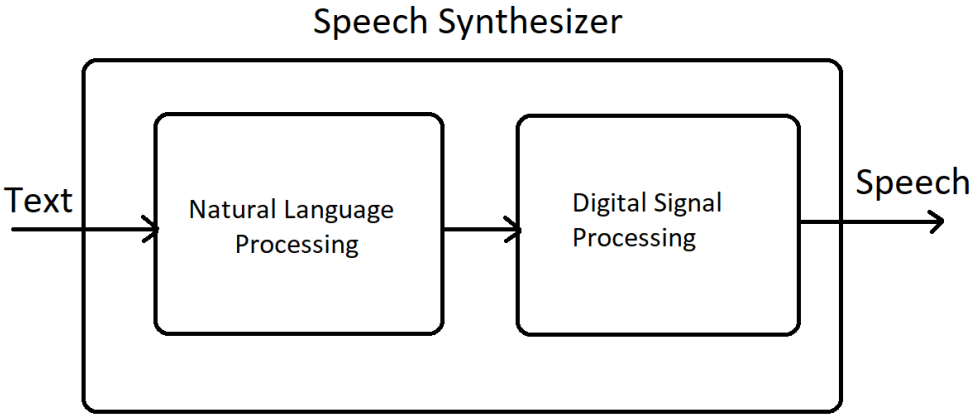
Text-to-Speech

Name: Manasi Manoj Tayade
UID: U01735870
Instructor Name: Altion Simo
Class: Artificial Intelligence CS-627

Introduction

Text-to-speech (TTS) is the mechanism by which words are translated into verbal audio form. The program, tool, or software takes an input text from the user and understands the linguistics of the language being used methods of natural language processing and performs logical inference on text. This processed text is moved into the next block where the processing of digital signals is carried out on the processed text. This processed text is eventually transformed using several algorithms and transformations. The process of translating text input into audio data is called *synthesis* and the output of synthesis is called *synthetic speech*.



NLP:
Natural Language Processing, also shortened to NLP, is a subfield of artificial intelligence (AI). It helps machines process and understand the human language in any given context so that they automatically can carry out repetitive tasks such as machine translation, summarization, ticket classification, and more.

DSP:
Digital Signal Processors (DSP) take real-world signals like voice, audio, video, temperature, pressure, or position that have been digitized and then mathematically manipulate them. A DSP is designed for performing mathematical functions like "add", "subtract", "multiply" and "divide" very quickly. Signals need to be processed so that the information that they contain can be displayed, analyzed, or converted to another type of signal that may be of use.

Installation of gTTS module:

- The installation of the gTTS module is simple and can be done using the following command in the command prompt terminal —

```
!pip install gTTS
```

- We can use gTTS module to convert our typed text into a speech converted output. Open the python file and give it a name of your choice and make sure it ends with the .py format.

- IPython.display lets you play audio directly in an IPython notebook.

```
from gtts import gTTS #Import Google Text to Speech
from IPython.display import Audio #Import Audio method from IPython's Display
Class
```

```
text = str(input('Enter a sentence'))
accent = 'en'
tts = gTTS(text=text, lang=accent, slow=False)
tts.save('10.wav')
Audio('10.wav', autoplay=True)
```

Here,

- The **text** variable is a string used to store the user's input. The text can be replaced by anything of your choice within the quotes. Another alternative can be to use the input statement for the user to type their own desired input each time the program is run.
- The **tts** variable is used to perform the Google text-to-speech translation on the user's input. The output of the converted text is stored in the form of speech in the tts variable.
- The **tts.save** function allows us to save the converted speech in a format that allows us to play sounds. I have saved it in a file called 10 and, in a format, called .wav. Other formats like .mp3 format can also be used.
- **Audio** provide bindings for the PortAudio library for cross-platform playback of WAV files.

gTTS in Python:

gTTS (Google Text-to-Speech) is a Python library and CLI tool to interface with Google Translate text-to-speech API.

Conclusion:

It is also possible top use the gTTS module extensively in other languages such as French, German, Spanish, Hindi, etc. If there is a communication obstacle, this is incredibly useful, and the user is unable to communicate his messages to people. Text-to-speech is a great aid for people with visual disability or people with other disabilities, as it will help them by converting text to speech.