

Embodied Intelligence

[Embodied Intelligence](#)
[Function Introduction](#)
[Function Experience](#)
[Program Source Code](#)
[How to Add More Commands and Actions](#)
[How to Replace the Large Model Interface](#)
[How to run this example in the terminal](#)
[How to change the recording duration](#)
[Functional Principle](#)

Function Introduction

This tutorial is exclusive to CM5 version; CM4 version users can skip this step.

This example demonstrates how to use a large application. It requires configuring the API-KEY related to the large model to function correctly.

This function requires an internet connection to operate.

Function Experience

1. Power on the RiderPI. Press the button in the upper right corner of the screen to enter example mode, then select the Embodied Intelligence function.



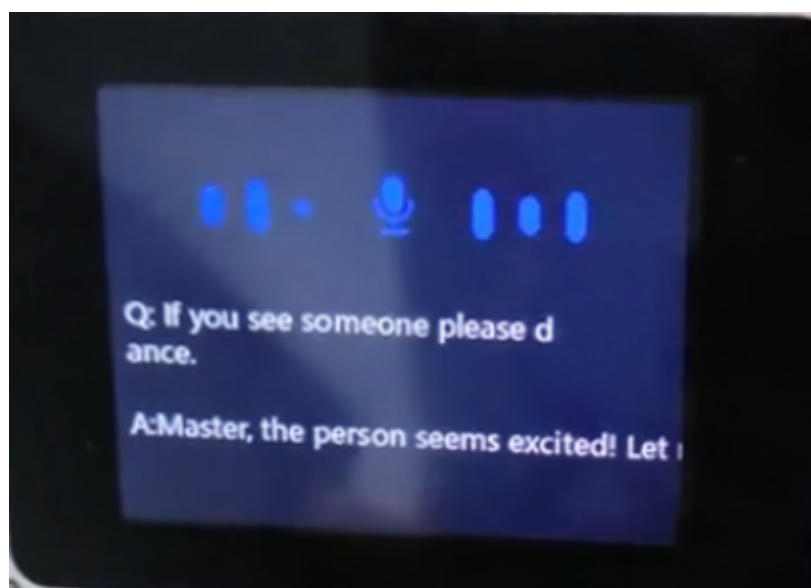
2. After entering the Embodied Intelligence function, wake it up with your voice, say "lulu".



3. After hearing a "ding" sound, you can speak the meaning of the command.



4. RiderPI will then respond and execute an action based on the user's semantics and a captured image of the scene.



Program Source Code

1. First, log in to the RiderPI system via VNC.
2. Then, in the terminal, enter:


```
cd /home/pi/RaspberryPi-CM5/demos/car_agent/
tree -L 1
```

3. Directory Structure Description (Only files related to this project are described)

- ├─ Alagent_go_image.py # Main function of the agent - Chinese version
- ├─ AlMImage_en.py # Main function of the agent - English version
- ├─ auto_platform.py # Environment dependencies
- ├─ DAgent_Image_en.py # Embodied intelligence agent - English version
- ├─ dog_agent_Image.py # Embodied intelligence agent - Chinese version
- ├─ dog_API_en.py # Large Interface for Embodied Intelligence - English Version
- ├─ dog_base_control.py # Basic Action Control Interface
- ├─ dog_ImageAPI.py # Image API Interface
- ├─ dog_record.py # Recording Interface
- ├─ dog_speakiat_en.py # Speech Recognition Interface - Chinese Version
- ├─ dog_speak_iat.py # Speech Recognition Interface - English Version
- ├─ dog_tongyiAPI.py # Large Interface for Embodied Intelligence - Chinese Version
- ├─ dog_tts.py # Audio Synthesis Interface
- └─ libnymaya.py # Voice Wake-up File

How to Add More Commands and Actions

- First, open the dog_agent_Image.py file to add an example command. The English version is DAgent_Image_en.py



```

19 Turn Around:Dog_Turn_Around()
20 Crawl:Dog_Crawl()
21 Squat:Dog_Squat()
22 Three-axis rotation:Dog_3_Axis()
23 pee:Dog_Pee()
24 sit down:Dog_Sit_Down()
25 wave/To greet:Dog_Wave_Hand()
26 stretch:Dog_Stretch()
27 Wave motion:Dog_Wave_Body()
28 Rocking motion:Dog_Swing()
29 handshake:Dog_Handshake()
30 dance:Dog_Dance()
31 Climb Stairs:Climb_The_Stairs()
32 push-up:Dog_push_up()
33 Display robotic arm:Dog_show_arm()
34 The robotic arm moves upwards:arm_up()
35 Robot arm grasping:arm_middle()
36 The robotic arm moves downwards:arm_down()
37 Pick up wooden blocks of the specified color, with a total of four colors: red, yellow, blue, and green,For example, pickin
38 Pick up wooden blocks of the specified color and place them in their corresponding positions. There are a total of four col
39 Kick away the balls of the designated color,There are a total of four colors for the balls: red, yellow, blue, and green. F
40 Scream (Surprise Scream):play_sound_surprised()
41 Scream (Angry Scream):play_sound_anger()
42 Introduce yourself: play_myself()
43 Rest and wait, such as waiting for two seconds:time.sleep(2)
44 There are also some color related meanings: for example, the sky color is blue, apples are red, bananas are yellow, and lea
45 Here are some executable action groups
46 【Output JSON format】
47 You can directly output JSON, starting from {, do not output the beginning or end of JSON containing ```
48 In the 'function' key, output a list of function names, where each element is a string representing the name and parameters
49 In the 'response' key, according to my instructions and your choreographed actions, output your reply to me in the first pe
50 【Here are some specific examples】 Here are some examples of command statements
51 My instructions: Move forward for 3 seconds, then lie down, show the robotic arm, and finally pee. You output:{"function":["Dog_Turn_Ar
52 My instructions:Start exercising.You output:{"function":["Dog_Squat()", "Dog_Squat()", "Dog_push_up()", "Dog_push_up()", "Dog_Wav
53 My instructions:Turn around and help me pick up the yellow wooden block. You output:{"function":["Dog_Turn_Around()", "caw_c
54 My instructions:First, perform three-axis rotation, and then kick the green ball away. You output:{"function":["Dog_Turn_Ar
55 My instructions:Just describe what you saw, then scream a few times and lie down. You output:{"function":["play_sound_surpr
56 My instructions:Move forward for 3 seconds, then move the robotic arm a few times, and finally climb the stairs.You output:
57 My instructions:Show the robotic arm upwards, then turn it around, and finally show the robotic arm downwards. You output:
58 My instructions:If you see yellow, turn around; otherwise, dance and finally lie down. You output:{"function":["Dog_Turn_Ar
59 My instruction: Put the small ball in the color of an apple that I grabbed onto the trash can on the right. You output: {"f
60 My instructions: Take two steps forward, then sit down and introduce yourself. You output: {"function": ["Dog_forward (2) "
61 Assuming there are two colors in the picture, my instructions are: if there is only one color, rotate in circles; if there
62 Assuming a male stranger appears in the picture, my instructions are: help me keep an eye on the door. If a stranger is fou
63
64
65

```

- Then, you need to encapsulate some action execution functions and save them in the dog_base_control.py file.

For example, encapsulate a forward movement function, as shown in the image below:

```
15  ##前进 forward
16  def Dog_forward(delay_time):
17      xgo.move_x(20)
18      time.sleep(delay_time)
19      xgo.stop()
20
21  ## 后退 back
22  def Dog_back(delay_time):
23      xgo.move_x(-20)
24      time.sleep(delay_time)
25      xgo.stop()
--
```

- Finally, add the interface and usage method of the encapsulated function to the `dog_agent_Image.py` file (DAgent_Image_en.py in the English version).

```
1  from dog_API_en import *
2
3  #DOGZILLA lite Action choreography agent description
4
5  AGENT_SYS_PROMPT = '''
6  You are my mechanical dog butler. Please output the corresponding function to be run and your reply to me in JSON format according to my instructions
7
8  [Here is an introduction to all built-in functions]
9  Forward movement:Dog_forward(time) #Among them, time represents the number of seconds of the action,Advance 1 second:Dog_forward(1)
10 Step back action:Dog_back(time) #Among them, time represents the number of seconds of the action,Step back for 1 second:Dog_forward(1)
11 Left translation action:Dog_Left_move(time) #Among them, time represents the number of seconds of the action,Left shift for 1 second:Dog_Left_move(1)
12 Right translation action:Dog_Rihgt_move(time) #Among them, time represents the number of seconds of the action,Right translation for 1 second:Dog_Rihgt_move(1)
13 Left rotation action:Dog_LeftTurn(time) #Among them, time represents the number of seconds of the action,Rotate left for 1 second:Dog_Rihgt_move(1)
14 Right rotation action:Dog_RightTurn(time) #Among them, time represents the number of seconds of the action,Rotate right for 1 second:Dog_Rihgt_move(1)
15 Looking up action:Dog_Looking_up()
```

How to Replace the Large Model Interface

The large model used in this function is **Xunfei Xinghuo in Chinese**, and **OpenRouter in English**.

- You can first find the corresponding Python version of the interface for your platform and fill in the necessary information according to the platform's interface and instructions.
- Then, encapsulate the executable file into a function. You can refer to the example of "xinghou_UltraAPI.py" and place it in the directory mentioned in point 2, for example, adding a file named "mychatgpt.py". For the English version, refer to the example of "dog_API_en.py". Replace the large model.

Chinese Large Model Interface (Text):

- For the Chinese version, open the file **dog_agent_Image.py**; for the English version, open **DAgent_Image_en.py**. Add `from mychatgpt import *` at the beginning.
- Then find this section and replace it with your encapsulated API function interface.

```
1 from dog_ImageAPI import *
2 from dog_tongyiAPI import *
```

- Then restart the app and re-enter this function. You should be able to run the replaced model platform. If it fails to run, there is an error, and you need to check the syntax and logic of the newly added file.

How to run this example in the terminal

- First, terminate the large program to prevent screen flickering. For instructions on how to terminate a large program, please refer to the tutorial in Chapter 1.
- Enter the following command in the terminal:

```
cd ~/RaspberryPi-CM5
python3 demos/car_agent/AIMImage_en.py
```

How to change the recording duration

1. Enter the following in the terminal:

```
nano ~/RaspberryPi-CM5/demos/car_agent/dog_record.py
```

2. Locate this section and modify the setting shown in the image below. This will allow you to adjust the recording duration according to your environment.

```
229 def start_recording(timer = 3, save_file=SAVE_FILE):
230     global automark, quitmark
231     start_threshold = 120000
232     end_threshold = 100000
233     endlast = 15
234     max_record_time = 5
```

Voice threshold during speaking, adjust according to your environment.

Maximum recording time, set to 5 seconds here.

Note: `start_threshold > end_threshold`. Adjusting these two values requires specific adjustments based on your environment.

Functional Principle

The detailed flowchart is as follows:

