**ASA Chat Bot**

**With Speech Recognition**

**for Luna Cat**

## Luna\_Cat\_Chat\_Main

A programme created to become the central system of the Luna Cat Chat Bot

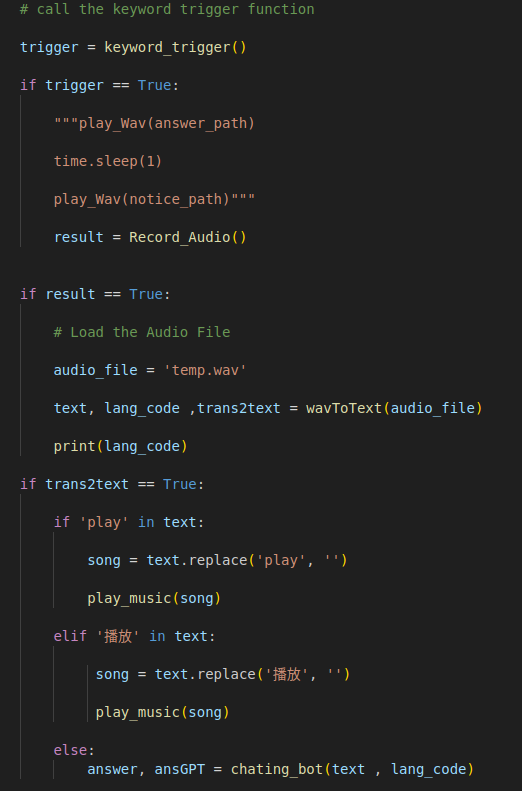
When called, the programme will wait for the users’ input and execute further with the users’ input.

This programme will play some audio as return and keep looping.

**Code Explain:**

****

This function uses the python package - speech\_recognition to check the users’ audio input. If the keyword is in the audio file, the function will trigger and return True.



With the keyword trigger, the programme will start to record the users’ input and turn the audio data to text with the help of the “wavToText” function and return the text to the programme.

The programme then will analyze the text return and decide which function it needs to execute.

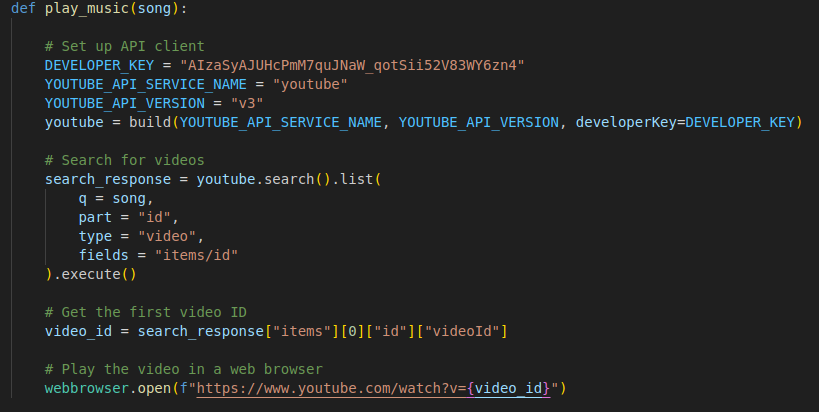
If the users’ input has the keyword play, the programme will execute the “play\_music” function, Else, the programme will execute the “chating\_bot” function.

## Play\_VT

This function is created for playing the video or music from Youtube with the users’ input.

This function required the Google Cloud Api key and a Youtube Premium Subscription to execute.

**Code Explain:**

****

Using Youtube Data API from Google Cloud to search the url link with the users’ input.

The programme will pop a web browser to play the searched video.

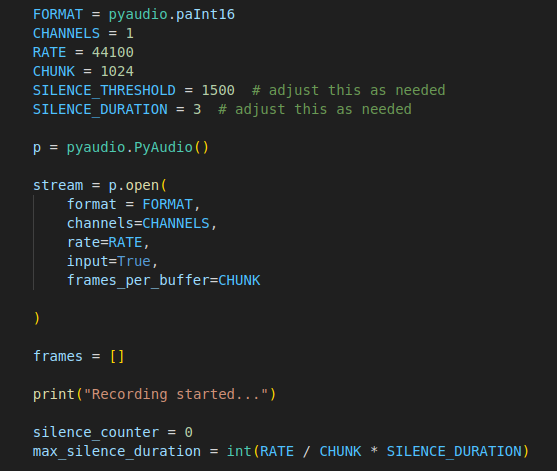
## Record\_Speech

This programme is created to record the audio from the user.

This programme will keep recording the sound until the user keeps silent for the whole 3 seconds.

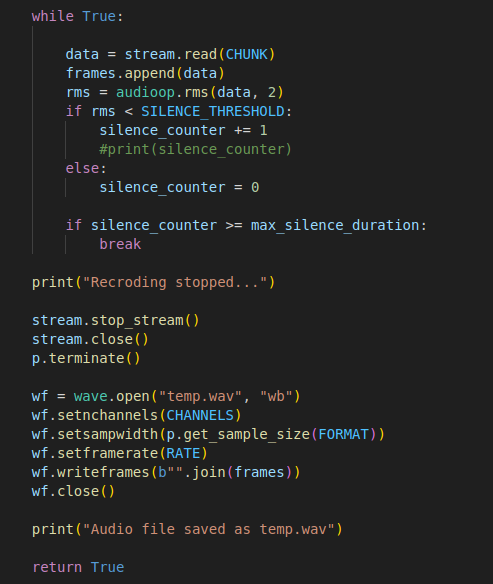
This program saves and reads the file in Wav format.

**Code Explain:**



The above image is the setting of the recording environment.

Parameters silence\_threshold and silence\_duration can be customized for different needs. Parameter “SILENCE\_THRESHOLD” represents the minimum hz the recording system will accept. And parameter “SILENCE\_DURATION” represents the maximum duration that the recording sound is below the threshold.



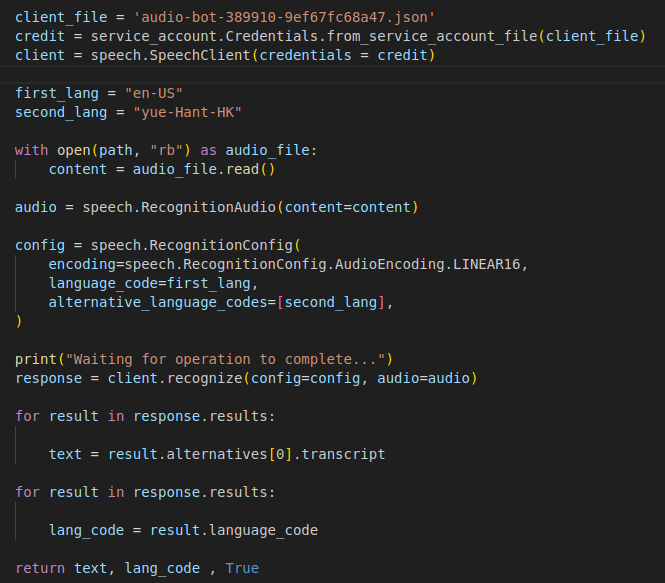
The program will record the users’ input and after the counter hits the maximum silence duration. The program will save the audio file as temp.wav for further instruction.

## SpeechToText

This program create to translate users’ recorded audio to text for further instruction.

This program can detect English input and also Cantonese input.

**Code Explain:**



Remember to check the location of the Api client json file.

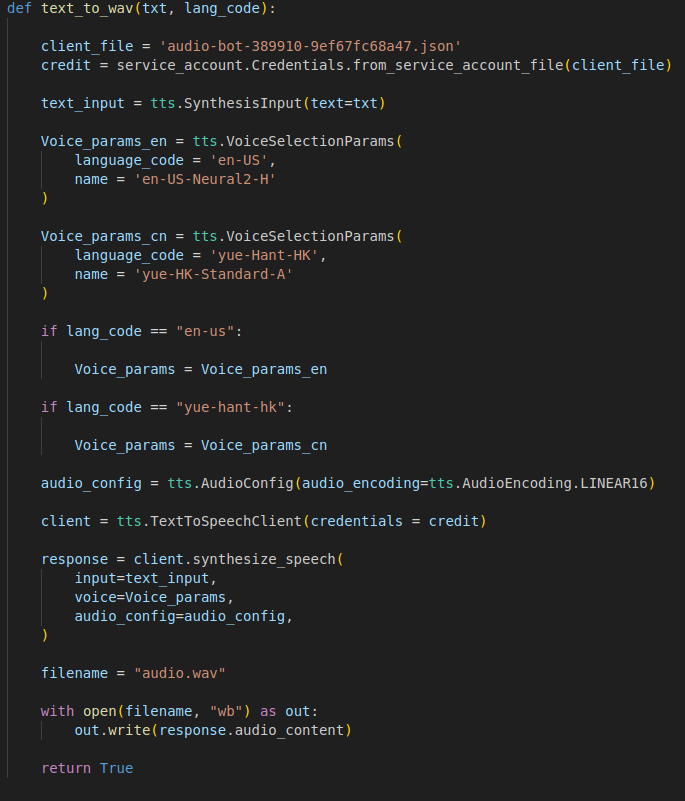
## Text\_to\_Speech

This program was created to turn the text input to audio file.

The Audio file will be in .WAV format.

With the language code given by the speech to text, this program will generate an audio file with the help of google cloud texttospeech api.

**Code Explain:**

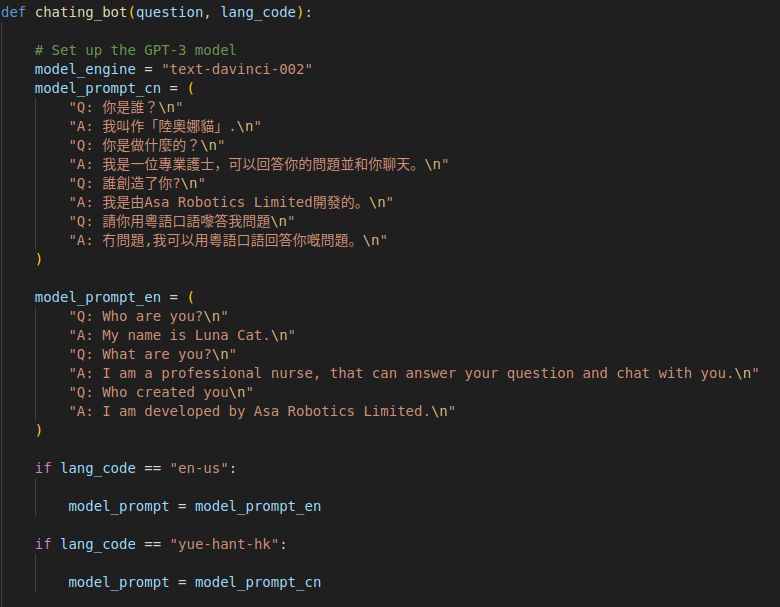


## GPT\_Chat

This program was created to provide a communication function with the users’ input.

This program uses the ChatGPT to answer the users’ questions.

**Code Explain:**



Input the configuration for the ChatGPT to answer the question with certain roles.

With the given language code by the program “SpeechToText”, the program will answer users’ questions in a certain language.

## Further Development Direction

* Play the video from Youtube with a external media player

To execute the function on Luna Cat, the program “PlayYT” has to use a media player to play the video instead of using a web browser pop up window.

* Using Vits text to speech for more humanize voice

To avoid the audio answer return too robotics, using a better texttospeech module to produce a more humanize audio file is the solution

<https://www.youtube.com/watch?v=riYOD_EFKDE>

* ChatGpt Configuration

Update the ChatGPT configuration for the AI robot for a better and humanize answer return

* Keyword detection improve

For the Cantonese keyword detection, the accuracy is not great enough. For improvement, may be we need to apply AI training in the detection to improve the performance.