Proceedings of HLT-NAACL 2003
Short Papers , pp. 37-39
Edmonton, May-June 2003

Automatic Expansion of Equivalent Sentence Set Based on Syntactic Substitution

Kenji Imamura, Yasuhiro Akiba and **Eiichiro Sumita**
ATR Spoken Language Translation Research Laboratories
2-2-2 Hikaridai, "Keihanna Science City", Kyoto, 619-0288, Japan
{kenji.imamura,yasuhiro.akiba,eiichiro.sumita}@atr.co.jp

Abstract

In this paper, we propose an automatic quantitative expansion method for a sentence set that contains sentences of the same meaning (called an equivalent sentence set). This task is regarded as paraphrasing. The features of our method are: 1) The paraphrasing rules are dynamically acquired by Hierarchical Phrase Alignment from the equivalent sentence set, and 2) A large equivalent sentence set is generated by substituting source syntactic structures. Our experiments show that 561 sentences on average are correctly generated from 8.48 equivalent sentences.

1 Introduction

Sentences can be represented by various expressions even though they have the same meaning. Paraphrasing that transfer from sentence to sentence (Barzilay and McKeown, 2001) is a technique that generates such various expressions.

In this paper, we propose an automatic quantitative expansion method for a sentence set that contains sentences of the same meaning (called an equivalent sentence set), as a paraphrasing technique. Our method is roughly structured from the following two phases.

1. Extract phrasal correspondences that have the same meaning (called equivalent phrases) from the source equivalent sentence set (acquisition phase).
2. Based on the parse tree of the sentence selected from the source set, generate target sentences by recursively substituting the equivalent phrases for the source phrases (generation phase).

Paraphrasing is regarded as machine translation into the same language. In this paper, we apply syntactic substitution for generating sentences, which corresponds to transfer-based machine translation. In addition, Hierarchical Phrase Alignment (HPA) (Imamura, 2001), which is an automatic acquisition method for machine translation rules, is applied to acquire the paraphrasing rules. Namely, two equivalent sentences are regarded as bilingual sentences, and simplified machine translation is carried out.

Paraphrasing by our method has the following characteristics.

- Not only lexical paraphrasing but also phrasal paraphrasing can be generated because our method is based on structural substitution.
- Equivalent phrases extracted by HPA are not only semantically but also grammatically equivalent. Thus, our method rarely generates ungrammatical sentences by substitution.

Expansion of the equivalent sentence set can be applied to automatic evaluation of machine translation quality (Papineni et al., 2002; Akiba et al., 2001), for example. These methods evaluate the quality of the translation by measuring the similarity between machine translation results and translations done by humans (called references). However, the accuracy increases when multiple references are applied because one source sentence can be translated into multiple target expressions. Our method generates multiple sentences that are suitable for this purpose.

2 Acquisition of Paraphrasing Rules: Hierarchical Phrase Alignment

Hierarchical Phrase Alignment is based on the assumption that an "equivalent phrase pair has the same information and the same grammatical role." We decompose this assumption into the following two conditions for computation.

- The words in the phrase pair correspond, with no deficiency and no excess.
- The phrases are of the same syntactic category.

Therefore, HPA is a task to extract phrase pairs that satisfy the above two conditions. The procedure of HPA

is summarized as follows. [1]

1. Tag and parse two equivalent sentences.
2. Extract corresponding words (called word links) between the sentences. In this paper, we regard identical words and words that belong to the same group in a thesaurus as word links.
3. Check all combinations of syntactic nodes between the sentences. If the node pair satisfies the above two conditions, then output the pair as an equivalent phrase. Namely, if no words in the phrase link to the outside of the other phrase, and the nodes have the same category, the phrase pair is regarded as equivalent.

Figure 1 shows an example of equivalent phrase extraction from source equivalent sentences. The upper sentence is interrogative, the lower sentence is imperative, and they have the same meaning. For example, focusing on the upper phrase "*get me*," this phrase is VP and contains two word links. However, no nodes contain only the links '*get*', and '*me*' in the lower sentence. On the other hand, focusing on the upper phrase "*get me a taxi*," it contains four word links that correspond to the lower phrase "*get a taxi for me*", and they have the same syntactic category. Therefore, the node pair VP(4) is regarded as an equivalent phrase.

By iterating the above process, HPA consequently extracts eight nodes as equivalents from the source sentences shown in Figure 1. Excluding the identical phrases, the following three phrases are acquired as equivalent phrases.

- "*get me a taxi*" and "*get a taxi for me*"
- "*10 in the morning*" and "*10 a.m.*"
- "*at 10 in the morning*" and "*at 10 a.m.*"

HPA can extract phrasal correspondences from source equivalent sentences even if their sentence structures are significantly different. In addition, because node pairs have to be in the same syntactic category, unparaphrasable correspondences, such as "*morning*" and "*a.m.*," are ignored even though they have word links.

3 Expansion of Equivalent Sentence Set

The equivalent phrases extracted by HPA are substitutable with one another because they are semantically and grammatically equivalent. Therefore, they are regarded as bi-directional paraphrasing rules. When we paraphrase from any N sentences, target equivalent sentences are generated by the following procedure, where the range from Step 1 to Step 3 corresponds to the acquisition phase, and Steps 4 and 5 correspond to the generation phase.

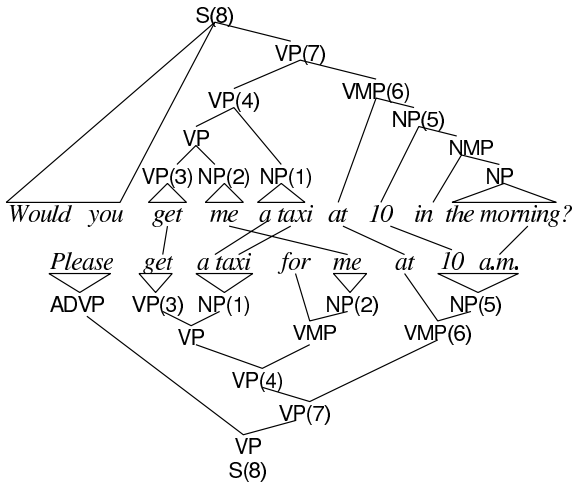

Figure 1: Example of Equivalent Phrase Extraction from English Equivalent Sentences (The lines between the sentences denote word links, the trees denote parsing results, and the numbers on the nodes denote corresponding equivalent phrases.)

1. First, select one sentence from the source equivalent sentence set.
2. Process HPA with the remaining $(N-1)$ sentences, and extract equivalent phrases.
3. Repeat Steps 1 and 2 for all combinations of the source sentences. All phrases that construct the source set and their paraphrasing rules are acquired.
4. Next, select one tree created by HPA from the source equivalent sentence set, and trace the tree top-down. If a node registered in the paraphrasing rules is found, substitute the equivalent phrase for the node. Substitution is recursively done until it reaches a leaf.
5. Repeat Step 4 with all sentences in the source set.

For example, when the source equivalent sentence set contains only the two sentences shown in Figure 1, the following six sentences are generated. Our method generates all sentences constructed from the phrases of N sentences.

Would you get a taxi for me at 10 a.m.?
Would you get a taxi for me at 10 in the morning?
Would you get me a taxi at 10 a.m.?
Please get me a taxi at 10 in the morning
Please get me a taxi at 10 a.m.
Please get a taxi for me at 10 in the morning.

[1] The original method of HPA has two additional features. 1) Ambiguity of parsing is resolved by comparing parse trees of input sentences. 2) It employs partial parsing to analyze irregular sentences. Details are described in (Imamura, 2001).

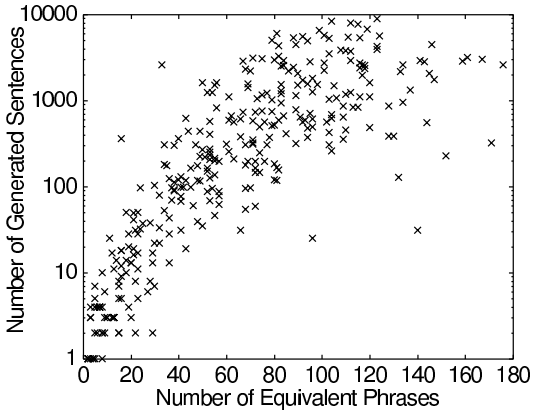

Figure 2: Relationship between Number of Equivalent Phrases and Number of Generated Sentences

		Japanese Correctness		
		OK	NG	Total
Translation Effectiveness	OK	892 (61%)	382 (26%)	1274 (86%)
	NG	87 (6%)	112 (8%)	199 (14%)
Total		979 (66%)	494 (34%)	1473 (100%)

Table 1: Quality of Generated Sentences

4 Experiments

Expansion experiments of Japanese equivalent sentences were carried out. We used 339 source equivalent sentence sets selected from ATR corpus (Furuse et al., 1994). The sets were created by having ten Japanese native speakers translate English sentences into Japanese. The number of different sentences was 8.48 sentences per English sentence on average.

Number of Generated Sentences Figure 2 is a graph plotting the number of equivalent phrases and the number of of generated sentences. Each point denotes a source equivalent sentence set. Consequently 60.2 equivalent phrases on average were acquired, and 920 sentences on average were generated from a source set.

Quality of Generated Sentences We randomly selected five sentences per set from above generated sentences and showed them to a Japanese native speaker together with the English sentences. One-by-one, he/she judged whether the sentences were good or not from the viewpoints of Japanese correctness (grammatically and pragmatically correct or not) and translation effectiveness (understandable or not). The results are shown in Table 1.

Consequently, approximately 61 percent of the generated sentences were judged good from the viewpoints of both Japanese correctness and translation effectiveness. In other words, 561 sentences on average were correctly generated, and the source equivalent sentence sets were expanded about 66 times. About 39% of the generated sentences contain errors. However, we believe that our method is effective when we make a large equivalent sentence set because eliminating error sentences is easier than creating a large set manually.

Our error analysis found that major errors are caused by inconsistency in the modality. Our method does not consider pragmatical correctness although it generates syntactically correct sentences.

5 Conclusion

We proposed an expansion method for an equivalent sentence set based on syntactic substitution. Our method dynamically acquires paraphrasing rules by using HPA, and generates many sentences by applying the rules to a parse tree recursively. One application of our method is the automatic evaluation of machine translation quality. We are planning to integrate this method into a form of automatic evaluation.

Acknowledgment

The research reported here is supported in part by a contract with the Telecommunications Advancement Organization of Japan entitled, "A study of speech dialogue translation technology based on a large corpus."

References

Yasuhiro Akiba, Kenji Imamura, and Eiichiro Sumita. 2001. Using multiple edit distances to automatically rank machine translation output. In *Proceedings of Machine Translation Summit VIII*, pages 15–20.

Regina Barzilay and Kathleen R. McKeown. 2001. Extracting paraphrases from a parallel corpus. In *Proceedings of the 39th Annual Meeting of the Association for Computational Linguistics*, pages 50–57.

Osamu Furuse, Y. Sobashima, Toshiyuki Takezawa, and N. Uratani. 1994. Bilingual corpus for speech translation. In *Proceedings of the AAAI'94 Workshop 'Integration of Natural Language and Speech Processing'*, pages 84–91.

Kenji Imamura. 2001. Hierarchical phrase alignment harmonized with parsing. In *Proceedings of the 6th Natural Language Processing Pacific Rim Symposium (NLPRS 2001)*, pages 377–384.

Kishore Papineni, Salim Roukos, Todd Ward, and Wei-Jing Zhu. 2002. Bleu: a method for automatic evaluation of machine translation. In *Procee*