Audio Speech
To
Indian Sign Language
Conversion

Made By:
Kiran Dapkar (170050020)
Ansh Verma (170050014)

Guided By:
Prof. Preethi Jyoti
IIT Bombay
Course: Automatic Speech Recognition (CS 753)

Problem Statement:

To design an architecture for differently abled people with hearing impact to convert an audio speech as an input from the user into the Indian sign language

Abstract:

Deaf people always miss out on the fun that a normal person does, may it be communication, playing computer games, attending seminars or video conferences etc. So, Sign language is a natural way of communication for challenged people with speaking and hearing disabilities. The aim of our project is to develop a communication system for the deaf people. The scarcity of any sign language corpus caused lesser development of sign language conversion systems. This field has been much more focused on American sign language(ASL). Our project converts the audio message into the text and displays the relevant Indian Sign Language Videos/GIFs of Avtar.

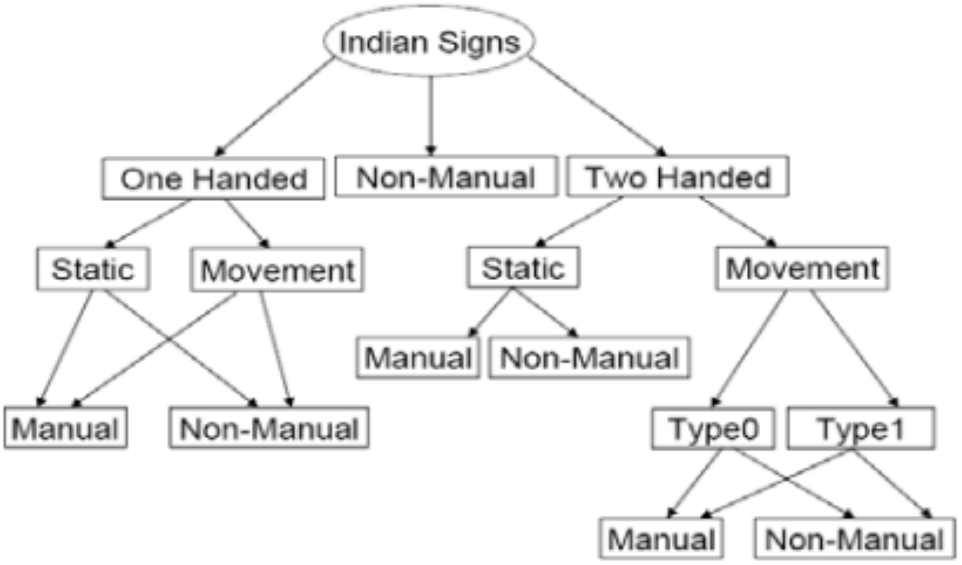
About Sign Language:

Sign language is a natural language which has some facts which the people are not aware of. Some of the facts of the sign language are:

- NOT the same all over the world.
- NOT just gestures and pantomime, but do have their own grammar.
- Have a much smaller dictionary than the other spoken natural languages.
- Finger-spelling for the unknown words.
- Words may be joined e.g. to represent dinner, one might show the sign of Night and then Food.
- Most of the sign languages put the adjective after the noun e.g. Car Red.
- Never use am/is/are/was/were/ (linking verbs).

- Never use word-endings/suffixes.
- Always sign in the Present Tense.
- Do not use articles. (a, an, some, the).
- Do not use I, but use me.
- WH-questions are at the END e.g. “You go where?”
- Have no gerunds. (-ing).
- Use non-manual expressions as well e.g. use of eyebrows, eyelids, facial expressions, head and shoulders movement.
- NOT been invented by hearing people.

ISL Type Hierarchy:



Methodology:

1. Speech is taken as input through a microphone and also as a text.
2. Used JavaScript Web speech API for speech recognition i.e. *webkitSpeechRecognition*.
3. The text is then pre-processed using NLP (Natural Language Processing).
4. Lastly, playing a video of avtar that displays sign language for each particular word from the parsed text.

Algorithm Overview:

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 the dictionary.

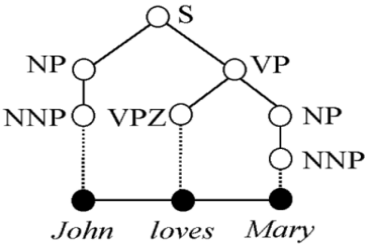
1. Parsing of the Input English Text:

To carry out rule based conversion of one language to another, grammatical structure of both the source and target language must be known. This can be done using Parser. So, we used a Stanford parser which is capable of producing three different outputs, part-of-speech tagged text, CFG (context free grammar) representation of phrase structure and type dependency representation. Reason behind using stanford parser is because this parser provides universal dependencies. It also creates parsing trees, namely Penn trees with tags and which is used to determine the relationship between words and sentences rather than a normal syntactic parser.

Penn trees example:

Sentence: John loves Mary

(S (NP (NNP John))
(VP (VPZ loves)
(NP (NNP Mary)))
(. .))



Source: wikipedia

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 has increased many folds. Comparison between English Grammar and ISL grammar is given below:

English Grammar	ISL Grammar
English grammar follows the subject-verb-object order.	The structure of sentences of ISL follows the subject-object-verb order.
English language uses various forms of verbs and adjectives depending upon the type of the sentence.	ISL does not use any inflections (gerund,suffixes, or other forms), it uses the root form of the word.
Question word in interrogative sentences is at the start in English	In Indian sign language, the question word is always sentence final
English language has much larger dictionary	Indian sign language has a very limited dictionary, approximately 1800 words
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

Therefore ISL grammar rules require all the verb patterns (20 patterns) [3] being shifted after the corresponding noun occurrence. The parsed sentence is the input to this module where the noun phrase and the prepositional phrase are freezed but if there is any verb phrase

present in the sentence, it is checked recursively because the verb phrase may further composed of noun phrase, prepositional phrase, verb phrase or even the sentence. Some of the rules of conversion are given in the 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 (VBD 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 + indirect object + direct object	NP V NP NP	I lent her my pen	NP (FW i)) (VP (VBD lent) (NP (PRP her)) (NP (PRP$ my) (NN pen)))	I her my lent pen
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

3. **<u>Elimination of Stop Words:</u>**

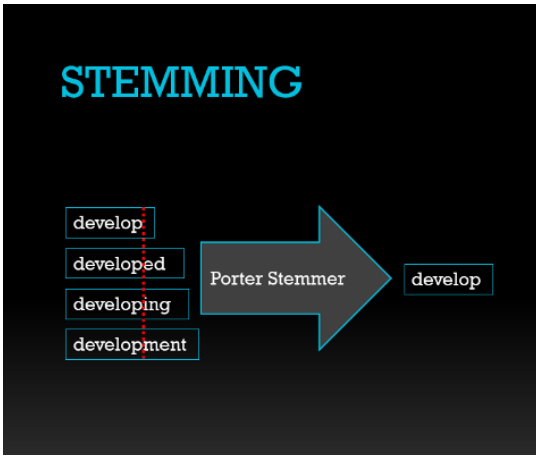
Indian sign language sentences are formed of main words. All the words like linking verbs, suffixes, articles are not used. After applying the grammar rules, the ISL sentence is generated of which all the unwanted words are required to be removed. The parts of speech which are not the part of ISL sentences are detected and eliminated from the sentence. Out of 36 POS tags, the various part of speech which do not form the part of ISL sentence are 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. These above mentioned unwanted words are removed from the ISL sentence.
Below are the examples in which unwanted words are removed.

Input English Sentence	**Sentence after reordering**	**Output after elimination**
Go to school	School to go	School go
she made coffee for all of u	She coffee for all of us made	She coffe all us made
we are waiting for suresh	we for suresh are waiting	We suresh waiting

4. **Lemmatization:**

Indian sign language uses the root words in their sentences. All the words used must not contain suffixes, gerund or it should not be an inflexion of a word. If a word in the ISL sentence is not a root word, it is converted into the root word after passing it to the stemmer and applying lemmatization rules. The porter stemmer is used for stemming.
Ex: laughing -> laugh



Dataset:

We came across lots of data majorly showing images of every alphabet of text. Some dataset(website https://indiansignlanguage.org/) contains videos of words, which requires manual dataset creation and labelling which is not possible due to time constraint, and also, it is high memory consumption. Since, we found a little dataset of around top 150 words which is created using blender animation application.

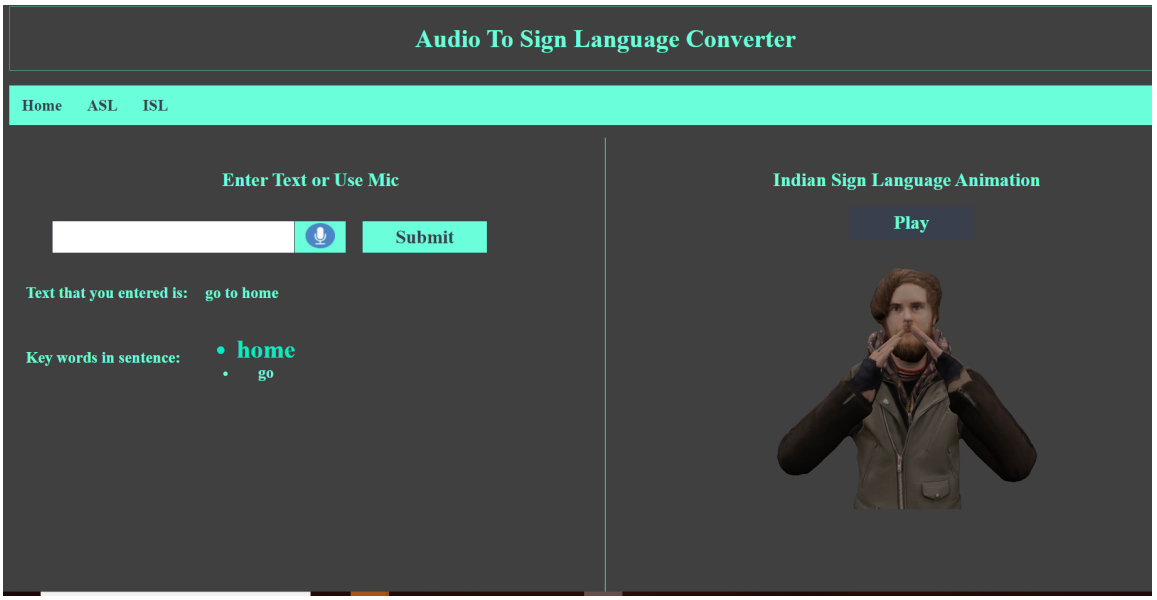
Deployment:

We used Django framework to deploy our project
Used HTML, CSS, JS for front-end
Specifications:

-we display both ASL as well as ISL avatars
-input can be given through speech or can be typed
-processed text and words shown side-by-side
-those words not present in vocab represented by dividing them into single alphabet

Demo:

The below figure displays an output for sentence “Go to home” with parsed text and reordering grammar it displays “home go” in Indian sign language.



Future Work:

- Dataset can be created using blender tool or any other tool to cover more vocabulary
  To animate the sentence, you can use an animation tool SiGML Player [4]. To generate the animation through this tool, the input must be the tags of SiGML(Signing Gesture Markup Language) and the output of the tool is animated characters.

- Performance of ISL algorithm can be improved

- Continuous Conversion without pause can be developed. It can be used to convert live communication ex: Display of live news with side-by-side sign language avatar.

References used:

1. Audio to Sign Language Translation for Deaf People
   https://www.ijeit.com/Vol%209/Issue%2010/IJEIT1412202004_05.pdf

2. English Text to Indian Sign Language Translator
   https://www.cse.scu.edu/~mwang2/projects/NLP_English2IndianSignLanuguage_18w.pdf

3. Wren, P. C., & Martin, H. (1999). High school English grammar and composition. From (page number 663-674)

4. JASigning. (2015). Retrieved june 15, 2016, from http://vh.cmp.uea.ac.uk/index.php/JASigning

My github repo:

https://github.com/kirandapkar/Speech-to-Indian-SIgn-Language-Conversion