Mutuality Measures Corresponding to Subjective Judgment of Similarity and Matching

Ayumi Yoshikawa
Faculty of Education, Okayama University
1-1, Naka 3-chome, Tsushima, Okayama 700-8530, Japan
ayumi@sip.ed.okayama-u.ac.jp

Abstract

In this paper matching measures corresponding to subjective judgment, which are calculated from two fuzzy sets, are examined through psychological experiments, in conjunction with examination of the similarity measures that have already been studied. We execute three separate psychological experiments that consist of subjective rating tasks of similarity and matching, and identification of membership functions.

Results show the following two conclusions clearly.

1) Indices related to distance between two fuzzy sets along element axis correspond to subjective judgment of similarity. This result confirms Yoshikawa's results in [5].
2) Same indices also correspond to subjective judgment of matching.

Moreover, the results reveal the following two tendencies. They, however, need to be reviewed through more experiments.

3) Human can distinguish between two different concepts, similarity and matching.
4) All the indices adopted in this study are functional as matching measures rather as similarity measures.

1. Introduction

Background: Similarity measures have played important role in various real applications. Many researchers have already studied this topic. Most recent studies on similarity measures laid weight on the side of mathematical properties [1-3]. Although their mathematical properties are significant, correspondence of these measures to subjective judgment of similarity is more important to obtain successful results in real applications

Zwick et al. [4] and Yoshikawa et al. [5] pointed out importance of considering subjective similarity, and evaluated the correspondence through psychological experiments. They then recommended some indices respectively based on their own results.

Point of this study: Matching is a concept similar to similarity. Indeed, both concepts have been used to express mutual relationship between two objects in many real applications. There is, however, no explicit distinction in usage between two concepts. Also, it is questionable whether we need to distinguish them.

Aims: This paper aims to evaluate following four issues through three psychological experiments.

1) Investigate if human can distinguish between two concepts of mutuality; similarity and matching.
2) Regress Yoshikawa's results [5]; indices related to distance along element axis correspond to subjective judgment of similarity.
3) Find out which indices correspond to subjective judgment of matching.
4) Clarify functionality of each index; which concepts each index works as.

We first compare between subjective ratings of similarity and matching for same stimuli to solve the first issue. The other issues are examined by correlation analysis between the subjective ratings and index values calculated from fuzzy sets identified by subjects.

2. Similarity and matching measures

Indices used in this study: 30 indices adopted in this study are exactly same as those used in Yoshikawa's study [5]. The indices are divided into following two categories; 1) related to overlap between two fuzzy sets, and 2) related to distance between two fuzzy sets along element axis.

Indices related to overlap between two fuzzy sets: In this category, these index values are calculated based on two membership degrees at each element. We adopted four major t-norms and two difference operators. Equations (1) through (6) denote their definitions.

Logical product:

$$\mu_C(x) = \mu_{A \wedge B}(x) = \min \{\mu_A(x), \mu_B(x)\} \tag{1}$$

Algebraic product:

$$\mu_C(x) = \mu_{A \bullet B}(x) = \mu_A(x) \cdot \mu_B(x) \tag{2}$$

Boundary product:

$$\mu_C(x) = \mu_{A \odot B}(x) = \max \{0, \mu_A(x) + \mu_B(x) - 1\} \tag{3}$$

Drastic product:

$$\mu_C(x) = \mu_{A \triangle B}(x) = \begin{cases} \mu_A(x) & \mu_B(x) = 1 \\ \mu_B(x) & \mu_A(x) = 1 \\ 0 & \textit{otherwise} \end{cases} \tag{4}$$

Symmetric difference:

$$\mu_C(x) = \mu_{A \Delta B}(x) = \max\{\min\{\mu_A(x), 1-\mu_B(x)\}, \min\{1-\mu_A(x), \mu_B(x)\}\} \quad (5)$$

Absolute difference:

$$\mu_C(x) = \mu_{A \nabla B}(x) = |\mu_A(x) - \mu_B(x)| \quad (6)$$

Where, A, B and C are fuzzy subsets on universe of discourse of X, and x belongs to X.

Then, equations (7) through (9) are used to defuzzify the fuzzy set of C calculated from above six operators. The equations mean height of fuzzy set, cardinality of it and cardinality ratio of it to logical sum set, in order.

Height:

$$\mathrm{hgt}(C) = \max_{x \in X} \mu_C(x) \quad (7)$$

Cardinality:

$$\mathrm{power}(C) = \int_{x \in X} \mu_C(x)\, dx \quad (8)$$

Cardinality ratio:

$$\mathrm{cr}(C) = \mathrm{power}(C) / \mathrm{power}(A \vee B) \quad (9)$$

Where, *Logical sum*: $\mu_{A \vee B}(x) = \max\{\mu_A(x), \mu_B(x)\}$.

Furthermore, we adopted the generalized Euclidean metric shown as (10).

Generalized Euclidean metric:

$$d_2(A,B) = \left(\int |\mu_A(x) - \mu_B(x)|^2 dx\right)^{1/2} \quad (10)$$

The number of indices in this category, therefore, becomes 19 in total.

Indices related to distance between two fuzzy sets along element axis: Meanwhile, in this category, index values are in relational to the distance between two elements on universe of discourse. Referring to Zwick et al. [4], we adopted the generalized Housdorff metric [6] and the generalized dissemblance index [7] in (11) and (12).

Housdorff metric:

$$\psi(A,B) = q(A,B) = \max\{|a_L - b_L|, |a_U - b_U|\} \quad (11)$$

Dissemblance index:

$$\psi(A,B) = \Delta(A,B) = \{|a_L - b_L| + |a_U - b_U|\} / 2 \cdot \{\beta_2 - \beta_1\} \quad (12)$$

Then, three values are calculated with (13) through (15) for each index.

Average:

$$\psi_1(A,B) = \int_0^1 \psi(A_\alpha, B_\alpha)\, d\alpha \quad (13)$$

Supremum:

$$\psi_\infty(A,B) = \sup_{\alpha \geq 0} \psi(A_\alpha, B_\alpha) \quad (14)$$

1 *level set*:

$$\psi_*(A,B) = \psi(A_{1.0}, B_{1.0}) \quad (15)$$

Where A_α and B_α are α-level sets of fuzzy sets of A and B. a_L and a_U are infimum and supremum of α-level set of A. β_1 and β_2 are arbitrary constant to normalize the index value.

Moreover, distance between gravity centers of fuzzy sets and fuzzy distance in type-2 fuzzy set in (16) and (17) were used. For the fuzzy distance, we calculated three values; infimum and supremum of 1 level set for the fuzzy distance, and gravity center of it, as shown in (18) through (20).

Distance between gravity centers:

$$\mathrm{Dist\text{–}gc}(A,B) = |\mathrm{gc}(A) - \mathrm{gc}(B)| \quad (16)$$

Where, $\mathrm{gc}(A) = \int_{x \in X} x \cdot \mu_A(x)\, dx / \mathrm{power}(A)$.

Fuzzy distance:

$$\mu_{\tilde{d}(A,B)}(\delta) = \sup_{\delta = |u - v|} \min\{\mu_A(u), \mu_B(v)\} \quad (17)$$

Infimum:

$$\mathrm{GLB\text{–}FDL1}(A,B) = \inf_{\mu_{\tilde{d}(A,B)}(\delta) = 1.0} \delta = \inf_{a \in A_{1.0}, b \in B_{1.0}} |a - b| \quad (18)$$

Supremum:

$$\mathrm{LUB\text{–}FDL1}(A,B) = \sup_{\mu_{\tilde{d}(A,B)}(\delta) = 1.0} \delta = \sup_{a \in A_{1.0}, b \in B_{1.0}} |a - b| \quad (19)$$

Gravity center:

$$\mathrm{GC\text{–}FD}(A,B) = \mathrm{gc}(\tilde{d}(A,B)) \quad (20)$$

Lastly, we adopted distance between vectors that consist of cardinality and gravity center of fuzzy set, in (21).

$$\mathrm{V}(A,B) = \sqrt{(\mathrm{power}(A) - \mathrm{power}(B))^2 + (\mathrm{gc}(A) - \mathrm{gc}(B))^2} \quad (21)$$

The number of indices in this category, therefore, becomes eleven in total.

Translation of indices: Note that indices from the two groups, i.e., all indices that belong to the later category, and six indices related to symmetrical and absolute difference, show dissimilarity or disagreement. We therefore translated them into degrees of similarity and matching by subtracting their values from unity, upon normalizing the universe of discourse to unity.

3. Experimental conditions and methodologies

3.1 Preparatory experiment: PE

Stimuli: In PE, five of verbal heaviness were used. They consisted of verbal hedges and "heavy," such as "very heavy." We used ten pairs combined from two of the five as stimuli.

Procedures: This experiment consists of two sub-tasks. In the first task, our subjects were asked to rate subjective similarity of each pair using six-graded rating scale along with six figures from 1 through 6. The ten stimuli and their rating scales were displayed in same page, and their order was same for all subjects. The second task was same as the first task

except that the subjects were asked to rate subjective matching. The subjects were instructed that they imagine to express heaviness of Japanese adult male verbally prior to both tasks. The two tasks were executed on Web page through internet browser. Thus, there was no limitation of answering time.

Subjects: The subjects were 28 volunteers recruited through internet. All of them were native Japanese speakers. Accordingly, all materials used in the experiments were displayed in Japanese.

3.2 Experiment 1: E1

Stimuli: The stimuli used in E1 were absolutely same as those used in the PE, namely, five of verbal heaviness and the ten pairs.

Procedures: This experiment consists of three sub-tasks. In the first task, subjects identified a membership function for each verbal heaviness using the Fuzzy graphical rating scale. They drew two ranges on graph scale; one is a most suitable range to express the degree of verbal heaviness, and the other is a possible range to express it. The other two tasks were absolutely same as the two tasks in the PE except for using printed questionnaires instead of Web page in the PE. The questionnaires were distributed, and were collected after a few days.

Subjects: The subjects were 29 students at Kyoto Institute of Technology. All of them were native Japanese speakers. All materials used in the E1, thus, were printed in Japanese. They participated voluntarily and had little technical knowledge on fuzzy set theory.

3.3 Experiment 2: E2

Stimuli: In E2, seven of verbal tallness, such as "very tall," were used. The five hedges used in the PE and E1 were remained, and two new ones were added. We used twenty-one pairs as stimuli.

Procedures: This experiment had two major differences from both the PE and the E1. The first difference is experimental order of two subjective rating tasks. In this experiment, task of subjective rating of matching was followed by rating task of similarity. Consequently the order of three tasks was as follows; identification of the membership functions, rating of similarity, and rating of matching. The other difference is the number of stimuli per one page. In this experiment, only one stimulus was displayed on the screen at every question, unlike both of the prior experiments, the PE and the E1. Correction of rating value was not allowed after the next stimulus was displayed. Another minor change was that seven-graded rating scale was used in both subjective rating tasks. These three tasks were executed on Web page using JAVA applets and internet browser. The BASE method [8] was used to identify the trapezoidal membership functions, instead of the Fuzzy graphic rating scale. Other conditions were same as the PE and the E1.

Subjects: The subjects were 21 students at Kyoto Institute of Technology and one more participant from outside the Institute. Each subject participated in only one experiment. All of them were native Japanese speakers. Consequently all materials used in the E2 were displayed in Japanese. They participated voluntarily and had little technical knowledge on fuzzy set theory. Results obtained from ten subjects had problems on their subjective rating and/or membership functions, and were excluded from the analysis. The number of effective subjects was twelve.

4. Results and discussions

4.1 Difference between the two concepts, similarity and matching

PE and E1: We first make scatter table based on rating values of similarity and matching, RVS and RVM, for same stimuli. ***Table 1*** shows cumulative results for all subject. Results from the PE and the E1 reveal similar tendency as follows. 1) The highest frequencies among same RVS take at

Table 1: Comparison of subjective rating results among three experiments.

(a) PE

rating values		rmatching: RVM 1	2	3	4	5	6	sum
similarity: RVS	6	1	4	4	6	22	14	51
	5	3	4	11	10	14	3	45
	4	1	5	12	13	8	1	40
	3	7	16	8	5	1		37
	2	18	22	8	2			50
	1	47	6	3	1			57
	s.	77	57	46	37	45	18	280

(b) E1

rating values		matching: RVM 1	2	3	4	5	6	sum
similarity: RVS	6		4	2	6	42	20	74
	5	2	3	6	14	12	4	41
	4	1	8	15	8	4		36
	3	5	27	11	2			45
	2	20	18	5				43
	1	41	8			2		51
	s.	69	68	39	30	60	24	290

(c) E2

rating values		matching: RVM 1	2	3	4	5	6	7	sum
similarity: RVS	7			1			7	13	21
	6		2	1	3	11	29	9	55
	5		1	3	2	18	12	2	38
	4	2	3	2	1	3	4	1	16
	3		5	15	7	3	3		33
	2	10	21	13	2				46
	1	26	13	4					43
	s.	38	45	39	15	35	55	25	252

the left cells of diagonal cells, i.e., the cell surrounded by square in Table 1(a) and (b). 2) The cumulative frequency of upper-left triangular area is greater than that of lower-right triangular area. Namely, the ratio of "RVS > RVM" versus "RVM > RVS" is 124 : 38 and 155 : 25 in Table 1(a) and (b) respectively. These results show that the subjects recognize similarity and matching as distinct concepts.

E2: Result from the E2, however, is different from the others. As seen from Table 1(c), the highest frequencies among same RVS take at diagonal cells surrounded by square. The ratio of "RVS > RVM" versus "RVM > RVS" is 53 : 76. This means there is a tendency that the RVM is greater than or equal to the RVS. The tendency, however, disagrees with our intuitive recognition of mutual relationship.

Causes for differences: Possible causes for above differences are as follows; 1) differences of subjects, and 2) differences of experimental conditions in both rating tasks. For the subjects, there is no subject who participated in more than two experiments. However, it is difficult to think that the differences of subjects among three experiments cause the differences of results, since the results from the PE and the E1 show close resemblances.

Meanwhile, as for experimental conditions, order of rating tasks and the number of displayed rating scale per one time in the E2 differs from those in the PE and the E1. As mentioned in section 3, in the E2, only one rating scale was displayed on screen per one page, and the subjects rated each stimulus in sequence and in one direction. The subjects, therefore, did not know a whole set of the stimuli, at least until they finished the rating task of similarity, which preceded that of matching. It can be thought that the lack of knowing the whole set of stimuli suppresses the RVS. Although we need extra experiments to clear this difference, we conclude 1) the RVS in the E2 is biased because of the experimental conditions, and 2) the subjects recognize similarity and matching as distinct concepts based on the results from the PE and the E1.

4.2 Similarity measures corresponding to subjective judgment of similarity

Preparations: We first calculated 30 index values for each pair of stimuli from membership functions identified by the subjects. Next, the Pearson's correlation coefficients between the index values and RVS or RVM were computed for each index and each subject. Strictly speaking, both RVS and RVM satisfied only up to level of ordinal scale. We, however, regarded that they were in level of interval scale, considering

Table 2: Pearson's correlation coefficients between RVS, RVM and indices

(a) E1

Indices	Matching			Similarity		
	Ave.	SD	Sig.	Ave.	SD	Sig.
1-Euclidean distance. of vectors(cardinality, gravity center)	0.844	0.066	18	0.816	0.092	17
1-Dissemblance index (maximum value over all level sets)	0.843	0.070	18	0.813	0.088	17
1-distance between gravity centers of fuzzy sets	0.842	0.070	18	0.813	0.093	17
Average for 11 indices related to distance along element axis	0.832	0.074	17	0.810	0.095	16
height of logical product set	0.793	0.151	14	0.769	0.129	13
cardinality ratio of logical product set to logical sum set	0.764	0.163	14	0.683	0.146	10
relative cardinality of logical product set	0.749	0.189	12	0.693	0.169	11
Average for 19 indices to express overlap of fuzzy set	0.732	0.187	12	0.665	0.183	9

• "Sig." means the number of significant correlation at 1% level in one-tail

(b) E2

Indices	Matching			Similarity		
	Ave.	SD	O.	Ave.	SD	O.
1-Dissemblance index (for 1 level set)	0.870	0.070	1	0.870	0.056	1
1-Dissemblance index (maximum value over all level sets)	0.864	0.075	3	0.869	0.049	2
1-distance between gravity centers of fuzzy sets	0.863	0.070	4	0.858	0.052	4
Average for 11 indices related to distance along element axis	0.838	0.084		0.834	0.079	
1-relative cardinality of symmetric difference set	0.821	0.107	10	0.826	0.094	9
1-relative cardinality of absolute difference set	0.819	0.107	11	0.824	0.092	11
1-Euclidean metric (Minkowski metric)	0.797	0.109	13	0.805	0.088	12
Average for 19 indices to express overlap of fuzzy set	0.748	0.136		0.755	0.122	

• "O." means order of each index among all indices based on average.

that these rating scales were in proportion to their labels. Ten subjects in the E1 excluded from following analysis, since their maximum value of the correlation coefficient over 30 indices was less than 1 % point in one-tail, 0.716 (*df.* = 8). Consequently the number of effective subjects becomes 18. Meanwhile, for 12 effective subjects in the E2, the maximum value for each subject was greater than 0.7, since ill subjects were excluded prior to this analysis. Then, mean values and SDs over all subjects were calculated. ***Table 2*** shows summarized results from the E1 and the E2.

Comparison between indices: As seen from columns of "Similarity" in Table 2(a) and (b), 11 indices related to distance along element axis, IDE, strongly more correlate with the RVS than 19 indices to express overlap of fuzzy sets, IOF. These results agree with the results in Yoshikawa et al. [5]. From the results of the E1, minimum value of average among 11 IDE, LUB-FD1 (0.799), is larger than maximum value of average among 19 IOF, hgt-LP (0.769). For the E2, moreover, similar order is found, except for two indices of the IDE. In addition, there is no significant difference of the average among the 11 IDE at 5% in two-tail, except for GLB-FD1, V(A,B) in E2.

4.3 Matching measures corresponding to subjective judgment of matching

Comparison among indices: As seen from Table 2(a) and (b), correlation between the IDE and the RVM is apparently larger than correlation between the IOF and the RVM. The order of indices based on averages shows close resemblance to the order of the RVS. The indices that correspond to subjective judgment of similarity in Zwick et al. [4] and Yoshikawa et al. [5], therefore, correspond to subjective judgment of matching. Moreover, there is no significant difference of averages of correlation coefficients among the IDE at significance level of 5% in two-tail, except for GLB-FD1 in E2. Consequently, these results mean that we have a choice for a matching index to utilize in real applications out of the ten indices, with room of considering cost and computation complexity of each index in those applications.

4.4 Functionality of each index; which concept the index works as

Comparison between RVS and RVM: As seen from Table 2(a) and (b), both results from the E1 and the E2 differ each other. The results from the E1 show averaged correlation coefficient with the RVM is greater than the coefficient with the RVS, comparing between both for each index. These mean all the indices work as matching measures, rather as similarity measures that has been so far ordinarily used. Contrarily, the results from the E2 imply all indices work as similarity and matching measures, and there is no difference in their functionality. We, however, need to recall the results of the RVS and the RVM from the E2 mentioned in 4.1. The RVS in the E2 were biased because of the experimental conditions, then the RVS and the RVM are almost equal. Here we regard excluding the results of the E2 is adequate. Consequently, it is concluded that all the indices adopted in this study function as matching measures than similarity measures. Yet, we need to execute more experiments to review the conclusion.

5. Conclusion

In this study, we have examined the relationships among subjective judgment of similarity and matching and indices calculated from fuzzy sets through three separate psychological experiments. The results have shown following two conclusions clearly.

1) Indices related to distance between two fuzzy sets along element axis correspond to subjective judgment of similarity. This result confirms Yoshikawa's results in [5].
2) Same indices also correspond to subjective judgment of matching.

Moreover, the results reveal the following two tendencies. They, however, need to be reviewed through more experiments.

3) Human can distinguish between two different concepts, similarity and matching.
4) All the indices adopted in this study are functional as matching measures rather as similarity measures.

References

[1] C. P. Pappis, "Values approximation of fuzzy systems variables", Fuzzy Sets and Systems, 39 (1991) 111-115.

[2] S. M. Chen, et al., "A comparison of similarity measures of fuzzy values", Fuzzy Sets and Syst., 72 (1995) 79-89.

[3] J. Fan and W. Xie, "Some notes on similarity measure and proximity measure", Fuzzy Sets and Systems, 101 (1999) 403-412.

[4] R. Zwick, et al., "Measures of Similarity Among Fuzzy Concepts: A Comparative Analysis", International Journal of Approximation Reasoning, 1 (1987) 221-242.

[5] A. Yoshikawa and T. Nishimura, "Relationship between Subjective Degree of Similarity and Some Similarity Indices of Fuzzy Sets", in Methodologies for the Conception, Design, and Application of Intelligent Systems, eds. T. Yamakawa and G. Matsumoto (World Scientific, Singapore 1996) 818-821.

[6] A. L. Ralescu and D. A. Ralescu, "Probability and Fuzziness", Information Science, 34 (1984) 85-92.

[7] A. Kaufman and M. M. Gupta, Introduction to Fuzzy Arithmetic (Van Nostrand Reinhold, New York, 1985).

[8] A. Yoshikawa, "Improvement of membership function identification method in usability and precision", in Advanced in Soft Computing – Engineering Design and Manufacturing, eds. R. Roy, et al. (Springer-Verlag London, 1998) 248-259.