A Multi-Level Reordering Model for Statistical Machine Translation Using Source Side Parse Trees

Mohammad Mehdi Kholoosi †, Abdoreza Rezapour††, Mohammad Hadi Sadreddini and Seyed Mostafa Fakhrahmad

Department of Computer Science & IT, School of Electrical and Computer Engineering, Shiraz University, Shiraz, Iran.

Summary

In translating between a pair of languages, reordering is a major task, which is roughly defined as finding the right order of words in target language. Word reordering is a key element affecting the machine translation quality and one of its serious difficulties as well. In this paper, we present a new reordering model based on POS tags and syntactical information that exists in source sentence's parse trees. In order to use this information properly, we proposed an innovative method that reorders sentences on two different levels (i.e., phrase and word level). This method considers relationships among the words in a sentence and performs reordering with respect to the sentence structure, unlike only POS-based models. We examined this model on English-Persian language pair. Our experiments showed that this model can improve the measure of precision and reorder sentences more reliably than previous approaches.

Key words:

Reordering, Machine Translation, Parse Trees, POS-based reordering model.

1. Introduction

Non-similar word orders in different languages is a major challenge in statistical machine translation. If a machine translator doesn't have a suitable reordering mechanism, the produced translation will not have adequate quality, no matter how good the machine is in other aspects of MT[1] such as WSD[2]. Recently, due to the importance of machine translation in modern life, the reordering problem has drawn the attention of many researchers, and hence, different approaches have already been proposed to overcome this problem.

The reordering models which only rely on rule collections could not achieve a good performance, because every POS tag in a rule represents an independent word. Therefore, these models move words in a sentence without considering the relationships among them. On the other hand, Parse trees are syntactic structures which demonstrate the relationships among components of a sentence. In present study, we propose a novel method which makes use of a big reordering rule collection as well as syntactic information presented by parse trees. In this method, first, words of each phrase are reordered separately, and then resulting phrases are reordered to constitute the whole sentence in target language. Throughout this procedure, related components will stay together and reordering will be accomplished, more accurately.

This paper is organized as follows. Section 2 reviews the related work on the reordering problem. Section 3 is devoted to introduction of the new reordering model. In Section 4, experiments on the English-Persian language pair are described, and the results are discussed. Finally, Section 5 concludes the paper.

2. Related Work

Several reordering methods have already been proposed in machine translation (SMT[3]) studies. The set of reordering methods can be divided into three major categories:

(a) Jointly carrying out word selection and reordering. These include phrase-based SMT[1], hierarchical phrase-based SMT[2], and syntax-based SMT[3,4,5,6]

In a phrase-based SMT system reordering can be achieved during decoding by allowing swaps of words within a defined window. Lexicalized reordering models[7,8] include information about the orientation of adjacent phrases that is learned during phrase extraction. This reordering method, which affects the scoring of translation hypotheses but does not generate new reorderings, is for example used in Moses[9] which is an open source machine translation system. Syntax-based or syntax-augmented MT systems address the reordering problem by embedding syntactic analysis in the decoding process. Hierarchical MT systems construct a syntactic hierarchy during decoding, which is independent from linguistic categories.

(b) Pre ordering: Pre ordering is a popular approach in overcoming the word ordering problem. These

[1] Machine Translation
[2] Word Sense Disambiguation

E-mail:
† mehdi.kh2012@gmail.com
†† A.R.Rezapour@asaiau.ac.ir

[3] Statistical Machine Translation

Manuscript received July 5, 2017
Manuscript revised July 20, 2017

methods reorder the source language into the target language order before translating it. Many different pre-ordering strategies have been proposed: Deterministic preordering aims at finding a single optimal reordering for each input sentence, which is then translated monotonically or with a low distortion limit[10,11,12,13,14,15].

In[16], Lerner and Petrov present a simple classifier-based preordering approach. They combine the strengths of lexical reordering and syntactic preordering models by performing long-distance reorderings using the structure of the parse tree, while utilizing a discriminative model with a rich set of features, including lexical features.

Yang et al in[17] present a ranking based reordering method to reorder source language to match the word order of target language given the source side parse tree. Reordering is formulated as a task to rank different nodes in the source side syntax tree according to their relative position in the target language. The ranking model is automatically trained to minimize the mis-ordering of tree nodes in the training data.

Non-deterministic preordering encodes multiple alternative reorderings into a word lattice (or forest) and lets a monotonic (usually n-gram based) decoder choose the best path according to its models[18,19,20,21,22].

In[23], Elming and Habash extend a pre-translation syntactic reordering approach developed on a close language pair (English-Danish) to the distant language pair, English-Arabic. They achieve improvements in translation quality over related approaches, measured by manual as well as automatic evaluations. They also examined the effect of the alignment method on learning reordering rules. Their experiments produced better translation using rules learned from automatic alignments than using rules learned from manual alignments.

A hybrid approach is presented in[24,25,26]. In this approach, rules are used to generate multiple likely pre-orderings, but only for a specific language phenomenon that is responsible for difficult (long-range) reordering patterns. The sparse reordering lattices produced by these techniques are then translated by a decoder performing additional phrase-based reordering. Bisazza and Federico in[25] introduce another way to encode multiple pre orderings of the input: instead of generating a word lattice, pre-computed permutations are used to reduce the distortion cost and create ‘shortcuts’ between selected pairs of input positions (i. e. modified distortion matrices). At the price of some approximations, this technique allows for a more compact input representation because, unlike lattices, it does not involve the creation of multiple nodes for the same source word.

(c) Post ordering: Post ordering is the other approach which translate source words into target words monotonously. Then, the translated words are reordered into the target language word order[27,28,29,30].

Goto et al in[27] propose the post-ordering framework for Japanese to English machine translation. They reorder the sequence of target words thereby parsing the translated words to obtain syntax structures and then transferring the obtained syntax structures into the syntax structures of the target language.

In[31], Farzi et al present a translation system and syntactically-informed post reordering models. They exploit sophisticated syntactic-based features for re-ranking the N-best translation candidates provided by a phrase-based statistical machine translation system. Their sophisticated reordering features are based on an innovative structure, named, phrasal dependency tree inspired from dependency relations between contiguous non-syntactic phrases. The features gain benefits from phrase dependencies, translation directions and translation distance.

3. The Proposed Reordering Model

The main process of the proposed model consists of two major phases. Figure 1 presents the block diagram of the proposed scheme.

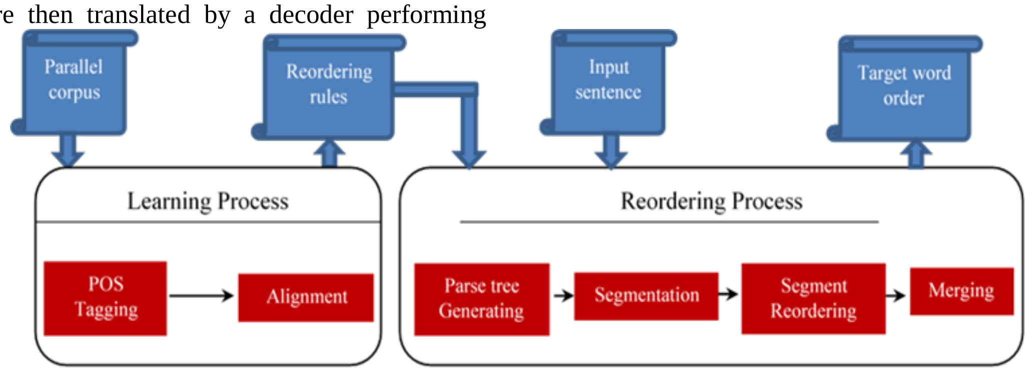


Fig. 1 A high-level view of the proposed method

3.1 Making a Big Reordering Rule Collection

In order to construct the desirable rule collection, POS tags of the words should be determined in both sides of the parallel corpus. For the English-Persian language pair, we used Mizan English-Persian corpus[32], which contains about one million sentences. Both sides of the corpus have been tagged using part of speech taggers. For the English side we used Stanford Log-linear Part-Of-Speech Tagger[33] with left3words-distim.tagger model rather than bidirectional model because it is much faster. For the Persian side, we used Ferdowsi University Persian POS Tagger[34] that to best of our knowledge, it is the best tagger available for Persian language.

We extracted tag sequences from the output of taggers as shown in Example 1.

EXAMPLE 1. Consider the following source sentence:

Source sentence: this system has a good performance.

The xml output of the POS-tagger will be as follows:

<sentence id="0">
 <word wid="0" pos="**DT**">this</word>
 <word wid="1" pos="**NN**">system</word>
 <word wid="2" pos="**VBZ**">has</word>
 <word wid="3" pos="**DT**">a</word>
 <word wid="4" pos="**JJ**">good</word>
 <word wid="5" pos="**NN**">performance</word>
 <word wid="6" pos=".">.</word>
</sentence>

Hence, the extracted tag sequence is: DT+NN+VBZ+DT+JJ+NN

All of the generated sequences from both sides of the corpus are then inserted into a database such that every tuple includes English-Persian tag sequences of the same sentence. Some sample sentences from the parallel corpus as well as their tag sequences in both languages are presented in Tables 1 and 2, respectively.

Table 1: A set of sample sentences from the parallel corpus

English sentence	Persian sentence
They were but wonderful	اما خیلی شگرف بودند
My proper share was a minor one	سهم واقعی من ناچیز بود
We were fond together	ما به یکدیگر علاقمند بودیم
The men were young and sturdy	افراد همه جوان و قوی بودند

Table 2: Tag sequences of the sentences presented in Table-1

English tag sequence	Persian tag sequence
PRP+VBD+CC+JJ	CC+RB+JJ+VB
PRP$+JJ+NN+VBD+DT+JJ+CD	NN+JJ+PRP+JJ+VB
PRP+VBD+JJ+RB	PRP+IN+PRP+NN+VB
DT+NNS+VBD+JJ+CC+JJ	PSUS+NN+JJ+CC+JJ+VB

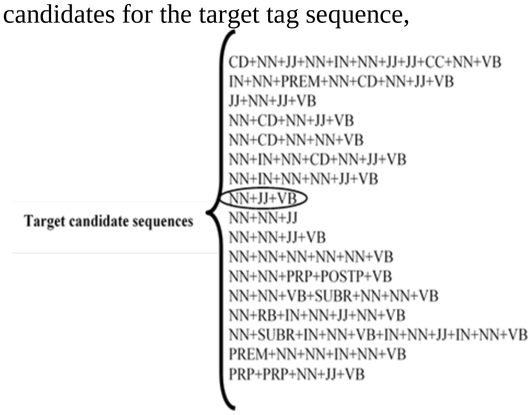
Automatic Reordering. In order to reorder a new sentence, first, we tag it by the source side tagger and generate the tag sequence. Then, we conduct a matching search on the source side of the database. It is likely either no tuple or numerous tuples (target sequence) will be retrieved according to the searched sequence (source sequence). If more than one tuple is retrieved from the database, then the model has to choose the best sequence. We consider two simple principles for selecting the winner sequence. The first principle is that the target sequence should have the maximum number of tokens (POS tag) in common with the source sequence, and the second principle is that the target sequence should have the minimum length among all candidates, as shown in example 2.

EXAMPLE2. Remember the sentence and its generated tag sequence, given in example 1:

Source sentence: This system has a good performance

Source tag sequence: DT+NN+VBZ+DT+JJ+NN

If the following sequences have been retrieved as candidates for the target tag sequence,



The sequence [NN+JJ+VB] is selected since it has three tokens in common with the source sequence, and it is the shortest one too. After the best target language sequence is selected, we reorder the words of the source sentence according to the order of the selected target sequence using a queue data structure and generate a fluent sentence in target language. In execution of the mentioned process, four situations would happen. As shown in Figure 2, the sequence already exists as an exact match or by adding some tokens to the beginning or the end of it.

Fig. 2 Availability situations for a sequence in the rule collection

However, an important question is that what happens if the sequence does not exist in the database, and none of the four situations has occurred. To encounter this problem, we use a split function which removes a token from the end of sequence and then repeats the search process until finding a proper matching sequence. In our experiments, we observed that if the sentence is real and meaningful, finally there will be an appropriate target sequence due to our vast different sentence structures in the database.

3.2 The Multilevel Parsing Method

The main idea of this method is that reordering has to be done at different structural levels, i.e., word and phrase level. In this work, we used Stanford lexical parser[35] to generate parse trees. This method has four stages, as will be discussed.

3.2.1 Breaking source side sentence to smaller segments

After the parse tree is generated, we traverse it top down in order to find the boundaries of sub-sentences. The node "S" defines beginning of a sentence, so every node after it until the next "S" node is a member of the first segment.

As an example, the sentence "This was not at all the story I had expected him to tell me" consists of three segments. The segments are shown in Figure 3.

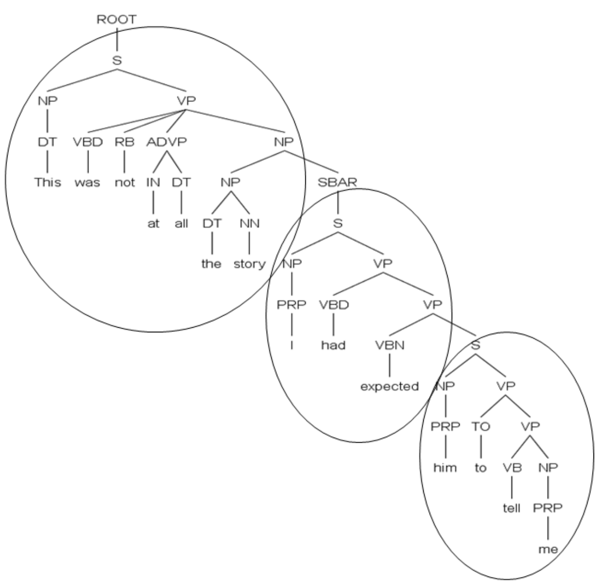


Fig. 3. Segmentation of the sentence "This was not at all the story I had expected him to tell me"

3.2.2.Reordering phrases at the word (leaf) level in each segment

The following four-step process is carried out in this stage:

(a) Phrases that have only one child are marked

(b) Siblings of each phrase (except the marked ones in step (a)) are marked if they are leaves

(c) Leaves of each phrase except the leaves marked in step (b) are read and recorded as a sequence

(d) The generated sequence in step (c) is given to the rule collection in order to map to the best target sequence

The mentioned steps are illustrated for the previous sample sentence as follows.

The first segment of the sentence in stage 1 is "This was not at all the story".

(a) As shown in figure 4, "NP" is marked, since it is the only node having one child (i.e., DT).

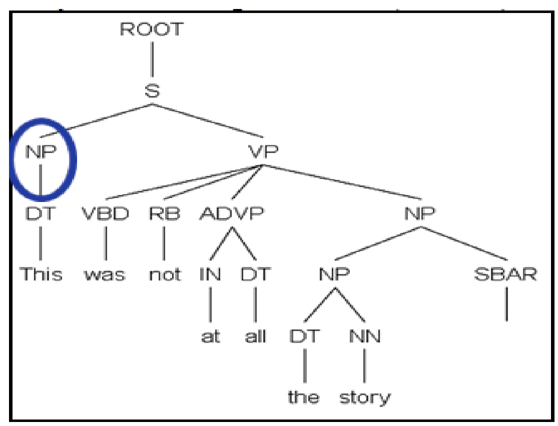


Fig. 4. The first step of running stage 2

(b) As shown in Figure 5, "VBD" and "RB" are siblings of "ADVP" and "NP". They are leaves and haven't been marked in step (a).

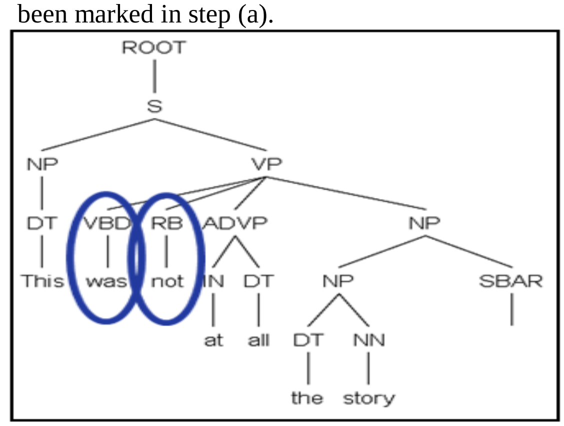


Fig. 5. The second step of running stage 2

(c) As shown in Figure 6, "DT", "IN+DT" and "DT+NN" are leaves and children of "NP", "ADVP" and "NP", respectively. They haven't been marked in step (b).

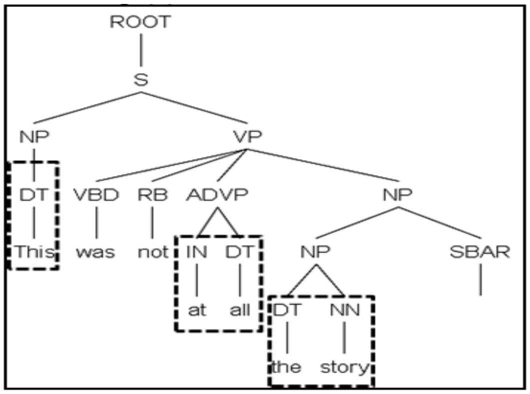
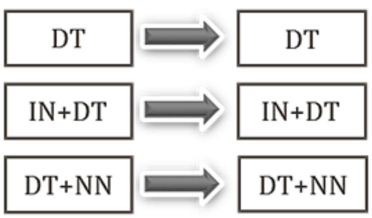
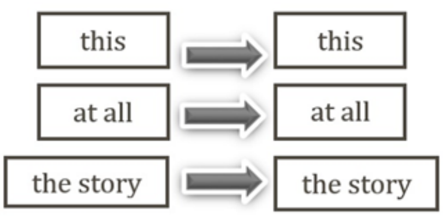


Fig. 6. The third step of running stage 2

(d) As shown in Figure 7 (a), the sequence “DT”, “IN+DT”, “DT+NN” is given to the rule collection and the best target language sequence is retrieved. Figure 7 (b) shows the order of words before and after reordering.



(a)



(b)

Fig. 7. The forth step of running stage 2
(a) Sending the generated sequence to the rule collection and retrieving the target sequence
(b) Order of words before and after of reordering

3.2.3. Reordering phrases at the phrase level in each segment

In this stage, each phrase is first transformed to its equivalent POS tag. For example, “NP” and “VP will be transformed to “NN” and “V”, respectively. Then, the transformed phrases are reordered level by level, from top to bottom of the tree. If any of the nodes had been marked by stage 2.b, they would be added to the generated sequence.
As shown in Figure 8, this segment has two levels. In the first level “NP” and “VP” will be transformed to “NN” and “V” and in the second level “ADVP” and “NP” will be transformed to “RB” and “NN”, respectively.

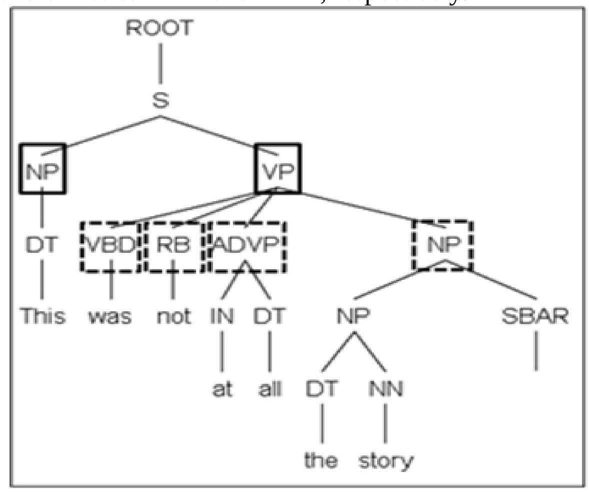


Fig. 8. Two levels of phrases in the parse tree

The best target language sequences retrieved for both levels are shown in Figure 9.

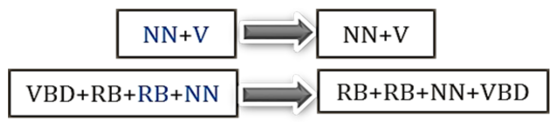


Fig. 9. Sending and retrieving sequences of both levels

From the second level, “VBD” and “RB” were added to the generated sequence since they both have been marked by stage 2.b.
“VBD” moved to the end of the sequence, and one reordering occurred here.
Figure 10 shows the order of the phrases before and after reordering.

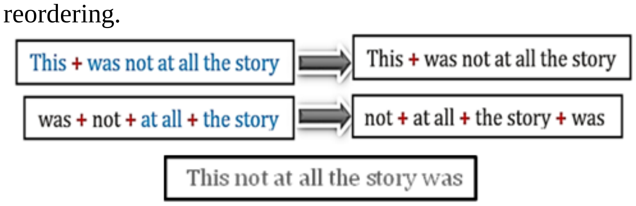


Fig. 10. Order of phrases before and after of reordering

3.2.4. Merging the results together

In this step, the generated results of each segment are merged in order to build a sentence with a meaningful word order in target language.
Figure 11 shows the merged sentence in target language word order. If we translate it to target language word by word, we will end up with a fluent sentence.

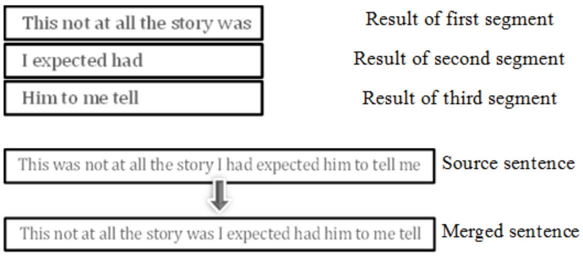


Fig. 11. The sentence before and after of reordering

If we translate this sentence to Persian language, we will have:
" این به هیچ وجه داستانی نبود که من انتظار داشتم او به من بگوید "

4. Experimental Results

We used PCTS [1] [36] as the test dataset. This corpus contains 400 English-Persian sentences. First of all, alignments were extracted using GIZA++ toolkit[37] and reviewed manually. These alignments were used in creating numerical order sequences that represent the correct word order of source sentences in target language.
As shown in Table 3 the correct word orders for translating the English sentence "he goes to school" to Persian is "1-3-2-4".

Table 3: Alignment table

	Alignments	Persian sentence	Alignments
He	1	او	1
Goes	2	به	3
To	3	مدرسه	4
School	4	میرود	2

In this way, we obtained word orders for all the sentences exactly as they appeared in the source (corpus).
Afterwards, we reordered all the sentences and generated numerical sequences but this time by using the proposed model.
In order to evaluate the performance of the proposed reordering model, precision and recall measures of the reorderings that were produced by this model were calculated[38] using the following equations:

$$\text{Precision} = \frac{\text{Number of correct reorderings generated by model}}{\text{Total number of reorderings generated by model}} \quad (1)$$

$$\text{Recall} = \frac{\text{Number of correct reorderings generated by model}}{\text{Total number of reorderings in source}} \quad (2)$$

To distinguish correctness of the reorderings, we compared numerical order sequences that had been generated by this model with ones extracted from alignments.
Table-4 demonstrates the total precision and recall of the reordering task on PCTS test data.

Table 4: Total results on PCTS dataset

Model	Precision	Recall
Dependency tree model[39]	0.33	0.32
Proposed model (rule collection)	0.293	0.27
Proposed model (rule collection + parse tree method)	**0.365**	**0.32**

As can be seen in Table 4 the reordering model that was only based on the rule collection could not reach an acceptable result. However, when we combined the rule collection with the multilevel parse tree method, precision and recall measures were improved about +0.072 and +0.05 points. Furthermore, the proposed model improves precision about 0.035 points compared to the Dependency tree model.

5. Conclusion

In this paper, we presented a new reordering model for statistical machine translation based on a reordering rule collection and sentences' parse trees. The main difference between this model and its counterparts is that we developed a novel method that uses syntactic information of parse trees properly. This method has a multilevel point of view. At first, the sentence is reordered at the word level, and then at the phrase level. In this way, not only related words stayed together but also sentence structure will be preserved, and many more essential reorderings will be performed.
Experiments on the English-Persian language pair showed significant (+3.5%) improvements in precision by using the multilevel reordering method compared to the previous approaches under the same test data and conditions.
Stanford POS tagger and lexical parser are available in different languages with reliable performances. Therefore, this reordering model is easily applicable to other languages, as well.

References

[1] Philipp Koehn, Franz J. Och, and Daniel Marcu. 2003.Statistical Phrase-Based Translation. In Proceedingsof the 2003 HLT-NAACL, pages 48–54.
[2] David Chiang. 2007. Hierarchical Phrase-Based Translation. Computational Linguistics, 33(2):201–228.
[3] Michel Galley, Mark Hopkins, Kevin Knight, and Daniel Marcu. 2004. What's in a translation rule? In Daniel Marcu Susan Dumais and Salim Roukos, editors,HLT-NAACL 2004: Main Proceedings, pages 273–280, Boston, Massachusetts, USA, May 2 – May 7. Association for Computational Linguistics.

[1] Parallel Corpus Test Set

[4] Yuan Ding and Martha Palmer. 2005. Machine Translation Using Probabilistic Synchronous Dependency Insertion Grammars. In Proceedings of the 43rd ACL,pages 541–548, Ann Arbor, Michigan, June. Association for Computational Linguistics.
[5] Yang Liu, Qun Liu, and Shouxun Lin. 2006. Tree to String Alignment Template for Statistical Machine Translation. In Proceedings of the 21st ACL, pages 609–616, Sydney, Australia, July. Association for Computational Linguistics.
[6] Yang Liu, Yajuan L¨u, and Qun Liu. 2009. Improving Tree-to-Tree Translation with Packed Forests. In Proceedings of the Joint Conference of the 47th Annual Meeting of the ACL and the 4th International Joint Conference on Natural Language Processing of the AFNLP, pages 558–566, Suntec, Singapore, August. Association for Computational Linguistics.
[7] Philipp Koehn, Amittai Axelrod, Alexandra BirchMayne, 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.
[8] Christoph Tillmann. 2004. A unigram orientation model for statistical machine translation. In Proceedings of HLT-NAACL 2004: Short Papers, pages 101–104.
[9] Philipp Koehn, Hieu Hoang, Alexandra Birch, Chris Callison-Burch, Marcello Federico, Nicola Bertoldi, Brooke Cowan, Wade Shen, Christine Moran, Richard Zens, et al. 2007. Moses: Open source toolkit for statistical machine translation. In Annual meeting association for computational linguistics, volume 45, page 2.
[10] Nießen, Sonja and Hermann Ney. 2001. Morpho-syntactic analysis for reordering in statistical machine translation. In Proceedings of the MT Summit VIII: Machine Translation in the Information Age, pages 247–252, Santiago de Compostela, Spain.
[11] Khalilov, Maxim and José A. R. Fonollosa. 2011. Syntax-based reordering for statistical machine translation. Computer Speech and Language journal (Elsevier), 25, October.
[12] Khalilov, Maxim and Khalil Sima'an. 2011. Context-sensitive syntactic source-reordering by statistical transduction. In Proceedings of 5th International Joint Conference on Natural Language Processing, pages 38–46, Chiang Mai, Thailand, November. Asian Federation of Natural Language Processing.
[13] Visweswariah, Karthik, Rajakrishnan Rajkumar, Ankur Gandhe, Ananthakrishnan Ramanathan, and Jiri Navratil. 2011. A word reordering model for improved machine translation. In Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing, pages 486–496, Edinburgh, Scotland, UK., July. Association for Computational Linguistics.
[14] Gojun, Anita and Alexander Fraser. 2012. Determining the placement of German verbs in English-to-German SMT. In Proceedings of the 13th Conference of the European Chapter of the Association for Computational Linguistics, pages 726–735, Avignon, France, April. Association for Computational Linguistics.
[15] Jehl, Laura, Adrià de Gispert, Mark Hopkins, and Bill Byrne. 2014. Source-side preordering for translation using logistic regression and depth-first branch-and-bound search. In Proceedings of the 14th Conference of the European Chapter of the Association for Computational Linguistics, pages 239–248, Gothenburg, Sweden, April. Association for Computational Linguistics.
[16] Lerner, Uri and Slav Petrov. 2013. Source-side classifier preordering for machine translation. In Proceedings of the EMNLP 2013. Association for Computational Linguistics, October.
[17] Yang, Nan, Mu Li, Dongdong Zhang, and Nenghai Yu. 2012. A ranking-based approach to word reordering for statistical machine translation. In Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), pages 912–920, Jeju Island, Korea, July. Association for Computational Linguistics.
[18] Zens, Richard, Franz Josef Och, and Hermann Ney. 2002. Phrase-based statistical machine translation. In 25th German Conference on Artificial Intelligence (KI2002), pages 18–32, Aachen, Germany. Springer Verlag.
[19] Zhang, Yuqi, Richard Zens, and Hermann Ney. 2007. Chunk-level reordering of source language sentences with automatically learned rules for statistical machine translation. In Proceedings of SSST, NAACL-HLT 2007 / AMTA Workshop on Syntax and Structure in Statistical Translation, pages 1–8, Rochester, New York, April. Association for Computational Linguistics.
[20] Rottmann, Kay and Stephan Vogel. 2007. Word reordering in statistical machine translation with a pos-based distortion model. In Theoretical and Methodological Issues in Machine Translation (TMI), Skövde, Sweden.
[21] Crego, Josep M. and Nizar Habash. 2008. Using shallow syntax information to improve word alignment and reordering for SMT. In StatMT '08: Proceedings of the Third Workshop on Statistical Machine Translation, pages 53–61, Morristown, NJ, USA. Association for Computational Linguistics.
[22] Niehues, Jan and Muntsin Kolss. 2009. A POS-based model for long-range reorderings in SMT. In Proceedings of the Fourth Workshop on Statistical Machine Translation, pages 206–214, Athens, Greece, March. Association for Computational Linguistics.
[23] Elming, Jakob and Nizar Habash. 2009. Syntactic reordering for English-Arabic phrase-based machine translation. In Proceedings of the EACL 2009 Workshop on Computational Approaches to Semitic Languages, pages 69–77, Athens, Greece, March.Association for Computational Linguistics.
[24] Andreas, Jacob, Nizar Habash, and Owen Rambow. 2011. Fuzzy syntactic reordering for phrase-based statistical machine translation. In Proceedings of the Sixth Workshop on Statistical Machine Translation, pages 227–236, Edinburgh, Scotland, July. Association for Computational Linguistics.
[25] Bisazza, Arianna, Daniele Pighin, and Marcello Federico. 2012. Chunk-lattices for verb reordering in Arabic-English statistical machine translation. Machine Translation, Special Issue on MT for Arabic, 26(1-2):85–103.
[26] Hardmeier, Christian, Arianna Bisazza, and Marcello Federico. 2010. FBK at WMT 2010:Word lattices for morphological reduction and chunk-based reordering. In Proceedings of the Joint Fifth Workshop on Statistical

Machine Translation and Metrics MATR, pages 88–92, Uppsala, Sweden, July. Association for Computational Linguistics.

[27] Goto, I., Utiyama, M., and Sumita, E.," Post-Ordering by Parsing for Japanese-English Statistical Machine Translation," in Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics: Short Papers-Volume 2, pp. 311-316,(2013).

[28] Sudoh, K., Wu, X., Duh, K., Tsukada, H., and Nagata, M.," Post-Ordering in Statistical Machine Translation," in Proc. MT Summit,(2011).

[29] Sudoh, K., Wu, X., Duh, K., Tsukada, H., and Nagata, M., Syntax-Based Post-Ordering for Efficient Japanese-to-English Translation, ACM Transactions on Asian Language Information Processing (TALIP), vol. 12, p. 12, (2013)

[30] Goto, I., Utiyama, M., and Sumita, E., Post-Ordering by Parsing with ITG for Japanese-English Statistical Machine Translation, ACM Transactions on Asian Language Information Processing (TALIP), vol. 12, p. 17, (2013)

[31] Farzi, S., & Faili, H. (2015). Improving Statistical Machine Translation using Syntax-based Learning-to-Rank System. Digital Scholarship in the Humanities, fqv032.

[32] Supreme Council of Information and Communication Technology. (2013). Mizan English-Persian Parallel Corpus. Tehran, I.R. Iran. Retrieved from http://dadegan.ir/catalog/mizan.

[33] Kristina Toutanova, Dan Klein, Christopher Manning, and Yoram Singer. 2003. Feature-Rich Part-of-Speech Tagging with a Cyclic Dependency Network. In Proceedings of HLT-NAACL 2003, pp. 252-259.

[34] ابزارهای پردازش متون زبان فارسی، آزمایشگاه فناوری وب دانشگاه فردوسی مشهد، 1391. (wtlab.um.ac.ir)

[35] Dan Klein and Christopher D. Manning. 2003. Accurate Unlexicalized Parsing. Proceedings of the 41st Meeting of the Association for Computational Linguistics, pp. 423-430.

[36] A. Mansoori, H. Faili, "State-of –the-art English to Persian Statistical Machine Translation System", in Proceeding of 16th CSI International Symposiums on Artificial Intelligence and Signal Processing (AISP 2012), 2-3 May 2012, pp. 174-179, 2012.

[37] F. J. Och and H. Ney, "A Systematic Comparison of Various Statistical Alignment Models", Computational Linguistics, vol. 29, no. 1,pp. 19-51, March 2003

[38] Gao, Y., Koehn, P., and Birch, A.: Soft dependency constraints for reordering in hierarchical phrase-based translation. In: Proceedings of the Conference on Empirical Methods in Natural Language Processing, pp. 857–868 (2011)

[39] Farzi, Saeed, Hesham Faili, Shahram Khadivi, and Jalal Maleki. "A Novel Reordering Model for Statistical Machine Translation." research in computing science 1, no. 65 (2013): 51-64.