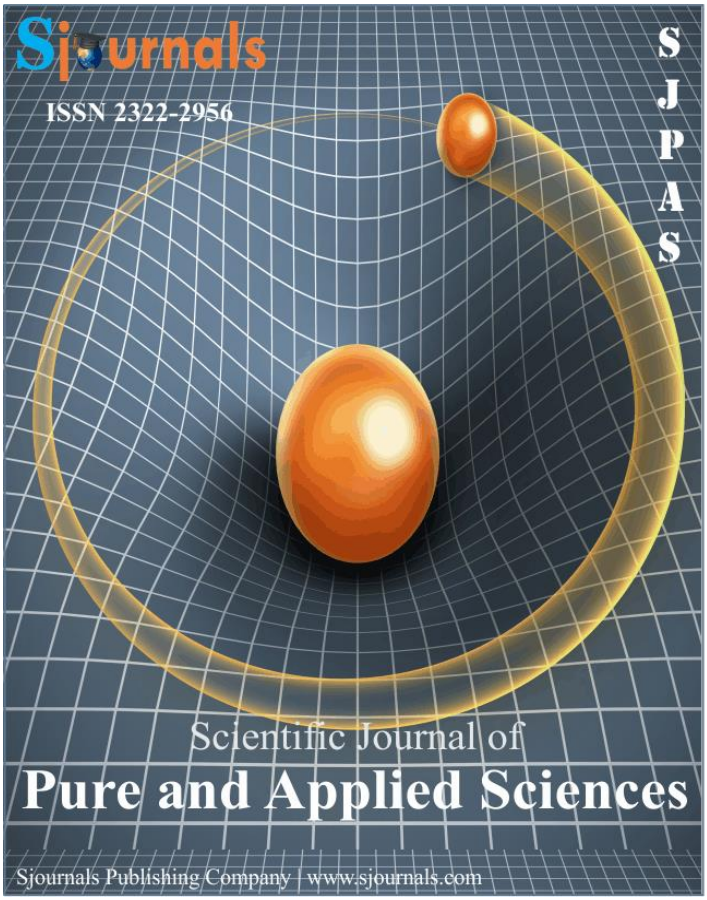
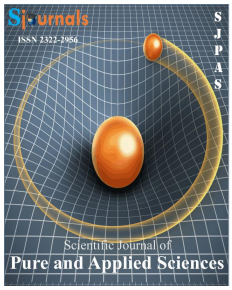


Sjournals
Scientific Journals
Scientific Journal of Pure and Applied Sciences (2020) 9(3) 915-925
ISSN 2322-2956
doi: 10.14196/sjpas.v9i3.2637



Contents lists available at Sjournals
Scientific Journal of **Pure and Applied Sciences**
Journal homepage: www.sjournals.com

Review article

Development of the lexico-conceptual knowledge resource for Myanmar NLP applications

Soe Lai Phyue*
University of Computer Studies, Mandalay, Myanmar.

*Corresponding author: soelaiphyue@gmail.com

ARTICLE INFO

Article history,
Received 12 February 2020
Accepted 13 March 2020
Available online 20 March 2020
iThenticate screening 14 February 2020
English editing 11 March 2020
Quality control 18 March 2020

Keywords,
Lexico-conceptual
Myanmar
NLP

ABSTRACT

The computational lexicons and semantic resources as WordNets of knowledge resources are mostly important for Natural Language Processing (NLP). The importance of computational lexicon is even greater in languages with rich morphology, where the lexicon is expected to provide morphological analyzers with enough information to enable them to correctly process intricately inflected forms. This paper presents for the semiautomatic construction of a multipurpose lexico-conceptual knowledge resources for NLP systems by acquiring the lexical and conceptual knowledge from WordNet and Myanmar<->English Machine Readable Dictionaries (MRDs). To construct the WordNet like lexicon, to collect the translation links, semantic meaning and synset links, the existing bilingual MRDs and English WordNet are used for building the knowledge resources. The collected of links and their meaning are manually verified. In addition, morphological processors (analyzers and generators) are support to handle cases of inflectional morphology. These resources are very efficient and effected for NLP applications, the building methodology of constructing Myanmar<->English Computational Lexicon and Myanmar WordNet lexical database reduced the human effort and time consuming.

1. Introduction

Natural Language Processing (NLP) is the one of the essential research for many productivity tools in many computer applications. The development of language resources and its availability is a must for enhancing language processing capabilities. A bilingual concept Lexicon is of increasingly great significance for Machine Translation (MT), Information Retrieval (IR). And it is for sure that the computational linguists would find such a lexicon indispensable and useful as semantic information when facing ambiguities in languages in their applications. The ongoing development of public knowledge bases such as WordNet, FrameNet, CYC, etc. has the potential to support domain independent solutions to NLP (Artale et al., 1997).

The Republic of the Union of Myanmar used to Myanmar Language as official language and there is no doubt in the necessity of constructing basic language processing resources for it. On the other hand, one of the most urgent problems in language technology is the lexical semantics bottleneck, the unavailability of domain-independent lexica with rich semantic information on lexical items. There are a number of semantic lexicons for English and some other languages but Myanmar lacks such a complete resource to be used in NLP works.

The development of Myanmar language resources are presented in this paper. Since the Myanmar language is complex language, the resource is poor for the language. The building of Myanmar WordNet which is widely useful for Word Sense Disambiguation(WSD) and computational bilingual lexicon which has a feature analysis of lexical entities that is useful for language translation, cross language information retrieval and like that. Construction of wordNet lexical database, particularly the Myanmar WordNet, is a long term project. The manual construction requires a large number of lexicographers to hand-build the lexical database. The origin of WordNet is a semantic word database based on psycholinguistic principles. It is a large-scale resource like LDOCE, but its information is organized in a completely different manner. WordNet groups' synonymous word senses into single units ("synsets"). Noun senses are organized into a deep hierarchy, and the database also contains part-of links, antonym links, and others (Pianta et al., 1990). Therefore, we considered the WordNet to collect the semantic meaning. Myanmar<-> English MRDs that provides translation relations between Myanmar and English words is used for word marching. The computational lexicon stores the word according their part of speech. Through a detailed study of the Myanmar language, we have been able to develop an analyzer that incorporates many of the unique features and challenges present in Myanmar.

This paper is organized as follows. The related module of language resources is described in section 2. The component of language resources that morphocon, Myanmar WordNet and Computational bilingual lexion is presented in section 3 and 4. In section 5, the statistical result and evaluation method of language resources is explained. Finally, Section 5 concludes the remarks.

2. Background environment of language resources

The lexicon presented in this paper has been developed by semiautomatic methodology of a multipurpose lexico-conceptual knowledge base Myanmar natural language processing (NLP) systems, and also useful for Natural Language Understanding. Moreover, the component of the system is separated component, so it can be sense as multifunctional. Thus, it has been designed to be potentially reused in many NLP tasks (e.g. Information Retrieval (IR) and extraction (IE), machine translation, dialogue-based systems, etc). This system comprises three major resources which are morphocon, Myanmar WordNet, bilingual computational lexicon as shown in Figure 1. These resources are formed several independent but interrelated modules.

2.1. Myanmar English bilingual computational lexicon

Myanmar English computational lexicons are among the most important resources for Myanmar natural language processing (NLP) application. Their importance is even greater in Myanmar language that with rich morphology, where the lexicon is expected to provide morphological analyzers with enough information to enable them to correctly process intricately inflected forms. It contained the 71 tag set for 8 part of speech.

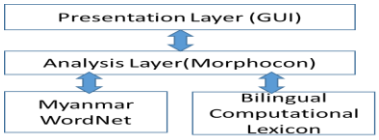


Fig. 1. Methodology of language resources.

2.2. Myanmar WordNet

The most ambitious feature of Myanmar WordNet, however, is its attempt to organize lexical information in terms of word meanings, rather than word forms. Therefore, Myanmar WordNet which contained semantic concept is crucial role in NLP because WordNet gives not only lexical information of words but also semantic meaning of each word.

2.3. Morphocon

The quality of a morphological analyzer greatly depends on the quality of the lexicon. A morphological analyzer must consult with the lexicon to check whether a theoretical analysis of a word indeed belongs to the language. This system facilitates a modular development of morphological analysis and disambiguation systems. The morphological analyzer interacts with, but is separated from the lexicon.

3. How to work and how to build morphocon

As the morphological relations between word forms is one of the important class of lexical relations Computational lexicon and WordNet as a language resource became increasingly obvious that have to deal with inflectional morphology (Fellbaum, 1998). For example, in WordNet, the word of input sentence or phrase include as "သစ်ပင်များ" and clicked a request for information, WordNet should not reply that the word was not in the database. A program was needed to strip off the plural suffix and then to look up "သစ်ပင်" which certainly is in the database. On the side of lexicon, we need to generate the relevant information as "trees" in English word. This need led to the development of a program for dealing with inflectional morphology. Although the inflectional morphology of Myanmar is relatively simple, writing a computer program to deal with it proved to be a more complex task than had been expected.

Our current morphological analyzer performs analysis by generation: this is basically the rule of the morpheme of the Myanmar word which contain in for WordNet and grammar pattern relation between Myanmar and English word for lexicon. The basic idea is to first generate all the inflected forms is useful for NLP application and induced by the lexicon. It is common to think that for languages with rich morphology such a method is impractical. While this may have been the case in the past, contemporary computers can efficiently store and retrieve millions of inflected forms. Of course, this method would break in the face of an infinite lexicon (which can easily be represented with FST), but for most practical purposes it is safe to assume that natural language lexicons are finite. The morphological analyzer is obtained by inflecting the base forms in the lexicon. The number of inflected forms is used by the analysis program. The morphological analyzer is shown in Figure2. The analysis start with application of a decomposition system defined by morphological grammar, to each word of the text to identify it's radical and affixes. In the second step, grammars (finite-state transducers) produce lexical constraints checking the validity of segmentation thanks to a dictionary lookup. So, these grammars associate the recognition of a word to lexical constraints, working only with valid combinations of the various components of the form. Typically, there are several output strings, each representing a possible analysis of the input word (Phyue and Thida, 2012).

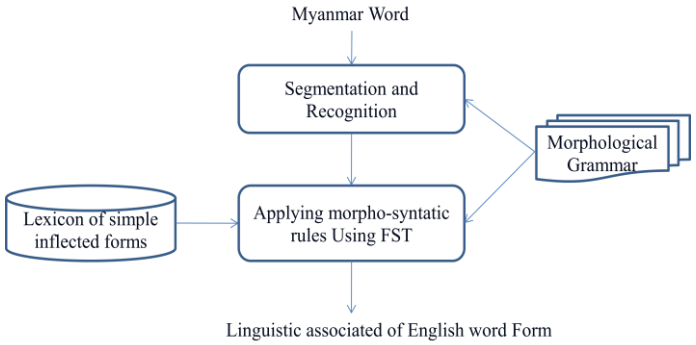


Fig. 2. Chain of a text form morphological analysis.

A. Relationship between Myanmar and English noun form

The noun in Myanmar word can have a suffix indicating plurality. It can be pluralized by suffixing the particle ”တွေ” in colloquial Burmese or “များ” in formal Burmese. The particle “တွေ” which indicates a group of persons or things, is also suffixed to the modified noun. Therefore, a word ‘စာအုပ်’ and it followed by ‘များ’ it has one translation word as English “books”. We ignored the suffix as ‘များ’and ‘တို့’ in Myanmar WordNet however we should consider it as lexicon. There plural form of (‘စာအုပ်’) PL(‘များ’) is translated by English as ‘Books”. There we used the Finite State Translator (FST) as followed in Figure 3.

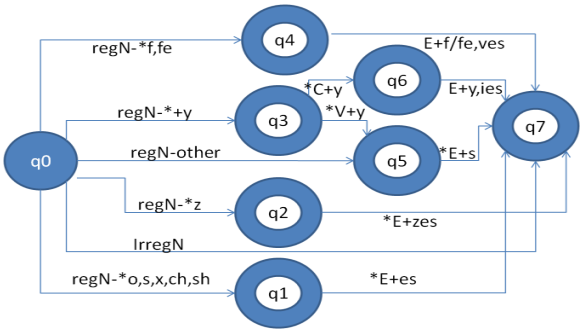


Fig. 3. Finite state automata for Singular noun to plural form.

B. Relationship between Myanmar and English verb form

In the affirmative, the order of elements is V [one or more roots, possibly compounded] (+auxiliary verb) + aspect particle + modal ending. The most commonly used verb particles and their usage are shown below with an example verb root “စား”. Alone, the statement “စား” is imperative. The suffix “တယ်” (literary form: သည်) can be viewed as a particle marking the present tense and/or a factual statement. The suffix ခဲ့ denotes that the action took place in the past. Note that the suffix “သည်” in this case denotes a factual statement rather than the present. The particle “နေ” is used to denote an action in progression. It is equivalent to the English '-ing'". This particle “ပြီ” which is used when an action that had been expected to be performed by the subject is now finally being performed, has no equivalent in English. So in the above example, if someone had been expecting you to eat and you have finally started eating, the particle ပြီ is used. The particle “မည်” “မယ်” “တော့မည်” ”တော့မယ်” “လိမ့်မယ်” “လိမ့်မည်” are used to indicate the future tense or an action which is yet to be performed. An FSA for English derivational morphology structure for English word is as shown in Figure 4.

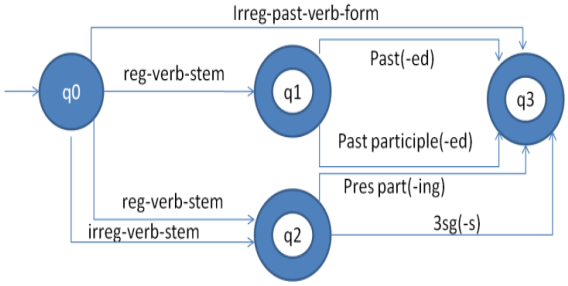


Fig. 4. Finite state automata for verb tense.

C. Relationship between Myanmar and English adjective form

In Myanmar word, adjective is defined by the word with suffix “သော”, “သည့်”, “မည့်”. Beside then, it has verbs that carry the meaning "to be X", where X is an English adjective. Reversely, some of the verb can transform

as adjective by following the adjective particles. Comparatives are usually ordered: X + "ထက်ပို" "ပို" "ပို၍" + adjective, where X is the object being compared to. Superlatives are indicated with the prefix အ + adjective + ဆုံး. The FST for Myanmar word to English degree transformation is as shown in Figure5.

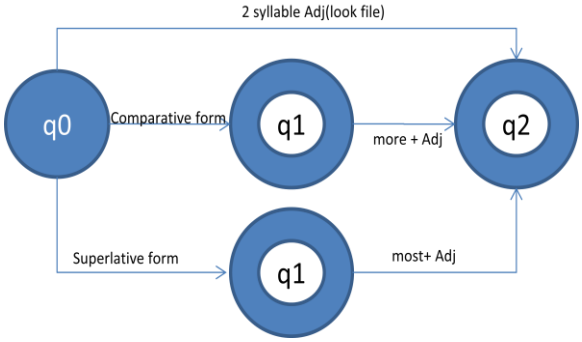


Fig. 5. Finite state automata for degree of adjactive.

4. How to work and how to build Myanmar WordNet and computational lexicon

The most ambitious feature of Myanmar WordNet, however, is its attempt to organize lexical information in terms of word meanings, rather than word forms. Therefore, Myanmar WordNet which contained semantic concept is crucial role in NLP because WordNet gives not only lexical information of words but also semantic meaning of each word. Manual construction of WordNet is the most reliable technique for obtaining structured lexicons but it is costly and highly time-consuming (Sathapornrungkij, 2004). Myanmar WordNet is constructed by using the origin of WordNet lexical database and Myanmar<->English MRDs with the semiautomatic methodology (Phyue, 2011). The relations WordNet applied to the noun and verb concepts are synonymy, antonymy, hypernymy, holonymy, entailment, cause and etc., among which synonymy and hypernymy are the most important. Synonymy and hypernymy help to form the SynSets and their hierarchies respectively. The hypernymy tree, as the hierarchy of Concepts, provides a common way of making induction for the NLP researchers (Neff et al., 1993). The words in Myanmar WordNet are stem word and it does not contain in inflected form of noun as plural and verb of various tense form.

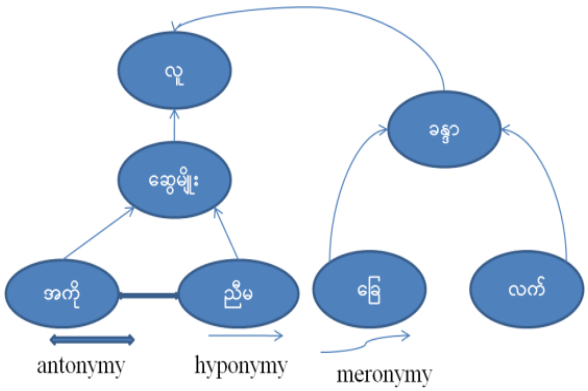


Fig. 6. Network representation of noun concept.

According the origin of WordNet, synsets which based on syntactic category and semantic coherence for each noun synset is assigned one out of 26 broad categories. A graphical representation of a fragment of the noun network is shown in Figure 6 which included all three kinds of semantic relations-hyponymy, meronymy, and antonymy and the result is a highly interconnected network of nouns. There is enough structure to hold each lexical concept in its appropriate place relative to the others.

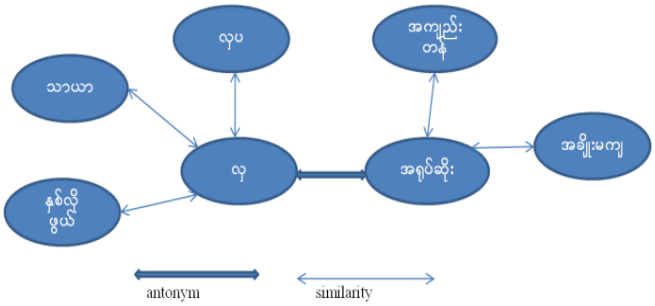


Fig. 7. Network representation of adjective concept.

WordNet divides adjectives into two major classes: descriptive and relational. Descriptive adjectives that do not have direct antonyms are said to have indirect antonyms by virtue of their semantic similarity to adjectives that do have direct antonyms. The configuration that results is illustrated in Figure7 for the cluster of adjectives around the direct antonyms, "ရုပ်ဆိုး"and "လှ".

Verb categories in WordNet divided into 15 class. For the verb in WordNet described by entailment is included. For example, the word "အိပ်" and which entailment included is "ဟောက်" as shown in Figure 8. The labels of Myanmar verb are also divided according to the WordNet.

Fig. 8. Entailment representation of verb concept

In computational lexicon, POS are specified for their function and their class according the WordNet structure. But it also lists several lexical properties which are specifically targeted at morphological analysis. A typical example is the plural suffix as "များ" "တို့. " "တွေ" for nouns: while by default, this suffix is -s or –es or others of lexical items are idiosyncratic. The lexical representation of verbs is more involved. Here, the lexicon stores two main pieces of information: a root and an inflection pattern. Therefore, the design is greatly depending on Myanmar WordNet lexical database structure and information. In this lexicon defined the noun as 26 tag set, verb as 15 tag set, adjective as 3 tag set and adverb has 2 tag set, proposition as 17 and conjunction as 8 tag set. Beside then particle are used as indicator for defining the definite POS and produced the inflected form of word (Phyue, 2011). The computational lexicon contains the following information as shown in Table 1.

Table 1
Summary information of bilingual computational lexicon.

	Noun	Verb	Adjective	Adverb	Pronoun	Preposition	Conjunction
Myanmar	•	•	•	•	•	•	•
English	•	•	•	•	•	•	•
categories	•	•	•	•	•	•	•
lexname	•	•	•	•	•	•	•
POS description	•	•	•	•	-	-	-
Definitions (Synsets)	•	•	•	•	-	-	-
Sample	•	•	-	-	-	-	-
Adj-modifier	-	-	•	-	-	-	-
Semantic-Frame	-	•	-	-	-	-	-
Lexical-Frame	-	•	-	-	-	-	-
Inflected Sense	•	•	•	-	-	-	-

To build a Myanmar WordNet (Phyue, 2011) in the same route just as Princeton had taken and then to construct the mapping between these two WordNets may be not a satisfying idea. So, it is crucial that we had

better find an approach to reusing the English common knowledge already described in WordNet as the semantic basis for Myanmar when building the Myanmar WordNet and bilingual computational lexicon. And this kind of reusing should contain some capabilities of adjustments to the bilingual concepts besides word-for-word translations. By reusing the existing resources and manage this resources, not only the building of the monolingual Myanmar WordNet but also Myanmar-English computational lexicon benefits. The architecture of the system is as shown in Figure 9.

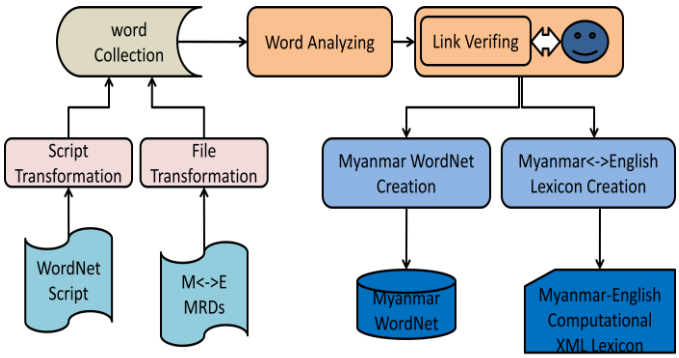


Fig. 9. Architecture of the building Myanmar WordNet and Mynamar-English computational lexicon.

WordNet2Sql is used to collect the semantic link between English and their synsets and Myanmar<->English MRDs are used to collect the translation links. The script of WordNet structure and Myanmar<->English word are stored in a relational database to align format and further processes. In this phase, data from Monolingual (WordNet) and Bilingual MRDs (Myanmar<->English MRDs) are extracted for link analyzing phase. These are done with separated processes for each MRDs. WordNet2Sql is used to collect the semantic link between English and their synsets and Myanmar<->English MRDs are used to collect the translation links. It provides access to the underlying machine-readable dictionaries and support data restructuring and data cleansing. The English-Myanmar and Myanmar-English MRDs are aligned format of font and storage structure and needed to remove noise, erroneous and duplicate data. The existing resources of translation word of Bilinguals are combined and then the duplicate records are removed from combining result. Our system used the wordnet2sql relational database WordNet. The extracted data is used in Link analyzing phrase to drive the links between Myanmar words and synsets.

The result of each process is as shown in Table 2 and this data are used in link analyzing phase to create the candidate of translation link.

Table 2
Result of data statistic.

				Grouping						
Dictionary	**Original**	**Cleaning**	**Merging**	**Noun**	**Verb**	**Adj**	**Adv**	**Pron**	**Prep**	**Conj**
E-M Dic	29625	23855	-	-	-	-	-	-	-	-
M-E Dic	30722	29049	-	-	-	-	-	-	-	-
Merge	-	-	55277	31243	12706	6819	3369	170	108	185
WordNet2 Sql	148730	148730	148730	-	-	-	-	-	-	-

The analyzer supports classifying translation links with respect to semantic links and supports deriving candidate links. The translation links are referred to as the relationships between Myanmar and English words. The semantic links are referred to as the relationships between English words and their meaning. The candidate links are referred to as the relationships between Myanmar words and their meaning. In English WordNet, words are organized by synset which is a set of words having the same meaning. They can be classified in two main methods according to the kind of knowledge sources involved in the process which is class and structure method. In this research, we aim at constructing Myanmar WordNet in which Myanmar words are organized by synsets. The synsets in Myanmar WordNet can be derived from the English WordNet. The candidate links between Myanmar

words and synsets can be derived from the semantic links which are obtained from WordNet and the translation links which are obtained from Bilingual MRDs. The candidate of translation links are sampled and then verifies candidates by taking advantage of the statistic approach to reduce human intervention and time consuming. Candidate set of translation link is sample by using the random stratified sampling technique. For our system, the candidate links are already classified into several stratums as criteria. Then the candidate links are randomly sampled with the number of the 400 translation links in each criterion. In this research, the stratified sampling technique is applied with 95% confidence level which is the most common confidence level for research purposes (Krejcie and Morgan, 1970). The link verifier verified the links which result from analyzer with the help of human being. If the word and synset relation is wrong, the link is discarded.

The probability that a link would be correct can be estimated by *P(OK) = NOK/NTOT*. In this research, we use the logistic regression model to predict the correctness of the remaining links. In general, the logistic regression is used to predict a discrete outcome from a set of binary variables. The linear logistic regression model can be defined as

$$P(OK) = \log(pi) = ln\left(\frac{pi}{1-pi}\right) = \beta_0 + \beta_1 c_{01} + \beta_2 c_{02} + \cdots + \beta_{13} c_{13} \quad (1)$$

Where *nok* is the number of correct evaluation for the set of solutions of every group of methods, *ntot* accumulates the total number of evaluations, *Ci* is a boolean variable representing the existence of link in the *i*th criterion, and β*i* is unknown parameter which required the least square criterion. Figure 10 shows the correctness (accuracy) of each criterion over the 400 sampling link.

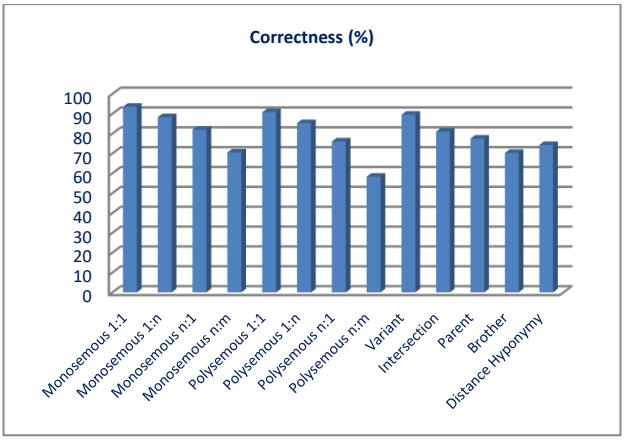


Fig. 10. Correctness of Link verification.

To evaluate the model, a statistical approach is used. By using Pearson goodness-of-fit, determines the significance of model. A P-value of each criterion describes the significant of that criterion in the model by explaining the probability of a link of being correct. For a P-value lower than 0.05, the criterion is significant in the model. The model is deployed to construct the Myanmar WordNet. By applying the model to all remaining translation links, the semantic relations between Myanmar words and synsets are constructed according to each criterion.

The result of verified links is useful not only for building Myanmar WordNet but also for building the bilingual computational lexicon. The structure of noun, verb, adjective and adverb are built by their structure in Myanmar WordNet. Utilize English WordNet as a skeleton, the logical structure of Myanmar WordNet (relation of each word) and bilingual computational lexicon (frame of each word) is translated. The theory of transitive relation (in basic discrete mathematics) will be adapted for inference the relations and the frame of English WordNet.

"Transitivity is a mathematical property of binary relations such that if A and B are related, and B and C are related, then it follows that A and C are also related, for all A, B and C for which the relation may apply. This relation is said to be transitive. In construction phase, transitivity relation, equation (2) is taken into account.

$$A\ R\ B \wedge B\ R\ C \Rightarrow A\ R\ C \quad (2)$$

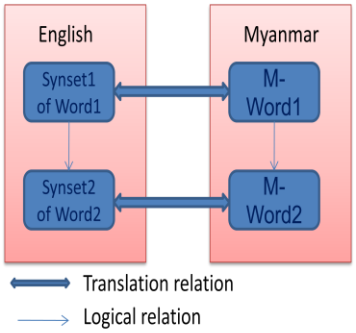


Fig. 11. Inference logical relations.

The relations between words and synsets in English WordNet can also be applied with this property, then relations in English WordNet will be inferred to Myanmar internal relation that mean English word in sysnset1 has internal relation with another English word in synset2, if both English words has translation link to Myanmar words then the internal relation will infer to Myanmar as well. That means Myanmar word1 and Myanmar word2 will have logical relation similar to English Word1 and English Word2 as shown in Figure11. All possible relations will be inferred and constructed.

5. Statistical results and evaluation

The dictionary statistics and set of link derivation for Myanmar WordNet according to their part of speech is as shown in Table 3. The quality of WordNet depends on number of synset link and the link which covered the meaning for each word. According the result, the set of Myanmar word and their relation yielded a satisfactory result in Myanmar WordNet, it is also useful for computational lexicon.

Table 3
Dictionary statistic of the resources and coverage of Myanmar WordNet.

	Noun		Adjective		Adverb		Verb	
	Words	**Synsets**	**Words**	**Synsets**	**Words**	**Synsets**	**Words**	**Synsets**
Bilinguals dictionary	31243	-	6859	-	3369	-	12706	-
WordNet	119034	82115	21538	18156	4481	3621	11531	13767
Maximum coverage	15752	20094	2458	4495	1398	1275	2434	6746
Coverage of bilinguals	50.41%	-	35.83%	-	41.49%	-	19.15%	-
Coverage of WordNet	13.23%	24.47%	11.41%	24.76%	31.2%	35.21%	21.11	49.00%

A. Test case

There are two alternatives to evaluate the performance of the Myanmar language resources. The first is to examine the mapping by human directly. The second is to employ the linguistic resource to an application, and show its performance indirectly. In this study, we adopt the second approach. Myanmar Word Segmentation and translation based on the Myanmar WordNet and Myanmar-English WordNet like lexicon is experimented. We employ text collection of 50 sentences. Moreover, these sentences queries into English are translated by human. The translator is independent of the persons who conduct the query experiments.

B. Word segmentation

To process text computationally, syllables have to be determined first. Since Myanmar linguistic tradition there is not a clear-cut, we need to pass the tokenize syllable as a one of the step. Word tokenization is done by rules based approach. Segmentation and pattern merging is done by proposed algorithm of Phyue (2012). The merging of tokenized word is use as input to the algorithm. The output of the algorithm is the segmentation word and possible pattern. Overview of the segmentation system that consumes our language resource is as shown in

Figure 12. There are 1,017 words in 50 sentences. After removing stop words, there remain 703 words. Total 484 words can be found in synsets.

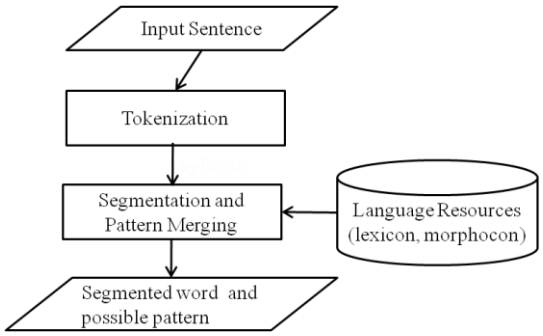


Fig. 12. Overview of segmentation process.

There are still several rooms to improve the Myanmar-English lexicon and Myanmar WordNet. The coverage of the bilingual and monolingual WordNet is not so good enough. Based on the approach in this paper, over 50% word entries in WxpyDic are put into the WordNet. This is because English translations of a Myanmar word may not be found in the Myanmar-English dictionary, and the WordNet may not gather the English translations even dictionary look-up is successful. In particular, WordNet lacks proper names. Besides, the bilingual dictionary, which acts as a bridge to link the different resources, is not a sense-based dictionary. That is, we do not know the exact sense of each dictionary entry. Thus the method we proposed is served as a computer-aided tool to construct a knowledge base in a more effective way.

Table 4 shows the distribution of these parts of speech in the two formats: The first column is the distribution of the root forms in the lexicon files, and the second column is tile distribution for the inflected forms derived from the morphocon.

Table 4
Statistic of the lexicon with morphocon.

Parts of speech	No. of root form	No. of inflected form
Noun	31243	110846
Verb	12720	44979
Pronoun	170	170
Adverb	3369	3369
Adjective	6819	1726
Preposition	108	108
Conjunction	185	185

6. Conclusion

Myanmar WordNet and Myanmar English Computational lexicon are new state for Myanmar lexical resources. The morphological analysis and morphological generation of morphocon is very pleasurable for Myanmar NLP applications. This paper described a method of constructing Myanmar WordNet, a lexical database in which Myanmar words are organized by their meanings and computational lexicon for Myanmar to English with inflected form. This system takes WordNet and Bilingual machine-readable dictionaries into account. The semantic relations between English words in WordNet and the translation relations between English and Myanmar words in English<-> Myanmar machine readable dictionary are considered.

Although automatic dictionary merging is an old line of research and the dictionaries merged were roughly similar, while in our work, we have chosen the very different resources. Another motivation for merging dictionaries is to get several definitions for the same sense, to maximize the information that can be extracted by analyzing those definitions. We have not yet extracted information from WordNet definitions, though this is a clear source of knowledge for enriching the lexicon, and there is a great deal of network to build on Myanmar WordNet

and bilingual computational lexicon. In the later, we will evaluate the effectiveness of our proposed Myanmar WordNet and computational lexicon with morphocon by using the other NLP applications.

References

Artale, A., Magnini, B., Strapparava, C., 1997. Lexical discrimination with the Italian version of WordNet. In Proceedings of ACL Workshop on Automatic Information Extraction and Building of Lexical Semantic Resources.

Dzikovska, O., Swift, D., Allen, F., 2004. Building a computational lexicon and ontology with FrameNet. Proceeding of the Workshop on Building Lexical Resources from Semantically Annotated Corpora, LREC.

Fellbaum, C., 1998. WordNet: An electronic lexical database. MIT Press, Cambridge, Massachusetts.

Krejcie, R.V., Morgan, D.W., 1970. Determining sample size for research activities. Educ. Psychol. Meas., 30, 607-610.

Neff, M.S., Blaser, B., Lange, J.M., Lehmann, H., Dominguez, I.Z., 1993. Get it where you can: Acquiring and maintaining bilingual lexicons for machine translation. In Proceeding of AAAI.

Phyue, S.L., 2011. Building Myanmar language resources. Appl. Inform. Comput. Technol.

Phyue, S.L., 2011. Construction of Myanmar WordNet Lexical Database. IEEE student Conference on Research and Development, Malaysia.

Phyue, S.L., 2012. Lexical analyzer for Myanmar language. Proceedings of the International Conference on Computer Applications, Yangon, 95-101.

Phyue, S.L., Thida, A., 2012. Morphological processor for inflectional case of multipurpose lexico-conceptual knowledge resource. Int. J. Adv. Res. Comput. Eng. Technol., 1(7), September.

Pianta, E., Bentivogli, L., Girardi, C., 1990. MultiWordNet developing an aligned multilingual database. ITC-irst, Centro per la Ricerca Scientifica e Tecnologica.

Prószéky, G., Mihăltz, M., 2002. Semi-automatic development of the Hungarian WordNet. Proceedings of the LREC 2002 Workshop on WordNet Structures and Standardization, Las Palmas.

Sathapornrungkij, P., 2004. A semi-automatic construction of Thai WordNet lexical database from machine-readable dictionaries. Master's Thesis, Mahidol University.

How to cite this article: Phyue, S.L., 2020. Development of the lexico-conceptual knowledge resource for Myanmar NLP applications. Scientific Journal of Pure and Applied Sciences, 9(3), 915-925.