MACHINE TRANSLATION WITH SPECIAL REFERENCE TO MALAYALAM LANGUAGE

Jomy Jose

Asst. Librarian
Asian School of Business, Technocity,Trivandrum ,Kerala
Email:mail2jomichan@gmail.com

Abstract

Google Translate gives machine translation services in online particularly for written content. In the 24th phase of the Google Translate project undertaking started on June 2011 five new Indic languages Bengali, Gujarati, Kannada, Tamil, Telugu were included yet Malayalam – the Classical languages spoken 38 million individuals in the state of Kerala has been barred. This paper delineates the significance of machine translations with special reference Malayalam language.

Keywords : Malayalam, Machine translation , Google Translate

INTRODUCTION

Machine Translation got paramount in the field of correspondence around the world. Generally individuals are intrigued to peruse, compose and chat on their own local dialect. Machine translation is one of the examination ranges under computational semantics. Different techniques have been proposed to mechanize the interpretation process. Since the time that the appearance of Malayalam computing on a bigger scale around the turn of the thousand years, the absence of exact English to Malayalam interpreter has been a bug. Google have discovered the script and grammar testing and they are not yet to include the language in the translator list.

Merriam Webster dictionary defined translation as, its is an act or process of translating something into a different language.[1] Harinder and Vijay Laxmi mentioned machine translation as the study of designing the systems that can translate one human language into another.[2] These systems take input in one natural language and convert it into another human language. The language that is given as an input is called Source Language and the language in which we get the output is called Target language.

Machine Translation is computerized systems responsible for the production of translations from one natural language into another with or without human assistance. It is a sub-field of computational linguistics that investigates the use of computer software to translate text or speech from one natural language to another.

REVIEW OF LITERATURE

Anju and Kumar (2014) prescribed a Machine Translation framework for transaction from Malayalam to English language. The translation framework is dependent upon Example Based Machine Translation (EBMT) approach. The info to the translation framework is Malayalam sentence and the relating English sentence is created as yield. Case Based machine translation is dependent upon the thought of reusing the effectively interpreted samples. Illustration based translation includes three real steps -Example obtaining, Matching and Recombination. It is established that the translation framework works well for the basic sentence in Malayalam language.[3]

Antony (2013) portrayed different methodologies of significant machine translation improvements in India. The literary works demonstrates that there have been numerous endeavors in MT for English to Indian languages and Indian languages to Indian languages. At present, various government and private division tasks are working towards creating a full-fledged MT for Indian languages. Despite the fact that there has been exertion towards building English to Indian language and Indian language to Indian language translation framework, shockingly, we don't have an effective translation framework starting today. [4]

Chéragu (2010) clarified the interest for language translation has enormously expanded lately because of expanding cross-local correspondence and the need for data trade. Most material needs to be deciphered, including investigative and specialized documentation, guideline manuals, authoritative archives, course books, reputation flyers, daily paper reports and so on. Some of this work is testing and troublesome however basically it is repetitive and dull and obliges consistency and exactness. It is getting challenging for expert interpreters to meet the expanding requests of translation. In such a circumstance the machine translation could be utilized as a substitute.[5]

Dwived and Sukhadeve(2010) gave a concise thought on the machine translation systems scenario in India through information and past exploration on machine interpretation.Natural Language Processing (NLP)

and Machine Translation (MT) devices are approaching regions of study the field of computational semantics. Machine interpretation is the provision of workstations to the interpretation of writings from from one natural language into another natural language. It is an essential sub-control of the more extensive field of artificial intelligence.[6]

Naskar and Bandyopadhyay (2005) led a study of the machine translation frameworks created in India for translation from English to Indian languages and around Indian languages uncovers that the MT programming projects are utilized as a part of field testing or are accessible as web translation administration. These frameworks are additionally utilized for showing machine translation to the scholars and scientists. The majority of these frameworks are in the English-Hindi or Indian language-Indian language area. The translation spaces are basically government documents/reports and news stories.[7]

MALAYALAM LANGUAGE

Malayalam, spoken in India, is a language prevalent in the state of Kerala. It is one of the 22 formally distinguished languages of India exceptionally in the state of Kerala and in the union domains of Lakshadweep and Puducherry. It was proclaimed a classic language by the Government of India in 2013. Fitting in with the Dravidian group of dialect it is spoken and loved by more or less 38 million individuals. Malayalam is likewise spoken in the neighboring states of Tamil Nadu and Karnataka; with additional people in the Nilgiris, Kanyakumari and Coimbatore regions of Tamil Nadu and the Dakshinakannada and Kodagu areas of Karnataka. Many words of Malayalam have been acquired from Sanskrit. There are 37 consonants and 16 vowels in the script[8].Malayalam has a composed conventional going back from the late ninth century and the soonest work dates from thirteenth century.

NEED OF MACHINE TRANSLATION

In an expansive multilingual social order like India, there is an extraordinary interest for interpretation of archives starting with one language then onto the next language. The constitution gives that all incidents in the Supreme Court of India, the nation's most elevated court and the High Courts, might be in English. Subject to the procurements of articles 346 and 347, the Legislature of a State might by law embrace any one or a greater amount of the languages being used in the State or Hindi as the language or languages to be utilized for all or any of the authority purposes of that State. Gave that, until the Legislature of the State generally gives by law, the English dialect might keep on being utilized for those authority purposes inside the State for which it was being utilized instantly before the beginning of this Constitution. The vast majority of the state government works in there commonplace languages, though the central government's authority records and reports are in English and Hindi. To have a proper correspondence between the state and central government there is a necessity to interpret these records and reports in the particular state languages. [9]

GOOGLE TRANSLATE

Google Translate provides machine translation services in online especially for written text.[10] In the 24th stage of the Google Translate project launched on June 2011 five new Indic languages Bengali, Gujarati, Kannada, Tamil, Telugu were added but Malayalam – the Classical language spoken 38 million people in the state of Kerala has been excluded.

ADVANTAGES OF GOOGLE TRANSLATE

- A Universal Communicator in your hand
- Convenience of use
- Fast translation
- Simple user interface for all
- To translate all websites, web pages
- All Wikipedia articles can easily translate to local language and vice versa
- Translate English PDF e-Books to local language

COMPARISON :-LANGUAGE AVAILABLE IN GOOGLE TRANSLATE AND NATIVE SPEAKERS

Table -1: Language available in Google Translate and Native Speakers

Sl.No	Language	Native speakers
1	Afrikaans	7.1 million
2	Albanian	7.4 million
3	Armenian	6 million
4	Azerbaijani	23 million
5	Basque	7.2 Lakhs
6	Belarusian	4 million
7	Bosnian	3.5 million
8	Bulgarian	10 million
9	Catalan	7.2 million
10	Cebuano	21 million
11	Croatian	5.5 million
12	Czech	10 million
13	Danish	5.6 million
14	Dutch	23 million
15	Esperanto	2 million
16	Estonian	1.05 million
17	Filipino	28 million
18	Finnish	5 million
19	Galician	3.2 million
20	Georgian	7 million
21	Greek	13 million
22	Haitian Creole	9.6 million
23	Hebrew	5.3 million
24	Hmong	4 million
25	Hungarian	14 million
26	Icelandic	3.2 Lakhs
27	Indonesian	23 million
28	Irish	1.80 million
29	Khmer	16 million
30	Lao	25 million
31	Latin	Dead Language
32	Latvian	1.3 million
33	Lithuanian	3.1 million
34	Macedonian	2.5 million
35	Malayalam	38 million
36	Maltese	4.3Lakhs
37	Norwegian	5 million
38	Romanian	25 million
39	Serbian	10.2 million
40	Slovak	5 million
41	Slovenian	2.5 million
42	Swedish	8.5 million
43	Thai	20 million
44	Ukrainian	37 million
45	Welsh	7.2 Lakhs
46	Yiddish	1.5 million

Table -1 clearly reveals that such a variety of languages accessible in Google Translate in which local speakers were not exactly that of Classical Malayalam language. Individuals of Kerala appreciate the most astounding extent of Internet access office and Internet use than in whatever available Indian state. Legislature of Kerala has launched measures to enhance machine and ICT abilities of individuals living in Kerala through the Akshaya Project and the IT@school Project. [11]These undertakings have assumed a significant part in improving e-ability by channeling preparing projects in cooperation with Kerala State IT Mission and different IT companies. Both activities point at spanning the advanced separation, giving preparing in essential workstation expertise

and empowering accessibility of important substance through the utilization of Malayalam dialect. Akshaya Centers spotted all over Kerala have risen as a significant channel between government and nationals, administrations which now fill in as a spine of the administration framework in Kerala. In this way, all exercises of the project achieve their objective through Malayalam Language and Malayalam Computing. Besides, Kerala additionally has the most elevated number of clients and promoters of free and open source virtual products. The better utilization of sites and data requires an interpretation administration. Hence, its exceedingly prescribed that Google ought to begin translation services for Malayalam in Google Translate.

CONCLUSION

Numerous endeavors are continuously made everywhere throughout the world to create machine translation frameworks for different languages utilizing rule-based as well as statistically based approaches. Such a variety of different languages on the planet with less local speaker have machine interpretation benefits in Google translate. Paribhashika programming, mutually created by C-DAC and Kerala Bhasha Institute, can interpret complex English sentences into Malayalam with a high level of correctness.[12] It is prepared to interpret complex sentences with various structures. There are numerous machine interpretation programming projects accessible on the web, Google translator is recognized as the best. A joined together exertion of Google translator aw well as government offices will give an unreservedly accessible Malayalam to English and vive versa to the regular clients of information and workstation engineering.

REFERENCE

[1] Translate. (n.d.). Merriam-Webster. Retrieved January 27, 2014, from http://www.merriam-webster.com/dictionary/translate
[2] Harjinder,Kaur and Vijay Laxmi A survey of machine translation approaches,International Journal of Science, Engineering and Technology Research,2013 Vol.2
[3] Anju E S and Kumar , Manoj K V,Malayalam To English Machine Translation:An EBMT System IOSR Journal of Engineering ,2014,Vol.4 .pp18-23
[4] Antony, P. J.Machine Translation Approaches and Survey for Indian Languages Computational Linguistics and Chinese Language Processing .2013 Vol.18.pp.47-78
[5] Chéragui, Mohamed AmineTheoretical Overview of Machine translation 4th International Conference on Web and Information Technologies. ICWIT 2012, Sidi Bel-Abbes, April 29-30 2012,pp.160-169
[6] Dwivedi, Sanjay Kumar and Sukhadeve, Pramod Premdas Machine Translation System in Indian Perspectives Journal of Computer Science,2010, Vol .6.pp.1111-1116,
[7] Naskar, Sudip and Bandyopadhyay, Sivaji ,Use of Machine Translation in India: Current Status 465-490 http://www.mt-archive.info/MTS-2005-Naskar-2.pdf
[8] History of malayalam. (n.d.). chintha. Retrieved March 7, 2014, from http://chintha.com/keralam/malayalam/language-history.html
[9] Government of India ,The Official Languages 1976 (As Amended, 1987, 2007, 2011)
[10] Google Translate. Wikipedia. Retrieved March 7, 2014, from http://en.wikipedia.org/wiki/Google_Translate
[11] Malayalam translation services in Google translate. (n.d.). Causes. Retrieved March 7, 2014, from https://www.causes.com/actions/1760822-a-petition-to-google
[12] Eliminating translation bugs. (n.d.). The Hindu. Retrieved March 5, 2014, from http://www.thehindu.com/news/cities/Thiruvananthapuram/eliminating-translation-bugs/article5305420.ece