Reordering Model Using Syntactic Information of a Source Tree for Statistical Machine Translation

Kei Hashimoto[*1], **Hirohumi Yamamoto**[*2*3], **Hideo Okuma**[*2*4],
Eiichiro Sumita[*2*4], **and Keiichi Tokuda**[*1*2]
[*1]Nagoya Institute of Technology Department of Computer Science and Engineering
/ Gokiso-cho Syouwa-ku Nagoya-city Aichi Japan
[*2]National Institute of Information and Communications Technology
[*3]Kinki University School of Science and Engineering Department of Informaiton
[*4]ATR Spoken Language Communication Research Labs.

Abstract

This paper presents a reordering model using syntactic information of a source tree for phrase-based statistical machine translation. The proposed model is an extension of IST-ITG (imposing source tree on inversion transduction grammar) constraints. In the proposed method, the target-side word order is obtained by rotating nodes of the source-side parse-tree. We modeled the node rotation, monotone or swap, using word alignments based on a training parallel corpus and source-side parse-trees. The model efficiently suppresses erroneous target word orderings, especially global orderings. Furthermore, the proposed method conducts a probabilistic evaluation of target word reorderings. In English-to-Japanese and English-to-Chinese translation experiments, the proposed method resulted in a 0.49-point improvement (29.31 to 29.80) and a 0.33-point improvement (18.60 to 18.93) in word BLEU-4 compared with IST-ITG constraints, respectively. This indicates the validity of the proposed reordering model.

1 Introduction

Statistical machine translation has been wiedely applied in many state-of-the-art translation systems. A popular statistical machine translation paradigms is the phrase-based model (Koehn et al., 2003; Och and Ney, 2004). In phrase-based statistical machine translation, errors in word reordering, especially global reordering, are one of the most serious problems. To resolve this problem, many word-reordering constraint techniques have been proposed. These techniques are categorized into two types. The first type is linguistically syntax-based. In this approach, tree structures for the source (Quirk et al., 2005; Huang et al., 2006), target (Yamada and Knight, 2000; Marcu et al., 2006), or both (Melamed, 2004) are used for model training. The second type is formal constraints on word permutations. IBM constraints (Berger et al., 1996), the lexical word reordering model (Tillmann, 2004), and inversion transduction grammar (ITG) constraints (Wu, 1995; Wu, 1997) belong to this type of approach. For ITG constraints, the target-side word order is obtained by rotating nodes of the source-side binary tree. In these node rotations, the source binary tree instance is not considered. Imposing a source tree on ITG (IST-ITG) constraints (Yamamoto et al., 2008) is an extension of ITG constraints and a hybrid of the first and second type of approach. IST-ITG constraints directly introduce a source sentence tree structure. Therefore, IST-ITG can obtain stronger constraints for word reordering than the original ITG constraints. For example, IST-ITG constraints allows only eight word orderings for a four-word sentence, even though twenty-two word orderings are possible with respect to the original ITG constraints. Although IST-ITG constraints efficiently suppress erroneous target word orderings, the method cannot assign the probability to the target word orderings.

This paper presents a reordering model using syntactic information of a source tree for phrase-based statistical machine translation. The proposed reordering model is an extension of IST-ITG con-

Proceedings of SSST-3, Third Workshop on Syntax and Structure in Statistical Translation, pages 69–77,
Boulder, Colorado, June 2009.

straints. In the proposed method, the target-side word order is obtained by rotating nodes of a source-side parse-tree in a similar fashion to IST-ITG constraints. We modeled the rotating positions, monotone or swap, from word alignments of a training parallel corpus and source-side parse-trees. The proposed method conducts a probabilistic evaluation of target word orderings using syntactic information of the source tree.

The rest of this paper is organized as follows. Section 2 describes the previous approach to resolving erroneous word reordering. In Section 3, the reordering model using syntactic information of a source tree is presented. Section 4 shows experimental results. Finally, Section 5 presnts the summary and some concluding remarks and future works.

2 Previous Works

First, we introduce two previous studies on related word reordering constraints, ITG and IST-ITG constraints.

2.1 ITG Constraints

In one-to-one word-alignment, the source word f_i is translated into the target word e_i. The source sentence $[f_1, f_2, \cdots, f_N]$ is translated into the target sentence which is the reordered target word sequence $[e_1, e_2, \cdots, e_N]$. The number of reorderings is $N!$. When ITG constraints are introduced, this combination $N!$ can be reduced in accordance with the following constraints.

- All possible binary tree structures are generated from the source word sequence.
- The target sentence is obtained by rotating any node of the binary trees.

When $N = 4$, the ITG constraints can reduce the number of combinations from $4! = 24$ to 22 by rejecting the combinations $[e_3, e_1, e_4, e_2]$ and $[e_2, e_4, e_1, e_3]$. For a four-word sentence, the search space is reduced to 92% (22/24), but for a 10-word sentence, the search space is only 6% (206,098/3,628,800) of the original full space.

2.2 IST-ITG Constraints

In ITG constraints, the source-side binary tree instance is not considered. Therefore, if a source sentence tree structure is utilized, stronger constraints than the original ITG constraints can be created. IST-ITG constraints directly introduce a source sentence tree structure. The target sentence is obtained with the following constraints.

- A source sentence tree structure is generated from the source sentence.
- The target sentence is obtained by rotating any node of the source sentence tree structure.

By parsing the source sentence, the parse-tree is obtained. After parsing the source sentence, a bracketed sentence is obtained by removing the node syntactic labels; this bracketed sentence can then be converted into a tree structure. For example, the parse-tree "(S1 (S (NP (DT This)) (VP (AUX is) (NP (DT a) (NN pen)))))" is obtained from the source sentence "This is a pen," which consists of four words. By removing the node syntactic labels, the bracketed sentence "((This) ((is) ((a) (pen))))" is obtained. Such a bracketed sentence can be used to produce constraints. If IST-ITG constraints is applied, the number of word orderings in $N = 4$ is reduced to 8, down from 22 with ITG cnstraints. For example, for the source-side bracketed tree "$((f_1 f_2)\ (f_3 f_4))$," the eight target sequences $[e_1, e_2, e_3, e_4]$, $[e_2, e_1, e_3, e_4]$, $[e_1, e_2, e_4, e_3]$, $[e_2, e_1, e_4, e_3]$, $[e_3, e_4, e_1, e_2]$, $[e_3, e_4, e_2, e_1]$, $[e_4, e_3, e_1, e_2]$, and $[e_4, e_3, e_2, e_1]$ are accepted. For the source-side bracketed tree "$(((f_1 f_2)\ f_3)\ f_4)$," the eight sequences $[e_1, e_2, e_3, e_4]$, $[e_2, e_1, e_3, e_4]$, $[e_3, e_1, e_2, e_4]$, $[e_3, e_2, e_1, e_4]$, $[e_4, e_1, e_2, e_3]$, $[e_4, e_2, e_1, e_3]$, $[e_4, e_3, e_1, e_2]$, and $[e_4, e_3, e_2, e_1]$ are accepted. When the source sentence tree structure is a binary tree, the number of word orderings is reduced to 2^{N-1}. The parsing results sometimes do not produce binary trees. In this case, some subtrees have more than two child nodes. For a non-binary subtree, any reordering of child nodes is allowed. If a subtree has three child nodes, six reorderings of the nodes are accepted.

In phrase-based statistical machine translation, a source "phrase" is translated into a target "phrase". However, with IST-ITG constraints, "word" must be used for the constraint unit since the parse unit is a "word". To absorb different units between translation models and IST-ITG constraints, a new limitation for word reordering is applied.

- Word ordering that destroys a phrase is not allowed.

When this limitation is applied, the translated word ordering is obtained from the bracketed source sentence tree by reordering the nodes in the tree, which is the same as for one-to-one word-alignment.

3 Reordering Model Using Syntactic Information of the Source Tree

In this section, we present a new reordering model using syntactic information of a source-side parse-tree.

3.1 Abstract of Proposed Method

The IST-ITG constraints method efficiently suppresses erroneous target word orderings. However, IST-ITG constraints cannot evaluate the accuracy of the target word orderings; i.e., IST-ITG constraints assign an equal probability to all target word orderings. This paper proposes a reordering model using syntactic information of the source tree as an extension of IST-ITG constraints. The proposed reordering model conducts a probabilistic evaluation of target word orderings using syntactic information of the source-side parse-tree.

In the proposed method, the target-side word order is obtained by rotating nodes of the source-side parse-tree in a similar fashion to IST-ITG constraints. Reordering probabilities are assigned to each subtree of source-side parse-tree S by reordering the positions into two types: monotone and swap. If the subtree has more than two child nodes, the number of child node order is more than two. However, we assume the child node order other than monotone to be swap. The source-side parse-tree S consists of subtrees $\{s_1, s_2, \cdots, s_K\}$, where K is the number of subtrees included in the source-side parse-tree. The subtree s_k is which is represented by the parent node's syntactic label and the order, from sentence head to sentence tail, of the child node's syntactic labels. For example, Figure 1 shows a source-side parse-tree for a four-word

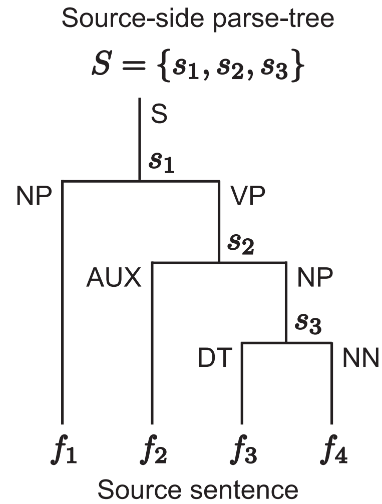

Figure 1: Example of a source-side parse-tree fo a four-word source sentence consisting of three subtrees.

source sentence consisting of three subtrees. In Figure 1, the subtrees s_1, s_2, and s_3 are represented by **S+NP+VP**, **VP+AUX+NP**, and **NP+DT+NN**, respectively. Each subtree has a probability $P(t \mid s_k)$, where t is monotone (m) or swap (s). The probability of the target word reordering is calculated as follows.

$$P_r = \prod_{k=1}^{K} P(t \mid s_k) \tag{1}$$

Each target candidate is assigned the different reordering probability by Equation (1). Since the proposed reordering model uses the syntactic labels, which is not considered in IST-ITG constraints, the different parse-tree assigns the different reordering probability. The proposed model is effective for global word reordering, because reordering probabilities are also assigned to higher-level subtrees of the source-side parse-tree.

3.2 Training of the Proposed Model

We modeled monotone or swap node rotating automatically from word alignments of a training parallel corpus and source-side parse-trees. The training algorithm for the proposed reordering model is as follows.

1. The training process begins with a word-aligned corpus. We obtained the word alignments using Koehn et al.'s method (2003),

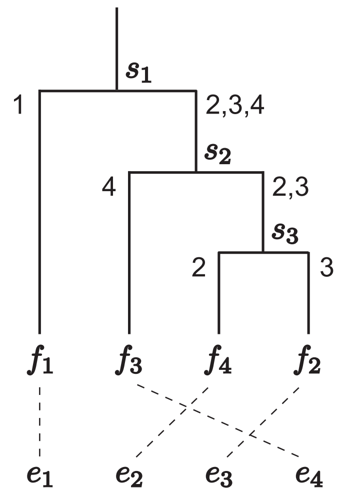

Figure 2: Example of a source-side parse-tree with word alignments using the training algorithm of the proposed model.

which is based on Och and Ney's work (2004). This involves running GIZA++ (Och and Ney, 2003) on the corpus in both directions, and applying refinement rules (the variant they designate is "final-and") to obtain a single many-to-many word alignment for each sentence.

2. Source-side parse-trees are created using a source language phrase structure parser, which annotates each node with a syntactic label. A source-side parse-tree consists of several subtrees with syntactic labels. For example, the parse-tree "(S1 (S (NP (DT This)) (VP (AUX is) (NP (DT a) (NN pen)))))" is obtained from the source sentence "This is a pen" which consists of four words.

3. Word alignments and source-side parse-trees are combined. Leaf nodes are assigned target word positions obtained from word alignments. Via the bottom-up process, target word positions are assigned to all nodes. For example, in Figure 2, the left-side (sentence head) child node of subtree s_2 is assigned the target word position "4," and the right-side (sentence tail) child node is assigned the target word positions "2" and "3," which are assigned to the child nodes of subtree s_3.

4. The monotone and swap reordering positions are checked and counted for each subtree. By comparing the target word positions, which are assigned in the above step, the reordering position is determined. If the target word position of the left-side child node is smaller than one of the right-side child node, the reordering position determined as monotone. For example, in Figure 2, the subtrees s_1, s_2 and s_3 are monotone, swap, and monotone, respectively.

5. The reordering probability of the subtree can be directly estimated by counting the reordering positions in the training data.

$$P(t \mid s) = \frac{c_t(s)}{\sum_t c_t(s)} \tag{2}$$

where $c_t(s)$ is the count of reordering positon t included all training samples for the subtree s.

Subtree type	Monotone probability
S+PP+,+NP+VP+.	0.764
PP+IN+NP	0.816
NP+DT+NN+NN	0.664
VP+AUX+VP	0.864
VP+VBN+PP	0.837
NP+NP+PP	0.805
NP+DT+JJ+NN	0.653
NP+DT+JJ+VBP+NN	0.412
NP+DT+NN+CC+VB	0.357

Table 1: Example of proposed reordering models.

The parsing results sometimes do not produce binary trees. For a non-binary subtree, any reordering of child nodes is allowed. However, the proposed reordering model assumes that reordering positions are only two, monotone and swap. That is, the reordering position which the order of child nodes do not change is monotone, and the other positions are swap. Therefore, the probability of swap $P(\mathrm{s} \mid s_k)$ is derived from the probability of monotone $P(\mathrm{m} \mid s_k)$ as follows.

$$P(\mathrm{s} \mid s_k) = 1.0 - P(\mathrm{m} \mid s_k) \tag{3}$$

Table 1 shows the example of proposed reordering models.

If a subtree is represented by a binary-tree, there are L^3 possible subtrees, where L is the number of

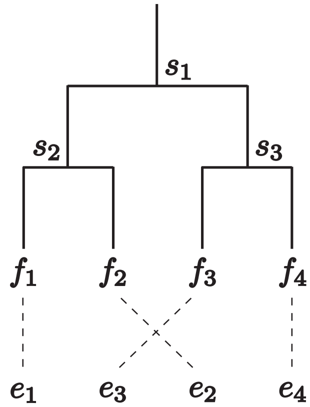

Figure 3: Example of a target word order which is not derived from rotating the nodes of source-side parse trees.

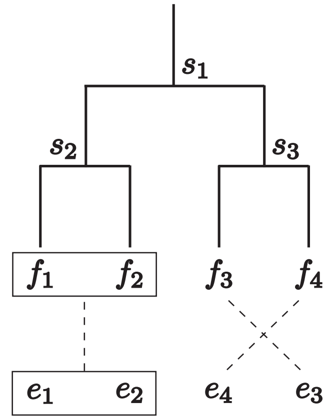

Figure 4: Example of a target candidate including a phrase.

syntactic labels. However, in the possible subtrees, there are subtrees observed only a few times in training sentences, especially when the subtree consists of more than three child nodes. Although a large number of subtree models can capture variations in the training samples, too many models lead to the over-fitting problem. Therefore, subtrees where the number of training samples is less than a heuristic threshold and unseen subtrees are clustered to deal with the data sparseness problem for robust model estimations.

After creating word alignments of a training parallel corpus, there are target word orders which are not derived from rotating nodes of source-side parse-trees. Figure 3 shows a sample which is not derived from rotating nodes. Some are due to linguistic reasons, structural differences such as negation (French "ne...pas" and English "not"), adverb, modal and so on. Others are due to non-linguistic reasons, errors of automatic word alignments, syntactic analysis, or human translation (Fox, 2002). The proposed method discards such problematic cases. In Figure 3, the subtree s_1 is then removed from training samples, and the subtrees s_2 and s_3 are used as training samples.

3.3 Decoding Using the Proposed Reordering Model

In this section, we describe a one-pass phrase-based decoding algorithm that uses the proposed reordering model in the decoder. The translation target sentence is sequentially generated from left (sentence head) to right (sentence tail), and all reordering is conducted on the source side. To introduce the proposed reordering model into the decoder, the target candidate must be checked for whether the reordering position of a subtree is either monotone or swap whenever a new phrase is selected to extend a target candidate. The checking algorithm is as follows.

1. For old translation candidates, the subtree s, which includes both translated and untranslated words, and its untranslated part u are calculated.
2. When a new target phrase $\bar{e}$ is generated, the source phrase $\bar{f}$ and the untranslated part u calculated in the above step are compared. If the source phrase $\bar{f}$ does not include the untranslated part u and is not included u, the new candidate is rejected.
3. In the accepted candidate, the reordering positions for all subtrees included the source side parse-tree are checked by comparing the source phrase $\bar{f}$ with the source phrase sequence used before.

Subtrees checked reordering positions are assigned a probability–monotone or swap–by the proposed reordering model, and the target word order is evaluated by Equation (1).

Phrase-based statistical machine translation uses a "phrase" as the translation unit. However, the proposed reordering model needs a "word" order. Because "word" alignments form the source phrase to target phrase are not clear, we cannot determine the

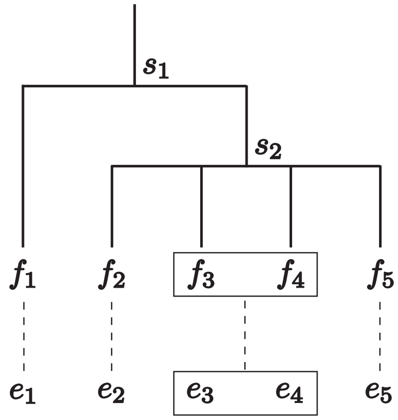

Figure 5: Example of a non-binary subtree including a phrase.

reordering position of subtree included in a phrase. Therefore, in the decoding process using the proposed reordering model, we define that higher probability, monotone or swap, are assigned to subtrees included in a source phrase. For example, in Figure 4, the source sentence $[[f_1, f_2], f_3, f_4]$ is translated into the target sentence $[[e_1, e_2], e_4, e_3]$, where $[f_1, f_2]$ and $[e_1, e_2]$ are used as phrases. Then, the source phrase $[f_1, f_2]$ includes the subtree s_2. If the monotone probabilities of subtrees s_1, s_2, and s_3 are 0.8, 0.4 and 0.7, the proposed reordering probability is $0.8 \times 0.6 \times 0.3 = 0.144$. If a source phrase is $[f_1, f_2, f_3, f_4]$ and a source-side parse-tree has the same tree structure used in Figure 4, the subtrees s_1, s_2, and s_3 are assigned higher reordering probabilities. If the source phrase $[f_1, f_2, f_3, f_4]$ used in Figure 4, the subtrees s_1, s_2, and s_3 are assigned higher reordering probabilities.

Non-binary subtrees are often observed in the source-side parse-tree. When a source phrase $\bar{f}$ is included in a non-binary subtree and does not include a non-binary subtree, we cannot determine the reordering position. For example, the reordering position of subtree s_2 in Figure 5, which includes the phrase $[f_3, f_4]$, can not be determined. In this case, we define that such subtrees are also to be assigned a higher probability.

4 Experiments

To evaluate the proposed model, we conducted two experiments: English-to-Japanese and English-to-Chinese translation.

		English	Japanese
Train	Sentences	1.0M	
	Words	24.6M	24.6M
Dev	Sentences	2.0K	
	Words	50.1K	58.7K
Test	Sentences	2.0K	
	Words	49.5K	58.0K

Table 2: Statistics of training, development and test corpus for E-J translation.

4.1 English-to-Japanese Paper Abstract Translation Experiments

The first experiment was the English-to-Japanese (E-J) translation. Table 2 shows the training, development and test corpus statistics. JST Japanese-English paper abstract corpus consists of 1.0M parallel sentences were used for model training. This corpus was constructed from 2.0M Japanese-English paper abstract corpus belongs to JST by NICT using the method of Uchiyama and Isahara (2007). For phrase-based translation model training, we used the GIZA++ toolkit (Och and Ney, 2003), and 1.0M bilingual sentences. For language model training, we used the SRI language model toolkit (Stolcke, 2002), and 1.0M sentences for the translation model training. The language model type was word 5-gram smoothed by Kneser-Ney discounting (Kneser and Ney, 1995). To tune the decoder parameters, we conducted minimum error rate training (Och, 2003) with respect to the word BLEU score (Papineni et al., 2002) using 2.0K development sentence pairs. The test set with 2.0K sentences is used. In the evaluation and development sets, a single reference was used. For the creation of English sentence parse trees and segmentation of the English, we used the Charniak parser (Charniak, 2000). We used Chasen for segmentation of the Japanese sentences. For decoding, we used an in-house decoder that is a close relative of the Moses decoder. The performance of this decoder was configured to be the same as Moses. Other conditions were the same as the default conditions of the Moses decoder.

In this experiment, the following three methods were compared.

- Baseline : The IBM constraints and the lexical reordering model were used for target word

	Baseline	IST-ITG	Proposed
BLEU	27.87	29.31	29.80

Table 3: BLEU score results for E-J translation. (1-reference)

reordering.

- IST-ITG : The IST-ITG constraints, the IBM constraints, and the lexical reordering model were used for target word reordering.
- Proposed : The proposed reordering model, the IBM constraints, and the lexical reordering model were used for target word reordering.

During minimum error training, each method used each reordering model and reordering constraint.

The proposed reordering model are trained from 1.0M bilingual sentences for the translation model training. The amount of available training samples represented by subtrees was 9.8M. In the available training samples, there were 54K subtree types. The heuristic threshold was 10, and subtrees with training samples of less than 10 were clustered. The proposed reordering model consisted of 5,960 subtrees types and one clustered model "other". The models not including "other" covered 99.29% of all training samples.

The BLEU scores are presented in Table 3. In comparing "Baseline" method with "IST-ITG" method, the improvement in BLEU was a 1.44-point. Furthermore, in comparing "IST-ITG" method with "Proposed" method, the improvement in BLEU was a 0.49-point. Both the IST-ITG constraints and the proposed reordering model fixed the phrase position for the global reorderings. However, the proposed method can conduct a probabilistic evaluation of target word reorderings which the IST-ITG constraints cannot. Therefore, "Proposed" method resulted in a better BLEU.

4.2 NIST MT08 English-to-Chinese Translation Experiments

Next, we conducted English-to-Chinese (E-C) newspaper translation experiments for different language pairs. The NIST MT08 evaluation campaign English-to-Chinese translation track was used for the training and evaluation corpora. Table 4 shows the training, development and test corpus statistics. For the translation model training, we used 4.6M bilingual sentences. For the language model training, we used 4.6M sentences which are used for the translation model training. The language model type was word 3-gram smoothed by Kneser-Ney discounting. A development set with 1.6K sentences was used as evaluation data in the Chinese-to-English translation track for the NIST MT07 evaluation campaign. A single reference was used in the development set. The evaluation set with 1.9K sentences is the same as the MT08 evaluation data, with 4 references. In this experiment, the compared methods were the same as in the E-J experiment.

		English	Chinese
Train	Sentences	4.6M	
	Words	79.6M	73.4M
Dev	Sentences	1.6K	
	Words	46.4K	39.0K
Test	Sentences	1.9K	
	Words	45.7K	47.0K (Ave.)

Table 4: Statistics of training, development and test corpus for E-C translation.

	Baseline	IST-ITG	Proposed
BLEU	17.54	18.60	18.93

Table 5: BLEU score results for E-C translation. (4-reference)

The proposed reordering model are trained from 4.6M bilingual sentences for the translation model training. The amount of available training samples represented by subtrees was 39.6M. In the available training samples, there were 193K subtree types. As in the E-J experiments, the heuristic threshold was 10. The proposed reordering model consisted of 18,955 subtree types and one clustered model "other." The models not including "other" covered 99.45% of all training samples.

The BLEU scores are presented in Table 5. In comparing "Baseline" method with "IST-ITG" method, the improvement in BLEU was a 1.06-point. In comparing "IST-ITG" method with "Proposed" method, the improvement in BLEU was a 0.33-point. As in the E-J experiments, "Proposed" method performed the highest BLEU. We demonstrated that the proposed method is effective for multiple language pairs. However, the improvement of BLEU score in E-C translation is smaller than the improvement in E-J translation, because English and Chinese are similar sentence structures, such as SVO-languages (Japanese is SOV-language). When the sentence structures are different, the proposed reordering model is effective.

5 Conclusion

This paper proposed a new word reordering model using syntactic information of a source tree for phrase-based statistical machine translation. The proposed model is an extension of the IST-ITG constraints. In both IST-ITG constraints and the proposed method, the target-side word order is obtained by rotating nodes of the source-side tree structure. Both the IST-ITG constraints and the proposed reordering model fix the phrase position for the global reorderings. However, the proposed method can conduct a probabilistic evaluation of target word reorderings which the IST-ITG constraints cannot. In E-J and E-C translation experiments, the proposed method resulted in a 0.49-point improvement (29.31 to 29.80) and a 0.33-point improvement (18.60 to 18.93) in word BLEU-4 compared with IST-ITG constraints, respectively. This indicates the validity of the proposed reordering model.

Future work will focus on a reduction of computational cost of decoding including the proposed reordering model, and a simultaneous training of translation and reordering models. Moreover, we will deal with difference between source and target in multi level like in Gally et al. (2004).

The improvement could clearly be seen from visual inspection of the output, a few examples of which are presented in the following Appendix.

A Samples from the English-to-Japanese Translation

A.1 Sentence 1

Source: Aggravation was obvious from the latter half of March to the end of April, and he contracted the disease in February to the beginning of May.
Baseline: ４月末に３月後半から５月上旬に２月に疾患を発症し，著明な増悪した。
Reference: ３月後半から４月末に増悪が著明で，２～５月上旬に発症した。
Proposed: ３月後半から４月末に著明な増悪し，５月上旬に２月に疾患を発症した。

A.2 Sentence 2

Source: The value of TF, on the other hand, was higher in the reverse order, indicating that high oxidation rate causes severe defects on the surface of Ni crystallites.
Baseline: 一方，重症の表面上の欠陥の原因となることを示し，逆順に高かったが，ＴＦの値は高い酸化速度はＮｉの微結晶た。
Reference: 一方，ＴＦの値は逆の順序で高く，酸化速度が高いことはＮｉ結晶の表面欠陥の原因になることを示した。
Proposed: 一方，ＴＦの値は逆の順序で高かったことを示し，高い酸化速度は，Ｎｉの微結晶表面に重篤な欠陥の原因となる。

A.3 Sentence 3

Source: After diagnosing the pleural effusion and ascites, vein catheter was left in place under the echo guide, and after removing the pleural effusion and ascites, OK-432 was administered locally.
Baseline: 診断後，胸水，腹水，胸水・腹水を除去した後，エコーガイド下で，静脈カテーテルを左に代わってＯＫ‐４３２を投与した。
Reference: 胸水・腹水の診断を行った後にエコーガイド下に静脈カテーテルを留置し，胸水・腹水を除去し，ＯＫ‐４３２を局所投与した。
Proposed: 胸水・腹水の診断後，静脈カテーテルを残したエコーガイド下で代わりに，胸水・腹水を除去した後，ＯＫ‐４３２，局所的に投与した。

A.4 Sentence 4

Source: From result of the consideration, it was pointed that radiation from the loop elements was weak.
Baseline: 考察の結果からことを指摘し，ループ素子からの放射は弱かった。
Reference: 考察結果より，ループ素子からの放射が弱いことを指摘する。
Proposed: 考察の結果から，ループ素子からの放射は弱いことを示した。

References

Adam L. Berger, Peter F. Brown, Stephen A. Della Pietra, Vincent J. Della Pietra, Andrew S. Kehler, and Robert L. Mercer 1996. Language translation apparatus and method of using context-based translation models. United States patent, patent number 5510981.

Eugene Charniak. 2000. A Maximum-Entropy-Inspired Parser. In *Proceedings of NAACL 2000*, pages 132–139.

Chasen
http://chasen-legacy.sourceforge.jp/

Heidi J. Fox, 2002. Phrasal cohesion and statistical machine translation. In *Proceedings of EMNLP*, pages 304–311.

Michel Galley, Mark Hopkins, Kevin Knight, and Daniel Marcu. 2004. What's in a translation rule? In *Proceedings of HLT/NAACL-04*.

Liang Huang, Kevin Knight, and Aravind Joshi. 2006. Statistical Syntax-Directed Translation with Extended Domain of Locality. In *Proceedings of AMTA*.

Japanese-English paper abstract corpus
http://www.jst.go.jp

Reinhard Kneser and Hermann Ney. 1995. Improved backing-off for m-gram language model In *Proceedings of ICASSP 1995*, pages 181–184.

Philipp Koehn, Franz Josef Och, and Daniel Marcu. 2003. Statistical phrase-based translation. In *Proceedings of HLT-NAACL 2003*, pages 127–133.

Daniel Marcu, Wei Wang, Abdessamad Echihabi, and Kevin Knight. 2006. SPMT: Statistical Machine Translation with Syntactified Target Language Phrases In *Proceedings of EMNLP2006*, pages 44–52.

Dan Melamed. 2004. Statistical machine translation by parsing In *Proceedings of ACL*, pages 653–660.

Moses
http://www.statmt.org/moses/

Franz josef Och and Hermann Ney. 2003. A Systematic Comparison of Various Statistical Alignment Models. *Computational Linguistics*, 29(1), pages 19–51.

Franz josef Och. 2003. Minimum error rate training for statistical machine translation. In *Proceedings of ACL*, pages 160–167.

Franz josef Och and Hermann Ney. 2004. The alignment template approach to statistical machine translation. *Computational Linguistics*, 30(4), pages 417–449.

Chris Quirk, Arul Menezes, and Colin Cherry. 2005. Dependency treelet translation: Syntactically informed phrasal SMT. In *Proceedings of ACL*, pages 271–279.

Kishore Papineni, Salim Roukos, Todd Ward, and Wei-Jing Zhu. 2002. Bleu: a method for automatic evaluation of machine translation. In *Proceedings of ACL*, pages 311–318.

Andreas Stolcke. 2002. SRILM - An Extensible Language Model Toolkit In *Proceedings of ICSLP2002*, pages 901–904. http://www.speech.sri.com/projects/srilm/

Christopher Tillmann. 2004. A unigram orientation model for statistical machine translation. In *Proceedings of HLT-NAACL*, pages 101–104.

Masao Uchiyama and Hitoshi Isahara. 2007. 2007. A japanese-english patent parallel corpus. In MT summit XI, pages 475–482.

Dekai Wu. 1995. Stochastic inversion transduction grammars, with application to segmentation, bracketing, and alignment of parallel corpora. In *Proceedings of IJCAI*, pages 1328–1334.

Dekai Wu. 1997. Stochastic inversion transduction grammars and bilingual parsing of parallel corpora. *Computational Linguiatics*, 23(3), pages 377–403.

Kenji Yamada and Kevin Knight. 2000. A syntax-based statistical translation model In *Proceedings of ACL*, pages 523–530.

Hirofumi Yamamoto, Hideo Okuma, and Eiichiro Sumita. 2008. Imposing Constraints from the Source Tree on ITG Constraints for SMT. In *Proceedings of ACL : HLT Second Workshop on Syntax and Structure in Statistical Translation (SSST-2)*, pages 1–9.