

# A Word Sense Disambiguation System Using Naïve Bayesian Algorithm for Myanmar Language

Nyein Thwet Thwet Aung, Khin Mar Soe, Ni Lar Thein

**Abstract**— Natural Language Processing has been developed to allow human-machine communication to take place in a natural-language. Word Sense Disambiguation (WSD) has always been a key problem in Natural Language Processing. WSD is defined as the task of finding the correct sense of a word in a specific context. Several methodological issues come up with the context of WSD. These are supervised and unsupervised WSD approaches. Supervised WSD approaches have obtained better results than unsupervised WSD approaches. There is not any cited work for resolving ambiguity of words in Myanmar language. Using Naïve Bayesian (NB) classifiers is known as one of the best method for supervised approaches for WSD. In this paper, we use Naïve Bayesian Classifier to disambiguate ambiguous Myanmar words with part-of-speech 'noun' and 'verb'. The system also uses Myanmar-English Parallel Corpus as training data. The WSD module developed here will be used as a complement to improve Myanmar-English machine translation system. As an advantage, the system can improve the accuracy of Myanmar to English language translation.

**Index Terms**— Myanmar-English machine translation, Myanmar-English parallel corpus, Naïve Bayes Classifier, Natural Language Processing, supervised approach, unsupervised approach, Word Sense Disambiguation.

## 1 INTRODUCTION

Word sense disambiguation (WSD) is one of the most critical and widely studied Natural Language Processing tasks, which is used in order to increase the success rates of NLP applications like machine translation, information search and information extract, natural language understanding (such as man-machine conversation system, interrogator-responder system), text auto-proofreading, speech recognition, sound-character transformation, syntax structure recognition and the language study etc [8].

WSD can be defined as the process of identifying the correct sense or meaning of a word in a particular context. When a human being is encountered with a word with multiple senses, he easily identifies the exact sense of the word with the help of context without giving a single thought to the other senses. But when the same situation is provided to a computer it is not an easy task to correctly identify the desired sense. WSD process helps in resolving such ambiguity issues [1]. Sometimes a word differs in meaning when its Part-of-Speech (POS) is different. For example butter can be a verb or a noun as it can be seen in the following example:  
Will you spread butter [Noun] on toast?  
Don't think you can butter [Verb] me up that easily.

In one sentence butter as a noun means "a solid yellow food made from milk or cream" [3], while in the other sentence butter as a verb means "to say nice things to someone so that they will do what you want" [3]. As such ambiguities can easily be resolved with the help of Part Of Speech (POS), WSD does not entertain such words. The word with different meanings having same POS needs some WSD process to conclude the accurate sense. For example, Chair in English can be "a separate seat for one person" or "the person in charge of a meeting or an organization".

Three main approaches have been applied in the WSD field. These are knowledge-based approaches, corpus based approaches and hybrid approach. Knowledge based approaches use Machine Readable Dictionaries (MRD). It relies on information provided by MRD. Corpus based approaches can be divided into two types, supervised and unsupervised learning approaches. Supervised learning approaches use information gathered from training on a corpus that has sense-tagged for semantic disambiguation. The classification approach of WSD makes use of statistical approaches either referring lexicons or using corpus for training. Thesauri, lexicons and corpus are the main source of training in the supervised approach. Unsupervised learning approaches determine the class membership of each object to be classified in a sample without using sense-tagged training examples. Hybrid approach combines aspects of fore mentioned methodologies [11].

All approaches mentioned above have been used by different researchers for different languages. Among

- Nyein Thwet Thwet Aung is currently pursuing Ph.D degree program in University of Computer Studies, Yangon, Myanmar. I got M.C.Sc in 2008. I am also an assistant lecturer. My current research interest is Natural Language Processing. PH-95-095094134. E-mail: thwet.nyein@gmail.com, nyeinthwetaung@gmail.com

them, corpus based approaches select a target word using statistic information that is automatically extracted from corpora. Corpus based method is one of the successful lines of research on WSD. In this paper, we focus on implementing WSD process for Myanmar language. We aim an application of WSD for machine translation (MT), where the system has to select the correct translation equivalent in the target language of a polysemous item in the source language. The current work is an initial step to resolve the ambiguity of words in Myanmar context. The technique that is implemented to resolve ambiguity is Bayesian Classification.

The remainder of this paper is organized as follows: We discuss the related work in section 2. Section 3 shows the ambiguity of Myanmar Language. Section 4 and 5 describe about Naïve Bayesian Classification and Naïve Bayesian Classifier for WSD. Section 6 shows the overview of the proposed system. Section 7 discusses the execution of proposed WSD Algorithm and section 8 shows the implementation of the system. The evaluation of the system is described in section 9. The paper is concluded in Section 10.

## 2 RELATED WORK

Many researchers have been work for word sense disambiguation in English Language. For the research reported in this paper, we will emphasis on the ambiguity of the Myanmar words because it is still now open in Machine Translation. In the following paragraphs, we discuss briefly some of the related work and history in the area of Word Sense Disambiguation.

Cuong Anh Le and Akira Shimazu (2004) performed to obtain High WSD accuracy using Naive Bayesian classifier with rich features [2]. Ishizaki (2006) performed a word sense disambiguation system using modified Bayesian algorithms for Indonesian language [9]. Samir El-mougy, Taher Hamza and Hatem M.Noaman (2008) discussed rooting algorithm with Naïve Bayes Classifier for Arabic Word Sense Disambiguation [10]. Farag Ahmed and Andreas Nurnberger (2008) proposed Arabic/English Word translation disambiguation using parallel corpora and matching schemes [4].

Yu Zheng-tao, Deng Bin, Hou Bo, Han Lu and Guo Jian-yi (2009) discussed word sense disambiguation based on Bayes model and information gain [13]. Asma Naseer and Sarmad Hussain (2009) proposed Supervised Word Sense Disambiguation for Urdu Using Bayesian Classification [1]. Zhang Zheng and Zhu Shu (2009) presented a new approach to Word Sense Disambiguation in MT System [14]. Laroussi Merhbene, Anis Zouaghi and Mounir Zrigui (2010) discussed Ambiguous Arabic Words Disambiguation [7]. They used context matching algorithm. The system achieved a precision of 78% and recall of 65%, using roots and signatures identifying each sense.

## 3 AMBIGUITY OF MYANMAR LANGUAGE

Myanmar language is the official language of the Union of Myanmar. It is written from left to right and no spaces between words, although informal writing often contains spaces after each clause. It is syllabic alphabet and written in circular shape. It has sentence boundary mark. It is a free-word-order language, which usually follows the subject-object-verb (SOV) order. In particular, preposition adjuncts can appear in several different places of the sentence. However, English Language has a rigid subject-verb-object (SVO) order.

Like English, Myanmar language has semantic ambiguity problem. Although using statistical methods has been very successful for some of important problems in Myanmar Natural Language Processing such as Part Of Speech tagging, segmentation and alignment of parallel translation, an effective method for solving semantic ambiguity problem does not exist yet. In table 1 and 2 show some examples of Myanmar ambiguous nouns and verbs and their senses.

TABLE 1. SOME AMBIGUOUS NOUNS AND THEIR SENSES

Ambiguous Word	No: of Sense	Sense 1	Sense 2	Sense 3	Sense 4
တု(tue)	3	Hammer	Chopsticks	Nephew	-
ဈေး(zaye)	2	Market	Price	-	-
ကျွန်း(kjun)	2	Island	Teak	-	-
နာရီ(narye)	4	O'clock	Hour	Watch	Clock
ငွေ(ngwe)	2	Silver	Money	-	-

TABLE 2. SOME AMBIGUOUS VERBS AND THEIR SENSES

Ambiguous Word	No: of Sense	Sense 1	Sense 2	Sense 3	Sense 4
ကပ်သည်(kapthe)	4	Stick	Offer	Approach	Come
ခေါက်သည်(koukthe)	4	Knock	Play	Fold	Strike
ချုပ်သည်(chokethe)	2	Sew	Stitch	-	-
စားသည်(sarthe)	3	Eat	Divide	Exceed	-
မှူးသည်(kuthe)	2	Pluck	Ladle	-	-

## 4 NAÏVE BAYESIAN FOR CLASSIFICATION

A learning algorithm or induction algorithm is the forms of concept descriptions from example data. Concept descriptions are often referred to as the knowledge or model that the learning algorithm has induced from the data. Knowledge may be represented differently from one algorithm to another. For example, C4.5 represented knowledge as a decision tree; Naïve Bayes represented knowledge in the form of probabilistic summaries Naive Bayes Classification (NBC) is a machine learning method, particularly popular in medical applications. NBC assumes that the attributes are mutually independent. Although in practice this assumption is not quite true, experience shows that the NBC approach is effective and gives rela-

tively good classification accuracy in comparison with other, more, elaborate learning methods.

### 4.1 Bayes' Theorem

Bayesian classifiers are statistical classifier. They can predict class membership probabilities, such as the probability that a given sample belongs to a particular class. Bayesian classification is based on Bayes' Theorem. Bayesian classifiers have also exhibited high accuracy and speed when applied to large databases. Naïve Bayesian classifiers assume that the effect of an attribute value on a given class is independent of the values of the other attributes. This assumption is known as "class conditional independence".

Let X be a data sample whose class label is unknown

Let H be a hypothesis that X belongs to class C

For classification problems, determine P(H/X): the probability that the hypothesis holds given the observed data sample X.

P(H): prior probability of hypothesis H (i.e. the initial probability before we observe any data, reflects the background knowledge)

P(X): probability that sample data is observed

P(X|H) : probability of observing the sample X, given that the hypothesis holds

Given training data X, posteriori probability of a hypothesis H, P(H|X) follows the Bayes theorem:

$$P(H / X) = \frac{P(X / H).P(H)}{P(X)}$$

Informally, this can be written as

Posterior = likelihood x prior/evidence

MAP (maximum posteriori) hypothesis:

$$h_{Map} = \arg \max_{h \in H} P(h / D) = \arg \max_{h \in H} P(D / h)P(h)$$

Practical difficulty: require initial knowledge of many probabilities, significant computational cost.

## 5 NAÏVE BAYESIAN CLASSIFICATION FOR WORD SENSE DISAMBIGUATION

### 5.1 Word Sense Disambiguation

A word sense is one of the meanings of a word. A word is called ambiguous if it can be interpreted in more than one way, i.e., if it has multiple senses. Disambiguation determines a specific sense of an ambiguous word. Word Sense Disambiguation (WSD) is the process of selecting the appropriate meaning or sense for a given word in a document. WSD is one of the fundamental and important processes needed for many Natural Language Processing (NLP) applications, especially for language translation. Many WSD algorithms rely on contextual similarity to help choose the proper sense of a word in context. Several important methodological issues come up in the context of word sense disambiguation. These

are: All words approach or unsupervised and Supervised or lexical sample approach.

Many Word Sense Disambiguation approaches use the following as sources:

- Dictionaries and thesauri
- Word Net
- Automatic, corpus-based; apply heuristics
- Variation or combination of above.

### 5.2 Bayesian classification

Naïve Bayes methods have been used in most classification work and were first used for WSD by Gale et al. (1992). Naïve Bayesian (NB) classifier for word sense disambiguation is that it looks at the words around an ambiguous word in a large context window. Each content word contributes potentially useful information about which sense of the ambiguous word is likely to be used with it. The supervised training of the classifier assumes that we have a corpus where each use of ambiguous words is labeled with its correct sense. These context windows can be presented in two classes:

Bag-of-word feature vectors –These are unordered set words with their exact position ignored and

Collocation feature vectors –A collocation is a word or phrase in a position specific relationship to a target word.

It is based on the assumption that all features representing the problem are conditionally independent given the value of classification variables. For a word sense disambiguation tasks, giving a word w, candidate classification variables  $S=(s_1, s_2, \dots, s_k)$  that represent the sense of the ambiguous word, and the feature  $F=(f_1, f_2, \dots, f_n)$  that describe the context in which an ambiguous word occurs, the Naïve Bayesian finds the proper sense  $s_i$  for the ambiguous word w by selecting the sense that maximizes the conditional probability  $P(w=s_i | F)$ .

Suppose C is the context of the target word w, and  $F=(f_1, f_2, \dots, f_n)$  is the set of features extracted from context C, to find the right sense  $s'$  of w given context C, we have:

$$\begin{aligned} s' &= \arg \max_{s_i} P(w = s_i | F) \\ &= \arg \max_{s_i} \frac{P(F | w = s_i) P(w = s_i)}{P(F)} \\ &= \arg \max_{s_i} P(F | w = s_i) P(w = s_i) \end{aligned}$$

The NB classifier works with the assumption that the features are conditional independent, so that we have:

$$\begin{aligned} s' &= \arg \max_{s_i} \prod_{f_j \in C} P(f_j | w = s_i) P(w = s_i) \\ &= \arg \max_{s_i} [ \sum_{f_j \in C} \log( P(f_j | w = s_i) ) + \log P(w = s_i) ] \end{aligned}$$

The features for WSD using a NB algorithm are words which are extracted from the context of the ambiguous word. The probability of sense  $s_i$ ,  $P(s_i)$ , and the conditional probability of feature  $f_j$  with observation of sense  $s_i$ ,  $P(f_j | s_i)$ , are computed via Maximum-Likelihood Estimation:

$$P(s_i) = C(s_i) / N$$

$$P(f_j | w = s_i) = C(f_j, s_i) / C(s_i)$$

Where  $C(f_j, s_i)$  is the number of occurrences of  $f_j$  in a context of sense  $s_i$  in the training corpus,  $C(s_i)$  is the number of occurrences of  $s_i$  in the training corpus, and  $N$  is the total number of occurrences of the ambiguous word  $w$  or the size of the training dataset. To avoid the effects of zero counts when estimating the conditional probabilities of the model, when meeting a new feature  $f_j$  in a context of the test dataset, for each sense  $s_i$  we set  $P(f_j | w=s_i)$  equal  $1/N$ .

### 5.3 Myanmar-English Parallel Corpus

Corpora are also called bilingual corpora, one serving as primary language, and the other working as a secondary language. A bilingual corpus was used since different senses of some words often translate differently in another language. In our experiments, we use Myanmar-English parallel corpus as training. There is no Myanmar-English parallel corpus which contains Myanmar polysemous words in public. So, we create Myanmar-English parallel corpus that contain ambiguous words manually. It contains various sense meanings of ambiguous Myanmar word. We present the following aligned sentences as part of the training corpus. The corpus structure of the following example sentences are as follows.

- (1) သူသည်တူဖြင့်ခေါက်ဆွဲစားသည်။  
 He eats the noodle with **chopsticks**.  
 သူ စားသည် ခေါက်ဆွဲ ဖြင့် တူ ။
- (2) သူ့မှာတူသုံးယောက်ရှိသည်။  
 He has three **nephews**.  
 သူ မှာရှိသည် သုံးယောက် တူ ။
- (3) လက်သမားသည်တူကိုသုံးသည်။  
 Carpenter uses the **hammer**.  
 လက်သမား သုံးသည် တူ ။

- [0]သူ/[0]he[NN]	[1]တူ/[4]chopsticks[NN]
[2]ဖြင့်/[3]with[IN]	[3]ခေါက်ဆွဲ/[2]noodle[NN]
[4]စားသည်/[1]eats[VBZ]	
- [0]သူ/[0]he[NN]	[1]တူ/[3]nephew[NN]
[2]သုံးယောက်/[2]three[CD]	[3]မှာရှိသည်/[1]has[VBZ]
- [0]လက်သမား/[0]carpenter[NN]	[1]တူ/[2]hammer[NN]
[2]သုံးသည်/[1]uses[VBZ]	

Figure 1. The Structure of Myanmar-English Parallel Corpus

Figure 1 shows the structure of Myanmar-English parallel corpus. As it is clear, the Myanmar word “တူ (tu)” are mapped into three different English words “chopsticks”, “nephew” and “hammer” based on its sense. From the corpus, we extract the possible English meanings of the target ambiguous word and collect data concerning each sense meaning.

## 6 OVERVIEW OF PROPOSED SYSTEM

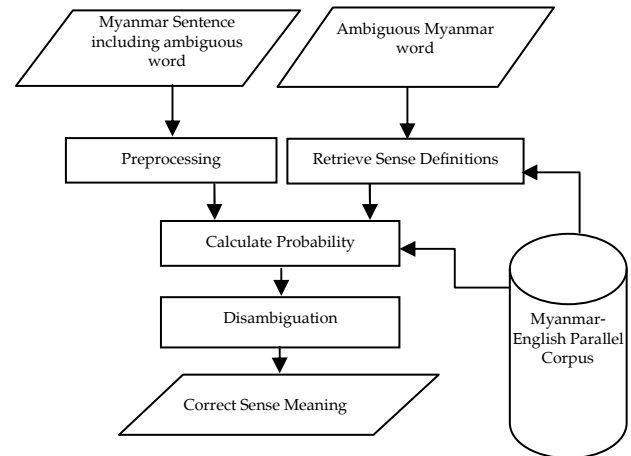


Figure 2. General Overview of proposed System

Figure 2 describes the overview of the proposed system. The proposed system consists of four main parts. These are

- 1) Preprocessing
- 2) Multi-Senses look-up on Corpus
- 3) Calculating Probability based on Bayes Theorem
- 4) Disambiguation (Calculating maximum scores using Bayes decision rule)

Firstly, the system takes the Myanmar sentence including ambiguous words and the ambiguous Myanmar word that want to disambiguate as input. In preprocessing stage, it segments the input sentence by using Myanmar word segmenter and removes all words that can be stop words which are a list of common or general terms from the input sentence. The segmenter uses maximum matching. Stop words include pronouns, prepositions, conjunctions, particles etc. After gathering information in the preprocessing step, the system uses the remaining words in the input sentence as features. Secondly, the system also retrieves the possible English sense definitions of the ambiguous word from the corpus. Thirdly, the system calculates the prior probability and the likelihood based on Bayes Theorem. Finally, disambiguation process is performed using Bayes decision rule. The disambiguation process computes the score of each sense of ambiguous word and decides the most appropriate sense for a given word in the test sentence. Our proposed algorithm is shown in the following figure 3.

- 1) Preprocessing
  - a. Segment input sentence
  - b. Remove stop words from input
- 2) Multi sense lookup
  - Lookup possible sense meanings of the ambiguous word from the corpus

```

3) Calculating Probability
  for all senses si of W do
    for all words fi in the vocabulary do
      P(fi | si) = C(fi,si)/C(si)
    end
  end
  for all senses si of W do
    P(si) = C(si) / N
  end
4) Disambiguation
  for all senses si of W do
    score(si) = log P(si)
    for all words fi in the context window c do
      score(si) = score(si) + log P(fi | si)
    end
  end
  Choose s' = arg max score(si)
    
```

Figure 3. Naive Bayes Algorithm for Myanmar WSD

**7 EXECUTION OF PROPOSED WSD ALGORITHM**

We give an example of the execution of our system and we try to disambiguate the word “ကျွန်း”(kjun) in the sentence: For example: Input sentence: “ကျွန်းသည်အလွန်အသုံးဝင်သောသစ်မာဖြစ်သည်။” (Teak is a very useful hardwood.) Input word: “ကျွန်း(kjun)”

**1) Preprocessing**

In the preprocessing, we first segment the input sentence by using Myanmar word segmenter. After segmentation: we get the following sentence.

“ကျွန်း\_သည်\_အလွန်\_အသုံးဝင်သော\_သစ်မာ\_ဖြစ်သည်\_။”  
(Teak is a very useful hardwood.)

Then, we remove all the function words (stop words). Stop words include pronouns, prepositions, conjunctions and particles. After Removing stop words: we obtain the following sentence: “ကျွန်း\_အလွန်\_အသုံးဝင်သော\_သစ်မာ\_ဖြစ်သည်”.

**2) Multi-sense loop up**

Secondly, we find the English meanings of Myanmar ambiguous word from the corpus. The word “ကျွန်း(kjun)” has two senses, teak and island. We have Bag-of-words: “အလွန်\_အသုံးဝင်သော\_သစ်မာ\_ဖြစ်သည်”.

**3) Calculating Probabilities for each sense**

Thirdly, we find the prior probabilities and the likelihood of each sense. Assume the total word count of “ကျွန်း(kjun)” in corpus is 10 (4 times for teak and 6 times for island).

$$P(\text{ကျွန်း}=\text{teak}) = 0.4, P(\text{ကျွန်း}=\text{island}) = 0.6$$

For P(F<sub>i</sub>/S=teak),

$$P(\text{အလွန်} / \text{teak}) = 0.25,$$

$$P(\text{အသုံးဝင်သော} / \text{teak}) = 0.25,$$

$$P(\text{သစ်မာ} / \text{teak}) = 0.5,$$

$$P(\text{ဖြစ်သည်} / \text{teak}) = 0.5,$$

For P(F<sub>i</sub>/S=island),

$$P(\text{အလွန်} / \text{island}) = 0.1,$$

$$P(\text{အသုံးဝင်သော} / \text{island}) = 0.1,$$

$$P(\text{သစ်မာ} / \text{island}) = 0.1,$$

$$P(\text{ဖြစ်သည်} / \text{island}) = 0.33,$$

**4) Disambiguation**

Finally, we compute the score of each sense. For P (F<sub>1</sub>,F<sub>2</sub>,...F<sub>n</sub>/S) , we multiply P(S) and P(F<sub>i</sub>/S) for each sense :

$$P(F_1,F_2,...F_n/\text{teak}) = 0.4 * 0.25 * 0.25 * 0.5 * 0.5 = 0.00625$$

$$P(F_1,F_2,...F_n/\text{island}) = 0.6 * 0.1 * 0.1 * 0.1 * 0.33 = 0.000198$$

After calculating the score of each sense, we can assign the sense with the highest probabilities to the word. So, we choose “teak” for the ambiguous word “ကျွန်း(kjun)”. By the way, we can disambiguate a word with multiple senses in a given context.

**8 IMPLEMENTAION OF THE SYSTEM**

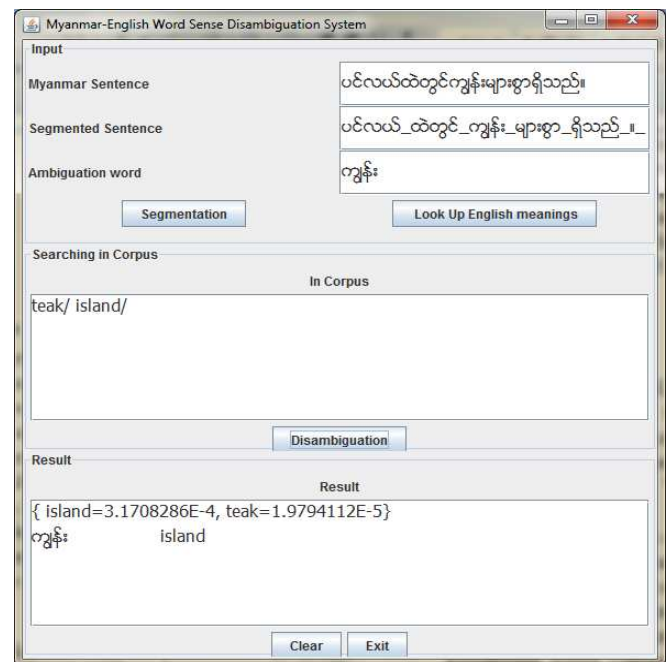


Figure 4. Implementation of the system

The implementation of our system is shown in figure 4. The system takes Myanmar sentence as input as shown in figure. The system shows segmentation result, ambiguous word, the possible English meanings of the ambiguous word and the correct English translation of the target ambiguous word in the sentence as output.

## 9 EVALUATION OF THE SYSTEM

There is not any cited work for resolving ambiguity of words in Myanmar language. Therefore, to measure the rate of disambiguation of our system, we use the most common evaluation techniques, which select a small sample of words and compare the results of the system with a human judge. We use the metric of the precision P, recall R and finally the balanced F-score which determines the weighted harmonic mean of precision and recall.

$P = \text{Correct answers provided} / \text{Answers provided}$

$R = \text{Correct answers provided} / \text{Total answers to provide}$

$F\text{-score} = 2 (P \times R) / (P + R)$

The experiments are conducted using data drawn from the training corpus, which contains sentences used in various domains. The system uses Zawgyi-one Myanmar font. Our approach relies on supervised learning. The training set consists of 1000 sentence pairs. We collected 60 ambiguous nouns and 100 ambiguous verbs for experiment. We used only the pure text data, and not the speech transcriptions. The sense of the ambiguous words was obtained from the Myanmar-English dictionary. The number of senses per test word ranged from 2 to 9 and the average was 4. Our system gets 89% precision, 92% recall and 90% f-score.

TABLE 3 PRECISION, RECALL AND F-SCORE

Precision	Recall	F-score
89%	92%	90%

## 10 CONCLUSIONS AND FUTURE EXTENSIONS

### 10.1 Conclusion

This paper shows that word sense disambiguation can be performed by using Naïve Bayesian Classifier which is based on bag-of-words feature. This approach is evaluated by using nouns and verb words in Myanmar-English parallel corpus. This research was the first attempt to create a word sense disambiguation system for Myanmar Language. The system achieves 89% precision. Therefore, the system can improve the accuracy of Myanmar to English language translation system.

### 10.2 Future Works

A number of issues have arisen in the course of this work. Addition to bag-of-words feature, the system can also use either co-occurrence feature or collocation feature, which can disambiguate the two same target words in the same sentence. This system disambiguates only words with part of speech 'Noun' and 'Verb'. We can also implement this system for words with other part of speech such as 'Adjective' and 'Adverb'. We plan to in-

vestigate the suitability of other algorithms for Myanmar word sense disambiguation such as Support Vector Machine, Decision Lists and Trees and various feature types. Our plan also is to use this work in the areas that must have word sense disambiguation algorithm before it such as machine translation, grammatical analysis, speech processing and text processing. Hence, our proposed system of disambiguation senses can be considered to be useful and applicable for other research efforts in natural language processing.

## REFERENCES

- [1] A.Naseer and S.Hussain, "Supervised Word Sense Disambiguation for Urdu Using Bayesian Classification", 2009, unpublished.
- [2] C.A. Le and A. Shimazu, "High WSD accuracy using Naïve Bayesian classifier with rich features", In *Proceedings of the PACLIC 18*, Waseda University, Tokyo, December 8th-10th, 2004.
- [3] Compaq Oxford Dictionary and Thesaurus.
- [4] F. Ahmed and A. Numberger, "Arabic/English Word Translation Disambiguation using Parallel Corpora and Matching Schemes", In *Proceedings of the 12th EAMT conference*, Hamburg, Germany, 22-23 September 2008.
- [5] F. Ahmed and A. Numberger, "Corpora based Approach for Arabic/English Word Translation Disambiguation", *Speech and Language Technology*. Volume 11, 2009.
- [6] Ide and veronis, "Word Sense Disambiguation: The State of the Art." *Computational Linguistics*, 1998.
- [7] L.Merhbene, A.Zouaghi and M.Zrigui, "Ambiguous Arabic Words Disambiguation", *11th ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing*, 2010.
- [8] N.Ide and J.Veronis, "Word Sense Disambiguation Computational Linguistics", 1998, vol.24(1), 1-42.
- [9] M.T. Uliniansyah and S. Ishizaki, "A Word Sense Disambiguation System Using Modified Naïve Bayesian Algorithms for Indonesian Language", *Information and Media Technologies* 1(1): 257-274(2006).
- [10] S. Elmougy, T. Hamza and H.M. Noaman, "Naïve Bayes Classifier for Arabic Word Sense Disambiguation", In *Proceedings of the INFOS2008*, Cairo-Egypt, March 27-29, 2008.
- [11] S. Pongpinigpinyo and W. Rivepiboon, "Distributional Semantics Approach to Thai Word Sense Disambiguation", In *Proceedings of the International Journal of Computational Intelligence* 2:3 2006.
- [12] T.M. Ma and N.L. Thein, "MASE Framework for Selecting Most Appropriate Sense of English Content Words in support of English-Myanmar Translation", In *Proceedings of the sixth international conference on Computer Applications*, 2008.
- [13] Y. Zheng-tao, D. Bin, H. Bo, H. Lu. and G. Jian-yi, "Word Sense Disambiguation Based on Bayes Model and Information Gain", In *the Proceedings of the International Journal of Advanced Science and Technology*, Vol.3, February, 2009.
- [14] Z.Zheng and Z.Shu, "A New Approach to Word Sense Disambiguation in MT System", *World Congress on Computer Science and Information Engineering*, 2009.



Nyein Thwet Thwet Aung is currently pursuing Ph.D degree program in University of Computer Studies, Yangon, Myanmar. I got M.C.Sc from University of Computer Studies, Yangon in 2008. I am also an assistant lecturer. My current research interest is Natural Language Processing.  
E-mail: [thwet.nyein@gmail.com](mailto:thwet.nyein@gmail.com)