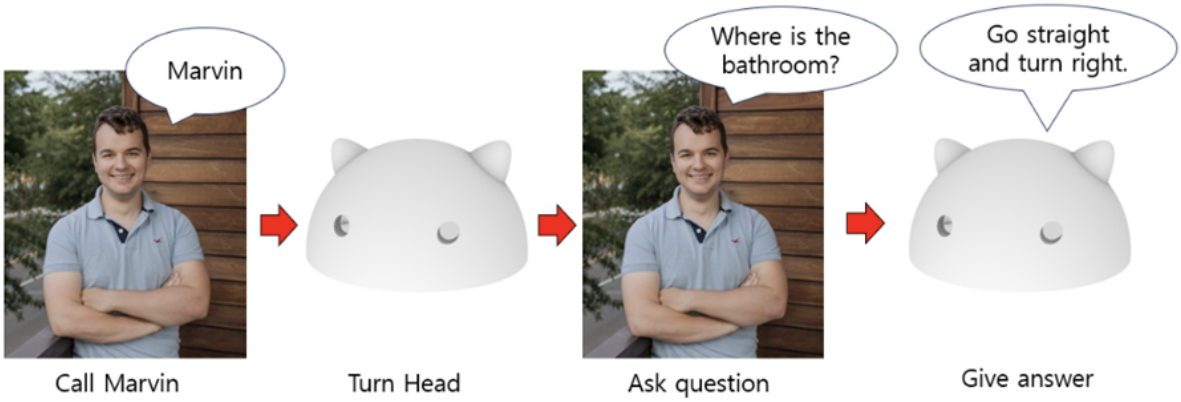
Final Report

Group4
9212420231 Will Webster
2018007938 MinKwan Kim(김민관)
2021029443 Yeeun Jeon(전예은)

1. Introduction

1.1. Short introduction of Marvin

In our project, we built a speech based receptionist robot, Marvin, that answers people's questions. For example, if a person calls "Marvin", Marvin will turn his head to the person, and then when a person asks a question such as "Where is the bathroom?" Marvin will give the answer "Go straight and turn right".



How Marvin Works

Marvin is a receptionist robot, so it can be used in places like museums, hospitals, and businesses and help by giving responses to people's questions. It can be in the information desk and provide information on things people are curious about, such as guiding the location of specific amenities.

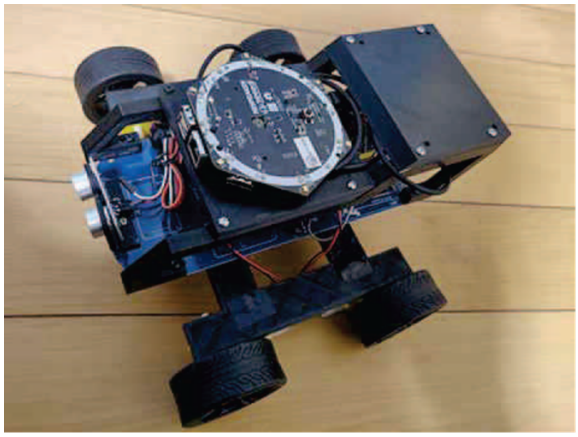
1.2. Motivation of doing the project

'ChatGPT' is a user-friendly Large Language Model (LLM) where users ask questions by text and get answers by text. We wanted to build a friendly receptionist robot that when asked questions by speech, we get the answers by speech. Our first goal was to build a robot that

operates with a LLM, so that if we ask any questions, the robot can give any answers to the question; however, due to limited time and hardware, we built a robot that gives predefined answers.

Traditional receptionists are often constrained by time limitations, human errors, and varying workloads. A receptionist robot will address this problem because it can work for twenty-four hours and will not forget any information about the building. Receptionist robots will also free up human receptionists to focus on more complex and value-added responsibilities, rather than responding to repetitive questions from people.

1.3. Literature Review

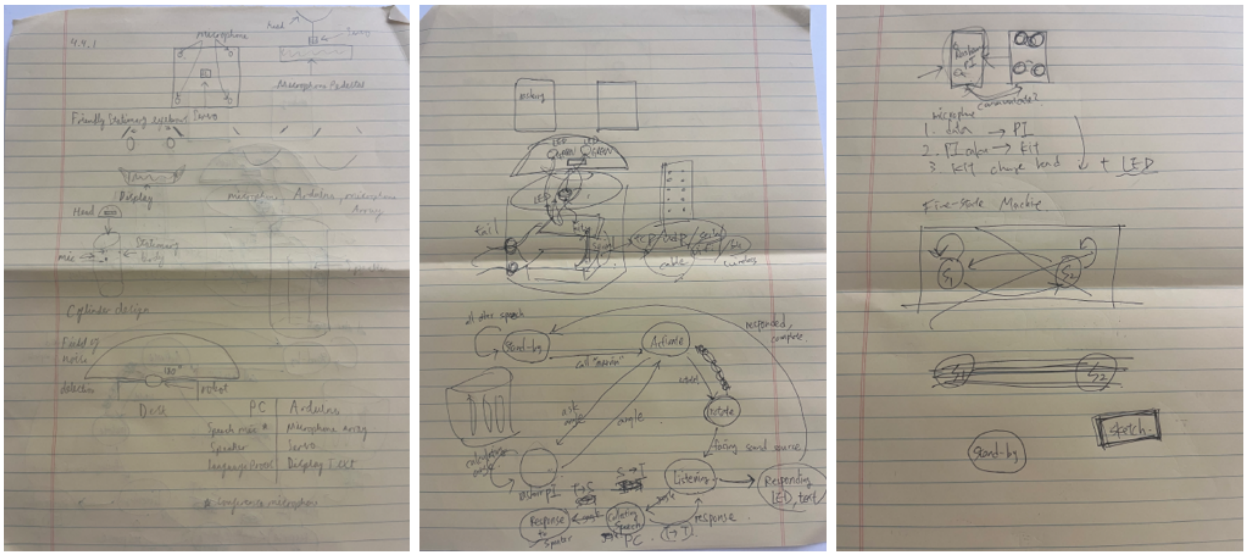


In our literature review, the most helpful research article was, “Portable keyword spotting and sound source detection system design on [a] mobile robot with Mini Microphone Array” [1]. The researchers used an Arduino for Sound Source Detection (SSD) and keyword spotting, which we intended to do as well. They were mostly successful and a great starting point for our research. They used the ReSpeaker Core v2.0 Mic Array, which is where we started in our search for a microphone array. We settled on a cheaper alternative to the ReSpeaker Core v2.0 that was also from Seeed Studio since we had a reduced budget.

2. Design and construction process

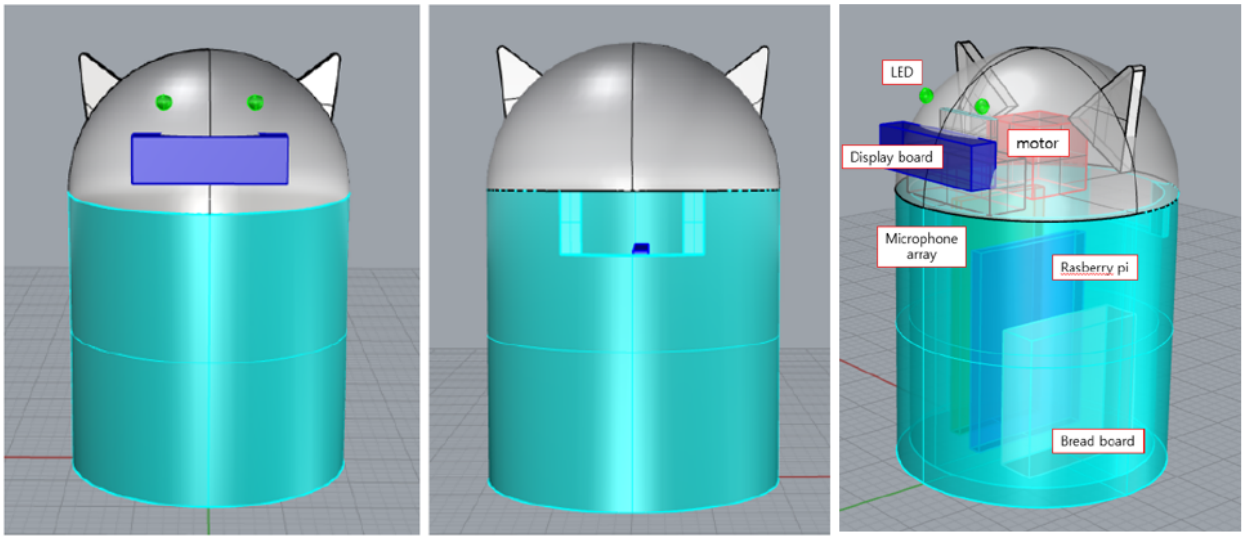
2.1. Design

We wanted to build a robot that can contain all the necessary hardware (Arduino board, LEDs, etc.), that was also friendly looking. We first drew sketches of the robot design by hand.



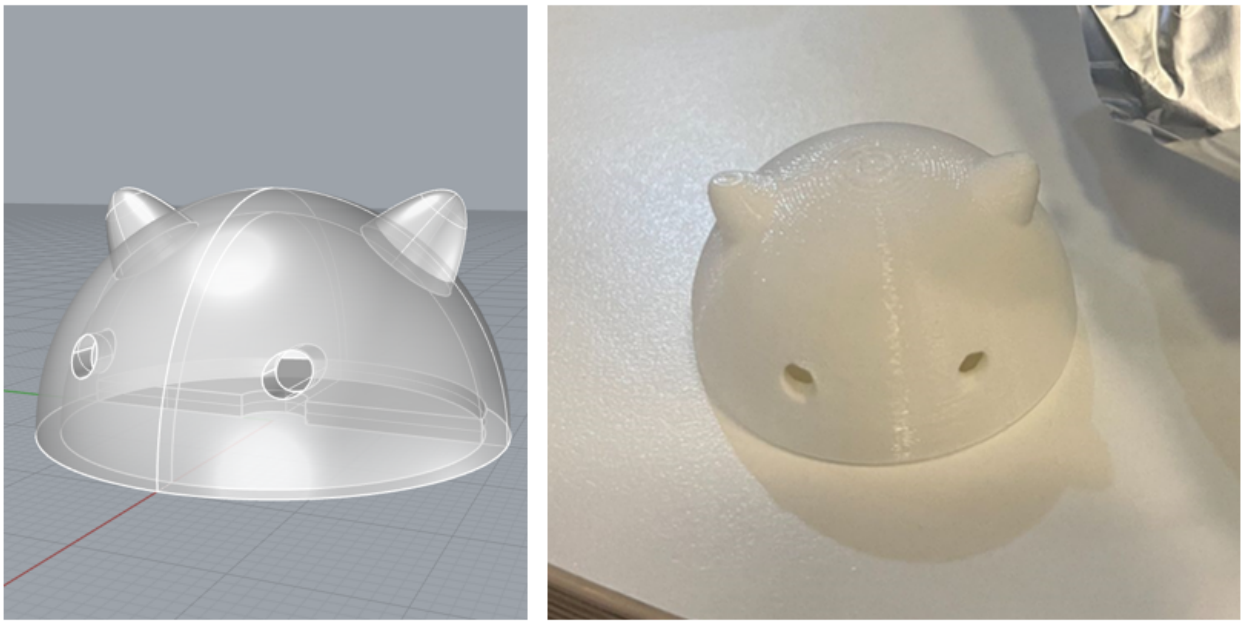
Initial Sketches by Hand

We then designed the robot in 3D using the software: Rhino. The initial design had two parts: head and body, where the head included LED and display board to indicate things. The body included boards for operating the robot.



Initial Design

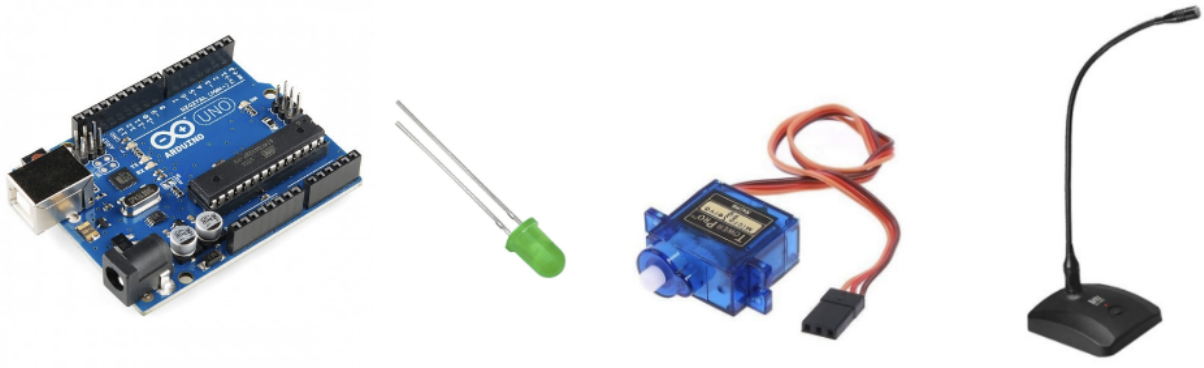
Due to time limits and technical difficulties, we only made the robot's head. We added the ears so that the robot looks more friendly. We made holes that looked like eyes, to put LEDs inside the robot and make the light come out of it. There is also a hole at the bottom to attach the motor easily. We then 3D printed the design. The radius of the robot bottom is about 2 centimeters.



Changed Design

2.2. Hardware of the robot

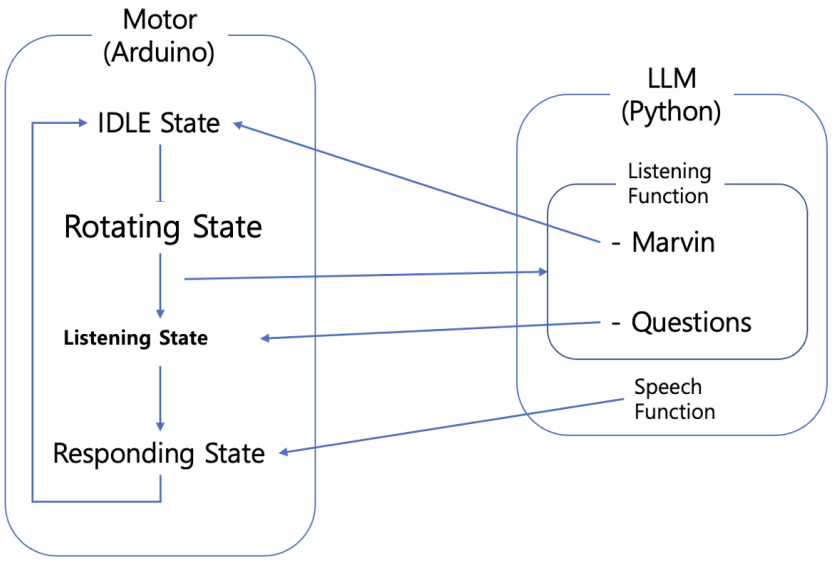
For the hardware of the robot, we used Arduino Uno board, 2 LEDs, Servo motor, and microphone. We used LEDs for turning lights on, and blinking when robot is in some particular state, and we used servo motor to turn the robot's head towards the person asking question. And we used microphone for better speech recognition from far distances.

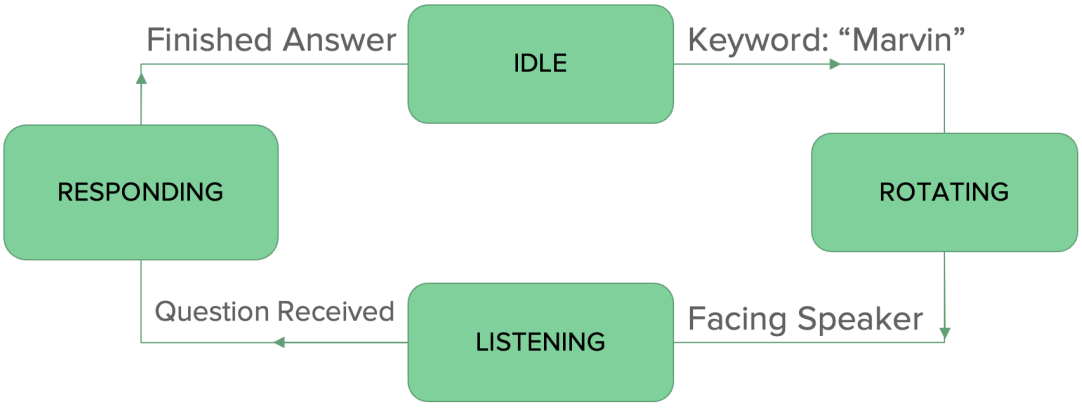


Hardware of the robot

2.3. Code (Explanation of how Marvin works)

The code for operating Marvin can be broadly divided into two parts. One is the code that controls the motor, corresponding to Marvin's head (Arduino), and the other is the code that manages the simulated LLM (Large Language Model) in Python. These two codes communicate via serial communication, continually checking the robot's current state. Based on the observed state, appropriate actions are taken.



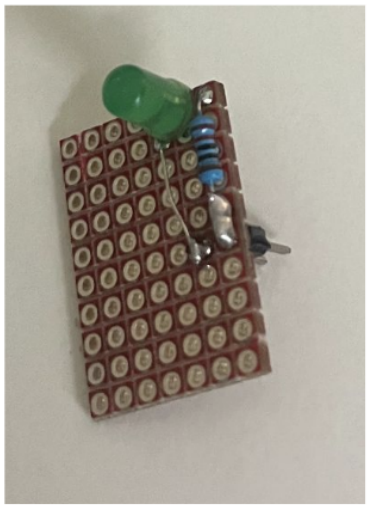


The motor part has four main states: IDLE, Activate, Rotating, and Responding. Meanwhile, the simulated LLM section consists of the Listening Function and the Speech Function. When implementing the code, we tried to speculate on how Marvin would behave as if it were a person. Therefore, the Listening Function of the simulated LLM code is always running in the background, constantly listening to all sounds. If the Listening Function recognizes the word 'Marvin,' similar to turning one's head when someone calls their name, it requests a transition to the Activate state from the Motor. After receiving confirmation that the motor has successfully rotated, the simulated LLM listens to a question from the user. The Speech Function then

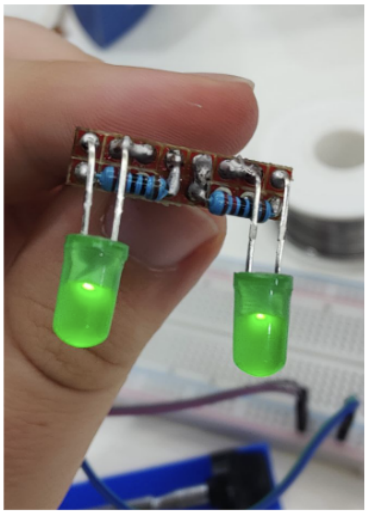
answers the question and instructs the Motor to face forward again. It waits until Marvin is called again before repeating the process.

2.4. Electrical

Then we tried to make the LED lights come out of the holes. Our original design had a single LED that would illuminate the inside of the head, so the eyes would only be holes; however, the LED was not bright enough to be noticeable, so we created a new board with a LED for each eye. A soldering iron was used on the eye holes to allow the LEDs to fit perfectly in place. The board fits securely inside without any adhesives.



Original Board



Final Board

Complete Head

3. Experimental data collection

We did some experiments to check if the robot operates properly as we designed. We had four parts to test: 1) Recognizing “Marvin”, 2) Recognizing questions and giving answers to questions, 3) LED, 4) Servo Motor. For each part, we tested if each part operates properly. We had three people in our team, so if two out of three people succeed for each experiment, we assume it is ‘success’ for each experiment. First part of the experiment was about whether the robot recognized Marvin correctly or not, and operating properly for each situation. We could see that all the results were successful.

Experiment	Minkwan	William	Yeeun	Result
If person calls “Marvin”, the robot has to say “Hello, I am Marvin. Do you need any help?”	O	O	O	Success
If person calls name other than “Marvin”, the robot has to say "My name is Marvin. If you have questions, call my name."	O	O	O	Success
If person doesn’t say anything in 20 seconds, the robot has to say "My name is Marvin. Do you need any help?"	O	O	O	Success

Recognizing “Marvin”

The second part of the experiment was to recognize people’s questions and give correct answers for each situation. We could see that all the results were successful.

Experiment	Minkwan	William	Yeeun	Result
If ‘bathroom’ is in the question, the robot has to say “Go straight and turn right.”	O	O	O	Success
If ‘professor’ is in the question, the robot has to say “He is on the 5th floor in the FTC building.”	O	O	O	Success
If person says something other not including ‘bathroom’ or ‘professor’, the robot has to say “Sorry, I could not understand what you were talking about.”	O	O	O	Success
If person doesn’t say anything in 20 seconds, the robot has to say “Sorry. I could not hear anything.”	O	O	O	Success
If person doesn’t say anything in 40 seconds, the robot has to say “Sorry, Try again.”	O	O	O	Success

Recognizing questions and giving answers to questions

The third part of the experiment was to see if LED works properly in each situation. We could see that all the results were successful.

Experiment	Minkwan	William	Yeeun	Result
If the robot is in IDLE state, LEDs should be turned off.	O	O	O	Success
If the robot is in state other than IDLE state, LEDs should be turned on.	O	O	O	Success
If the robot is in LISTENING state, LEDs should blink.	O	O	O	Success

Experiments of LEDs

The last experiment was to see if the servo motor works properly in each situation. We could see that all results were successful for experiments.

Experiment	Minkwan	William	Yeeun	Result
If the robot is in ACTIVATED state, the robot should turn head 30 degrees in the first time and 150 degrees in the second time.	O	O	O	Success
If person says something other not including 'bathroom' or 'professor', the robot has to shake its head side to side.	O	O	O	Success

Experiments of Servo Motor

4. Data analysis

Although the experimental data may be limited, each team member has a different voice, varying in tone, volume, and accent; therefore, the experimental results are sufficiently reliable for the scope of this project; however, experiments regarding the distance and angle from the microphone are still lacking, and all tests were conducted in quiet environments without noise. The presence of noise can pose challenges in recognizing keywords or questions, which is an aspect we need to address and improve upon.

5. Conclusion

5.1. Lessons learned about HRI or robotics

- The robot did not work like what we expected. It is more complicated to get it right. For this reason, making a plan B is important.
- When encountering a roadblock, be willing to adapt and change the plan; otherwise, time will be wasted trying to force something that will not work.
- When designing a robot(technical/aesthetic design), we should think about users and design towards how users will use it more conveniently. We also have to think about more of the potential use cases of the robot.

5.2. Limitation & Future research

Limitation:
Due to hardware limitations, we were unable to implement sound detection; therefore, we instead worked on moving the robot's head to predefined angles. For the LLM, our initial goal was real-time learning, but due to time constraints, we opted for a keyword-based approach. We selected specific keywords such as 'Marvin,' 'Professor,' and 'Bathroom' for recognition. When these keywords are detected, the system responds appropriately with predefined answers.

Future Research:
With sufficient time and utilization of advanced hardware, this robot has the potential for diverse applications in the future. Especially in the context of an information desk, it could respond to a variety of customer inquiries. Additionally, even when faced with unfamiliar questions, it could learn from previous interactions and provide answers based on acquired knowledge. OpenAI has introduced custom GPTs, enabling the creation of a GPT tailored for building information. Users can now run the language model locally, enhancing security and privacy. The update includes additional states for nuanced dialogue, and there is a focus on designing a physical body for the GPT. Additionally, a mobile robot has been developed to guide users around a building.

References

[1] M. B. Andra and T. Usagawa, “Portable keyword spotting and sound source detection system design on mobile robot with Mini Microphone Array,” 2020 6th International Conference on Control, Automation and Robotics (ICCAR), 2020. doi:10.1109/iccar49639.2020.9108086