RESEARCH ARTICLE OPEN ACCESS

Machine Translation: An Analytical Study

Baljinder Kaur, Brahmaleen Kaur Sidhu
Scholar, Assistant Professor, Department of Computer Engineering, Punjabi University, Patiala

ABSTRACT
Machine translation is one of the oldest subfields of artificial intelligence research, the recent shift towards large-scale empirical techniques has led to very significant improvements in translation quality. The field of Machine Translation is responsible for the conversion of data from one natural language to other without human intervention.
India, a potpourri of different cultures, religions, and beliefs, is home to not just one or two languages but to an uncountable number of different lingual families. More than 90 % of the population is speaking the languages belonging to two major language families - Indo Aryan and Dravidian. This research paper focuses on major approaches of MT, the developed as well as developing tools and the challenges faced during the translations and the need for an accurate MT system.

I. INTRODUCTION

India is rich in languages where the spoken language changes after every 50 miles. India has 22 officially recognized languages and around 2000 dialects [1]. A number of translation systems have been developed and many systems are still developing for these languages Because of the challenges faced in the field of machine translation (MT), a completely automated system has not been developed till now.

MT is a sub-field of computational linguistics that investigates the use of software to translate text or speech from one natural language to another [2]. The aim of machine translation systems is to produce the possible translation without taking any help from human. The machine translations systems should work independently and produce all the possible results of their own without any human intervention. Basically every machine translation system is a combination of programs, automated dictionaries and grammars were programs are responsible for translations [3]. MT consists of a number of phases. In the first phase, MT performs simple substitution of words in one human or natural language to another human language. In the next phase MT approaches like rule based MT, corpus based MT, statistical based MT etc are used for the translations. These approaches has improved the quality of translations and has also provided better results in the translation of idioms. Some approaches are discussed in the following sections.

II. MACHINE TRANSLATION APPROACHES AND TOOLS

MT includes a number of approaches and based upon these approaches a number of tools have been developed. Both approaches and its tools are explained below:

1. RULE BASED APPROACH

This approach associates the structure of the source language with the structure of the target language, in order to preserve the meaning of the input sentence. This approach is said to be suitable for the languages with different word orders. For example majority of the Indian languages are of the order S O V , on the other hand English is of the order S V O. For these type of language pairs this approach is widely used. A Rule-Based Machine Translation (RBMT) system generally consists of grammar rules, a bilingual or multilingual lexicon, and software programs to process the rules. Nevertheless, building RBMT systems entails a huge human effort to code all of the linguistic resources, such as source side part-of-speech taggers and syntactic parsers, bilingual dictionaries, source to target transliteration, TL morphological generator, structural transfer, and reordering rules. Rules play a major role in various stages of translation, such as syntactic processing, semantic interpretation, and contextual processing of language. [4]. A number of tools have been developed based on this approach, some of them are discussed below.

Rule based Punjabi to Hindi MT for legal documents translates simple sentences in legal domain from

Punjabi to English. The need of the system arises from the translations of the legal documents transferred from District Courts of Punjab to the High Court in Chandigarh. The FIR which is written in Punjabi is translated to English before presenting it to the high court. The challenges faced by the developer lie with major problems, mainly related to the non-availability of lexical resources, spelling variations, part of speech tagging, transliteration, named entity recognition and collocations. Rules are also developed for combining phrases to form a target language sentence. Accuracy of the system was reported as 69.50% on the basis of intelligibility test and 70.54% on the basis of accuracy test performed on the results [5].

Web Based Hindi to Punjabi Machine Translation System is developed using rule based. The system includes lexicon based translation, transliteration and continuously improving the system through machine learning module. This system also takes care of basic word sense disambiguation. Approach used classifies the system into three major phases, Preprocessor, Tokenizer and Translation Engine. The System accuracy was proclaimed to be 95% [6].

Rule Based Approach was used for Machine Translation System for Related Languages, Punjabi to Hindi. System architecture consists of components like parsing, tokenization, rule application to DB, character mapping and target language generation. The accuracy of the system depends upon the rules that are the simple word for word translation rules. The major inaccuracies in the direct mapping of characters are due to poor word choice for confusing words [7].

Machine Translation of Idioms from English to Hindi was developed using rule based approach. The sentences containing idioms were translated using Google Translator. The system is implemented in Asp.net at front end and MS SQL 2008server at back end. A English to Hindi Idiom Translator System Interface is created whose object will accept a string in English language and returns its corresponding Hindi string. The accuracy of system was reported to be 70%. Analysis of results shows that the problems of bad translation are due to errors of different categories like-irrelevant idioms, grammar agreement, part of speech etc [8].

MaTra is an English to Indian languages Human-Assisted translation system based on a transfer approach using a frame-like structured representation that resolves the ambiguities using rule-based and heuristics approaches. MaTra is an innovative system, which provides an intuitive GUI, where the user can inspect the analysis of the system and can provide disambiguation information to produce a single correct translation [4].

ANGLABHARTI aims at developing a system where majority of the work is done by the machine and only about 10% of the task is left for human post-editing. Anglabharti is a pattern directed rule based system with context free grammar like structure for English (source language). It generates a `pseudo-target' (Pseudo-Interlingua) applicable to a group of Indian languages (target languages) such as Indo-Aryan family (Hindi, Bangla, samiya, Punjabi, Marathi, Oriya, Gujarati etc), Dravidian family (Tamil, Telugu, Kannada & Malayalam) and others [9].

2. DIRECT MACHINE TRANSLATION APPROACH

In the direct translation method, the source language text is analyzed structurally up to the morphological level and is designed for a specific source and target language pair. The performance of a direct MT system depends on the quality and quantity of the source-target language dictionaries, morphological analysis (a method for exploring all possible solutions of a problem), text processing software, and word-by-word translation with minor grammatical adjustments on word order and morphology [4]. A number of tools have been developed based on this approach, some of them are discussed below.

Punjabi to Hindi machine translation system architecture consist of stages like Text normalization, Tokenization, Translation Engine and Target Language Generation. Systems accuracy was claimed to be 90.67%, much higher than others . The accuracy of the translation achieved by the system justifies the hypothesis that word-for-word translation might also be a solution for language pair of Punjabi and Hindi. The major inaccuracies in the direct translation were due to poor word choice for ambiguous words and some corrections regarding post positions. The lack of information in glossaries and dictionaries sometimes causes an unnecessary translation error [10].

Direct Machine Translation System from Punjabi to Hindi was developed for newspapers headlines domain. The system architecture consists of Preprocessing Stage, Translation Engine and Transliteration Engine. Average ratings of %age for the system depends upon the intelligibility test and accuracy test. From the accuracy analysis total number of accurate sentence were calculated and then their %age was claimed to be 97%. Error rating of the system depends upon the wrongly translated word or expression, Un-translated words and Wrong choice of

words. The %age of error rating was depicted from total sentences which came out to be 3% [11].

3. INTERLINGUAL APPROACH

In this approach, the source language, is transformed into an interlingua, i.e., an abstract language-independent representation. The target language is then generated from the interlingua [12]. The advantages of this approach is that it requires fewer components in order to relate each source language to each target language, it takes fewer components to add a new language, it supports paraphrases of the input in the original language, it allows both the analyzers and generators to be written by monolingual system developers, and it handles languages that are very different from each other. The disadvantage is that the definition of an interlingual is difficult and maybe even impossible for a wider domain. A number of tools have been developed based on this approach, one of them is discussed below.

Language-to-Interlanguage-to-Language System Based on UNL was an interlingua-based human-aided multilingual machine translation web service. It was expected to provide end-to-end high-quality translations through semi-automatic analysis of the source text into the Universal Networking Language (UNL) and fully-automatic generation from the resulting UNL document into several different target languages. The basic components of LILY MTS involves Language Resources, Grammars, Dictionaries and Tools and Engines [13].

4. TRANSFER APPRAOCH

Transfer approach has three stages involved. In the first stage, source language (SL) texts are converted into abstract SL oriented representations. In the second stage, SL oriented representations are converted into equivalent target language (TL) oriented representations. Final texts are generated in the third stage. In transfer approach complete resolution of ambiguities of SL text is not required, but only the ambiguities inherent in the language itself are tackled [4]. Three types of dictionaries are required: SL dictionaries, TL dictionaries and a bilingual transfer dictionary. Transfer systems have separate grammars for SL analysis, TL analysis and for the transformation of SL structures into equivalent TL forms. It is possible with this translation strategy, a fairly high quality translations can be obtained with accuracy in the region of 90% which is highly dependent on the language pair in question. A number of tools have been developed based on this approach, one of them is discussed below.

Experiments with a Hindi-to-English Transfer based MT System under a Miserly Data Scenario was developed using Transfer approach. The researchers compare the performance of the Xfer approach with two corpus-based approaches, Statistical MT (SMT) and Example-based MT (EBMT) under the limited data scenario. The results indicate that the Xfer system significantly outperforms both EBMT and SMT in this scenario. Results also indicate that automatically learned transfer rules are effective in improving translation performance, compared with a baseline word-to-word translation version of the system. The system called Hindi to English Transfer MTS was developed using this approach which consist of components like Morphology, Grammar, Lexical Resources and Runtime Configuration [14].

5.STATISTICAL MACHINE TRANSLATION

Statistical machine translation is a data-oriented statistical framework for translating text from one natural language to another based on the knowledge and statistical models extracted from bilingual corpora. In statistical-based MT, bilingual or multilingual textual corpora of the source and target language or languages are required [4]. During translation, the collected statistical information is used to find the best translation for the input sentences, and this translation step is called the decoding process. There are three different statistical approaches in MT, Word-based Translation, Phrase-based Translation, and Hierarchical phrase based model.

- Word Based Translation: The words in an input sentence are translated word by word individually, and these words finally are arranged in a specific way to get the target sentence. The alignment between the words in the input and output sentences normally follows certain patterns in word based translation
- Phrase Based Translation: In this approach each source and target sentence is divided into separate phrases instead of words before translation. The alignment between the phrases in the input and output sentences normally follows certain patterns, which is very similar to word based translation.
- Hierarchical Phrase Based model: By considering the drawback of previous two methods, the hierarchical phrase based model was developed. The advantage of this approach is that hierarchical phrases have recursive structures instead of simple phrases. This higher level of abstraction approach further improved the accuracy of the SMT system.

A number of tools have been developed based on this approach, one of them is discussed below.

Google Translator is a multilingual statistical machine-translation service used to translate written text from one language into another. When Google Translate generates a translation, it looks for patterns in hundreds of millions of documents to help decide on the best translation [15]. By detecting patterns in documents that have already been translated by human translators, Google Translate makes intelligent guesses (AI) as to what an appropriate translation should be. Google Translate does not apply grammatical rules, since its algorithms are based on statistical analysis rather than traditional rule-based analysis. It is based on a method called statistical machine translation. Google does not translate from one language to another (L1 → L2), but often translates first to English and then to the target language (L1 → EN → L2).

6. EXAMPLE-BASED APPRAOCH

EBMT is characterized by its use of bilingual corpus with parallel texts as its main knowledge, in which translation by analogy is the main idea. There are four tasks in EBMT: example acquisition, example base and management, example application and synthesis. EBMT uses the bi-text as its primary data source, in which pre processing the data is optional and if the input is in the example set, the same translation is to occur. The example-based machine translation (EBMT) approach is based on analogical reasoning between two translations [4]. EBMT consists of components like, Example Acquisition, Similarity Measure, Alignment Algorithm, Evaluation of Alignment Result, Example Base, EB Management, EB as a Language Model, Example Application, Sentence Synthesis and Smoothing. A number of tools have been developed based on this approach, one of them is discussed below.

System called A Pure Example Based approach for English to Hindi Sentence MT Systems was developed using EBMT approach. It is a corpus based machine translation, which requires parallel-aligned 3 machine-readable corpora's. The already translated example serves as knowledge to the system for example based machine translation. This approach derives the information from the corpora for analysis, transfer and generation of translation. The adapted segments are recombined according to sentence structure of the source and target language. The training corpus is the parallel database containing 677 sentences. The corpus generated is not preprocessed. The examples contained are newspaper headlines. The training module forms the matrix; this matrix is used for matching in translation phase. Once trained the matrix remains same till the database is modified. The training matrix gives the number of occurrence of the word in the corpus [16].

7. PRINCIPLE BASED MT

Principle-Based Machine Translation (PBMT) Systems employ parsing methods based on the Principles & Parameters Theory of Chomsky's Generative Grammar [4]. The parser generates a detailed syntactic structure that contains lexical, phrasal, grammatical, and thematic information. It also focuses on robustness, language-neutral representations, and deep linguistic analyses. In the PBMT, the grammar is thought of as a set of language-independent, interactive well-formed principles and a set of language-dependent parameters. Thus, for a system that uses n languages, one must have n parameter modules and a principles module.

8. KNOWLEDGE-BASED MT

Knowledge-Based Machine Translation (KBMT) is characterized by a heavy emphasis on functionally complete understanding of the source text prior to the translation into the target text [4]. KBMT does not require total understanding, but assumes that an interpretation engine can achieve successful translation into several languages. KBMT is implemented on the Interlingua architecture; it differs from other interlingual techniques by the depth with which it analyzes the SL and its reliance on explicit knowledge of the world. KBMT must be supported by world knowledge and by linguistic semantic knowledge about meanings of words and their combinations. Thus, a specific language is needed to represent the meaning of languages. Once the SL is analyzed, it will run through the augmenter. It is the knowledge base that converts the source representation into an appropriate target representation before synthesizing into the target sentence. Thus, KBMT system provides high quality translations.

9. EMPIRICAL MT APPROACH

Empirical approach is the emerging one that uses large amount of raw data in the form of parallel corpora. Parallel corpora is a set of raw data consisting of text in source language and its corresponding translations in target language. Example-based MT, analogy-based MT, memory-based MT, and case-based MT are the techniques that use empirical approach. Basically all these techniques use a corpus or a database of translated examples.

10. CORPUS BASED MACHINE TRANSLATION

Corpus-based machine translation (CBMT) is generated on the analysis of bilingual text corpora. Corpus can be defined as a data on a specific subject. CBMT approach is based upon the conviction that there are no pre-established solutions to translation, but most possible solutions can be found in texts already translated by professionals. In other words, a large portion of a translator's competence is encoded in the language equivalencies that can be found in already translated texts. The corpus based machine translation is based upon concepts like Bilingual corpora, empirical bilingual corpus and Alignment of translated corpus. Moreover, a bilingual corpus is richer in information about the language than a monolingual corpus, since it provides situational equivalency information on the possibilities of the language system when in contact with a different linguistic system [17].

11. DICTIONARY BASED MT

Machine translation can use a method based on dictionary entries, which means that the words will be translated as they are by a dictionary .One of the major factors that can potentially degrade the effectiveness of dictionary-based cross-language information retrieval is the ambiguity in translating query words [18]. The Dictionary based method consists of dictionaries and multilingual thesauri.

12. HYBRID BASED TRANSLATION

By taking the advantage of both statistical and rule-based translation methodologies, a new approach was developed, called hybrid-based approach. The systems developed using this approach have proven to be of better efficiency in the area of MT systems [4].The hybrid approach can be used in a number of different ways. In some cases, translations are performed in the first stage using a rule-based approach followed by adjusting or correcting the output using statistical information. In the other way, rules are used to pre-process the input data as well as post-process the statistical output of a statistical-based translation system. This technique is better than the former approaches and has more power, flexibility, and control in translation. A number of tools have been developed based on this approach, some of them are discussed below.

ANUBHARTI System uses a hybrid approach called HEBMT (Hybrid Example Based Machine Translation System), which combines the essentials of the pattern-based approach with those of the example-based approach. It makes use of abstracted example-base instead of raw example-base. When an input sentence is fed for translation, after morphological analysis, it is passed on to the finite state machine which identifies the syntactic units like verb phrase, noun phrase etc of the sentence. Then, an appropriate partition of the abstracted example-base is searched based on the number and the type of syntactic units. If such a partition exists in the example-base, a distance matrix is computed for the input sentence and the example sentences of that partition, and the example with the minimum distance is invoked for target language translation. If such a partition does not exist, input sentence is entered as an example in the example-base and user/developer is asked to enter the target language pattern [15].

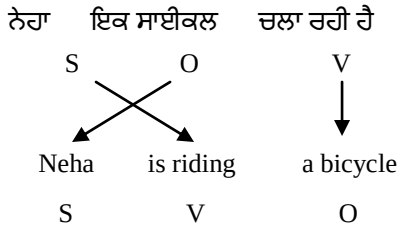
System called Hybrid Approach for English to Punjabi Translation System for News Paper Headlines in a Specific Domain was developed using hybrid approach. In the proposed system hybrid approach is used to translate the news headline written in English text into its equivalent Punjabi text. Hybrid approach in the proposed system is a combination of Rule based approach; Example based approach and direct mapping approach. Proposed system uses a corpus for translation purpose. Corpus consist of rules , examples , nouns , verbs and other entities stored within database and within the programming logic [19].

Web Based English to Punjabi MTS for News Headlines was developed using hybrid approach for translating English sentences into Punjabi was used. The approach is a combination of direct, example-based (EBMT) and Rule based approach. Using Direct Approach direct mapping of the source language with the target language is done. Using Example-Based Machine Translation (EBMT) the examples of existing translation were used and reused for new translations. Following steps are performed under EBMT. During this approach example matching, alignment and final recombination is done to make sure that the reusable parts identified during alignment are put together in a legitimate way. The system was successfully tested on 300 news headlines under the domain. The accuracy of the system has been found as 81.67 percent [20].

III. CHALLENGES IN MACHINE TRANSLATION

A number of challenges are faced in the field of MT, some of them are discussed here:

- Two languages which have completely different structures, for example:. English has SVO (Subject- Verb- Object) structure and Punjabi has SOV (Subject- Object- Verb) structure, make the translation process tedious.



- One of the biggest challenge faced in the development of an MTS is Multi Word Expressions (MWEs). Multiword Expressions are idiosyncratic word usages of a language which often have non compositional meaning . It generally includes things like Collocations(an expression consisting of two or more words that correspond to some conventional way of saying things.), Idioms(a group of words whose meaning cannot be predicted from the meanings of the constituent), Phrase (a group of related words within a sentence or we can say it is grammatical unit at a certain level between a word and a clause), Proverbs. Since all mentioned things above grow on changing and due to the lack of proper updated material, the translation quality is affected.

- Ambiguity has an adverse effect on translation process. Ambiguity means capable of being understood in more than one ways or a word having a variety of senses in a sentence and WSD(word–sense-disambiguation) governs the process of identifying which sense of a word (i.e. meaning) is used in a sentence, when the word has multiple meanings. When a word has more than one meaning, it is said to be lexically ambiguous. When a phrase or sentence have more than one structure it is said to be structurally ambiguous.

- Another one of the major challenges is Regional Dialects. Since, India is rich in languages where the spoken language changes after every 50 miles. India has 22 officially recognized languages and around 2000 dialects have been identified in India. Systems have been developed for regional languages but for regional dialects no system has been developed yet. Same translation rules cannot be applied to different dialects of same language. Thus a big challenge to deal with. System needs to be developed for these regional dialects.

- All words in one language may not have equivalents in other language. Thus a challenge to deal with.

- There are certain words in some languages where transliteration is required before translation .This is one of the biggest challenge faced. Some MTS fail to detect the noun, thus instead of transliterating, it directly translated. An example from Google Translator.
  (E) Where is rose?
  (H) गुलाब कहाँ है ?
  (P) ਗੁਲਾਬ ਕਿਥੇ ਹੈ ?

- The natural language is open and keeps on changing from time to time. So complete automatic simulation of natural language is almost impossible. Thus in the lack of the updated material a complete automatic simulation of the human language is nearly impossible.

IV. RESULTS AND DISCUSSIONS:

For studying the process of MTS, 785 sentences in English, Hindi and Punjabi were tested on 7 different tools. Majority of the data was retrieved from various online resources. According to the availability of the translation functions of the tools, the corpus was divided into following languages pairs :

E-H(English – Hindi) ; H-E(Hindi - English)
E-P(English - Punjabi) ; P-E(Punjabi - English)
H-P(Hindi - Punjabi) ; P-H(Punjabi - Hindi)

The testing was done on sentence level. The results were categorized as satisfactory results and unsatisfactory results.

$$\text{satisfactory result\%} = \frac{\text{number of satisfactory statemnts}}{\text{total number of satements}}$$

$$\text{unsatisfactory result\%} = \frac{\text{number of unsatisfactory statemnts}}{\text{total number of satements}}$$

The testing results of these tools are listed below in tables:

Google Translator	E-H	H-E	E-P	P-E	H-P	P-H
Total number of statements tested	87	50	64	25	51	29
Satisfactory Results	07	26	03	00	01	02
Unsatisfactor-y Results	78	24	61	25	50	27

Table 1: Google Translator results

Babylon	E-H	H-E
Total number of statements tested	63	50
Satisfactory Results	01	06
Unsatisfactory Results	62	44

Table 2: Babylon results

ImTranslator	E-H	H-E
Total number of statements tested	53	52
Satisfactory Results	06	22
Unsatisfactory Results	47	30

Table 3: ImTranslator results

Learn Punjabi	H-P	P-H
Total number of statements tested	51	49
Satisfactory Results	13	26
Unsatisfactory Results	38	23

Table 4: Learn Punjabi results

Anuvadaksh	E-H
Total number of statements tested	31
Satisfactory Results	01
Unsatisfactory Results	30

Table 5: Anuvadaksh results

Sampark	H-P	P-H
Total number of statements tested	51	48
Satisfactory Results	10	11
Unsatisfactory Results	41	37

Table 6: Sampark results

AnglaMT	E-P
Total number of statements tested	31
Satisfactory Results	00
Unsatisfactory Results	31

Table 7: AnglaMT results

132 English proverbs and idioms were tested on four tools namely, Google Translator (Tool 1), Babylon (Tool 2), ImTranslator (Tool 3) and Anuvadaksh (Tool 4). In majority of the cases the results were not satisfactory because the systems fails to detect the statement as an idiom or proverb instead, translate it as normal statement, resulting a unsatisfactory result. Every idiom and proverb has its own predefined output which never changes, neither can be modified. The testing results of all the four tools are bind up in one table shown below.

Tools/Results	Tool 1	Tool 2	Tool 3	Tool 4
Total number of statements tested	36	35	36	25
Satisfactory Results	11	0	1	0
Unsatisfactory Results	15	35	35	25

Table 8: Testing of proverbs and idioms

V. CONCLUSION

Through this research we can say that a large of Machine translation systems (MTS) have been developed and are still developing . In this research a number of approaches and systems based upon these approaches were studied. Through this research ,a number of challenges faced by the researchers in developing these machine translation systems were found. Due to the lack of data and lack of proper vocabulary and dictionaries , the tested systems and there results could not reach the expectations. Work should be done on ambiguity resolution and databases for better results. Every challenge that has been highlighted in this article, provides scope for future research.

REFERENCES

[1]. Retrieved from http: //www.webindia123.com / india/ people/language.htm on 11 april, 2014 .
[2]. Retrieved from http:/ /en.wikipedia.org/wiki/ Machine_ translation
[3]. Retrieved from http:// language.worldofcomputing.net/ category/ machine-translation.
[4]. Antony P. J. , "Machine Translation Approaches and Survey for Indian Languages", Computational Linguistics and Chinese Language Processing Vol. 18, No. 1, March 2013, pp. 47-78.
[5]. Kaur Kamaljeet, Lehal S. Gurpreet, "A Punjabi To English Machine Translation System for Legal Documents ", 2011, Internet: http://shodhganga.inflibnet .ac.in.
[6]. Goyal Vishal, Lehal S. Gurpreet, "Web Based Hindi to Punjabi Machine Translation System", Journal of Emerging Technologies in Web Intelligence, VOL. 2, NO. 2, MAY 2010.
[7]. Rani Sumita, Laxmi Vijay,"Rule Based Approach for Machine Translation System

for Related Languages : Punjabi to Hindi", International Journal of Computer Science And Technology, Vol. 4, Iss ue Spl - 1, Jan - March 2013 .

[8]. Gaule Monika and Josan S. Josan, "Machine Translation of Idioms from English to Hindi", International Journal Of Computational Engineering Research , Vol. 2 Issue. 6 , Oct 2012.

[9]. Retrieved from https:// sites.google.com/site/ profrmksinha / research-projects/machine-tran

[10]. Josan S. Gurpreet, Lehal S. Gurpreet, "A Punjabi to Hindi Machine Translation System" , Coling 2008: Companion volume – Posters and Demonstrations, pages 157–160 Manchester, August 2008.

[11]. Rani Sumita, Laxmi Vijay, " Direct Machine Translation System from Punjabi to Hindi for Newspapers headlines Domain", International journal for Computer and Technology, Vol 8, No 3, June 3 0 , 2 0 1 3.

[12]. Retrieved from http:// en.wikipedia.org/wiki/Interlingual _ machine_translation

[13]. S. Alansary, M. Nagi, " LILY: Language-to-Interlanguage - to-Language System Based on UNL", ", in Proceedings of the 13th Language Engineering Conference, (ESOLEC 2013), Ain Shams University, Cairo, Egypt, December 11 – 12, 2013.

[14]. http://en.wikipedia.org/wiki/Transfer-based_machine_ translation

[15].http://en.wikipedia.org/wiki/Google_Translate

[16]. http://en.wikipedia.org/wiki/Example-based_machine_ translation

[17]. Retrieved from http://www.translationjournal.net/ journal/19mt.htm on 14 may, 2014.

[18]. Dictionary Based Word Translation in CLIR Using Cohesion Method Mallamma V Reddy1 and M. Hanumanthappa Department of Computer Science and Applications, Bangalore University, Bangalore, INDIA

[19]. Singla Savita and Baghla Seema, " Hybrid Approach for English to Punjabi Translation System for News Paper Headlines in a Specific Domain" , International Journal of Engineering Research & Technology (IJERT) Vol. 2 Issue 11, November – 2013.

[20] Kaur Harjinder and Laxmi Vijay , "A Web Based English to Punjabi MT System for News Headlines ", International Journal of Advanced Research in Computer Science and Software Engineering, Volume 3, Issue 6, June 2013