

Source side pre-ordering using recurrent neural networks for English-Myanmar machine translation

May Kyi Nyein, Khin Mar Soe

Natural Language Processing Laboratory, University of Computer Studies, Yangon, Myanmar

Article Info

Article history:

Received Aug 5, 2020

Revised Mar 13, 2021

Accepted Mar 24, 2021

Keywords:

Asian language treebank

Dependency parse trees

Machine translation

Pre-ordering model

Recurrent neural networks

Word reordering

ABSTRACT

Word reordering has remained one of the challenging problems for machine translation when translating between language pairs with different word orders e.g. English and Myanmar. Without reordering between these languages, a source sentence may be translated directly with similar word order and translation can not be meaningful. Myanmar is a subject-object-verb (SOV) language and an effective reordering is essential for translation. In this paper, we applied a pre-ordering approach using recurrent neural networks to pre-order words of the source Myanmar sentence into target English's word order. This neural pre-ordering model is automatically derived from parallel word-aligned data with syntactic and lexical features based on dependency parse trees of the source sentences. This can generate arbitrary permutations that may be non-local on the sentence and can be combined into English-Myanmar machine translation. We exploited the model to reorder English sentences into Myanmar-like word order as a pre-processing stage for machine translation, obtaining improvements quality comparable to baseline rule-based pre-ordering approach on asian language treebank (ALT) corpus.

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



Corresponding Author:

May Kyi Nyein

Natural Language Processing Laboratory

University of Computer Studies, Yangon

No.4, Main Road, Shwe Pyi Thar Township, Yangon, Myanmar

Email: maykyinyein@ucsy.edu.mm

1. INTRODUCTION

Machine translation (MT) system has broadly focused on word ordering and translation in translation output. Generating a proper word order of the target language has become one of the main problems for MT. Phrase-based statistical machine translation (PBSMT) system relies on the target language model and can produce short-distance reordering. They have still difficulties related to long-distance reordering that happen between different language pairs. For empirical reasons, all decoders for PBSMT limit the amount of reordering and are unable to produce exact translations when the required movement is over a large distance. Moreover, the different syntactic structures have challenged to generate syntactically and semantically correct word order sequences for both efficiency and translation quality. Hence, the quality of PBSMT is enhanced by reordering the words in the source sentences like the target word orders as a preprocessing phase in MT.

The pre-ordering method has been successfully employed in various language pairs between English and subject-object-verb (SOV) languages such as Japanese, Urdu, Hindi, Turkish, and Korean [1]. The major issue is that it needs a high-level quality parser that may not be obtainable for the source language and considerable linguistic expertise for generating the reordering rules [2]. This method can be categorized by

pre-ordering rules that manually hand-coded rules on linguistic knowledge, or learned automatically from data [3], [4]. The approaches have the advantages that they are independent of the actual MT system used, are also fast to apply, and tend to decrease the time spent in actual decoding. Linguistically, the syntax-based pre-ordering method requires parse trees: the constituency or dependency parse trees that more powerful in capturing the sentence structure.

In this study, we employed a pre-ordering approach using recurrent neural networks (RNN) based on rich dependency syntax features for reordering the words in the source sentence according to the target word order. This approach is a sequence prediction machine learning task and produces permutations that may be long-range on the source sentences. The main purpose of this work is to construct English-to-Myanmar pre-ordering system to develop statistical machine translation (SMT). No work has been done pre-ordering using neural networks for English-to-Myanmar language. But we can't conduct the experiment on Myanmar-to-English because parsers are unavailable for Myanmar language.

The rest of the paper is designed is being as. Some representative workings on reordering approaches are illustrated in section 2, and section 3 presents Myanmar language and reordering issues for English-Myanmar translation. Section 4 describes recurrent neural networks pre-ordering model and section 5 presents details of the experiments and section 6 explains results and discussion. Finally, in section 7, we present the conclusion of this paper.

2. LITERATURE REVIEW

Several reordering methods have been proposed to address the reordering problems. The previous reordering on Myanmar was done by applying rule-based and statistical approaches. The preceding research made use of rules with small parallel sentences. Wai *et al.* [5] described automatic reordering rules generation for English to Myanmar translation applying part-of-speech (POS) tags and function tags rule extraction algorithms. First order Markov theory was used to implement a stochastic reordering model. Win [6] presented clause level reordering (CLR) technique by applying POS rules, article checking algorithm, and a pattern matching algorithm for Myanmar to English reordering. In this study, translated target English words were reordered to make the proper order of the target sentences by using 14 test sentence patterns.

The automatic rules learning approach for English-Myanmar pre-ordering was proposed in [7]. The training method extracted the preordering rules from dependency parse trees to decrease the number of alignment crossings in the training corpus. The results showed that this approach gained improvement by 1.8 BLEU points on ALT data over the baseline PBSMT. The dependency-based head finalization pre-ordering method for English-Myanmar was studied in [8]. The dependency structures are used to get the head finalization because the approach jumps a head word after all its modifiers. In this analysis, the source English sentences were reordered before neural machine translation (NMT) and statistical machine translation (SMT) translation systems. Pre-ordering improved translation task by 0.2 BLEU in the PBSMT but it did not improve in the NMT system.

Currently, reordering methods has received lots of observation by researchers. The earlier works used a source parser and manual reordering rules. Izozaki *et al.* [9] introduced a simple rule-based pre-ordering system to reorder in an English-to-Japanese task with a parser annotating syntactical heads. The rules were created by constituency tree using Enju, a head driven phrase structure grammar parser. Some works have used automatic reordering rules learning applying syntactic features of the parse trees for machine translation. Genzel [10] and Cai *et al.* [11] proposed the approaches which learn unlexicalized reordering rules from automatic word-aligned data by minimizing the amount of crossing alignments. The rules are automatically learned with a sequence of POS tags by dependency parse trees. These can capture several types of reordering occurrences. Each rule has a vector of features for a vertex with a context of its child nodes and defines the permutation of the child nodes.

Jehl *et al.* [12] developed a pre-ordering approach using a logistic regression model that predicts whether two child nodes in the source parse tree need to swap or not. This model used a feature-rich representation and a depth-first branch-and-bound search to make reordering predictions. They looked for the best permutation of its child nodes given the pairwise scores made by the model. Kawan *et al.* [13] proposed a pre-ordering method with recursive neural network (RvNN) that is free of manual feature design and made use of information in sub-trees. The method learned whether to reorder the nodes with a vector representation of sub-trees. The method achieved better translation quality compared with the top-down BTG preordering method on English-Japanese translation.

Pre-ordering as a visiting on the dependency trees, conduct by a recurrent neural network was introduced in [14]. This method applied a neural reordering model joining with beam search for multiple ordering alternatives and compared with other phrase-based reordering models. Hadiwinoto and Ng [15]

developed a reordering method exploiting sparse features related to dependency word pairs and dependency-based embedding to predict whether the translation outputs of two source words connected by a dependency relation should be remained in a similar order or swapped in the target sentence. Gispert *et al.* [16] used a feed-forward neural networks to result in faster the execution and accuracy of pre-ordering in MT. This task was applied in various language pairs and gave acceptable results compared with the state-of-the-art techniques.

3. MYANMAR LANGUAGE

Myanmar language is complex, rich morphology, and ambiguity. The sentences are made up of one or more phrases as verb phrases, and noun phrases, or clauses. Morphologically, it is analytic language without the inflection of morphemes. Syntactically, it is usually the head-final language that the functional morphemes follow content morphemes, and the verb always becomes at the end of a sentence [17]. The sentences are delimited by a sentence boundary marker, but phrases and words are rarely delimited with spaces. There are no guidelines on how to place the spaces in the sentences. So, word segmentation is needed to detect word boundary in Myanmar text and this segmentation result affects reordering and translation performance.

3.1. Reordering issues in English-Myanmar translation

Myanmar is SOV (or) OSV word order with the post-verbal position: the verbs and its inflections follow the arguments (subject, object, and indirect object), adverbials, and postpositional phrases. There are various syntactical differences not only in word-level but also in phrase-level. English is SVO language and puts auxiliary verbs close to the main verbs. Myanmar word orders differ from English mostly in the noun phrase, prepositional phrase, and verb phrase. Myanmar uses postpositionally inflected by many grammatical features but English is prepositions. In addition, there are multiple word orders in Myanmar sentence for one English sentence. This can be illustrated with an example English sentence “He carries the baggage to the train” and all Myanmar translation are valid sentences in Table 1. Without reordering between the languages, words in English sentence are directly translated and translation can't be meaningful. So, reordering is one of the problems for English-Myanmar translation.

Table 1. Different word order translation in Myanmar

	Myanmar translation	Pre-ordered English sentences
1	သူ အိတ်ကို ရထားပေါ်သို့ သယ်ဆောင်သည်။ ။	He the baggage the train to carries.
2	သူ ရထားပေါ်သို့ အိတ်ကို သယ်ဆောင်သည်။ ။	He the train to the baggage carries.
3	အိတ်ကို ရထားပေါ်သို့ သူ သယ်ဆောင်သည်။ ။	the baggage the train to He carries.

4. PRE-ORDERING MODEL

4.1. Dependency parsing

Dependency parsing can be defined as a directed graph over the words of a sentence to show the syntactic relations between the words in Figure 1. A parse tree rooted at the main verb as the head word with the subject (nouns or pronouns) and object as direct modifiers, subordinate verbs, and their adjectives as own modifiers and so on, resulting in hierarchical head-modifier relations. The edges are labeled by the syntactic relationships between a head word and a modifier word [18]. Each sub-tree is an adjacent substring of the sentence. There are two kinds of commonly used dependency parsers: Conference on natural language learning (CoNLL) typed dependency parser and stanford typed dependency parser [19].

4.2. Reordering on a dependency parse tree

Reordering is performed using a syntax-based method which operates source side dependency parse trees and is restricted to swaps between the nodes that are in parent-child or sibling relations. Let a source sentence be a list of words $f = \{f_1, f_2, \dots, f_l\}$ annotated by a dependency parse tree. Reordering is performed by passing the dependency tree beginning at the root. The process is defined that runs from word to word across the vertices of the parse tree and probably at each stage produces the current word that each word must be produced exactly once. The final sentence of the produced words f' is the permutation of the sentence f and any permutation can be produced by a suitable path on the parse tree. This reordering framework performs a sequence of movements on the trees and carries on until all the nodes have been produced. Figure 2 shows a bilingual parallel sentence with word alignments and the equivalent pre-ordered source sentence. The source English sentence is “Three people were killed due to the storm.” and Myanmar

translation is “မုန်တိုင်းကြောင့် လူ သုံးယောက် သေဆုံးခဲ့သည် ။”. When the English words are converted according to the Myanmar word order, the pre-ordered English is “the storm due to people Three killed were.” and will become monotonic translation.

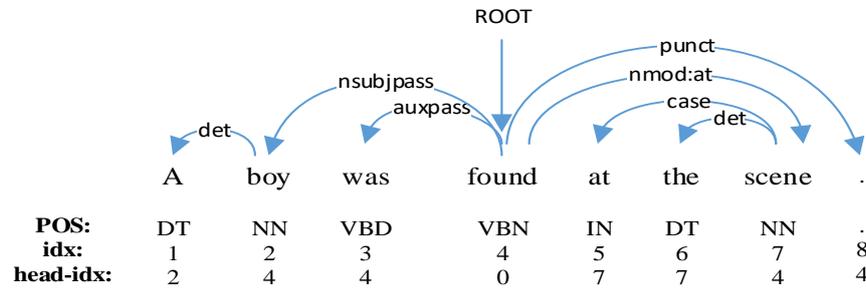


Figure 1. Dependency graph for example sentence that contains POS tags, dependency labels, word index, and head index

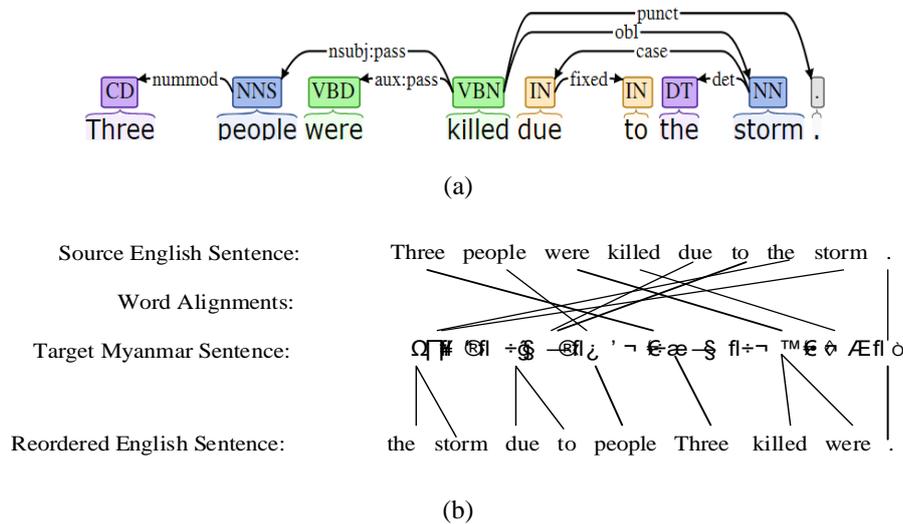


Figure 2. An example dependency parse tree and aligned sentence pair; (a) Dependency tree of an English sentence, (b) Pre-ordering for English-to-Myanmar translation

4.3. Recurrent neural networks pre-ordering model

Recurrent neural networks have been used to model long-range dependencies among words as they are not restricted to a fixed context length, like feed-forward neural networks. It allows easy incorporation of many features and can be trained on sequential data with varying lengths. This pre-ordering system is modeled using the syntax-based features and restricted to reorder between a pair of nodes that are in parent-child or siblings relations. Let $f = \{f_1, f_2, \dots, f_i\}$ be a source sentence and the state transition of the recurrent neural networks is defined as:

$$h(t) = \tanh(\theta^1 \cdot x(t) + \theta^2 \cdot h(t - 1)) \tag{1}$$

where $x(t) \in R^n$ is an input feature vector related to the t -th word at step t in a permutation f' and θ^1 and θ^2 are parameters. $\tanh(\cdot)$ is hyperbolic tangent function and input features are encoded with "one-hot" encoding. The logistic softmax function is used for normalization of scores:

$$P(I_t = j | f, i_{t-1}, t) = \frac{\exp(h_{j,t})}{\sum_{j' \in J(t)} \exp(h_{j',t})} \tag{2}$$

The probability of the total permutation f' can be computed by multiplying the probabilities of each word:

$$P(f'|f) = \prod_{t=1}^L P(I_t = i_t | f, t) \quad (3)$$

For decoding, recurrent neural network (RNN) pre-ordering model is needed to compute the most likely permutation f' of the source sentence. A local optimum problem is solved using a beam search strategy. The permutation is generated from left to right incrementally starting from an empty string and is generated all probable successor states at each step.

$$f' = \underset{f' \in \text{Gen}(f)}{\text{argmax}} P(f'|f) \quad (4)$$

Two types of feature structures are used for RNN training such as lexicalized and unlexicalized features. In the unlexicalized structure, the input feature function is made of the following. In unigram features, POS tags and dependency tags of current word, left and right child of dependency parent, and multiplication of these tags. Bigram features also contain a relationship between current word and previous released word in the permutation. All features are encoded by one-hot encoding. The features of lexicalized structure just contain the surface form of each word.

5. EXPERIMENTS

5.1. Corpus statistics

Myanmar Language is one of the low resource languages and manual aligned parallel large corpora are extremely rare. We used English-Myanmar parallel corpus from the asian language treebank (ALT) project [20] for training and testing of the pre-ordering model. The corpus consists of about 20 K parallel sentences from English wiki news and is translated into other languages, such as English, Bengali, Filipino, Hindi, Bahasa Indonesia, Japanese, Lao, Khmer, Malay, Myanmar, Thai, Vietnamese, and Chinese. Most of the sentences in the ALT corpus are very long, complex, and has some idioms. This is randomly divided into 19,098 sentences for training, 500 sentences for development and 200 sentences for testing to evaluate the reordering systems as shown in Table 2. The average English sentence length is about 15 words per sentence.

Table 2. Statistics of English-Myanmar corpus

Data Set	# Sentences	Number of Words	
		English	Myanmar
Training	19,098	441,901	511,765
Development	500	7,544	8,291
Testing	200	2,582	2,916

5.2. Training the RNN pre-ordering model

We used the pre-ordering framework of Miceli Barone [14] and tried to permute the word order of English sentences to that of the Myanmar sentences by adjusting the English word order. Manual word alignments are used from the corpus as the quality of alignment tool is still low for English-Myanmar language. But there are alignment errors and missed aligning in this manual process. So, we manually checked to get more correct alignment estimates. English sentences are parsed with Stanford dependency parser [21] in CoNLL format to obtain syntactic dependency features that represents the grammatical structure of English. The CoNLL format shows one word per line and each line represents tab-separated columns with a series of features: index, surface form, lemma, POS tag, head index, and the syntactic relation between the existing word and head word in Figure 3.

The training instances are obtained from the word-aligned corpus and the dependency tree corpus. Then, each source word is assigned to an integer index equivalent to the position of the target word that it is aligned to, joining unaligned words to the subsequent aligned word. The training problem aims to learn the parameters which minimize the cross-entropy function. We utilized the following architecture. Stochastic gradient descent is calculated using automatic differentiation of Theano that implements a backpropagation through time. L2-regularization and learning rates with the AdaDelta are used. Gradient clipping is used to avoid the exploding gradient problem. We exploited RNN pre-ordering approach with lexicalized (Lex-RNN) and unlexicalized (Unlex-RNN) features. It took about 17 hours for the Unlex-RNN training, 25 hours for the Lex-RNN training. Decoding for English corpus required about 15 hours for Unlex-RNN and 17 hours for Lex-RNN.

1	a	a	DT	2	det
2	boy	boy	NN	5	nsubjpass
3	was	be	VBD	5	auxpass
4	also	also	RB	5	advmod
5	found	found	VBN	0	root
6	at	at	IN	8	case
7	the	the	DT	8	det
8	scene	scene	NN	5	nmod
9	.	.	.	5	punct

Figure 3. Dependency features in the CoNLL format

5.3. Pre-ordering baseline

For baseline rule-based pre-ordering training, Otedama, an open-source pre-ordering tool [22], was used for comparison. Hyper-parameters are set as: the matching features is the maximum value of 10, the window size is 3, and the maximum waiting time is set to 30 minutes. The rules learning contains POS tags and dependency relations of a current node, a sequence of child nodes, and the parent of this node.

5.4. Reordering metrics

Current metrics calculate on indirect approaches for measuring the quality of the word order, and their capability to capture reordering performance has been demonstrated to be poor. To evaluate the pre-ordering models, we compared the reference permutations from manual word alignments to the hypothesis permutations generated from the models. There are different ways of measuring permutation distance. Hamming distance was presented by Ronald [23] and is known as the precise match permutation distance. It can perform to capture the number of disorders and compare the amount of disagreements between two permutations.

Kendall's tau distance [24] is the minimum number of swaps for two adjoining words to convert one permutation into another. This is named as the swap distance and can be taken as a probability of detecting discordant and concordant pairs. This metric is the most attractive for measuring word order as it is sensitive to all relative orderings. We also utilized a combined metric or LRscore [25] that was presented to evaluate reordering performance. The metric applied the distance scores in combining with BLEU to calculate the order of translation output. The combined metric connects with human assessment than BLEU only.

6. RESULTS AND DISCUSSION

The proposed pre-ordering model mainly tested for English-to-Myanmar pre-ordering. Table 3 shows the results of pre-ordering experiments in term of monolingual bilingual evaluation understudy (BLEU), Hamming distance, Kendall's tau distance, LRscore of 4-gram BLEU on Kendall's tau distance calculation (LR-KB4) and LRscore of 4-gram BLEU on Hamming distance calculation (LR-HB4) by comparing the reference reordering to the reordered output sentences for the two RNN pre-ordering and rule-based pre-ordering methods. The results proved that the two RNN pre-ordering methods are higher than the rule-based pre-ordering and then Unlex-RNN achieves improvement by 0.54 BLEU score over the rule-based method. Unexpectedly, the results showed that Unlex. RNN made worse than the Lex-RNN. This contradicts to expectation as neural models are commonly lexicalized features but often use unlexicalized features. We also observed that dependency parsing and word alignments accuracy effects on pre-ordering model training.

Table 3. Metric scores for English pre-ordering on ALT test set

Metric	BLEU	Hamming	Kendall's tau	LR-KB4	LR-HB4
Unlex-RNN	22.326	16.64	34.39	17.19	10.22
Lex-RNN	22.078	15.36	34.01	17.11	10.01
Rule-based	21.784	15.54	33.12	17.07	9.69

We also compared the results in the pre-ordered sentences by the crossing score [26]. The score computes the number of crossing links from the aligned sentence pairs. We would like this score to be zero, meaning a monotonic parallel corpus. If the corpus is more monotonic, the translation models will be better

and decoding will be faster as less distortion will be needed. It can be seen that the crossing scores percentage remain after applying the pre-ordering models in Table 4. Unlex-RNN and Otedama achieved crossing score reductions of about 30 percentage compared to the original crossing score (baseline). We also found that the neural pre-ordering method outperforms Otedama, achieving near 5 percentage points reduction. The lower point is better.

Table 4. Crossing scores for English-to-Myanmar alignments

System	# Crossing alignments	% of baseline
Baseline	6,088	100
Unlex-RNN	4056	66.62
Rule-based	4354	71.51

In the proposed RNN pre-ordering model, pre-translation reordering is carried out in the source language side. Figure 4 illustrates pre-ordering examples applying different pre-ordering approaches: Unlex-RNN pre-ordering and rule-based pre-ordering. It is found that the training data is long sentences and the sentences which have long word length may have more errors than those of short word length. Although the pre-ordering results of examples are not the same to the references exactly, their crossing scores reduced over the baseline.

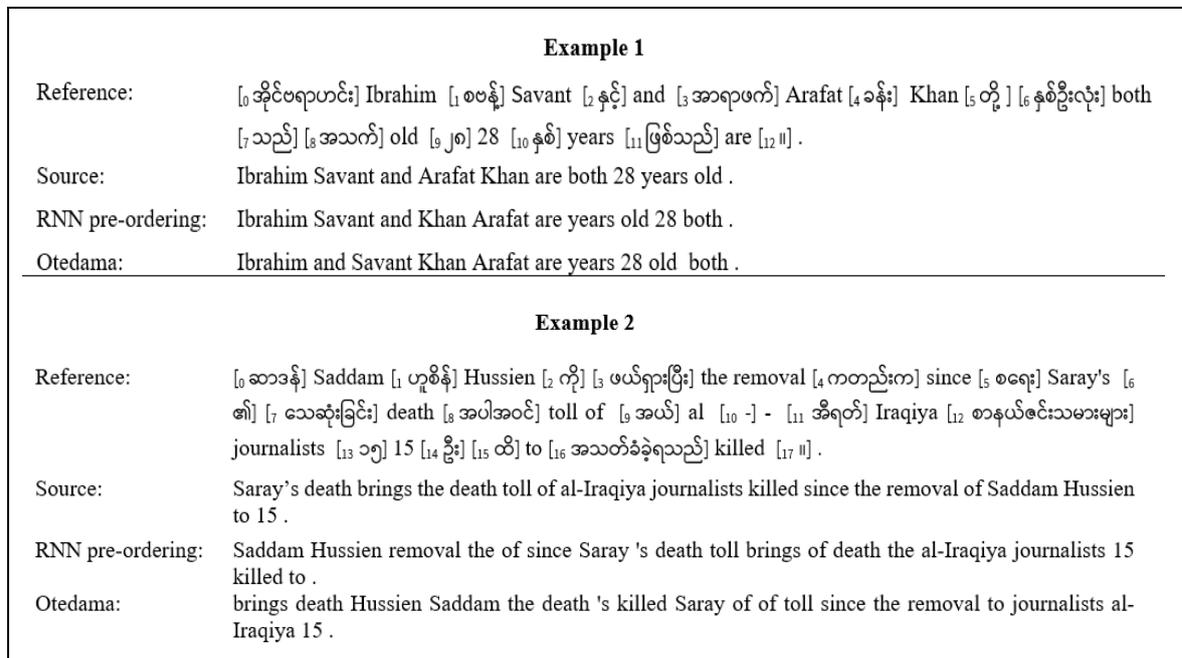


Figure 4. Examples English-to-Myanmar pre-ordering

7. CONCLUSION

To solve the reordering problem in English-Myanmar language pair, recurrent neural networks pre-ordering model is introduced. The model is trained with the machine learning approach and can generate arbitrary permutations without manual reordering rules. This studying of pre-ordering is the first work of RNN architectures on English-Myanmar pre-ordering. This approach combined the strength of lexical and syntactic reordering using the structure of the dependency parse trees and rich features. We utilized the effectiveness of neural networks and made a comparison between neural networks and rule-based on reordered source sentences. Experimental results showed that the performance of neural networks is quite promising in English-to-Myanmar pre-ordering. Currently, we primarily focus our experiments on English pre-ordering for English-Myanmar SMT. With better word alignments, better correct parsing, and more training, better reordering results can be achieved in the future as this work depends on the correct parsing and better word alignments. We will extend to incorporate the pre-ordering model to statistical machine translation.

REFERENCES

- [1] P. Xu, J. Kang, M. Ringgaard, and F. J. Och, "Using a dependency parser to improve SMT for subject-object-verb languages," *The 2009 annual conference of the North American chapter of the association for computational linguistics*, 2009, pp. 245-253.
- [2] V. H. Tran, H. T. Vu, V. V. Nguyen, and M. L. Nguyen, "A classifier-based preordering approach for English-Vietnamese statistical machine translation," In *International Conference on Intelligent Text Processing and Computational Linguistics*, Springer, Cham, 2016, pp. 74-87, doi: 10.1007/978-3-319-75487-1_7.
- [3] J. Srivastava, S. Sanyal, and A. K. Srivastava, "Extraction of reordering rules for statistical machine translation," *Journal of Intelligent and Fuzzy Systems*, vol. 36, no. 5, pp. 4809-4819, 2019, doi: 10.3233/JIFS-179029.
- [4] U. Lerner, S. Petrov, "Source-side classifier preordering for machine translation," In *Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing*, Washington, USA, 2013, pp. 513-523.
- [5] T. T. Wai, T. M. Htwe, and N. L. Thein, "Automatic Reordering Rule Generation and Application of Reordering Rules in Stochastic Reordering Model for English-Myanmar Machine Translation," *International Journal of Computer Applications*, vol. 27, no. 8, pp. 19-25, August 2011, doi: 10.1.1.259.3759&rep=rep1&type=pdf.
- [6] A. T. Win, "Clause level Reordering for Myanmar-English Translation System," *10th International Conference on Computer Applications*, UCSY, Yangon, Myanmar, Feb 2012.
- [7] M. K. Nyein and K. Mar Soe, "Exploiting Dependency-based Pre-ordering for English-Myanmar Statistical Machine Translation," *2019 23rd International Computer Science and Engineering Conference (ICSEC)*, Phuket, Thailand, 2019, pp. 192-196, doi: 10.1109/ICSEC47112.2019.8974760.
- [8] R. Wang, C. Ding, M. Utiyama, and E. Sumita, "English-Myanmar NMT and SMT with Pre-ordering: NICT's Machine Translation Systems at WAT-2018," In *32nd Pacific Asia Conference on Language, Information and Computation*, 2018, pp. 972-974.
- [9] H. Isozaki, K. Sudoh, H. Tsukada, and K. Duh, "HPSG based preprocessing for English-to-Japanese translation," *ACM Transactions on Asian Language Information Processing*, vol. 11, no. 3, pp. 1-16, 2012, doi: 10.1145/2334801.2334802.
- [10] D. Genzel, "Automatically learning source side reordering rules for large scale machine translation," In *Proceedings of the 23rd International Conference on Computational Linguistics (COLING'10)*, PA, USA, 2010, pp. 376-384.
- [11] J. Cai, Y. Zhang, H. Shan, and J. Xu, "System Description: Dependency-based Pre-ordering for Japanese-Chinese Machine Translation," *Proceedings of the 1st Workshop on Asian Translation*, Japan, pp. 39-43, October 2014.
- [12] L. Jehl, A. D. Gispert, M. Hopkins, and B. Byrne, "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*, 2014, pp. 239-248.
- [13] Y. Kawara, C. Chu, and Y. Arase, "Recursive neural network based preordering for English-to-Japanese machine translation," *arXiv preprint arXiv: 1805.10187*, May 25, 2018.
- [14] A. V. Miceli-Barone and G. Attardi, "Non-projective dependency-based pre-reordering with recurrent neural network for machine translation," In *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing*, 2015, pp. 846-856.
- [15] C. Hadiwinoto and H. T. Ng, "A dependency-based neural reordering model for statistical machine translation," In *Thirty-First AAAI Conference on Artificial Intelligence*, vol. 31, no. 1, 2017, pp. 109-115.
- [16] A. D. Gispert, G. Iglesias, and B. Byrne, "Fast and Accurate Preordering for SMT using Neural Networks," In *Proc. of the North American Chapter of the Association for Computational Linguistics*, pp. 1012-1017, 2015.
- [17] C. Ding *et al.*, "Towards Burmese (Myanmar) morphological analysis: Syllable-based tokenization and part-of-speech tagging," *ACM Transactions on Asian and Low-Resource Language Information Processing*, pp. 1-34, 2019, doi: 10.1145/3325885.
- [18] S. Kübler, R. McDonald, and J. Nivre, "Dependency parsing," *Synthesis lectures on human language technologies*, vol. 1, no. 1, pp. 1-127, 2009, doi: 10.2200/S00169ED1V01Y200901HLT002.
- [19] M. Wang, G. Xie and J. Du, "Dependency-enhanced reordering model for Chinese-English SMT," *2016 35th Chinese Control Conference (CCC)*, Chengdu, China, 2016, pp. 7094-7097, doi: 10.1109/ChiCC.2016.7554478.
- [20] H. Riza *et al.*, "Introduction of the Asian Language Treebank," *2016 Conference of The Oriental Chapter of International Committee for Coordination and Standardization of Speech Databases and Assessment Techniques (O-COCOSDA)*, Bali, Indonesia, 2016, pp. 1-6, doi: 10.1109/ICSDA.2016.7918974.
- [21] D. Manning, M. Surdeanu, J. Bauer, J. Finkel, S. J. Bethard, and D. McClosky, "The Stanford CoreNLP natural language processing toolkit," In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics*, pp. 55-60, 2014.
- [22] J. Hitschler, L. Jehl, S. Karimova, M. Ohta, B. Körner, S. Riezler, "Otedama: Fast Rule-Based Pre-Ordering for Machine Translation," *The Prague Bulletin of Mathematical Linguistics*, pp. 159-168, 2016, doi: 10.1515/pralin-2016-0015.
- [23] S. Ronald, "More distance functions for order-based encodings," *1998 IEEE International Conference on Evolutionary Computation Proceedings. IEEE World Congress on Computational Intelligence*, 1998, pp. 558-563, doi: 10.1109/ICEC.1998.700089.
- [24] M. Lapata, "Automatic evaluation of information ordering: Kendall's tau," *Computational Linguistics*, 32, no.4, 471-484, 2006, doi: 10.1162/coli.2006.32.4.471
- [25] A. Birch and M. Osborne, "LRscore for evaluating lexical and reordering quality in MT," In *Proceedings of the Joint Fifth Workshop on Statistical Machine Translation and Metrics MATR*, pp. 327-332, 2010.

- [26] N. Yang, M. Li, D. Zhang, and N. Yu, "A ranking-based approach to word reordering for statistical machine translation," In *Proceedings of ACL*, pp. 912-920, 2012.

BIOGRAPHIES OF AUTHORS



May Kyi Nyein received B.C. Sc (Hons) in 2005, followed by M.C. Sc in 2009. Currently, she is doing Ph. D research at Natural Language Processing Lab., in the University of Computer Studies, Yangon. Her research interest is Natural Language Processing (NLP) and Machine Learning.



Khin Mar Soe received a Ph.D (IT) degree from the University of Computer Studies, Yangon since 2005. She is currently working as a professor and in-charge of the Natural Language Processing lab. She supervises doctoral and master thesis in the field of NLP. She is a member of the project of ASEAN MT, the machine translation project for Southeast Asian languages.