Reordering Model with Recurrent Neural Networks for Myanmar-English Statistical Machine Translation

May Kyi Nyein, Khin Mar Soe
Natural Language Processing Lab, University of Computer Studies, Yangon, Myanmar
maykyinyein@ucsy.edu.mm, khinmarsoe@ucsy.edu.mm

Abstract

Word reordering is a problematic issue for language pairs with significantly different word orders, such as the translation between a subject-verb-object (SVO) language and a subject-object-verb (SOV) language. When translating between language pairs with high disparity in word order, reordering is extremely desirable for translation accuracy. In this paper, the future research directions of reordering models for Myanmar-English statistical machine translation (SMT) are also depicted. In this reordering model, the word order on source-side is arranged into the target side word order, before SMT system is applied. We propose the use of recurrent neural networks (RNNs) to model preordering for SMT.

Keywords: *Reordering, English-Myanmar statistical machine translation, recurrent neural networks.*

1. Introduction

Machine translation systems need to reorder words in the source sentence to produce fluent output in the target language that preserve the meaning of the source sentence. The system has to decide in which order to translate the given source words. Therefore, many reordering approaches have been proposed to solve word order differences for SMT systems. Most of them can solve the short-distance reordering, but long-distance reordering still remains a major challenging task in current study.

Word reordering is a preprocessing step and it makes the translation process easier. But, it is a main problem for language pairs with significantly different word orders, such as SVO-languages (English, French and Chinese) and SOV-languages (Japanese, Korean and Myanmar).

In recent years, neural network models have become increasingly popular in NLP. Initially, these models were primarily used to create n-gram neural network language models (NNLMs) for speech recognition and machine translation (Bengio et al., 2003; Schwenk, 2010). They have since been extended to translation modeling, parsing, and many other NLP tasks. The proposed reordering model will be trained using RNNs model as sequence prediction machine learning tasks.

The paper is organized as follows. Section 2 introduces related work on applying neural network to SMT and machine translation is introduced in Section 3. Afterwards, section 4 is the syntactic differences between Myanmar and English languages, reordering model framework is section 5 and the recurrent neural networks architecture is explained in section 6. Finally, section 7 concludes the paper.

2. Related Work

Myanmar-English machine translation is challenging because the grammatical forms of the two languages are totally dissimilar. For instance, English is a head-initial language, and utilizes subject-verb-object (SVO) word orders, while Myanmar is a pure head-final language, and utilizes subject-object-verb (SOV). There has been a lot of works on trying to improve the reordering model for machine translation system.

T. T. Wai et al. proposed automatic reordering rules generation for Myanmar-English machine translation. They firstly created parallel tagged aligned corpus. In addition, function tag and part-of-speech tag reordering rule extraction algorithms were proposed to generate reordering rules automatically. Among the possible reordering rules, optimal reordering rules are chosen according to the maximum probabilities [8].

Recently, neural network models have been applied into machine translation. Bengio et al. (2003) firstly proposed Feed-forward neural language model, which was a breakthrough in language modeling. Mikolov et al. (2011) proposed to use recurrent neural networks in language modeling, which can include an unbounded

sentence history and also include an unbounded future source words for predicting next word [2].

Li et al. (2014) proposed a neural reordering model that conditions reordering probabilities on the words of both the current and previous phrase pairs for phrase-based translation. Including the words of previous phrase pairs significantly improve context sensitivity and reduce reordering ambiguity to make the same classification decisions for orientation type [1].

Cui et al. (2016) presented a novel neural reordering feature by including much longer range context dependencies for predicting orientations. They utilized a long short-term memory recurrent neural network (LSTM-RNN) (1997), and directly models word pairs to predict its most probable orientation [2].

Barone et al. (2015) proposed a class of recurrent neural models which exploit source side dependency syntax features to reorder the words into a target-like order. They formulate a top-down and left-to-right walk through the dependency tree and make reordering decisions at any node. This model processes with a recurrent neural network that includes past decisions in the conditioning context. They performed German-to-English pre-reordering experiments with based RNN-RM, Fragmented RNN [4].

Recent studies on reordering model suggest that conditioning previous phrase pairs can improve context sensitivity and reduce reordering ambiguity.

3. Machine Translation

The goal of machine translation is to translate an input word sequence in the source language into a target language word sequence. In order to improve the translation process, it is possible to perform preprocessing steps before translation. In machine translation, reordering is one of the major problems, when different languages have different word order requirements. Correct word order is important for the overall quality of the translations as well as the fluency of output.

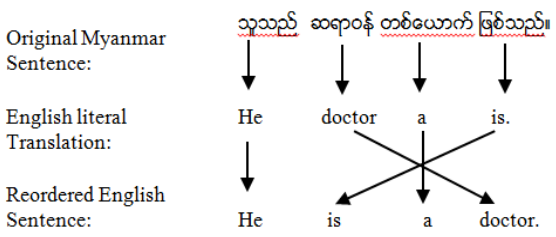

Figure 1. Example Myanmar to English Translation

In figure (1), when we translate word to word translation from Myanmar to English, we cannot get correct target sentence structure. So, we have to reorder the word's position to get the correct order for target translation.

4. Syntactic Differences between Myanmar and English Languages

The followings are the differences between Myanmar and English languages:

- English is a highly positional language with morphology, and default sentence structure as SVO.
- Myanmar is highly inflectional with a rich morphology, agglutinative language and default sentence structure as SOV.
- English uses prepositions while Myanmar is postpositionally inflected with various grammatical features.
- Although English language has restrictive word order, Myanmar language allows word order free.

There are also many syntactic differences not only in phrase level but also in word level. The basic syntactic structure in Myanmar is a head final language: the verb and its inflections follow all arguments (subject, object, indirect object) and all adjuncts (adverbials and postpositional phrases). The basic syntactic structure for English sentence is head initial language. Besides, other complements such as indirect object, place, time and adverb can be added to form a meaningful English sentence. When translating from Myanmar to English, the verb must be moved from the ending in Myanmar sentence to the beginning of the English sentence after the subject in English sentence.

Although there are many forms in declarative sentence that identified the verb as present tense, past tense, and future tense in English, Myanmar has no tense marking. Basically, there are two modal endings such as "the (သည်)" in past and present tense and "me (မယ်)" in future tense. In addition, Myanmar is free word order language and there are many possible Myanmar sentences for one English sentence as shown in table (1).

Table 1. Some Differences of Declarative Sentence between Myanmar and English

Myanmar Syntactic Structure	English Syntactic Structure
သူသည် စာအုပ်တစ်အုပ်ကို ဝယ်ခဲ့သည်။ [Subj] [Obj] [Verb]	He bought a book. [Subj] [Verb] [Obj]
(1)သူမ အခု အိမ်ကို သွားမယ်။ [Subj] [Time] [Place] [Verb] (2) အခု သူမ အိမ်ကို သွားမယ်။ [Time][Subj][Place] [Verb]	She will go home now. [Subj] [Verb] [Place] [Time]

In Myanmar, WH-questions are marked with a sentence-final "le (လဲ)"" and a yes-no questions by sentence-final "lar (လား)". In English, interrogatives are put at the beginning of the sentences as shown in in table (2).

Table 2. Some Differences of Interrogative Sentence between Myanmar and English

Myanmar Syntactic Structure	English Syntactic Structure
မင်း ဘာကို ဆိုလိုတာ လဲ။ [Subj][WH] [M-verb][SF:le]	What do you mean? [WH][Aux-V] [Subj][M-verb]
မင်းမှာ အစီအစဉ်တစ်ခုခု ရှိ လား။ [Subj] [Obj] [M-verb] [SF:la]	Do you have any plans? [Aux-V] [Subj] [M-Verb] [Obj]

In the imperative English sentences, these sentences begin verb that is followed by the respective object. In Myanmar, the order of verb and object are inversed. This syntactical difference is shown in table (3).

Table 3. Some Differences of Imperative Sentence between Myanmar and English

Myanmar Syntactic Structure	English Syntactic Structure
တံခါး ပိတ်ပါ။ [Obj] [Verb]	Lock the door. [Verb] [Obj]
ကျွန်တော့်ကို သော့ ပေးပါ။ [I-Obj] [Obj] [Verb]	Give me the key. [Verb] [I-Obj] [Obj]

In addition, there are also syntactic differences in word level for Myanmar-English reordering.

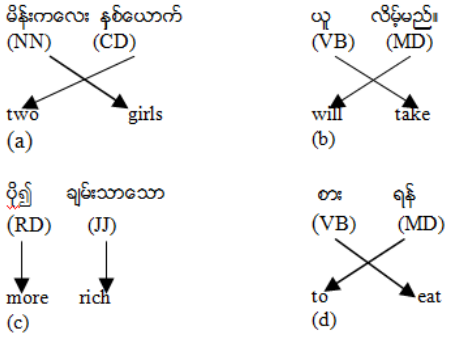

Figure 2. Some of Word Level Differences between Myanmar and English

Figure (2) shows some of the word movements in each chunk type between different language pairs. They are (a)noun chunk "မိန်းကလေး နှစ်ယောက်", (b) verb chunk "ယူလိမ့်မည်။", (c) adjective chunk "ပို၍ ချမ်းသာသော" and (d) infinitive chunk "စားရန်". Among these chunks, noun chunk, verb chunk and infinitive chunk have word order differences. In order to get correct translation from Myanmar sentences to the respective English sentences, we need to reorder these chunks. On the other hand, adjective chunk consists of adjective and adverb that modifies these adjective and appear before them. This adjective chunk in Myanmar has the same order of English.

Furthermore, a collocation of the negation word in English "can't do" is translated into Myanmar "မလုပ်နိုင်ဘူး" by negation prefix "မ" before the verb and a negation suffix "ဘူး" must succeed the verb. Finally, the prefix and suffix surround a verb to form a negation.

5. RNNs-based Reordering Model

Figure (3) illustrates the architecture of reordering model framework. Firstly, we will create Myanmar-English parallel tagged aligned corpus to generate reordering automatically in preprocessing steps.

In this system, input Myanmar sentences have no space and are written from left to right continuously. So, "Myanmar Word Segmenter" [10] From UCSY-NLP (University of Computer Studies, Yangon) lab is used for Myanmar word segmentation. After obtaining the segmented words, each segmented word in Myanmar sentence is tagged with its respective Part-of-Speech tags by

using bi-gram POS tagger [11]. Myanmar is postpositionally inflected with various grammatical features and English uses prepositions. Finally, word-based reordering with orientation within sentence is performed using recurrent neural networks model.

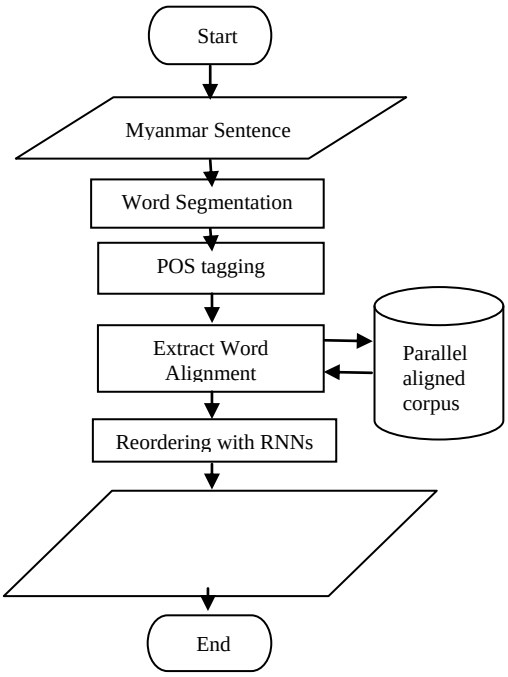

Figure 3. Myanmar to English Reordering System

6. Recurrent Neural Networks Architecture

Artificial neural networks are powerful models, which have been widely applied into many aspects of machine translation, such as language modeling (Mikolov et al., 2010) and translation modeling. In order to include more context information for determining reordering, RNNs can be used to perform better than feed-forward architectures for sequence prediction. It conditions the reordering probabilities on both the current and previous phrase pairs by including long context.

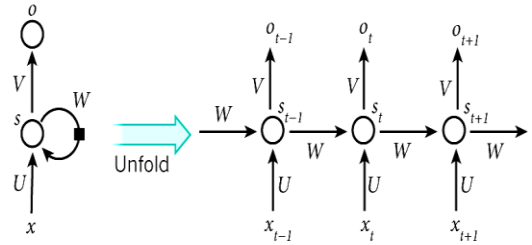

Figure 4. Recurrent Neural Networks Architecture

RNNs use unbounded history information inside the network for arbitrarily long time, and it has recurrent connections on hidden states as shown in figure (4) and corresponding equations are described in Equation (1) and (2) [5].

$$s_t = f(Ux_t + Ws_{t-1}) \quad (1)$$

$$o_t = soft\max(Vs_t) \quad (2)$$

The simple RNNs architecture consists of an input layer, a hiden layer with recurrent connections and the output layer, plus the corresponding weight matrices (U,V,W). In the input layer, the input vector represents input word encoded using 1-of-N coding (also called one-hot coding) and the output layer produces a probability distribution of a next word given previous words. The hidden layer maintains a representation of the sentence histories [6].

The source side phrases are fed into the neural network one word at a time. Only representation of the source side are finished, the probabilities for the target side are start estimated. In this way, the presentation of the next phrase of the source side starts after the prediction of the current target phrase is completed.

6.1. Training Data Processing

The 20K sentences provided by ASEAN IVO Project "Open Collaboration for Developing and using Asian Language Treebank" are used to construct Myanmar-English parallel tagged aligned corpus. The word alignment process is done by using IBM word alignment models and it produces the possible aligned words.

For reducing model complexity and easy implementation, RNNs reordering model is purely lexicalized and train on word-level. After building the parallel corpus, we utilize the word pairs and it orientation to train a neural reordering model. Given a bilingual sentence pairs and the associated word alignments, we make word-based reordering. When given source language sentence $f = \{f_1, \ldots, f_n\}$, target language sentence $e = \{e_1, \ldots e_n\}$ and word alignment $a = \{a_1, \ldots a_n\}$. RNNs reordering model can be illustrated in Equation (3) which conditions on previous and current alignment. The o_i represents the set of word orientation.

$$p(o/e,f) = \prod_{i=1}^{n} p(o_i / e_1^i, f_1^{a_i}, a_{i-1}, a_i) \quad (3)$$

Where $e_1^i = \{e_1, \ldots, e_i\}$, $f_1^{a_i} = \{f_1, \ldots, f_{a_i}\}$.

Inclusion of history word pairs is done with recurrent neural network for capability of learning history information.

6.2. Sample Input and Output

This section provides sample Myanmar sentence that was given as input to the reordering system along with the output reordered Myanmar sentence generated by the system as shown in table 4.

Table 4. Example Output of Reordering System

Step	Process	Result
1.	Input Myanmar Sentence	သူသည်နံနက်ဖြန်ရီတွင်အိပ်ရာမှ ထသည်။ (Ref. He gets up from the bed at 6 o'clock in the morning.)
2	Word Segmentation	သူ_ သည်_ နံနက်_ ၆_ နာရီ_ တွင်_ အိပ်ရာ_ မှ_ ထသည်_ ။
3.	POS tagging	သူ@PRN.Person သည်@PPM.Subject နံနက်@NN.Time ဖြန်ရီ@NN.Time တွင်@PPM.Time အိပ်ရာ@NN.Location မှ@PPM.Leave ထ@VB.Common သည်@SF.Declarative
4.	Reorder Myanmar Sentence	သူသည် ထသည် မှအိပ်ရာ တွင်ဖြန်ရီ နံနက် ။

Finally, the reordered Myanmar sentence "သူသည် ထသည် မှအိပ်ရာ တွင်ဖြန်ရီ နံနက် ။ " can be translated into "He gets up from the bed at 6 o'clock in the morning." through machine translation system. Therefore, proposed reordering system also serves as a pre-translation reordering system.

7. Conclusion

We propose word-based reordering model using recurrent neural networks, which is sensitive to change of context and introduce rich context information for reordering predictions. For reducing model complexity and easy implementation, our neural reordering model is intended purely lexicalized and trained on word-level. The ongoing research will be described for local reordering and global reordering that map various types of sentence and many English grammar patterns.

8. References

[1] P. Li, Y. Liu, M. Sun, T. Izuha and D. Zhang "A Neural Reordering Model for Phrase-based Translation", proceedings of COLING 2014, the 25th International Conference on Computational Linguistics, Dublin, Ireland, August 23-29 2014, pages 1897–1907.

[2] Y.Cui, S. Wang and Jianfeng "LSTM Neural Reordering Feature for Statistical Machine Translation", 16 Jun 2016.

[3] A.Guta, T. Alkhouli, J. Peter, J.Wuebker and H. Ney "A Comparison between Count and Neural Network Models Based on Joint Translation and Reordering Sequences", Aachen, Germany.

[4] A. V. Barone and G. Attardi "Non-projective Dependency-based Pre-Reordering with Recurrent Neural Network for Machine Translation", Proceedings of SSST-9, 9th Workshop on Syntax, Semantics and Structure in Statistical Translation, Denver, Colorado, June 4, 2015, pages 10–20.

[5] S. Liu, N. Yang, M. Li and M. Zhou "A Recursive Recurrent Neural Network for Statistical Machine Translation", Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics, Baltimore, Maryland, USA, June 23-25 2014, pages 1491–1500.

[6] M. Sundermeyer, I. Oparin,J.-L. Gauvain, B. Freiberg, R. Schliiter and H. Ney "Comparison of Feedforward and Recurrent Neural Network Language Models", 2013 IEEE.

[7] "Fast and Robust Neural Network Joint Models for Statistical Machine Translation"

[8] T.T.Wai," Morpho-syntactic reordering for English - Myanmar and Myanmar-English translation", PHD Thesis, University of Computer Studies, Yangon.

[9] M. Sundermeyer, T. Alkhouli, J.Wuebker, and H. Ney, "Translation Modeling with Bidirectional Recurrent Neural Networks", Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), Doha, Qatar, October 25-29, 2014, pages 14–25.

[10] W.P.Pa and N.L.Thein, "Myanmar Word Segmentation using Hybrid Approach" In Proc. 7th International Conference for Computer Application, Yangon, Myanmar, May 5-6, 2009.

[11] P.H.Myint, T.M.Htwe and N.L.Thein, "Bigram Part-of-Speech tagger for Myanmar Language" International Conference on Information Communication and Management, IPCSIT vol.16 (2011), IACSIT Press(2011), Singapore.

[12] A.T.Win, "Clause level Reordering for Myanmar-English Translation System", 10th International Conference ICCA, University of Computer Studies, Yangon, Myanmar, Feb 2012.