

English to ISL generation

Anurag Kumar Jha – 2020A7PS0128U
-under Dr. Sujala D. Shetty

Table of Contents

1. Abstract	5
2. Introduction	6
2.1 Objective	6
2.2 Problem Statement	6
2.3 Existing Approaches	6
2.4 Proposed Approach	6
2.5 Scope of Investigation	7
3. Theoretical Bases and Literature Review	8
3.1 Theoretical Background of the Problem	8
3.2 Related Research of the Problem	8
3.3 Advantages/Disadvantages of those research	10
3.4 Our Solution to the Problem	10
3.5 Where Our Solution is different	10
4. Hypothesis	11
5. Methodology	12
5.1 Data Collection	12
5.2 Algorithm design	12
5.2.1 Solution Structure	12
5.2.2 Language	16
5.2.3 Tools used	16
5.3 Output Generation	16
6. Data Analysis and Discussion	18
6.1 Output Generation	18
6.2 Output Analysis	19
6.3 Compare Output Against Hypothesis	19
6.4 Abnormal Case Explanation	20
7. Conclusions and Recommendations	21
7.1 Summary and Conclusions	21
7.2 Recommendations for future studies	21
8. Bibliography	22

List of Figures

1. ISL Type Hierarchy	9
2. Parts of Speech Tagging	12
3. Removing Stopwords	15
4. Stemming of Words	15

List of Tables

1. Media Comparison for representing sign	11
2. Comparison of Grammar of English and Indian Sign Language	13
3. Examples of Grammatical Reordering of Words of English Sentence	14

Acknowledgement

I would like to thank our honorable Director Prof. Srinivasan Madapusi, for giving me the opportunity to work on this Design Project, under the field of NLP. I would like to express my deepest appreciation to all those who provided me with the possibility to complete this report. A special gratitude and note of appreciation to my professor, mentor and guide, Prof. Dr. Sujala D. Shetty, whose contribution in stimulating suggestions and encouragement, helped me to coordinate my project and understand all things better, especially in writing this report. Furthermore, I would also like to acknowledge with much appreciation the crucial role of our seniors, who gave me an inspiration to think better in terms of content quality.

1. Abstract

Sign language is a natural way of communication for challenged people with speaking and hearing disabilities. There have been various mediums available to translate or to recognize sign language and convert them to text, but text to sign language conversion systems have been rarely developed, this is due to the scarcity of any sign language corpus. Our project aims at creating a translation system that consists of a parsing module which parses the input English sentence to phrase structure grammar representation on which Indian sign language grammar rules are applied. This is done by eliminating stopwords from the reordered sentence. Stemming is applied to convert the words to their root form as Indian sign language does not support for inflections of the word. All words of the sentence are then checked against the words in the dictionary containing videos representing each of the words. If the words are not found in the dictionary, its corresponding synonym is used to replace it.

The proposed system is innovative as the existing systems are limited to direct conversion of words into Indian sign language whereas our system aims to convert these sentences into Indian sign language grammar in real domain.

2. Introduction

2.1 Objective

The objective of the project is to convert English text language to Indian Sign Language using Natural Language Processing to enhance the communication capabilities of people with hearing disabilities.

2.2 Problem Statement

Sign language is a language that uses manual communication methods such as facial expressions, hand gestures and bodily movements to convey information. This project makes use of videos for specific words combined to translate the text language into sign language.

2.3 Existing Approach

Although sign language is used across the world to bridge the gap of communication for hearing or speech impaired which depend mostly on sign language for day to day communication, there are not efficient models that convert text to Indian sign language. There is a lack of proper and effective audio visual support for oral communication. While significant progress has already been made in computer recognition of sign languages of other countries but a very limited work has been done in ISL computerization.

Work done so far in this field has been much more focused on American sign language (ASL) or British sign language, but for Indian sign language, systems that have been developed are very few. The underlying architecture for most of the systems are based on:

1. Direct translation: Words from source language are directly transformed into target language words. Output may not be a desired one.
2. Statistical Machine Translation: It requires large parallel corpus, which is not readily available in case of sign language.
3. Transfer based architecture: Grammar rules are being applied so as to define a proper translation from one language system to another.

2.4 Proposed Approach

Few works have been done to generate a system that is based on the above concepts listed in the existing approaches section and cater to Indian sign language. Thus we propose to develop one for Indian sign language based on transfer based translation.

The success of this translation system will depend on the conversion of English text to Indian sign language bearing its lexical and syntactic knowledge.

2.5 Scope of investigation

Our project aims to encompass the domain of Indian Sign Language which roughly consists of 1500 words in its dictionary. For each of these words its corresponding video will be gathered. For words which do not fall in this dictionary, they would be replaced by their synonyms considering duplicacy of words as well as their parts of speech.

Translation of one spoken language to another spoken language is complex task if both the languages have different grammar rules. The complexity is increased many folds when source language is spoken language and the target language is sign language.

The target audience for this system is not limited to hearing impaired individuals as communication for hearing impaired people in common places like railway stations, bus stands, banks, hospitals etc is very difficult because a vocal person may not understand sign language and thus won't be able to convey any message to hearing impaired person. Thus, it is targeted to all those who wish to learn this translation, to facilitate better communication.

3. Theoretical Bases and Literature Review:

3.1 Theoretical Background of the Problem

According to 2011 census of India, there are 63 million people which sums up to 6.3% of the total population, who are suffering from hearing problems. Out of these people, 76-89% of the Indian hearing challenged people have no knowledge of language either signed, spoken or written. The reason behind this low literacy rate is either the lack of sign language interpreters, unavailability of Indian Sign Language tool or lack of researches on Indian sign language.

3.2 Related Research to Solve the Problem

In spite of the modern computer system being so advanced, there is a paucity of research in developing machine translation (MT) system on sign language particularly in India. Some of the MT systems used for other sign languages are:

1. Direct Translation System

This system is based on word to word conversion. The meaning and context of sentences are not taken in consideration. Neither the grammar is changed, direct transformation into target sign language. No syntactic analysis takes place on the original text and even the order of conversion is unaffected. However while converting from English text to ISL, the word order of ISL may not be same as given text. To overcome this problem, a system is needed that has strong knowledge of both source and target language.

2. Transfer Based Translation

In this system, plain text is given as input and this input then undergoes syntactic and semantic transformation and then finally transformed into a sign language. In this system, source language is transformed into some intermediate text and then after applying some linguistic rules it is transformed into target language. It also called as Rule based Translation.

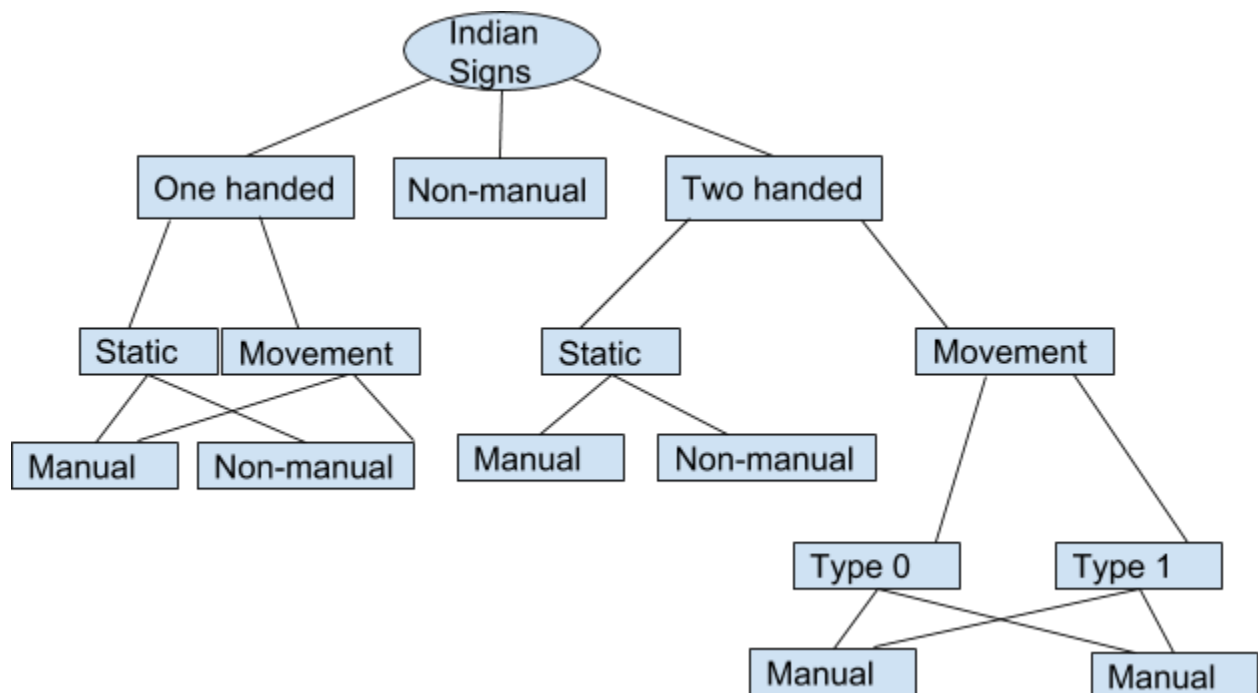
3. Interlingua Based Translation

In this system, a language independent semantic structure is produced by doing only semantic analysis on the original input text. This independent structure is called Interlingua. The target language is then generated from this Interlingua. It is thus an alternative of both direct and transfer based translation.

Facts about Sign Language:

Sign languages is not just 'natural language represented by sign' or not just hand representation of the words as it is but rather it is the representation of meaning. There are various facts associated with sign language, a natural language, which most of us are unaware. Some of them are listed below:

- NOT the same all over the world
- NOT just gestures and pantomime but do have their own grammar.
- Dictionary is smaller compared to other languages.
- Finger-spelling for unknown words.
- Adjectives are placed after the noun for most of the sign language.
- Never use suffixes
- Always sign in present tense
- Do not use articles.
- Do not uses I but uses me.
- Have no gerunds
- Use of eyebrows and non-manual expressions.
- NOT been invented by hearing people.



ISL Type Hierarchy

[Figure 1]

The one handed signs are represented by static as well as dynamic movement of one hand. The static and dynamic movements are then classified into manual and non-manual signs. Two handed signs use both the hands for gesturing. It can be further classified into Type 0 and

Type 1.

Type 0 signs are those where the signer makes use of both the hands

Type 1 signs are those where the use of one hand (dominant) is more compared to the other hand (non-dominant)

Indian Sign Language Grammar

This language has its own grammar and is not same as the manual representation of spoken English or Hindi. Certain distinct features that it has are:

1. Number representation are done with hand gestures for each hand.
2. Signs for family relationships are preceded by male or female.
3. In interrogative sentence, all the WH questions are placed in the back of the sentence.
4. It also consists of many non-manual gestures such as mouth pattern, mouth gestures, body postures, head position and eye gaze.
5. The past, present and future tense is represented by signs for before, then and after.

3.3 Advantages and disadvantages

The advantage of using the video technique is that it is more effective and can convey the message in the most effective way. Whereas, the other approaches are using the images which are way less effective as compared to videos.

3.4 Our Solution to the Problem

The module for translation will help hearing disabled people to understand in an efficient and easy way by providing them with a video to convey them the message of text.

3.5 Where our Solution is Different

Our solution is different as we are using the technique of videos while other techniques are using the images for message conveyance. Also the stopwords have been removed so text processing efforts are also reduced.

4 Hypothesis

Translation from source language to target language requires bilingual dictionary. Thus, we will be creating one such dictionary that contains the english word and its equivalent Indian sign. The counterpart for the English word could be in the format of images, videos or coded sign language text (gloss). All the approaches have their own pros and cons but the video approach is well suited as it has an edge over other systems. A comparison of all the media has been given in the table as shown below:

Kind of media	Pros	Cons
Video Signs	<ul style="list-style-type: none">• Realistic• Easy to create	<ul style="list-style-type: none">• Time consuming to create• High memory consumption• Not supported by translation system
Pictures	<ul style="list-style-type: none">• Very less memory consumption	<ul style="list-style-type: none">• Time consuming to create• Not realistic as compared to videos• Not supported by translation system
Code Sign Language Text	<ul style="list-style-type: none">• Minimal Memory Consumption• Supported by translation system as it is the written form and can be processed very easily	<ul style="list-style-type: none">• Very difficult to read and understand• Required to be learnt

Media Comparision for representing sign

[Table-1]

Because videos provide much realistic content as compared to image and coded text so we preferred using videos.

If time permits we wish to develop a system based on synthetic animations as it requires less memory consumption, can easily be reproduced, is supported by translation system and follows an avatar based approach which can be easily customized.

5 Methodology

5.1 Data Collection

We will be using the <http://www.indiansignlanguage.org/> to download the video clips of each and every word. We will manually label each of the video and remove videos that we will find irrelevant. We will like to maintain an unfiltered input that covers a wide range of words.

5.2 Algorithm Design

The system consists of 5 modules:

1. English parser for parsing the English text
2. Sentences reordering module based on ISL grammar rules
3. Eliminator for eliminating stopwords
4. Stemming for getting the root words of each word and synonym replacement for words not in dictionary.
5. Video conversion module.

The input to the system is a written English text which is parsed to create a phrase structure based on its grammar representation. Then reordering is done to meet ISL grammar needs since, English text follows Subject-Verb-Object structure whereas ISL follows Subject-Object-Verb structure along with variation of negative and interrogative sentences. After which unwanted words are removed, as ISL will only those words which have meanings and all helping words like linking verbs, articles etc are not used. The output of which is sent to lemmatization module which reduces each of the words to its root form. The words not present in dictionary are replaced with their synonyms.

5.2.1 Solution Structure

1. Parsing of the Input English Text

To carry out rule based conversion of one language to another, grammatical structure of both the source and target language must be known. Parsing is the answer to acquiring this grammatical structure. Stanford parser is capable to produce three different outputs, part-of-speech tagged text, context free grammar representation of phrase structure and type dependency representation. The parser uses Penn tree tags for parsing the English sentence.



Parts of Speech Tagging

[Figure 2]

2. Grammar rules for conversion from English to ISL

Since both spoken language and sign language have different grammar rules. The complexity to translate them is increased many folds. Comparison between ISL and English grammar is listed below:

English Grammar	Indian Sign Language Grammar
English grammar is well structured and a lot of research work has been carried out to define the rules for it. English grammar follows the subject-verb-object order.	ISL is invented by deaf and a little work has been done to study the grammar of this language. The structure of sentences of ISL follows the subject-object-verb order[13].
English language uses various forms of verbs and adjectives depending upon the type of the sentence. Also, a lot of inflections of the words are used in English sentences.	ISL does not use any inflections (gerund, suffixes, or other forms), it uses the root form of the word.
English language has much larger dictionary	Indian sign language has a very limited dictionary, approximately 1800 words[10].
Question word in interrogative sentences is at the start in English	In Indian sign language, the question word is always sentence final
A lot of helping verbs, articles, and conjunctions are used in the sentences of English	In Indian sign language, no conjunctions, articles or linking verbs are used

Comparison of Grammar of English and Indian Sign Language

[Table-2]

Therefore ISL grammar rules, require all the verb patterns (20 patterns) being shifted after the corresponding noun occurrence. Some of the rule conversion are given in the following table

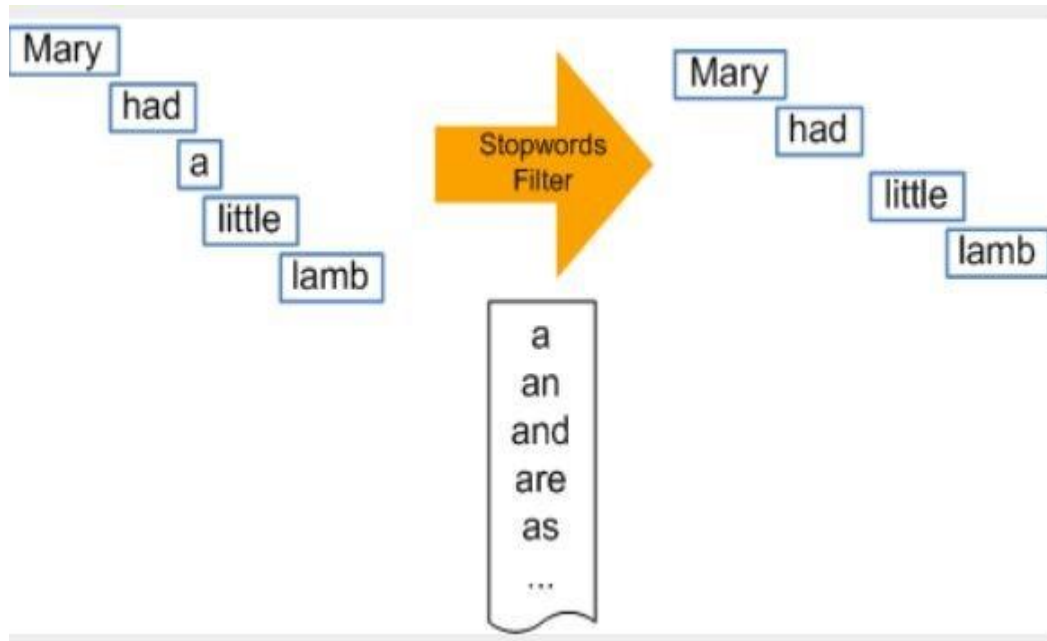
Verb Pattern	Rule	Input Sentence	Parsed Sentence	Output Sentence
verb+object	VP TO NP	Go to school	(VP (VB Go) (TO to) (NP (NN school)))	School to go
Subject + verb	NP V	Birds fly	(NP (NNS birds)) (VP (VBP fly))	Birds fly
subject + verb + subject complement	NP V NP	His brother became a soldier	(NP (PRP\$ his) (NN brother)) (VP (VBD became) (NP (DT a) (NN soldier)))	His brother a soldier became
subject + verb + direct object + preposition +prepositional object	NP V NP PP	She made coffee for all of us	(NP (PRP She)) (VP (VBD made)) (NP (NN coffee)) (PP (IN for) (NP (NP (DT all)) (PP (IN of) (NP (PRP us))))))	she coffee for all of us made

Examples of Grammatical Reordering of Words of English Sentence

[Table-3]

3. Elimination of Stop Words

Since ISL deals with words associated with some meaning, unwanted words are removed these include various parts of speech such as TO, POS(possessive ending), MD(Modals), FW(Foreign word), CC(coordinating conjunction), some DT(determiners like a, an, the), JJR, JJS(adjectives, comparative and superlative), NNS, NNPS(nouns plural, proper plural), RP(particles), SYM(symbols), Interjections, non-root verbs

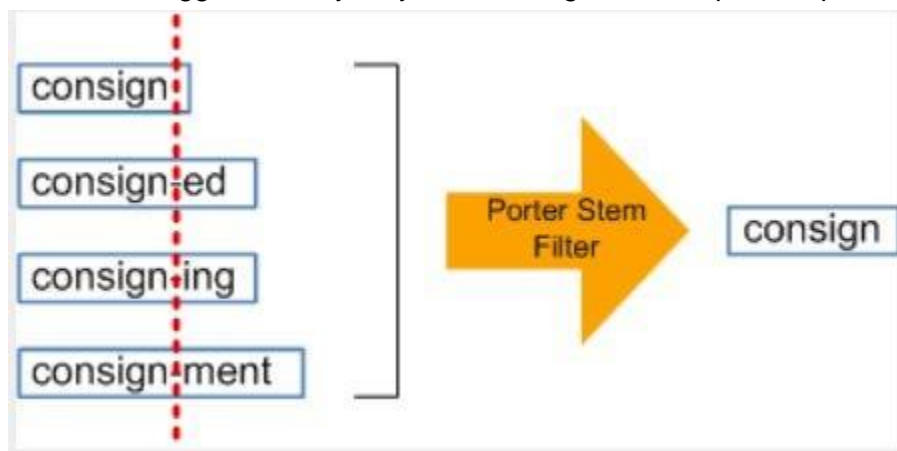


Removing Stopwords

[Figure 3]

4. Lemmatization and Synonym replacement

Indian sign language uses root words in their sentences. So we convert them to root form using Porter Stemmer rules. Along with this each word is checked in bilingual dictionary, if word does not exist, it is tagged to its synonym containing the same part of speech.



Stemming of Words

[Figure 4]

5. Video Conversion Stage

After the completion of above stages we get the ISL transformed text , the program will then find matches from the dataset available for each of the word. This will be based on the basic string matching algorithm between the processed input text and labels of videos. Finally a display of set of videos as a sequence one after the other can be seen on the screen.

5.2.2 Language

We will be using Python(version 3) for the development of the project. Also, we will be using many libraries of python for implementation of the project.

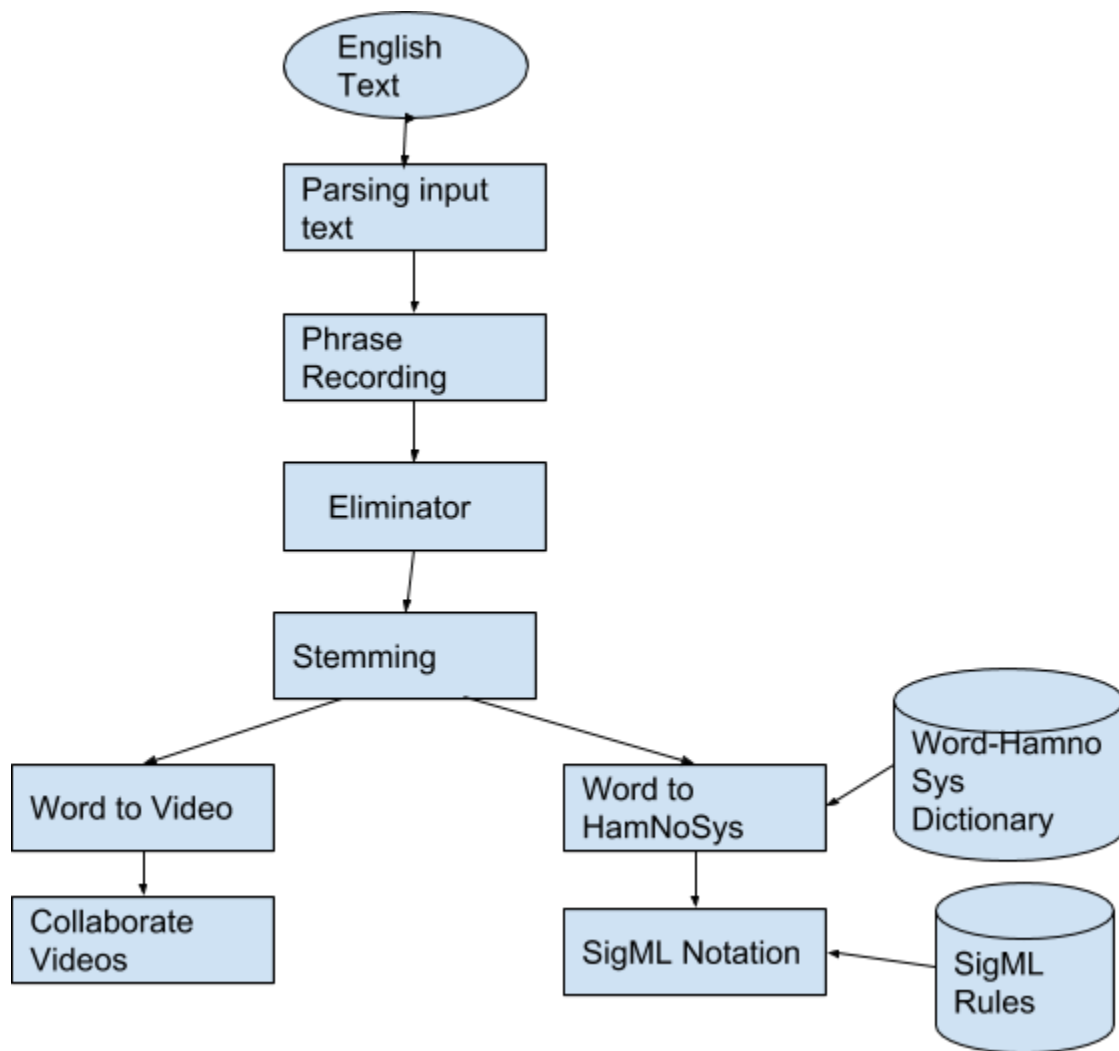
5.2.3 Tools/Libraries Used

- 1- Numpy
- 2- JASigning - Sigml file conversion into animated format
- 3- movies.py
- 4- nltk

5.3 Output Generation

The output of this module will be a movie clip of ISL translated words. The database will be having video for each and every separate words and the resultant video will be a merged video of such words.

5.1 Design Document and Flowchart



6. Data Analysis and Discussion

6.1 Output Generation

For a given english text the system aims at generating its equivalent sign language depiction. Our system generates these outputs in the following two ways:

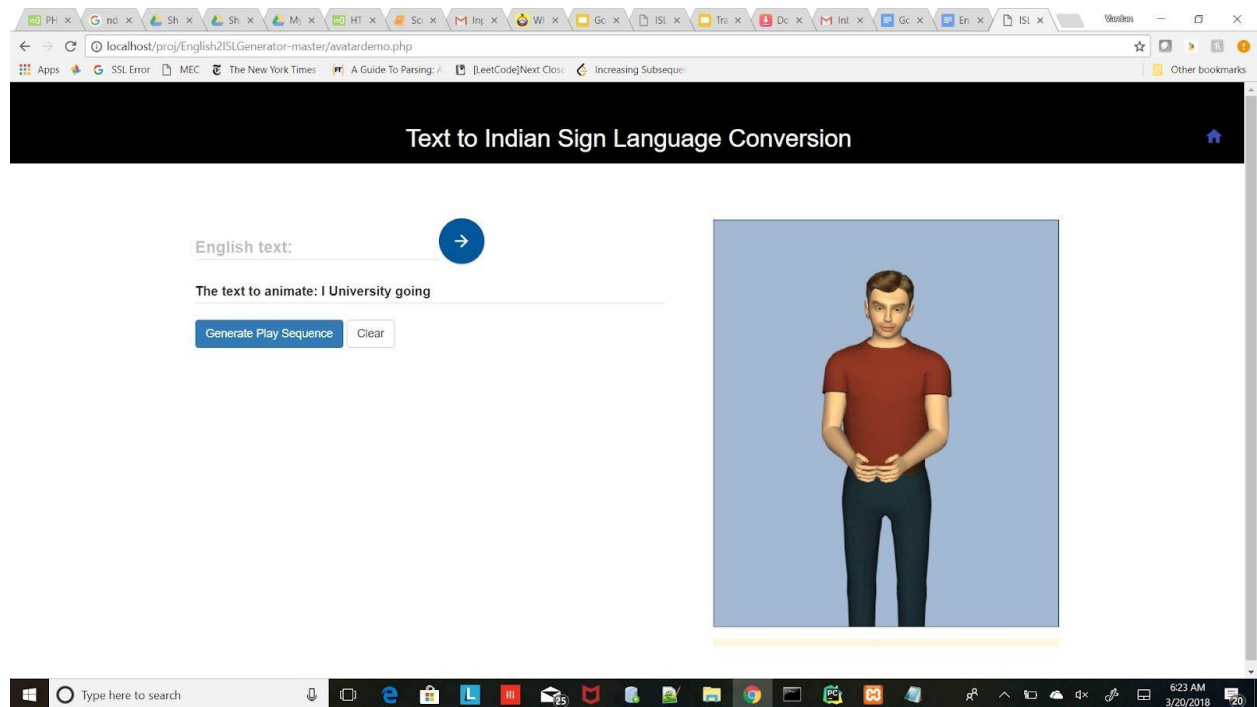
1. Video generation - Output from the ISL conversion phase of input sentence is passed to video generation phase wherein for each of the words in the sentence are looked up in the database for its corresponding video file and then these files are all concatenated to produce a more structured, informative and easy to understand visual depiction of Indian Sign Language.
2. Synthetic Animation generation - In this approach the ISL converted text is checked for its corresponding SigML file wherein these files are generated by a process called Hamnosys to SigML conversion to generate a markup language that is compatible with JASigning tool to generate equivalent synthetic animations.

Consider the following input sentence to which the generated ISL, video output and animation output are as follows:

English input text: I am going to the University.

ISL parsed text: I University going

Animation Generated Output:



6.2 Output Analysis

The given input text undergoes parsing using the StanfordParser to identify the subject, verb and object in the given sentences and which according to Indian Sign language grammar requires its verb and object positions to be swapped. This process fairly returns proper results and after this we require the system to lemmatize each of the words as sign language does not follow tense, however lemmatization follows a corpus and we chose wordnet corpus which not necessarily stem each of the verb forms, due to which there is a requirement to modify video and sigml files to support this issue. The generated output is then checked for stopwords and they are eliminated if present. After which the ISL sentence is then passed through two methods to generate sign language

Each of these methods were being analysed and we came up with the following analysis for each of them:

1. Video
 - o The process of finding the relevant videos and creating a collaboration of these videos is a time taking task.
 - o The video output couldn't ensure consistency due to the lack of research going on this field.
 - o All words not necessarily find their equivalent word file, thus such words are skipped.
2. Synthetic Animation

- The process of mapping the signml notations to generate animation is comparatively less time consuming
- This ensures consistency as the entire sentence is enacted by the same character throughout.
- However, these animations do not provide a realistic feel as is provided by videos.
- Finger spelling are used for those words which are not present in the database.

6.3 Compare Output Against Hypothesis

Our approach with the use of animation and light weight videos improves the overall performance by pre-processing the input text using NLP techniques and making use of the relevant term from the input sentence. Furthermore, we have used visual depiction for the entire input sentence and not for a particular word that is depicted in many of the previous projects. This ensures that our system is more robust and useful during live conversation. Videos and synthetic animations do provide a realistic appeal to the output. This system can be used readily and requires less prior knowledge about sign language.

6.4 Abnormal Case Explanation

Since much research hasn't been done on Indian Sign Language grammar, there are few words and types of sentences which still haven't found their respective translation.

Exclamation words like Oh! Alas! Or Hurray! don't have a direct translation in the Indian sign language, however these are words that are based on facial cues and expressions and thus need to be handled.

Imperative sentences such as "Book the flight", however they can be converted into ISL grammar format resulting in "Flight Book", they cannot be expressed in the imperative format owing to the sentence structure.

Also words having more than one parts of speech, example book has both verb and noun form cannot be handled well by the conversion system, since the original input sentence after being parsed and processed is converted into ISL grammar format which does not have a parsing structure, which indeed becomes difficult to identify the nouns and verbs from a given sentence, eventually causing the wrong form of the sentence being selected.

However, since less research has been carried out in this field, there are many pitfalls that needs to be worked on and some scenarios which needs to be thought of.

7 Conclusions and Recommendations

7.1 Summary and Conclusions

In this project, we have tried to develop a system that would be helpful for the disabled people having communication difficulties by developing a system that would help them in expressing themselves clearly and easily.

Our model successfully converts the entire input sentence into a single visual rather than depicting different words through a GIF/ picture giving the model a much realistic and lively appeal. Much more development on this track can be done as the ISL dictionary is still small and needs to grow eventually.

7.2 Recommendations for future studies

As the ISL is new and not much development has been done on this topic, many new videos for different words can be added to the dictionary to widen its scope and help people communicate better using this language. Moreover, text to speech integration can be performed in the project to enable better communication strategies which will allow users to convert text to Indian Sign Language with manually entering the input sentence.

8 References

- Regina Leven, Heiko Zienert, Thomas Hamke, and Jan Henning Siegmund Prillwitz.(1989). HamNoSys Version 2.0: Hamburg Notation System for Sign Languages: An Introductory Guide in Proceedings of International Studies on Sign Language and Communication of the Deaf. vol. 5, Hamburg, Germany.
- Ulrike Zeshan. (2003) . Indo-Pakistani Sign Language Grammar:A Typographical Outline. Sign Language Studies, vol. 3, no. 2, pp. 157-212,.
- E´va Safar Ian Marshall. (2003). A Prototype Text to British Sign Language (BSL) Translation.
- SystemProceedings of the 41st Annual Meeting on Association for Computational Linguistics - Volume 2 (ACL '03), vol. 2, pp. 113-116
- Matthew Huenerfauth. (2003). A Survey and Critique of American Sign Language Natural Language Generation and Machine Translation Systems.
- Computer and Information Sciences, University of Pennsylvania, Philadelphia, September.
- Pamela W.Jordan, John W.Benoit, Bonnie J.Dorr. (1999). A Survey of Current Paradigms in Machine Translation, Advances in Computers. vol. 49, pp. 1-68
- R. San-segundo , J. M. Montero , J. Macías-guarasa , R. Córdoba , J. Ferreiros and J. M. Pardo. (2004).
- Generating Gestures from Speech Xu, H, Abdel Rahman, S Jiang, M Fan, J.W. and Huang
- Y, (2011). November. An initial study of full parsing of clinical text using the Stanford

Parser. In 2011

- IEEE International Conference on Bioinformatics and Biomedicine Workshops (BIBMW) (pp. 607-614). IEEE.
- Yinhan Liu, Jiatao Gu, Naman Goyal, Xian Li, Sergey Edunov, Marjan Ghazvininejad, Mike Lewis, and Luke Zettlemoyer, (2020). Multilingual denoising pre-training for neural machine translation.