Automatic Detection and Manual Analysis of Inadequate Descriptions in a Thesis

Takuma Okada, Masaki Murata
Department of Information and Electronics
Tottori University
4-101 Koyama-Minami, Tottori 680-8552, Japan
Email: {s112011,murata}@ike.tottori-u.ac.jp

Qing Ma
Department of Applied Mathematics and Informatics
Ryukoku University
Seta, Otsu, Shiga 520-2194, Japan
Email: qma@math.ryukoku.ac.jp

***Abstract*—Some scholarly articles contain inadequate descriptions and are therefore likely to be difficult to read and understand. In this paper, we refer to information that must be described in a thesis as an "item requiring mention (IRM)," and we propose methods for detecting inadequate descriptions of IRMs in a paper using rules and machine learning. In our experimental results, the F-measure of the rule-based method was the highest for all IRMs. The F-measures for the IRMs, "comparison," "problem," and "purpose," were 0.6 to 0.7. The F-measure for the IRM, "example," was 0.86. We performed analysis to support automatic correction of inadequate descriptions. We were able to gather description patterns that are useful to consult when writing adequate descriptions. We also propose a method of using extracted patterns for a system that supports writing papers.**

I. Introduction

Some research papers lack some of the information requiring description, such as the research results or the necessity and effectiveness of the study. When the required information is not described in a paper, it becomes difficult for a reader to understand the contents of the study.

In this research, we refer to the information required in a thesis as an item requiring mention (IRM), and support writing papers by automatically determining whether an IRM is described.

In this study, we used rule-based and machine learning methods to automatically detect papers in which an IRM is not described. Furthermore, we manually analyzed sentences in a paper for automatic correction of a paper in which an IRM is not described. We conducted our study in Japanese.

The characteristics of our study are as follows:

- We examined papers using frequency and word meaning and determined six types of IRM, "necessity," "originality," "comparison," "problem," "purpose," and "example."
- In our experiments, we were able to extract papers with inadequate descriptions automatically for the four types of IRM,"comparison," "problem," "purpose," and "example," with F-measures of 0.6 to 0.8.
- We gathered description patterns that would be useful to consult when writing adequate descriptions by manually categorizing papers into five levels of descriptive quality. We propose a method of using the extracted patterns for a system that supports writing papers.

II. Related studies

Fukuda et al. [1] extracted and visualized technical trend information from research papers, and Ptaszynski et al. [2] developed a system to support writing research papers. Fukuda et al. constructed a system that extracts expressions indicating the effects of technology from research papers and visualized the extracted information. Ptaszynski et al. developed a system for the support of research and the writing of research papers. Their system prepares the data for experiments, automatically performs the experiments, and obtains accuracies. It creates tables in a LaTex template containing all the results and draws graphs showing the results. Their study is useful for surveying and preparing research papers. However, their studies were not intended to support the correction of inadequate parts of sentences in a paper. In contrast, our study provides support for correcting inadequate parts of a paper. This is the difference between their study and ours.

A study by Nadamoto et al. proposes a technique that recognizes missing parts [3]. In their paper, they observed that discussions in community-type contents, such as SNS or blogs, can concentrate on a small domain and thereby miss some viewpoints. They call a missed viewpoint a *content hole* and proposed a method of detecting such occurrences by comparing discussions in community-type contents and general information, such as what appears in Wikipedia. With respect to detecting missed information, Nadamoto et al.'s study and our study are similar. However, they differ in that Nadamoto et al.'s study deals with community-type contents, and ours deals with academic papers.

Tsudo et al. [4] conducted a study of automatically detecting redundant sentences to support the writing of sentences. They proposed a method of using machine learning to detect redundant sentences automatically. Many other studies to support writing sentences have also been conducted. However, among studies to support writing sentences, there are no papers to support writing sentences using IRMs.

III. Determination of IRMs and words useful for detecting them

A. Task

We determine IRMs and words useful for detecting them. When a paper does not have any word useful for detecting

(Category of quantity)		
	{Category of amount }	*shucyuryoku* (output), *nyuuryoku* (input), *sousuu* (total), *suuchi* (numerical), *hindo* (frequency), *bangou* (number), *kansuu* (function)
	{Category of number }	*ooku* (most), *tasuu* (a lot), *ooi* (many), *tairyou* (enormous), *juubun* (adequate), *sukunai* (a few)
	{Category of price }	*nagai* (long), *mijikai* (short), *shakudo* (degree), *takai* (high), *hikui* (low), *fukai* (deep), *chikai* (near), *kyori* (distance)
(Category of relation)		
	{Category of cause and effect }	*jouken* (condition), *yuukou* (effective), *zentei* (presupposition), *gen'in* (cause), *youin* factor, *kekka* result, *kouka* (effect), *eikyou* (influence)
	{Category of reason etc. }	*riyuu* (reason), *mokuteki* (purpose), *jitsuyou* (practical)
	{Category of difference }	*soutai* (relative), *sougo* (mutual), *oujiru* (accept), *taiou* (handling), *soutou* (correspond), *kuraberu* (compare), *hikaku* (comparison)
	{Category of relativeness }	*onaji* (same), *niru* (resemble), *douyou* (similarly), *ruiji* (similarity), *kotonaru* (different), *fukumu* (contain), *fukumeru* (include), *chigai* (difference), *kubetsu* (distinction)
	{Category of presence }	*sonzai* (existence), *kison* (already exist)
	{Category of appearance }	*arawareru* (appear), *shutsugen* (appearance), *jitsugen* realization, *teian* (proposition), *teiji* (present), *shimesu* (show), *dasu* (output)

Fig. 1: Partial meaning sort results

IRMs, we can judge that the paper does not contain an IRM.

B. Procedure of the determination

We determine IRMs and words useful for detecting them using the following procedure.

1) We extract words appearing in many papers (Section III-B1).
2) We display words similar to the extracted words using a meaning sort [5] (Section III-B2).
3) By manually examining the results in 2), we determine IRMs and words useful for detecting them. (Section III-B3)

We explain the details of the above procedure below.

1) Investigation of frequency: Words appearing in many papers would be likely to be IRMs. We calculate the occurrence ratio of a word by dividing the number of papers in which the word appears by the total number of papers. For example, when a word "Z" appears in 250 of 300 papers, the occurrence ratio of the word "Z" is 250/300.

2) Meaning sort: Words appearing in many papers would be likely to be words useful for detecting IRMs. Words similar to words appearing in many papers would also be likely to be words useful for detecting IRMs. For example, when a word "different" is useful for detecting IRMs, words such as "difference" would also be useful for that purpose. In this paper, we use meaning sort [5] to extract words similar to words useful for detecting IRMs. Meaning sort can display a table in which similar words are located near each other. Using meaning sort we can use words whose occurrence ratios are low as words useful for detecting IRMs.

3) Manual examination: Consulting the meaning sort results of Section III-B2, we manually examine and determine IRMs and words useful for detecting them.

TABLE I: Occurrence ratios of words appearing in papers

Word	Occurrence Ratio
hitsuyou (necessity)	0.994
juuyou (importance)	0.811
kotonaru (different)	0.951
chigau (differ)	0.631
hikaku (comparison)	0.895
kuraberu (compare)	0.727
tatoeba (for example)	0.858
mokuteki (purpose)	0.773
mondai (problem)	0.931

C. Data

We use 393 papers in Journal of Natural Language Processing (Japan, 1994 to 2013 yearly editions) in experiments determining IRMs.

D. Results of determination

1) Results of frequency investigation: Results of frequency investigation: In this paper, we performed a frequency investigation by using the method described in Section III-B1. Words whose occurrence ratios are high and that are manually judged as important expressions among the results of frequency investigation are shown in Table I.

2) Results of meaning sort: We displayed the top 500 words with the highest occurrence ratios to ensure that similar words are located near each other by using meaning sort. A part of the meaning sort results is shown in Figure 1.

3) Results of manual determination of IRMs and words useful for detecting them: From the results in Section III-D1, we found that words, such as *hitsuyou* (necessity), which can indicate necessity and effectiveness, and *kotonaru* (different), which can indicate originality, have high occurrence rates. Because it is difficult to understand the content of a paper

TABLE II: IRMs and words useful for detecting them

IRMs	Words useful for detection	Definition of IRMs
Necessity	*hitsuyou* (necessity) *juuyou* (importance)	The necessity of a study
Originality	*kotonaru* (different) *chigau* (differ) *chigai* (difference)	Originality of a study
Comparison	*hikaku* (comparison) *kuraberu* (compare)	Comparison between current and previous studies, or comparison among experimental results
Problem	*mondai* (problem)	Problems in the world (background of the study), or problems in previous studies
Purpose	*mokuteki* (purpose) *mokuhyou* (goal) *mezasu* (aim at)	The purpose for which the study is performed
Example	*tatoeba* (for example) *rei* (example) *gutai* (concrete)	Concrete examples

whose necessity and originality are not described, *hitsuyou* (necessity) and *kotonaru* (different) can be considered as IRMs.

As in the above, it is difficult to understand the problems and the purpose of a paper whose *mondai* (problems) and *mokuteki* (purpose) are not described, and furthermore, a paper where *tatoeba* (for example) is not described is likely to have no examples, and it can be difficult to understand its contents. Therefore, words such as *mondai* (problems), *mokuteki* (purpose), and *tatoeba* (for example) can be considered as IRMs.

By manually comparing the words that can be considered as IRMs in the above and the meaning sort results described in Section III-B2, we examined words similar to them and determined IRMs and words useful for detecting them. The results are shown in Table II. A paper that does not have any of the words useful for detecting IRMs will probably be a paper that does not have IRMs.

TABLE III: Number of data items in the 266 papers (2011 edition)

Items	Papers which an IRM is not described	Papers which an IRM is described	Total
Comparison	53	213	266
Problem	73	193	266
Purpose	83	183	266
Example	7	259	266

TABLE IV: Number of data items in the 305 papers (2012 edition)

Items	Papers which an IRM is not described	Papers which an IRM is described	Total
Comparison	59	246	305
Problem	114	191	305
Purpose	94	211	305
Example	9	296	305

IV. Automatic detection of IRMs using a rule-based method and a machine learning method

A. Task

The input to the system is a paper. The system judges whether an IRM is described in the paper or not. When the system can correctly detect a paper in which an IRM is not described, the detection is considered to be useful for the support of paper writing.

B. Proposed method

In this paper, we propose two types of methods, a rule-based method and a machine learning method.

1) Rule-based method: The rule-based method automatically judges that a paper in which a word that is useful for detection of an IRM does not appear is a paper in which an IRM is not described. Words useful for detection of an IRM are shown in Table II.

2) Machine learning method: In the machine learning method, we use training data containing papers in which an IRM is described and papers in which an IRM is not described. The machine learning method can categorize a new paper into a paper in which an IRM is described or a paper in which an IRM is not described by using training data. In this study, we use the maximum entropy method [6] for machine learning.[1] We use all words appearing in a paper and words used in the rule-based method as features used in learning. In machine learning, when the number of data items for each category is very different, the performance is likely to be low. Therefore, in this study, we decrease the number of data items of a category to ensure that it is equal to the number of data items of the other category.

3) Data: We used 266 papers from the 2011 annual meeting and 305 papers from the 2012 annual meeting of the association for natural language processing in Japan as training data and test data, respectively. The number of data items is shown in Tables III and IV.

4) Procedure of evaluation: We manually judge that when a paper judged by a system to be a paper in which an IRM is not described is, in fact, a paper in which an IRM is not described, the output of the system is considered to be correct. We use recall rates, precision rates, and F-measures in the evaluation. We calculate the F-measure of extracting a paper in which an IRM is not described.

5) Criterion of manual judgment: We constructed criteria of manual judgment to ensure that a judge does not hesitate on performing a manual judgment. By examining the results

[1]The maximum entropy method was a well-known machine learning technique that presents good performance in many studies as well as a support vector machine [7]. In the experiments using the data set in this paper, the maximum entropy method obtained a higher performance than the support vector machine.

TABLE V: Criteria of manual judgment (partial)

Items	Judgment	Criterion
Comparison	Useful	A paper in which a comparison between the current study and a previous study or a comparison of results among plural methods in experiments is not described.
	Not useful	A paper in which a comparison between the current study and a previous study is described and has a description such as "A previous study proposed Method A. In contrast, we ..."
Problems	Useful	Papers in which explanation of problems in the world (the background of a study) and problems in a previous study is inadequate or unclear.
	Not useful	Papers in which explanation of the background of a study and problems in a previous study is described in detail and clearly.
Purpose	Useful	Papers in which we cannot understand the purpose of the study (the reason why the study is performed) unless we read a paper carefully.
	Not useful	Papers in which we can understand the purpose of a study if we read a paper once.
Example	Useful	Papers that do not include examples (including ones in figures etc.)
	Not useful	Papers that include examples (including ones in figures etc.)

TABLE VI: F-measures detecting inadequate descriptions for "comparison"

Method	Recall	Precision	F-measure
Baseline	1.00 (59/59)	0.19 (59/305)	0.32
Rule-based	0.58 (34/59)	0.60 (34/ 57)	0.59
Machine learning	0.61 (36/59)	0.21 (36/174)	0.31

TABLE VII: F-measures detecting inadequate descriptions in "problem"

Method	Recall	Precision	F-measure
Baseline	1.00 (114/114)	0.37 (114/305)	0.54
Rule-based	0.61 (70/114)	0.81 (70/ 86)	0.70
Machine learning	0.69 (79/114)	0.47 (79/169)	0.56

TABLE VIII: F-measures detecting inadequate descriptions for "purpose"

Method	Recall	Precision	F-measure
Baseline	1.00 (94/94)	0.31 (94/305)	0.47
Rule-based	0.53 (50/94)	0.60 (50/ 84)	0.56
Machine learning	0.44 (41/94)	0.32 (41/127)	0.37

of manual judgment on the 266 papers in the 2011 annual meeting of the association for natural language processing in Japan, we constructed criteria of manual judgment. We calculated the agreement (Kappa value) between a person A who judges all of the training and test data and a person B who is different from person A. We extracted 24 data items from the training data that are judged by person A. We used the 24 data items to calculate a Kappa value. The Kappa value was 0.67 (substantial agreement).

A portion of the criteria of manual judgment is shown in Table V. In the table, "Useful" indicates that a paper detected by a system as a paper with inadequate descriptions is judged to be useful for support of writing the paper. "Not useful" indicates that a paper detected by a system as a paper with inadequate descriptions is not judged to be useful for support of writing the paper.

6) Experimental results: We performed experiments detecting papers in which an IRM is not described. In the experiments, we used the four IRMs "comparison," "problem," "purpose," and "example." In a preliminary experiment, we found that it was difficult to detect a paper in which an IRM is not described against the two IRMs "necessity" and "originality" by using rule-based and machine learning methods that use only words in a paper. Therefore, we used the remaining four IRMs in the experiments. Detecting papers inadequate in the two IRMs "necessity" and "originality" is not dealt with here, but will be dealt with in a future study.

We show the experimental results of extracting a paper in which an IRM is not described by using the rule-based method and the machine learning method in Tables VI to IX. The baseline method judges that all of the papers are papers in which an IRM is not described.

7) Discussions: From Tables VI to IX, we see that the F-measures in the rule-based method are the highest among all of the methods in all of the types of IRMs. Comparing the rule-based method and the machine learning method, we see that the F-measures of the machine learning method were lower than those of the rule-based method.

The reason that the performance of machine learning is low is that the number of features is too large. When we performed additional experiments that we did not use all words appearing in a paper as features and we use only words used in the rule-based method as features, the F-measures were the same as the rule-based method.

Because the F-measures for the rule-based method are the highest among all of the methods in the experiments, we could confirm the effectiveness of that method.

In the rule-based method, the F-measures for the IRMs, "comparison," "problem," and "purpose," were 0.6 to 0.7. The F-measure for the IRM, "example," was 0.86.

V. Analysis for Automatic Correction of Inadequate Descriptions

We handled automatic detection of inadequate descriptions in the previous section. We would like to handle automatic correction of inadequate descriptions as the next stage. However, it is difficult to handle automatic correction of inadequate descriptions directly. Therefore, we performed an analysis that is useful for automatic correction or construction that supports manual correction. For the analysis, we set five levels for description of IRMs. A paper of a higher level has clearer descriptions of IRMs. We also examined a pattern of describing IRMs. We can use a pattern of describing IRMs in a paper of Level 5 to correct descriptions of IRMs in a paper of Level 1. In this section, we examined five levels for the two IRMs "purpose" and "problems."

TABLE IX: F-measures detecting inadequate descriptions for "example"

Method	Recall	Precision	F-measure
Baseline	1.00 (9/9)	0.03 (9/305)	0.06
Rule-based	1.00 (9/9)	0.75 (9/ 12)	0.86
Machine learning	0.33 (3/9)	0.02 (3/129)	0.04

TABLE X: Definition of five levels

Level	Definition
5	Clue expressions for exist an IRM in sentences. Anyone can recognize an IRM easily.
4	Even if a reader does not have expert knowledge, he can recognize an IRM from a context.
3	It is difficult for a reader to recognize an IRM from a context. He can recognize an IRM when he reads a paper carefully.
2	A reader can recognize an IRM using his expert knowledge and deep insight.
1	A reader cannot recognize an IRM at all.

A. Data

We used 266 papers in the 2011 annual meeting of the association for natural language processing in Japan. We randomly extracted 50 papers from the 266 papers. We examined the 50 papers to ensure that we would manually categorize them into five levels.

B. Definition of five levels

The definition of the five levels is shown in Table X. A paper with a higher level is a paper in which an IRM is described more clearly.

C. Analysis results and examples

We manually categorized 50 papers into five levels for the two IRMs "purpose" and "problems." The numbers of papers categorized into five levels are shown in Table XI.

Examples of papers categorized into Levels 5 to 1 for the IRM "purpose" are shown in Figures 2 through 6. The sentences in the figures are the English translations of original Japanese sentences.

In the paper of Figure 2, the clue expression of "the purpose of this study is" occurs. Therefore, anyone can recognize the purpose of the study very easily. By the analysis in Level 5, we obtained the following expressions as the clue expressions for "purpose": "the purpose of $\cdots$ is," "$\cdots$ is the goal of $\cdots$," and "perform $\cdots$ for $\cdots$."

In the paper shown in Figure 3, the expression "Therefore, we do $\cdots$" appears just after the descriptions of problems. We can understand that the purpose is to solve the problems from the context. This is a context-based pattern using expressions such as "therefore" after problems. From the context-based pattern, we can recognize the purpose easily.

In the paper shown in Figure 4, the expression of "therefore" does not appear just after the descriptions of problems, so we must think logically. The purpose can be understood logically by grasping and thinking of the contents as well as the context.

In the paper shown in Figure 5, the background (problems) of the study is not described, but the effectiveness of the study

TABLE XI: Number of papers in the five levels

	Level 5	Level 4	Level 3	Level 2	Level 1
Purpose	19	16	12	2	1
Problem	11	26	6	4	3

$\cdots$ Maintenance of personal vocabulary according to the knowledge level and the learning step and dictionary manpower is an effective plan for supporting facilitation of communication based on specialized knowledge. It is necessary to grasp the features of the vocabulary that actually exists by the knowledge level and the shape according to the learning step for it first. The purpose of this study is to clarify the features of the vocabulary system according to the stage of school by analyzing and comparing the construction of the knowledge in a textbook of a junior high school, a high school, and a university as a network structure of professional vocabulary.

Fig. 2: Portion of a paper with Level 5 for "purpose"

$\cdots$ Technology of documentary classification based on mechanical learning is used by the service indicated above. For example, the naive Bayes distinction method and a distinction method, such as the support vector machine (SVM), are well-known. We must prepare a corpus from a great deal of learning data to achieve a high classification accuracy at this time. Such a corpus is constructed through work that provides labels (annotation) to a great deal of non-labeled data items. At this time, the problems is that when the amount of annotation is larger, human costs and time are larger.

Therefore, we propose a construction method for the corpus for documentary classification using active learning based on clustering as the method for reduce the amount of annotation when constructing a corpus for documentary classification handling the service described.

Fig. 3: Portion of a paper with Level 4 for "purpose"

$\cdots$ Subtitles of a TV program (closed captions) are one big pillar of information security to a hearing-impaired person. Subtitles have begun to be given to almost all non-live broadcast programs in recent years. A target program of providing subtitles has expanded into all programs, including raw radio programs (except for a portion of them) according to a broadcasting guideline for persons with seeing and hearing disabilities by the Ministry of Public Management, Home Affairs, Posts and Telecommunications. The projects will be implemented by fiscal year 2017.

At present, subtitles in real time for live sports programs are provided. In this paper, we conducted a basic investigation of the number of characters and named entities and the presentation speed by comparing the scripted text from sounds and subtitles in terms of the subtitles provided to a sports raw radio program on television (program of the soccer and sumo tournament) actually in real time. $\cdots$

Fig. 4: Portion of a paper with Level 3 for "purpose"

··· Micro blog services such as Twitter have spread rapidly in recent years, and a lot is contributed every day. A micro blog is an information-sending tool with a very high-speed information communication speed compared with a conventional blog service.

This paper proposes a method of performing effective information recommendation to a user by using the real-time contribution peculiar to microblogging. It is practical and useful to guess at a user's taste by analyzing contributions in the past and performing information recommendation by effective timing using actual data of goods. ···

Fig. 5: Portion of a paper with Level 2 for "purpose"

··· The meaning and expressive form of a word are often a relation of many-to-many, therefore, natural language processing is a challenging domain. Knowledge acquisition of paraphrasing is the technique of acquiring knowledge to recognize and generate different plural expressive forms that have the same meaning. In this study, we propose a method for acquiring knowledge on paraphrases from definition sentences on the web. Paraphrases are defined as pairs of expressions for which a two-way implication relation is satisfied. There are many sentences defining the same concept on the web. The sentences are likely to be paraphrases. Therefore, the web is a treasure house of knowledge on paraphrases.

Fig. 6: A part of a paper with Level 1 in "purpose"

is described. In the paper, we cannot recognize the purpose unless we have expert knowledge and use deep insight.

In the paper shown in Figure 6, only the methods are described, and we cannot recognize the purpose of the study. Problems are not described. The reason they extract paraphrases or the importance of extracting paraphrases is not described. We cannot recognize for what purpose they performed the study using the method.

Analyzing the IRM "problem," we found the following expressions as the clue expressions "there are problems that," "it is difficult that ···," and "but ··· occurs."

D. Discussion

We manually analyzed the two IRMs, "purpose" and "problem," and also manually categorized papers into five levels. As a result, we found that there were patterns that were often found in each IRM.

In the current study, we categorized papers into five levels manually. However, in the future, we would like to try to categorize automatically.

The five levels are useful for supporting the writing of papers. Because Level 5 has the best papers, it can be used as a good sample. We can consider the following support system. The support system automatically determines the level of a paper. When the level is low, the system shows a writer the reasons the paper is at a low level. Moreover, it shows an example of a paper of Level 5. The writer can correct the inadequate descriptions by consulting a paper of Level 5. Although this support system is not fully automatic, it would still be useful for supporting writing. The examination of the five levels in this paper is useful for gathering example sentences of Level 5 used in the above method to support writing.

We manually attempted to analyze the five levels for the two IRMs, "comparison" and "example." However, in these cases, because the issue is whether a paper includes a comparison, or whether a paper includes an example, the five levels are not necessary, and two levels are adequate.

VI. Conclusions

In this paper, we used rules and machine learning to propose methods for detecting inadequate descriptions in a paper. In our experimental results, the F-measure of the rule-based method was the highest for all IRMs. The F-measures for the IRMs, "comparison," "problem," and "purpose," were 0.6 to 0.7. The F-measure for the IRM "example" was 0.86.

We performed analysis for automatic correction of the inadequate descriptions. We could gather description patterns that are useful to consult when writing adequate descriptions, such as "the purpose of ··· is," "is the goal of," and "perform ··· for ···"

In the future, we would like to construct a system to support writing by using automatic correction of inadequate descriptions and description patterns we gather from an analysis. We would also like to be able to categorize papers automatically into five levels based on the quality of descriptions in a paper.

Acknowledgment

This work was supported by JSPS KAKENHI Grant Number 26330252.

References

[1] S. Fukuda, H. Nanba, and T. Takezawa, "Extraction and visualization of technical trend information from research papers and patents," *D-Lib Magazine*, 2012.

[2] M. Ptaszynski and F. Masui, "Spass: A scientific paper writing support system," in *Proceedings of The Third International Conference on Informatics Engineering and Information Science (ICIEIS 2014)*, 2014, pp. 1–10.

[3] A. Nadamoto, E. Aramaki, T. Abekawa, and Y. Murakami, "Extracting content-holes by comparing community-type content with wikipedia," *The International Journal of Web Information Systems*, vol. 6, no. 3, pp. 248–260, 2010.

[4] S. Tsudo, M. Murata, M. Tokuhisa, and Q. Ma, "Machine learning for analysis and detection of redundant sentences toward development of writing support systems," in *Proceedings of the 13th International Symposium on Advanced Intelligent Systems*, 2012, pp. 2225–2228.

[5] M. Murata, K. Kanzaki, K. Uchimoto, Q. Ma, and H. Isahara, "Meaning sort — three examples: dictionary construction, tagged corpus construction, and information presentation system —," *Computational Linguistics and Intelligent Text Processing, Second International Conference, CICLing 2001, Mexico City, February 2001 Proceedings*, pp. 305–318, 2001.

[6] A. L. Berger, S. A. D. Pietra, and V. J. D. Pietra, "A Maximum Entropy Approach to Natural Language Processing," *Computational Linguistics*, vol. 22, no. 1, pp. 39–71, 1996.

[7] N. Cristianini and J. Shawe-Taylor, *An Introduction to Support Vector Machines and Other Kernel-based Learning Methods*. Cambridge University Press, 2000.