Rule-Based Preordering on Multiple Syntactic Levels in Statistical Machine Translation

Ge Wu, Yuqi Zhang

Institute for Anthropomatics
Karlsruhe Institute of Technoloy, Germany
`utcur@student.kit.edu`, `yuqi.zhang@kit.edu`

Abstract

We propose a novel data-driven rule-based preordering approach, which uses the information of syntax tree and word alignment to reorder the words in source sentences before decoding in a phrase-based SMT system between English and Chinese. The preorering algorithm extracts reordering patterns from multiple levels of the syntax trees of the training data and applies the rules on the source sentences in a similar manner. We've conducted experiments in English-to-Chinese and Chinese-to-English translation directions. Our results show that the approach has led to improved translation quality both when it was applied separately on the baseline or when it was combined with some other reordering approaches. As our reordering approach was used alone, it showed an improvement of 1.61 in BLEU score in the English-to-Chinese translation direction and an improvement of 2.16 in BLEU score in the Chinese-to-English translation direction, in comparison with the baseline, which used no preordering at all. As our preordering approach were combined with the short rule [1], long rule [2] and tree rule [3] based preordering approaches, it showed an improvement of 0.43 in BLEU score in the English-to-Chinese translation direction and an improvment of 0.3 in BLEU score in the Chinese-to-English translation direction. Through the translations that used our preordering approach, we've also found improvements in syntactic structure.

1. Introduction

Word order is a general issue when we want to translate text from one language to the other. Different languages normally have different word orders and the difference could be very huge. Among all the languages, Chinese is one language which is very different from English, because they belong to different language families and have long period of separately development. Both languages have a Subject-Verb-Object order, but they also have a lot of differences in word order. Especially sentences in both languages can sometimes have completely different syntactic structures. The differences may involve long-distance or unstructured position changes.

Most state-of-the-art phrase-based SMT systems use language model, phrase table or decoder to adjust the word order. Or an additional reordering model is used in the log-linear model for word reordering. However, these methods may have some disadvantages, such as some don't handle long-distance reordering, some don't handle unstructured reordering and some are rather time consuming.

Encouraged by the results from the paper [1], [2] and [3], we further propose a new data-driven, rule-based preordering method, which extracts and applies reordering rules based on syntax tree. The method is called Multi-Level-Tree (MLT) reordering, which orders the constituents on multiple levels of the syntax tree all together. This preordering method rearranges the words in source language into a similar order as they are supposed to be in the target language before translation. With the appropriate word order, better translation quality can be achieved. Especially, our preordering method is very suitable for translation between language pairs like English and Chinese, which have very different word orders. Besides, the method can also be combined with the above mentioned rule-based reordering methods to achieve better translation quality.

2. Related Work

Word reordering is an important problem for statistical machine translation, which has long been addressed.

In a phrase-based SMT system, there are several possibilities to change the word orders. Words can be reordered during the decoding phase by setting a window, which allows the decoder to choose the next word for translation. Reordering could also be influenced by the language model, because the language model give probability of how a certain word is likely to follow. Different language model may give different probability, which further influences the decision made by log-linear model. Other ways to change the word orders include using distance based reordering models or lexicalized reordering models [4, 5]. The lexicalized reordering model reorders the phrases by using information of how the neighboring phrases change orientations.

Another way to achieve word reordering is to detach it from decoding phase and do it separately in a pre-process before decoding, in order to reduce the time for translation.

This kind of preordering approaches use linguistic information to modify the word orders. Preordering can also be rule-based, which extracts different types of reordering rules by observing reordering patterns from the training data and apply the rules to the sentences to be translated. Depends on how the rules are defined, different information may be used such as word alignments, POS tags, syntax trees, etc.

Some early approaches use manually defined reordering rules based on the linguistic information for particular languages [6, 7, 8, 9]. Later come the data-driven methods [10, 11], which learn the reordering rules automatically.

Rottman and Vogel 2007 [1] introduced the idea of extracting reordering rules from the POS tag sequences of training data and use them for reordering. Niehues and Kolss 2009 [2] went further, and developed a method for long-distance word reordering, which works good on German-English translation task due to the long-distance shift of verbs. The method extracts discontinuous reordering rules in addition to the continuous ones, which contains a place-holder to match several words and enables the word to shift cross long distance.

Afterwards, Herrmann et al. 2013 [3] introduced a new approach to reorder the words based on syntax tree, which leads to further improvements on translation quality. The algorithm takes syntactic structure of the sentences into account and extract the rules from the syntax tree by detecting the reordering of child sequences. It also has the variant based only on part of the child sequences which is suitable for language with flat syntactic structure such as German.

However, these approaches which are bases on POS tag sequences or syntax trees are mostly designed for languages like German and are not especially adapted for languages like Chinese. As Chinese has very different word orders, a reordering approach, which can further explore the syntactic structure of Chinese and utilize this information for reordering, is desirable.

The hierarchical phrase-based translation model [12] is especially suitable for Chinese translation, and provide very good translation results. It extracts hierarchical rules by using information of the syntactic structure. Phrases from different hierarchies, or so-called phrases of phrases, are reordered during the decoding.

The idea of phrases on different hierarchies has inspired us to create this preordering method based on multiple levels of the syntax tree. Besides, we also hope to detach the reordering from decoding phase and do it separately in a pre-process before decoding, in order to reduce the time for translation. This kind of preordering approaches use linguistic information to modify the word orders.

Oracle reordering has also shown values for evaluating the potential of preordering. [13] introduced the permutation distance metrics which can be used to measure reordering quality. And [14] described how we can construct permutations from the word alignment as oracle reordering.

3. Motivation

The word order between English and Chinese differs very significantly. For one, the words in Chinese have generally different origins as those in English, which leads to very different vocabulary and word construction. Sometimes it is very hard to find corresponding words in the other language. For example, some prepositions in Chinese have very different usage than those prepositions in English. Also the continuous writing of Chinese without spaces makes this problem more severe, since word boundaries are not always so clear in Chinese. The text needs to be segmented first before translation. A word segmentation process is used to separate the words, but the results may not always be ideal.

For the other, both languages have sometimes very different sentence structures. Thus, a word-for-word translation between English and Chinese is often unnatural or difficult to understand. Each of them has some sentence patterns that don't exist or rarely used in the other. In Chinese, a modifier is often put before the part that it modifies. While in English, it is very common that the modifier is put after the part that it modifies. Besides, English sentences with a lot of long clauses may be more suitable to translate into several Chinese sentences, because in Chinese people don't tend to use long clauses in general.

Some typical problems of word orders between English and Chinese that we've found in the data are as follows:

- *Pre-modifier instead of post-modifier*
  In Chinese people tend to use pre-modifier rather than post-modifier. This involves the position change of adverbials, relative clauses and preposition phrases during translation.

- *Construction of questions*

- *Special sentence constructions*
  For example, *Bâ*-construction in Chinese and sentence constructions such as *there be* and inverted negative sentences in English don't have correspondence in the other language in general.

- *Long distance word position change*

From previous work of others, the following types of reordering rules have been researched:

- *Short rules* [1]
  Short rules are extracted based on the sequences of adjacent words or their POS tags in the sentence from training data. Sequences of adjacent words or POS tags are observed, rules are then extracted if the same reordering patterns appear frequently. Following is a reordering pattern of a short rule:

  ```
  WRB MD DT -> DT WRB MD
  ```

- *Long rules* [2]
  Long rules are specially designed to improve the long

distance word reordering for language pairs like English and German. The rules are based on POS tags of the text, and use a placeholder (∗) to represent one or more words. Following is a reordering pattern of a long rule:

```
NN * MD -> * MD NN
```

- *Tree rules* [3]
  While short rules and long rules are based on the flat structure of sentences, tree rules reorder sentences by using information from sentences' syntactic structure. The syntax tree and word alignment of the training corpus are used to train the reordering rules. The tree rules reorder the words both on the word level and on the constituent level. Following is a reordering pattern of a tree rule:

```
NP ( ADJP JJ NN ) -> JJ NN ADJP
```

The parenthesis in the example represents the hierarchy in the syntax tree. The left side of the rule corresponds a tree with the root labeled with *NP* and three children, each labeled with *ADJP*, *JJ* and *NN*.

However, not all these reordering problem of Chinese translation can be solved well by using these reordering rules based either on POS tag sequence or single level of syntax tree. In order to improve the reordering between English and Chinese, we need reordering method that can handle more complicated, unstructural word order change. Inspired by the ideas of reordering on syntax tree and hierarchical phrases, we created the MLT reordering, which reorders words based on multiple syntactic levels and can handle long distance word position change and complicated word position change very well.

4. Multi-Level-Tree Reordering

4.1. Reordering on Multiple Syntactic Levels

Reordering patterns are based on multiple levels of the syntax tree. Figure 1 illustrates how the reordering patterns are detected from the syntax tree. In the example, the detection starts from the root node, go downwards three levels and use the nodes in these levels to detect the reordering pattern. These nodes that are used for detecting the reordering pattern are colored gray and have an italic font. The leaf nodes in the syntax tree indicate words in the sentence, which has rectangle shape with angular corners.

According to the alignment information with the corresponding Chinese translation, the node labeled with *NP* is moved forward to the first place in the translation and the node labeled with *IN of* is moved forward to the second place in translation. So from the root node with a search depth 3, the following reordering pattern can be found:

NP (**CD**$_0$ NP (NP (**JJ**$_1$ **NNS**$_2$) PP (**IN**$_3$ **NP**$_4$))) -> NP IN CD JJ NNS
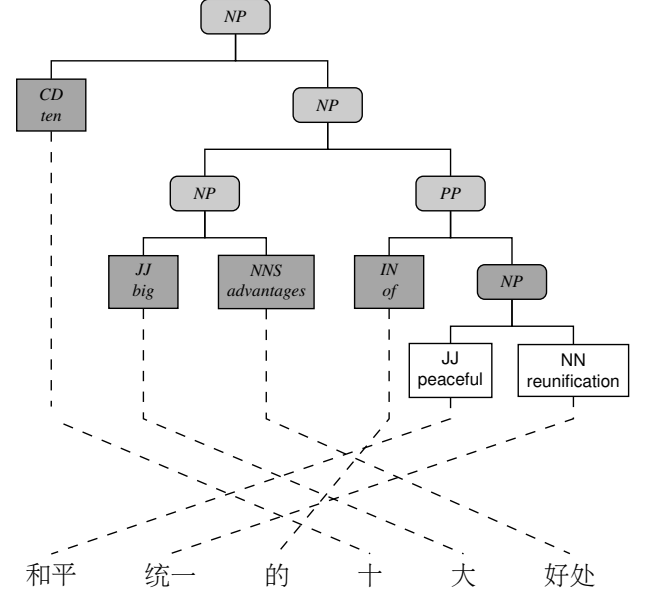

Figure 1: Detection of reordering pattern from multiple syntactic levels

```
or -> 4 3 0 1 2 (with index)
```

where the POS tags in bold corresponds to words in the leaf nodes of the syntax tree, which are parts that are actually reordered and have a darker gray color in the figure. The parentheses in the reordering pattern indicate the corresponding hierarchies in the syntax tree. The reordering can be alternatively represented with index, which is the actual internal representation to avoid ambiguity caused by POS tags with the same name and we use this representation throughout this paper.

4.2. Rule Extraction

In order to find as much information for reordering as possible. The algorithm of rule extraction detects the reordering patterns from all nodes in the syntax tree and it goes downwards for any number of hierarchies, until it reaches the lowest hierarchy in the subtrees.

In the implementation, the program conducts a depth-first search (DFS) to traverse every node in a syntax tree. Every time when a node is reached, the program conducts another iterative deepening depth-first search (IDDFS) in its subtree with depth-limit from 1 to the subtree's depth. And the program detects if there are any patterns of word position changes at the same time, by using the alignment for comparison.

The detected word position changes are checked for their validity for reordering rules. A valid reordering pattern should both has actual word reordering and clearly distinguishable new order on the side of the target language, i.e. no collision of aligned ranges on the target side.

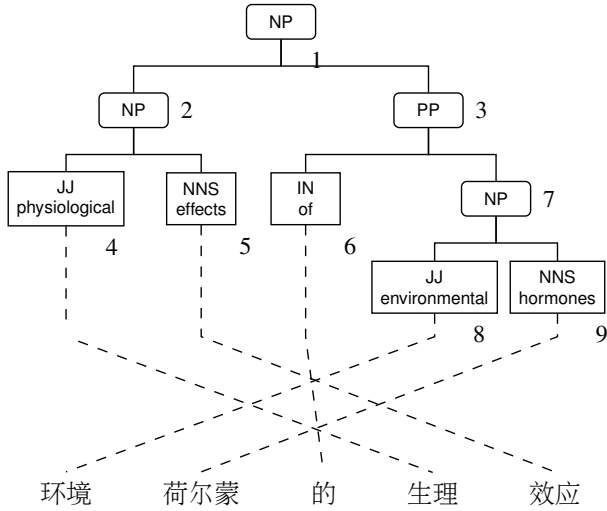Figure 2 shows a phrase to be translated, together with

Figure 2: A phrase with its syntax tree and alignment for rule extraction

its syntax tree and word alignment of parallel text. The leaf nodes are underlined and indexed according to the order they appear in the left side of the patterns. The index sequences on the right side of the patterns indicate the reordering. In this example, we can find the following reordering patterns:

From node 1:
```
NP ( NP PP ) -> 1 0                    [1 level]
NP ( NP ( JJ NNS ) PP ( IN NP ) ) -> 3 2
0 1                                    [2 levels]
NP ( NP ( JJ NNS ) PP ( IN NP ( JJ NNS )
) ) -> 3 4 2 0 1                       [3 levels]
```

From node 3:
```
PP ( IN NP ) -> 1 0                    [1 level]
PP ( IN NP ( JJ NNS ) ) -> 1 2 0   [2 levels]
```

The probability of the reordering patterns are calculated based on frequency of their occurrences in the training corpus. There are the left part and the right part of the reordering patterns separated by the arrow. The left part indicates the syntactic tags that should be reordered and the right part indicates how the new order should be like. The probability of the pattern is calculated by how often the left part is reordered into the right part among all its appearances in the training corpus. In addition, reordering patterns that appear less than a threshold are ignored to be used as reordering rules, in order to prevent too concrete rules without generalization capability and overfitting.

4.3. Rule Application

The syntax tree is traversed by DFS as the same in rule extraction. But from the root of each subtree, it's scanned with depth limit from its maximal levels, i.e. its depth, to 1. As it turns out, any rule can be applied for a subtree at some level,

a new path for this reordering will be added to the word lattice for decoding. As long as rules can be applied on a subtree for a certain depth, the rules are applied and the search for rule application on this subtree stops, and the search on the next subtree continues.

The reason for this is to prevent duplicate reorderings due to application of nested rules, which has overlapped effect with each other. These rules are normally patterns that are generated on the same subtree, but with different number of levels, which has different generalization effect on the same range of words in the text. For example, the following patterns can be detected from the syntax tree in figure 2:

```
PP ( IN NP ) -> 1 0
PP ( IN NP ( JJ NNS ) ) -> 1 2 0
```

Both patterns are detected from the same node, but the second pattern is detected by retrieving the nodes one level deeper and it's more concrete. So the first pattern can be seen as a generalization of the second pattern. Whenever a rule of the second pattern can be applied, a rule of the first pattern can be applied too. Because subtrees are checked from the highest number of levels in rule application, the more concrete rule is applied first. Because the more concrete rule fits the detected pattern better and contains more details of reordering, so it may be more suitable for rule application. In this example, the second rule is applied rather than the first rule.

Word reorderings are added to the word lattice as paths, which is further past to the decoder for translation. Paths with very low probability are removed, in order to save space for storing the lattice and reduce decoding time later, without compromising too much translation quality.

4.4. Rule Combination

In order to further explore the probability of improvement, the MLT reordering rules can be combined with other types of reordering rules to achieve further improvements. This is done by training the different types of rules separately and applying them on the monotone path of the sentence independently. They generate different paths in the word lattice.

5. Experiments

We've conducted experiments on both English-to-Chinese and Chinese-to-English translation directions to get a better overview of the MLT reordering's effect. In the experiments, the MLT reordering rules are also combined with other reordering rules that are introduced before, in order to show the improvement achieved by our approach.

In order to evaluate the potential of word reordering, we also used oracle reordering in the experiments. Oracle reordering is considered to be an optimally reordered sentence as input to the SMT system and do not allow additional reordering during decoding.[15] The oracle reordering is created by using the permutation of source sentences, which is

	BLEU (%)	Imprv.	TER (%)
Baseline	12.07		72.15
+Short Rules	12.50	0.43	71.41
+Long Rules	12.99	0.92	70.71
+Tree Rules	13.38	1.31	68.27
+MLT Rules	**13.81**	**1.74**	**68.20**
Oracle Reordering	18.58	6.51	62.13
Long Rules	12.31	0.24	71.81
Tree Rules	13.30	1.23	70.42
MLT Rules	**13.68**	**1.61**	**70.25**

Table 2: Result overview of English-to-Chinese system

	BLEU (%)	Imprv.	TER (%)
Baseline	21.80		62.09
+Short Rules	22.90	1.10	61.64
+Long Rules	23.13	1.33	61.43
+Tree Rules	23.84	2.04	60.95
+MLT Rules	**24.14**	**2.34**	**60.79**
Oracle Reordering	26.80	5.00	56.97
Long Rules	22.10	0.30	62.21
Tree Rules	23.35	1.55	61.52
MLT Rules	**23.96**	**2.16**	**60.83**

Table 4: Result overview of Chinese to English systems

extracted from the word alignment between the source text and reference.[14]

Because oracle reordering is considered as an optimally reordering, we can use it as input of the SMT system and the scores are the optimal results that can be achieved by word reordering, from which we can evaluate the potential of reordering methods.

5.1. English-to-Chinese System

We performed experiments with or without different reordering methods covering the English-to-Chinese translation direction. The reordering methods included our MLT reordering approach and the other reordering approaches with short rules, long rules and tree rules. The system was trained on news text from the LDC corpus and subtitles from TED talks. The development data and test data were both news text from the LDC corpus. The system was a phrase-based SMT system, which used a 6-gram language model with Knersey-Ney smoothing. Besides the preordering, no lexical reordering or other reordering method in decoding phase was used. The text was translated through a monotone decoder.

The reordering rules were extracted by using the word alignments, POS tags and syntax trees from the training data. One reference of the test data was used for evaluating the results. The threshold for rule extraction is set as 5 times and reordering paths with probability less than 0.1 are not added to word lattices. The decoder was a monotone decoder. Table 1 shows the size of data used in the system.

Table 2 shows the BLUE scores, absolute improvements of BLEU scores and TER scores for configurations with different reordering methods. The table consist of 2 sections. the first row of the top section shows results of the baseline, which used no preordering at all. In the following rows of the top section, different types of reordering rules are combined gradually, with each type per row, and the results are showed. For example, the row with +*MLT* Rules presents the configuration with all the rule types including MLT rules and all the other rules in the rows above. All improvements are absolute improvements of BLEU scores in comparison to the baseline. Each row with a certain reordering type presents all the different variations of this type and the best score un-

der these configurations are shown. For example, long rules include the left rules and right rules, and the tree rules include the partial rules and recursive application. The baseline used a monotone decoder and no preordering. The row with *oracle reordering* shows the results from the configuration that used the oracle reordering as input. The results of oracle reordering can be used for analyzing the potential of source sentence reordering. In the lower section of the table, different rule types are not combined and the effect of each rule type is shown.

5.2. Chinese-to-English System

The experiments for Chinese-to-English systems had a similar setup as described in the last section. The parallel data used in the English-to-Chinese system was also used in this experiment by switching the source language and the target language. We only used the LDC data set for training, and no TED data were used in this system. The test data had three English references for evaluating the results instead of one as in the previous system. The data used are summarized in table 3.

Table 4 shows the results for configurations with different reordering methods for the Chinese-to-English translation. The table can be interpreted in the same manner as table 2 in the previous section.

5.3. Evaluation

The results shows increasing scores as we used reordering methods from short rules, long rules, tree rules to MLT rules. And better BLEU scores were achieved as we combined the different reordering rules. The MLT rules achieved better BLEU scores and TERs in both translation directions, not only when it was used alone, but also it was added to the other reordering rules. As the MLT reordering rules were used alone, it showed an improvement of 1.61 in BLEU score in the English-to-Chinese translation direction and an improvement of 2.16 in BLEU score in the Chinese-to-English translation direction, in comparison with the baseline, which used no reordering at all. As the MLT reordering rules were combined with the other existing reordering rules, a further

Data Set		Sentence Count	Word Count		Size (Byte)	
			English	Chinese	English	Chinese
Training Data	LDC	303K	10.96M	8.56M	60.88M	47.27M
	TED	151K	2.58M	2.86M	14.24M	15.63K
Development Data		919	30K	25K	164K	142K
Test Data		1663	47K	38K	263K	220K

Table 1: Data details in English-to-Chinese system

Data Set	Sentence Count	Word Count		Size (Byte)	
		Chinese	English	Chinese	English
Training Data	303K	8.56M	10.96M	47.27M	60.88M
Development Data	919	25K	30K	142K	164K
Test Data	1663	38K	47K	220K	263K

Table 3: Data details in Chinese-to-English system

improvement of 0.43 in BLEU score (from 13.38 to 13.81) was shown in the English-to-Chinese translation direction, as well as a further improvement of 0.3 in BLEU score (from 23.84 to 24.14) in the Chinese-to-English translation direction.

We've also found improvements in the sentence structure. Table 5 shows some translation examples in both translation directions. Sections are separated by double lines in the table. Each section of this table shows one translation example with the source sentence (*source*), translation without using MLT reordering (*no MLT*), translation with MLT reordering (*MLT*) and the reference (*reference*). The translation without using MLT reordering comes from the configuration with highest BLEU score that didn't use MLT reordering. And the translation with MLT reordering comes from the configuration with highest BLEU score that used MLT reordering. From the examples, we can clearly see the improvements in sentence structure.

However, by taking a close look at the gap between the scores of oracle reordering and the best scores achieved by MLT reordering, we can also see, there's still potential for improvements of translation between English and Chinese through better reordering methods.

From these experiments we can draw the conclusion that our reordering method obviously improves the sentence structure and translation quality in both English-to-Chinese and Chinese-to-English translation directions, no matter when we apply it alone or when we combine it with short rules, long rules and tree rules.

6. Conclusions

We've present a new preordering approach for translation between English and Chinese. The algorithm detects and applies reordering rules uses information of word alignment and multiple syntactic levels in the syntax tree to reorder the source sentences. Reordering patterns are detected by checking if the nested tag sequences in subtrees with any number of search levels have clearly new orders in the aligned text in the target language.

We've conducted experiments in both translation directions between English and Chinese with different SMT configurations. From the results we can see the BLEU scores was improved no matter when we applied the SMT reordering method to the baseline directly or when we combined it with the other reordering methods we introduced before, i.e. short rules, long rules and tree rules based reordering methods.

When our approach was applied alone, it achieved an BLEU score improvement of 1.61 in the English-to-Chinese translation direction, and an BLEU score improvement of 2.16 in the Chinese-to-English translation direction.

When our approach was combined with the short rules, long rules and tree rules based reordering methods, further improvements were achieved for both translation direction. Our approach improved the BLEU score further by 0.43 in the English-to-Chinese translation direction and by 0.30 in the Chinese-to-English translation direction.

Besides the improvement in BLEU scores, our preordering approach also showed improvement in the sentence structure of the translation.

7. References

[1] Rottmann, K. and Vogel, S., "Word Reordering in Statistical Machine Translation with a POS-Based Distortion Model," 2007.

[2] Niehues, J. and Kolss, M., "A POS-Based Model for Long-Range Reorderings in SMT," in *Proceedings of the Fourth Workshop on Statistical Machine Translation*, (Athens, Greece), pp. 206–214, Association for Computational Linguistics, 2009.

[3] Herrmann, T., Weiner, J., Niehues, J., and Waibel, A., "Analyzing the Potential of Source Sentence Reordering in Statistical Machine Translation," 2013.

Source	hu jintao also extended deep condolences on the death of the chinese victims and expressed sincere sympathy to the bereaved families .
No MLT	胡锦涛 还 表示 深切 哀悼 的 受害者 家属 的 死亡 , 向 迁难者 家属 表示 诚挚 的 慰问 。
MLT	胡锦涛 还 对 中国 迁难者 表示 哀悼 , 向 迁难者 家属 表示 诚挚 的 慰问 。
Reference	胡锦涛 还 对 中方 不幸 遇难 人员 表示 深切 的 哀悼 , 并 向 遇难 着 的 亲属 致以 诚挚 的 慰问 。
Source	satisfying personal interests and expanding knowledge are also major reasons why hourly work appeals to people .
No MLT	满足 个人 利益 和 扩大 知识 也 是 主要 原因 小时 工作 吸引 人 。
MLT	满足 个人 利益 和 扩大 知识 也 是 为什么 学生 工作 吸引 人 的 主要 原因 。
Reference	满足 个人 兴趣 , 扩大 自己 的 知识面 也 是 兼职 小时 工 受 青睐 的 一 个 重要 原因 。
Source	the dalai lama will go to visit washington this month .
No MLT	达赖 喇嘛 将 访问 华盛顿 的 这 一 个 月 。
MLT	达赖 喇嘛 将 本 月 访问 华盛顿 。
Reference	达赖 喇嘛 将 在 本 月 前往 华盛顿 访问 。
Source	陈至立 说 , 古巴 是 拉美 和 加勒比 地区 有 重要 影响 的 国家 。
No MLT	chen zhili said : cuba is the latin america and the caribbean region has an important influence on the state .
MLT	chen zhili said : cuba is a country of important influence latin america and the caribbean region .
Reference	chen zhili said that cuba is a country of great influence in the latin american and caribbean region .
Source	近年 来 , 两 国 教育 交流 日益 密切 , 人员 来往 频繁 。
no MLT	in recent years , the two countries education have been increasingly close exchanges and personnel contacts have been frequent .
MLT	in recent years , the educational exchanges between the two countries have become increasingly frequent , and have had frequent contacts .
Reference	in recent years , the educational exchange between the two countries has become increasingly close with frequent personnel visits .

Table 5: Examples of translations

[4] Tillmann, C., "A Unigram Orientation Model for Statistical Machine Translation," in *Proceedings of HLT-NAACL 2004: Short Papers*, pp. 101–104, Association for Computational Linguistics, 2004.

[5] Koehn, P., Axelrod, A., Birch, A., Callison-Burch, C., Osborne, M., Talbot, D., and White, M., "Edinburgh System Description for the 2005 IWSLT Speech Translation Evaluation," in *IWSLT*, pp. 68–75, 2005.

[6] Collins, M., Koehn, P., and Kučerová, I., "Clause Restructuring for Statistical Machine Translation," in *Proceedings of the 43rd annual meeting on association for computational linguistics*, pp. 531–540, Association for Computational Linguistics, 2005.

[7] Popovic, M. and Ney, H., "POS-Based Word Reorderings for Statistical Machine Translation," in *International Conference on Language Resources and Evaluation*, pp. 1278–1283, 2006.

[8] Habash, N., "Syntactic Preprocessing for Statistical Machine Translation," *MT Summit XI*, pp. 215–222, 2007.

[9] Wang, C., Collins, M., and Koehn, P., "Chinese Syntactic Reordering for Statistical Machine Translation," in *EMNLP-CoNLL*, pp. 737–745, Citeseer, 2007.

[10] Zhang, Y., Zens, R., and Ney, H., "Chunk-Level Reordering of Source Language Sentences with Automatically Learned Rules for Statistical Machine Translation," in *Proceedings of the NAACL-HLT 2007/AMTA Workshop on Syntax and Structure in Statistical Translation*, pp. 1–8, Association for Computational Linguistics, 2007.

[11] Crego, J. M. and Habash, N., "Using Shallow Syntax Information to Improve Word Alignment and Reordering for SMT," in *Proceedings of the Third Workshop on Statistical Machine Translation*, pp. 53–61, Association for Computational Linguistics, 2008.

[12] Chiang, D., "Hierarchical Phrase-Based Translation," *computational linguistics*, vol. 33, no. 2, pp. 201–228, 2007.

[13] Birch, A., Osborne, M., and Blunsom, P., "Metrics for MT Evaluation: Evaluating Reordering," *Machine Translation*, vol. 24, Mar. 2010.

[14] Birch, A., "Reordering Metrics for Statistical Machine Translation," 2011.

[15] Herrmann, T., Niehues, J., and Waibel, A., "Combining Word Reordering Methods on Different Linguistic Abstraction Levels for Statistical Machine Translation," in *Proceedings of the Seventh Workshop on Syntax, Semantics and Structure in Statistical Translation*, (Atlanta, Georgia), pp. 39–47, Association for Computational Linguistics, June 2013.