

## Japanese Sentence Order Estimation using Supervised Machine Learning with Rich Linguistic Clues

YUYA HAYASHI, MASAKI MURATA, LIANGLIANG FAN, AND  
MASATO TOKUHISA

*Tottori University, Japan*

### ABSTRACT

*Estimation of sentence order (sometimes referred to as sentence ordering) is one of the problems that arise in sentence generation and sentence correction. When generating a text that consists of multiple sentences, it is necessary to arrange the sentences in an appropriate order so that the text can be understood easily. In this study, we proposed a new method using supervised machine learning with rich linguistic clues for Japanese sentence order estimation. As one of rich linguistic clues we used concepts on old information and new information. In Japanese, we can detect phrases containing old/new information by using Japanese topic-marking postpositional particles. In the experiments of sentence order estimation, the accuracies of our proposed method (0.72 to 0.77) were higher than those of the probabilistic method based on an existing method (0.58 to 0.61). We examined features using experiments and clarified which feature was important for sentence order estimation. We found that the feature using concepts on old information and new information was the most important.*

**KEYWORDS:** *Sentence order estimation, supervised machine learning, linguistic clue, old / new information*

## 1 INTRODUCTION

Estimation of sentence order (sometimes referred to as sentence ordering) is one of the problems that arise on sentence generation and sentence correction [1–6]. When generating a text that consists of multiple sentences, it is necessary to arrange the sentences in an appropriate order so that the text can be understood easily.

Most of the studies on sentence order estimation were for multi document summarization, and they used the information obtained from the original sentences before summarizing for estimating sentence order [7–21]. If we can estimate sentence order without the original sentences before summarizing, the technique of estimating sentence order can be utilized for a lot of applications (e.g., sentence correction). For example, a text where the order of sentences is not good can be modified into a text where the order of sentences is good. Furthermore, the grammatical knowledge on sentence order will be able to be obtained through the study on sentence order without the original sentences. For example, when we find that a feature using a linguistic clue is important in the study on sentence order estimation, we can acquire the grammatical knowledge that the linguistic clue is important in sentence order estimation. Therefore, in this study, we handle the sentence order estimation that does not use the information on the original sentences before summarizing. In a study about sentence order estimation without using the original sentences before summarizing, Lapata proposed a probabilistic model [22]. However, supervised machine learning has not been used for that estimation. Therefore, in this study, we use supervised machine learning for sentence order estimation without using the original sentences before summarizing. In this study, we use the support vector machine (SVM) as the supervised machine learning [23].

We propose a method of sentence order estimation using numerous linguistic clues besides supervised machine learning. It is difficult for a probabilistic model to use a lot of information. In contrast, when using supervised learning, we can very easily use a lot of information by preparing many features. Because our proposed method uses a lot of information, it can be expected that our proposed method outperforms the existing method based on a probabilistic model.

In this paper, we use a simple task for sentence order estimation. We consider that the phenomenon across multiple paragraphs is complicated. We handle the problem where we judge which sentence we should write

first among two sentences in a paragraph using the information in the paragraph.<sup>1</sup>

In this study, we handle sentence order estimation in Japanese.

We present the main points of this study as follows:

1. Our study has originality, and used supervised machine learning for sentence order estimation with rich linguistic clues for the first time. As one of rich linguistic clues we used features based on concepts of old information and new information.
2. We confirmed that the accuracy rates of our proposed method using supervised machine learning (0.72 to 0.77) was higher than those of the existing methods based on a probabilistic model (0.58 to 0.61). Our proposed method has a high usability because the performance accuracy was high.
3. Our proposed method using supervised learning can use a lot of features (information) easily. It is expected that our method improves the performance by using more features.
4. In our proposed method using supervised learning, we can find important features (information) in sentence order estimation by examining features. When we examined features in our experiments, we found that the feature based on the concept of old/new information. The feature checked the number of common content words between the subject in the second sentence and the part after the subject in the first sentence is the most important in sentence order estimation.

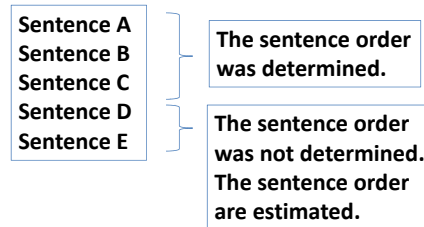
## 2 RELATED STUDIES

In a study [22] that is similar to ours, Lapata proposed a probabilistic model for sentence order estimation that did not use the original sentences before summarizing. Lapata calculated the probabilities of sentence occurrences using the probabilities of word occurrences, and estimated sentence orders by the probabilities of sentence occurrences.

Most of the studies on sentence order estimation are for multi document summarization, and they use the information obtained from the original sentences before summarizing for estimating sentence order [8, 9, 13, 19, 21]. Bollegala et al. performed sentence order estimation against the sentences that were extracted from multiple documents. They used

---

<sup>1</sup> An estimate of the order of all the sentences in a full text would be handled by combining estimated orders in pairs of two sentences.



**Fig. 1.** The model of the task

original documents before summarization for sentence order estimation. They focused on how the sentences, whose order would be estimated, were located in original documents before summarization. In addition, they used chronological information and topical-closeness. They used supervised machine learning for combining these kinds of information. However, they did not use linguistic clues such as POSs (parts of speech) of words and a concept on linguistic old/new information (related to subjects and Japanese postpositional particles) as features for machine learning.

Uchimoto et al. studied word order using supervised machine learning [24]. They used linguistic clues such as words and parts of speech as features for machine learning. They used machine learning for word order estimation. In contrast, we used machine learning for sentence order estimation. They estimated word order using word dependency information. Correct word orders are in corpora. Therefore, the training data on word order can be constructed from corpora automatically. In a similar way, the training data on sentence order can be constructed from corpora automatically. In our study, we use the training data that are constructed from corpora automatically.

### 3 THE TASK AND THE PROPOSED METHOD

#### 3.1 *The task*

The task in this study is as follows: a paragraph is input, the order of the first several sentences in the paragraph is determined, the order of the remaining sentences in the paragraph is not determined, and the estimation of the order of two sentences among the remaining sentences is the task. The information that can be used for estimation is the two sentences

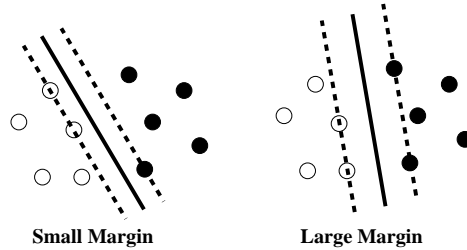


Fig. 2. Maximizing the margin

whose order will be estimated, and the sentences before one of the two sentences appears in the paragraph (see Figure 1).

### 3.2 Our proposed method

We assume that we need to estimate the order of two sentences,  $A$  and  $B$ . These sentences are input in the system and our method judges whether the order of “ $A$ - $B$ ” is correct by using supervised learning. In this study, we use SVM as machine learning. We use a quadratic polynomial kernel as a kernel function.

The training data is composed as follows: two sentences are extracted from a text that is used for training. From the two sentences, a sequence of the two sentences with the same order as in an original text, and a sequence of the two sentences with the reverse order are made. The two sentences with the same order are used as a positive example, and the two sentences with the reverse order are used as a negative example.

### 3.3 Support vector machine method

In this method, data consisting of two categories is classified by dividing space with a hyperplane. When the margin between examples which belong to one category and examples which belong to the other category in the training data is larger (see Figure 2<sup>2</sup>), the probability of incorrectly choosing categories in open data is thought to be smaller. The hyperplane

<sup>2</sup> In the figure, the white circles and black circles indicate examples which belong to one category and examples which belong to the other category, respectively. The solid line indicates the hyperplane dividing space, and the broken lines indicate planes at the boundaries of the margin regions.

maximizing the margin is determined, and classification is done by using this hyperplane. Although the basics of the method are as described above, for extended versions of the method, in general, the inner region of the margin in the training data can include a small number of examples, and the linearity of the hyperplane is changed to non-linearity by using kernel functions. Classification in the extended methods is equivalent to classification using the following discernment function, and the two categories can be classified on the basis of whether the output value of the function is positive or negative [23, 25]:

$$f(\mathbf{x}) = \text{sgn} \left( \sum_{i=1}^l \alpha_i y_i K(\mathbf{x}_i, \mathbf{x}) + b \right) \quad (1)$$

$$b = - \frac{\max_{i, y_i = -1} b_i + \min_{i, y_i = 1} b_i}{2}$$

$$b_i = \sum_{j=1}^l \alpha_j y_j K(\mathbf{x}_j, \mathbf{x}_i),$$

where  $\mathbf{x}$  is the context (a set of features) of an input example;  $\mathbf{x}_i$  and  $y_i$  ( $i = 1, \dots, l, y_i \in \{1, -1\}$ ) indicate the context of the training data and its category, respectively; and the function  $\text{sgn}$  is defined as

$$\text{sgn}(x) = \begin{cases} 1 & (x \geq 0), \\ -1 & (\text{otherwise}). \end{cases} \quad (2)$$

Each  $\alpha_i$  ( $i = 1, 2, \dots$ ) is fixed when the value of  $L(\alpha)$  in Equation (3) is maximum under the conditions of Equations (4) and (5).

$$L(\alpha) = \sum_{i=1}^l \alpha_i - \frac{1}{2} \sum_{i,j=1}^l \alpha_i \alpha_j y_i y_j K(\mathbf{x}_i, \mathbf{x}_j) \quad (3)$$

$$0 \leq \alpha_i \leq C \quad (i = 1, \dots, l) \quad (4)$$

$$\sum_{i=1}^l \alpha_i y_i = 0 \quad (5)$$

Although the function  $K$  is called a kernel function and various types of kernel functions can be used, this paper uses a polynomial function as follows:

$$K(\mathbf{x}, \mathbf{y}) = (\mathbf{x} \cdot \mathbf{y} + 1)^d, \quad (6)$$

where  $C$  and  $d$  are constants set by experimentation. In this paper,  $C$  and  $d$  are fixed as 1 and 2 for all experiments, respectively.<sup>3</sup> A set of  $\mathbf{x}_i$  that satisfies  $\alpha_i > 0$  is called a support vector, and the portion used to perform the sum in Equation (1) is calculated by only using examples that are support vectors.

We used the software TinySVM [25] developed by Kudoh as the support vector machine.

### 3.4 Features used in our proposed method

In this section, we explain features (information used in classification), which are required to use machine learning methods.

Features used in this study are shown in Table 1. Each feature has additional information of whether it appears in the first or second sentence. The first and the second sentence that are input are indicated with  $A$  and  $B$ , respectively.

Concretely speaking, we used a topic instead of a subject for F9. The part before a Japanese postpositional particle *wa* indicates a topic. We used the number of the common content words between the part before *wa* in the second sentence  $B$  and the part after *wa* in the first sentence for F9.

F9 is a feature based on a concept of old/new information. Because the part before a Japanese postpositional particle *wa* indicates a topic, it is likely to contain old information and the part after a Japanese postpositional particle *wa* is likely to contain new information. A Japanese postpositional particle *wa* in “Noun X *wa*” is similar to an English prepositional phrase “in terms of” in “in terms of Noun X” and indicates that “Noun X” is a topic. In correct sentence order, words in a part containing old information of the second sentence are likely to appear in a part containing new information of the first sentence. Based the above idea, we used F9.

<sup>3</sup> We confirmed that  $d = 2$  produced good performance in preliminary experiments.

**Table 1.** Feature

ID	Definition
F1	The words and their parts of speech (POS) in the sentence $A$ (or $B$ ).
F2	The POS of the words in the sentence $A$ (or $B$ ).
F3	Whether the subject is omitted in the sentence $A$ (or $B$ ).
F4	Whether a nominal is at the end of the sentence $A$ (or $B$ ).
F5	The words and their POS in the subject of the sentence $A$ (or $B$ ).
F6	The words and their POS in the part after the subject in the sentence $A$ (or $B$ ).
F7	The pair of the postpositional particles in the two sentences $A$ and $B$ .
F8	The number of common content words between the two sentences $A$ and $B$ .
F9	The number of common content words between the subject in the second sentence $B$ and the part after the subject in the first sentence $A$ .
F10	The words and their POS in all the sentences before the two sentences $A$ and $B$ in the paragraph.
F11	Whether a nominal is at the end of the sentence just before the two sentences $A$ and $B$ in the paragraph.
F12	Whether the subject is omitted in the sentence just before the two sentences $A$ and $B$ in the paragraph.
F13	The number of the common content words between the sentence just before the two sentences $A$ and $B$ in the paragraph and the sentence $A$ (or $B$ ).

#### 4 PROBABILISTIC METHOD (COMPARED METHOD)

We compare our proposed method based on machine learning with the probabilistic method. Here, the probabilistic method is based on Lapata's method using probabilistic models [22].

The detail of the probabilistic method is as follows: words that appear in two adjacent sentences are extracted from a text that is used for calculating probabilities. All the pairs of a word  $W_A$  in the first sentence, and a word  $W_B$  in the second sentence are made. Then the occurrence probability that when a word  $W_A$  appears in a first sentence, a word  $W_B$  appears in a second sentence is calculated for each word pair. The occurrence probability (that we call sentence occurrence probability) that the second sentence appears when the first sentence is given is calculated by multiplying the probabilities of all the word pairs. In this study, to estimate the order for two sentences  $A$  and  $B$ , a pair  $Pair_{AB}$  with the original order ( $A$ - $B$ ) and a pair  $Pair_{BA}$  with the reverse order ( $B$ - $A$ ) are generated. When the sentence occurrence probability of  $Pair_{AB}$  is



**Table 2.** The number of pairs of two sentences

	CASE1	CASE2	CASE3
Training data	33902	64290	130316
Test data	40386	82966	170376

larger than that of  $Pair_{BA}$ , the method judges that the order of  $Pair_{AB}$  is correct. Otherwise, it judges that the order of  $Pair_{BA}$  is correct.

$a_{\langle i,1 \rangle}, \dots, a_{\langle i,n \rangle}$  indicate to the words that appear in a sentence  $S_i$ . The probability that  $a_{\langle i,j \rangle}$  and  $a_{\langle i-1,k \rangle}$  appear in the two adjacent sentences are expressed in the following equation: equation:

$$P(a_{\langle i,j \rangle} | a_{\langle i-1,k \rangle}) = \frac{f(a_{\langle i,j \rangle}, a_{\langle i-1,k \rangle})}{\sum_{a_{\langle i,j \rangle}} f(a_{\langle i,j \rangle}, a_{\langle i-1,k \rangle})} \quad (7)$$

$f(a_{\langle i,j \rangle}, a_{\langle i-1,k \rangle})$  is the frequency that a word  $a_{\langle i,j \rangle}$  appears in the sentence just after the sentence having a word  $a_{\langle i-1,k \rangle}$ .

When there is a sentence  $C$  just before sentences whose order will be estimated, the sentence occurrence probability of  $Pair_{AB}$  is multiplied by the sentence occurrence probability of sentence  $A$  appearing just after sentence  $C$ .

## 5 EXPERIMENT

### 5.1 Experimental condition

We used Mainichi newspaper articles (May, 1991) for the machine learning of the training data. We used Mainichi newspaper articles (November, 1995) for the test data. We used Mainichi newspaper articles (1995) for the text that is used for calculating probabilities in the probabilistic method.

We used the following three kinds of cases for pairs of two sentences used in the experiments: CASE 1: We made pairs of two sentences by using only the first two sentences in a paragraph. CASE 2: We made pairs of two sentences by using all the adjacent two sentences in a paragraph. CASE 3: We made pairs of two sentences by using all the two sentence combinations in a paragraph.

The number of pairs of two sentences used in the training and test data are shown in Table 2.

**Table 3.** Accuracy

Machine learning (ML)			Probabilistic method (PM)		
CASE1	CASE2	CASE3	CASE1	CASE2	CASE3
0.7677	0.7246	0.7250	0.6059	0.5835	0.5775

**Table 4.** Comparison with accuracies of human subjects

	Subjects					Ave.	ML	PM
	A	B	C	D	E			
CASE1	0.75	0.70	0.75	0.95	0.95	0.82	0.79	0.65
CASE2	0.80	0.80	0.85	1.00	0.90	0.87	0.67	0.64
CASE3	0.65	0.75	0.85	0.65	0.70	0.72	0.71	0.56

## 5.2 Experimental results

The accuracies of our proposed method and the probabilistic method are shown in Table 3. As shown in Table 3, the accuracies of our proposed method (0.72 to 0.77) were higher than those of the probabilistic method (0.58 to 0.61).

## 5.3 Comparison with accuracies of manual sentence order estimation

We randomly extracted 100 pairs (each pair consists of two sentences) from Mainichi newspaper articles (November, 1995), and each of the five subjects estimated the order of 20 pairs among the 100 pairs for each of the CASEs 1 to 3. Our proposed method (ML) and the probabilistic method (PM) estimated the orders of 100 pairs. In CASE 2 and CASE 3, because the information on sentences was used in the supervised learning and the probabilistic methods, the sentences before two sentences whose orders will be estimated are shown to subjects.

Accuracies of subjects, ML, and PM are shown in Table 4. “A” to “E” in the table indicate the five subjects. “Average” indicates the average of accuracies of the five subjects.

When we compared the average accuracies of the subjects, and the accuracy of our proposed method (ML) in Table 4, we found that our proposed method could obtain accuracies that were very similar to the average accuracies of the subjects in CASEs 1 and 3.

**Table 5.** Accuracies of eliminating a feature

Eliminated feature	Accuracy	Difference
F1	0.7211	-0.0039
F2	0.7226	-0.0024
F3	0.7251	+0.0001
F4	0.7251	+0.0001
F5	0.7212	-0.0038
F6	0.7223	-0.0027
F7	0.7243	-0.0007
F8	0.7201	-0.0049
F9	<u>0.6587</u>	<u>-0.0663</u>
F10	0.7172	-0.0078
F11	0.7240	-0.0010
F12	0.7241	-0.0009
F13	0.7241	-0.0009

#### 5.4 Analysis of features

Among the features used in this study, we examined which feature was useful for sentence order estimation. We compared accuracies of eliminating a feature and the accuracy of using all the features in CASE 3. Table 5 shows the accuracies of eliminating a feature. It also shows the result of subtracting the accuracy using all the features from the accuracies after eliminating a feature.

From Table 5, we found that the accuracy went down heavily without feature F9. We found that feature F9 was particularly important in sentence order estimation. An example that the estimation succeeds when using F9 and the estimation fails when not using F9 is shown as follows:

Sentence 1:

*kotani-san-niwa hotondo chichi-no kioku-ga nai.*  
 (Kotani) (almost) (father) (recollection) (no)  
 (Kotani has very few recollection of his father.)

Sentence 2:

*chichi-ga byoushi-shita-no wa gosai-no toki-datta.*  
 (father) (died of a disease) (five years old) (was when)  
 (The time that his father died of a disease was when he was five years old.)

The correct order is “Sentence 1 to Sentence 2.” No use of F9 estimated that the order was “Sentence 2 to Sentence 1.” F9 is the feature

that checks the number of common content words between the subject in the second sentence and the part after the subject in the first sentence. Because “*chichi*” (father) appeared at the subject in the second sentence and the part after the subject in the first sentence, the use of F9 could estimate the correct order of the above example.

F9 is based on concepts of old/new information. In our method, we obtained good results on sentence order estimation by using the feature (F9) based on concepts of old/new information. A Japanese word *wa* in the phrase *byoushi-shita-no wa* (died of a disease) is a postpositional particle indicating a topic. A phrase *chichi-ga byoushi-shita-no wa* (father, died of a disease) is a topic part indicated by *wa* and corresponds to old information. Old information must appear in a previous part. “*chichi*” (father) appearing in a phrase corresponding to old information of Sentence 2 appears in Sentence 1. Therefore, the sentence order of “Sentence 1 to Sentence 2” is good. Our method using F9 can handle the concepts of old/new information and accurately judge the sentence order of the above example.

## 6 CONCLUSION

In this study, we proposed a new method of using supervised machine learning for sentence order estimation. In the experiments of sentence order estimation, the accuracies of our proposed method (0.72 to 0.77) were higher than those of the probabilistic method based on an existing method (0.58 to 0.61). When examining features, we found that the feature that checked the number of common content words between the subject in the second sentence, and the part after the subject in the first sentence was the most important in sentence order estimation. The feature is based on concepts of old/new information.

In the future, we would like to improve the performance of our method by using more features for machine learning. Furthermore, we would like to detect more useful features in addition to the feature based on concepts of old/new information. Useful detected features can be used as grammatical knowledge for sentence generation.

In this study, we handled the information within a paragraph. However, we should use information outside a paragraph when we handle orders of sentences in a full text. We should also consider sentence order estimation of two sentences across multiple paragraphs and estimation of the order of paragraphs. In the future, we would like to handle such things.

ACKNOWLEDGMENTS This work was supported by JSPS KAKENHI Grant Number 2350 0178.

## REFERENCES

1. Duboue, P.A., McKeown, K.R.: Content planner construction via evolutionary algorithms and a corpus-based fitness function. In: Proceedings of the second International Natural Language Generation Conference (INLG '02). (2002) 89–96
2. Karamanis, N., Manurung, H.M.: Stochastic text structuring using the principle of continuity. In: Proceedings of the second International Natural Language Generation Conference (INLG '02). (2002) 81–88
3. Mann, W.C., Thompson, S.A.: Rhetorical structure theory: Toward a functional theory of text organization. *Text* **8** (1988) 243–281
4. Marcu, D.: From local to global coherence: A bottom-up approach to text planning. In: Proceedings of the 14th National Conference on Artificial Intelligence. (1997) 629–635
5. Marcu, D.: The rhetorical parsing of unrestricted texts: A surface-based approach. *Computational Linguistics* **26** (2000) 395–448
6. Murata, M., Isahara, H.: Automatic detection of mis-spelled japanese expressions using a new method for automatic extraction of negative examples based on positive examples. *IEICE Transactions on Information and Systems* **E85-D** (2002) 1416–1424
7. Barzilay, R., Elhadad, N., McKeown, K.R.: Inferring strategies for sentence ordering in multidocument news summarization. *Journal of Artificial Intelligence Research* **17** (2002) 35–55
8. Barzilay, R., Lee, L.: Catching the drift: Probabilistic content models, with applications to generation and summarization. In: Proceedings of HLT-NAACL 2004. (2004) 113–120
9. Bollegala, D., Okazaki, N., Ishizuka, M.: A bottom-up approach to sentence ordering for multi-document summarization. In: Proceedings of the 44th Annual Meeting of the Association of Computational Linguistics. (2006) 385–392
10. Carbonell, J., Goldstein, J.: The use of MMR, diversity-based reranking for reordering documents and producing summaries. In: Proceedings of the 21st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval. (1998) 335–336
11. Duboue, P.A., McKeown, K.R.: Empirically estimating order constraints for content planning in generation. In: Proceedings of the 39th Annual Meeting of the Association for Computational Linguistics. (2001) 172–179
12. Elhadad, N., Mckeown, K.R.: Towards generating patient specific summaries of medical articles. In: Proceedings of the NAACL 2001 Workshop on Automatic Summarization. (2001)

13. Ji, P.D., Pulman, S.: Sentence ordering with manifold-based classification in multi-document summarization. In: Proceedings of Empirical Methods in Natural Language Processing. (2006) 526–533
14. Karamanis, N., Mellish, C.: Using a corpus of sentence orderings defined by many experts to evaluate metrics of coherence for text structuring. In: Proceedings of the 10th European Workshop on Natural Language Generation. (2005) 174–179
15. Madnani, N., Passonneau, R., Ayan, N.F., Conroy, J.M., Dorr, B.J., Klavans, J.L., O’Leary, D.P., Schlesinger, J.D.: Measuring variability in sentence ordering for news summarization. In: Proceedings of the 11th European Workshop on Natural Language Generation. (2007) 81–88
16. Mani, I., Schiffman, B., Zhang, J.: Inferring temporal ordering of events in news. In: Proceedings of North American Chapter of the ACL on Human Language Technology (HLT-NAACL 2003). (2003) 55–57
17. Mani, I., Wilson, G.: Robust temporal processing of news. In: The 38th Annual Meeting of the Association for Computational Linguistics. (2000) 69–76
18. McKeown, K.R., Klavans, J.L., Hatzivassiloglou, V., Barzilay, R., Eskin, E.: Towards multidocument summarization by reformulation: Progress and prospects. In: Proceedings of AAAI/IAAI. (1999) 453–460
19. Okazaki, N., Matsuo, Y., Ishizuka, M.: Improving chronological sentence ordering by precedence relation. In: Proceedings of the 20th International Conference on Computational Linguistics (COLING 04). (2004) 750–756
20. Radev, D.R., McKeown, K.R.: Generating natural language summaries from multiple on-line sources. *Computational Linguistics* **24** (1999) 469–500
21. Zhang, R., Li, W., Lu, Q.: Sentence ordering with event-enriched semantics and two-layered clustering for multi-document news summarization. In: Proceedings of COLING 2010. (2010) 1489–1497
22. Lapata, M.: Probabilistic text structuring: Experiments with sentence ordering. In: Proceedings of the 41st Annual Meeting of the Association of Computational Linguistics. (2003) 542–552
23. Cristianini, N., Shawe-Taylor, J.: An Introduction to Support Vector Machines and Other Kernel-based Learning Methods. Cambridge University Press (2000)
24. Uchimoto, K., Murata, M., Ma, Q., Sekine, S., Isahara, H.: Word order acquisition from corpora. In: Proceedings of COLING 2000. (2000) 871–877
25. Kudoh, T.: TinySVM: Support Vector Machines. <http://cl.aist-nara.ac.jp/taku-ku/software/TinySVM/index.html> (2000)

**YUYA HAYASHI**  
TOTTORI UNIVERSITY,  
4-101 KOYAMA-MINAMI, TOTTORI 680-8552, JAPAN  
E-MAIL: <S082043@IKE.TOTTORI-U.AC.JP>

**MASAKI MURATA**  
TOTTORI UNIVERSITY,  
4-101 KOYAMA-MINAMI, TOTTORI 680-8552, JAPAN  
E-MAIL: <MURATA@IKE.TOTTORI-U.AC.JP>

**LIANGLIANG FAN**  
TOTTORI UNIVERSITY,  
4-101 KOYAMA-MINAMI, TOTTORI 680-8552, JAPAN  
E-MAIL: <K112001@IKE.TOTTORI-U.AC.JP>

**MASATO TOKUHISA**  
TOTTORI UNIVERSITY,  
4-101 KOYAMA-MINAMI, TOTTORI 680-8552, JAPAN  
E-MAIL: <TOKUHISA@IKE.TOTTORI-U.AC.JP>