. May overrun by 36 pts
      iii. Allow asymmetric overrun = true
      iv. Char. Space = up to 200%
      v. Reduce font from 14 pts. to 10 pts. by 1 pt. increments
   c. Long
      i. Placement: Curved
      ii. May overrun by 12 pts
      iii. Char. Space = up to 300%
   d. Long and Skinny
      i. Placement: Boundary
      ii. May Place Outside = true
      iii. Offset = 4 pts
      iv. Char. Space = up to 240%
      v. Background Label = true
   e. Splotch
      i. Placement: Curved
      ii. Char. Space = up to 300%
      iii. Reduce Font from 14 pts. to 10 pts. by 1 pt. increments
   f. Snaky or Pronged
      i. Placement: Curved
      ii. May overrun by 12 pts
      iii. Char. Space = up to 400%
   g. Snaky or Pronged and Skinny
      i. Placement: Boundary
      ii. May Place Outside = true
      iii. Offset = 4 pts
      iv. Char. Space = up to 240%
      v. Background Label = true

The above placement rules often work very well, but in the skinny cases (cases d. and g.), the orientation of the label with respect to the feature on the page does not work well; however, this is the best option currently available.

Finally it is worth noting the distribution of these shape types in terms of their frequency of occurrence in a variety of datasets. Table 1 shows that in each of these datasets there are many shapes in the round-ish category, which require different Maplex rules than the oblong and long shapes, which are also well represented. Thus, by using this method the amount of effort, particularly manual editing of text, is roughly cut in half.

	Data				
Shape Type	Hydro 1M	Physio-graphic	Soils 25K	Hydro 5K	Vegetation 5K
Round-ish	364	278	445	1786	14380
Oblong	592	211	1261	950	870
Long	11	14	8	931	70
Long & Skinny	0	0	0	157	0
Splotch	35	2	30	26	14
Snaky or Pronged	33	7	39	135	17
Snaky or Pronged & Skinny	2	1	2	489	5

Table 1. Feature counts for each type for five data sets. The datasets are: (1) Hydro 1M: Hydrography for 1:1,000,000 of the northeastern United States, (2) Physiographic features of North America, (3) Soils for Ada County, Idaho, (4) Hydro Areas for Ada County, Idaho for 1:5,000 scale maps, and (5) Vegetation for Ada County, Idaho for 1:5,000 scale maps.

CONCLUSIONS & FUTURE WORK

This method of identifying shape types for labeling has worked well in creating general reference maps at scales ranging from 1:5,000 to 1:1,200,000, on natural and built hydrographic polygons, and on physiographic features. This method could also be very useful in labeling vegetation type, soil type, surface geology type, and many other such features. A more complicated adaptation of this method is also being tested as a basis for identifying features to be eliminated or generalized on maps at scales smaller than the data was originally intended. Initial results of this work are quite promising.

In general, the ability to enhance GIS data that were not captured with the intent of creating higher quality cartography in a highly automated fashion is valuable. Many cartographic operations in GIS are conducted by attempting to directly and often simplistically translate GIS features that were captured independently of any cartographic product requirements into a product-specific semantic and graphical context. The result, not surprisingly, is an awkward mix that defies stylistic and semantic expectations. The method described in this article successfully adds additional meaning to the GIS features before attempting a requirements-based transformation into a cartographic solution.

REFERENCES

Abdelmoty, A. and El-Geresy, B., 2004. Schema Visualization Using a Metadata Approach for GIS. *Proceedings of the GIS Research UK 12th Annual Conference. Norwich, UK. 160-162.*

Papadias, D., Sellis, T., Theodoridis, Y., and Egenhofer, M., 1995. Topological Relations in the World of Minimum Bounding Rectangles: A Study with R-trees. *Procedings of the ACM SIGMOD Conference 1995*: San Jose, California, 92-103.

Zhou, X., Krumm-Heller, A., and Gaede, V., 1999. Generalization of Spatial Data for Web Presentation. 2nd Asia Pacific Web Conference (APWeb99). Hong Kong. September.

reviews

Applied Environmental Economics: A GIS Approach to Cost-Benefit Analysis

Ian J. Bateman, Andrew A. Lovett and Julii S. Brainard
2005; Paperback edition, 335 pages, $43
New York: Cambridge University Press
ISBN-13 978-0-521-67158-3

Reviewed by Grace Wong
Advisor, Corridor Economics and Strategies
People, Protected Areas and Conservation Corridors
Conservation International

In a decision-making process, economics plays a role to finding the most efficient and cost-effective solution amongst various options. The efficient solution is typically one where resources will be properly allocated based on their economic value in markets. This very basic economic assumption has proven to be a thorn, particularly in decisions relating to land use and land cover change, as these changes impact the natural landscape and have wide ranging environmental consequences that often cannot be adequately measured nor traded in markets.

The book by Ian Bateman and his colleagues from the University of East Anglia seeks to address this issue by incorporating the non-market environmental values of land use and land cover change into standard cost-benefit analysis (CBA) to support decision-making. In addition, they push the analytical boundaries further by incorporating Geographic Information Systems (GIS) in the analysis to account for spatial and geophysical differences that are likely to impact on those values. This book demonstrates a number of ways that GIS can be employed to improve the way in which real world complexities are incorporated into CBA, thus reducing the need for simplifying assumptions.