Lab1: Automatic Speech Recognition

@author 1851055 Mingjie Wang

Installation

- To run the code, you should install `conda` environment by the following:

  ```
  pip install -r requirements.txt
  ```

  After installing the essential package, you should replace the file `pronounciation-dictionary.dict` in `D:\Anaconda\envs\hci-asr\Lib\site-packages\speech_recognition\pocketsphinx-data\en-US`

- Then you can simply run the code:

  ```
  python main.py
  ```

Modifications to GUI

To make user use it more conveniently and easily, I have carefully modified the original user interface as follows:

```
self.introLabel = QLabel(self)
self.introLabel.setFont(self.font)
self.introLabel.setText("Hi, what can I do for you? You can start from saying 'help
me'.")
self.introLabel.setWordWrap(True)
self.introLabel.move(130, 100)
self.introLabel.setFixedWidth(300)
self.introLabel.setMinimumHeight(30)
self.introLabel.setStyleSheet("background-color: #00CCCC; "
                              "color: white; "
                              "font-size: 18px; "
                              "border-radius: 10px; "
                              "padding: 10px;"
                              )
self.introLabel.adjustSize()

'''
初始头像
```

```
'''
self.introProfile = QLabel(self)
self.introProfile.resize(100, 50)
self.introProfile.move(50, 100)
self.introProfile.setStyleSheet("background-color: transparent;"
                                "background-image:url(icon/root.png);"
                                "background-repeat: no-repeat;")

'''
用户头像
'''
self.userProfile = QLabel(self)
self.userProfile.resize(150, 50)
self.userProfile.move(720, 100 + self.introLabel.height() + 25)
self.userProfile.setStyleSheet("background-color: transparent;"
                               "background-image:url(icon/user.png);"
                               "background-repeat: no-repeat;")

self.speaker = pyttsx3.init()

# 说话状态
self.isSpeaking = False
self.speakingGif = QMovie("icon/speaking.gif")

# 对话框按钮
self.speakMovie = QMovie("icon/speak.gif")
self.speakMovie.setScaledSize(QSize(120, 120))
self.speakLabel = MyQLabel(self)
self.speakLabel.setMovie(self.speakMovie)
self.speakLabel.setStyleSheet("background-color:transparent;")
self.speakLabel.resize(120, 120)
self.speakLabel.move(370, 430)
self.speakMovie.start()

'''
回复头像
'''
self.responseProfile = QLabel(self)
self.responseProfile.resize(100, 50)
self.responseProfile.move(50, 260)
self.responseProfile.setStyleSheet("background-color: transparent;"
                                   "background-image:url(icon/root.png);"
                                   "background-repeat: no-repeat;")
```

In above code, I have tried 5 main ways:

- Redesign the interface layout to make the overall structure more compact
- Use **bubble dialog box** to make the information look more prominent
- The use of **voice output library** can obtain information more effectively
- When voice input, new **threads** are used to prevent interface locking

- Use `QSS` to modify the UI more flexibly

Accuracy

To improve the accuracy of recognition, I use the following ways:

- Edit pronunciation dictionary

  In our system, we'll only a few words like `weather`, `file` and so on. So we can modify the `pronounciation-dictionary.dict` which is used for recognizing our words:

  ```
  music M Y UW Z IH K
  play P L EY
  open OW P AH N
  weather W EH DH ER
  help HH EH L P
  text T EH K S T
  file F AY L
  me M IY
  a AH
  ```

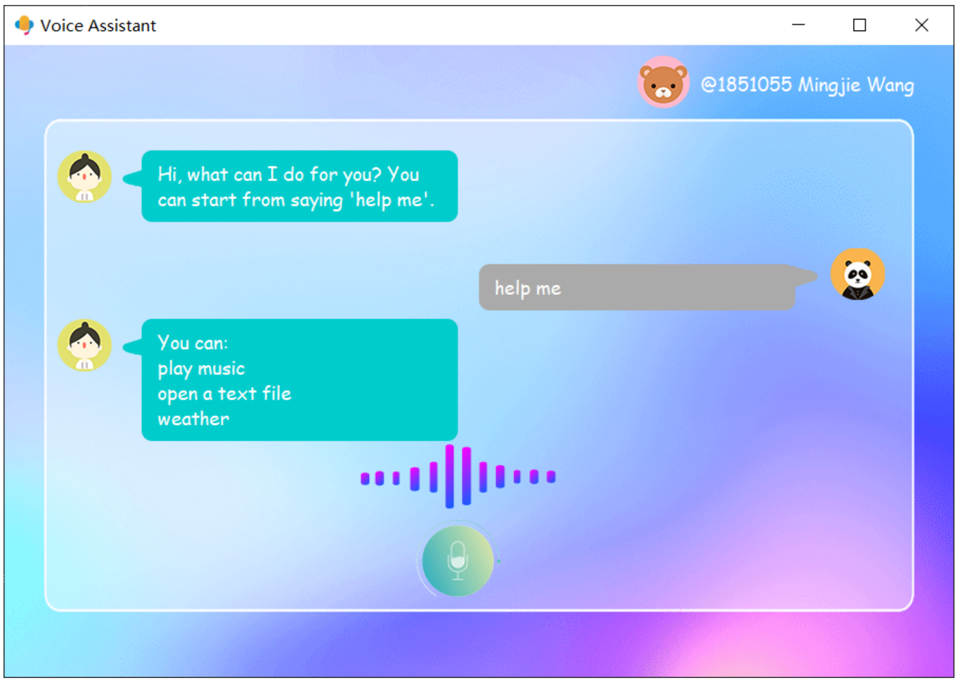
- Compare to find the most likely instruction

  It is a simple way by comparing the input with each of the instruction:

  ```
  maxScore = 0
          maxIndex = -1
          for index, item in enumerate(self.supportCommand):
              score = self.calculateSimilarity(self.myCommand, item)
              # 分数过低则不考虑
              if score <= 0.2:
                  continue
              if score > maxScore:
                  maxScore = score
  ```

By the following two ways, the accuracy can be nearly 100%.

Interface

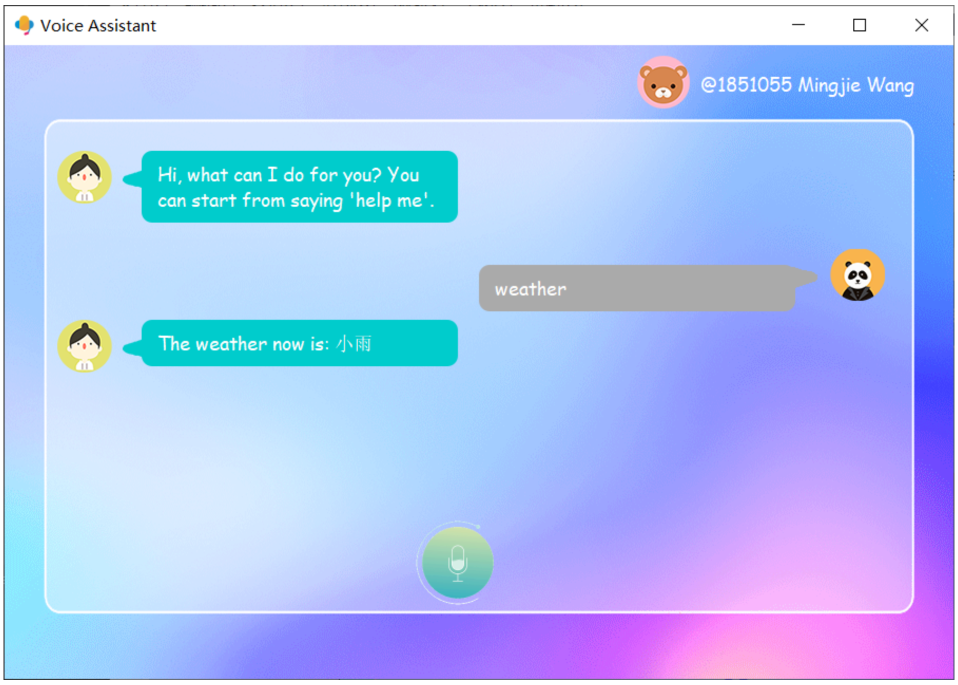

Voice Assistant

@1851055 Mingjie Wang

Hi, what can I do for you? You can start from saying 'help me'.

help me

You can:
play music
open a text file
weather


Voice Assistant

@1851055 Mingjie Wang

Hi, what can I do for you? You can start from saying 'help me'.

weather

The weather now is: 小雨