Metrics of Textual Coherence for Structured Document

Mana Ihori, Atsuko Tominaga, Yoshinari Takegawa, and Keiji Hirata

***Abstract*—In this study, we propose the metrics to measure textual coherence by considering the document structure such as section, paragraph, and sentence. We suppose that when textual coherence is captured by considering the document structure, it is possible to fit human intuition. In this metrics, we first make graph structures at each document structure base on sentence similarity. Next, we recursively aggregate the coherence values of structures at a layer in a bottom-up manner. We produce test data employing the sentence ordering task for individual layers. We assess the proposed metrics employing sentence ordering tasks: discrimination and insertion. We compare the performance of our metrics with the conventional graph-based local coherence model. The our metrics outperforms, in particular, at sentence ordering tasks conducted of the section layer; it follows that the proposed metrics works well for a structured document as containing sections, paragraphs such as a technical paper and a product manual.**

***Index Terms*—textual coherence, layered structure, sentence similarity graph, text evaluation metrics, graph structure.**

I. INTRODUCTION

ACCORDING to Richard [1], coherence is "the relationships which link the meaning of utterances in a discourse or of the sentences in a text". Coherence concerns the content of a text, and if a document has high coherence, it is easy to understand its content [2]. Moreover, a document that has high coherence can convince the assertion to readers and can be communicated to them flexible. Therefore, we should always write a document has high coherence.

A document is composed of a lot of sentences. Moreover, a paragraph of the document is composed of some sentences, and a section of the document is composed of some paragraphs. These units are composed of sentences that are related semantically. In this way, the document is composed of a lot of structures that have different granularity. We suppose that these units should have coherence in each of them. Thus, it is inferred that we should consider the document structure to capture textual coherence.

In this paper, we propose a metrics of textual coherence for logical and structured documents such as a technical paper and a product manual (Section III). We make the graph structure at every document structure to capture textual coherence of a structured document (Section III-A). Moreover, we recursively aggregate coherence values of structures at a layer in a bottom-up manner (Section III-B). We carry out experiments that are discrimination and insertion to compare the performance of the metrics with the conventional graph-based local coherence model (Section IV). The result of experiments shows that the importance of considering the document structure (Section V). We discuss that reason (Section VI).

Manuscript received July 02, 2018; revised July 30, 2018.

M. Ihori is with Future University Hakodate Graduate School, Japan, Hakodate, e-mail: g2117008@fun.ac.jp.

A. Tominaga, Y. Takegawa and K. Hirata are with Future University Hakodate, e-mail: { tominaga, yoshi, hirata } @fun.ac.jp

TABLE I
ENTITY GRID

	Excel	substitution	function	character	you
S_1	S	O	O	X	-
S_2	-	X	X	O	S
S_3	-	X	S	-	-

S = Subject, O = Object, X = Other, - = No edge

II. RELATED WORK

A. Graph-based Coherence Model

There are many works about the textual coherence since there is much demand for it in fields such as automatic evaluation and automatic generation. The centering theory that attempts to model one aspect constituting coherence of discourse is a representative work among them [3]. In this theory, coherence is captured by focusing on updating of the central part (a topic of a discourse) in each discourse based on the hypothesis that the appearance of an entity in a document that has high coherence is the regularity.

However, Barzilay and Lapata pointed out that centering theory relies on manually tagged input too much, and they proposed the entity grid model that is easier to input [4]. Entity grid is the matrix that represents a distribution of discourse entities across sentences in a document. Table. I shows the entity grid of the following sentences (S1-S3).

- S1 : Excel has a substitution function of character.
- S2 : You can substitute a character with the substitution function.
- S3 : However, the function of "all substitution" needs to beware.

The entity grid model captures local coherence with ranking function, after computing the transition probability of the syntactic role of an entity in adjacent sentences with entity grid. The performance of this model works well. However, it relies on computationally expensive training phase and faces data sparsity problems.

Guinaudeau and Strube proposed the graph representation of the entity grid in order to overcome these problems [5]. The graph-based local coherence model captures local coherence by applying centrality measures to the nodes in the graph. This model presents a text with a graph in which the syntactic role of entity grid is replaced by numerical value (Subject: 3, Object: 2, Other: 1, No edge: 0). Moreover, this model captures local coherence by using each edge in the graph associated with a weight that depends on the syntactic role of the entity. For example, Table. II shows that an

ISBN: 978-988-14048-1-7
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

TABLE II
ADJACENCY MATRIX IN GRAPH REPRESENTATION

	S_1	S_2	S_3
S_1	0	X_1	X_2
S_2	0	0	X_3
S_3	0	0	0

Weighted :$X_1 = 2$, $X_2 = 0.5$, $X_3 = 1$
Acc :$X_1 = 6$, $X_2 = 8$, $X_3 = 4$

adjacency matrix in graph representation of Table. I. It is proposed that two kinds of methods to calculate weights of the graph; (1) Weighted: Edges are weighted according to the number of entities shared by two sentences. (2) Acc: Edges are weighted according to the sum of the product of the value of entities shared by two sentences. These weights are divided by distance of two sentences. Formally, the local coherence of a document D that has N sentences is equal to

$$LocalCoherence(D) = \frac{1}{N}\sum_{i=1}^{N} OutDegree(S_i),$$

where $OutDegree(S_i)$ is the sum of the weights associated to edges of $i-th$ sentence. The higher this value is, the higher coherence is.

The performance is comparable to the entity grid model with a simple calculation, and this model became possible to consider coherence of non-adjacent sentences. However, this model does not consider the document structures that are described in the introduction since it only focuses on the distance between sentences.

B. Cosine Similarity of Sentences

Foltz *et al.* proposed the method that captures coherence by using semantic relevance between adjacent sentences based on the hypothesis that a document that has high coherence contains many semantically related words [6]. In this model, coherence is capture by calculating cosine similarity between adjacent sentences with Latent Semantic Analysis (LSA) [7]–[9] to consider semantic relevance. The coherence value is the sum of cosine similarity is divided by the total number of sentences in a document. Formally, the local coherence of a document D that has N sentences is equal to

$$Coherence(D) = \frac{\sum_{i=1}^{N-1} cos(S_i, S_{i+1})}{N-1},$$

$$cos(S_x, S_y) = \frac{\sum_{i=1}^{z} S_{xi} \cdot S_{yi}}{\sqrt{\sum S_{xi}^2} \cdot \sqrt{\sum S_{yi}^2}}, \quad (1)$$

where S_i and S_{i+1} is adjacent sentences, z is the number of all words in a document D, and S_{xi}, S_{yi} is the value of $i-th$ word of sentence vector S_x and S_y. This model shows coherence could be captured by calculating cosine similarity between adjacent sentences with LSA. However, this model does not consider non-adjacent sentences and the document structures.

Also, Erkan and Radev proposed to the method to summarize multiple documents by making sentence similarity graph that uses cosine similarity between sentences [10]. In this way, cosine similarity between sentences is used for a summary of a document. Therefore, cosine similarity between sentences is used to search relating information of sentences in a document.

III. METHOD

A. Graph Considering Document Structures

Paragraph-writing[1] is a common international method to write a logical document. In this method, coherence is main components of a paragraph since the ideas in the paragraph must be presented in logical order [11]. Hence, a document that is written by the method has high coherence and tends to have similar writing at every document layer. In other words, the similar writing method is recursively repeated along the layered structure of a document in a bottom-up manner.

We focus on the distance of a similarity sentence pair in our metrics. In centering theory, textual coherence is high when the center of a sentence continues [3]. In the other word, when the distance of a sentence pair is short, the sentence pair has similar topic and it is related. The more the sentence pairs a document has, the higher coherence is. Moreover, when a text is written by paragraph-writing, it is inferred that a information of the sentence pair in paragraphs or sections with short distance also is related.

In this study, we propose the metrics that capture textual coherence by considering the document structure. We suppose that this metrics is possible to fit human intuition since it is based on insights of layered structure of a document and paragraph-writing. First, we make graph structures to consider non-adjacent sentences [5]. Here, we use sentence similarity since cosine similarity between sentences is suitable for capturing the related information [6], [10]. Next, we capture coherence by using features of each graph. For example, in Fig. 1, for (C) in section 1, it is inferred that coherence of the graph of subsection 1 is higher than that of the graph of subsection 2. Because of this, in same structure, the more sentence pairs are and the shorter the distance of the edge is, the higher coherence is.

We describe our method to make the graph at every document structure. First, we make sentence vector constituted by tf values from a document and calculate cosine similarity of all sentence pairs with equation (1). Here, we calculate cosine similarity without LSA since our most important idea is to propose the metrics to capture coherence of the structured document. We adopt tf value that can be calculated correctly from few words since sentence vectors are made from one sentence. Second, we make graphs from all sentence pairs whose cosine similarity is 0.1 or more. In the graph-based coherence model, graphs are made when one discourse entity is shared by two sentences at least [5]. Hence, we employed the low cosine similarity (0.1) as the threshold.

For example, in Fig. 1, for (A), it is assumed that the cosine similarity of sentence pairs that connected by a red line is 0.1 and more. In this time, we make graphs in a bottom-up manner. First, we focus on paragraph layer and make graphs from sentence similarity pairs in same paragraphs. Next, we focus on subsection layer and make graphs from sentence similarity pairs across paragraphs in

[1]Scott, Fred Newton. Paragraph-writing. Boston Allyn And Bacon, 1909.

ISBN: 978-988-14048-1-7
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

same sections. In this way, we decide a target layer and make graphs from sentence similarity pairs across layers that is under the target layer. We repeat this method until the top layer and make graphs in every document structure. It should be noted here that we make as many graphs as structures in a document instead of making as many graphs as layers in a document.

B. Coherence Metrics of Each Graph

Our metrics capture coherence from features of the graph at each document structure. We focus on features of the graph: structure coverage ratio, sentence similarity, the distance of the edge, and the number of edges. Structure coverage ratio is the ratio of the number of structures the graph has. For example, in Fig. 1, for (C) in section 1, the graph of subsection 1 is made from two paragraphs. However, subsection 1 has three paragraph. Here, structure coverage ratio of this graph is equal to $2/3 = 0.67$. We suppose that the more structures the graph has, the higher coherence is. Hence, when this value closes 1, coherence is high.

The edge of the graph is weighted according to cosine similarity between two sentences. In this study, these weights are divided by the distance of the edge (the distance between structures that have similarity sentences). For example, in Fig. 1, for (c) in section 2, the graph of subsection 1 is made across two paragraph. The distance of the edge is 1 since these are adjacent paragraphs. Here, to evaluate each graph fairly, these values are divided by the number of edges in the graph. Formally, the coherence of graph G that have n edges is equal to

$$Coherence(G) = R \times \frac{1}{n} \sum_{i=1}^{n} \frac{cos_i}{dis_i}, \qquad (2)$$

where R is structure coverage ratio, cos_i is cosine similarity of $i-th$ edge, dis_i is a distance of $i-th$ edge.

C. Method of Capture Textual Coherence

We propose a method to aggregate the coherence value of each structure since a document consists of more than one structure. Here, a bottom layer in the document structure is a paragraph. The coherence value of a paragraph layer is calculated by equation (2). However, the layer except for paragraphs has a layer of the different level. For example, a section has some subsections, and a subsection has some subsubsections or paragraphs. Thus, the coherence value in the layer except for paragraph is calculated by multiplying the coherence value of that layer and the average of coherence values of a layer under that layer. Formally, the value of coherence of the layer (i) except for paragraph is equal to

$$X^i = Coherence(i)^{\alpha} \times (\frac{1}{k} \sum_{n=1}^{k} X_n^{i-1})^{1-\alpha}, \qquad (3)$$

where k is the number of graphs of the $i-1$ layer, $i-1$ layer is the layer under i layer. The value of textual coherence is calculated by repeating equation (3) recursively in a bottom-up manner. It should be noted here that parameter α ($0 < \alpha < 1$) is set in equation (3). It is thought that we can choose the coherence value of the layer that is prioritized by setting parameter α. When the value of α is set low, our metrics priors the coherence value of an under layer. In this study, the value of textual coherence is calculated by equations (2), (3). The higher this value is, the higher coherence is.

TABLE III
AVERAGE NUMBER OF EACH LAYER FROM ACADEMIC PAPERS

	Data A	Data B
Section	5.9	5.9
Subsection	13.6	7.0
Subsubsection	1.1	13.7
Paragraph	49.4	47.9
Sentence	182.6	169.1

D. Feature of Graph that is Made from Academic Papers

We investigate the graph at every layer that is made from a technical papers to confirm if we can evaluate textual coherence with the proposed metrics. Our corpora are journal papers in Japanese of 7 educational engineering (hereinafter, referred to as Data A) and 7 information engineering (hereinafter, referred to as Data B). It is expected that coherence of journal papers is higher than general papers since they are reviewed exactly. Also, these papers have layers: sections, subsections, subsubsections, paragraphs, sentences. Table. III shows the average number of each layer that each data have.

In this investigation, we focus on the cosine similarity of each edge and the distance of each edge in all graphs at each layer. Fig. 2 shows that the ratio of the number of edges for each cosine similarity at each layer. Fig. 3 shows that the ratio of the number of edges for each the distance of the edge at each layer. Here, these graphs show the result of Data A as an example since these results of the two datasets were similar. As a result, we found that the lower sentence similarity is or the shorter distance of the edge is, the more edges the graph has. Moreover, these results were similar in every layer. Therefore, it is inferred that our metrics is adequate for a technical paper since it uses the same evaluation formula for all structures.

Our corpora have been written in Japanese. We investigated a journal paper[2] in English too to confirm our metrics can be used in other languages. As a result, these features of the graph that is made from an English corpus were similar behavior with Japanese corpora. Therefore, we suppose that our metrics can be used by every language since it does not use the feature such as the syntactic role that is dependent on specific language but only sentence similarity.

IV. EXPERIMENT

We compare the performance of our metrics with that of the conventional graph-based coherence model. We evaluate the ability of them with two sentence ordering tasks: discrimination [4] and insertion [12]. Moreover, we permute paragraphs and sections in a document as well as these sentence ordering tasks. These tasks are the relative evaluation method since it is difficult to evaluate textual coherence absolutely.

[2]This paper is best of CHI 2016 that is selected in the top 1% total submissions by the separate Best Paper committee.

ISBN: 978-988-14048-1-7
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

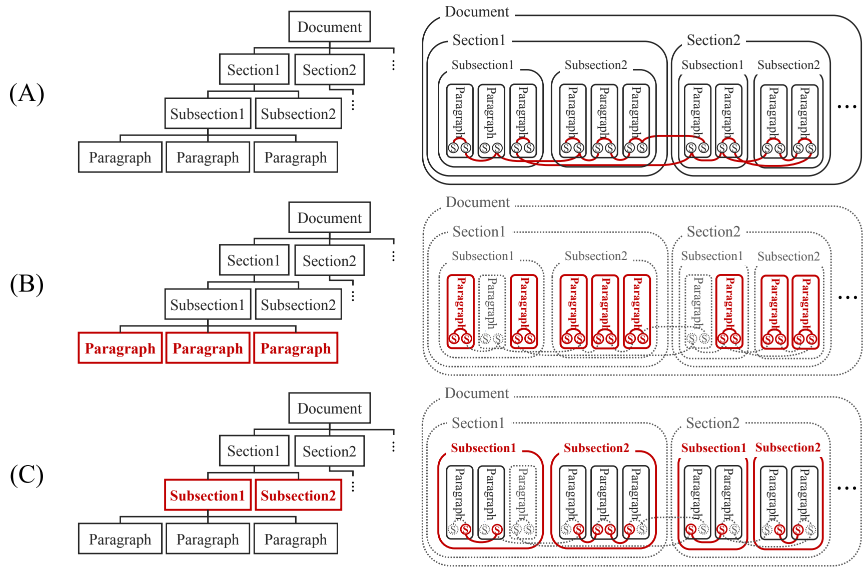

Fig. 1. Graph Considering Document Structure

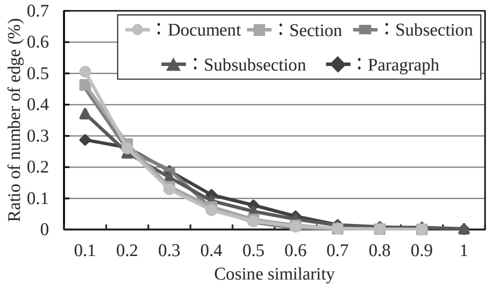

Fig. 2. Tendency of Cosine Similarity

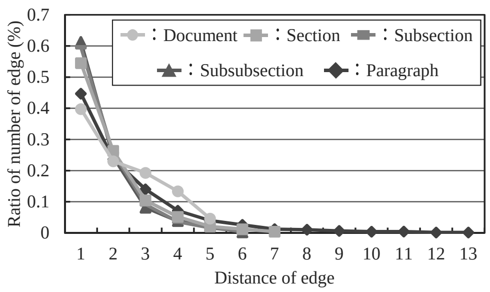

Fig. 3. Tendency of Distance

Discrimination consists of comparing a document to a random permutation of its sentences. We make 20 data that are random permutations according to the previous method. When the score of the original document is higher than the score of its permutation, an output is considered as correct.

Insertion is a method that evaluates the ability of our metrics to retrieve the original position of a sentence previously removed from a document. Here, each sentence is removed in turn and a coherence score is computed for every possible reinsertion position. When the document associated with the highest coherence score is the one in which the sentence is reinserted in the correct position, an output is considered as correct. In complement to accuracy, we use the insertion score introduced by Elsner and Charniak [12] for evaluation. This score -the higher, the better- is calculated by the proximity between the initial and the proposed position of

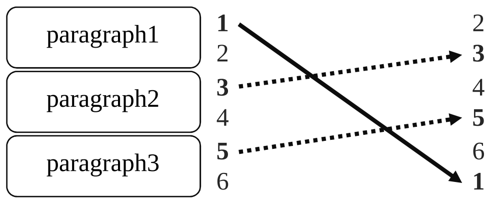

Fig. 4. Insert 1st sentence in 6th sentence position.

a sentence and assign the value that is reciprocal of it that is divided by the number of sentences to the sigmoid function. In this paper, Ins stand for insertion score. Formally, this score of a document D that has S sentences is equal to

$$InsertionScore(D) = \sigma(\frac{S}{\sum_{i=1}^{S} |c_i - m_i|}), \qquad (4)$$

$$\sigma(a) = \frac{1}{1 + exp(-a)},$$

where c is the initial position, m is the proposed position of a sentence.

It should be noted here that we repeat them tasks that are carried out in previous method fidelity. For example, Fig. 5 shows that when we insert 1st sentence of a document to 6th sentence position, paragraphs of a 3rd sentence and a 5th sentence are changed. Also, we carry out these experiments with these corpora that are same as it in the above section.

V. Results

A. Accuracy and Insertion Score

Table. IV, V, VI shows that all results of experiments. Here, values that are bold in these tables are the highest value in our metrics. We test the significance of these values and the output of the previous method by using the Student's t-test.

Table. IV shows that results of experiments for sentences. For discrimination, the accuracy of our metrics was comparable to the previous method when α is set other than

ISBN: 978-988-14048-1-7
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

TABLE IV
RESULT OF SENTENCE

	Discrimination Accuracy		Insertion Accuracy		Ins.	
	A	B	A	B	A	B
Weighted	1.0	1.0	0.099*	0.104†	0.509	0.508
Acc	1.0	1.0	0.084†	0.089	0.508	0.508
$\alpha = 0.1$	**1.0**	**0.993**	**0.039**	**0.046**	**0.508**	**0.507**
$\alpha = 0.5$	0.907	0.921	0.011	0.010	0.505	0.504
$\alpha = 0.9$	0.336	0.250	0	0	0.503	0.503

(†: $p < .10$, *: $p < .05$, **: $p < .01$)

TABLE V
RESULT OF PARAGRAPH

	Discrimination Accuracy		Insertion Accuracy		Ins.	
	A	B	A	B	A	B
Weighted	1.0*	1.0**	0.134†	0.125*	0.527†	0.524†
Acc	1.0*	1.0**	0.124†	0.145*	0.532*	0.524†
$\alpha = 0.1$	0.457	0.229	0.013	0.009	**0.517**	0.514
$\alpha = 0.5$	**0.650**	**0.400**	**0.041**	0.015	0.515	**0.517**
$\alpha = 0.9$	0.350	0.365	0.020	**0.029**	0.514	**0.517**

(†: $p < .10$, *: $p < .05$, **: $p < .01$)

0.9. However, when α is set 0.9, the accuracy of our metrics was significantly lower than that of the previous method. For insertion, the accuracy of our metrics was significantly lower than that of the previous method. However, the insertion score of our metrics and that of the previous method were no significant difference.

Table. V shows that results of experiments for paragraphs. For both discrimination and insertion, accuracies of our metrics were significantly lower than that of the previous method. However, the insertion scores of our metrics and the previous method were no significant difference.

Table. VI shows that results of experiments for sections. For discrimination, the accuracy of our metrics was higher than that of the previous method; however, they were no significant difference. For insertion, the accuracy of our metrics was higher than that of the previous method, and they were the significant difference in Data B. The insertion score of our output was higher than that of the previous method; however, they were no significant difference.

B. Value of Parameter α

We set three parameters as arbitrary values for our metrics: 0.1, 0.5, 0.9. We tested the significance of each accuracy in Table. IV, V of the value of α in sentences and paragraphs using Fisher's multiple comparisons by LSD method. Here, we did not test the significance of results of sections since Table. VI shows that it did not vary by each value of parameter α.

Table. VII shows that for discrimination in sentences, the accuracy of α that is set 0.1 or 0.5 was significantly higher than that of α that is set 0.9. Moreover, for insertion in sentences, the accuracy of α that is set 0.1 was significantly higher than that of α that is set 0.5 or 0.9. In other words, when α is set 0.1, the performance of our metrics works well. However, Table. VIII shows that there was no significant difference for each parameter for both discrimination and insertion in paragraphs.

TABLE VI
RESULT OF SECTION

	Discrimination Accuracy		Insertion Accuracy		Ins.	
	A	B	A	B	A	B
Weighted	0.650	0.443	0.286	0.137*	0.686	0.627
Acc	0.664	0.456	0.324	0.149*	0.668	0.620†
$\alpha = 0.1$	0.707	0.586	**0.425**	**0.348**	**0.741**	**0.706**
$\alpha = 0.5$	0.686	**0.650**	**0.425**	**0.348**	**0.741**	**0.706**
$\alpha = 0.9$	**0.721**	0.593	**0.425**	**0.348**	**0.741**	**0.706**

(†: $p < .10$, *: $p < .05$, **: $p < .01$)

TABLE VII
COMPARISON OF EACH ACCURACY OF PARAMETER α FOR SENTENCE

	Data A	Data B
Discrimination	0.1 = 0.5	0.1 = 0.5
	0.1 > 0.9*	0.1 > 0.9*
	0.5 > 0.9*	0.5 > 0.9*
Insertion	0.1 > 0.5*	0.1 > 0.5*
	0.1 > 0.9*	0.1 > 0.9*
	0.5 = 0.9	0.5 = 0.9

(*: $p < .05$, **: $p < .01$)

VI. DISCUSSION

A. Result of Discrimination

It is thought that the result of discrimination is due to the probability. Table. VI shows that the result of insertion was not varied by the value of parameter α. However, Table. VI shows that the accuracy of discrimination was varied by the value of α. This shows that these results are due to the probability, not because the value of α affect experiments. Moreover, corpora of the previous method are composed of 36.1 sentences on average; however, our corpora are composed of 4.9 times sentences on average. Thus, we should increase the number of trials to obtain an accurate result since discrimination task relies on the number of trials.

The tendency of the graph in Fig. 3 affects the result for discrimination. Table. IV, V shows that results of the previous method for discrimination in sentences and paragraphs was the best. The result is due to a tendency that is shown in Fig. 3. This is the tendency of the graph that is made by our method; however, the graph that is made by the previous method has the similar tendency. It is assumed that the coherence value become low since the probability that the number of edges whose distance is 1 decreases is high for discrimination task. Thus, the performance of the previous method works well.

However, our result is inferior in spite of that the graph that is made by our method have a similar tendency to the previous method. This is due to experiments without considering document structures. For example, Fig. 5 shows that when two paragraphs that have sentences that are designed by black are exchanged, the graph that is made in the same subsection vary to the graph that is made across subsections. It follows that the permutation of sentences affects the formation of the graph at every layer. Moreover, Fig. 5 shows that when the position of a sentence varies, structures in other two sentences vary too. It is thought that even if the coherence value in a target layer is low, the coherence value in an above layer from the target layer may be high. Thus, it is inferred

ISBN: 978-988-14048-1-7
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

TABLE VIII
COMPARISON OF EACH ACCURACY OF PARAMETER α FOR PARAGRAPH

	Data A	Data B
Discrimination	0.1 = 0.5	0.1 = 0.5
	0.1 = 0.9	0.1 = 0.9
	0.5 = 0.9	0.5 = 0.9
Insertion	0.1 = 0.5	0.1 = 0.5
	0.1 = 0.9	0.1 = 0.9
	0.5 = 0.9	0.5 = 0.9

(*: $p < .05$, **: $p < .01$)

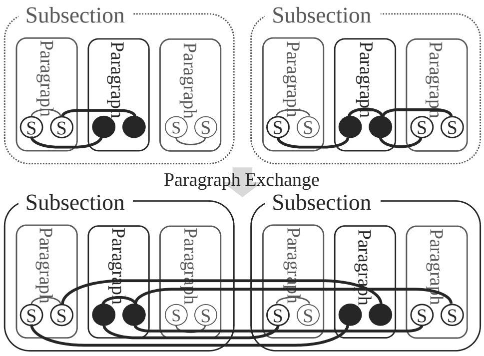

Fig. 5. The graph formation is varied by exchanging paragraphs.

that we do not obtain an accurate result, when we carry out the experiment without considering document structure.

B. Validity of Definition of Metrics

Table. VI shows that for insertion in sections, the accuracy and the insertion score were higher than that of the previous method. Moreover, the accuracy of our metrics in Data B was significantly higher than that of the previous method. We think that this result is due to the validity of our metrics. For the experiment in sections, the coherence value is calculated by varying only the distance of sections without considering other layers. It is assumed that the performance of our metrics for insertion in sections is better than the previous method since the insertion score is similar with that of the previous method, whose accuracy is lower than ours. Therefore, it is confirmed that our metrics that use the distance between structures instead of sentences are valid.

Table. IV, V shows that for insertion in sentences and paragraphs, the accuracy of our output was significantly lower than that of the previous method. We think that these results are due to the same problem as is the case with discrimination. However, the insertion score of our metrics was similar to that of the previous method. Thus, our metrics are not remarkably inferior to the previous method.

C. Validity of Parameter α

Table. VII shows that when we set 0.1 as a parameter α, the performance works well for all experiments in sentences. When we set 0.1 to α, our metrics computes the coherence that is prioritized a value of the under layer. Thus, it is supposed that the setting the parameter is valid.

However, Table. VIII shows that accuracies in paragraphs were no significant difference for both discrimination and insertion. The reason for this, it is thought that we decided the value of α as an arbitrary value. It is inferred that the values of α (0.1, 0.5, 0.9) are not optimum for the paragraph layer. We think that this is a case of that results for experiments in paragraphs are more inferior than that of the previous method. Therefore, we need to search optimum the value of α.

VII. CONCLUSION

In this paper, we proposed the metrics of textual coherence for a structured document. The proposed metrics is based on the insight that an easy-to-understand document tends to repeat similar writing method along the layered structure recursively in a bottom-up manner. Our method makes graph structures at every document structure and we capture coherence by features of the graph. The proposed metrics outperforms in particular at the section layer; it follows that the proposed metrics works well for a structured document as containing sections, paragraphs, such as a technical paper and a product manual. However, we think that experiments are carried out in this paper were not suited to evaluate our metrics. This is cause that the graph that is made by our method is affected by varying the structure of sentences.

In future work, we need to carry out the experiment by considering document structures. Moreover, we should search optimum the value of parameter α since we decided the value of α as arbitrary value. We make the graph with weights using by sentence similarity; however, we do not consider word similarity. It is important to consider word similarity to capture textual coherence [6]. Thus, we should need to incorporate word similarity with LSA for our calculation of cosine similarity between two sentences.

REFERENCES

[1] J. C. Richards, W. S. Richard. “Logman Dictionary of Language Teaching and Applied Linguistics”, *Routledge*, 1985.
[2] J. Guan. “Theme-Rheme Theory and the Textual Coherence of College Students’ English Writing,” *International Coherence on Economics, Social Science, Arts, Education and Management Engineering*, pp. 344-345, 2015.
[3] B. J. Grosz, S. Weinstein, and A. K. Joshi, “Centering: A framework for modeling the local coherence of discourse,” *Computational Linguistics*, vol.21, no. 2, pp. 203-225, 1995.
[4] R.Barzilay, M.Lapata. “Modeling local coherence: An entity-based approach,” *Computational Linguistics*, vol.34, no. 1, pp. 1-34, 2008.
[5] C. Guinaudeau, M. Strube. “Graph-based local coherence modeling,” *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, vol. 1, pp. 93-103, 2013.
[6] P. W. Foltz, W. Kintsch, and T. K. Landauer. “The measurement of textual coherence using with latent semantic analysis,” *Discourse Processes*, vol.25, no. 2-3, pp. 285-307, 1998.
[7] S. Deerwester, S. T. Dumais, G. W. Furnas, T. K. Landauer, and R. Harshman. “Indexing by latent semantic analysis,” *Journal of the American society for information science*, vol. 41, no. 6, pp. 391-407, 1990.
[8] T. K. Landauer, S. T. Dumais. “A solution to Plato’s problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge,” *Psychological Review*, vol. 104, no. 2, pp. 211-240, 1997.
[9] X. Hu, Z. Cai, P. Wiermer-Hastinfs, A. C. Graesser, and D. S. McNamara. “Strengths, limitations, and extensions of LSA,” *The handbook of latent semantic analysis*, pp. 401-426, 2007.
[10] G. Erkan, D. R. Radev, “Lexrank: Graph-based lexical centrality as salience in text summarization,” *Journal of artificial intelligence research*, vol. 22, pp. 457-479, 2004.
[11] A. Sattayatham, P. Ratanapin. “Analysis of errors in paragraph writing in English by first year medical students from the four medical schools at Mahidol University,” *Silpakorn University International Journal*, vol. 8, pp. 17-38, 2008.
[12] M. Elsner, E. Charniak. “Extending the entity grid with entity-specific features,” In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies: short papers-Volume 2*, Association for Computational Linguistics, pp. 125-129, 2011.

ISBN: 978-988-14048-1-7
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)