

Free dialogue

Free dialogue

1. Experimental Purpose
 2. Experimental steps
 3. Analysis of the main source code of the experiment
 4. Experimental summary
- Notes

1. Experimental Purpose

This chapter will teach you how to combine the online large language model to realize the process of free conversation with the robot dog. This case only requires audio equipment.

Notes:

1. **Before running this case, you need to close the startup program**, please refer to the Raspberry Pi system configuration chapter **9. Close and open the APP/handle control program** This tutorial ends the startup program.
2. You need to fill in the API_KEY of the large model, please refer to the operation method of **AI large model chapter "1. Prerequisites for using the large model"**

2. Experimental steps

(This tutorial takes the Chinese version of the effect diagram as an example)

1. Terminal input

```
cd /home/pi/DOGZILLA/Samples/4_Big_Modle
python3 Free_QA/chatgpt_main_en.py
```

2. Wake-up operation
Before wake-up:

The wake-up word is: **hello yahboom**

```
pi@raspberrypi:~/DOGZILLA/Samples/4_Big_Modle $ python3 Free_QA/chatgpt_main.py
serial /dev/myspeech open
Waiting for keyword...
```

After wake-up, you will hear a "ding" sound, and then you can express your thoughts to the robot dog.

```
pi@raspberrypi:~/DOGZILLA/Samples/4_Big_Modle $ python3 Free_QA/chatgpt_main.py
serial /dev/myspeech open
Waiting for keyword...
Keyword detected: 05.Jun 2025 07:53:58
Playing WAVE './ding.wav' : Signed 16 bit Little Endian, Rate 16000 Hz, Mono
```

3. The robot dog will be processed by the big model to understand the customer's expression of ideas, and then feedback the corresponding text results and audio playback.

```

###speak iat closed ###
Q: Who are you?
A:I am Qwen, a large language model developed by Alibaba Cloud. I can help you with various tasks such as answering questions, generating text, solving problems, and more. I'm designed to assist with a wide range of topics, including technology, science, culture, humanities, and more. How can I help you today?

```

4. At this point, the interactive process ends. If you need to express your ideas again, just wake up again.

At the same time, in this case, the previous interactive dialogue can be interrupted by waking up.

3. Analysis of the main source code of the experiment

In the path "/home/pi/DOGZILLA/Samples/4_Big_Modle/Free_QA/chatgpt_main_en.py" is a main function entry

```

# Main function flow
while True:

    if detect_keyword():

        os.system("pkill mplayer")
        time.sleep(.2)

        start_recording()
        time.sleep(1)

        content = rec_wav_music_en()

        if content != "":
            print("Q:"+content)
            re = QAGPT_en(content+"English reply") #English

            print("A:"+re)
            try:
                response = re
                tts_thread = threading.Thread(target=Speak_Vioce)
                tts_thread.daemon = True
                tts_thread.start()
            except:
                pass
        if content == 0:
            break

    time.sleep(0.1)

```

1. Program flow: detect wake-up word->listen to expression semantics->understand with big model->feedback answer

2. Ultra_gpt: This is the interface combined with big model. The big model used in this function is **OpenRouter**

If you want to change other big models, you can refer to **xinghou_UltraAPI.py**, or write a py file according to the source code provided by the big model you want to use. For example, if the file name is: **mychatgpt.py**, then introduce this file at the beginning of chatgpt_main_en.py. The writing method is: from mychatgpt.py import *, and then replace the Ultra_gpt interface provided by the routine with the interface function of the big model

that can be called in mychatgpt.py. There are many DIY operations involved here, and it is not recommended for novices to change the model.

3. If you want to change the threshold for recording start and the duration of recording, you can change this file. Enter in the terminal

```
nano /home/pi/DOGGZILLA/Samples/4_Big_Modle/audio.py
```

Change the recording part of this file as shown in the figure below

```
quitmark = 0
automark = True
def start_recording(timer = 3, save_file=SAVE_FILE):
    global automark, quitmark
    start_threshold = 3000 #30000 Start threshold
    end_threshold = 1500 #2000 Stop threshold
    endlast = 15 # The total number of times detected is lower than the stop threshold, and the recording is automatically stopped
    max_record_time = 10 #The maximum recording duration
    CHUNK = 1024
    FORMAT = pyaudio.paInt16
    CHANNELS = 1
    RATE = 16000
    WAVE_OUTPUT_FILENAME = save_file
```

Parameter meaning:

- start_threshold = 3000 #Start recording when a sound louder than this value is detected. This value changes according to the environment
- end_threshold = 1500 #Sounds lower than this value are detected. This value changes according to the environment
- endlast = 15 #Stop recording when the number of sounds lower than end_threshold is detected. Here it is 15 times
- max_record_time = 5 #The duration of the recorded audio. Here it is 5

Note: start_threshold must be greater than end_threshold (start_threshold>end_threshold)

In general, the ideal value of end_threshold is half of start_threshold, which can be adjusted according to your own environment.

4. If you feel that the recorded audio cannot be recognized by the online large model because the sound is too small, you can adjust the value here to amplify the recorded audio.

Terminal input

```
nano /home/pi/DOGGZILLA/Samples/4_Big_Modle/audio.py
```

```
269 wf.close()
270 print(f"The recording has been saved as: {WAVE_OUTPUT_FILENAME}")
271
272 amplify_audio_librosa("recorded_audio.wav", "recorded_audio.wav", gain_factor=5.0) #放大它 Enlarge it
273
274
```

Here it is amplified 5 times, here you can make an adjustment according to the distance of the sound source.

Note: If the distance is too far to record audio at all, adjusting the parameters here will be meaningless.

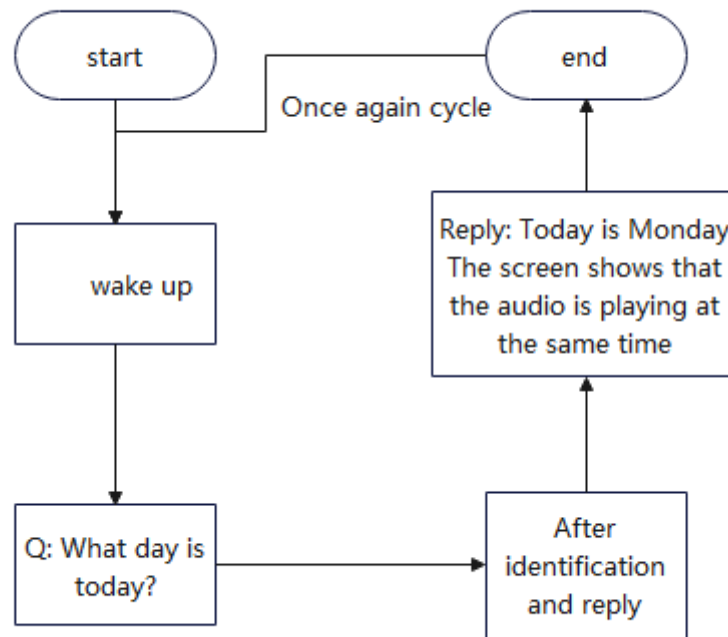
It is recommended that the distance from the sound source of the recorded audio should not be greater than 1.5m.

5. In the path of "/home/pi/DOGGZILLA/Samples/4_Big_Modle/Free_QA/" Directory structure description

```
|— chatgpt_main_en.py # English moderator program entry
|— chatgpt_main.py # Chinese moderator program entry
|— xinghou_speak_iat.py # Spark speech recognition interface file
|— xinghou_tts.py # Spark speech synthesis interface file
|— xinghou_UltraAPI.py # Spark large model interface file
```

4. Experimental summary

Through the above description, the flowchart of this case is as follows:



If you don't know what this case means, here are some reference examples

Example:

1. Location + time + real-time status? For example: What is the weather like in New York today? / What is the news today? / What day is it today? Etc.
 2. Request + content type For example: Tell me a bedtime story
 3. You can also ask some knowledge questions For example: What is 100+100?
- There are other questions in other directions, you can use your imagination, I will not elaborate here.

Notes

1. If this error occurs when the program starts, you can press "ctrl+C" to end the program and then restart the program.

```
python3 chatgpt_main_en.py
serial /dev/myspeech open
Network check failed: HTTPConnectionPool(host='www.baidu.com', port=80): Max retries exc
eeded with url: / (Caused by NewConnectionError('<urllib3.connection.HTTPConnection obje
ct at 0x7fff84058610>: Failed to establish a new connection: [Errno -3] Temporary failur
e in name resolution'))
检测网络没连上, 请重启网络
```

2. If you want to terminate the case, press "ctrl+C" to end the program.

