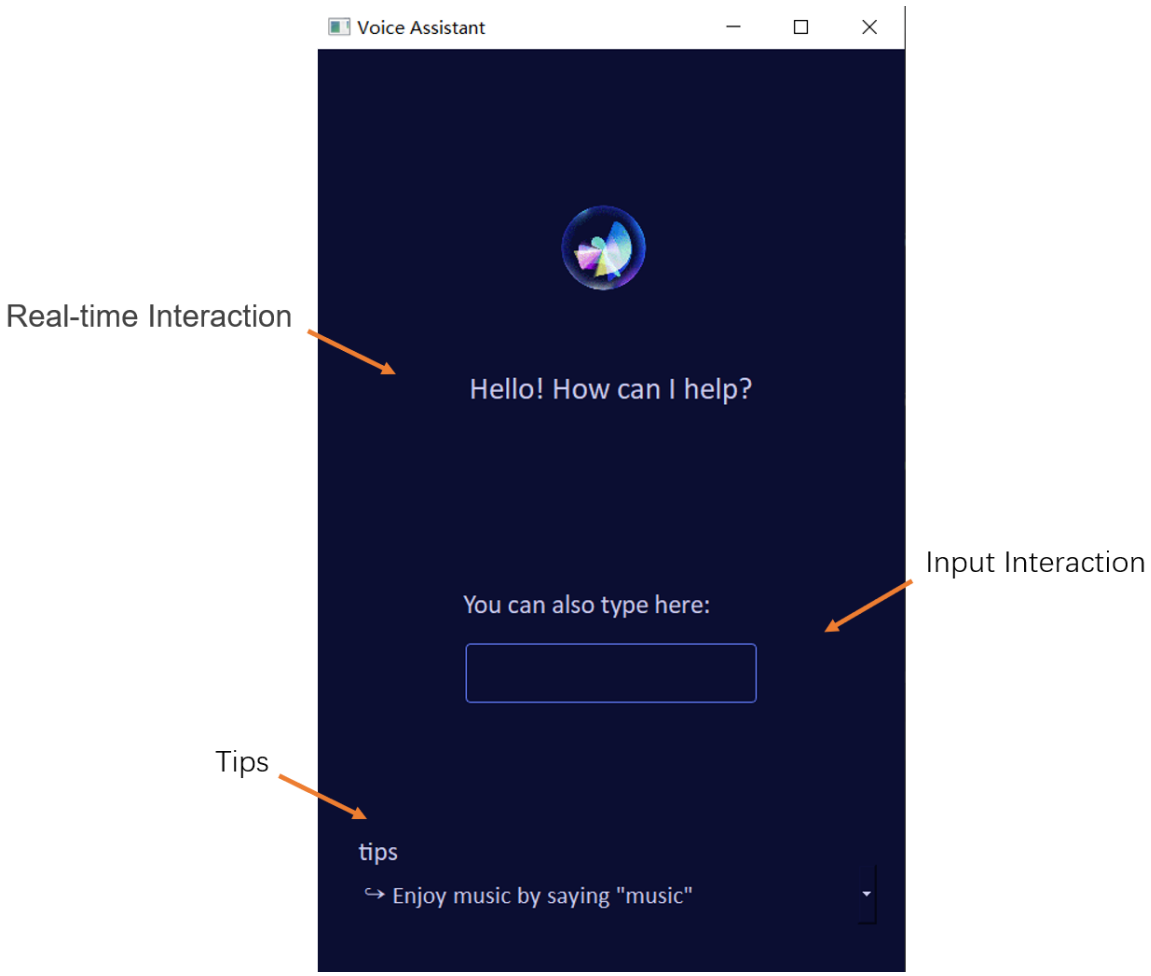
Assignment 1: ASR

UI



Added Functions

1. Enjoy music by saying "music"
2. Take some notes by saying "note pad"
3. Browse a web page by saying "browser"
4. Change the volume by saying "up" or "down"
5. Switch music by saying "next" or "previous"

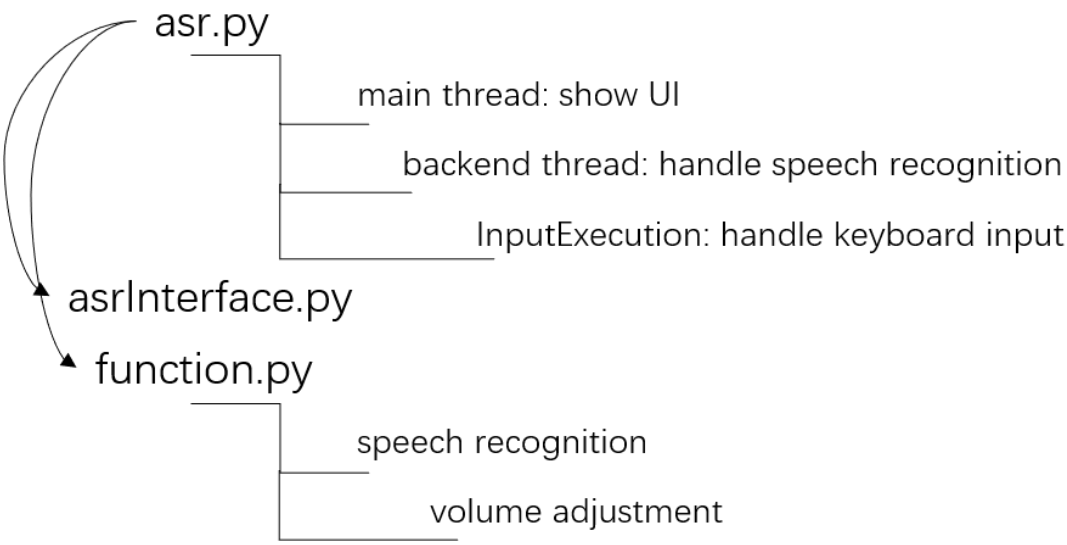
Modifications to GUI and Codes

GUI

1. Replace the original voice icon with a new one.
2. Remove the instructions which were placed in the middle of the interface to ensure enough space, enhancing the focus on interaction.
3. Add an input box to provide an alternative to interact.
4. Put tips in the form of ComboBox at the bottom of the interface .

Codes

Structure



Illustration

There are three .py files in all——asrInterface.py for UI, function.py for functions of speech recognition and volume adjustment, asr.py importing the first two .py for executing the UI app.

There are to threads in asr.py. The main thread runs app.show() and sys.exit(), while the backend thread runs the program of speech recognition. A class called InputExecution is responsible for handling the keyboard input, providing the same funcitons as that of speech recognition.

Accuracy and Improvement of Speech Recognition

It's a shame that the package speech_recognition doesn't work well.

To improve the accuracy:

1. Make sure that there is no noise in the speaking environment and the pronunciation is correct(online dictionary pronunciation recommended).
2. For this code, convert `recognition=='music'` into `recognition.find('music')`
3. Change API, such as iFlyTek, Baidu or Tencent.

How to run

Environment: python 3.8.8

1. Decompress the file.
2. Install the concerned packages if needed.
3. Terminal run: python asr.py