

Sentence Type Based Reordering Model for Statistical Machine Translation

Jiajun Zhang, Chengqing Zong, Shoushan Li

National Laboratory of Pattern Recognition

Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China
{jjzhang, cqzong, sshanli}@nlpr.ia.ac.cn

Abstract

Many reordering approaches have been proposed for the statistical machine translation (SMT) system. However, the information about the type of source sentence is ignored in the previous works. In this paper, we propose a group of novel reordering models based on the source sentence type for Chinese-to-English translation. In our approach, an SVM-based classifier is employed to classify the given Chinese sentences into three types: special interrogative sentences, other interrogative sentences, and non-question sentences. The different reordering models are developed oriented to the different sentence types. Our experiments show that the novel reordering models have obtained an improvement of more than 2.65% in BLEU for a phrase-based spoken language translation system.

1 Introduction

The phrase-based translation approach has been the popular and widely used strategy to the statistical machine translation (SMT) since Och, *et al.* (2002) proposed the log-linear model. However, reordering is always a key issue in the decoding process. A number of models have been developed to deal with the problem of reordering. The existing reordering approaches could be divided into two categories: one is integrated into the decoder and the other is employed as a preprocessing module.

Many reordering methods belong to the former category. Distortion model was first employed by Koehn *et al.* (2003); a lexicalized reordering model was proposed by Och *et al.* (2004) and Koehn *et al.* (2005); and the formal syntax-based reordering models were proposed by Chiang (2005) and Xiong *et al.* (2006). It is worthy to note that little syntactic knowledge is used in the models mentioned above.

Compared to the reordering models that are integrated into the decoder, the reordering at the source side can utilize more syntactic knowledge, with the goal of adjusting the source language sentence to make its word order closer to that of the target language. The most notable models are given by Xia and McCord (2004), Collins *et al.* (2005), Li *et al.* (2007) and Wang *et al.* (2007). Xia and McCord (2004) parsed the source and target sides of the training data and then automatically extracted the rewriting patterns. The rewriting patterns are employed on the input source sentence to make the word order more accordant to target language. Collins *et al.* (2005) described an approach to reorder German in German-to-English translation. The method concentrates on the German clauses and six types of transforming rules are applied to the parsed source sentence. However, all the rules are manually built. Li *et al.* (2007) used a parser to get the syntactic tree of the source language sentence. In this method, a maximum entropy model is developed to determine how probable the children of a node are to be reordered. Obviously, there is also disadvantage in this method because the parsing tree is obtained by a full parser and contains too many nodes that are not involved in desired reorderings. Wang *et al.* (2007) discussed three categories which are considered to be the most prominent candidates for reordering in Chinese-to-English translation, including verb phrases (VPs), noun phrases (NPs), and localizer phrases (LCPs). The

© 2008. Licensed under the *Creative Commons Attribution-NonCommercial-Share Alike 3.0 Unported* license (<http://creativecommons.org/licenses/by-nc-sa/3.0/>). Some rights reserved.

method deals with some special modifiers of VPs and NPs because they have the property that some specific modifiers appear before VPs or NPs in Chinese but occur after VPs or NPs in its English translation. We observe that all the transformation rules in this method are hard crafted. Furthermore, there are some other related works, such as Costa-jussa and Fonollosa's work (2006) and Zhang *et al.*'s work (2007). Costa-jussa and Fonollosa (2006) considered the source reordering as a translation task which translates the source sentence into reordered source sentence. A chunk-level reordering model was first proposed by Zhang *et al.* (2007).

However, all the existing models make no distinction between the different types of the source sentence. Intuitively, we have different reordering information in different sentence type. Taking Chinese special interrogative sentence as an example, there is a fixed phrase that usually occurs at the end of Chinese sentence but appears at the beginning part of its English translation. See the following Chinese to English translation:

Chinese: 你想要什么样的座位 ?

English: What kind of seats do you like ?

Obviously, the Chinese question phrase “什么样的座位 (What kind of seats)” should be put at the beginning of its English translation. However, many phrase-based systems fail to do this.

In this paper, we are interested in investigating the value of Chinese sentence types in reordering for Chinese-to-English spoken language translation. Due to the syntactic difference between Chinese and English, different sentence type provides different reordering information. A phrase-ahead model is developed to exploit and utilize the reordering information of special interrogative sentences. A phrase-back model is employed to catch and make use of the reordering information of other sentence types. However, the sentence type should be first identified by an SVM-based classifier before reordering the source sentence. The method overall is used as a preprocessing module for translation. We will introduce our method in detail later.

The remainder of this paper is organized as follows: Section 2 introduces our motivations; Section 3 gives the details on the implementation of our approach; the experiments are shown in Section 4; and the final concluding remarks are given in Section 5.

2 Our Motivations

In this section, before we analyze the Chinese-to-English spoken language translation corpus, some definitions are given first.

2.1 Definitions

- Special interrogative sentence / other interrogative sentence / non-question sentence

Chinese sentence can be divided into question sentence and non-question sentence. If a Chinese question sentence is translated into the English sentence of wh-questions, the sentence is named as a Chinese special interrogative sentence; otherwise, it is called the Chinese other interrogative sentence. Figure 1-3 show some examples for the three sentence types respectively.

- SQP / TP / SP

In Chinese special interrogative sentence, the question phrase is always moved ahead while it is translated into English. Correspondingly, the question phrase is named as the special question phrase (SQP). For example, the question phrase “什么样的座位 (What kind of seats)” in the example mentioned above is an SQP.

A few quantifier phrases (QPs) like “多次 (many times)”, “很多年 (many years)” in Chinese and some LCPs like “事故发生后 (after the accident happened)”, “会议结束前 (before the meeting ends)” together with some NPs like temporal phrases are named temporal phrase (TP) in our model. Some LCPs like “宾馆前 (at the front of the hotel)”, “桌子旁 (near the table)” and a few NPs like spatial phrases are called spatial phrase (SP) in our model. As PPs¹, TPs and SPs are the most prominent candidates for reordering in Chinese other interrogative sentences and non-question sentences, they will be handled in the phrase-back reordering model.

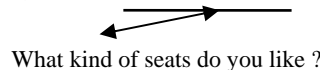
你想要什么样的座位 ?


Figure 1. An example of Chinese special interrogative sentence with its English translation.

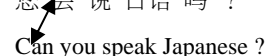
您会说日语吗 ?


Figure 2. An example of Chinese other interrogative sentence with its English translation.

¹ PPs here mean prepositional phrases

我钱包在地铁里被偷了。
 My wallet was stolen in the subway .

Figure 3. An example of Chinese non-question sentence with its English translation.

2.2 Analysis of Corpus and Our Motivations

In order to have an overview of the distribution of the Chinese sentence types, we have made a survey based on our training set for translation, which contains about 277k Chinese and English sentence pairs. We found that about 17.2% of the sentences are special interrogative sentences, about 25.5% of sentences are other interrogative sentences and the remainders are all non-question sentences.

Each sentence type has its own reordering strategy, as demonstrated in Figures 1-3. There is a settled phrase (SQP) in Chinese special interrogative sentence which usually appears at the end but will be translated first in English, just as Figure 1 illustrates. For other interrogative sentences, some specific Chinese words like “会、能、可以” will just be translated into “Can” or “Do” and come first in English. At present, this information is not used in our approach. Figure 2 gives an example. For non-questions, the reordering candidates usually need to be moved back during translation. An example is shown in Figure 3.

According to the analysis above, it is meaningful to develop reordering models based on the source sentence types.

2.3 Framework

As we mentioned above, our framework is illustrated as follows:

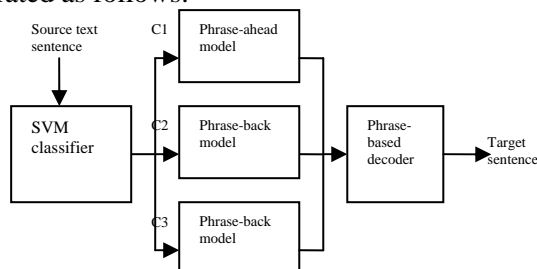


Figure 4. Architecture of the framework, where C1 means the special interrogative sentence, C2 is other interrogative sentence and C3 is non-question sentence.

Conventional preprocessing approaches divide the translation into two phases:

$$S \rightarrow S' \rightarrow T \quad (1)$$

Reordering is first done in the source side which changes the source sentence S into reordered one S' , and then a standard phrase-based translation engine is used to translate the reordered source sentence S' into target language sentence T .

In our method, to utilize the information of sentence types, a new approach is proposed to improve the translation performance by developing a hybrid model as follows:

$$S \rightarrow S^c \rightarrow S' \rightarrow T \quad (2)$$

Before the source sentence is reordered, an SVM-based classifier is first employed to determine its sentence type S^c , then, different reordering model is used to reorder the source sentence with the specific sentence type S^c . After getting the reordered source sentence S' , we use our phrase-based SMT to obtain the optimal target language sentence.

The contribution of this paper is embodied in the first two steps of our method.

In the first step, an SVM classifier is used to identify the type of source sentence².

In the second step, two reordering models are built according to the different sentence types. A phrase-ahead reordering model is developed for the special interrogative sentences which uses shallow parsing technology to recognize the most prominent candidates for reordering (special question phrase) and extracts reordering templates from bilingual corpus. For other sentence types, we build a phrase-back reordering model which uses shallow parsing technology to identify the phrases that are almost always moved back during translation and applies maximum entropy algorithm to determine whether we should reorder them.

3 Models and Algorithms

In this section, we first introduce the sentence type classifier model, and then we describe in detail the two reordering models, phrase-ahead reordering model and phrase-back reordering model.

3.1 Sentence Type Identification

Many models are used for classification such as Naïve Bayes, decision tree and maximum entropy. In our approach, we use an SVM-based classifier to classify the sentence types. SVM

² There are three sentence types: special interrogative sentence, other interrogative sentence and non-question sentence, which are defined in sub-section 2.1.

has been shown to be highly effective at traditional text categorization. For our problem, we regard a sentence as a text. The decision boundary in SVM is a hyperplane, represented by vector \vec{w} , which separates the two classes, leaving the largest margin between the vectors of the two classes (Vapnik, 1998). The search of margin corresponds to a constrained optimization problem. Suppose $c_j \in \{1, -1\}$ (positive and negative) be the correct class of sentence s_j , the solution can be formalized as:

$$\vec{w} := \sum_j \alpha_j c_j \vec{s}_j \quad \alpha_j \geq 0 \quad (3)$$

Where the \vec{s}_j is feature vector of our sentence s_j . We get α_j s through solving a dual optimization problem. Identifying the type of a sentence is just to determine which side of \vec{w} 's hyperplane it will fall in.

Feature selection is an important issue. We directly use all the words occurring in the sentence as features.

Some readers may argue that the features to distinguish the sentence types are very obvious in Chinese. For example, “?” can easily separate the interrogative sentences from non-question sentences. In this case, a simple classifier like decision tree will work. It is true when the punctuation always appears in the sentence. However, sometimes there is no punctuation in the spoken language text. Under this situation, the decision tree will lose the most powerful features, but the performance of SVM is not affected by the punctuations. The experimental results verifying this will be given in Section 4.

3.2 Phrase-ahead Reordering Model

As we mentioned above, about 17.2% of the spoken language sentences are special interrogative sentences. Furthermore, we note that each Chinese special interrogative sentence has one or more special question phrases (SQP) that we defined in section 2.1. Due to the difference between Chinese and English word order, the SQP needs to be moved ahead³ when it is translated into English.

Let S be a Chinese special interrogative sentence, our first problem is to recognize the SQPs in S . If we have known the SQP, namely S becomes $S^0 \text{ SQP } S^1$ (S^0 is the left part of the

sentence before SQP, and S^1 is the right part of the sentence after SQP), our second problem is to find the correct position in S^0 where SQP will be moved to.

For the first problem, because each syntactic component is possible a SQP, for example, “什么样的座位” in Figure 1 is NP, “在哪里 (Where)” in Chinese sentence “我在哪里能买到票? (Where can I buy the ticket?)” is PP (also a VP modifier), “怎么走 (How to go)” in “去海滨怎么走? (How to go to the beach?)” is VP, it is very difficult to find the SQP by syntax. In our model, we first find out all the key words, which we list below, in the special interrogative sentences through mutual information. Then, we define the syntactic component containing the key word as an SQP. Instead of full syntactic parser, we utilize a CRF toolkit named FlexCrfs⁴ to train, test and predict the SQPs chunking.

什么	What
哪(哪里/哪儿...)	Where
多(多大/多长...)	How much/many/old...
怎(怎样/怎么办...)	What about/How
谁(谁的/是谁...)	Who/whose/whom
几(几点/几个...)	How many/old When...
为什么	Why
何(何时/何地...)	When/where

Table 1. The special key words set

For the second problem, we note that there are only three positions where the SQP will be moved to: (1) the beginning of the sentence; (2) just after the rightmost punctuation (“;”, “,” or “:”) before the SQP; (3) or after a regular phrase such as “请问 (May I ask)” and “告诉我 (Please tell me)”. Therefore, we can learn the reordering templates from bilingual corpus⁵. The simple algorithm is illustrated in Figure 5, and some reordering templates are shown in Table 2.

On the whole, When we reorder the special interrogative sentence, we first identify the SQP, then we find out whether there are punctuations (“;”, “,” or “:”) before SQP; if any, we keep the rightmost punctuation index, otherwise we keep the index 0 (beginning of sentence). In the third

³ There is a specific situation that the SQP don't have to be moved. In this case, we suppose it needs to be moved, but the distance is 0.

⁴ See <http://flexCRF.sourceforge.net>

⁵ The bilingual corpus is the corpus combined by the training corpus for chunking SQPs and its corresponding English translation.

step, if we find that a reordering template like some one given in Table 2 can match the sentence, we just apply the template, otherwise we just move the SQP after the index that we kept before (0 or punctuation index).

```

1: Input: special interrogative sentence pair (s, t) in which
SQP is labeled and their alignment M is given
2: R={}
3: Find the rightmost punctuation index c_punc_index before
SQP and English index e_punc_index aligned to
c_punc_index
4: Find the smallest index e_smallest_index of English which
align to the SQP
5: Get the Chinese phrase C_Phrase which aligns to
[e_punc_index+1, e_smallest_index-1]
6: if C_Phrase is NONE then
7: Continue ;
8: end if
9: if C_Phrase in R then
10: Count(C_Phrase)++;
11: else
12: Insert C_Phrase into R;
13: Count(C_Phrase)=1;
14: end if
15: remove C_Phrase if Count(C_Phrase)<N
16: output R

```

Figure 5. Reordering template extraction algorithm. The empirical value N is 10 in our experiment.

X1 请问 X2 SQP	X1 请问 SQP X2
X1 告诉我 X2 SQP	X1 告诉我 SQP X2
X1 我想知道 X2 SQP	X1 我想知道 SQP X2
X1 然而 X2 SQP	X1 然而 SQP X2
.....

Table 2. Some reordering templates

3.3 Phrase-back Reordering Model

In this paper, we employ the phrase-back reordering model for Chinese other interrogative sentences and non-question sentences.

Inspired by the work of Wang *et al.* (2007), we only consider the most prominent candidates for reordering. The VP modifiers like PP, TP, and SP which we defined in sub-section 2.1 are typically in pre-verb position in Chinese but almost always appear after the verb in its corresponding English translation. Wang *et al.* (2007) concentrate on VP, NP, then determine whether their modifiers should be moved back. Instead, our interests are focused on the modifiers: PP, TP and SP; namely, we consider the modifiers PP, TP and SP as triggers, and the first VP occurring after triggers will be the candidate position where the triggers may be moved to.

Changing the focus gives us the ability to handle a specific situation that there is no VP after the triggers for recognition error or other

reasons. As the example in Figure 6, there is no VP after PP (“在这里”) because the phrase “签名” next to PP is wrongly recognized to be a NP. To deal with the case, we will further define a fake verb phrase (FVP): the phrase after PP (TP or SP) until the punctuation (“;”, “,” or “.”). The phrase “签名 (sign your name)” in Figure 6 is an FVP. Here, FVP is given the same function with VP, thus it makes our model suitable for more situations.

请在这里 签名。

Figure 6. An example of FVP. In our model the whole sentence is recognized as a VP, “在这里 (here)” is a PP, and “签名 (sign your name)” is identified as a NP.

Unlike hard reordering rules of Wang *et al.* (2007), we develop a probabilistic reordering model to alleviate the impact of the errors caused by the parser when recognizing PPs, TPs, SPs and VPs. We believe that no reordering is better than bad reordering. The rule forms and the probabilistic model will be given as follows:

- The form of phrase-back reordering rules:

$$A : A_1 X A_2 \Rightarrow \begin{cases} A_1 X A_2 & \text{straight} \\ X A_2 A_1 & \text{inverted} \end{cases}$$

Where, $A_1 \in \{PP, TP, SP\}$ $A_2 \in \{VP, FVP\}$
 $X \in \{\text{phrases between } A_1 \text{ and } A_2\}$

- We use the Maximum Entropy Model which is implemented by Zhang⁶. The model is trained from bilingual spoken language corpus to determine whether A_1 should be moved after A_2 . The features that we investigated include the leftmost, rightmost, and their POSs of A_1 and A_2 . It leads to the following formula:

$$P(O|A) = \frac{\exp(\sum_i \lambda_i h_i(O, A))}{\sum_o \exp(\sum_i \lambda_i h_i(o, A))} \quad (4)$$

Where, $O \in \{\text{straight, inverted}\}$, $h_i(O, A)$ is a feature, and λ_i is the weight of the feature.

When applying the rules, we first identify pairs like ($A_1 X A_2$) in the sentence, and then from beginning to end of the sentence, we move A_1 behind A_2 if $P(\text{inverted} | A) > P(\text{straight} | A)$. After all the pairs are processed, we will get the reordered source result.

⁶http://homepages.inf.ed.ac.uk/s0450736/maxent_toolkit.html

4 Experiments

We have conducted several experiments to evaluate the models. In this section, we first introduce the corpora, and then we discuss the performance of the SVM-based classifier, chunking and reordering models respectively.

4.1 Corpora

We perform our experiments on Chinese-to-English speech translation task. The statistics of the corpus is given in Table 3 where CE_train means the Chinese-to-English training data released by IWSLT 2007; CE_sent_filtered means the bilingual sentence pairs filtered from the open resources of the bilingual sentences on the website; CE_dict_filtered means the bilingual dictionary filtered from the open resources of the bilingual dictionaries on the website; CE_dev123 denotes the bilingual sentence pairs obtained by the combination of the development data IWSLT07_CE_devset1, IWSLT07_CE_devset2 and IWSLT07_CE_devset3 which are released by the IWSLT 2007; CE_dev4 and CE_dev5 are the remainder of development data released by IWSLT 2007; CE_test means the final test set released by IWSLT 2007.

We combine the data from the top four rows as our training set. We use CE_dev4 as our development set. CE_dev5 and CE_test are our two test data. The test data released by IWSLT 2007 is based on the clean text with punctuation information, so we add the punctuation information on the Chinese sentences of CE_dev4 and CE_dev5 by our SVM sentence type classifier to form the final development set. The detailed statistics are given in Table 4.

4.2 Classification Result

To evaluate the performance of SVM-based classifier on classifying the sentence types, we first use a simple decision tree to divide the Chinese sentences of our training data for translation into three sentence types. Then we clean them by hand in order to remove the errors. At last, 10k sentences for each sentence type are randomly selected as the experiment data. For each sentence type, 80% of the data are used as training data, 20% as test data. Table 5 gives the classification results.

Punctuation in Table 5 means the punctuation which occurs at the end of the sentence such as “。” and “?””. We can see from the table that SVM classifier performs very well even if we remove the punctuations at the end of every sen-

tence. Therefore, almost no errors will be passed to the reordering stage.

Data	Chinese	English
CE_train	39,953	39,953
CE_sent_filtered	188,282	188,282
CE_dict_filtered	31,132	31,132
CE_dev123	24,192	24,192
CE_dev4	489	3,423
CE_dev5	500	3,500
CE_test	489	2,934

Table 3. Statistics of training data, development data and test data

		Chinese	English
Train set	sentences	276,633	
	words	1,665,073	1,198,984
Dev set CE_dev4	sentences	489	489*7
	words	6241	47609
Test set CE_dev5	sentences	500	500*7
	words	6596	52567
Test set CE_test	sentences	489	489*6
	words	3166	22574

Table 4. Detailed statistics of training data on development set

	Accuracy (%)
With punctuation	99.80
Without punctuation	98.00

Table 5. The accuracy of SVM classifier

4.3 Chunking Results

In our experiment, except that VPs are obtained by a syntactic parser (Klein and Manning, 2003), SQPs, PPs, TPs, SPs are all chunked by the FlexCrfs.

The chunking data used for training and test in Table 6 are annotated by ourselves. Every chunk is annotated according to the definition that we define in sub-section 2.1. The raw training and test data are all extracted from our training set for translation. TPs, SPs are annotated together; SQPs, PPs are annotated respectively. The statistics of the training and test data are shown in Table 6. Table 7 gives the chunking results.

The precision, recall and F-Measure are metrics for the chunking results. F-Measure follows the criteria of CoNLL-2000⁷.

$$F - Measure = \frac{2 * (precision * recall)}{precision + recall}$$

⁷ See <http://www.cnts.ua.ac.be/conll2000/chunking/>

Because the SQPs have the regularity that each one contains a key word listed in Table 1, the result of SQPs chunking is quite good. Moreover, the chunking of PPs, TPs and SPs also performs well.

		Train	Test
SQP	sentences	10,000	500
	chunks	10030	501
PP	sentences	10,000	500
	chunks	10106	512
SP and TP	sentences	11,000	500
	chunks	10342	523

Table 6. Statistics of train and test data

	Precision (%)	Recall (%)	F-Measure (%)
SQP	95.52	95.52	95.52
PP	94.65	93.31	93.98
SP and TP	93.92	92.68	93.25

Table 7. Chunking results on test set

4.4 Translation Results

For the translation experiments, BLEU-4 and NIST are used as the evaluation metric. The baseline SMT uses the standard phrase-based decoder that applies the log-linear model (Och and Ney, 2002).

In the preprocessing module, all the Chinese words are segmented by the free software toolkit ICTCLAS3.0⁸, and the POS tags are obtained by using the Stanford parser with its POS parsing function. For the decoder, the phrase table is obtained as described in (Koehn *et al.*, 2005), and our 4-gram language model is trained by the open SRILM⁹ toolkit. It should be noted that we use monotone decoding in translation.

We have done three groups of experiments for translation. The first one is to test the effect of phrase-ahead reordering model, the result of which is shown in Table 8. Compared to the baseline system, phrase-ahead reordering model improves the results of the two test sets by 0.41% and 1.87% in BLEU respectively. The difference in the performance gains can be attributed to the fact that there are 100 Chinese special interrogative sentences in Test 2, while only 30 are found in Test 1. Accordingly, the reordering candidates of Test 1 are much fewer than that of Test 2. Thus, we can conclude that the more special interrogative sentences the better performance of the translation. Furthermore,

the results show that the reordering on special interrogative sentences is a good try.

The second experiment is conducted to test the effect of phrase-back reordering model. Table 8 gives the results. For the two test sets, the model brings an improvement to the baseline by 2.24% and 0.93% in BLEU respectively. However, the difference between them is still very big. We think there are two reasons: firstly, there are much more special interrogative sentences in Test 2 than in Test 1, so the sentences of other sentence types in Test 2 are much fewer than that in Test 1. Thus, fewer candidates are found in Test 2 than in Test 1. Secondly, the average sentence length of Test 2 (6.5 words) is much shorter than that of Test 1 (13.2 words). We know that if the sentence is very short, the PP, TP, and SP will seldom occur. Naturally, only 89 candidates are found in Test 2 but 366 in Test 1. Regardless of the difference, the phrase-back reordering model indeed improves the translation quality significantly.

The last experiment merges the two reordering model together. The results in Table 8 show that the overall reordering model has done very well in both test sets: it improves the two test sets by 2.65% and 2.78% in BLEU score respectively. It demonstrates that every reordering model has a positive effect on translation. Therefore, our reordering model based on the sentence type is quite successful.

5 Conclusions and Future Work

In this paper, we have investigated the effect of the Chinese sentence types on reordering problem for Chinese-to-English statistical machine translation. We have succeeded in applying a phrase-ahead reordering model to process the special interrogative sentences and a phrase-back reordering model to deal with other sentence types. Experiments show that our reordering model obtains a significant improvement in BLEU score on the IWSLT-07 task.

With the encouraging experimental results, we believe that we can mine more reordering information from the Chinese sentence types. In this paper, we only apply a phrase-back model to reorder Chinese other interrogative sentences. In the next step, we will try to develop a special reordering model for this sentence type. Furthermore, we plan to integrate the phrase-back model into phrase-ahead model for special interrogative sentences and investigate the value of this integration.

⁸ See <http://www.nlp.org.cn>

⁹ See <http://www.speech.sri.com/projects/srilm>

		BLEU (%)	NIST	Sentences	Candidates
Test 1 CE_dev5	Baseline	32.16	6.4844	500	
	Phrase-ahead reordering	32.57	6.5579	30	31
	Phrase-back reordering	34.40	6.6857	470	366
	Phrase-ahead+phrase-back	34.81	6.7584	500	397
Test 2 CE_test	Baseline	34.04	5.8340	489	
	Phrase-ahead reordering	35.91	6.0693	100	97
	Phrase-back reordering	34.97	5.9172	389	89
	Phrase-ahead+phrase-back	36.82	6.1535	489	186

Table 8. Statistics of translation results

Notes: candidates here mean how many candidate reordering phrases are recognized for each model. Sentences mean the number of sentences belonging to the specific sentence type, i.e. for phrase-ahead reordering in Test 1, 31 special question phrases (SQP) are recognized in 30 Chinese special interrogative sentences.

Acknowledgments

The research work described in this paper has been partially supported by the Natural Science Foundation of China under Grant No. 60575043 and 60736014, the National High-Tech Research and Development Program (863 Program) of China under Grant No. 2006AA01Z194 and 2006AA010108, the National Key Technologies R&D Program of China under Grant No. 2006BAH03B02, and Nokia (China) Co. Ltd as well.

References

- Cao Wang, Michael Collins and Philipp Koehn. 2007. Chinese syntactic reordering for statistical machine translation. *In Proceedings of joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning, 2007.*
- Chi-Ho Li, Dongdong Zhang, Mu Li, Ming Zhou Minghui Li and Yi Guan. 2007. A probabilistic approach to syntax-based reordering for statistical machine translation. *In Proceedings of 45th Meeting of the Association for Computational Linguistics .*
- Dan Klein and Christopher D. Manning. 2003. Accurate unlexicalized parsing. *In Proceedings of 41st Meeting of the Association for Computational Linguistics.*
- David Chiang. 2005. A hierarchical phrase-based model for statistical machine translation. *In Proceedings of 43rd Meeting of Association for Computational Linguistics.*
- Deyi Xiong, Qun Liu and Shouxun Lin. 2006. maximum entropy based phrase reordering model for statistical machine translation. *In Proceedings of the joint conference of the International Committee on Computational Linguistics and the Association for Computational Linguistics 2006.*
- Fei Xia and Michael McCord. 2004. Improving a Statistical MT system with automatically learned rewrite patterns. *In Proceedings of 20th International Conference on Computational Linguistics.*
- Franz Josef Och and Hermann Ney. 2002. Discriminative training and maximum entropy models for statistical machine translation. *In Proceedings of 40th Meeting of Association for Computational Linguistics.*
- Franz Josef Och and Hermann Ney. 2004. The alignment template approach to statistical machine translation. *Computational Linguistics, 30:417-449*
- Marta R. Costa-jussa and Jose A.R. Fonollosa. 2006. Statistical machine reordering. *In proceedings of Conference on Empirical Methods in Natural Language Processing 2006.*
- Michael Collins, Philipp Koehn, and Ivona Kucerova. 2005. Clause restructuring for statistical machine translation. *In proceedings of 43rd Meeting of the Association for Computational Linguistics.*
- Philipp Koehn, Franz J. Och. and Daniel Marcu. 2003. Statistical Phrase-based Translation. *In proceedings of HLT-NAACL 2003.*
- Philipp Koehn, Amittai Axelrod, Alexandra Birch Mayne, Chris Callison-Burch, Miles Osborne and David Talbot. 2005. Edinburgh System Description for the 2005 IWSLT Speech Translation Evaluation. *In International Workshop on Spoken Language Translation.*
- Yuqi Zhang, Richard Zens and Hermann Ney. 2007. Chunk-level reordering of source language sentence with automatically learned rules for statistical machine translation. *In Proceedings of SSST, NAACL-HLT 2007/AMTA Workshop on Syntax and Structure in Statistical Translation.*
- Vladimir Naumovich Vapnik. 1998. Statistical Learning Theory. John Wiley and Sons, Inc.