A Comparison of Naïve Bayesian and Nearest Neighbor Cosine Classifiers for Myanmar Word Sense Disambiguation

Nyein Thwet Thwet Aung, Khin Mar Soe, Ni Lar Thein
University of Computer Studies, Yangon, Myanmar
thwet.nyein@gmail.com, nilarthein@gmail.com

Abstract

This paper presents Word Sense Disambiguation for Myanmar Language. Word Sense Disambiguation (WSD) is an intermediate but an important step in Natural Language processing. WSD is defined as the task of finding the correct sense of a word in a specific context. WSD systems can help to improve the performance of statistical machine translation (MT) system. In the most used classifiers, Nearest Neighbor Cosine (NNC) model has excellent performance, and Naïve Bayesian (NB) is preferred by researchers for it is simple and useful. In this paper, we choose NNC and NB as classifiers to disambiguate ambiguous Myanmar words with part-of-speech 'noun', 'verb' and 'adjective'. 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. We present a comparison of two methods in our experiments.

1. Introduction

Word sense disambiguation (WSD) is considered as an AI-complete problem. Word sense ambiguity can be thought of as the most serious problem in machine translation systems. A human being may also automatically consider a group of words, rather than just one word, in order to understand the meaning of a sentence, even if the words of the group are not relevant.

In order to simulate this behavior in a machine, a huge amount of data will be required as input and the output may still not be free from errors [1].

The objective of word sense disambiguation (WSD) is to identify the correct sense of a word in context. It is one of the most critical tasks in most natural language processing (NLP) applications, including information retrieval (IR), information extraction (IE) and machine translation (MT). 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?
(a solid yellow food made from milk or cream)
-Don't think you can **butter [Verb]** me up that easily. (to say nice things to someone so that they will do what you want)

As such ambiguities can easily be resolved with the help of 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,

"Take a seat on this **chair [Noun]**."
(a separate seat for one person)
"He is a **chair [Noun]** of the Mathematics Department." (the person in charge of a meeting or an organization)

For these sentences, WSD is needed to perform. WSD algorithms can be broadly classified into three categories:

Supervised Approaches: these approaches use machine-learning and data mining techniques

to train a classifier from sense-tagged corpora. The success of supervised learning approaches to word sense disambiguation is largely dependent on the features used to represent the context in which an ambiguous word occurs.

Unsupervised approaches: these approaches do not use a training corpus and are based on unlabeled corpora.

Semi-supervised approaches: A hybrid of the two other categories [2].

Supervised WSD approaches have obtained better results than unsupervised WSD approaches. In this paper, we focus on implementing WSD process for Myanmar language using supervised approach. 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 techniques that are implemented to resolve ambiguity are Bayesian Classification and Nearest Neighbor Cosine classifier. All the processes in our system are developed by Java Programming.

The remainder of this paper is organized as follows: We discuss the related work in section 2. Section 3 shows ambiguity in Myanmar Language and section 4 presents about the Myanmar-English parallel corpus. Section 5 and 6 show Naïve Bayesian Classifier and Nearest Neighbor Cosine Classifier for WSD. Overall system design is presented in Section 7. Section 8 discusses Algorithms for Myanmar WSD and section 9 shows Execution of proposed algorithms. Evaluation result and Error analysis is shown in section 10. The paper is concluded in Section 11.

2. Related Work

In this section, previous works in word sense disambiguation on different languages are reviewed. Nameh et al. (2011) presented a supervised learning method for WSD, which is based on Cosine Similarity [1]. They extract two sets of features; the set of words that have occurred frequently in the text and the set of words surrounding the ambiguous word. Naseer and Hussain (2009) discussed Supervised Word Sense Disambiguation for Urdu Using Bayesian Classification [2]. They used Bayesian classifier. Parameswarappa and Narayana (2011) proposed Kannada Word Sense Disambiguation for Machine Translation [3]. They used the compound words clue, syntactic features, argument structure and semantic information. Marianna Apidianaki (2008) described Translation-oriented word sense induction based on Parallel Corpora [4].

Heshaam Faili (2008) demonstrated an experiment of WSD in a machine translation (MT) system for two different languages, English and Persian [5].He uses decision trees automatically learned from a training data set. Marton Mihaltz and Gabor Pohl (2006) presented an experiment to automatically generate annotated training corpora for a supervised word sense disambiguation module operating in an English-Hungarian and a Hungarian-English machine translation system [6]. Their algorithm relies on monolingual and bilingual lexicons and dictionaries.

Yasaman Motazedi and Mehnoush Shamsfard (2009) proposed a new WSD method by presenting a hybrid measure to score different senses of a word [7]. They use WordNet glosses and hierarchy extended WordNet to extract WSD tags which makes the proposed work unique. Lim Lian Tze and Tang Enya Kong (2004) focus on WSD in the context of machine translation [8]. They propose a hybrid model, using the corpus (corpus-based) and a lexical ontology (knowledge-based) as their knowledge source. Ebony Domingo and Rachel Edita Roxas (2006) presented resolving target-word selection, based on "word-to-sense" and "sense-to-word" relationship between source words and their translations, using syntactic relationships [9]. Jong-Hoon Oh and Key-Sun Choi (2002) reported on word sense disambiguation of English words using static and dynamic sense vectors [10].

3. Ambiguity in 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 adjunctions can appear in several different places of the sentence.

Like English, Myanmar language has semantic ambiguity problem. For example, the Myanmar noun "ငွေ"(ngwe) would translate to two different English words (**money** for a medium of exchange in the form of coins and banknotes, or **silver** for a precious shiny grayish-white metal) in the following two sentences:

a. "သူမသည်ဘဏ်၌**ငွေ**အများအပြားစုသည်။"
(She saves a lot of **money** at bank.)   and
b. "သူသည်ပြိုင်ပွဲ၌**ငွေ**တံဆိပ်ဆုရသည်။"
(He gets **silver** medal in the competition.)

In order to translate this word to correct English word, Word Sense Disambiguation is needed to perform. Table 1, 2 and 3 show some examples of Myanmar ambiguous nouns, verbs, adjectives and its possible English meanings.

Table 1. Ambiguous nouns and their senses

Ambiguous word	No: of Sense	Sense 1	Sense 2	Sense 3	Sense 4
ဝက်ခြံ (watchan)	2	Acne	Pigsty	-	-
အခန်း (akhan)	3	Room	Chapter	Role	-
အဆောင် (asung)	2	Charm	Hostel	-	-
အဆက် (asat)	4	Joint	Continuation	Descendant	Sweet-heart
အရပ် (ayard)	2	Height	Region	-	-

Table 2. Ambiguous Verbs and their senses

Ambiguous Word	No: of Sense	Sense 1	Sense 2	Sense 3	Sense 4
နောက်သည် (noukthe)	2	(be) Muddy	Tease	-	-
ပစ်သည် (pitthe)	3	Throw	Shoot	Fire	-
မောင်းသည် (maungthe)	3	Goad	Drive	Operate	-
မြှောက်သည် (mhaukthe)	4	Raise	Toss	Flatter	Multiply
ချိန်သည် (chaethe)	2	Weigh	Aim	-	-

Table 3. Ambiguous adjectives and their senses

Ambiguous Word	No: of Sense	Sense 1	Sense 2	Sense 3	Sense 4
ပါးလွှာသော (parlwarthaw)	2	Gauzy	Sparse	-	-
ပွသော (pwathaw)	4	Puffy	Pulpy	Dishevelled	Baggy
အလွတ် (alut)	3	Blank	Free	Empty	-
အများအပြား (amyarapyar)	3	Much	Many	A Lot of	-
ကြီးသော (kyeethaw)	2	Older	Larger	-	-

4. Myanmar-English Parallel Corpus

Parallel Corpus is called bilingual corpora, one serving as primary language, and the other working as a secondary language. There is not available Myanmar-English sense tagged corpus in public. So, the corpus is created manually and which contains the various usages of Myanmar ambiguous words. In our corpus, Myanmar words are aligned with respective English words. "/" is put between these words. It contains the sentences of the newspapers, Myanmar historical books, Myanmar text books and example sentences from the Babylon dictionary. The corpus structure is shown in figure 1. The training corpus has approximately 1500 sentences. As an average, there are 10 example

instances per each ambiguous word in the corpus.

> [0]သူ/[0]he[PP]   [1]တူ/[4]**chopsticks**[NN]
> [2]ဖြင့်/[3]with[IN][3]ခေါက်ဆွဲ/[2]noodle[NN]
> [4]စားသည်/[1]eat[VB]
> [0]သူ/[0]he[PP]   [1]တူ/[5]**nephew**[NN]
> [2]သုံးယောက်/[4]three[CD]   [3]နှင့်အတူ/[3]with[IN]
> [4]ထမင်း/[2]rice[NN]   [5]စားသည်/[1]eat[VB]
> [0]တူ/[0]**hammer**[NN]   [1]သံ/[3]nail[NN]
> [2]ရိုက်သောအခါ/[2]strike[JJ]   [3]သုံးသည်/[1]use[VB]

Figure1. Sample of Myanmar-English Parallel corpus format

5. Naïve Bayesian Classification

Bayesian classification is an algorithm proposed by Gale et al. to determine a sense of a polysemous word (Gale et al. 1992) [11]. 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, s_3,…,s_n) that represent the senses of the ambiguous word, and the features F= (f_1,f_2,f_3,…,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 given F and S.

A Bayes classifier applies Bayes decision rule when choosing a class.

Decide s' if $P(s'|c) > P(s_k | c)$ for $s_k \neq s'$ (1)

Where the value of $P(s_k | c)$ is computed by the following equation:

$$P(s_k | c) = \frac{P(c | s_k)}{P(c)} P(s_k) \qquad (2)$$

$P(c)$ is a constant for all senses and therefore does not influence the value of $P(s_k | c)$. The sense s' of w is then:

$$s' = \arg\max_{s_k} P(s_k | c)$$

$$= \arg\max_{s_k} \frac{P(c | s_k)}{P(c)} P(s_k)$$

$$= \arg\max_{s_k} P(c | s_k) P(s_k)$$

$$= \arg\max_{s_k} [\log P(c | s_k) + \log P(s_k)] \qquad (3)$$

Where values of $P(c | s_k)$ and $P(s_k)$ are computed by using the following equations:

$$P(c | s_k) = \frac{C(c, s_k)}{C(s_k)} \qquad (4)$$

$$P(s_k) = \frac{C(s_k)}{C(w)} \qquad (5)$$

Where $C(c, s_k)$ is the number of occurrences of c in a context of sense in a training corpus, $C(s_k)$ is the number of appearances of s_k in a training corpus, and $C(w)$ is the number of occurrences of polysemous word w. To avoid the effects of zero counts when estimating the conditional probabilities of the model, when meeting a new feature c_j in a context of the test dataset, for each sense s_k we set $P(c_j | s_k)$ equals $\frac{1}{C(w)}$.

6. Nearest Neighbor Cosine Classification

The nearest neighbor cosine classifier is a supervised learning algorithm in which the classification is accomplished based on learning by analogy, that is, by comparing a given test vector with training vectors that are similar to it. It uses the context vectors created for each sense during training and for the ambiguous instance during testing. The cosines between the ambiguous vector and each of the context vectors are calculated, and the sense that is the "nearest" (largest cosine/smallest angle) is selected by the classifier.

In this algorithm we make use of the concept of inner product of vectors. Each of two vectors

towards each other has an angle that can be calculated using the inner product of the vectors. In this paper, after converting each context to a vector of words, we use this idea to measure the similarity between a new context and each existing context in the training corpus.

Supposing two typical vectors a = (a,b,c, …, z) and b= (A,B,C , … ,Z), the Cosine Similarity of the vectors is defined as follows:

$$\cos\theta = \frac{|a.b|}{|a|.|b|} = \frac{\sum_{i=1}^{n} a_i b_i}{\sqrt{\sum_{i=1}^{n} a_i^2}\sqrt{\sum_{i=1}^{n} b_i^2}} \quad (6)$$

Where |a| stands for the length of vector a and it is defined as:

$$|a| = \sqrt{a^2 + b^2 + c^2 + \ldots + z^2} \qquad (7)$$

And a.b is the inner product of the vectors a and b which is defined as:

a.b=(a*A,b*B,c*C,……,z*Z)   (8)
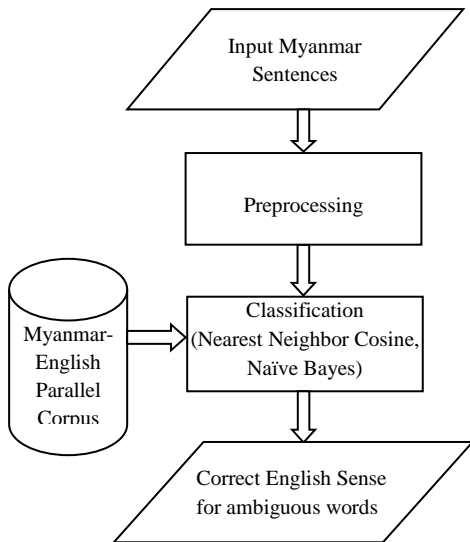
7. Overall System Description

Figure 2. Overall System Description

Figure 2 shows the overall system description. The system takes the Myanmar sentences including ambiguous words as input. In preprocessing, it segments the input sentences by using the Myanmar word segmenter. In our system, we use the existing Myanmar word segmenter for segmentation process. This segmenter is based on maximum matching scheme.

Then, the system removes all words that can be stop words which are a list of common or general terms from the input sentence. 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. The system uses topical features that represent co-occurring words in bag-of-word feature. The system also uses Myanmar-English parallel corpus as a training data. In classification, Naïve Bayes classifier (NB) and Nearest Neighbor Cosine classifier(NNC) are used. The output of the system is the correct English meanings of the ambiguous words.

8. Algorithms for Myanmar WSD

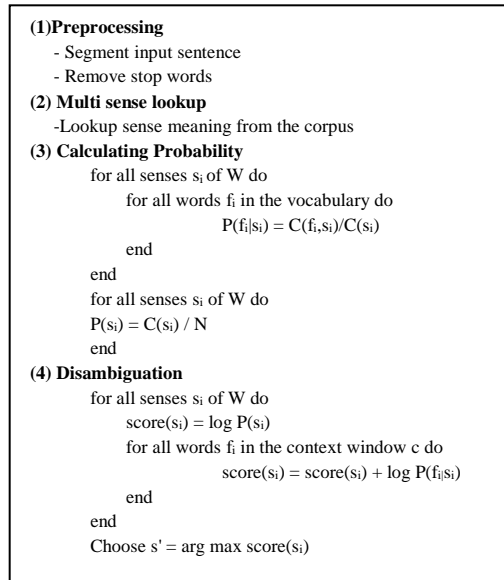We use the algorithms shown in the following figure 3 and 4 in our experiments.

(1)Preprocessing
 - Segment input sentence
 - Remove stop words
(2) Multi sense lookup
 -Lookup sense meaning from the corpus
(3) Calculating Probability
      for all senses s_i of W do
            for all words f_i in the vocabulary do
                      $P(f_i|s_i) = C(f_i,s_i)/C(s_i)$
            end
      end
      for all senses s_i of W do
      $P(s_i) = C(s_i) / N$
      end
(4) Disambiguation
      for all senses s_i of W do
          score(s_i) = log P(s_i)
          for all words f_i in the context window c do
                    score(s_i) = score(s_i) + log P($f_i|s_i$)
          end
      end
      Choose s' = arg max score(s_i)

Figure 3. Naïve Bayes(NB) algorithm for Myanmar WSD

```
(1) Preprocessing
  -Segment input sentence
  -Remove stop words from input sentence and create ambiguous
  vector
(2) Multi-sense Look-up
  -Retrieve all possible sense meanings of ambiguous word from
   the corpus
  -Collect training data concerning with these senses from
   corpus (find Nearest features)
(3)Build context vectors for each sense based on collected
   training data
  -For all context vectors do
    -Remove stop words
    -Remove redundant words (make unique features)
  -End For
(4) Calculate the cosines between ambiguous vector and each
    of the context vectors
```

$$\cos \theta = \frac{A.B}{\|A\|.\|B\|} = \frac{\sum\limits_{i=1}^{n} A_i . B_i}{\sqrt{\sum\limits_{i=1}^{n} A_i^2} \sqrt{\sum\limits_{i=1}^{n} B_i^2}}$$

```
(5) Choose correct sense of the target word.
    s' = arg max score(s_i)
```

Figure 4. Nearest Neighbor Cosine(NNC) algorithm for Myanmar WSD

9.Execution of Proposed Algorithms

Input sentence:

"ကျွန်းသည်အလွန်အသုံးဝင်သောသစ်မာဖြစ်သည်။"

(Teak is a very useful hardwood.)

Ambiguous word: " ကျွန်း(kjun)"

After segmentation:

"ကျွန်း_သည်_အလွန်_အသုံးဝင်သော_သစ်မာ_ဖြစ်သည်_။_".

After removing stop words:

"ကျွန်း,အလွန်,အသုံးဝင်သော,သစ်မာ,ဖြစ်သည်" .

Bag-of-Words:"အလွန်,အသုံးဝင်သော,သစ်မာ,ဖြစ်သည်".

We find the English meanings of Myanmar ambiguous word from the corpus. The word "ကျွန်း(kjun)" has two senses, **teak** and **island**.

By using Naïve Bayes algorithm:

We find prior probabilities and 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(ကျွန်း=teak) = 0.4, P(ကျွန်း=island) = 0.6

For P(Fi/S=teak),

   P(အလွန်/teak) =0.25,

   P(အသုံးဝင်သော/teak) =0.25,

P(သစ်မာ/teak) =0.5,

P(ဖြစ်သည်/teak) =0.5,

For P(Fi/S=island),

   P(အလွန်/island) =0.1,

   P(အသုံးဝင်သော/island) =0.1,

   P(သစ်မာ/island) =0.1,

   P(ဖြစ်သည်/island) =0.33,

Finally, we compute the score of each sense. For P(F1,F2,…Fn/S) , we multiply P(S) and P(Fi/S) for each sense :

P(F1,F2,…Fn/teak) = 0.4* 0.25 *0.25* 0.5 * 0.5 = **0.00625**

P(F1,F2,…Fn/island) = 0.6* 0.1 *0.1* 0.1 * 0.33 = 0.000198

Therefore, the correct answer of the word "ကျွန်း(kjun)" is **teak** for the given sentence.

By using Nearest Neighbor Cosine algorithm:

Build Training Vector:

teak=[တံတား, အဖိုးတန်သည်, ဖြစ်သည်, တည်ဆောက်ထားသည်, အိမ်, နေထိုင်သည်, မာသည်, အ လွန်, မြ မြန်မာ့, ပ ရိဘောဂ, သစ်မာ, အဖိုးတန်သော, အသံ အသုံးဝင်သော, ကမ္ဘာပေါ်],

island=[ပတ်လည်, ရေ, ဘေးတွင်, ပေါ် ပေါ်တွင်, ထဲတွင်, ထ င်ရှားသည်, အိမ်, မြန်မာနိုင်ငံ, ရှိသ ရှိသည်, ပင်လယ်, မြန်မာနိုင်ငံ, အများအပြား, ကိုကိုး]

Create Input Feature Vector:

[အလွန်,အသုံးဝင်သော,သစ်မာ,ဖြစ်သည်]

We compute the similarities between the input feature vector and the training vector by using cosine similarity.After calculating the score of each sense, we can assign the sense with the highest similarity to the word.
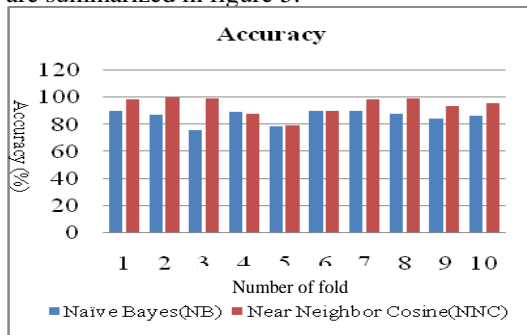
10. Evaluation and Error analysis

The experiments are conducted using data drawn from "Myanmar-English Parallel Corpus", which contains sentences used in various domains. Our approach relies on supervised learning. The system uses Zawgyi-One Myanmar font. For analysis, we found 150 ambiguous words so far. The sense of the ambiguous words was obtained from the Myanmar-English dictionary. The number of senses per test word ranged from 2 to 11. As a highly used methodology in machine learning and data

mining, we used 10 fold cross validation to estimate the performance of the algorithms. Thus, for each ambiguous word, the set of all related samples were divided into ten equal folds. Nine folds were used to extract the features and to train inner product classifier, while the remaining folds were used as test data. The above procedure is repeated 10 times so that each fold is used as the test data once. The average accuracy of the proposed method across the 10 fold cross validation is reported in table 4.

Table 4. 10 fold cross validation Accuracy

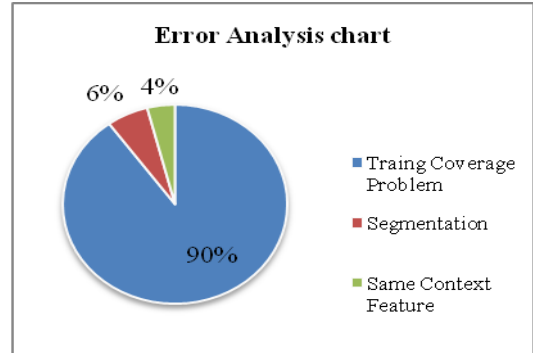| No : of fold	No: of Ambiguous word	Accuracy (%)	
		Naïve Bayes (NB)	Nearest Neighbor Cosine(NNC)
1	121	90.4	98.7
2	122	87.4	100
3	118	75.9	99.5
4	125	89.6	87.9
5	150	78.9	79.8
6	127	89.9	89.9
7	100	90.5	98.7
8	130	87.9	99.5
9	125	84.5	93.8
10	120	86.5	95.7
Average		86.15	94.35

The Naïve Bayesian Classifier gets 86.15% and the Nearest Neighbor Cosine Classifier gets 94.35% average accuracy. From the evaluation results, we can conclude that a Nearest Neighbor Cosine Classifier approach has been effective and practical.The accuracy of the two classifiers are summarized in figure 5.

Figure 5. Two Classifiers accuracy

10.1 Error analysis

Figure 6 shows the distribution of error categories for Myanmar WSD. They can be categorized into three groups: missing common words in the training corpus, segmentation and same context features.

Figure 6. Error Analysis Chart for Myanmar word sense disambiguation

It can be observed that 90% of the errors come from training corpus coverage problems, while the remaining 10% can be eliminated by making some improvements to the current. The limitation comes mainly from the coverage of the training corpus.

10.2 Discussions

During the process of building and testing the proposed systems, the following observations are made.
- Myanmar Word segmenter will play a major role during disambiguation process. Due to wrong analysis by the word segmenter the word will be assigned incorrect sense.
- The creation of Sense tagged parallel corpus file will play a critical role in the sense disambiguation process.
- If this file provides the enough information then the performance of the proposed system is guaranteed to be high.
- Because of the insufficient context information, the system can assign the incorrect sense.

10.3 Limitation

The system has the following limitations:
- The system cannot disambiguate the same target word with different meanings in the same sentence because of bag-of-word condition.
- This system can disambiguate senses of words which are only in the training corpus.
- If probabilities or similarity of bag-of-word are the same, the system allocates the sense which has greater probabilities to the target word.

11. Conclusion and Future Work

In this paper, we described a comparison of Naïve Bayesian and Nearest Neighbor Cosine Classifiers for solving the ambiguity words in Myanmar language. The experiments show that Nearest Neighbor Cosine classifier has generated a better result than Naïve Bayes classifier. Due to unavailability of the earlier systems for the same tasks, we are not able to do the performance comparison of the proposed system. The system can improve the accuracy of Myanmar to English language translation. Moreover, we believe that our aims, thoughts, ideas and endeavors can be valuable in the areas that must have word sense disambiguation algorithm before it such as machine translation, grammatical analysis, speech processing and text processing.

As a future work, we plan to investigate the suitability of other algorithms for Myanmar word sense disambiguation such as Support Vector Machine, Decision Lists and Trees. The system use bag of word features only. Syntactic and collocation features may be useful to improve the performance of our system.

References

[1] M. Nameh, S.M. Fakhrahmad, and M. Zolghadri Jahromi, "A New Approach to Word Sense Disambiguation Based on Context Similarity", *Proceedings of the World Congress on Engineering 2011 Vol I, WCE 2011, July 6-8, 2011, London, U.K.*

[2] A.Naseer and S.Hussain, "Supervised Word Sense Disambiguation for Urdu Using Bayesian Classification", 2009.

[3] S. Parameswarappa and V.N.Narayana, "Kannada Word Sense Disambiguation for Machine Translation", *International Journal of Computer Applications* (0975 – 8887) Volume 34– No.10, November 2011.

[4] M. Apidianaki, "Translation-oriented Word Sense Induction Based on Parallel Corpora", 2008.

[5] Heshaam FAILI, "An Experiment of Word Sense Disambiguation in a Machine Translation System", 2008 IEEE.

[6] M. Miháltz and G. Pohl, "Exploiting Parallel Corpora for Supervised Word-Sense Disambiguation in English-Hungarian Machine Translation", 2006.

[7] Y. Motazedi and M. Shamsfard, "English to Persian Machine Translation exploiting Semantic Word Sense Disambiguation", *14th International CSI conference* (CSICC2009), July 1-2, 2009, Tehran, Iran.

[8] L. L. Tze and T. E. Kong, "Lost in Translation: Word Sense Disambiguation", *USM Computer Science Postgraduate Colloquium* 2004, 25-26th May 2004,Penang, Malaysia.

[9] E. Domingo and R. E. Roxas, "Utilizing Clues in Syntactic Relationship for Automatic Target Word Sense Disambiguation", *Journal of Research in Science, Computing and Engineering (JRSCE)* December 2006 Vol. 3 No.3.

[10] J.H.Oh and K.S. Choi "Word Sense Disambiguation using Static and Dynamic Sense Vectors", 373-1, Guseong-dong, Yuseong-gu, Daejeon, 305-701, Korea, 2002.

[11] W.A. Gale, K.W. Church and D. Yarowsky, "A Method for Disambiguation Word Senses in A Large Corpus", Computers and the Humanities, 26, pp.415-439, 1992.

[12] Myanmar Language Commission, "*Myanmar-English Dictionary*", 2008.

[13] N.T.T.Aung, K.M. Soe and N.L.Thein, "Ambiguous Myanmar Word Disambiguation System for Myanmar-English Statistical Machine Translation", *International Journal of Computer Applications* (0975-8887) volume 26-No.2, July 2011.

[14] N.T.T.Aung, K.M.Soe and N.L.Thein, "A Word Sense Disambiguation System Using Naïve Bayesian Algorithm for Myanmar Language", *International Journal of Scientific & Engineering Research Volume* 2, Issue 9, September-2011, ISSN 2229-5518.