

ATTENTION-BASED SYLLABLE LEVEL NEURAL MACHINE TRANSLATION SYSTEM FOR MYANMAR TO ENGLISH LANGUAGE PAIR

Yi Mon Shwe Sin¹ and Khin Mar Soe²

^{1,2}Natural Language Processing Lab., University of Computer Studies, Yangon, Myanmar

ABSTRACT

Neural machine translation is a new approach to machine translation that has shown the effective results for high-resource languages. Recently, the attention-based neural machine translation with the large scale parallel corpus plays an important role to achieve high performance for translation results. In this research, a parallel corpus for Myanmar-English language pair is prepared and attention-based neural machine translation models are introduced based on word to word level, character to word level, and syllable to word level. We do the experiments of the proposed model to translate the long sentences and to address morphological problems. To decrease the low resource problem, source side monolingual data are also used. So, this work investigates to improve Myanmar to English neural machine translation system. The experimental results show that syllable to word level neural machine translation model obtains an improvement over the baseline systems.

KEYWORDS

Attention-based NMT, Syllable to word level NMT, Low resource language, Myanmar language

1. INTRODUCTION

In recent year, neural machine translation (NMT) has been proposed and got the superior performance results in many language pairs. Early NMT models based on the encoder and decoder framework, it consists of one encoder RNN (Recurrent Neural Network) and one decoder RNN. A well-known flaw of early encoder-decoder NMT models was the inability to properly translate long sentences [1]. The introduction of the attention model emerged to address this problem somewhat. Attention mechanism was first introduced in NMT by [2]. Neural machine translation (NMT) with the attention-based encoder-decoder framework has achieved significant improvements in translation quality of many language pairs. And, attention-based techniques prove to be effective in modeling long-distance dependencies and complicated alignment relations in the building of translation process.

For Myanmar-English machine translation, the previous research mostly learnt by using rule-based as well as statistically based approach with the small amount of parallel corpus. Actually, the quality of translation depends on the availability of the training data. However, the small scale of Myanmar-English parallel corpus is not sufficient to build a neural translation model. Myanmar language is one of the low resource languages and there were not many resources that include parallel sentences. So, there are some difficulties in preparation of parallel corpus. Hence, we concentrated on building a huge amount of Myanmar-English parallel corpus which consist nearly 230K Myanmar- English sentences pair. And then, we apply a word to word level

attention-based neural machine translation on the Myanmar to English machine translation with this corpus.

Neural machine translation has achieved the successful performance in several language pairs. However, one of the main challenges that NMT still faces is to address the morphological problems. Like other low-resource languages, Myanmar language has a lot of morphology information. Traditionally, morphology is classified into three main types: (1) Inflectional process, (2) Derivational process and (3) Compound process. Myanmar language has a large number of inflectional features. Inflectional process is a word which is derived from another word form. It acquires the certain grammatical features but different from the same part of speech or category “e.g. walk, walks”. Many inflectional features appear on words to express agreement purposes as well as to express case, aspect, mood and tense. For example, number has four main contracts: singular, dual, trial and plural. English makes distinction only between singular and plural. In Myanmar language, number particles of noun phrases are “များတို့တွေ”. “ကလေး” is singular form and “ကလေးများ”, “ကလေးတို့”, “ကလေးတွေ” are plural form of Myanmar noun “ကလေး”. This issue leads to increase ambiguity and poses a challenge for Myanmar to English machine translation.

Like the other languages, one Myanmar sentences has one or more Myanmar Words. Unlike the other languages, Myanmar word is composed of one or more syllables. And, the syllables are composed of one or more characters. Detailed information of Myanmar language is shown in Section 3. Hence, using Myanmar characters and syllables rather than words as basic units try to solve the morphological problem. Based on the attention-based NMT, we further apply character to word level and syllable to word level attention-based neural machine translation on the Myanmar to English machine translation with the same corpus. By using source-side monolingual data for low-resource languages, it is better to boost the performance of the low-resource Myanmar to English neural machine translation model. The experiments show that Syllable to word attention-based neural machine translation gets the 26.50 BLEU and obtains an improvement of 4.22 BLEU points over the word to word level neural machine translation baseline on an open training set. The BLEU of source side monolingual data gets the 28.51 and obtains an improvement of 6.43 BLEU over the word to word level neural machine translation system.

Rests of the paper are organized as follows: section 2 describes the related work; section 3 describes about Myanmar Language; section 4 describes the model of attention-based NMT; section 5 describes the experimental setting and section 6 is the conclusion and future work.

2. RELATED WORK

In this section, previous works on Myanmar language translation are first introduced. Next, some existing works on neural machine translation are reviewed. Although the research in Myanmar-English machine translation has started since 2010, the study of automatic translation of Myanmar to English is quite few. Most of all previous research is based on rules with the use of small thousands of parallel sentences.

Myanmar phrases translation model with morphological analysis was presented and Bayes rule was used to reformulate the translation probability in [4]. Experiment results showed the improvement of translation quality by applying morphological analysis on Myanmar language. In [5], phrase-based, hierarchical phrase-based, the operational sequence model, string-to-tree, tree-to-string statistical machine translation methods were studied. It showed that the phrase-based statistical machine translation gave the highest BLEU scores than the other methods.

Recently, neural machine translation is proposed to machine translation. NMT models are a popular research field and it achieves good results in some languages. In [19], large Myanmar to English parallel corpus was created firstly and introduced neural machine translation system. The parallel corpus is general domain and consists of 230K parallel sentences mostly from spoken language available online. This work did the experiment on the word-level model and character-level model based on neural method for Myanmar to English translation. The evaluation results showed that neural machine translation models lead to improve the performance of Myanmar to English translation. This paper was the first large-scale evaluation of neural machine translation system for the Myanmar language.

In [6], character-to-character model was proposed and showed the advantages of character-level models that are better than a subword-level encoder. Most recently, [3] proposed a hybrid scheme that combines the character-level information and word level information for Czech and English language pairs. The character level model can handle unknown words and gave a comparable performance to the word-level baseline.

In [2], attention mechanism was proposed to extend the basic encoder–decoder architecture. It allowed a model to automatically (soft-) search in the predicting a target word given the source sentence. This achieved a translation performance comparable with the existing system such as statistical phrase based system. In [7], it researched on attention mechanism, sub-words translation, monolingual data and a NMT correction model. It proposed a sub-words model to address the out-of-vocabulary (OOV) problem. Both bilingual data and monolingual data were used. To enhance the translation performance, a Chinese NMT correction model was further explored. The experiments showed that attention-based neural machine translation obtained an improvement.

In this paper, attention-based neural machine translation models for Myanmar language are introduced based on word to word level, character to word level, and syllable to word level. The experiments of the proposed model do to translate the long sentences and to address morphological problems. So, this work investigates to improve Myanmar to English neural machine translation system. The experimental results show that syllable to word level neural machine translation model improves over the word to word level and character to word level.

3. MYANMAR LANGUAGE

Myanmar language is the official language and native language of Union of Myanmar. Myanmar language is recognized as one of the Tibeto-Burman groups. According to the history, the Myanmar script was originally adapted from the Mon script which was derived from Pali, the ancient India language of the text of Theravada Buddhism. Myanmar script consists of consonants, medials, vowels, numerals, punctuation marks and other special characters. Myanmar language is spoken by 32 million people as a first language and as a second language by 10 million, particularly ethnic minorities in Myanmar and those in neighbouring countries. Like some other Asian languages such as Thai and Vietnamese, Myanmar is a tonal language [9].

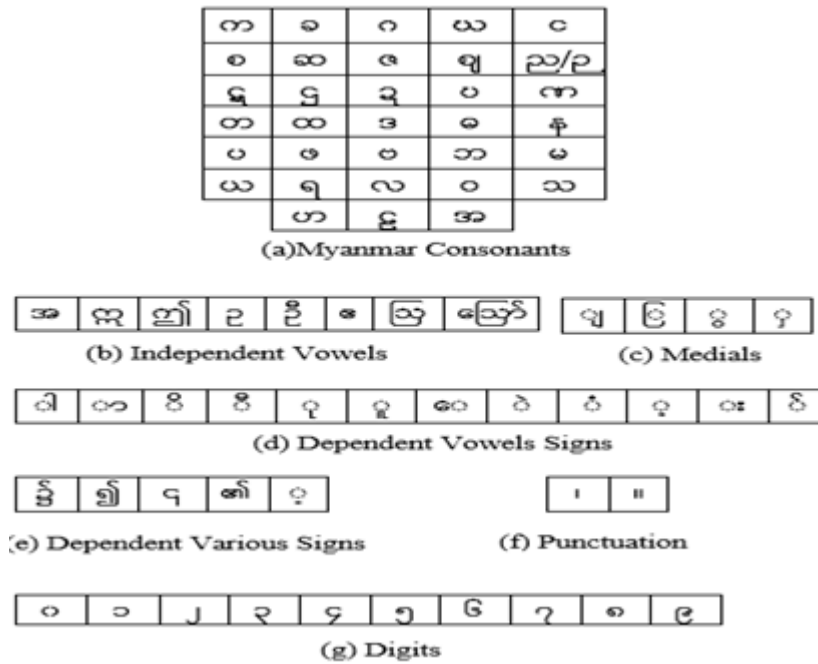


Figure 1. Myanmar character pattern

One Myanmar script is written form left to right and there are two languages styles such as spoken styles and written styles. Myanmar language is a tonal and pitch-register, largely monosyllabic and analytic language, with a Subject-Object-Verb(SOV) word order. Myanmar language characters are rounded in shape. Myanmar language has its own script and it is syllabic based. Myanmar script consists of (33) Consonants, Independent vowels, Dependent consonant signs (also known as Medials), Dependent vowels signs, Dependent various signs (also known as Parli Word), punctuation and digits. They can be seen in Figure 1.

Generally, sentence is composed of one or more phrases or clauses such as noun phrases and a verb phrase. A phrase is composed of one or more words and a word is composed of one or more syllables. Syllable is the smallest unit of the language. A syllable is generally defined as a unit of pronunciation having one vowel sound, with or without surrounding consonants which are forming all or part of word. Syllables are composed of consonants and vowels. But sometimes, a word may be only consonant without any vowel. Characters are building blocks if syllables and a Myanmar syllable is formed by joining the consonant letters with optional media, vowels and various tones. A group of words, which makes sense, but does not complete sense, is called a phrase. The formation of Myanmar sentence is shown in Figure 2.

English Sentence	The doctor is given me this prescription.												
Myanmar Sentence	ဒီဆေးညွှန်းကဆရာဝန်ငါ့ကိုပေးလိုက်တာ။												
Myanmar Phrases or clauses	Noun Phrase			Noun Phrase		Noun Phrase	Verb Phrase		Punctuation				
	ဒီဆေးညွှန်းက			ဆရာဝန်		ငါ့ကို	ပေးလိုက်တာ		။				
Myanmar Word	ဒီ	ဆေးညွှန်းက		ဆရာဝန်		ငါ့	ကို	ပေး	လိုက်	တာ	။		
Myanmar Syllables	ဒီ	ဆေး	ညွှန်း	က	ဆ	ရာ	ဝန်	ငါ့	ကို	ပေး	လိုက်	တာ	။
Myanmar Characters	ဒ ဆ ည န က ဆ ရ ဝ န ငါ ကို ပေး လိုက် တာ ။												

Figure 2. Formation of Myanmar sentence

Myanmar language has mainly eight parts of speech: noun, pronoun, verb, adjective, adverb, particle, conjunction and post-positional marker. In Myanmar script, sentences are clearly delimited by a sentence boundary marker but words are not always delimited by spaces. Although there is a general tendency to insert spaces between phrases, inserting spaces is more of a convenience rather than a rule. Spaces may sometimes be inserted between words and event between a root word and the associated post-position. However, there is no rules how to put the spaces in Myanmar language.

4. NEURAL MACHINE TRANSLATION SYSTEM

In neural machine translation system, a translation model learns the conditional probability by searching the target sentence, given a source sentence. Recently, many researchers have proposed a number of papers with the use of neural networks. Typically, this neural networks approach consists of one encoder and one decoder. The first one encodes a source sentence and the second one decodes to a target sentence.

Although the network architectures of NMT models differ in various respects, they generally follow the encoder-decoder framework and have already shown promising results. In recent NMT models, soft attention mechanism [2] has been a key extension to ensure high performance. In [2], a bidirectional recurrent neural network is used as one encoder RNN with attention mechanism and one decoder RNN. In the Encoder-Decoder architecture, an encoder reads the input sentence, $a = (a_1, \dots, a_{T_x})$, into a vector c , like that:

$$h_t = f(a_t, h_{t-1}) \tag{1}$$

where $h_t \in R^n$ is a hidden state at time t , f is a kind of nonlinear functions.

In attention-based mechanism, the annotation of each word summarizes not only the preceding words, but also the following words. Hence, bidirectional RNN (BIRNN) was used, which consists of forward RNN and backward RNN. The forward RNN \vec{f} reads the input sentence as it is ordered from a_1 to a_{T_a} and calculates a sequence of forward hidden states $(\vec{h}_1, \dots, \vec{h}_{T_a})$. The backward RNN \overleftarrow{f} reads the sequence in the reverse order from a_{T_a} to a_1 and calculates the backward hidden states $(\overleftarrow{h}_1, \dots, \overleftarrow{h}_{T_a})$. By concatenating the forward hidden state and backward hidden state, the annotation h_j is obtained.

Then, the decoder RNN computes the conditional distribution as in equation 2, the next state based on the embedding of the previously generated word, and a context vector given by the attention mechanism. Then, the decoding is performed using the attended context vector for the current time step.

$$p(b) = \prod_{t=1}^T P(b_t | \{b_1, \dots, b_{t-1}\}, c_t) = g(b_{t-1}, s_t, c_t) \quad (2)$$

where s_i is an RNN hidden state for time i , computed by

$$s_i = f(s_{i-1}, b_{i-1}, c_i).$$

Then a context vector c_i is computed as a weighted sum of these annotations h_j :

$$c_i = \sum_{j=1}^{T_x} \alpha_{ij} h_j$$

The weight of each annotation h_j is computed through an alignment model a_{ij} , which models the probability that y_i is aligned to x_j . The weight α_{ij} of each annotation h_j is computed by:

$$\alpha_{ij} = \frac{\exp(e_{ij})}{\sum_{k=1}^{T_x} \exp(e_{ik})}$$

where $e_{ij} = a(s_{i-1}, h_j)$ is an alignment model where the first i represents the output time step and the second j represents the output time step. The sample graph of attention-based syllable to word level NMT is shown in Figure 3.

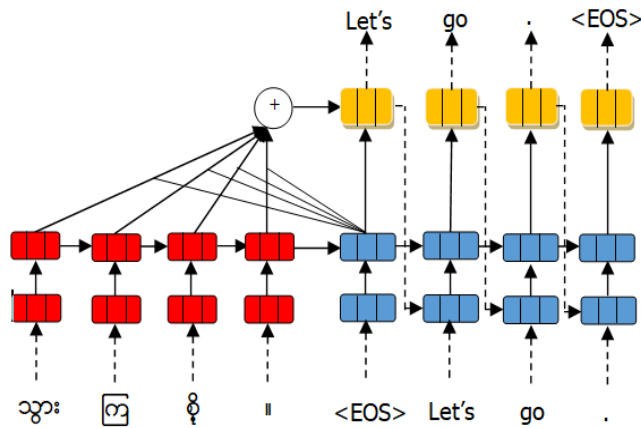


Figure 3. Attention-based syllable to word level Myanmar to English neural machine translation system

5. THE EXPERIMENTAL SETTING

We evaluate the attention-based neural machine translation approach on the task of Myanmar-to-English translation. We use the parallel corpus and monolingual corpus.

5.1. Dataset and Preprocessing Tools

Myanmar language is one of the low resources Asian languages and there were not many available parallel corpora between Myanmar and English. In our experiments, we use two parallel corpus: the ALT corpus and the UCSY corpus[20]. The ALT corpus is one part from the Asian Language Treebank(ALT) Project. It consists of twenty thousand Myanmar-English parallel sentences from Wiki news articles. The UCSY corpus is created by the NLP Lab, University of Computer Studies, Yangon(UCSY), Myanmar. It aims to promote machine translation research on Myanmar language. This corpus consists of 200K Myanmar-English parallel sentences collected from different domains, including local news articles and textbooks. The portions of the UCSY corpus and the ALT corpus are randomly divided into training data, development data and test data. Therefore, the data for Myanmar-English and English-Myanmar translation tasks is a mix domain data collected from different sources. And we also use source side monolingual corpus (150K sentences) which is local news data. Therefore, this corpus is general corpus covering different domains. Table-1 shows data statistics used for the experiments.

Table 1. Statistic of parallel sentences.

Data Type	Parallel Sentences	Training	Validation	Testing
ALT	18,965	18,525	300	140
UCSY	208,638	205,608	1,900	1,130
Total	227,603	224,133	2,200	1,270

In the preprocessing step, we use Moses's tokenization script[21] to segment the English sentences into word level. Myanmar Language is an unsegmented language and there is no clear definition of word boundaries. Therefore, proper text segmentation is essential for Myanmar Language. And I tested three models in my experiments: word to word level NMT, character to word level NMT and syllable to word level NMT. For word to word level NMT, we use UCSY_NLP lab segmenter [18] to segment the Myanmar sentence into word level. For character to word level NMT, we use Python Programming code. For Syllables to word level NMT, we use syllable segmenter[17].

5.2. Models

We build an attention-based NMT system which is a re-implementation of word to word level with Pytorch openNMT toolkit [16] as Baseline 1 and character to word level as Baseline 2. We use the same settings of this toolkit. We train Baseline 1 and Baseline 2 and apply the best BLEU scores, which are shown in Table 2.

Table 2. BLEU of the word to word level and character to word level NMT.

System(Myanmar/English)	BLEU
Attention-based Words to words level(Baseline 1)	21.88
Attention-based Characters to words level(Baseline 2)	20.71

These systems were built by using the Pytorch OpenNMT available in GitHub[16]. All weights are initialized from a uniform distribution. Each model is trained on Tesla K80 GPU. Based on different parameter settings, the training time is different. We generally used the same settings: 2-layer long short-term memory, 500 hidden units, 13 epochs, 1.0 learning rate, 64 batches size and

0.3 dropout. We used a vocabulary size of 30,000 words and 56,000 words for Myanmar language and English language respectively. For character to word level, 300 characters and 56,000 words are used. We further investigated the usage of Syllable to Word level Myanmar to English NMT with the same setting. We used the vocabulary size 6600 syllables and 56,000 words for Myanmar language and English language respectively. The BLEU of Syllable to word level NMT is shown in Table 3.

Table 3. BLEU of the Syllable to Word level NMT.

System(Myanmar/English)	BLEU
Attention-based Syllables to words level	26.50

As Table 3 shows, the syllable to word level attention-based NMT is 26.50 which are 4.62 points higher than the baseline 1. The experiments show that using syllable to word level NMT model can improve the translation performance significantly. Besides, we can see that compared with words level translation result, the structure and translation results of the syllable-words model are more flexible.

We evaluate the source-side Monolingual data based on the syllable to words level attention-based NMT model to improve the performance of Myanmar to English NMT. For the Myanmar to English machine translation, the parallel corpus is in very small scale. Hence, we directly add Myanmar monolingual data in the source side of the syllable-word level attention NMT. We apply 150K Myanmar monolingual language sentences getting from local news data to the source training set. To fill the absence in the target side, we translate the source Myanmar sentences to English sentences by using syllable to word level NMT system and add the translations to the target training set. After mixing the monolingual data and the filled data with the training set, we get a new training set of 380k parallel training corpus to train the attention-based NMT system. From the experiment result as Table 4 shows, the BLEU score increase to 28.51. The improvements of 2.01 BLEU points show that reasonable use of monolingual can enhance the low-resource Myanmar to English translation.

Table 4. BLEU of source side monolingual data.

System(Myanmar/English)	BLEU
Attention-based Syllables to words level+ Monolingual Data	28.51

It can be seen that only syllable to word level NMT is better than the word to word level NMT and character to word level NMT. The syllable to word level NMT is much better than the baseline systems at translating long sentences. In syllable to word level NMT, it knows the usage of verb and sentence reordering. It decreases the morphological problems. Therefore, syllable to word level NMT is the best for Myanmar to English neural machine translation system.

5.3. Evaluation Details

In Myanmar language, syllable is a basic unit and it has a Myanmar meaning. Therefore, the syllable-level neural machine translation is investigated and the evaluation results are the best. Generally, it improves the performance of machine translation up to 4 BLEU over baseline systems. Figure 4 gives a translation example from Myanmar-English task. Word-NMT, Char-

NMT and Syllable-NMT refer to the translation results of word to word neural machine translation, character to word neural machine translation and syllable to word neural machine translation systems respectively. It is clear that the translations from Syllable level neural machine translation are much better than the others. Compared with Word-NMT and Char-NMT, Syllable-NMT generates a better translation. Though Work-NMT keeps most meaning of the source sentence, it fails to identify the subject of the whole sentence. In the example 1,3,4,6, the translation results are correct exactly. Although the translation results of example 2 and 5 are not same to the references, their results are meaningful.

Example 1	Source Reference Char-NMT Word-NMT Syllable-NMT	ရုတ်တရက်ငါ့အချုပ်ကားတွေကိုမြင်ခဲ့တယ်။ Suddenly I saw prison cars. Suddenly I saw the movies. Suddenly I saw the cars. Suddenly I saw the prison cars.
Example 2	Source Reference Char-NMT Word-NMT Syllable-NMT	ကျောက်ကြီးမြို့နယ်မှရွာသားများသည်ပိုကောင်းမွန်သောအခြေအနေများကိုမျှော်လင့်သည်။ The villagers from Kyaukkayi township hope the better condition. The cliffs are looking forward to the better. The villagers from kyi township hope better conditions. Villagers from Kyaukkayi township expect the better condition.
Example 3	Source Reference Char-NMT Word-NMT Syllable-NMT	လယ်သမားများသည်အစိုးရဆီကဆွေးနွေးမှုများလက်ခံရရှိသည်။ Farmers receive loans from government. Farmers have to accept the government. Farmers receive loans from government. Farmers receive loans from government.
Example 4	Source Reference Char-NMT Word-NMT Syllable-NMT	သူတို့တွေထဲကအများစုဟာသူတို့မိသားစုတွေကိုထောက်ပံ့ရန်ရန်ကန်နေရသောအမျိုးသမီးတွေဖြစ်ပါတယ်။ Many of them are women struggling to support their families. Most of them are to be their families. Most of them are women who are struggling to support their families. Many of them are women struggling to support their families.
Example 5	Source Reference Char-NMT Word-NMT Syllable-NMT	ငြိမ်းချမ်းရေးဆွေးနွေးပွဲ၏ ရလဒ်သည်အမျိုးသမီးအခွင့်အရေးလှုပ်ရှားသူတွေအတွက်ကြီးမားတဲ့အောင်မြင်မှုတစ်ခုပဲဖြစ်ပါပြီ။ The outcome of the peace conference is a big success for women's rights activists. Peace is a major achievement for women. The outcome of the peace conference is a great achievement for women's rights activists. The outcome of the peace conference is a great achievement for women's rights activists.
Example 6	Source Reference Char-NMT Word-NMT Syllable-NMT	သမ္မတဦးလင်ကျော်ဟာနိုင်ငံတော်အတွင်းပင်မပုဂ္ဂိုလ်၏အောင်ဆန်းစုကြည်၏အနီးကပ်အားကိုးရသူတစ်ဦးပါ။ President U Htin Kyaw is a close confidant of State Councillor Aung San Suu Kyi. President U Htin Kyaw is a chief Indian Kyi Kyi. President U Htin Kyaw is a close researcher of State Councillor Daw Aung San Suu Kyi. President U Htin Kyaw is a close confidant of State Councillor Aung San Suu Kyi.

Figure 4. Translation results

6. CONCLUSION

In this paper, we presented attention-based neural machine translation models such as word to word level, character to word level and syllable to word level. Experimental results show the performance of Syllable-NMT is better than the other level such as words or characters for the

low resource languages like Myanmar language. We further investigated the usage of source side monolingual data in syllable to word Myanmar to English NMT to improve the performance of Myanmar to English Syllable-NMT. By incorporating the source side monolingual data into existing corpus, it is able to improve Myanmar-English Syllable NMT. In the future, existing bilingual corpus and monolingual corpus is not sufficient to train the translation model. So we will try to collect the more data and to do more work the other neural models to increase the translation performance.

REFERENCES

- [1] Kyunghyun Cho, Bart van Merriënboer, Caglar Gulcehre, Dzmitry Bahdanau, Fethi Bougares, Holger Schwenk, Yoshua Bengio, (2014) "Learning phrase representations using RNN encoder-decoder for statistical machine translation", In Proceedings of EMNLP.
- [2] Dzmitry Bahdanau, KyungHyun Cho, Yoshua Bengio, (2015) "Neural machine translation by jointly learning to align and translate", In *Proceedings of ACL – IJCNLP 2015*, Volume 1: Long Papers.
- [3] Minh-Thang Luong, Christopher D. Manning, (2016) "Achieving Open Vocabulary Neural Machine Translation with Hybrid Word-Character Models", Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics, pages 1054–1063.
- [4] Thet Thet Zin, Khin Mar Soe, Ni Lar Thein, (2011) "Myanmar Phrases Translation Model with Morphological Analysis for Statistical Myanmar to English Translation System", 25th Pacific Asia Conference on Language, Information and Computation, pages 130–139.
- [5] Win Pa Pa, Ye Kyaw Thu, Andrew Finch, Eiichiro Sumita, (2016) "A Study of Statistical Machine Translation Methods for Under Resourced Languages", 29th Pacific Asia Conference on Language, Information and Computation pages 259–269.
- [6] Jason Lee, Kyunghyun Cho, Thomas Hofmann, (2017) "Fully Character-Level Neural Machine Translation without Explicit Segmentation", Transactions of the Association for Computational Linguistics, vol. 5, pp. 365–378.
- [7] Jing Wu, Hongxu Hou, Zhipeng Shen, Jian Du, Jinting Li, (2011) "Adapting Attention-based Neural Network to Low-resource Mongolian-Chinese Machine Translation", Springer-Verlag Berlin Heidelberg.
- [8] Minh-Thang Luong, Hieu Pham, Christopher D. Manning, (2015) "Effective Approaches to Attention-based Neural Machine Translation", Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing, pages 1412–1421.
- [9] Myanmar Language Committee, (2005) "Myanmar Grammar", Myanmar Language Committee, Ministry of Education, Myanmar.
- [10] Junyoung Chung, Kyunghyun Cho, Yoshua Bengio, (2016) "A character-level decoder without explicit segmentation for neural machine translation", In Proceedings of ACL.
- [11] Patrik Lambert, Holger Schwenk, Christopher Servan, Sadaf Abdul-Rauf, (2011) "Investigations on translation model adaptation using monolingual data", In Proceedings of the Sixth Workshop on Statistical Machine Translation, pages 284–293, Edinburgh, Scotland.
- [12] Guillaume Klein, Yoon Kim, Yuntian Deng, Jean Senellart, Alexander M. Rush, (2017) "OpenNMT: Open-Source Toolkit for Neural Machine Translation", Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics, pages 67–72, Vancouver, Canada, July 30-August 4, 2017.
- [13] Khin Thandar Nwet, Khin Mar Soe, (2016) "Myanmar-English Machine Translation Model", International Conference on Genetic and Evolutionary Computing (ICGEC): Genetic and Evolutionary Computing, pp 195-203.

- [14] Marta R. Costa-Jussà, Jose' A.R. Fonollosa, (2016) "Character-based Neural Machine Translation", Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics, pages 357–361.
- [15] Rico Sennrich, Barry Haddow, Alexandra Birch, A, (2016) "Improving neural machine translation models with monolingual data", In Proceedings of the Annual Meeting of the Association for Computational Linguistics, pages 86–96.
- [16] Pytorch-OpenNMT, <http://github.com/OpenNMT/OpenNMT-py>.
- [17] Ye Kyaw Thu, (2017) Syllable segmentation tool for Myanmar language (Myanmar), <https://github.com/ye-kyaw-thu/sylbreak>.
- [18] UCSY_NLP lab segmenter, http://www.nlpresearch-ucsy.edu.mm/NLP_UCSY/wsandpos.html.
- [19] Yi Mon Shwe Sin, Khin Mar Soe, (2018) "Large Scale Myanmar to English Neural Machine Translation System", Proceeding of the IEEE 7th Global Conference on Consumer Electronic (GCCE 2018).
- [20] <http://lotus.kuee.kyoto-u.ac.jp/WAT/my-en-data>
- [21] <https://github.com/moses-smt/mosesdecoder/blob/master/scripts/tokenizer/tokenizer.perl>

Authors

Yi Mon Shwe Sin received M.C.Sc in Computer Science, from University of Computer Studies, Yangon (UCSY), Myanmar. She is currently pursuing Ph.D in University of Computer Studies, Yangon. Her research interests are Natural Language Processing and Machine Translation.



Dr. Khin Mar Soe got Ph.D(IT) in 2005. Currently, she is working as a Professor and also the Head of Natural Language processing Lab, at the University of Computer Studies, Yangon. She has been supervising Master thesis and Ph.D researches on Natural Language Processing. Moreover, she participated in the project of ASEAN MT, the machine translation project for South East Asian languages.

