

# MACHINE TRANSLATION DEVELOPMENT FOR INDIAN LANGUAGES AND ITS APPROACHES

Amruta Godase<sup>1</sup> and Sharvari Govilkar<sup>2</sup>

<sup>1</sup>Department of Information Technology (AI & Robotics), PIIT, Mumbai University, India

<sup>2</sup>Department of Computer Engineering, PIIT, Mumbai University, India

## ABSTRACT

*This paper presents a survey of Machine translation system for Indian Regional languages. Machine translation is one of the central areas of Natural language processing (NLP). Machine translation (henceforth referred as MT) is important for breaking the language barrier and facilitating inter-lingual communication. For a multilingual country like INDIA which is largest democratic country in whole world, there is a big requirement of automatic machine translation system. With the advent of Information Technology many documents and web pages are coming up in a local language so there is a large need of good MT systems to address all these issues in order to establish a proper communication between states and union governments to exchange information amongst the people of different states. This paper focuses on different Machine translation projects done in India along with their features and domain.*

## KEYWORDS

*Machine translation, computational linguistics, Indian Languages, Rule-based, Statistical, Empirical MT, Principle-based, Knowledge-based, Hybrid*

## 1.INTRODUCTION

Machine translation is one of the central areas of Natural language processing (NLP). Machine translation (MT) is important for breaking the language barrier among the people and for inter-lingual communication where translation is done from source language to target language. Many researchers, Institutions and organizations in India have started working on MT systems for Indian languages and have gained satisfactory results. The research scenario in India is relatively young and MT gained momentum in India with institutions like IIT Kanpur, IIT Mumbai, IIT Hyderabad, University of Hyderabad and CDAC Pune play a major role in developing these systems.

This paper is organized into 4 sections. Section 2 gives an introduction of MT & different approaches to build MT systems, Section 3 discuss major MT systems in India based on language with their features, domain etc. and finally we conclude the paper in the next section.

## 2.MACHINE TRANSLATION

Like translation done by human, MT does not simply substituting words but the application of complex linguistic knowledge; morphology, grammar, meaning all this things is taken into consideration. Generally, MT is classified into various categories: Direct based, rule-based, corpus based, statistical-based, hybrid-based, example-based, knowledge-based, principle-based, and online interactive based methods. At present, most of the MT related research is based on Rule based approaches because rule based is always extensible and maintainable.

### 1.Direct Translation

Direct Machine Translation is the one of the simplest machine translation approach in which a direct word to word translation is done with the help of a bilingual dictionary.

### 2.Rule Based Translation

A Rule-Based Machine Translation (RBMT) system consists of collection of various rules, called grammar rules, a bilingual lexicon or dictionary, and software programs to process the rules.

### 3.Interlingua Based Translation

In this approach, the translation consists of two stages, where the source Language (SL) is first converted in to the Interlingua (IL) form. The main advantage of Interlingua approach is that the analyzer and parser of SL is independent of the generator for the Target Language (TL) and this requires complete resolution of ambiguity in source language text.

### 4.Statistical-based Approach

Statistical machine translation (SMT) is a data-oriented statistical framework which is based on the knowledge and statistical models which are extracted from bilingual corpora. In this MT, bilingual or multilingual corpora of the languages are required. In SMT, a document is translated according to the probability distribution function which is indicated by  $p(e/f)$ . Finding the best translation is done by picking the highest probability, as shown in Equation 1.

$$e = \operatorname{argmax} p(e | f) = \operatorname{argmax} p(f | e) p(e) \dots \dots \dots (i)$$

### 5.Example-based translation

Basic idea of this MT is to reuse the examples of already existing translations. An example-based translation is uses a bilingual corpus as its main knowledge base and it is essentially translation by analogy.

### 6.Knowledge-Based MT

Knowledge-Based Machine Translation (KBMT) requires complete understanding of the source text prior to the translation into the target text. KBMT is implemented on the Interlingua

architecture. KBMT must be supported by world knowledge and by linguistic semantic knowledge about meanings of words and their combinations.

### 7.Principle-Based MT

Principle-Based Machine Translation (PBMT) Systems are based on the Principles & Parameters Theory of Chomsky’s Generative Grammar and which employs parsing method. In this, the parser generates a detailed syntactic structure which contains lexical, phrasal, grammatical information.

### 8.Online Interactive Systems

In this online interactive translation system, the user has authority to give suggestion for the correct translation. This approach is very useful, where the context of a word is not that much clear or unambiguous and where multiple possible meanings for a particular word.

### 9.Hybrid-based Translation

By taking the advantage of statistical MT and rule-based MT methodologies, a new approach was developed, which is termed as “hybrid-based approach”. The hybrid approach used in a number of different ways. Translations are performed in the first stage using a rule-based approach which is followed by adjusting or correcting the output using statistical information. Second way in which rules are used to pre-process the input data and for post-process the statistical output of a statistical-based translation system.

## 3.LITERATURE SURVEY

In this section we now look at some major English-Indian language MT project. The parameters we look at are: language pair(s), approaches used for handling problems, year of publication & application domain of each MT system. The scope of this paper is restricted to Hindi, Punjabi, Bengali, English and Marathi languages as source/target language.

#### A.Translation system for Hindi language as a Source or Target language

SR No	Machine Translation Systems	Year	Languages for translation	Domain / Application	Approach used	Observations
1.	ANUSAARAK MT [1][47]	1995	IL-IL	For translating children’s stories	Direct based	The focus is on language access between Indian languages. Works on the principles of Paninian Grammar (PG).

2.	MANTRA MT [2]	1997	English-Hindi	General	Transfer based	It uses TAG & XTAG. Uses tagger and light dependency analyzer for performing the analysis of I/P English text. It distributes a load on man and machine in equal way.
3.	MANTRA RAJYASABHA [2][48]	1999	English-Hindi	Office administration documents	Transfer based	System uses TAG & LTAG to represent a grammar. Can preserve the formatting of input word document. Currently working on Hindi to English and Hindi- Bengali.
4.	ANUBHARTI-I [2]	2003	Hindi-English	General	Hybrid	Combination of example based, corpus based & some grammatical analysis. It reduces the requirement of large example base and it depends on target language.
5.	ANUBHARTI-II [2]	2004	Hindi-English	General	Hybrid	It emulates human-learning process for storing knowledge from past experience to use it in future. Shallow chunker is used for fragmentation of input sentences.
6.	Hinglish MT System [49]	2004	Hindi-English	General	Example based	Based on Anubharti-II & Anglabharti-II. It produces satisfactory results in more than 90 cases. It performs shallow grammatical analysis.
7.	An English-Hindi Translation System	2002	English-Hindi	Weather narration	Transfer based	Translation modules consist

	[3]					of preprocessing and post processing of English tree. Also include generation of Hindi tree.
8.	UNL-based English-Hindi MT System [50]	2001	English-Hindi	General	Interlingua	Based on UNL grammar. Easy to add new language for translation.
9.	MaTra: A Practical Approach to Fully-Automatic Indicative English-Hindi Machine Translation [4][51]	2004-2006	English-Hindi	News, annual report, technical phrases	Transfer based	Based on MSIR. It uses transfer like structure representation & also uses heuristics to resolved ambiguities.
10.	GB Theory Based Hindi To English Translation System [5]	2009	Hindi-English	General	Example Based	System makes use of Government and Binding (GB) theory for undertaking translation. It consists of a parsing module and generating module.
10.	A Pure EBMT Approach for English to Hindi Sentence Translation System [6]	2014	English-Hindi	comparing sentence to extract the translation	Example based	This system uses parallel corpora for translation. It contains various modules such as similarity matrix, training matrix & tagging matrix.
11.	Interlingua based English-Hindi Machine Translation system and Language Divergence [7]	-	English-Hindi	General	Interlingua and Transfer based	Here UNL represents information sentence by sentence. Lexical-semantic divergence is handled in L-UW dictionary. Syntactic divergence is primarily tackled by analyzers. The amenability to generation is tested through Marathi Language.

## B. Translation system for Marathi language as a Source or Target language

SR No	Machine Translation Systems	Year	Languages for translation	Domain / Application	Approach used	Observations
1.	English to Devnagari Translation for UI Labels of Commercial web based Interactive Applications [8]	2011	English-Devnagari	Web based Applications	Hybrid	Used banking glossary available on the web site of RBI to create multilingual dictionary. For lexical analyzer rules are written by C languages. Used Bison tools for running system.
2.	Extending capabilities of English to Marathi Machine Translator [9]	2012	English-Marathi	General	Rule-based	Much functionality can be added for improving the performance of translation. It can be expanded by including spelling and grammatical checks, sentiment analysis modules.
3.	Rule based English to Marathi translation of Assertive sentence [10]	2013	English-Marathi	General	Rule-based	Database of set of rules maintained for mapping. Bilingual-Dictionary database plays very important role which is endless. Open-nlp tools performing different processes.
4.	A novel approach for Interlingual example-based translation of English to Marathi [11]	2014	English-Marathi	General	Hybrid	System is trained from bilingual parallel corpora. Sentence pairs contain sentence in one language with their translation into another. Uses parsing techniques.
5.	Transmuter: An approach to rule based English-Marathi machine Translation [12]	2014	English-Marathi	General	Rule-based	The focus is on grammar structure of target language that produces better & smoother translation. Lexicon is built for morphological & semantic properties.

## C.Translation system for Sanskrit language as a Source or Target language

SR No	Machine Translation Systems	Year	Languages for translation	Domain / Application	Approach used	Observations
1.	ANN & Rule based model for English to Sanskrit Translation (EST) [13]	2010	English-Sanskrit	General	Rule based	The system makes use of feed forward ANN to make selection of Sanskrit words and adjectives from English to Sanskrit User Data Vector (UDV). This system employs only morphological markings to identify Subject, Object, Verb, Preposition etc.
2.	English To Sanskrit Machine Translator ( Lexical parser & Semantic Mapper) [14]	2010	English-Sanskrit	General	Rule based	This system consists of 4 main modules: Lexical parser, Semantic Mapper, Translator & Composer. In this system the information can be obtained through introspection and analysis.
3.	Etrans- A complete Framework for English To Sanskrit Machine Translation [15]	2012	English-Sanskrit	General	Rule based	The translation model is primarily based on formulation of Synchronous Free Context Free Grammar (SCFG), a sub set of Context Free Grammar (CFG). Top-Down parsing algorithm is used for generating possibilities list.
4.	English to Sanskrit Translator and synthesizer [16]	2012	English-Sanskrit	General	Rule based	In the proposed algorithm system integrate traditional dictionary rule based approach for translation. It contains 2 models Text to Text Translator and Text to speech synthesizer module.
5.	English-to-Sanskrit Machine translation	2012	English-Sanskrit	General	Statistical based	Described the of a ubiquitous

	with ubiquitous applications [17]					translation and language learning framework, on Mini 2440 SBC, a growing cellular phone operating system with internet capabilities. Translation task is done by statistical machine decoder.
6.	TranSish [18]	2014	Sanskrit-English	General	Rule-based	With the help of Artificial Intelligence system provide an interface which converts Sanskrit sentences to English with a rule based model of parser and a semantic Mapper.

D.Translation system for Bengali language as a Source or Target language

SR No	Machine Translation Systems	Year	Languages for translation	Domain / Application	Approach used	Observations
1.	ANUBAAD [19]	2000-2004	English-Bengali	News headlines	Example based	If the headline is found in Generalized Tagged Example-base after synthesis then Bengali headline is generated. If the headline cannot be translated using Example-base, then Generalized Tagged example-base or Phrasal example-base is going to be used then after the heuristic translation strategy is used.
2.	VAASAANUBAADA [20]	2002	Bengali-Assamese	News text	Example based	Bilingual corpus is constructed and aligned manually. Longer sentences are fragmented at punctuation to obtain better quality translation.

3.	Exploiting Alignment Techniques in MATREX: the DCU Machine Translation System [21]	2008	English-Bengali	Conference papers	Example based	Makes use of marker-based chunking, psycholinguistic constraint which signifies context. System uses an “edit distance style” dynamic programming alignment algorithm for alignment purpose.
4.	A hybrid approach for Bengali to Hindi MT System [22]	2009	Bengali-Hindi	General	Hybrid	Multi-engine Machine Translation approach which Uses an integration of SMT with a lexical transfer based system (RBMT). The BLEU scores of SMT and lexical transfer based system are 0.1745 and .0424 respectively. The BLEU score of hybrid system is better which is 0.2275
5.	Lattice Based Lexical Transfer in Bengali Hindi MT Framework [23]	2011	Bengali-Hindi	General	Hybrid	Lattice based combined with transfer based Uses a lattice-based data structure which is a weighted directed acyclic graph.
6.	English To Bengali Machine Translation Using Context Free Grammars [24]	2013	English-Bengali	General	Rule based	Proposed system uses sentence construction rules in the form of CGF grammars. English-Bengali dictionary has been designed which support efficient translation.

E.Translation system for Punjabi language as a Source or Target language

SR No	Machine Translation Systems	Year	Languages for translation	Domain / Application	Approach used	Observations
1.	A Web Based English to Punjabi MT System for News Headlines [25]	2013	English-Punjabi	News Headline	Hybrid approach	Using Rule Based Approach, system parses the source text and produces an intermediate representation. The accuracy of the system is near about 81.67%.
2.	Statistical Machine Translation Based Punjabi to English Transliteration systems for proper noun [26]	2013	Punjabi-English	Government documents	Statistical based	The proposed system is divided to two parts – learning and transliteration. System Accuracy is depends on data stored into the database and overall accuracy of the system is measured upto 97%.
3.	Rule Based Machine Translation of Noun Phrases from Punjabi to English [27]	2010	Punjabi-English	To convert a noun phrase	Transfer	The system has analysis, translation and synthesis component. A Punjabi morph analyzer is being used for analyzing the exact grammatical structure of the word. Overall accuracy of translation is 85.33%.
4.	Web Based Hindi to Punjabi Machine Translation System [28]	2010	Hindi-Punjabi	News paper	Direct based	The present system is translating any complex sentence. The System accuracy is 95%.
5.	Hindi To Punjabi Machine Translation System [29]	2011	Hindi-Punjabi	General	Hybrid approach	Morphological analyzer developed by IIT-H has been ported which finds the root word for the token and its morphological features. The accuracy percentage for the system is 87.60%.

## F.Translation system for Urdu language as a Source or Target language

SR No	Machine Translation Systems	Year	Languages for translation	Domain / Application	Approach used	Observations
1.	Rule Based English to Urdu Machine Translation [30]	--	English-Urdu	Handle case phrases and verb post position	Transfer Approach	It describes the usage of Paninian theory in Urdu translation, which can handle case phrases and verb post position very efficiently. Moreover, this framework can also be used for other constructs and handle conditional and comparative sentences by designing their TAM.
2.	Model for English-Urdu Statistical Machine Translation [31]	2013	English-Urdu	General	Statistical Approach	The model is trained on TrainSet using Moses with language modeling toolkit IRSTLM. TestSet gives the BLEU score of 32.11.
3.	Urdu to English Machine Translation using Bilingual Evaluation Understudy [32]	2013	Urdu-English	General	Corpus based	Explained methodology of each system and found their comparison based on their respective outputs using BLEU. The EBMT approach produced accuracy of 84.21% whereas the accuracy of the online SMT system is 62.68%.
4.	Developing English-Urdu Machine Translation Via Hindi [33]	--	English-Urdu	General	Interlingua based Rule-based approach	The English-Hindi lexical database is used for collecting all possible Hindi words and also phrases. Paradigm file is used for generating different morphological forms from the root and entered

						based on correspondence with the tags associated with Hindi word. These are then manually validated.
5.	English to Urdu Hierarchical Phrase based statistical MT [34]	2013	English-Urdu	General	Hierarchical Phrase based	Uses EMILLE corpus. Language model is built on monolingual Urdu corpus and it is implemented as N-Gram model using SRILM toolkit.
6.	VERB TRANSFER FOR ENGLISH TO URDU MACHINE TRANSLATION (THESIS) [35]	2006	English-Urdu	Verb Transfer	Rule-based	Uses LFG for the analysis and generation modules of machine translation systems. Different senses and sub categorization frames were analyzed. Ditransitive Conversion rules are described to handle mismatch of ditransitive verb analysis in English and Urdu.
7.	Translation rules and ANN based model for English to Urdu Machine Translation [36]	2011	English-Urdu	General	ANN & Rule based	System uses feed-forward back-propagation artificial neural network for the selection of Urdu words/tokens and translation rules for grammar structure equivalent to English words/tokens and grammar structure rules. The n-gram blue score achieved is 0.6954; METEOR score achieved is 0.8583 and F-score of 0.8650.
8.	Hindi to Urdu Machine Translation	--	Hindi- Urdu	General	Character-based transliteration	Novel approach proposes two

	Through Transliteration [37]				model with a word-based translation model.	probabilistic models, based on conditional and joint probability. Use transliteration for disambiguation of Hindi homonyms which can be translated or transliterated or transliterated based on different contexts.
9.	Development of Parallel Corpus and English to Urdu Statistical Machine Translation [38]	2010	English-Urdu	General	Statistical approach	The whole corpus was divided into partitions for the purpose of cross-validation. The alignment of phrases is computed based on word-to-word alignment. Then after these translated phrases are sequenced using n-gram language model.

## G. Translation system for Malayalam language as a Source or Target language

SR No	Machine Translation Systems	Year	Languages for translation	Domain / Application	Approach used	Observations
1.	English to Malayalam Translation [39]	2008	English-Malayalam	General	Statistical Approach	Monolingual corpus of Malayalam is used and bilingual is used for English language. The structural difference between English Malayalam pair is resolved applying order conversion rules.
2.	A Hybrid approach to English to Malayalam Machine Translation [40]	2013	English-Malayalam	General	Hybrid Approach	A statistical machine translator performs translation by applying machine learning techniques on the corpus.
3..	Design & Development of a Malayalam To English Translator	2012	Malayalam-English	General	Transfer Approach	The system consists of a pre-processor for splitting the compound words, a

	[41]					morphological parser for context disambiguation and a bilingual dictionary. The system is designed using artificial intelligence techniques and can easily be modified to build translation systems for other language pairs.
4.	Malayalam to English Machine Translation [42]	2014	Malayalam-English	General	Example based	Consist of 3 phases: Acquisition, Matching & Recombination. The system searches in the corpus for each Malayalam fragments. About 75% of the test gives good quality translation.

H. Translation system for Kannada language as a Source or Target language

SR No	Machine Translation Systems	Year	Languages for translation	Domain / Application	Approach used	Observations
1.	MAT [43]	2002	English-Kannada	Government Circulars	Transfer based	Based on UCSG. 40-60% fully accuracy. Post editing tool is provided which outputs the number, type & inter-relationships amongst various clauses in sentences. For each word suitable target equivalence is obtained from bilingual dictionary.
2.	A Typical Machine Translation System for English-Kannada [44]	2014	English-Kannada	General	Rule based	In this system, Syntax reordering module does syntactic differences. Morphological generator which

						handle the complex morphology of the target language.
--	--	--	--	--	--	---

## I. Machine Translation for English to Indian language

SR No	Machine Translation Systems	Year	Languages for translation	Domain / Application	Approach used	Observations
1.	ANGLABHARTI-I [2]	2001	English-IL	Public health	Interlingua	Creates a PLIL intermediate structure. The effort of PLIL is 70% and text generation is 30%. Only with 30% new system can be built. In this 90% translation work is done by machine & 10% left to the human post-editing.
2.	ANGLABHARTI-II [2]	2004	English-IL	General	Example based	Provides provisions for automated pre-editing & paraphrasing, conditional multiword expressions as well as recognition of named-entities. Contains module for an error analysis, statistical language module.
3.	Shakti [52][53]	2003	English-IL	General	Transfer based	Linguistic rule based with statistical processing. Consist of various modules for analyzing the source languages, performing the bilingual task and generating target Indian language.
4.	Shiva and Shakti MT System [52][53]	2003	IL-IL	General	Example based	Easy to extend this system for new target

						language. Rules uses are mostly linguistic in nature. Semantic information is also used by some module.
5.	AnglaHindi [45]	2003	English-Hindi	General	Interlingua	Pseudo interlingua based. Uses all modules of Anglabharti. Use an abstracted example base. Accuracy is 90%.
6.	“English to Indian Languages MT System (E-ILMT)” [46]	2006	English-IL	Tourism and healthcare	Statistical based	The engine was developed using statistical techniques and tools such as fnTBL, Bikel, Pharaoh. Pre-processing phase does the syntactic re-ordering on the source language to reduce long distance movements. The corpus was morphologically processed and used for training to tackle the problem of degradation. A rule based suffix separation approach was used for separating the root word and the affixes.
7.	Google Translator [54]	2007	IL-IL	Geneal	Statistical based	Google Translate is a multilingual service supports 91 languages. The system depends on a solid corpus. The accuracy of translation is better to understand the translated text.
8.	ANUVAADAK by IIT-B [55]		IL-IL	General	Statistical Based	The system enables translation

						between 11 different Indian languages (including English), and provides transliteration support for Indic script input. Users can post-edit the translated text which provides feedback for improvement of the translation systems
9.	Sampark by IIT-H [56]	2009	IL-IL	General	Hybrid approach	Sampark is a hybrid system which is combination of traditional rules-based algorithms and dictionaries and various statistical machine-learning techniques.

#### 4. CONCLUSION

The present paper discusses the various language translation systems developed in the India which follows different approaches. The systems are developed for various languages which include various language pairs such as English-Hindi, English-Marathi, English-Sanskrit, English-Malayalam, Malayalam-English, Bengali-Hindi, Bengali-Aasamese, English-Punjabi, Hindi-Punjabi, Punjabi-English/Hindi, English-Kannada, English-Urdu, Urdu-English and IL-IL.

It is concluded that direct approach for Machine Translation is most suitable for closely related languages i.e. the languages with similar structure. The indirect and statistical approach is suitable for languages with different structures.

#### REFERENCES

- [1] Akshar Bharti, Chaitanya Vineet, Amba P. Kulkarni & Rajiv Sangal, (1997) ANUSAARAKA: Machine Translation in stages, Vivek, a quarterly in Artificial Intelligence, Vol. 10, No. 3, NCST Mumbai, pp. 22-25.
- [2] Sudip Naskar & Shivaji Bandyopadhyay, (2005) "Use of Machine Translation in India: Current status" AAMT Journal, pp. 25-31
- [3] Lata Gore & Nishigandha Patil, (2002) "English to Hindi - Translation System", In proceedings of Symposium on Translation Support Systems. IIT Kanpur. pp. 178-184

- [4] Ananthakrishnan R, Kavitha M, Jayprasad J Hegde, Chandra Shekhar, Ritesh Shah, Sawani Bade & Sasikumar M., (2006) “MaTra: A Practical Approach to Fully- Automatic Indicative EnglishHindi Machine Translation”, In the proceedings of MSPIL-06
- [5] Choudhary, A. Singh, M. (2009) “GB theory based Hindi to english translation system”, Computer Science and Information Technology, 2009. ICCSIT 2009. 2nd IEEE International Conference PP.293 – 297
- [6] Ruchika A. Sinhal, Kapil O. Gupta (2014) “A Pure EBMT Approach for English to Hindi Sentence Translation System” I.J.Modern Education and Computer Science, 2014, 7, 1-8 Published Online July 2014 in MECS (<http://www.mecs-press.org/>)
- [7] Shachi Dave, Jignashu Parikh And Pushpak Bhattacharyya Department of Computer Science and Engineering, Indian Institute of Technology, Mumbai, India, “Interlingua based English-Hindi Machine Translation system and Language Divergence.”
- [8] M.L.Dhore & S.X.Dixit (2011) “English to Devnagari Translation for UI Labels of Commercial web based Interactive Applications” International Journal of Computer Applications ( 0975-8887) Volume 35-No.10 , December 2011
- [9] Devika Pishartoy, Priya, Sayli Wandkar (2012) “Extending capabilities of English to Marathi machine Translator” , International journal of Computer Science Issues, Vol.9, Issues 3, No. 3, May 2012 ISSN (Online): 1694-0814
- [10] Abhay A, Anita G, Purnima T, Prajakta G (2013), “Rule based English to Marathi translation of Assertive sentence” , International Journal of Scientific & Engineering Research, Volume 4, Issues 5, May- 2013 ISSN 2259-5518 pp. 1754-1756
- [11] Krushnadeo B, Vinod W, S.V.Phulari, B.S.Kankate (2014), “A novel approach for Interlingual example-based translation of English to Marathi”, International Journal of Emerging Technology and Advanced Engineering Website: [www.ijetae.com](http://www.ijetae.com) (ISSN 2250-2459, ISO 9001:2008 Certified Journal, Volume 4, Issue 3, March 2014)
- [12] G.V.Gajre, G.Kharate, H. Kulkarni (2014), “Transmuter: An approach to Rule-based English to Marathi Machine Translation” , International Journal of Computer Applications (0975 – 8887) Volume 98 – No.21, July 2014
- [13] Vimal Mishra, R.B.Mishra Research Scholar, Department of Computer Engineering, Institute of Technology, Banaras, Hindu University, (IT-BHU), Varanasi-221005, U.P., India, “ANN and Rule Based Model for English to Sanskrit Machine Translation”
- [14] Ms.Vaishali.M.Barkade, Prof. Prakash R. Devale , Dr. Suhas H. Patil, “ENGLISH TO SANSKRIT MACHINE TRANSLATOR LEXICAL PARSER AND SEMANTIC MAPPER”, National Conference On "Information and Communication Technology" NCICT-10
- [15] Promila Bahadur ,D.S.Chauhan, , A.K.Jain , Indian Institute of Technology Kanpur, India, “ EtranS- A Complete Framework for English To Sanskrit Machine Translation”, IJACSA Special Issue on Selected Papers from International Conference & Workshop On Emerging Trends In Technology 2012 pp. 52-59
- [16] Sarita G. Rathod, Shanta Sondur, Information Technology Department, VESIT, Mumbai University, Maharashtra, India, “English to Sanskrit Translator and Synthesizer”, International Journal of Emerging Technology and Advanced Engineering Website: [www.ijetae.com](http://www.ijetae.com) (ISSN 2250-2459, ISO 9001:2008 Certified Journal, Volume 2, Issue 12, December 2012)
- [17] Sandeep R. Warhade, Prakash R. Devale ,Suhas H. Patil , “ English-to-Sanskrit Statistical Machine Translation with Ubiquitous Application” , International Journal of Computer Applications (0975 – 8887) Volume 51– No.1, August 2012
- [18] Pankaj Upadhyay, Umesh Chandra Jaiswal, Kumar Ashish, “ TranSish: Translator from Sanskrit to English-A Rule based Machine Translation” , International Journal of Current Engineering and Technology E-ISSN 2277 – 4106, P-ISSN 2347 - 5161 ©2014 INPRESSCO®, All Rights Reserved Available at <http://inpressco.com/category/ijcet>
- [19] S. Bandyopadhyay, (2004) "ANUBAAD - The Translator from English to Indian Languages", in proceedings of the VIIth State Science and Technology Congress. Calcutta. India. pp. 43-51

- [20] Kommaluri Vijayanand, Sirajul Islam Choudhury & Pranab Ratna “VAASAANUBAADA - Automatic Machine Translation of Bilingual Bengali-Assamese News Texts”, in proceedings of Language Engineering Conference-2002, Hyderabad, India © IEEE Computer Society.
- [21] Yanjun Ma, John Tinsley, Hany Hassan, Jinhua Du & Andy Way, (2008) “Exploiting Alignment Techniques in MATREX: the DCU Machine Translation System for IWSLT 2008”, in proceedings of IWSLT 2008, Hawaii, USA.
- [22] Sanjay Chatterji, Devshri Roy, Sudeshna Sarkar, Anupam Basu, 2009, “A Hybrid Approach for Bengali to Hindi Machine Translation” In proceedings of ICON 2009, 7th International Conference on Natural Language Processing. pp. 83-91.
- [23] Sanjay Chatterji, Praveen Sonare, Sudeshna Sarkar, and Anupam Basu, 2011, “Lattice Based Lexical Transfer in Bengali Hindi Machine Translation Framework”, In Proceedings of ICON-2011: 9th International Conference on Natural Language Processing, Macmillan Publishers, India.
- [24] Shibli A, Humayun K, Musfique A, K.M.Noman, 2013, “English To Bengali Machine Translation Using Context Free Grammars”, International journal of Computer Science Issues, vol.10, Issues 3, No.2, May 2013 ISSN: 1694-0814 pp. 144-153
- [25] Harjinder Kaur, Dr. Vijay Laxmi, 2013 “A Web Based English to Punjabi MT System for News Headlines” In International Journal of Advanced Research in Computer Science and Software Engineering 3(6), June - 2013, pp. 1092-1094
- [26] Pankaj Kumar and Er.Vinod Kumar, 2013, “Statistical Machine Translation Based Punjabi to English Transliteration System for Proper Nouns”, In International Journal of Application or- Innovation in Engineering & Management (IJAIEEM) Volume 2, Issue 8, August 2013 ISSN 2319 – 4847, pp . 318-320
- [27] Kamaljeet Kaur Batra and G S Lehal, 2010, “Rule Based Machine Translation of Noun Phrases from Punjabi to English”, In IJCSI International Journal of Computer Science Issues, Vol. 7, Issue 5, September 2010 ISSN (Online): 1694-0814, pp. 409-413
- [28] Vishal Goyal and Gurpreet Singh Lehal , 2010, “Web Based Hindi to Punjabi Machine Translation System”, JOURNAL OF EMERGING TECHNOLOGIES IN WEB INTELLIGENCE, VOL. 2, NO. 2, MAY 2010, pp.148-151
- [29] Vishal Goyal and Gurpreet Singh Lehal , 2011, “HINDI TO PUNJABI MACHINE TRANSLATION SYSTEM”, Proceedings of the ACL-HLT 2011 System Demonstrations, pages 1–6, Portland, Oregon, USA, 21 June 2011. Association for Computational Linguistics
- [30] Naila Ata, Bushra Jawaid, Amir Kamran “Rule Based English to Urdu Machine Translation”
- [31] Aasim Ali, Arshad , Hussain and Muhammad Kamran Malik, 2013, “Model for English-Urdu Statistical Machine Translation”, World Applied Sciences Journal 24 (10): 1362-1367, 2013 ISSN 1818-4952 © IDOSI Publications, 2013 DOI: 10.5829/idosi.wasj.2013.24.10.760
- [32] Asad Habib, Asad Abdul Malik ,Kohat University of Science and Technology, Kohat, Pakistan , 2013, “Urdu to English Machine Translation using Bilingual Evaluation Understudy” International Journal of Computer Applications (0975 – 8887) Volume 82 – No 7, November 2013, pp. 5-12
- [33] R. Mahesh K. Sinha Department of Computer Science & Engineering, Indian Institute of Technology, Kanpur, India, “Developing English-Urdu Machine Translation Via Hindi”
- [34] Nadeem khan, Waqas Anwar, Nadir Durrani, 2013, “English to Urdu Hierarchical Phrase based statistical Machine translation”The 4th Workshop on South and Southeast Asian NLP (WSSANLP), International Joint Conference on Natural Language Processing, pages 72–76
- [35] Nayyara Karamat December, 2006, “VERB TRANSFER FOR ENGLISH TO URDU MACHINE TRANSLATION (Using Lexical Functional Grammar (LFG)) – MS Thesis” at the National University of Computer & Emerging Sciences
- [36] SHAHNAWAZ, R. B. MISHRA, 2011, “Translation Rules and ANN based model for English to Urdu Machine Translation”, INFOCOMP, v. 10, no. 3, p. 36-47, September of 2011
- [37] Nadir Durrani ,Hassan Sajjad ,Alexander Fraser & Helmut Schmid, Institute for Natural Language Processing University of Stuttgart, “Hindi to Urdu Machine Translation Through Transliteration”

- [38] Aasim Ali, Shahid, Muhammad Malik, 2010, “ Development of Parallel Corpus and English to Urdu Statistical Machine Translation”, International Journal of Engineering & Technology IJET-IJENS Vol:10 No:05, pp.31-33
- [39] Mary Priya Sebastian, Sheena Kurian K, G. Santhosh Kumar, 2009 “ English to Malayalam Translation: A Statistical Approach”
- [40] Nithya B, Shibily Joseph, 2013 “ A Hybrid Approach to English to Malayalam Machine Translation”, International Journal of Computer Applications (0975 – 8887) Volume 81 – No.8, November 2013, pp. 11-15
- [41] Latha R Nair, David Peter & Renjith P Ravindran, 2012, “Design and Development of a Malayalam to English Translator-A Transfer Based Approach”, International Journal of Computational Linguistics (IJCL), Volume (3) : Issue (1) : 2012
- [42] Anju E S, Manoj Kumar K V, 2014, “ Malayalam To English Machine Translation: An EBMT System”, IOSR Journal of Engineering (IOSRJEN) ISSN (e): 2250-3021, ISSN (p): 2278-8719 Vol. 04, Issue 01 (January. 2014), ||V1|| PP 18-23
- [43] Murthy. K, (2002) “MAT: A Machine Assisted Translation System”, In Proceedings of Symposium on Translation Support System( STRANS-2002), IIT Kanpur. pp. 134-139
- [44] Mr. Chethan Chandra S Basavaraddi , Dr. H. L. Shashirekha, 2014, “A Typical Machine Translation System for English to Kannada” , International Journal of Scientific & Engineering Research, Volume 5, Issue 4, April-2014 ISSN 2229-5518
- [45] R.M.K. Sinha & A. Jain, (2002) “AnglaHindi: An English to Hindi Machine-Aided Translation System”, International Conference AMTA(Association of Machine Translation in the Americas)
- [46] Vishal Goyal & Gurpreet Singh Lehal, (2009) “Advances in Machine Translation Systems”, National Open Access Journal, Volume 9, ISSN 1930-2940 <http://www.languageinindia>
- [47] [http://www.iiit.net/ltrc/Anusaaraka/anu\\_home.html](http://www.iiit.net/ltrc/Anusaaraka/anu_home.html)
- [48] <http://www.cdac.in/html/aai/mantra.asp>
- [49] [http://www.academia.edu/7986160/Machine\\_Translation\\_of\\_Bilingual\\_Hindi-English\\_Hinglish\\_Text](http://www.academia.edu/7986160/Machine_Translation_of_Bilingual_Hindi-English_Hinglish_Text)
- [50] <http://www.cflit.iitb.ac.in/machine-translation/eng-hindi-mt>
- [51] <http://www.ncst.ernet.in/matra/>
- [52] <http://ebmt.serc.iisc.ernet.in/mt/login.html>
- [53] <http://shakti.iiit.net>
- [54] <https://translate.google.co.in>
- [55] [www.cflit.iitb.ac.in/indic-translator/](http://www.cflit.iitb.ac.in/indic-translator/)
- [56] <http://sampark.iiit.ac.in/sampark/>

## Authors

**Ms. Amruta Godase** is Pursuing M.E. in Information Technology with Specialization in Artificial Intelligence & Robotics from Mumbai University. She received her polytechnic from MSBTE & B.E from Mumbai University.



**Sharvari Govilkar** is Associate professor in Computer Engineering Department, at PIIT, New Panvel, University of Mumbai, India. She has received her M.E in Computer Engineering from University of Mumbai. Currently She is pursuing her PhD in Information Technology from University of Mumbai. She is having Sixteen years of experience in teaching. Her areas of interest are Text Mining, Natural language processing, System programming & Compiler Design etc.

