Interactive Visualization of Fuzzy Set Operations

Yeseul Park*[a], Jinah Park[a]
[a]Computer Graphics and Visualization Laboratory, Korea Advanced Institute of Science and Technology, 119 Munji-ro, Yusung-gu, Daejeon, 305-732, South Korea

ABSTRACT

Fuzzy sets refer to the datasets which do not have separate, distinct clusters, and they contain data elements whose membership values are between 0 and 1. Often each element belongs to several groups. We design a disk diagram with an augmented frequency graph to visualize elements in a fuzzy set. It shows the distribution of the elements within the set with respect to the membership function value. Fuzzy set operations-like intersection-are used to find the elements of interests with certain degree of confidence. However, few techniques exist for visualization of fuzzy set operations whose process and result provide both the relationship and the distribution of individual data point between sets. With the proposed disk diagram representing a fuzzy set, we suggest an interactive visualization system to analyze fuzzy data and grasp the relationship among the sets. By interactively overlapping the disk diagrams, a user can not only bring out the individual elements of interests, but also overview the whole data elements associated with the sets. Furthermore, we investigated two different visualization scenarios for intuitive interpretation of the visualization results. We demonstrate the proposed visualization system with a terrorist analysis output dataset.

Keywords: visualization, fuzzy data, fuzzy set operation, diagram

1. INTRODUCTION

In many cases of real world problems, the answer is not often binary but in-between yes or no. For example, the concepts of health, illness, disease are difficult to be explained in a classical binary logic. An individual may be healthy and not healthy simultaneously, having membership degrees of 'health'. Also, an individual can be healthy in spite of having a disease so that the concepts of health and disease are not mutually exclusive. This fuzzy nature of the reality can be tagged onto the data of fuzzy logic. Fuzzy sets carry membership function whose value ranges from 0.0 to 1.0 for each element belonging to the sets. Given the fuzzy sets, we can perform standard fuzzy set operations-like intersection or union-to select the elements of interests or to grasp the composition of elements.

The dataset in a fuzzy system has been visualized as an effective method of analyzing data and providing insights in problems. We should visualize this data by satisfying specific requirements of user types as well as common requirements for an effective and generic design of a visualization framework. Specific requirements can be decided by users of fuzzy systems, designers of fuzzy systems and designers of visualization framework [1]. Among common requirements, an interactive exploration into a distribution of data, individual data and the results of fuzzy operations are required [1].

Based on common requirements, if users can see a distribution of data with membership functions in each set and fuzzy set operations, an overview of whole data and the relationship among the sets take part in our decision making. An interactive visualization of fuzzy set operations delivers intuitive interpretations of fuzzy data. We focus on this interactive visualization that provides resulting individual data and whole data elements of fuzzy set operations.

In this paper, we design a disk diagram as interactive visualization for grasping a relationship of fuzzy sets. By our visualization, we provide recognizable single fuzzy membership function of each data in sets. We also visualize fuzzy operators (intersection, union, complementary set) on our diagram by users' direct interactions to overlap or attach. And then, the larger two or more disks of diagram are overlapped, the only higher intersection data are determined, so that the less results of intersection are represented on the disk diagram. While running, fuzzy operations, the causing composition of data elements in each set, are shown.

*ye-seul@kaist.ac.kr; phone +82 42 350 6257

Visualization and Data Analysis 2010, edited by Jinah Park, Ming C. Hao, Pak Chung Wong,
Chaomei Chen, Proc. of SPIE-IS&T Electronic Imaging, SPIE Vol. 7530, 753002

This paper provides a way of data mining for deriving new information in a fuzzy system. At the R&D agenda for visual analytics [2], visual analytics tool is primarily needed that will support analysts and emergency responders to facilitate advanced analytical insights in complex and uncertain datasets. Our visualization aims at this visual analytics for deriving insights from fuzzy data.

The section 3 describes our proposed disk diagram. First, we show one disk diagram for distribution of data. Second, we explain the relationship of sets by moving two disk diagrams.

2. RELATED WORK

As one of researches providing insights in a distribution of fuzzy dataset, Fua, Y. H. et al. focused on the interactive visualization of large multivariate datasets using the parallel coordinates display technique [3]. Parallel coordinates is a technique applying to a various set of multidimensional problems dominant in 1980's. This technique doesn't matter the issue of dimensionality. Moreover, we can recognize the correlations among variables in the dataset. But this technique has difficulty in distinguishing over-plotting large amount of data. Therefore, [3] imposes the amount of clutter using hierarchical organization by statistical aggregation to summarize over-plotting large amount of data. As a scheme for displaying, proximity-based coloring translucency is used.

While [3] displays the color of clusters declining the degree of shading linearly toward the standard deviation of the cluster spread, Berthold, M. R. and Hall, L. O. [4] suggested visualizing data in 2D parallel coordinates. In [4], iris data are represented by two sets, iris-setosa and iris-versicolor, using colored-fuzzy points along each data, sepal length, sepal width, petal length and petal width. Users can see a form of a distribution of each set as a summary by visualizing clusters [4].

For a summary of fuzzy data, Berthold, M. R. and Holve, R. [5] presented an approach to visualize high dimensional sets by multidimensional scaling of sets using a distance function in 2D. Using iris data and breast data, [5] shows their sets with rectangles separated by a calculated distance. It provides a way to visualize whole sets simultaneously and tries to give user interactions offering intelligent data analysis. This is another way of representing clusters using a distance among fuzzy points. While [3] and [4] display clusters with the area of fuzzy points along parallel coordinates to see distributions of data easily, [5] shows separated fuzzy sets to find differences among fuzzy data through their membership functions.

The quality of results of a fuzzy system relies on maintenance of useful information for intermediate stages [1]. It means that the visualization of a fuzzy system is required to prevent loss of useful information when users need to make decisions using visualizations only. To check which datum is important or not, visualizations often have to show individual information of each datum. However, [3, 4, 5] are difficult to find individual membership function of each datum, and users cannot recognize which data have high weights or low weights in the set at the same time. To recognize individual membership function of each datum, Cox, Z. et al. [6] visualized data with membership functions into clusters by c-means clustering with convex hulls as thresholds, glyphs as data points, coloring schemes as membership functions, and 3D plots.

As another method to show single datum with a membership function, Pham, B. and Brown, R. [1] suggested 3D parallel coordinates visualization displaying x-y plane displays as 2D parallel coordinates, and the z-axis displays as a membership function. This technique is similar with [3, 4] as an extension of parallel coordinates, but different at providing individual membership function of each datum. Above studies [1, 6] show individual membership function of each datum as well as a distribution of data on sets. However, those studies do not provide effects of fuzzy operations performed on two or more sets. A fuzzy operation of data means intersection, union, complementary set of a fuzzy logic. Therefore, a fuzzy relationship of sets is derived by the effects of fuzzy operations.

As an effort to visualize fuzzy relationship, Pham, B. and Brown, R. [1] derived new decision by fuzzy operations using a flowchart in a form of truncated double cone structure. In this way, the results of fuzzy relationship can derive new decision. But [1] cannot show both the distribution of data in each set and the effects of fuzzy operations performed on two or more sets.

3. VISUALIZATION OF FUZZY SETS AND OPERATION

3.1 Visual Representation for Fuzzy sets

3.1.1 Visual Features

First, visual features need to be decided to provide fuzzy set information and operations of fuzzy sets. Factors of fuzzy set we will visualize are (1) data as elements of a set, (2) individual membership function of single datum, (3) frequency of data on the same membership function and (4) a set itself, as listed in Table 1. Those are critical factors for visual features to represent fuzzy set and its operations. As visual features, we presented a set by a disk with a specific color for classifying data, and indicated data by dots. The membership function of each datum is drawn with a scalar component, a 'radius' of rings in a disk. The last feature, the number of data of the same membership function is represented by a bar graph.

In order to explain the features specifically, we create three simple fuzzy data sets as shown in Table 2. In Table 2, each Set A, B, and C includes data with the membership functions between 0(0%) and 1(100%). For example, Data 1 belongs to Set A with membership function 15%, Set B with membership function 58%, and Set C with membership function 45% at the same time. Each Set A, B, and C is drawn as a disk with a specific color. Membership function values are mapped onto the disk from its core to the rim in decreasing order from 1.0 to 0.0. Individual data is represented as a dot and is placed within the disk according to its membership function. That is, the membership function is used to calculate the distance from the center of the disk. For example, Data 1 in Set A will be placed 0.85 unit (=1.0-0.15) away from the center of the disk A, and Data 1 in Set B will be placed 0.42 unit away from the center of the disk B. A ring of the computed radius is drawn as a concentric circle on the disk in order to indicate that there is an element belonging to the set with the specific degree of membership. If there are two elements with the same membership value, two dots will be placed random position on the ring. Finally, we also visualize data frequency per each membership function in a disk by a frequency graph as a form of bar chart. The representation of data on a frequency graph will be covered in section 3.1.3. In this paper, we call the disk with rings, dots, and a frequency graph as "disk diagram".

On each disk diagram, the distribution of data in each set is represented. Interactive visualization for operations of fuzzy sets is designed as moving two disk diagrams. Operators of fuzzy sets [7] widely used are standard fuzzy unions, standard fuzzy intersections, and standard fuzzy complements. Therefore, we defined user interactions of moving two disks for union, intersections, and complements. Details on our visualization of fuzzy set operations will be discussed later in section 3.2.

Table 1. Fuzzy data factors and corresponding visual features in a disk diagram

Fuzzy Data Factors	Corresponding visual features in a disk diagram
Data	Dot
Membership function	Ring
Frequency of data	Bar
Set	Disk

Table 2. Simple dataset

Data	Set A	Set B	Set C
Data 1	15%	58%	45%
Data 2	15%	0	0
Data 3	76%	65%	73%
Data 4	23%	47%	22%
Data 5	36%	47%	0
Data 6	43%	67%	35%

Data 7	76%	19%	86%
Data 8	15%	43%	28%

3.1.2 Displaying Data on the Disk

In Figure 1[1], in case of the data which have degree of membership 0.65 to set B of Table 2, a ring of the data is drawn with its radius of 1 - 0.65 = 0.35. If the data which have degree of membership 0.22 exits in set C, 1 – 0.22 = 0.78 is a radius of a ring to be drawn. As the disk is a set of data, the disk is represented by one color. Each ring in one set has same color with the color of the set. To highlight rings comparing with the color of a set, the transparency of rings is set to 255 and the transparency of a set is fixed to 50. In positioning data dots on their rings, the color of dots is equal to the color of their rings, and the direction of the dots is randomly decided.

3.1.3 Frequency Graph

In previous section 3.1.2, we positioned the data with membership function on the ring. To recognize the number of data per ring, a disk is accompanied with a 2D frequency graph. The number of data per ring is presented in y-axis of the graph and the membership functions of data are expressed in x-axis. For example, in Figure 1, the two data of membership function, 0.47, exist in set B and other data are distributed once each. Two data of a ring on the disk is indicated by the height '2' of a bar, while one data is drawn by the height '1' of bar. By this frequency graph in a form of bar chart, numerical distribution of data every ring is displayed at a glance. A frequency graph is attached to each disk by connecting with the line as a bridge between the set and its bar chart. The bridge line connects each ring to each bar in a frequency graph. As effects of fuzzy operations, the results data changes on a disk. Because the updated data of a disk are connected to a frequency graph, the frequency graph is updated, too.

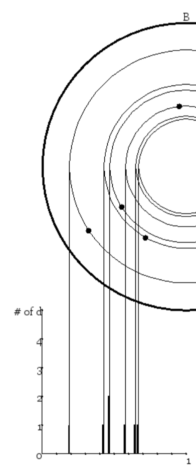

Figure 1. Distribution of data in a set B of disk diagram

3.2 Visual Representation for Fuzzy Set Operations

We mentioned how distribution of data is shown using our disk diagram in previous section 3.1. Now, we focus on how the relationship among fuzzy sets can be expressed with this proposed visualization. To visualize relationship of different fuzzy sets, standard fuzzy set operations are presented in a disk diagram. Standard fuzzy set operations are union, intersection and complement. While disks overlap or touch each other by users' mouse dragging, the result of the set operations changes on our visualization.

3.2.1 Standard Fuzzy Union

Standard fuzzy union is a collection of all data of several sets. In order to visualize union, the maximum membership function 'μ' of the same data 'u' included into different sets A and B is selected by fuzzy set theory, the equation (1) [7].

[1] The colored figures are converted to black & white for printing.

$$\mu_{A\cup B}(u) = \mathrm{Max}\{\mu_A(u), \mu_B(u)\}\ . \tag{1}$$

where A and B are fuzzy sets, and u is data of set A and B.

As a way of representing union, we defined a user interaction of dragging one set to another set for fuzzy union until they come in contact with each other. During union, like Figure 2, the results of union, the whole data included in set A and B, are shown with rings and their dots for overall distribution of union results. A dot as well as a radius of rings can represent individual weight of the union results.

When we represent rings and dots by union, the '$\mu_A(u)$' in the equation (1) becomes a ring on the set A and 'u' in '$\mu_A(u)$' becomes a dot on the set A. Especially, we divide a ring into a colored ring and a white ring. In the set A, the colored ring is derived by '$\mu_{A\square B}(u)$' and the white ring is derived by minimum of '$\mu_A(u)$' and '$\mu_B(u)$'. Therefore, colored dots on colored rings are union data and white dots (open circles) on white rings are called as ghost data in each set. Numeric distribution of union data is shown at the frequency graph of set A and B. Each bar in a frequency graph of set A connects to each colored ring in a set. The number of colored dots on each colored ring indicates the height of a bar in a frequency graph.

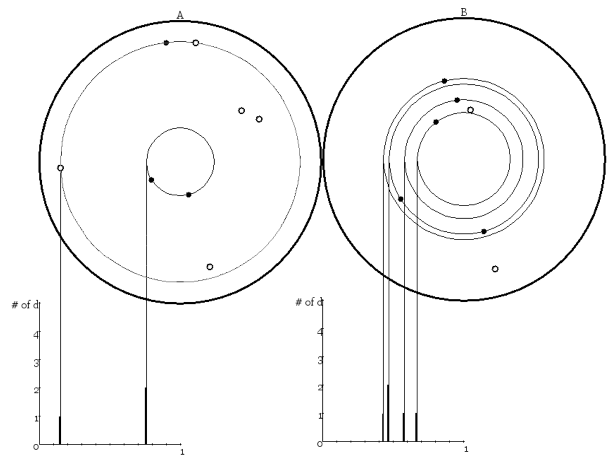

Figure 2. The union data and ghost data of union

3.2.2 Standard Fuzzy Intersection

Standard fuzzy intersection is a collection of all data included in several sets at the same time. For visualizing intersection, the minimum membership function 'μ' among the same data 'u' included into different sets A and B is selected by fuzzy set theory. Equation (2) defines the intersection of fuzzy set theory [7].

$$\mu_{A\cap B}(u) = \mathrm{Min}\{\mu_A(u), \mu_B(u)\}\ . \tag{2}$$

where A and B are fuzzy sets, and u is data of set A and B.

We mapped intersection into overlapping disks using a mouse dragging. If two sets are overlapped by 30% each other by user's mouse interaction, the result data which belong to two sets at least 30% at the same time are shown with each ring and each result dot on its ring. This overlapped percentage is used as a threshold for presenting the intersection data over the threshold.

The '$\mu_A(u)$' in the equation (2) is drawn by a ring on the set A and 'u' in '$\mu_A(u)$' is drawn by a dot on the set A. Like in union, a ring is divided into a colored ring and a white ring, as well as a dot is divided into a colored dot and a white dot. In the set A, the colored ring means '$\mu_{A\cap B}(u)$' and the white ring means the maximum of '$\mu_A(u)$' and '$\mu_B(u)$'. Colored dots on colored rings represent intersection data and white dots on white rings are called as ghost data except for intersection data in each set. The frequency graph of set A and B shows numeric distribution of intersection data. Each bar in a frequency graph of set A is coupled to each colored ring in a disk. The number of colored dots on each colored ring denotes the height of a bar in a frequency graph.

During the intersection, the colored rings of intersection data and the white rings of ghost data are shown at disks. The colored dots on colored rings and the white dots on white rings are represented simultaneously. And then, as changing the threshold by overlapping disks, only white dots and colored dots over the threshold are shown. Therefore, by changing overlapped percentage of disk, larger membership functions of intersection data than the threshold changes the result of the frequency graph. For example, Figure 3 shows the results of overlapped percentage 14%.

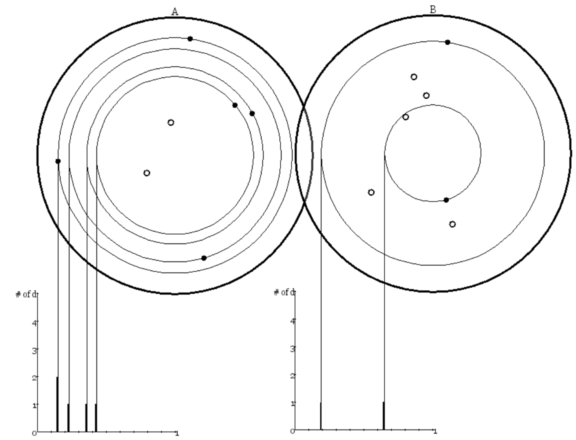

Figure 3. The intersection data and ghost data of intersection (Overlapped: 14%)

3.2.3 Standard Fuzzy Complement

Standard fuzzy complementary set is data not included in a specific set. Complementary set $\bar{A}$ has opposite meaning of correlation between data 'u' and its set A. Equation (3) defines the complement of fuzzy set theory [7].

$$\mu_{\bar{A}}(u) = 1 - \mu_A(u) . \tag{3}$$

where A is a fuzzy set, and u is data of set A.

With an updated membership function '$\mu_{\bar{A}}(u)$', the new ring is updated in set A as well as the new data dots 'u' in '$\mu_{\bar{A}}(u)$' on the new ring.

The user interaction of complementary set is when pressing mouse right button on the sets of which complementary membership need to be known. Its visualization is newly drawing of updated rings, dots and frequency graph like Figure 4. This way makes it possible to carry out the visual operation directly for recognizing how a specific set has its complementary set in universal set U. When a user need to release complementary set, pressing mouse middle button makes the complementary set turn to original membership function. Complementary set of a specific set A derives new bar heights in frequency graph.

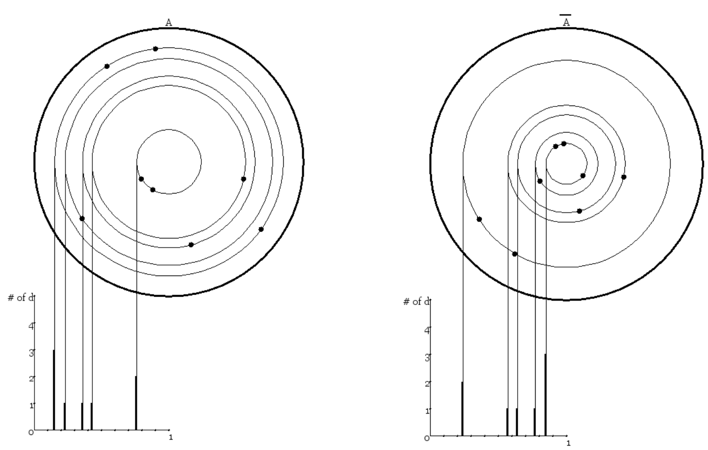

Figure 4. Fuzzy set A and its complement $\bar{A}$

4. SCENARIO AND ITS ANALYSIS

When we apply our disk diagram to the specific dataset, Table 2, the analysis of cases on the visualization is required. To visualize important data at the cases, we provide two scenarios according to which need-to-know results are represented. And then, we analyze results of our scenarios to derive new information from meanings of the results. At first scenario, results of fuzzy set operations are positioned at the set where the results are included. Therefore, results of union data and intersection data are shown at their set. At second scenario, the significant data of fuzzy set operations are positioned at the set where the significant data are included. For analyzing these two scenarios, we use simple dataset, Table 2.

4.1 Scenario 1: When results of fuzzy set operations are positioned at the set of the result data

4.1.1 Representation

4.1.1.1 Intersection

Intersection data are results of lower membership function between two sets A and B. If the intersection data are placed at their set, the intersection is represented like Figure 3 with following data.

Set A = intersection data: {(d1, 15), (d4, 23), (d5, 36), (d6, 43), (d8, 15)} ghost data: {(d3, 76), (d7, 76)}
Set B = intersection data: {(d3, 65), (d7, 19)} ghost data: {(d1, 58), (d4, 47), (d5, 47), (d6, 67), (d8, 43)}

4.1.1.2 Union

Union data are results of higher membership function between two sets A and B. If the union data are placed at their set, the union is represented like Figure 2 with following data.

Set A = union data: {(d2, 15), (d3, 76), (d7, 76)} ghost data: {(d1, 15), (d4, 23), (d5, 36), (d6, 43), (d8, 15)}
Set B = union data: {(d1, 58), (d4, 47), (d5, 47), (d6, 67), (d8, 43)} ghost data: {(d2, 0), (d3, 65), (d7, 19)}

4.1.2 Analysis on Transition of Union and Intersection

During the intersection, the results are intersection data of A and B while the ghost data in A and B are the maximum data. If the disks transit into union, intersection data of A and B become the ghost data in union. The ghost data of A and B in intersection turn into the union data. The data from intersection to union are reversed. Additional data exclusive of the reversed data in each set are the data included only in each set. During the union, the union data in A and B are the results when the ghost data in A and B are the minimum data. On the conditions above, scenario 1 can make trends for specific cases of results like the following lists.

(1) Intersection data gathered in one set A: $A \subset B$

(2) Intersection data positioned near the edge of a set A: The intersection data related to the set A are generally light.

(3) Intersection data placed near the core of a set A: These data are much related to the set A and they are much more related to the set B. Therefore, the intersection data of A and B have a large relationship with A, and data in A are strongly associated with B.

(4) Union data gathered in one set A: $B \subset A$

(5) Union data positioned near the edge of a set A: These data are related lightly to the set A, and they are much lightly related to the set B. Therefore, the union data of A and B have a low relationship with A, and data in A are weakly associated with B.

(6) Union data placed near the core of a set A: The union data generally have a high relationship with A.

4.2 Scenario 2: When the significant data of fuzzy set operations are positioned at the set of the significant data

4.2.1 Representation

4.2.1.1 Intersection

The significant data of intersection are compared data to intersection data between two sets A and B in Table 2. This compared data has a higher membership function than intersection data because the intersection data has minimum membership function among A and B. If the significant data are placed at their set, intersection operation is represented like Figure 5 with following data.

Set A = significant data: {(d3, 76), (d7, 76)} ghost data: {(d1, 15), (d4, 23), (d5, 36), (d6, 43), (d8, 15)}
Set B = significant data: {(d1, 58), (d4, 47), (d5, 47), (d6, 67), (d8, 43)} ghost data: {(d3, 65), (d7, 19)}

4.2.1.2 Union

The significant data are the data of higher membership function between two sets A and B. Therefore, significant data in union are the same with union data. If the significant data in union are placed at their set, union operation is represented like Figure 6 with following data.

Set A = significant data: {(d2, 15), (d3, 76), (d7, 76)} ghost data: {(d1, 15), (d4, 23), (d5, 36), (d6, 43), (d8, 15)}
Set B = significant data: {(d1, 58), (d4, 47), (d5, 47), (d6, 67), (d8, 43)} ghost data: {(d2, 0), (d3, 65), (d7, 19)}

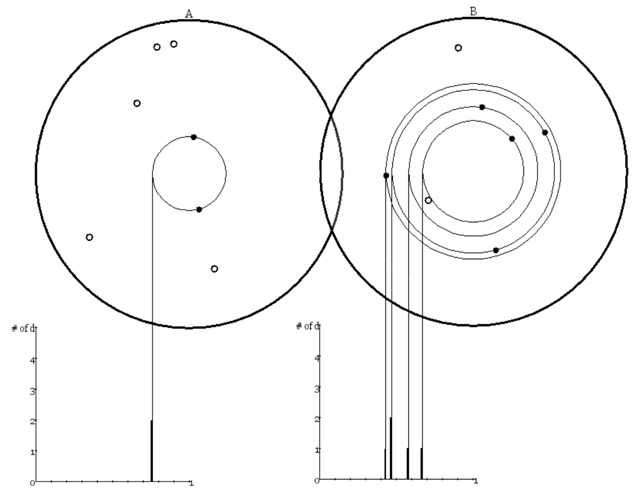

Figure 5. The significant data and ghost data of intersection

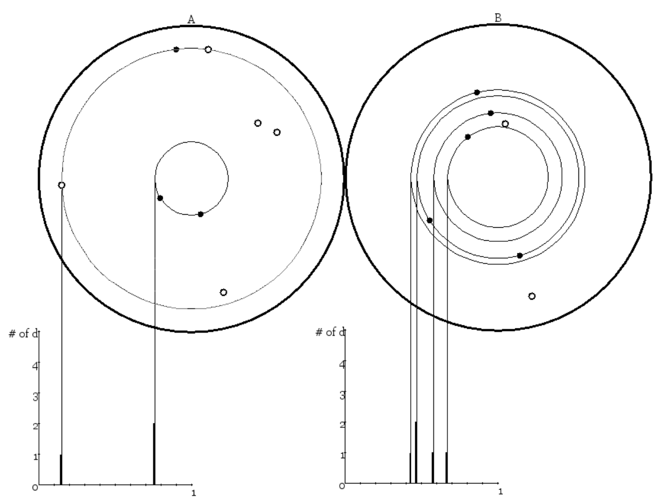

Figure 6. The significant data and ghost data of union

4.2.2 Analysis on Transition of Union and Intersection

In the representation of union, significant data in A and B are union when the ghost data in A and B are the minimum data. The difference between union and intersection is the data of each set only seen at union. These are included only in the set. During the intersection, significant data in A are the elements more related to A than B as ghost data of A and B are the intersection. By the explanation of scenario 2 above, trends of the results are differently analyzed from scenario 1. The following lists are the analysis about a trend of results on intersection and union set A and B in Table 2.

(1) On union, significant data gathered in one set A: $B \subset A$

(2) On union, significant data positioned near the edge of a set A: The data related to the set A are generally light, and they are much lightly related to the set B. Therefore, the union of A and B have a low relationship with A, and data in A are weakly associated with B.

(3) On union, significant data placed near the core of a set A: The significant data of union generally have a high relationship with A.

(4) On intersection, significant data gathered in one set A: $B \subset A$

(5) On intersection, significant data positioned near the edge of a set A: The data are not the intersection data of the set A, but they are more related to the set A than B. If the data are near the edge of A, they are lightly related to A and much lightly related to B. Therefore, these data of A are strongly associated with B.

(6) On intersection, significant data placed near the core of a set A: The significant data of intersection generally have a high relationship with A.

5. APPLICATION TO WORDS SET

Now, we provide an application to our scenarios using disk diagram. Section 4 describes two scenarios and their analysis using simple dataset. In this section, we use words set as a dataset. Fuzzy set operations, especially intersection and union, are covered with scenario 1 and 2 at section 5.2. In section 5.2, we compare scenario 1 and 2 by fuzzy set operation: intersection and union of two sets of words.

5.1 Dataset

The disk diagram visualization for fuzzy set operations is possible to incorporate into the dataset. ‘Many Eyes’ [8] provides dataset and lists of visualizations which represents the dataset. As fuzzy data, the neural network dataset has semantic weight between each word in a document. Especially using semantic dataset of 311 words [8] for a terror analysis output, we illustrate the relationship between words set by applying our disk diagram. The 3 features: a topic, each word in a document, a weight between a topic and each word are possible to match with a set, data of a set and a membership function. The words set chosen is composed up of five sets: Terrorist, Antiterrorism, Treatment, Spain, Al qaeda and they are topics of words. 16 words are semantically related to Antiterrorism, 148 words to Spain, 45 words to Terrorist, 53 words to Treatment, 49 words to Al qaeda. In the topic Terrorist, Terrorist is semantically related to the word such as ‘detention’ with its weight 85%, ‘custody’ with its weight 36%. These weights called membership functions of the data mean how the data is semantically related to their set. Using this semantic dataset of terror related word, our disk diagram provide the results of fuzzy set operations. Result data are different by using different scenario. Therefore, we will compare scenario 1 and 2 using words set to know the different meanings, in section 5.2.

5.2 Comparison between Scenario 1 and 2

Table 3. Comparison between scenario 1 and 2 by fuzzy set operations of the set ‘Terrorist’ and the set ‘Spain’

Operation	Scenario 1		Operation	Scenario 2	
Intersection	Terrorist	There is no intersection data of Terrorist.	Intersection	Terrorist	The significant data of intersection are only in Terrorist.
	Spain	All intersection data are in Spain.		Spain	The significant data are not in Spain, so that the intersection is not significant at Spain.
Transition of Intersection to Union	Terrorist	The ghost data of intersection in Terrorist are reversed to the union data of Terrorist.	Transition of Intersection to Union	Terrorist	All significant data of intersection in Terrorist belong to significant data of union in Terrorist.
		In addition to the reversed data, the remaining data of union are included only Terrorist.			Additionally, data which are included into only Terrorist are represented.

	Spain	The ghost data of intersection in Spain are reversed to the union data of Spain.		Spain	All significant data of intersection in Spain belong to significant data of union in Spain.
		In addition to the reversed data, the remaining data of union belong to only Spain.			Additionally, data which belong to only Spain are shown.
Union	Terrorist	There is no ghost data of Terrorist, so that Terrorist has much related data than Spain.	Union	Terrorist	There is no ghost data of Terrorist, so that Terrorist has much significant data than Spain.
	Spain	The ghost data of Spain mean that the data are less related at Spain than Terrorist. The union data in Spain are included into only Spain. All data (union and ghost) have low relationship of Spain.		Spain	The ghost data of Spain mean that the data are less significant at Spain than Terrorist. The union data in Spain are included into only Spain. Spain has low significant data.

Table 3 compares the representations of two scenarios of intersection, union and transition of intersection to union. Scenario 1 shows where the result data of intersection and union are positioned. Scenario 2 represents in which the significant data of intersection and union are positioned. That is, scenario 1 can deliver result membership function of intersection and union, but scenario 2 doesn't show the results of intersection with colored rings and dots. Instead, scenario 2 is intuitive to see the significant data in intersection and union. Both scenario 1 and 2 deliver how strongly data of Terrorist are related to Spain. The user can grasp the grouping of the words that belong to only Spain and Terrorist or to both sets. The following Figure 7, 8 are the results on two scenarios by intersection of the set 'Terrorist' and the set 'Spain' and Figure 9 is the same result of two scenarios by union.

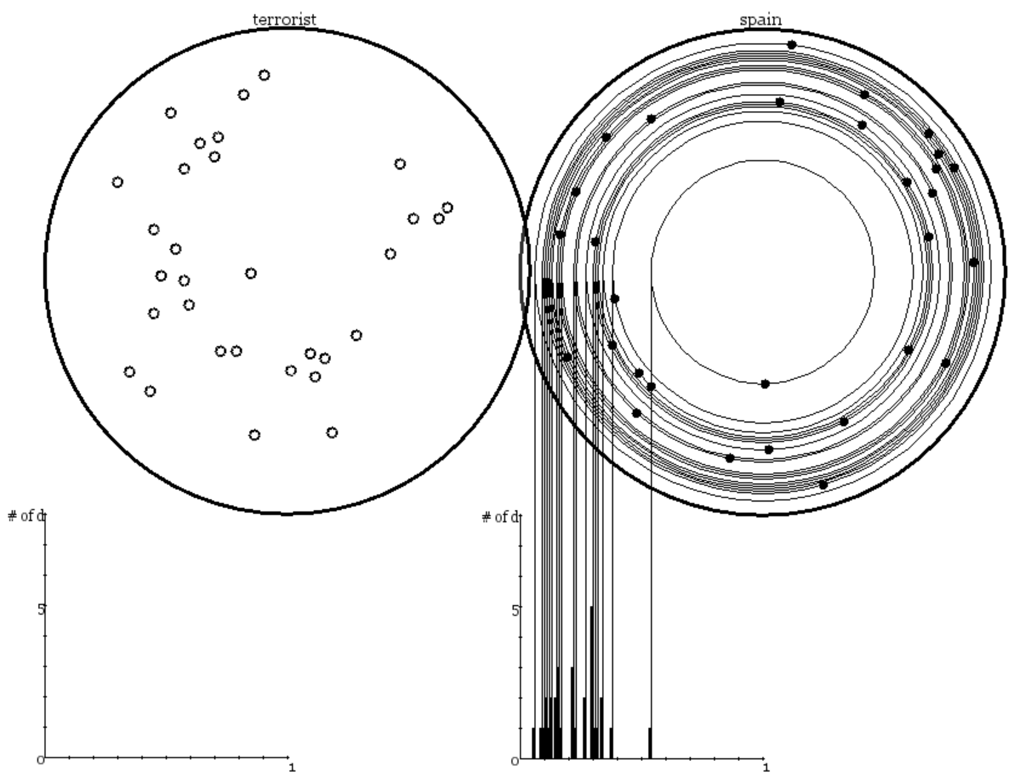

Figure 7. Intersection of scenario 1

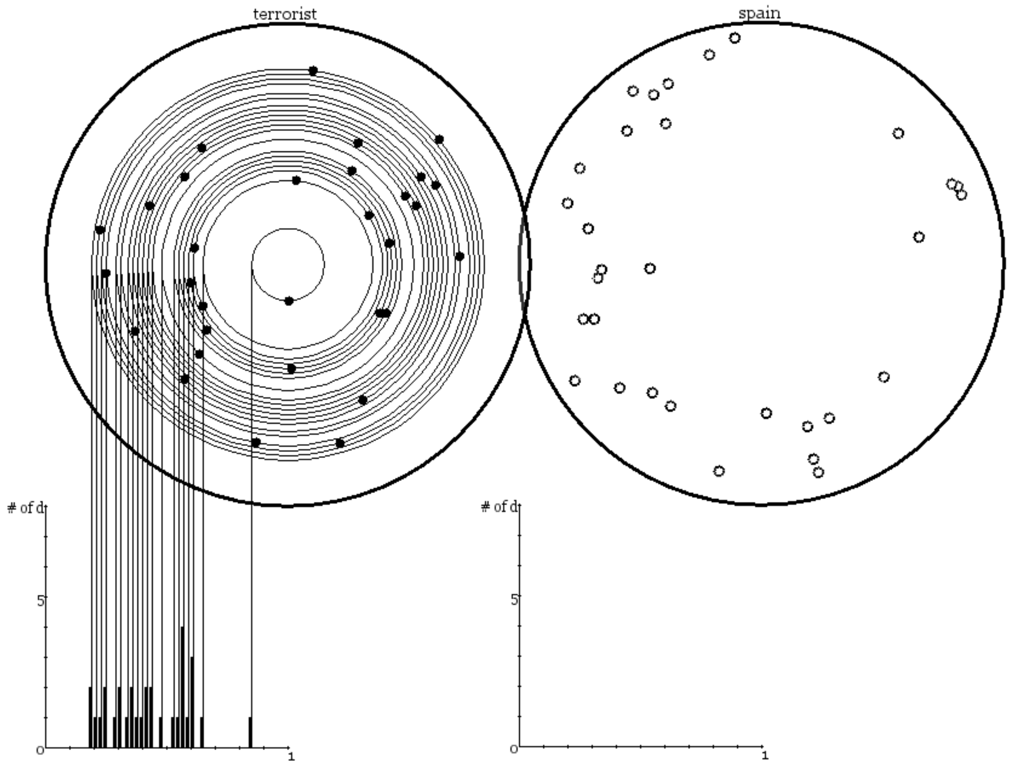

Figure 8. Intersection of scenario 2

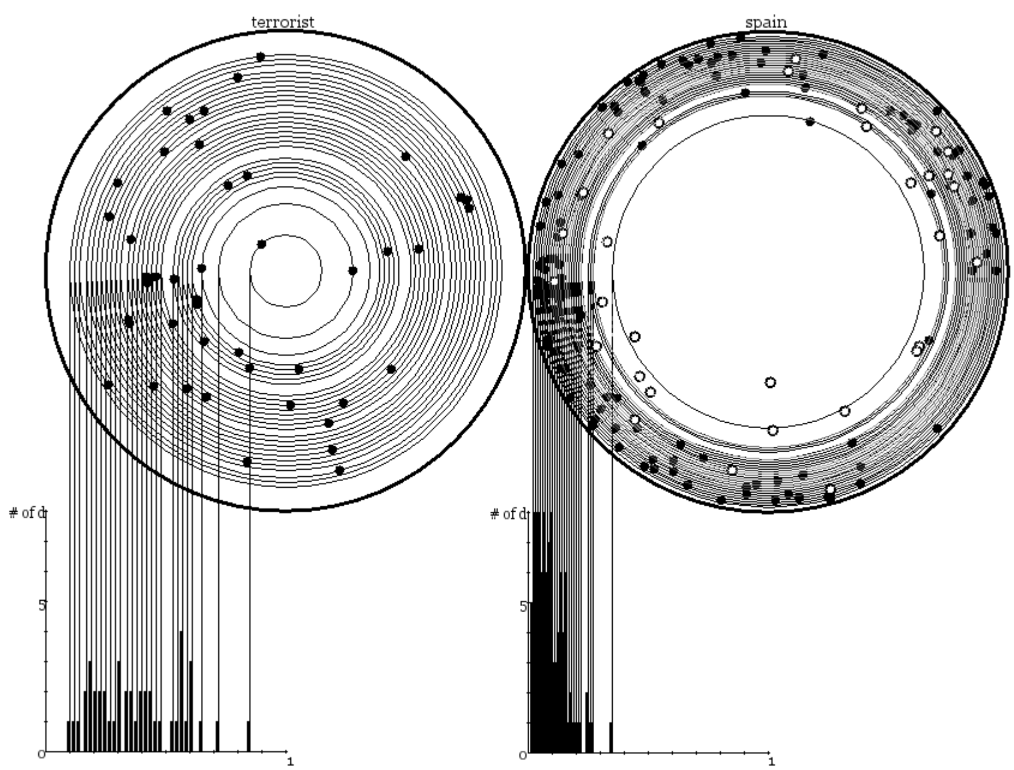

Figure 9. Union of scenario 1 and 2

6. IMPLEMENTATION

This system is implemented in JAVA based environment “Processing”, one of effective open source programming language for visualization to express visual context, developed in MIT [9].

7. CONCLUSION

We present new interactive visualization to grasp the relationship of fuzzy sets as well as to show distribution of membership functions in all sets. A fuzzy set is represented as a disk diagram while users interact with disk diagram for fuzzy intersection, union, complement. Then, we present two different scenarios for intuitive interpretation of visualization results. By applying each scenario to our disk diagram, we demonstrated how the disk diagram can be utilized to understand fuzzy set using word data. First, we visualize the distribution of words belonging to a particular topic set with respect to their membership function values. This provides a user an insight how much the topic is associated with the word data. Second, we visualize the standard fuzzy set operations by interactively moving the disk diagrams. In this process, the user can grasp the grouping of the words that belong to only one set or to both sets. Moreover, the user can interactively pick up more influential data during the process. Finally, the visualization results provide a user a new insight on how the fuzzy sets or topics are related each other. In our example, we have observed one topic (‘spain’) is subordinate to another topic (‘terrorist’). As another application, this visualization can be applied for web search engine that gives the relationship of web documents with membership function to users’ query. By user interaction of reaching or overlapping disks, users can find how the result web documents of query A and query B are related. To sum up, these applications give trends of all data and relationships among datasets.

The objective of our visualization is to find insight of a complex problem in the long run. As a future work, we need to advance our diagram which derives important information to users with suitable domains. Furthermore, user evaluation with experts on those domains is required whether the results are found to be important and intuitive when the experts use our visualization model. Accordingly, the heuristic evaluation of Jacob Nielson will be considered [10] for criteria of visual features.

ACKNOWLEDGMENTS

The Korea Science and Engineering Foundation (KOSEF) grant funded by the Korea government (MEST R01-2008-000-11678-0).

REFERENCES

1. Pham, B. and Brown, R., "Visualization of Fuzzy Systems: Requirements, Techniques and Framework", Future Generation Computer Systems 21(7), 1199-1212 (2005).
2. Thomas, J. J. and Cook, K. A., [Illuminating the Path: The Research and Development Agenda for Visual Analytics], IEEE Computer Society, USA, (2005).
3. Fua, Y. H., Ward, M. O. and Rundersteiner, E. A., “Hierarchical Parallel Coordinates for Exploration of Large Datasets”, Proc. IEEE Visualization, 43-50 (1999).
4. Berthold, M. R. and Hall, L. O., "Visualizing Fuzzy Points in Parallel Coordinates", IEEE Transactions on Fuzzy Systems 11(3), (2003).
5. Berthold, M. R. and Holve, R., “Visualizing high dimensional fuzzy rules”, Proc. Fuzzy Information Processing Society NAFIPS. 19th International Conference of the North American, Dept. of Electr. Eng. & Comput. Sci., California Univ., Berkeley, CA, USA, 64-68 (2000).
6. Cox, Z., Dickerson, J. A. and Cook, D., “Visualizing Membership in Multiple Clusters after Fuzzy C-means Clustering”, Proc. Visual Data Exploration and Analysis VIII, 60-68 (2001).
7. Zadeh, L. A., “Fuzzy sets”, Information Control 8(3), 338-353 (1965).
8. http://manyeyes.alphaworks.ibm.com/manyeyes/datasets/neural-network-ta
9. Reas, C. and Fry, B., [Processing: a programming handbook for visual designers and artists], The MIT Press, Cambridge, USA, (2007).
10. Nielson, J., [Usability Engineering], Morgan Kaufmann, San Francisco, USA, (1993).