International Journal of Scientific & Engineering Research, Volume 4, Issue 8, August 2013
ISSN 2229-5518

Morphological Analysis of Punjabi Verbs for UNL Based Machine Translation System

Parteek Kumar, RK Sharma

Abstract— Morphological analysis of verbs is an important activity in order to design a machine translation system for a language. This paper discusses morphological analysis of Punjabi verbs for developing Universal Networking Language (UNL) based Machine Translation (MT) system for this Language. Verbs play a significant role in extracting Gender, Number, Person, Tense, Aspect and Modality (GNPTAM) information from a given sentence. An analysis of Punjabi verbs has been carried out on the basis of verb paradigms in this work. There are approximately one hundred verb paradigms defined on the basis of their vowel ending and GNPTAM information in Punjabi language. Based on these paradigms, one thousand rules have been created with respect to UNL attributes in this analysis. This analysis has been carried out for simple as well as conjunct verbs. It has been found that a good number of conjunct verbs are formed in Punjabi language by adding the verbs kar (do) or ho (be) to nouns or adjectives. Verb morphology in UNL based EnConverter and DeConverter for Punjabi language has also been implemented in this paper.

Index Terms— Universal Networking Language, EnConverter, DeConverter, Vowel Ending Verb Root, Verb paradigms, Verb Morphology.

1 INTRODUCTION

Verbs are the most important categories in parts of speech owing to their distinctive behavior. These form the pivot of a sentence and their usage conveys a good amount of information about the activities in the sentence. The lexical knowledge of the sentence is generally characterized with respect to verbs being used. Grammatical information like tense, number, gender etc. can be extracted from the sentence on the basis of verbs. These are some of the reasons for immense importance given to verb analysis in Linguistics [4]. In this paper, analysis of Punjabi verb morphology for the development of Universal Networking Language (UNL) based Machine Translation (MT) system for Punjabi Language has been presented. Punjabi Language is a member of Indic branch of the Indo-European family [18]. According to the Ethnologue (2005) estimates, there are 88 million native speakers of Punjabi language, which makes it the 13th most widely spoken language in the world.
This paper is divided into 6 sections. This section highlights the role of verb morphology in UNL and also provides introductory details of UNL framework. Related work on morphology and UNL is described in Section 2. Detailed verb morphology for Punjabi Language is presented in Section 3. Section 4 of this paper discusses the verb morphology for UNL attributes and implementation of Punjabi EnConverter and DeConverter is illustrated in Section 5. Section 6 provides the results of implementation and summarizes the work carried out in this paper.

1.1 UNL BASED MACHINE TRANSLATION SYSTEM: THE FRAMEWORK

UNL is an Interlingua for knowledge representation in the context of machine translation. UNL is an electronic language for computers to express and exchange information [24]. Three building blocks of UNL are (i) Semantic Relations, (ii) Attributes and (iii) Universal Words. UNL representation of a sentence is expressed in the form of a semantic net called UNL graph. Consider the following English sentence for its UNL representation:
'Sukwinder played football in the garden'. (1)
UNL representation of (1) shall be:
{unl}
agt(play(icl>do).@entry.@past, Sukhwinder(iof>person))
obj(play(icl>do).@entry.@past, football(icl>game))
plc(play(icl>do).@entry.@past, garden(icl>place))
{/unl} (2)
In this expression, *agt (agent), obj (object)* and *plc (place)* are the semantic relations and the words *play(icl>do), Sukhwinder(iof >person), football(icl>game),* and *garden(icl>place*) are the Universal Words (UWs).

1.2 UNL ATTRIBUTES FOR REPRESENTATION OF INFORMATION

UNL attributes are used to describe the subjectivity information of sentences. These store the information about what is said from the speaker's point of view. UNL has 87 primary attributes (this number can be augmented by user defined ones) to express the semantic content of a sentence [25][23]. Table 1 describes some of the UNL attributes used to represent the knowledge extracted from an input sentence.

1.3 ROLE OF WORD MORPHOLOGY IN UNL BASED MT SYSTEM

UNL based MT system has two major components namely EnConverter and DeConverter. The process of converting a source language (natural language) expression into the UNL expression is referred to as EnConversion and the process of converting UNL expressions into a target language representation is called DeConversion. The process of EnConversion consists of four main stages: parsing of input sentence, extraction of universal words from bi-lingual dictionary, resolution of UNL relations and generation of UNL attributes [14][8]. Generation of UNL attributes can be carried out by creating morphological rules for the target language on the basis of verb mor-

- *Parteek Kumar is working as an Assistant Professor in Computer Science and Engineering Department, Thapar University, Patiala, India. E-mail: parteek.bhatia@gmail.com*
- *RK Sharma is working as a Professor in SMCA Department, Thapar University, Patiala, India. E-mail: rksharma@thapar.edu*

phology. Similarly, DeConversion process consists of three main stages: morphological generation of lexical words, function word insertion and syntax planning [23]. In order to perform morphological generation, analysis of verb morphology of a language is very important. This has been attempted for Punjabi language in this paper.

TABLE 1
SOME OF UNL ATTRIBUTES

Concept	UNL Attributes
Time with respect to the speaker	@past, @present, @future
Speaker's view of Aspect	@begin, @complete, @continue
Speaker's view of Reference	@generic, @def, @indef
Speaker's Focus	@emphasis, @entry, @qfocus
Speaker's attitudes	@confirmation, @exclamation, @interrogative
Speaker's view point	@ability, @although, @conclusion, @doubt
Speaker's attitudes	@polite, @request

2 RELATED WORK

Analysis of Punjabi grammar has earlier been carried out by Chander (1964), Gill et al. (1986), Singh (1988), Puar (1990), Joshi (2000), and Singh (2002). Their studies form the basis for the Natural Language Processing (NLP) systems for Punjabi language. Gill (2008) has developed a rule based part of speech tagger for Punjabi. He has also developed the morphological analyzer, morphological generator and a grammar checker for Punjabi language.

Analysis of Hindi grammar for part of speech tagger has been performed by Chakrabarti and Bhattacharyya 2002, Chakrabarti et al. 2006 and Singh et al. 2006. Shrivastava et al. (2005) have identified 622 unique verb paradigms in the TAM-GNP matrix. The analysis of Tamil morphology for the development of Tamil EnConverter for EnConversion of Tamil to UNL has been performed by Dhanabalan et al. 2002. Bengali morphology has been analyzed with respect to UNL for Bengali EnConverter by Ali et al. 2008. UNL based analysis and generation of Bengali case structure constructs have also been performed by Dey and Bhattacharyya 2003. Arabic grammar generator has been proposed for the development of Arabic MT System based on UNL by Adly and Alansary 2009. Hindi grammar has been analyzed to create UNL based MT system for Hindi language. Hindi generation rules have been created for Hindi DeConverter by Dwivedi 2002, Smriti et al. 2006 and Nalawade 2007. The UNL generation rules for Hindi EnConverter have been created by analyzing Hindi grammar by Giri 2000, Dave et al. 2001, Dave and Bhattacharyya 2001. The relevant work on Punjabi language also leads to establish that UNL related work has not been done for Punjabi language.

3 VERB MORPHOLOGY IN PUNJABI LANGUAGE

Punjabi verbs are classified as main verbs and auxiliary verbs. The main verbs have assigned transitivity and causativity. Transitive verbs are those that require an object in a sentence unlike intransitive verbs. For example, ਪੜ *paṛ* 'study' is transitive as in ਉਸਨੇ ਕਿਤਾਬ ਪੜ੍ਹੀ *usnē kitāb paṛhī* 'She read book' and verb ਦੌੜ *dauṛ* 'run' is intransitive as in ਮੁੰਡਾ ਦੌੜ ਰਿਹਾ ਸੀ *muṇḍā dauṛ rihā sī* 'The boy was running'.

There are two types of causatives used in Punjabi. These are simple causative and double causative. In general, the simple causative is formed by adding ਆ *ā* and double causative is formed by adding ਵਾ *vā*. For example, causative forms of verb root ਪੜ *paṛ* 'study' are ਪੜਾ *paṛā* (simple causal) and ਪੜਵਾ *paṛvā* (double causal). All transitive, intransitive and causative verbs are also known as simple verbs. There are three auxiliary verbs used in Punjabi. These are ਹੈ *hai* for present tense, ਸੀ *sī* for past tense and ਰਹੇਗਾ *rahēgā* for future tense.

3.1 CONJUNCT VERBS

Expressing the concept of a given word, in source language, may require two or more words in target language. Many verbs in English can be translated into Punjabi, only by using a noun-verb sequence (*e.g.*, ਸ਼ੁਰੂ ਕਰਨਾ *srū karnā* (start), ਦਿਖਾਈ ਦੇਨਾ *dikhāī dēnā* (visible)) or only by using an adjective-verb sequence (*e.g.*, ਮੀਠਾ ਲਗਨਾ *mīṭhā lagnā* (sweet), ਚੰਗਾ ਲਗਨਾ *caṅgā lagnā* (like)) or only by using an adverb-verb sequence (*e.g.*, ਦੂਰ ਹਟਾਨਾ *dūr haṭānā* (remove)) [5][22].

3.2 GRAMMATICAL PROPERTIES OF VERBS

Verbs play a vital role in extracting the Gender, Number, Person, Tense, Aspect and Modality (GNPTAM) information from a given sentence [22]. It provides the gender information like masculine, feminine or non-specific; number information like singular, plural or non-specific; person information like 1st, 2nd or 3rd; tense information like past, present or future about the sentence. It also provides the information about the aspect *i.e.* perfective, completive, frequentative, habitual, durative, inceptive or stative and modality information like imperative, probabilitive, subjunctive, conditional, deontic, ab ilitive or permissive about the sentence.In the next section, verb paradigm for Punjabi language is discussed.

3.3 VERB PARADIGMS

In order to capture the morphological variations of verbs, they are categorized into various paradigms to identify uninflected forms of words that share similar inflections [22]. There are approximately one hundred verb paradigms classified based on their vowel ending and GNPTAM information [12]. Punjabi has ten vowels, represented as ਾ (ਆ 'A'), ਿ (ਇ 'i'), ੀ (ਈ 'I'), ੁ (ਉ 'u'), ੂ (ਊ 'U'), ੇ (ਏ 'y'), ੈ (ਐ 'Y'), ੋ (ਓ 'o'), ੌ (ਔ 'O') and Mukta (ਅ 'a') which has no sign. Vowels other than ਅ (ਮੁਕਤਾ) are represented by accessory signs written around (*i.e.*, below, above, to the right or to the left) the consonant signs, popularly known as signs for *matras* [18].

Some of the important verb paradigms are given in Table 2.

TABLE 2
SOME IMPORTANT VERB PARADIGMS

Sr. No	Verb paradigm	Verb Root	Present	Past	Future
1	ਪੌਣਾ (ਪਾਉਣਾ) pauṇā (pāuṇā) to put	ਪਾ pā	ਪਾਉਨਾ pāunā	ਪਾਇਆ pāiā	ਪਾਊਂਗਾ pāūṅgā
2	ਲੌਹਣਾ (ਲਾਹੁਣਾ) lauhṇā (lāhuṇā) to pull down	ਲਾਹ lāh	ਲਾਹੁਣਾ lāhuṇā	ਲਾਹਿਆ lāhiā	ਲਾਹੂੰਣਾ lāhūṇṇā
3	ਨੂੰਣਾ nūṇṇā to bath	ਨ੍ਹਾ nhā	ਨ੍ਹਾਉਣਾ nhāuṇā	ਨ੍ਹਾਤਾ nhātā	ਨ੍ਹਾਉਂਗਾ nhāuṅgā
4	ਭੌਣਾ bhauṇā to wonder	ਭੌ bhau	ਭੌਨਾਂ bhaunāṃ	ਭੌਵਿਆ bhauviā	ਭੌਊਂਗਾ bhauūṅgā

4 VERB MORPHOLOGY FOR UNL ATTRIBUTES

In this section, verb morphology of UNL attributes has been discussed. Rules for verb morphology can generally be classified into three types as given below.

(a) General verb morphology rules,
(b) Verb morphology rules for passive sentence and
(c) Verb morphology rules for conjunct verbs.

The information about GNPTAM is generally extracted from subject of a sentence. For passive sentences this information is extracted from the object of the sentence. In this situation, the UNL attribute @passive is associated with the main verb and @topic is associated with the object of the sentence [17].

Some examples of verb morphology for simple verbs with respect to UNL attributes are shown in Table 3. The Verb morphology for conjunct verbs is given in Table 4 and Table 5. Here, morphology is formed by adding nouns or adjectives to the verb, for example, the verb ਕਰ (do) or ਹੋ (be) can be added to nouns or adjectives, Table 4 and Table 5, respectively. Table 6, contains the verb morphology for auxiliary verbs.

TABLE 3
SOME EXAMPLES OF VERB MORPHOLOGY FOR SIMPLE WORDS

Morphology	Person	SG/PL	UNL Attributes	Example
ਚੁੱਕਿਆ ਸੀ cukkiā sī	1	SG	@past @complete, 1per, @sg	ਮੈਂ ਖਾਣਾ ਖਾ ਚੁੱਕਿਆ ਸੀ maiṃ khāṇā khā cukkiā sī 'I had eaten food'
	2	SG	@past, @complete, 2per, @sg	ਤੂੰ ਖਾਣਾ ਖਾ ਚੁੱਕਿਆ ਸੀ tūṃ khāṇā khā cukkiā sī 'You had eaten food'
	3	SG	@past @complete, 3per, @sg	ਮੁੰਡਾ ਖਾਣਾ ਖਾ ਚੁੱਕਿਆ ਸੀ muṇḍā khāṇā khā cukkiā sī 'Boy had eaten food'
ਚੁੱਕੇ ਸੀ cukkē sī	1	PL	@past, @complete, 1per, @pl	ਅਸੀਂ ਖਾਣਾ ਖਾ ਚੁੱਕੇ ਸੀ asīṃ khāṇā khā cukkē sī 'We had eaten food'
	2	PL	@past, @complete, 2per, @pl	ਤੁਸੀਂ ਖਾਣਾ ਖਾ ਚੁੱਕੇ ਸੀ tusīṃ khāṇā khā cukkē sī 'You had eaten food'

TABLE 4
SOME EXAMPLES OF CONJUNCT VERB MORPHOLOGY FOR ਕਰ (do) TYPE OF VERBS

UNL Attributes	Punjabi Morphology
@custom, @present, @sg, @male,@link	ਕਰਦਾ ਹੈ karadā hai
@past, @sg, @male, @link	ਕੀਤਾ kītā
@future, @sg, @male, @link	ਕਰੇਗਾ karēgā
@custom, @past, @sg, @male,@link	ਕਰਦਾ ਸੀ karadā sī

TABLE 5
SOME EXAMPLES OF CONJUNCT VERB MORPHOLOGY FOR ਹੋ (be) TYPE OF VERBS

U NL Attributes	Punjabi Morphology
@custom, @present, @sg, @male,@lnk	ਹੁੰਦਾ ਹੈ hundā hai
@past, @sg, @male, @lnk	ਹੋਇਆ Hōiā
@future, @sg, @male, @lnk	ਹੋਵੇਗਾ hundā sī

When a sentence does not have a main verb in it, then auxiliary verb acts like the main verb. These sentences have a predicative adjective in it and they require a verb terminator in the end. They have *aoj* (defines a thing that is in a state or has an attribute) relation between UW1 and UW2; and @pred attribute is used with the UW1 for its UNL representation [11]. Table 6, provides some possible cases of auxiliary verb morphology.

TABLE 6
SOME EXAMPLES OF AUXILIARY VERB MORPHOLOGY

Tense	Person	SG/PL	Respect/ Polite	Verb Terminator	U NL Attributes	Example
Past	1	SG		ਸੀ sī	@past, 1per,@sg	ਮੁੰਡਾ ਸੌਹਣਾ ਸੀ muṇḍā sauhṇā sī
	1	PL		ਸਾਂ sāṃ	@past, 1per, @pl	ਅਸੀਂ ਸੌਹਣੇ ਸਾਂ asī sauhṇē sāṃ
	2	SG		ਸੀ Sī	@past, 2per, @sg	ਤੂੰ ਸੌਹਣਾ ਸੀ tūṃ sauhṇā sī

5 Implementation of Punjabi EnConverter and Punjabi DeConverter

Punjabi EnConverter and DeConverter have been implemented in this work with the help of rules discussed in previous section. Architecture of Punjabi EnConverter can be divided into six phases. These phases include: processing of input Punjabi sentence by Punjabi shallow parser, creation of linked list of nodes on the basis of output of shallow parser, extraction of UWs from lexicon, processing of nodes for case markers, handling of unknown words and generation of UNL representation of the input sentence.

Punjabi EnConverter developed in this work uses Punjabi Shallow Parser (by IIIT Hyderabad, India) for processing the input Punjabi sentence [15].

5.1 Punjabi DeConverter

Punjabi DeConverter takes UNL representation as input and generates equivalent Punjabi language sentence as output with the help of various database files, including Punjabi-UW dictionary, Function Word Insertion Database and Generation Rules database. The generation process can be divided into four phases. These phases include: parsing of UNL file for the creation of node-net, morphological generation of lexical words, function words insertion, and syntax planning [3].

For example, in order to generate Punjabi Sentence from UNL representation (2) given in Section 1, a machine has to generate the form ਖੇਡਿਆ (played) from root verb ਖੇਡ (play) by using the information of tense (past), number (singular), and gender (masculine) associated with the verb. For this example sentence, case markers ਨੇ and ਵਿੱਚ need to be inserted after the subject ਸੁੱਖਵਿੰਦਰ 'Sukwinder' and the place ਬਾਗ 'garden', respectively. All the words are finally arranged by the system to construct equivalent Punjabi sentence: 'ਸੁੱਖਵਿੰਦਰ ਨੇ ਬਾਗ ਵਿੱਚ ਫੁਟਬਾਲ ਖੇਡਿਆ '. (3)

In this work, a Punjabi-UW dictionary consisting of 1,15,000 words has been created. There are 3000 analysis rules and 3500 generation rules that have also been implemented in this work.

6 Results and Discussion

We have evaluated performance of proposed EnConverter *vis-a-vis* the performance of a gold standard EnConverter. The EnConverter available at Spanish UNL language server has been taken as gold standard in the present work. We have manually translated English sentences given at Spanish language server into equivalent Punjabi sentences and then inputted those equivalent Punjabi sentences to the proposed Punjabi-UNL EnConverter system. We have compared UNL expressions generated by our system with the UNL expressions generated by Spanish UNL language server. We consider the two UNL representations matched with each other if the relations, including associated UWs, present in the representation are same [15]. Proposed EnConversion system has been tested on ninety one Punjabi sentences. It has been seen that the system successfully handles the resolution of UNL relations and generation of attributes for these sentences. Out of the ninety one sentences, outputs of the two systems match for eighty seven sentences. Results of this evaluation on the basis of UNL relations are given in Figure 1.

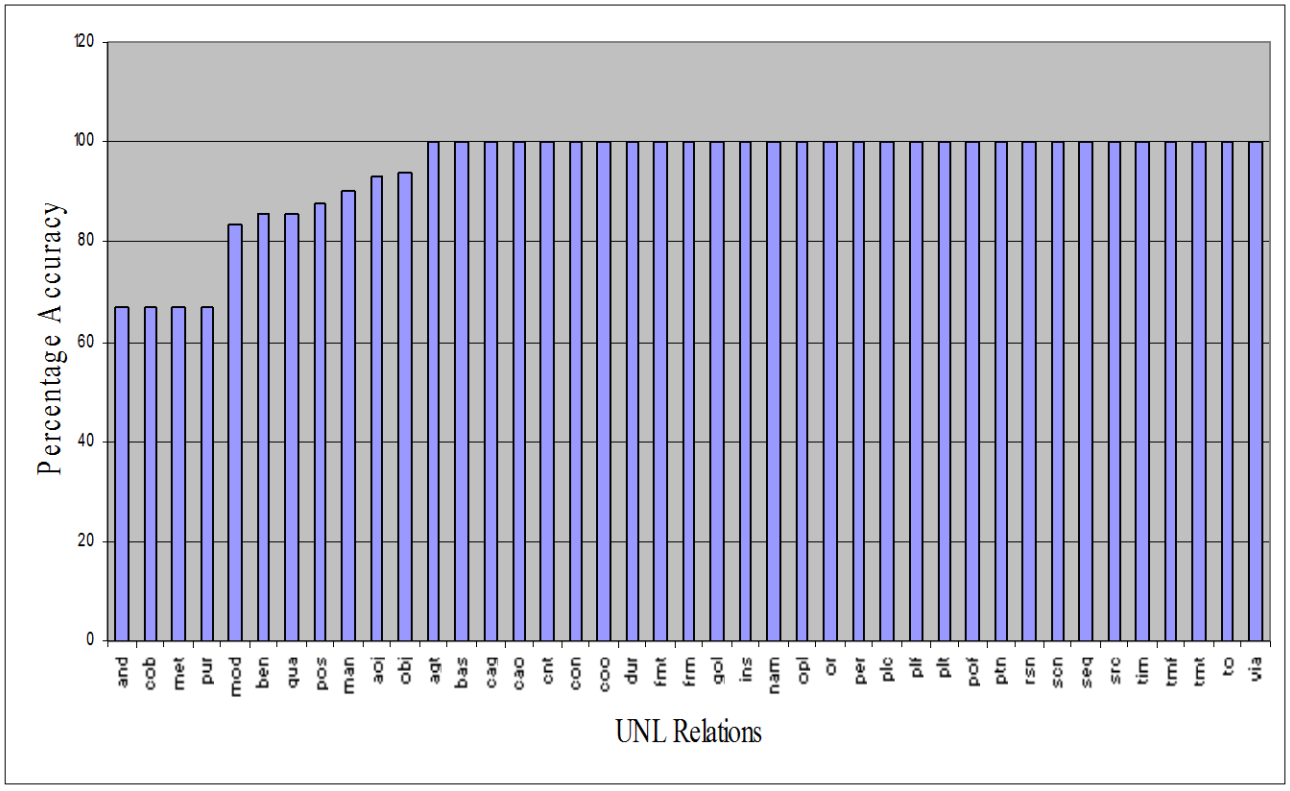

Fig. 1. Results of evaluation of Punjabi EnConveter for UNL relations

We have also evaluated the proposed Punjabi DeConverter by inputting the UNL representation generated by the Punjabi EnConverter to In this work, the evaluation methodology proposed by Singh et al. 2007 has also been used. Following them, a 4-point scale has been considered for the evaluation. In order to evaluate proposed UNL based MT system for Punjabi language, one hundred sentences have been considered. Some of these sentences are taken from Spanish UNL language server. Other sentences are taken from agricultural domain threads developed by IIT Bombay. Agriculture domain threads are taken from the script of questions-answers between farmers and agriculture experts. It has Hindi sentences and their equivalent UNL representation. These Hindi sentences have manually been translated to Punjabi and then processed in Punjabi EnConverter for their equivalent UNL representation. This UNL is then processed by Punjabi DeConverter. We have measured the fluency and adequacy of the machine generated Punjabi sentences with respect to the reference Punjabi sentences given to the proposed EnConverter. In this way, we have indirectly measured the faithfulness of the Punjabi EnConverter as well. Fluency score of the proposed system is 2.63 on the scale of 4 (*i.e.*, 65.75%) and adequacy score of the proposed system is 2.51 on the scale of 4 (*i.e.*, 63%).

REFERENCES

[1] Adly, N & S Alansary," Evaluation of Arabic Machine Translation System based on the Universal Networking Language," *Natural Language Processing and Information Systems*, LNCS (5723/2010): 243-257, 2009

[2] Ali, Yousuf, Jugal Krishna Das, S. M. Abdullah Al-Mamun, Md Ershadul & H. Choudhury," Specific Features of a Converter of Web Documents from Bengali to Universal Networking Language," *International Conference on Computer and Communication Engineering*, Kuala Lumpur, Malaysia, 13-15 May. 2008.

[3] Bhatia, Parteek & R. K. Sharma," Role of Punjabi morphology in designing Punjabi-UNL EnConverter," *International Conference on Advances in Computing, Communication and Control*, Bombay. 2009.

[4] Chakrabarti, Debasri & Pushpak Bhattacharyya, "Syntactic Alternations of Hindi Verbs with Reference to Morphological Paradigm," *Language Engineering Conference* (LEC2002), Hyderabad, India, December. 2002.

[5] Chakrabarti, Debasri, Vaijayanthi Sarma & Pushpak Bhattacharyya, "Hindi Verb Knowledge Base and Noun Incorporation in Hindi," *3rd Global Wordnet Conference (GWC 06)*, Jeju Island, Korea, January. 2006.

[6] Chandar, Duni. *Punjabi Bhasha Da Viyakaran (Punjabi)*. Chandigarh: *Panjab University Publication Bureau*, India. 2002.

[7] Dave, Shachi & Pushpak Bhattacharyya,"Knowledge extraction from Hindi text," *Journal of Institution of Electronic and Telecommunication Engineers*. 18 (4). 2001.

[8] Dave, Shachi, Jignashu Parikh & Pushpak Bhattacharyya, "Interlingua Based English Hindi Machine Translation and Language Divergence," *Journal of Machine Translation (JMT)*. 16(4): 251-304. 2001.

[9] Dey, K & Pushpak Bhattacharyya, "Universal Networking Language Based Analysis and Generation of Bengali Case Structure Constructs," *International conference on the Convergence of Knowledge, Culture, Language and Information Technologies*, Egypt. 2003.

it. Output of Punjabi DeConverter is compared with input Punjabi sentence given to Punjabi EnConverter.

7 CONCLUSION

In this paper, an analysis has been carried out for Punjabi verbs on the basis of root verbs ending on a vowel and GNPTAM information. We have created approximately one thousand verb morphology rules on the basis of verb paradigms. These rules are implemented in Punjabi EnConverter and DeConverter and a UNL based MT system for Punjabi has been proposed. The proposed EnConverter has been tested for its performance *vis-a-vis* an EnConverter available in public domain on the Spanish Language Server. The Punjabi DeConverter has also been tested on the output generated by the Punjabi EnConverter. Fluency score achieved for the proposed system is 2.63 while the adequacy score is 2.51 on the scale of 4. These scores can further be improved by refining the rules, proposing new rules and introducing word sense disambiguation module at EnConverter level.

[10] Dhanabalan, T, K. Saravanan & T. V. Geetha., "Tamil to UNL EnConverter," *International Conference on Universal Knowledge and Language*, Goa, India. 2002.

[11] Dwivedi Vijay. Generation of Hindi from Universal Networking Language. *IIT Bombay M Tech Thesis.* 2002.

[12] Gill, Harjeet S., A Henry & Gleason. A Reference Grammar of Punjabi. Patiala: *Publication Bureau, Punjabi University*. 1986.

[13] Gill, Mandeep Singh. Development of a Punjabi Grammar Checker. *Punjabi University, Patiala, PhD Thesis.* 2008.

[14] Giri, L. Semantic Net Like Knowledge Structure Generation from Natural Languages. *IIT Bombay B Tech Dissertation.* 2000.

[15] Jain, Manoj & Om P. Damani, "English to UNL (Interlingua) EnConversion." *Second Conference on Language and Technology.* 2009.

[16] Joshi, Shiv Sharma. Punjabi Bhasha, "Viyakaran Ate Bantar (Punjabi)," Patiala: *Publication Bureau, Punjabi University.* 2000.

[17] Nalawade, Ajay. Natural Language Generation from Universal Networking Language. *IIT Bombay M Tech Thesis.* 2007.

[18] Puar, Joginder S. The Punjabi verb form and function. Patiala: *Publication Bureau, Punjabi University*. 1990.

[19] Shrivastava, Manish, Nitin Agrawal, Bibhuti Mohapatra, Smriti Singh & Pushpak Bhattacharya, "Morphology Based Natural Language Processing tools for Indian Languages," *Workshop on Morphology, IIT Bombay*. 2005.

[20] Singh, Harkirat. Punjabi Shabad-Roop te Shabab-Jor Kosh (Punjabi). Patiala: *Shabad Jor Cell, Punjabi University*, India. 1988.

[21] Singh, Prem Parkash. Morphology and Punjabi Word-formation (Punjabi). Patiala: *Madan Publications*. 2002.

[22] Singh, Smriti, Kuhoo Gupta, Manish Shrivastava & Pushpak Bhattacharyya. "Morphological Richness Offsets Resource Poverty- an Experience in Building a POS Tagger for Hindi," *COLING/ACL-2006*, Sydney, Australia, July. 2006.

[23] Singh, Smriti, Mrugank Dalal, Vishal Vachhani, Pushpak Bhattacharyya & Om P. Damani," Hindi Generation from Interlingua," *Machine Translation Summit*, Copenhagen. 2007.

[24] Uchida, Hiroshi, Zhu Meiying & Tarcisio Della Senta, "The UNL, A Gift for a Millennium. Japan," *Institute of Advanced Studies*, The United Nations University. 1999.

[25] Universal Networking Language (UNL) Specifications, Version 2005. http://www.undl.org/unlsys/unl/unl2005/ (accessed 12 August 2010).