2014 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)
July 6-11, 2014, Beijing, China

Determining Interval Type-2 Fuzzy Set Models for Words Using Data Collected From One Subject: *Person FOUs*

Jerry M. Mendel
Signal and Image Processing Institute
University of Southern California
Los Angeles, CA 90089-2564, USA
mendel@sipi.usc.edu

Dongrui Wu
Machine Learning Laboratory
GE Global Research
Niskayuna, NY 12309, USA
drwu09@gmail.com

***Abstract*—This paper provides a new methodology for determining a word's interval type-2 fuzzy set model using *only one* subject, a Person FOU. It uses *interval end-point uncertainty intervals* instead of only the end-point intervals. Such uncertainty intervals are relatively easy to collect and they do not introduce methodological uncertainties during the data-collection process. This new method is applied to ten probability words. Person FOUs are obtained for these words, and the robustness of this new method to the choice of the probability distribution that is assigned to the interval end-point uncertainty intervals is examined and demonstrated.**

I. Introduction

IN a computing with words (CWW) product, one should use fuzzy set (FS) models for words that are commensurate with the linguistic uncertainty of its end-users. As is stated in [7]:

> We believe that fuzzy set models for words must be derived from data that are collected from a group of subjects. … Because words mean different things to different people, and so are uncertain, a fuzzy set model is needed for a word that has the potential to capture its uncertainties, and that an interval type-2 fuzzy set[1] (IT2 FS) should be used as a FS model of a word, because it is characterized by its footprint of uncertainty (FOU) and, therefore, has the potential to capture word uncertainties.

Recall that uncertainty about a word is of two kinds [5], [7], namely: (1) *intra-uncertainty*, the uncertainty a person has about the word; and (2) *inter-uncertainty*, the uncertainty that a group of people have about the word. For CWW products that are used by more than one person the FS model for a word incorporates both kinds of uncertainty; however, when the CWW product is a *Personal Advisor* (PA), the FS model for a word only needs to incorporate the uncertainty that the user of the PA has about the word (his/her intra-uncertainty). For example, in the *love-matching problem* that is considered in [3], a male subject wants only to use his knowledge and opinions to select a partner and does not care how others evaluate the qualities of candidate partners. One of the criteria that he evaluates a woman on is beauty. It is obvious that *beauty is in the eyes of the beholder* (a good looking person may be homely in someone else's eyes); hence, only his uncertainty about the terms that are used to describe beauty (the terms used are Hideous, Ugly, Homely, Normal, Good Looking, Pretty, Gorgeous) matter to him.

Methods have already been developed for collecting data from a group of subjects and then mapping those data into a word's interval type-2 fuzzy set model (i.e., its footprint of uncertainty (FOU)). Although one method has been published for trying to do the same for a single subject, it is, in our opinion, not a successful method (as will be explained in Section II.B).

The purpose of this paper is to provide a new methodology for determining a word's FOU using only one subject, the result being a *Person FOU*.

The rest of this paper is organized as follows: Section II provides the background for the rest of the paper; Section III provides our new methodology for obtaining a Person FOU from data that are collected from just one subject; Section IV describes the words that are used in the rest of the paper as well as the data that were collected from ten subjects; Section V explains different ways in which the word data were processed and provides many FOU results; and, Section VI draws conclusions and makes some suggestions for future works.

II. Background

This section briefly reviews the *Enhanced Interval Approach* (EIA) [11] for mapping a set of n data intervals about a word, that are collected from a group of n subjects, into an IT2 FS model, and the *Individual Interval Approach* [2] for mapping data from a single subject into an IT2 FS model. The EIA is reviewed because it is used in Section III to obtain a Person FOU. The latter is reviewed for historical reasons.

[1] IT2 FSs whose FOUs have no gaps (and never have had) are equivalent to interval-valued fuzzy sets (e.g., Bustince [1]).

978-1-4799-2072-3/14/$31.00 ©2014 IEEE

A. IT2 FS Word Models Computed by Using Data Collected From a Group of Subject: the EIA

We continue to believe [7] it is very important that methods for collecting data from a group of subjects or even from an individual should *not introduce methodological uncertainties* into the data collection procedure. Most people do not know what a fuzzy set is, and so a method that asks an individual to provide a membership function (MF) for a word has methodological uncertainty associated with it that becomes co-mingled with the word uncertainty, and the two kinds of uncertainty cannot be separated. Consequently, *we do not advocate asking subjects to provide MFs*.

The *Enhanced Interval Approach* (EIA) [11] is a multi-step procedure for mapping a set of n data intervals about a word, that are collected from a group of n subjects, into an IT2 FS model. It is an enhanced version of the *Interval Approach* (IA) [4], and replaces the IA.

The data are obtained by means of surveys in response to questions such as: *On a scale of 0–10, what are the endpoints of an interval that you associate with the word?* Of course, if the variable has an actual physical scale that is associated with it (e.g., [*l, r*]), that scale is used. We have administered many data collection surveys that use such simple questions and have found that everyone seems to be able to easily answer them.

The EIA has two parts, *Data Part* and *Fuzzy Set Part*. The *Data Part* uses statistics and probability to:

1) Remove bad data
2) Remove outliers
3) Keep only the data intervals that are within an acceptable tolerance limit
4) Remove data intervals that have no overlap or too little overlap with other data intervals (thereby also enforcing the maxim that *words must mean similar things to different people* for effective communication to occur)
5) Assign a uniform distribution to the remaining data intervals and then compute the mean and standard deviation of those intervals

At the end of the Data Part, the original n data intervals have been reduced to a set of m (surviving) data intervals, where $m \le n$.

The *Fuzzy Set Part* uses these m data intervals to:

6) Determine whether a word should be modeled as either a left-shoulder, interior or right-shoulder FOU by using a classification procedure that makes use of group statistics for all m data intervals (this is done by use of data and not a priori)
7) Map each of the m data intervals into the parameters of a T1 FS by equating the mean and standard deviation of that data interval to the mean and standard deviation of the word's corresponding T1 FS—the resulting T1 FS is called an *embedded T1 FS*
8) Delete those embedded T1 FSs whose end-points fall outside of the 0–10 scale, reducing the number of embedded T1 FSs from m to m^*
9) Construct lower and upper bounds of the m^* embedded T1 FSs that become the lower and upper MFs, respectively, of the word's FOU

Note that in [11] Wu, et al. have demonstrated that mean-square convergence of an EIA FOU occurs when data are collected from at least 30 subjects.

Software for implementing the EIA is available at: http://sipi.usc.edu/~mendel (Publications/Software/Type-2 Fuzzy Logic Software—"EIA" folder); or, at https://sites.google.com/site/drwu09/publications/EIA.zip

B. Individual Interval Approach (IIA)

Joo and Mendel [2] introduced an approach for determining a word's FOU from a single subject, called the *Individual Interval Approach (IIA)*, in which the following is stated:

> Because data are collected from a single subject in the IIA, more data are needed from the subjects than are needed in the EIA[2]. Instead of asking a subject to provide only the two end-points of an interval that they associate with a word, as is done in the EIA, a subject is now asked to provide four data, namely two interval end-points and two end-point uncertainty bands. … Uncertainty bands are in percentage and indicate how uncertain one is about the end-points, e.g., a 20% left uncertainty band indicates that one is 20% uncertain, both to the left and the right of the left end point.

Although this data collection procedure seemed very promising, the rest of the IIA is not a very successful method for obtaining a Person FOU because it creates only nine (somewhat contrived) intervals that are related to the collected data, and then maps only those intervals into T1 FSs that are finally bounded from above and below to obtain a word's FOU. Nine is not enough, as the FOU mean-square convergence studies have demonstrated in [11]. In addition, the IIA does not use the interval end-points and the end-point uncertainty bands that are collected from a subject *simultaneously*; and, it relies too heavily on the asymptotic classification diagram that is used in Step 6 of the EIA. The result of all of this is that many times the IIA leads to very strange looking FOUs.

Not using the four data simultaneously seems like an unsatisfactory way to obtain a Person FOU. In this paper, we show how to obtain a Person FOU using all four data simultaneously.

III. New Methodology

Our new methodology for obtaining a word FOU from a single subject, a Person FOU, is to:

1) Establish end-point intervals for a word, by asking a subject the following two similar questions: *Suppose that a word can be located on a scale of l to r, and you want to locate the end-points of the interval that you associate with the word on that scale, but you are unsure of these two end-points: (1)[(2)] On the scale of l to r, what are the endpoints of an* ***interval*** *of numbers that you associate with the left [right] end-point of the word? (If you are absolutely certain about the two end-points you do not have to provide end-point intervals,*

[2] [2] refers only to the IA; however, since the EIA (published three years after [2]) has supplanted the IA we have replaced "IA" by "EIA" in this quote.

you only have to provide the left and right end-points. Do not overlap the left and right end-point intervals.)[3]

2) Assume that each of the end-point intervals is uniformly distributed, and then compute the mean and variance for both of them.
3) Assign the mean and variance of the left and right intervals from Step 2 to probability distributions and generate 100 random numbers
$$(L_1, L_2, ..., L_{50}; R_1, R_2, ..., R_{50}) .$$
Form 50 end-point pairs from these random numbers[4] $\{(L_1, R_1), ..., (L_{50}, R_{50})\}$.
4) Assume each pair of end-points has been collected from a different subject (or the same subject who is sampled 50 times, where the spacing of the samples is long enough so that the subject does not remember his/her past responses).
5) Apply the EIA to the 50 intervals to obtain the Person FOU for the word.

Some comments are given next about Steps 2-5.

Step 2: Without further information from a subject, there is no reason to support a belief that the subject applies more weight to some portions of an uncertainty interval than to others when they are asked to provide the end-points of an uncertainty interval; hence, it seems quite reasonable (to us) to assume that the two uncertainty intervals are uniformly distributed (more about this in Step 3). The Step 2 distribution is solely for the purpose of computing the mean and variance for each of the end-point intervals. If one can establish an actual distribution of a subject's end-point intervals then that distribution would be used to compute the mean and variance in Step 2. In order to do this, additional information would have to be extracted from the subject.

Step 3: To see the effects of different distributions on the final Person FOUs, we ran all of our processing for uniform, normal and triangle distributions (see Section V.B); however, they all had the same Step 2 first and second-order moments. We chose to use 50 intervals because the convergence results given in [11] have demonstrated that FOUs converge in a mean-square sense when 30 or more intervals are used. Any number ≥ 30 should be adequate. The results from this step are a set of 50 independent data intervals that are specified by their left and right end-points.

Step 4: This is a conceptual step that is preparatory to using the EIA in Step 5, because the EIA is formulated using intervals of data that are collected from a group of subjects.

Step 5: This is done using the existing EIA as described in Section II.

We wish to point out that Steps 1-4 reduce the computation of a Person FOU to that of computing a group's FOU. By these steps we are able to again use the EIA; hence this paper can be viewed as an application paper.

[3] An illustrative example usually accompanies these questions, e.g. for the word *Pretty woman* the two end-point intervals on the scale 0-10 might be [7, 7.7] and [8.2, 8.7].

[4] It is also possible to match L_i to R_j $(i \neq j)$. The final results may be differ when the number of pairs is small, but should not be very different for 50 pairs.

IV. Words and Data

Wallsten and Budescu [10] focus on a variable that they called *qualitative probability expression*, whose terms are: *Almost Impossible*, *Improbable*, *Doubtful*, *Unlikely*, *Tossup*, *Possible*, *Likely*, *Probable*, *Good Chance* and *Almost Certain*. These words have a natural scale of [0,1]. We collected data from 10 subjects for these 10 words on that scale. So that others may repeat our experiments, we provide this raw data in Table A.1 in Appendix A. Although all of our results that are described in Section V are for these 10 words, it is straightforward to redo all of our simulations for other vocabularies.

V. Processing the Data

In this section we explain different ways in which the Table A.1 data were processed and provide many Person FOU results. Recall that in Step 3 of our new methodology one must choose a probability distribution for the left and right end-point intervals. To begin, we describe results when those intervals were assumed uniformly distributed. Then we examine the robustness of the resulting Person FOUs to that assumption by also showing results for normal and triangular distributions.

A. End-Point Intervals are Uniformly-Distributed

The end-point intervals that are given in Table A.1 were processed using the five-step procedure of Section III. Tables B.1 and B.2 in Appendix B summarize the results for Person 1. In Table B.1, Stages 1-4 correspond to Steps 1-4 of the EIA, as explained in Section II.A; m^* occurs at the end of Step 8; and, the left-end and right-end group statistics (mean and standard deviation) are used in Step 6 to establish the nature of the FOU.

In Table B.2, the trapezoidal UMF of the Person FOU is described (see Fig. 1) by four parameters (a, b, c, d), whereas the triangular LMF of the Person FOU is described (see Fig. 1) by five parameters ($e, f, g = f, i, h$). Centroids [7] provide a measure of uncertainty for each person's FOU, and the Center of Centroid is provided just for the ranking (ordering) of the words relative to one another [12].

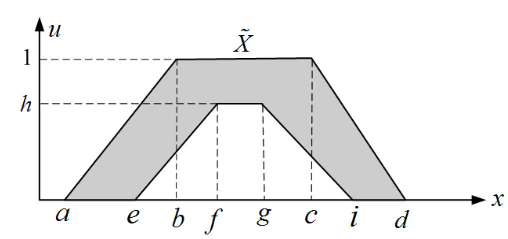

Fig. 1. Word FOU with the parameters that define its LMF and UMF.

Fig. 2 (located after the references) depicts the Person FOUs for each of the 10 persons and 10 words. Examining these FOUs, observe that: (1) For each word, usually the 10 Person FOUs look different, indicating significant inter-personal uncertainties for each word; and, (2) for many of the words, the FOUs for each person look different, indicating that many words are sufficiently different so that they will not be confused with one another.

These two observations are very qualitative and indicate the difficulty one encounters when trying to make comparative statements by looking at FOU pictures. We have therefore computed the Jaccard similarity measure [7, Ch. 4] across the 10 persons for each word. Our feeling is that if this similarity measure is greater than 0.50 then persons are more likely to be interpreting a word in the same way than not. Of course, one can raise this threshold and draw different conclusions.

Let $\tilde{A}$ and $\tilde{B}$ be IT2 FSs, with lower and upper MFs $\underline{\mu}_{\tilde{A}}(x)$, $\underline{\mu}_{\tilde{B}}(x)$, $\bar{\mu}_{\tilde{A}}(x)$, $\bar{\mu}_{\tilde{B}}(x)$, where $x \in X$. The formula for the Jaccard similarity measure used by us is based on average cardinality, was introduced in [12], and is:

$$sm_J(\tilde{A},\tilde{B}) = \frac{\sum_{i=1}^{N}\min\left(\bar{\mu}_{\tilde{A}}(x_i),\bar{\mu}_{\tilde{B}}(x_i)\right)+\sum_{i=1}^{N}\min\left(\underline{\mu}_{\tilde{A}}(x_i),\underline{\mu}_{\tilde{B}}(x_i)\right)}{\sum_{i=1}^{N}\max\left(\bar{\mu}_{\tilde{A}}(x_i),\bar{\mu}_{\tilde{B}}(x_i)\right)+\sum_{i=1}^{N}\max\left(\underline{\mu}_{\tilde{A}}(x_i),\underline{\mu}_{\tilde{B}}(x_i)\right)}$$

Tables C.1-C.10 in Appendix C provide the similarity matrices for the 10 words across the 10 subjects. Of the 45 unique similarity values in each of these 10 tables (the 10 diagonal self similarities are excluded), the counts of how many of the 45 were greater than 0.50 are:

Almost Impossible – 9/45, *Improbable* – 7/45, *Doubtful* – 10/45, *Unlikely* – 6/45, *Tossup* – 14/45, *Possible* – 9/45, *Likely* – 14/45, *Probable* – 5/45, *Good Chance* – 8/45 and *Almost Certain* – 8/45.

One observes from these numbers that *Probable* has the greatest inter-person uncertainty among the 10 words, whereas *Tossup* and *Likely* have the smallest inter-person uncertainty among the 10 words. All of these ratios are considerably less than 0.50 and so we can again conclude that there are very significant inter-personal uncertainties for all 10 of these words, i.e. *the FOUs are very person-dependent.* This is what we would expect when a CWW product is to be used as a Personal Advisor (and has nothing to do with robustness).

B. *End-Point Intervals are Normal- or Triangular-Distributed*

Everything just described in Section A, for uniformly distributed end-point intervals, was repeated by us for normal and triangular distributed end-point intervals. Figs. 3 and 4 (located after the references) depict the Person FOUs for each of the 10 individuals and 10 words for normal and triangular end-point interval distributions, respectively.

Comparing Figs. 2-4, observe that some Person FOUs can *look* quite different under the different distribution assumptions, but looks can be quite deceiving. Tables I-III give Jaccard similarities between Person FOUs under uniform and normal, uniform and triangular, and normal and triangular interval end-point uncertainty interval distributions, and show that: all similarities are greater than 0.50 and most are substantially greater than 0.50, indicating that *Person FOUs are robust to the kind of distribution that is assumed for the interval end-point data*.

VI. CONCLUSIONS

This paper has provided a new methodology for determining Person FOUs using only one subject. It uses interval end-point uncertainty intervals instead of only the end-point intervals. Such uncertainty intervals are relatively easy to obtain because they are obtained from and by the same person who is designing/using his /her own PA. Additionally, our new methodology for collecting the uncertainty intervals does not introduce methodological uncertainties during the data-collection process.

The EIA is still used by us for obtaining a Person's FOU, but it is applied to intervals that are generated by our new method from the collected end-point uncertainty intervals.

TABLE I
SIMILARITIES BETWEEN THE PERSON FOUS FOR UNIFORM AND NORMAL DISTRIBUTIONS. NUMBERS 1-10 ARE THE PERSON INDICES

Word	1	2	3	4	5	6	7	8	9	10
Almost Impossible	0.85	0.98	0.98	0.92	0.97	0.99	0.92	0.94	0.88	0.84
Improbable	0.96	0.96	0.97	0.95	0.5	0.92	0.81	0.92	0.94	0.89
Doubtful	0.92	0.92	0.94	0.89	0.98	0.91	0.94	0.92	0.96	0.93
Unlikely	0.93	0.95	0.97	0.94	0.93	0.89	0.93	0.95	0.92	0.91
Tossup	0.99	0.95	0.97	0.93	0.87	0.93	0.91	0.94	0.87	0.92
Possible	0.88	0.98	0.86	0.93	0.93	0.96	0.68	0.89	0.92	0.87
Likely	0.82	0.95	0.88	0.96	0.95	0.95	0.85	0.86	0.93	0.96
Probable	0.92	0.95	0.91	0.82	0.93	0.92	0.93	0.9	0.88	0.89
Good Chance	0.99	0.97	0.89	0.9	0.94	0.82	0.87	0.93	0.93	0.94
Almost Certain	0.97	0.99	0.89	0.92	0.92	0.91	0.93	0.94	0.89	0.96

TABLE II
SIMILARITIES BETWEEN THE PERSON FOUS FOR UNIFORM AND TRIANGULAR DISTRIBUTIONS. NUMBERS 1-10 ARE THE PERSON INDICES

Word	1	2	3	4	5	6	7	8	9	10
Almost Impossible	0.91	0.96	0.92	0.84	0.91	0.79	0.9	0.76	0.81	0.82
Improbable	0.88	0.92	0.91	0.83	0.52	0.91	0.79	0.85	0.81	0.85
Doubtful	0.83	0.85	0.81	0.88	0.96	0.89	0.74	0.87	0.87	0.8
Unlikely	0.86	0.87	0.94	0.89	0.81	0.76	0.92	0.78	0.87	0.79
Tossup	0.88	0.96	0.92	0.87	0.77	0.83	0.84	0.82	0.85	0.84
Possible	0.8	0.93	0.85	0.81	0.89	0.87	0.77	0.9	0.87	0.82
Likely	0.89	0.84	0.77	0.89	0.81	0.81	0.83	0.81	0.87	0.94
Probable	0.88	0.87	0.78	0.8	0.78	0.81	0.84	0.88	0.89	0.84
Good Chance	0.99	0.96	0.9	0.87	0.85	0.83	0.66	0.86	0.79	0.87
Almost Certain	0.98	0.95	0.74	0.88	0.9	0.83	0.9	0.82	0.85	0.87

TABLE III
SIMILARITIES BETWEEN THE PERSON FOUS FOR NORMAL AND TRIANGULAR DISTRIBUTIONS. NUMBERS 1-10 ARE THE PERSON INDICES

Word	1	2	3	4	5	6	7	8	9	10
Almost Impossible	0.87	0.97	0.94	0.91	0.93	0.79	0.97	0.81	0.88	0.92
Improbable	0.92	0.91	0.92	0.87	0.81	0.84	0.96	0.78	0.81	0.93
Doubtful	0.91	0.89	0.79	0.83	0.97	0.95	0.76	0.92	0.84	0.85
Unlikely	0.93	0.92	0.97	0.92	0.81	0.73	0.87	0.8	0.9	0.81
Tossup	0.87	0.95	0.91	0.86	0.86	0.88	0.78	0.86	0.9	0.91
Possible	0.9	0.95	0.97	0.79	0.86	0.89	0.88	0.96	0.85	0.91
Likely	0.9	0.87	0.87	0.87	0.79	0.82	0.84	0.84	0.82	0.94
Probable	0.92	0.91	0.87	0.87	0.73	0.84	0.8	0.96	0.84	0.94
Good Chance	0.98	0.97	0.94	0.84	0.85	0.83	0.73	0.91	0.84	0.9
Almost Certain	0.98	0.96	0.79	0.85	0.91	0.85	0.93	0.84	0.85	0.89

One may feel that needing two interval end-point bounds is a weakness of our method. We disagree and instead view this requirement as a novelty of our method. It is impossible to obtain an FOU from only one data interval, but it is now possible to obtain an FOU by collecting two interval end-point bounds from one person.

We applied our new method to ten probability words and found that: *Person FOUs are robust to the kind of* *distribution that is assumed for the interval end-point uncertainty intervals*.

Recently Miller, et al. [8] and Wagner et al. [9] have collected interval end-point data a few times from the same subject and from a small group of subjects. They then show how to obtain a general zSlice T2 FS model for a word where each subject's model resides at a different z-level. In this way they are able to preserve the nature of each subject's intra-uncertainty about a word as well as the inter-uncertainty about the word across the group of subjects. In a later (as yet unpublished) work they allow each subject to provide interval values for the two end-points. In all cases their final model is non-parametric, whereas our model is parametric. Since our goal was only to obtain a Person FOU, a GT2 FS model was not needed to do this.

We would also like to point out that to-date no theory or method exists to establish superiority of one kind of word model over another (e.g., [6] suggests using a Turing test, but, at present we do not know how to do this for word models). This is another area where new research would be most welcome.

REFERENCES

[1] H. Bustince, "Indicator of inclusion grade for interval-valued fuzzy sets: Applications in approximate reasoning based on interval-valued fuzzy sets," *Int. J. Approx. Reason.*, vol. 23, pp. 137-209, 2000.

[2] J. Joo and J. M. Mendel, "Obtaining an FOU for a word from a single subject by and *Individual Interval Approach*," *Proc. IEEE SMC Conference*, San Antonio, TX, October 2009.

[3] M. M. Korjani and J. M. Mendel, "Fuzzy love selection by means of perceptual computing," *Proc. of IFSA*, Edmonton, CA, June 2013.

[4] F. Liu and J. M. Mendel, "Encoding words into interval type-2 fuzzy sets using an *interval approach*, " *IEEE Trans. on Fuzzy Systems*, vol. 16, pp. 1503-1521, Dec. 2008.

[5] J. M. Mendel, "Computing with words and its relationships with fuzzistics," *Information Sciences*, vol. 177, pp. 988-1006, 2007.

[6] J. M. Mendel, "Computing with words: Zadeh, Turing Popper and Occam," *IEEE Computational Intelligence Magazine*, vol. 2, pp. 10-17, Nov. 2007.

[7] J. M. Mendel and D. Wu, *Perceptual Computing: Aiding People in Making Subjective Judgments*, John Wiley and IEEE Press, Hoboken, NJ, 2010.

[8] S. Miller, C. Wagner, J. Garibaldi and S. Appleby, "Constructing general type-2 fuzzy sets from interval-valued data," in *Proc. of FUZZ-IEEE 2012*, Brisbane, Australia, June 2012.

[9] C. Wagner, S. Miller and J. Garibaldi, "Similarity based Applications for Data-Driven Concept and Word Models based on Type-1 and Type-2 Fuzzy Sets," *Proc. of FUZZ-IEEE 2013*, Hyderabad, India, July 2013.

[10] T. S. Wallsten and D. V. Budescu, "A review of linguistic probability processing: general principles and empirical evidence," *The Knowledge Engineering Review*, vol. 10, no. 1, pp. 43-62, 1995.

[11] D. Wu, S. Coupland and J.M. Mendel, "Enhanced Interval Approach for Encoding Words into Interval Type-2 Fuzzy Sets and Its Convergence Analysis," *IEEE Trans. on Fuzzy Systems,* vol. 20, pp. 499-513, 2012.

[12] D. Wu and J. M. Mendel, "A comparative study of ranking methods, similarity measures and uncertainty measures for interval type-2 fuzzy sets," *Information Sciences*, vol. 179, no. 8, pp. 1169-1192, 2009.

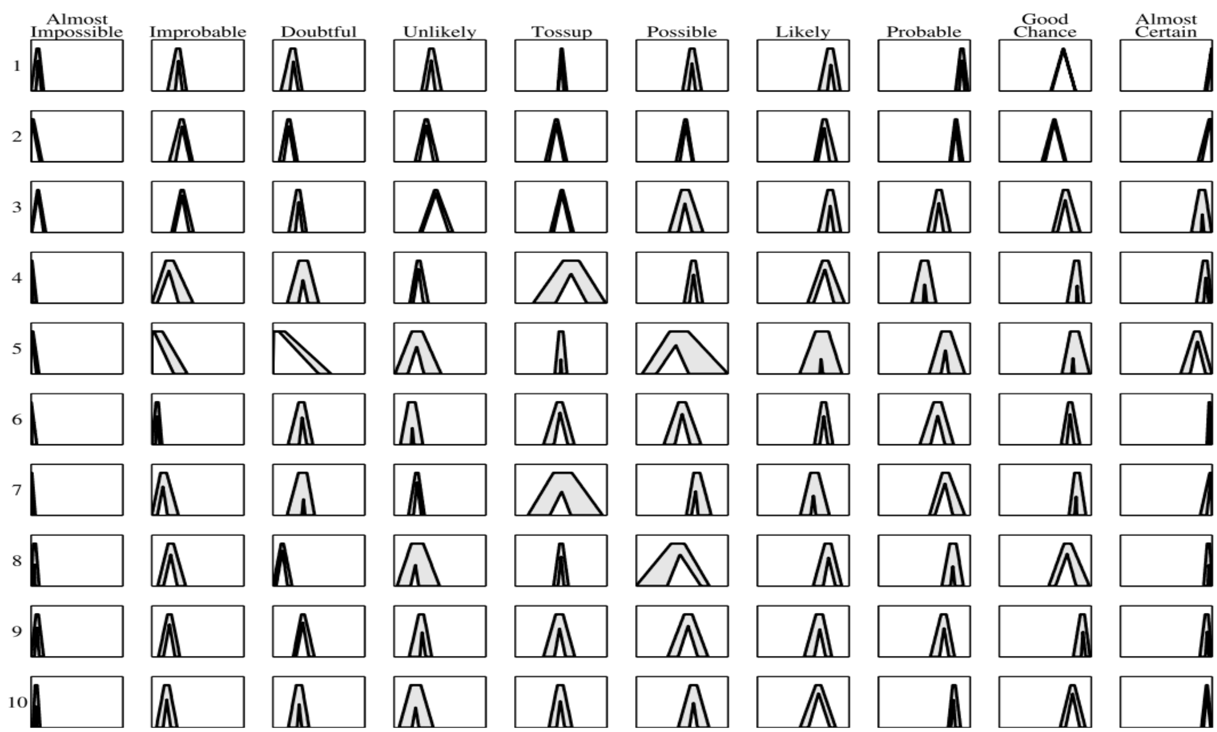

Fig.2. The 10 X 10 Person FOUs using uniformly distributed end-point intervals. Each numbered row represents a different individual.

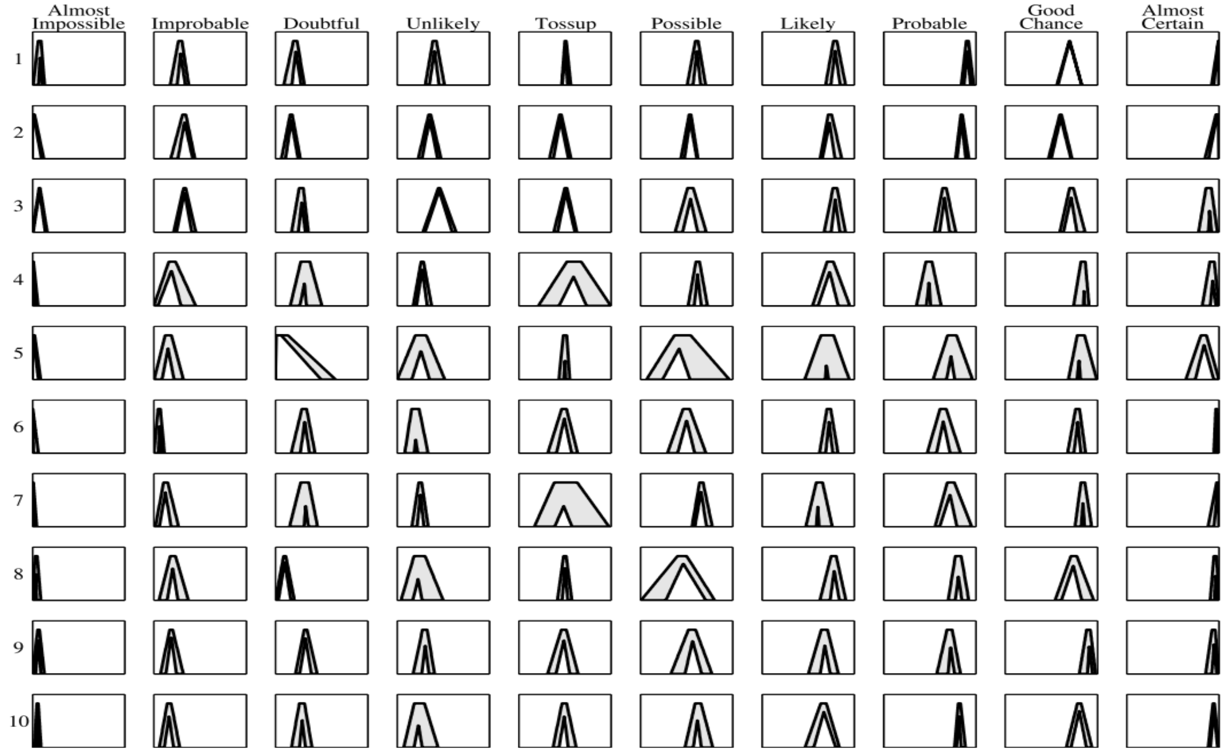

Fig. 3. The 10 X 10 Person FOUs using normally distributed end-point intervals. Each numbered row represents a different individual.

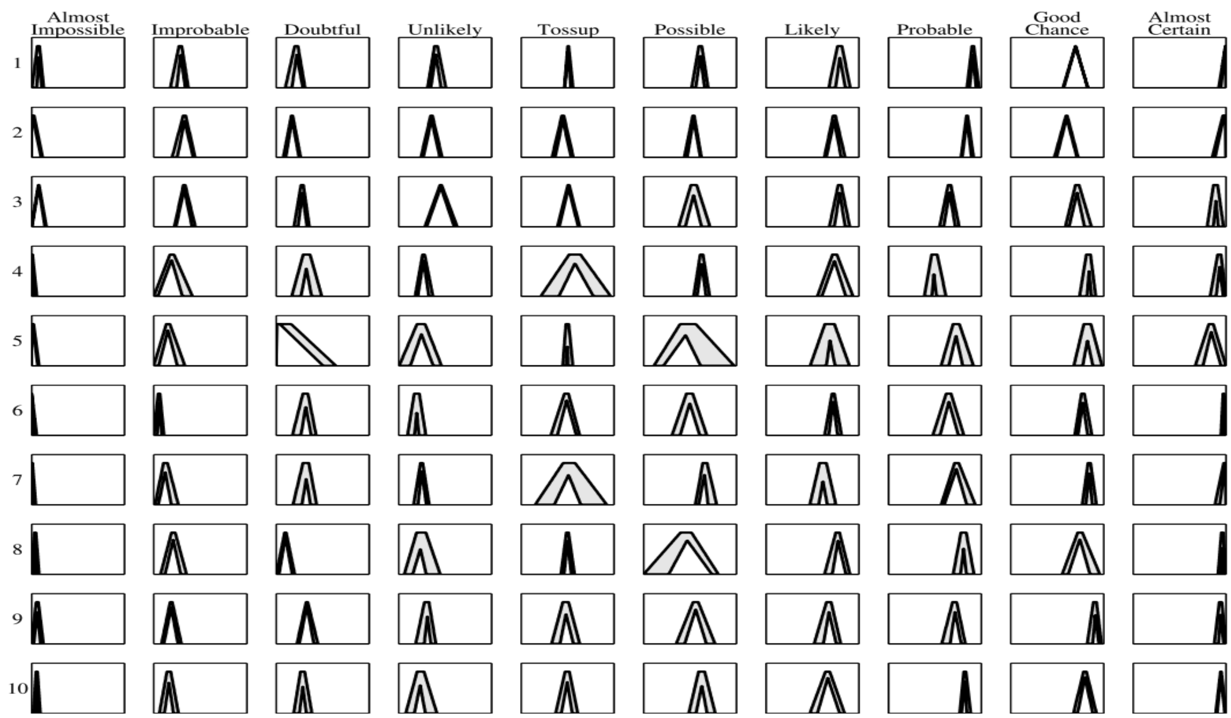

Fig. 4. The 10 X 10 Person FOUs using triangularly end-point intervals. Each numbered row represents a different individual.

APPENDIX A. INTERVAL END-POINTS DATA

TABLE A.1
INTERVAL END-POINTS DATA FROM TEN PERSONS

Person	Almost Impossible		Improbable		Doubtful		Unlikely		Tossup	
1	[0.02, 0.07]	[0.1, 0.12]	[0.2, 0.27]	[0.32, 0.35]	[0.12, 0.2]	[0.26, 0.29]	[0.33, 0.37]	[0.43, 0.49]	[0.48, 0.48]	[0.52, 0.55]
2	[0, 0.023]	[0.078, 0.096]	[0.225, 0.286]	[0.392, 0.411]	[0.097, 0.135]	[0.219, 0.243]	[0.265, 0.299]	[0.409, 0.444]	[0.363, 0.398]	[0.509, 0.534]
3	[0.02, 0.03]	[0.12, 0.14]	[0.25, 0.27]	[0.38, 0.42]	[0.2, 0.26]	[0.31, 0.34]	[0.33, 0.35]	[0.55, 0.59]	[0.42, 0.45]	[0.58, 0.6]
4	[0, 0.01]	[0.02, 0.05]	[0.05, 0.1]	[0.25, 0.39]	[0.2, 0.3]	[0.35, 0.47]	[0.2, 0.23]	[0.3, 0.35]	[0.3, 0.5]	[0.7, 0.9]
5	[0, 0.02]	[0.05, 0.07]	[0, 0.1]	[0.2, 0.3]	[0.1, 0.03]	[0.4, 0.5]	[0, 0.2]	[0.3, 0.45]	[0.45, 0.5]	[0.5, 0.55]
6	[0, 0]	[0.01, 0.05]	[0, 0.05]	[0.075, 0.1]	[0.2, 0.3]	[0.35, 0.4]	[0.1, 0.2]	[0.2, 0.3]	[0.35, 0.45]	[0.55, 0.6]
7	[0, 0.01]	[0.02, 0.04]	[0.04, 0.08]	[0.15, 0.25]	[0.18, 0.33]	[0.34, 0.42]	[0.18, 0.23]	[0.28, 0.31]	[0.25, 0.45]	[0.55, 0.85]
8	[0.01, 0.03]	[0.04, 0.08]	[0.1, 0.16]	[0.25, 0.33]	[0.02, 0.06]	[0.14, 0.18]	[0.1, 0.2]	[0.25, 0.45]	[0.44, 0.48]	[0.52, 0.56]
9	[0, 0.05]	[0.08, 0.12]	[0.1, 0.15]	[0.23, 0.28]	[0.25, 0.28]	[0.37, 0.42]	[0.2, 0.3]	[0.33, 0.38]	[0.35, 0.45]	[0.52, 0.61]
10	[0.03, 0.05]	[0.06, 0.09]	[0.08, 0.14]	[0.19, 0.25]	[0.2, 0.28]	[0.3, 0.37]	[0.1, 0.22]	[0.25, 0.38]	[0.4, 0.47]	[0.52, 0.59]
Person	Possible		Likely		Probable		Good Chance		Almost Certain	
1	[0.53, 0.59]	[0.63, 0.69]	[0.7, 0.78]	[0.83, 0.88]	[0.86, 0.89]	[0.93, 0.96]	[0.6, 0.61]	[0.79, 0.79]	[0.94, 0.96]	[1, 1]
2	[0.46, 0.499]	[0.589, 0.601]	[0.657, 0.681]	[0.768, 0.833]	[0.798, 0.812]	[0.872, 0.902]	[0.503, 0.531]	[0.677, 0.692]	[0.879, 0.912]	[0.973, 0.991]
3	[0.41, 0.49]	[0.57, 0.68]	[0.7, 0.77]	[0.82, 0.88]	[0.57, 0.63]	[0.69, 0.75]	[0.62, 0.66]	[0.76, 0.84]	[0.8, 0.89]	[0.9, 0.97]
4	[0.55, 0.6]	[0.65, 0.7]	[0.6, 0.65]	[0.8, 0.9]	[0.4, 0.5]	[0.5, 0.6]	[0.75, 0.85]	[0.85, 0.9]	[0.85, 0.92]	[0.95, 1]
5	[0.2, 0.3]	[0.5, 1]	[0.5, 0.7]	[0.7, 0.9]	[0.6, 0.7]	[0.75, 0.9]	[0.7, 0.8]	[0.8, 1]	[0.7, 0.8]	[0.9, 1]
6	[0.35, 0.45]	[0.55, 0.65]	[0.65, 0.7]	[0.75, 0.8]	[0.5, 0.6]	[0.7, 0.78]	[0.7, 0.75]	[0.8, 0.85]	[0.95, 0.97]	[0.98, 1]
7	[0.58, 0.62]	[0.66, 0.78]	[0.5, 0.6]	[0.62, 0.75]	[0.61, 0.65]	[0.78, 0.89]	[0.78, 0.83]	[0.84, 0.92]	[0.9, 0.96]	[0.98, 1]
8	[0.04, 0.4]	[0.64, 0.7]	[0.65, 0.75]	[0.82, 0.88]	[0.72, 0.8]	[0.82, 0.9]	[0.6, 0.66]	[0.8, 0.92]	[0.92, 0.96]	[0.97, 0.99]
9	[0.4, 0.5]	[0.63, 0.73]	[0.55, 0.65]	[0.72, 0.78]	[0.6, 0.7]	[0.75, 0.8]	[0.83, 0.9]	[0.93, 0.96]	[0.88, 0.94]	[0.96, 1]
10	[0.5, 0.6]	[0.65, 0.75]	[0.52, 0.6]	[0.75, 0.8]	[0.78, 0.8]	[0.83, 0.88]	[0.7, 0.75]	[0.85, 0.9]	[0.9, 0.92]	[0.97, 0.99]

APPENDIX B. RESULTS FOR PERSON 1

TABLE B.1
PERSON 1: REMAINING DATA INTERVALS AND THEIR END-POINT STATISTICS FOR 50 DATA INTERVALS

Word	Pre-processing				FS Part	Left-end Statistics		Right-end Statistics	
	Stage 1 n'	Stage 2 m'	Stage 3 m''	Stage 4 m	m^*	m_l	s_l	m_r	s_r
Almost Impossible	50	50	47	47	47	0.04	0.012	0.11	0.005
Improbable	50	50	49	49	49	0.23	0.021	0.34	0.009
Doubtful	50	50	50	50	50	0.16	0.026	0.27	0.008
Unlikely	50	50	48	48	48	0.35	0.01	0.46	0.017
Tossup	50	50	50	50	50	0.48	0	0.54	0.008
Possible	50	50	46	46	46	0.56	0.016	0.66	0.016
Likely	50	50	48	48	48	0.74	0.022	0.86	0.014
Probable	50	50	50	50	50	0.87	0.009	0.94	0.009
Good Chance	50	50	50	50	50	0.61	0.003	0.79	0
Almost Certain	50	50	50	50	50	0.95	0.006	1	0

TABLE B.2
PERSON 1: FOU DATA, WHERE EACH UMF AND LMF IS REPRESENTED AS A TRAPEZOID. THE 5^{TH} PARAMETER FOR THE LMF IS ITS HEIGHT.

Word	UMF	LMF	Centroid	Center of Centroid
Almost Impossible	[0,0.06,0.09,0.14]	[0.06,0.08,0.08,0.11,0.72]	[0.06,0.1]	0.08
Improbable	[0.17,0.27,0.31,0.38]	[0.26,0.29,0.29,0.33,0.72]	[0.25,0.32]	0.29
Doubtful	[0.09,0.19,0.24,0.32]	[0.19,0.22,0.22,0.28,0.69]	[0.18,0.25]	0.22
Unlikely	[0.3,0.39,0.43,0.52]	[0.35,0.41,0.41,0.45,0.73]	[0.38,0.43]	0.41
Tossup	[0.47,0.5,0.51,0.56]	[0.47,0.51,0.51,0.53,0.82]	[0.5,0.51]	0.51
Possible	[0.51,0.59,0.63,0.71]	[0.57,0.61,0.61,0.65,0.65]	[0.58,0.64]	0.61
Likely	[0.67,0.77,0.82,0.91]	[0.76,0.8,0.8,0.84,0.63]	[0.76,0.83]	0.8
Probable	[0.84,0.9,0.92,0.98]	[0.88,0.91,0.91,0.94,0.73]	[0.89,0.93]	0.91
Good Chance	[0.56,0.7,0.7,0.83]	[0.57,0.7,0.7,0.83,0.98]	[0.70,0.70]	0.70
Almost Certain	[0.92,0.99,1,1]	[0.95,1,1,1,1]	[0.97,0.98]	0.98

APPENDIX C. SIMILARITY MATRICES

TABLE C.1
ALMOST IMPOSSIBLE: SIMILARITY MATRIX AMONG THE TEN PERSONS

Person	1	2	3	4	5	6	7	8	9	10
1	1	0.31	0.66	0.13	0.19	0.13	0.08	0.31	0.76	0.44
2	0.31	1	0.33	0.4	0.7	0.32	0.34	0.44	0.38	0.32
3	0.66	0.33	1	0.1	0.2	0.09	0.07	0.26	0.61	0.33
4	0.13	0.4	0.1	1	0.57	0.8	0.83	0.29	0.14	0.13
5	0.19	0.7	0.2	0.57	1	0.46	0.49	0.46	0.23	0.24
6	0.13	0.32	0.09	0.8	0.46	1	0.73	0.29	0.14	0.13
7	0.08	0.34	0.07	0.83	0.49	0.73	1	0.21	0.1	0.07
8	0.31	0.44	0.26	0.29	0.46	0.29	0.21	1	0.37	0.47
9	0.76	0.38	0.61	0.14	0.23	0.14	0.1	0.37	1	0.54
10	0.44	0.32	0.33	0.13	0.24	0.13	0.07	0.47	0.54	1

TABLE C.2
DOUBTFUL: SIMILARITY MATRIX AMONG THE TEN PERSONS

Person	1	2	3	4	5	6	7	8	9	10
1	1	0.41	0.28	0.18	0.24	0.2	0.22	0.12	0.09	0.27
2	0.41	1	0.1	0.08	0.24	0.08	0.09	0.22	0.02	0.1
3	0.28	0.1	1	0.41	0.18	0.5	0.46	0.01	0.33	0.74
4	0.18	0.08	0.41	1	0.25	0.75	0.8	0.02	0.58	0.52
5	0.24	0.24	0.18	0.25	1	0.22	0.21	0.26	0.2	0.19
6	0.2	0.08	0.5	0.75	0.22	1	0.77	0.01	0.64	0.63
7	0.22	0.09	0.46	0.8	0.21	0.77	1	0.02	0.55	0.6
8	0.12	0.22	0.01	0.02	0.26	0.01	0.02	1	0	0.01
9	0.09	0.02	0.33	0.58	0.2	0.64	0.55	0	1	0.41
10	0.27	0.1	0.74	0.52	0.19	0.63	0.6	0.01	0.41	1

TABLE C.3
IMPROBABLE: SIMILARITY MATRIX AMONG THE TEN PERSONS

Person	1	2	3	4	5	6	7	8	9	10
1	1	0.41	0.28	0.18	0.24	0.2	0.22	0.12	0.09	0.27
2	0.41	1	0.1	0.08	0.24	0.08	0.09	0.22	0.02	0.1
3	0.28	0.1	1	0.41	0.18	0.5	0.46	0.01	0.33	0.74
4	0.18	0.08	0.41	1	0.25	0.75	0.8	0.02	0.58	0.52
5	0.24	0.24	0.18	0.25	1	0.22	0.21	0.26	0.2	0.19
6	0.2	0.08	0.5	0.75	0.22	1	0.77	0.01	0.64	0.63
7	0.22	0.09	0.46	0.8	0.21	0.77	1	0.02	0.55	0.6
8	0.12	0.22	0.01	0.02	0.26	0.01	0.02	1	0	0.01
9	0.09	0.02	0.33	0.58	0.2	0.64	0.55	0	1	0.41
10	0.27	0.1	0.74	0.52	0.19	0.63	0.6	0.01	0.41	1

TABLE C.4
UNLIKELY: SIMILARITY MATRIX AMONG THE TEN PERSONS

Person	1	2	3	4	5	6	7	8	9	10
1	1	0.36	0.41	0.05	0.17	0	0.01	0.18	0.12	0.09
2	0.36	1	0.24	0.19	0.29	0.05	0.1	0.3	0.37	0.21
3	0.41	0.24	1	0.04	0.12	0.01	0.01	0.12	0.08	0.07
4	0.05	0.19	0.04	1	0.41	0.24	0.64	0.37	0.55	0.45
5	0.17	0.29	0.12	0.41	1	0.43	0.36	0.86	0.41	0.7
6	0	0.05	0.01	0.24	0.43	1	0.35	0.48	0.22	0.57
7	0.01	0.1	0.01	0.64	0.36	0.35	1	0.34	0.37	0.42
8	0.18	0.3	0.12	0.37	0.86	0.48	0.34	1	0.44	0.8
9	0.12	0.37	0.08	0.55	0.41	0.22	0.37	0.44	1	0.48
10	0.09	0.21	0.07	0.45	0.7	0.57	0.42	0.8	0.48	1

TABLE C.5
TOSSUP: SIMILARITY MATRIX AMONG THE TEN PERSONS

Person	1	2	3	4	5	6	7	8	9	10
1	1	0.22	0.36	0.11	0.53	0.31	0.14	0.63	0.3	0.4
2	0.22	1	0.34	0.19	0.26	0.54	0.28	0.32	0.55	0.45
3	0.36	0.34	1	0.27	0.39	0.67	0.32	0.52	0.6	0.63
4	0.11	0.19	0.27	1	0.15	0.33	0.64	0.18	0.33	0.27
5	0.53	0.26	0.39	0.15	1	0.36	0.16	0.74	0.36	0.49
6	0.31	0.54	0.67	0.33	0.36	1	0.42	0.47	0.89	0.72
7	0.14	0.28	0.32	0.64	0.16	0.42	1	0.21	0.43	0.32
8	0.63	0.32	0.52	0.18	0.74	0.47	0.21	1	0.47	0.63
9	0.3	0.55	0.6	0.33	0.36	0.89	0.43	0.47	1	0.74
10	0.4	0.45	0.63	0.27	0.49	0.72	0.32	0.63	0.74	1

TABLE C.6
POSSIBLE: SIMILARITY MATRIX AMONG THE TEN PERSONS

Person	1	2	3	4	5	6	7	8	9	10
1	1	0.19	0.34	0.73	0.17	0.22	0.37	0.19	0.42	0.6
2	0.19	1	0.52	0.13	0.17	0.42	0.06	0.25	0.44	0.22
3	0.34	0.52	1	0.28	0.34	0.68	0.18	0.43	0.72	0.35
4	0.73	0.13	0.28	1	0.16	0.18	0.49	0.17	0.38	0.63
5	0.17	0.17	0.34	0.16	1	0.39	0.2	0.66	0.37	0.26
6	0.22	0.42	0.68	0.18	0.39	1	0.12	0.49	0.53	0.24
7	0.37	0.06	0.18	0.49	0.2	0.12	1	0.14	0.29	0.58
8	0.19	0.25	0.43	0.17	0.66	0.49	0.14	1	0.44	0.24
9	0.42	0.44	0.72	0.38	0.37	0.53	0.29	0.44	1	0.5
10	0.6	0.22	0.35	0.63	0.26	0.24	0.58	0.24	0.5	1

TABLE C.7
GOOD CHANCE: SIMILARITY MATRIX AMONG THE TEN PERSONS

Person	1	2	3	4	5	6	7	8	9	10
1	1	0.23	0.6	0.06	0.13	0.24	0.04	0.54	0	0.2
2	0.23	1	0.14	0	0.01	0.02	0	0.16	0	0.02
3	0.6	0.14	1	0.16	0.25	0.41	0.12	0.66	0.04	0.31
4	0.06	0	0.16	1	0.56	0.28	0.74	0.25	0.25	0.43
5	0.13	0.01	0.25	0.56	1	0.46	0.57	0.41	0.32	0.59
6	0.24	0.02	0.41	0.28	0.46	1	0.21	0.42	0.06	0.6
7	0.04	0	0.12	0.74	0.57	0.21	1	0.24	0.34	0.38
8	0.54	0.16	0.66	0.25	0.41	0.42	0.24	1	0.13	0.47
9	0	0	0.04	0.25	0.32	0.06	0.34	0.13	1	0.15
10	0.2	0.02	0.31	0.43	0.59	0.6	0.38	0.47	0.15	1

TABLE C.8
LIKELY: SIMILARITY MATRIX AMONG THE TEN PERSONS

Person	1	2	3	4	5	6	7	8	9	10
1	1	0.34	0.93	0.42	0.37	0.26	0.09	0.67	0.15	0.16
2	0.34	1	0.36	0.59	0.41	0.74	0.18	0.54	0.37	0.39
3	0.93	0.36	1	0.45	0.37	0.27	0.09	0.7	0.15	0.17
4	0.42	0.59	0.45	1	0.56	0.47	0.23	0.63	0.44	0.45
5	0.37	0.41	0.37	0.56	1	0.39	0.52	0.48	0.56	0.58
6	0.26	0.74	0.27	0.47	0.39	1	0.22	0.44	0.42	0.39
7	0.09	0.18	0.09	0.23	0.52	0.22	1	0.16	0.52	0.54
8	0.67	0.54	0.7	0.63	0.48	0.44	0.16	1	0.25	0.27
9	0.15	0.37	0.15	0.44	0.56	0.42	0.52	0.25	1	0.7
10	0.16	0.39	0.17	0.45	0.58	0.39	0.54	0.27	0.7	1

TABLE C.9
PROBABLE: SIMILARITY MATRIX AMONG THE TEN PERSONS

Person	1	2	3	4	5	6	7	8	9	10
1	1	0.14	0	0	0.07	0	0.06	0.12	0	0.07
2	0.14	1	0	0	0.19	0.01	0.16	0.44	0.03	0.48
3	0	0	1	0.07	0.35	0.65	0.36	0.07	0.44	0.01
4	0	0	0.07	1	0.03	0.17	0.02	0	0.03	0
5	0.07	0.19	0.35	0.03	1	0.33	0.79	0.42	0.65	0.23
6	0	0.01	0.65	0.17	0.33	1	0.35	0.09	0.4	0.03
7	0.06	0.16	0.36	0.02	0.79	0.35	1	0.35	0.63	0.2
8	0.12	0.44	0.07	0	0.42	0.09	0.35	1	0.18	0.55
9	0	0.03	0.44	0.03	0.65	0.4	0.63	0.18	1	0.06
10	0.07	0.48	0.01	0	0.23	0.03	0.2	0.55	0.06	1

TABLE C.10
ALMOST CERTAIN: SIMILARITY MATRIX AMONG THE TEN PERSONS

Person	1	2	3	4	5	6	7	8	9	10
1	1	0.44	0.09	0.14	0.04	0.43	0.64	0.36	0.21	0.23
2	0.44	1	0.24	0.41	0.13	0.26	0.7	0.4	0.47	0.51
3	0.09	0.24	1	0.52	0.42	0.08	0.2	0.19	0.36	0.25
4	0.14	0.41	0.52	1	0.22	0.14	0.3	0.32	0.64	0.55
5	0.04	0.13	0.42	0.22	1	0.03	0.1	0.08	0.15	0.11
6	0.43	0.26	0.08	0.14	0.03	1	0.36	0.47	0.21	0.21
7	0.64	0.7	0.2	0.3	0.1	0.36	1	0.51	0.41	0.41
8	0.36	0.4	0.19	0.32	0.08	0.47	0.51	1	0.48	0.47
9	0.21	0.47	0.36	0.64	0.15	0.21	0.41	0.48	1	0.65
10	0.23	0.51	0.25	0.55	0.11	0.21	0.41	0.47	0.65	1