Improvement of Translation Quality by Application of Translation Rules Based on Generalization of Bilingual Sentences to Phrase-Based SMT

Ryo Terashima
Graduate School of Information Science and Technology
Hokkaido University
N 14-Jo, W 9-Chome, Kita-ku
Sapporo, 060-0814 Japan
terashima@media.eng.hokudai.ac.jp

Hiroshi Echizen-ya
Faculty of Engineering
Hokkai-Gakuen University
S 26-Jo, W 11-Chome, Chuo-ku
Sapporo, 064-0926 Japan
echi@eli.hokkai-s-u.ac.jp

Kenji Araki
Graduate School of Information Science and Technology
Hokkaido University
N 14-Jo, W 9-Chome, Kita-ku
Sapporo, 060-0814 Japan
araki@media.eng.hokudai.ac.jp

***Abstract*—A phrase-based statistical machine translation (SMT) system can generate translation results for various input sentences. However, the phrase-based SMT system may generate unnatural translation results, because it cannot deal with structures of sentences sufficiently. On the other hand, in translation methods using translation rules based on generalization of bilingual sentences, a system acquires translation rules that possess structures of sentences. In this paper, we propose a method that combines phrase-based SMT with translation rules based on generalization of bilingual sentences. We are aiming at the improvement of translation quality by using acquired translation rules without any analytical tools in the phrase-based SMT. Evaluation experiments indicated that the translation quality improved by combining phrase-based SMT with translation rules based on generalization of bilingual sentences.**

***Keywords*-Machine translation; Translation rules; Bilingual sentences;**

I. Introduction

Recently, phrase-based statistical machine translations have been actively researched[1][2][3]. A phrase-based statistical machine translation (SMT) system can generate translation results for various input sentences. However, the phrase-based SMT system may generate unnatural translation results, because it cannot deal with structures of sentences sufficiently. It would appear that the phrase-based SMT uses N-gram model as a language model. N-gram model deals only with local parts in sentences. Therefore, it is difficult for it to deal with structures of sentences.

On the other hand, in the translation methods using translation rules based on generalization of bilingual sentences[4][5], a system acquires translation rules by replacing parts of the bilingual sentences with variables. These acquired translation rules possess structures of sentences. For example, a translation rule (Hello , this is @0 speaking . ; *kochira ha* @0 *desu* .) shows that "Hello , this is @0 speaking ." in the English sentence corresponds to "*kochira ha* @0 *desu* ." in the Japanese sentence. Then, "@0" becomes a variable, and "@0" in the English sentence corresponds to "@0" in the Japanese sentence.

We propose a method that combines phrase-based SMT with translation rules based on generalization of bilingual sentences. When the bilingual sentences that structures are similar to the input sentences appear in the training data, our system acquires translation rules based on generalization of bilingual sentences. The goal of our method is that system using it improves the translation quality by using acquired translation rules without any analytical tools in the phrase-based SMT.

Moreover, we performed the evaluation experiments by using various automatic evaluation methods. In addition, we performed the evaluation based on the pair comparison method by human evaluator. On the other hand, we performed preliminary experiments to tuning parameter of our system using automatic evaluation methods. The results of the evaluation experiments indicated that the translation quality improved by combining phrase-based SMT with translation rules based on generalization of bilingual sentences.

In this paper, we describe improvement of the translation quality by application of translation rules based on generalization of bilingual sentences for the phrase-based SMT, and performance evaluation of the proposed system.

II. Translation rules based on generalization of bilingual sentences

A. Outline of the proposed system

At first, we construct the phrase-based SMT system based on Moses[1], GIZA++[2], and SRILM(SRI Language Modeling Toolkit)[3]. Next, we combine processing that uses translation rules based on generalization of bilingual sentences with the phrase-based SMT system.

Figure 1 shows structure of the proposed system. At first, the system acquires the translation rules for the input sentences by using bilingual sentences in training data. When the translation rules are acquired, the translation results are

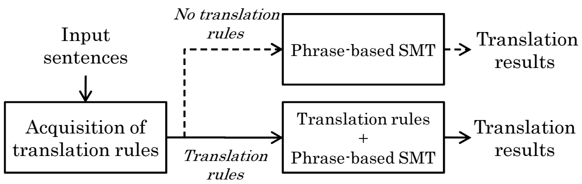

Figure 1. Structure of the proposed system.

generated by using translation rules and the phrase-based SMT. Otherwise, the translation results are generated by using only phrase-based SMT.

Moreover, in the processing that acquires and applies translation rules based on generalization of bilingual sentences, the system does not need any analytical tools.

B. Process

Figure 2 shows an example of acquisition of translation rule and translation. In Figure 2, the source language is English and the target language is Japanese.

1) Selection of the bilingual sentences: The system acquires translation rules based on the bilingual sentences that structures are similar to the input sentence.

The system determines common parts and different parts between the input sentence and source language parts of the bilingual sentences from training data based on LCS[6]. When there is only one different part between the input sentence and source language part of bilingual sentence, the system selects these bilingual sentences. However, when the number of words included in the different part is larger than 4, the system does not select those bilingual sentences. This threshold was determined by preliminary experiments described in section III–*D*.

In Figure 2, the input sentence is “Hello , this is Tanaka speaking .” The system selects the bilingual sentence (Hello , this is Richard speaking . ; *kochira ha richâdo desu*[1] .) that the different part is only “Richard” for the input sentence.

2) Generalization of the bilingual sentences: The system searches for the words that correspond to the words included in the different part, from the target language part of the selected bilingual sentences. The system determines the words where the alignment is done in both directions of source language → target language and target language → source language by GIZA++, correspond to each other between the source language part and the target language part in the bilingual sentences. GIZA++ is a toolkit acquiring word alignment using statistics information.

In Figure 2, the system determines the word “Richard” included in the different part corresponds to the word “*richâdo*” in the target language part of the bilingual sentence. Because these words are aligned in both directions of English → Japanese and Japanese → English by GIZA++.

[1]Italics expresses the pronunciation in Japanese.

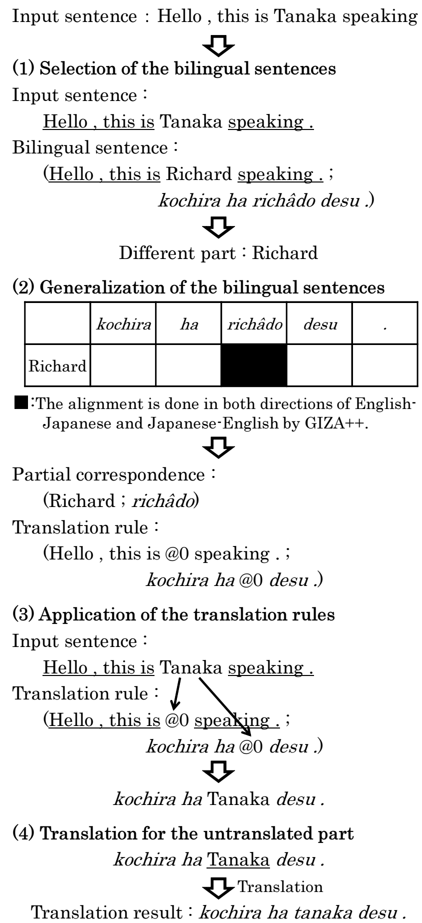

Figure 2. Example of acquisition of translation rule and translation.

Moreover, the system acquires a translation rule (Hello , this is @0 speaking . ; *kochira ha* @0 *desu* .) by replacing the partial correspondence (Richard ; *richâdo*) with variables “@0” from the bilingual sentence (Hello , this is Richard speaking . ; *kochira ha richâdo desu* .).

Then, “@0” becomes a variable, and “@0” in the English sentence corresponds to “@0” in the Japanese sentence.

3) Application of the translation rules: The system applies the acquired translation rules to an input sentence. As the result, translation of the part except the variable part is completed.

In Figure 2, the system applies the acquired translation rule (Hello , this is @0 speaking . ; *kochira ha* @0 *desu* .) to the input sentence “Hello , this is Tanaka speaking .” As the result, the system generates “*Hello , this is* Tanaka *speaking* .'' In this case, “Tanaka” is the untranslated part.

4) Translation of the untranslated parts: The system generates the translation results by translating the untranslated parts. In translation of the untranslated parts, the phrase-based SMT is used.

In Figure 2, the system generates "*Hello , this is Tanaka speaking* ." as a translation result.

On the other hand, when using only phrase-based SMT, a erroneous translation result "*moshimoshi, kore* speaking *Tanaka desu.*" was generated in the evaluation experiments described in section III.

When the translation rules were not acquired by processing *1)–2)*, the translation results are generated by using only phrase-based SMT. In the evaluation experiments, among the translation results for all evaluation data described in section III, 44.7% of the translation results were generated by using only phrase-based SMT.

III. Experiments

In this paper we performed the evaluation experiments to confirm the effectiveness of the proposed system. In that case, we used two MT systems. One is the proposed system described in section II, and the other is phrase-based SMT system based on Moses, GIZA++, and SRILM (the baseline system). We are aiming at the improvement of the translation quality by increasing fluent translations using acquired translation rules.

These systems translated English sentences into Japanese sentences using an experimental data. There are two kinds of experimental data: learning data and evaluation data. In these experiments, 1368 bilingual sentences were used as learning data. As well, 342 bilingual sentences were used as evaluation data. These bilingual sentences were taken from an English–Japanese travel phrase book.

A. Evaluation Method

We used the pair comparison method performed by human evaluator and the automatic evaluation to evaluate both MT systems.

B. Pair comparison method

In the pair comparison method performed by human evaluator, the first author of this paper compared the translation results from the proposed system with the translation results from the baseline system by one sentence and classified them in three of the following for all the evaluation data.

1. **Improved**: The proposed system is better than the baseline system.
2. **Deteriorated**: The baseline system is better than the proposed system.
3. **Same level**: Two MT systems are at the same level of the translation quality.

In the proposed system, translation quality in the relative evaluation comparing to the baseline system is calculated as the following.

$$\text{Translation quality} = \frac{N_i - N_d}{N_e} \qquad (1)$$

In this equation, N_i is the number of improved sentences, and N_d is the number of deteriorated sentences in comparing the proposed system with the baseline system. N_e is the number of the bilingual sentences in the evaluation data.

C. Automatic evaluation

We used BLEU, IMPACT[6], METEOR, ROUGE-L, mWER, and mPER as the automatic evaluation to compare two MT systems. In this case, four Japanese sentences translated by bilingual humans were used for each translated Japanese sentence to calculate the scores. In case of mWER and mPER, the lower score shows the higher evaluation.

D. Preliminary experiments to tuning parameter

When the system selects the bilingual sentences described in section II–*B–1)*, the limitation is given to the number of words included in the different part. We determine the threshold of the number of words included in the different part by using BLEU and IMPACT. Table I shows the BLEU and IMPACT scores when the threshold is changed. In Table I, when the threshold is 3 and 4, the scores are the highest. Therefore, we determine 4 as the threshold.

E. Experimental results

Table II shows experimental results of the pair comparison method performed by human evaluator. In Table II, the number of improved sentences is 28 and the number of deteriorated sentences is 14. That is, in the proposed system, 4.09% of the all translation results were improved comparing to the baseline system.

Table III shows the automatic evaluation scores of the proposed system and the baseline system respectively. In Table III, the scores of the proposed system improved from 0.008 to 0.016 comparing to the baseline system.

F. Discussion

Table IV shows examples of the translation results. In example 1 and 2 of Table IV, the proposed system was higher quality than the baseline system. In these input sentences, the proposed system could generate the correct translation results by using translation rules based on generalization of bilingual sentences.

Table I
Variation of the automatic evaluation scores.

Threshold	1	2	3	4	5
BLEU	0.3211	0.3274	**0.3314**	**0.3314**	0.3307
IMPACT	0.5711	0.5762	**0.5800**	**0.5800**	0.5795

Table II
Experimental results of the pair comparison method.

Number of improved sentences	28
Number of deteriorated sentences	14
Translation quality in the relative evaluation	+4.09%

Table III
EXPERIMENTAL RESULTS OF THE AUTOMATIC EVALUATION.

System	BLEU	IMPACT	METEOR	ROUGE-L	mWER	mPER
The proposed system	0.331	0.580	0.600	0.583	0.570	0.376
The baseline system	0.322	0.572	0.584	0.575	0.579	0.392

Table IV
EXAMPLE OF THE TRANSLATION RESULTS.

	Input sentence	The baseline system	The proposed system
1	Hello, this is Tanaka speaking.	*moshimoshi, kore* speaking *Tanaka desu.*	*kochira ha tanaka desu.*
2	I'd like a room with shower.	*shawâ no heya wo onegai si masu.*	*shawâ tsuki no heya wo onegai si masu.*
3	I'd like to make an overseas call to Tokyo, Japan.	*tokyo he kokusai denwa wo kake tai no desu ga.*	*tokyo he* overseas *denwa wo kake tai no desu ga.*

Next, we describe the deteriorated sentences. Main causes of errors are acquisition of the erroneous translation rules and failure of translation for the untranslated parts by the phrase-based SMT. In example 3 of Table IV shows example of the erroneous translation result by the proposed system. In the translation result of the proposed system, "overseas" was not able to be translated. The proposed system has translation process of two ways. One is using translation rules based on generalization of bilingual sentences, and the other is using only phrase-based SMT. In the current proposed system, when the system generated the translation results with translation rules based on generalization of bilingual sentences, the system gives priority to these translation results over the translation results using only phrase-based SMT. However, when the translation results using only phrase-based SMT are better, the system should give priority to these translation results. In a two-way translation process, it is important that system selects the better translation results. This is one of our future tasks.

Uchino[7], Akiba[8] and Sumita[9] proposed a method of combining plural MT systems. However, these methods depend on static language knowledge. On the other hand, our proposed system can perform acquisition translation rules and translation without any analytical tools. This means that it is easier for our proposed system to apply to many languages.

IV. CONCLUSION

In this paper, we propose a method that combines phrase-based SMT with translation rules based on generalization of bilingual sentences. We are aiming at the improvement of the translation quality using acquired translation rules without any analytical tools in the phrase-based SMT.

From the results of the evaluation experiments, in the pair comparison method performed by human evaluator, 4.09% of the all translation results were improved in comparing the proposed system with the baseline system. Moreover, by using various automatic evaluation methods, we understood that the scores of our proposed system improved from 0.008 to 0.016 comparing to the baseline system. We confirmed that the translation quality improved by combining phrase-based SMT with translation rules based on generalization of the bilingual sentences. In the future, we plan to perform experiments with large-scale bilingual corpus.

REFERENCES

[1] P. Koehn, H. Hoang, A. Birch, C. Callison-Burch, M. Federico, N. Bertoldi, B. Cowan, W. Shen, C. Moran, R. Zens, C. Dyer, O. Bojar, A. Constantin and E. Herbst: Moses: Open Source Toolkit for Statistical Machine Translation, Proceedings of ACL 2007, pp.177-180 (2007).

[2] F.J. Och and H. Ney: A Systematic Comparison of Various Statistical Alignment Models, Computational Linguistics, vol.29, No.1, pp.19-51 (2003).

[3] A. Stolcke: SRILM - an Extensible Language Modeling Toolkit, 7th International Conference on Spoken Language Processing, pp.901-904 (2002).

[4] R. Terashima, H. Echizen-ya and K. Araki: Learning Method for Extraction of Partial Correspondence from Parallel Corpus, Proceedings of the International Conference on Asian Language Processing 2009, pp.293-298 (2009)

[5] M. Sasayama, F. Ren and S. Kuroiwa: Automatic Super-Function Extraction for Translation of Spoken Dialogue, Natural Language Processing and Knowledge Engineering 2007, pp.141-148 (2007).

[6] H. Echizen-ya and K. Araki: Automatic Evaluation of Machine Translation based on Recursive Acquisition of an Intuitive Common Parts Continuum, Proceedings of the MT SUMMIT XI, pp.151-158 (2007).

[7] H. Uchino, Y. Ooyama and O. Furuse: Furuse: ALTFLASH: A Japanese-to-English Machine Translation System for Market Flash Reports, Proceedings of the MT Summit VII, pp.438-443 (1999).

[8] Y. Akiba, T. Watanabe and E. Sumita: Using language and translation models to select the best among outputs from multiple MT systems, Proceedings of COLING '02, pp. 8-14 (2002).

[9] E. Sumita, Y. Akiba, T. Doi, A. Finch, K. Imamura, M. Paul, M. Shimohata, and T. Watanabe: A Corpus-Centered Approach to Spoken Language Translation, Proceedings of EACL '03, pp.171-174 (2003).