

Setting up the whole environment

The following project can be used to control several agents with voice commands (translated using Google Speech API).

The game environment was used from an open-source project by Microsoft called AirSim. The necessary details to set up this game environment are mentioned below. This game environment can be used to control both Drones and Cars. However, because of simplicity of car controls I preferred using the car game environment. This open source code also has Unity code available. However, it does not support multi-agent system currently. So, the project uses Unreal Engine for the same. Changing the number of agents in the game environment only requires a few changes in settings.json created by the code.

However, the open source code does not use Speech command as of now. To implement this, I used python scripts to implement the same which currently is set up for three agents only. But with very little changes in the script more agents can be easily added. Also, the code is set for basic car commands like (go right, left, forward, backward and apply brakes). Using some simple classes from open source environment, more commands can also be added.

It is advised, for better Speech-to-text translation, use the paid Google Speech-To-Text API.

This code can further be used to study if adding more agents can affect the translation process and hence affect the performance of the multi-agent system.

A) Setting up the AirSim Environment

Follow the steps mentioned in the following link exactly.

<https://microsoft.github.io/AirSim/>

It mentions how to:

- a) Set up Unreal Engine
- b) Building AirSim Project
- c) Building Unreal Project
- d) Setting up Blocks Game Environment for the project

For a 3-car game environment update the settings.json file (which should be found here: **Documents\AirSim**) to the settings.json file in the repository. For more details about the json file:

<https://github.com/Microsoft/AirSim/blob/master/docs/settings.md>

B) Setting up the Google Speech-to-Text API

Follow the steps mentioned in the following link exactly.

<https://cloud.google.com/speech-to-text/docs/quickstart-client-libraries>

Ensure, you do set the proper path for the API key json file, as mentioned in step 2 in the above link, before running the python scripts. Rename the downloaded json key as “key.json” and save it in the **AirSim\PythonClient\car** directory for convenience. This directory will be created when AirSim is properly setup. It contains some simple Python examples for Car game environment.

C) Python Scripts

There are two Python scripts in the github repository. Save them in the **AirSim\PythonClient\car** directory too (for convenience).

- 1) `speech_command.py`: This contains functions for controlling the car. (For example: `goForward`, `goLeft`, `applyBrakes` etc)
- 2) `main.py`: Sets up the microphone streaming (to actively listen for commands), implements to speech-to-text translation and uses translated text to control the cars. This script also sets up certain variables the controls the number of agents. These variables along with changes in `settings.json` (which should be found here: **Documents\AirSim**) can be used for changing the number of agents.

Ensure both part A and part B are properly set up before running `main.py`

The video (**demo.mp4**) included in the repository gives a demonstration of the same. However, the camera in the game only follows one agent. The open source code is set up that way.