HexaPolo Bot: Revised Team Project Proposal

Zakaria Laribi, Nagyung Kim, Roméo Maignal, Tuan Dang Nguyen

1. Introduction

Robots that can sense and react to their surroundings are becoming more common, and combining different types of sensors can make these robots smarter and more interactive. This project explores how a walking robot can use sound and vision together to find and follow a person. The idea is inspired by the classic game *Marco Polo*, where one player has to find another by listening to their voice.

Instead of relying only on one sensor, the Hexapolo Bot uses microphones, a camera, and a lidar scanner to understand its environment. This project is a way to experiment with how different technologies—such as sound direction detection and computer vision—can be used to build a robot that moves and reacts to people in real time. The focus is not only on making the robot move but also on making it respond to human interaction in a playful and intelligent way.

2. Extended Project Description

2.1. Project Overview

The Hexapolo Bot is a six-legged robot that can find and follow a person by listening to their voice and looking at them. It starts by saying “Marco” and waits for someone to reply “Polo.” When it hears the reply, the robot uses four microphones to figure out which direction the sound came from. Then, it turns to face that direction and checks with its camera if a person is there. If it sees someone, it walks toward them and stops when it is about one meter away. After that, the robot will follow the person as they move.

The project’s goal is to combine sound detection, camera vision, and motor control so the robot can interact with people. The robot will also have optional features, like working in a noisy room or avoiding obstacles using a lidar scanner. The work is divided into small steps to make sure each part works well before moving to the next.

2.2. Functional Specifications

The Hexapolo Bot will be able to do several things to find and follow a person.

These are the main features it needs to have:

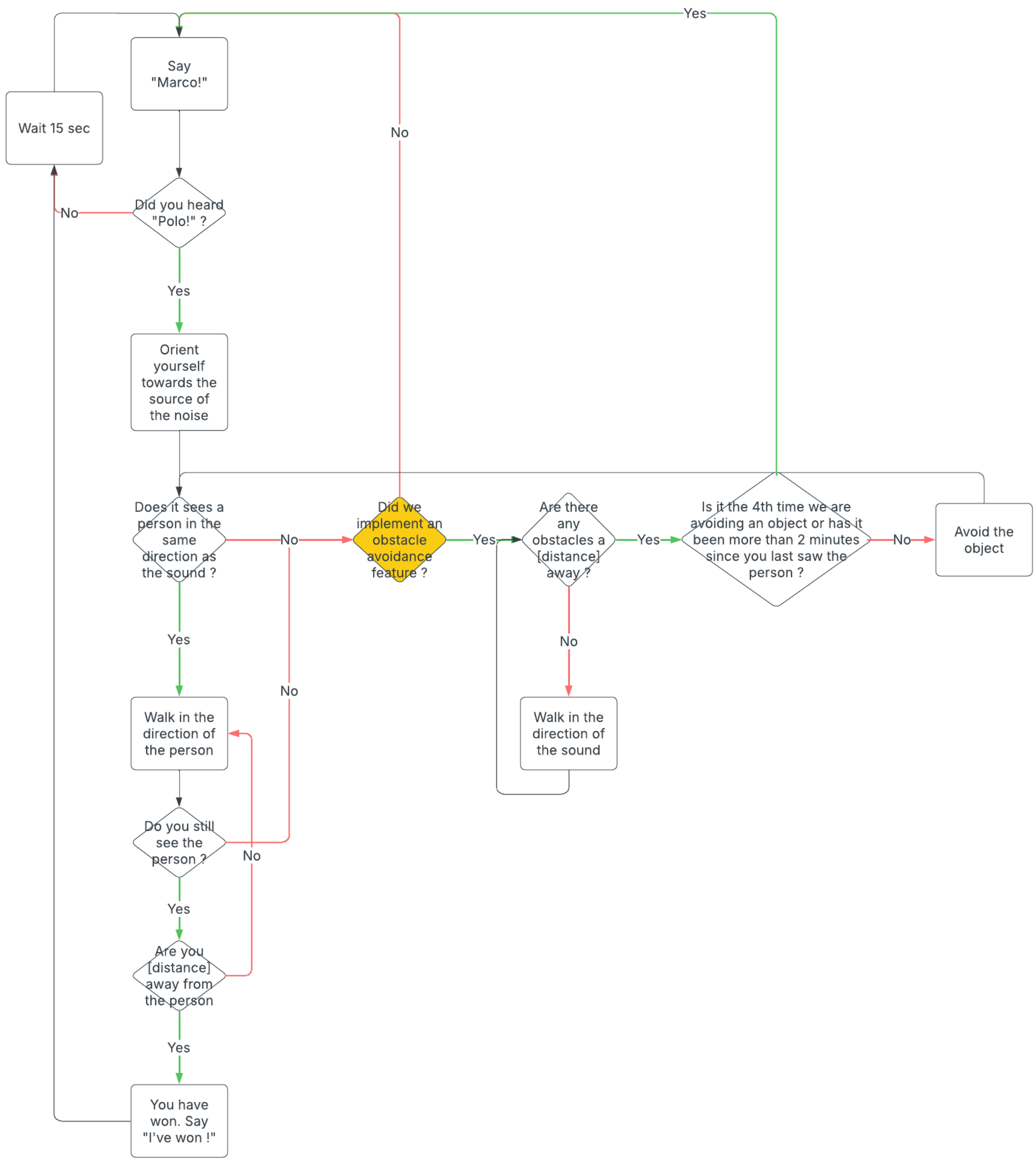
- **Keyword Detection**: The robot will listen for the word “Polo” and react when it hears it.
- **Sound Direction Estimation**: It will use four microphones and a time difference of arrival (TDoA) method to figure out which direction the sound came from.
- **Turning Toward the Sound**: After detecting the direction of the sound, the robot will rotate to face that direction.

- **Visual Detection**: The robot will use its camera to check if there is a person in front of it.
- **Short Verbal Response**: Once the robot detects a person, it will say a short phrase to confirm it.
- **Walking Toward the Person**: The robot will move toward the person and stop when it is about one meter away.
- **Following the Person**: After reaching the person, the robot will follow them as they move around.

Optional Features:

- **Noise Handling**: The robot will be able to detect the “Polo” keyword even in a noisy room.
- **Obstacle Avoidance**: The robot will use a lidar scanner to detect obstacles and avoid them while moving.

Here is a decision tree illustrating the functional specifications :



2.3. Technical Approach

The robot's system uses a combination of microphones, a camera, sensors, and motors, all connected to a Raspberry Pi and ESP32-CAM, as shown in the wiring diagram below.

The robot's main task is to find and follow a person based on sound and visual information. For this, the following hardware and techniques will be used:

- **Sound Direction Detection**:
  The robot uses a **ReSpeaker Mic Array** (4-microphone setup) connected to a Raspberry Pi. When someone says "Polo," the robot compares the tiny time differences of when the sound reaches each microphone. This method is called **TDoA (Time Difference of Arrival)**. It helps the robot calculate which direction the sound is coming from. For sound recognition and direction detection, the project will use the open-source library provided by Seeed Studio: https://github.com/respeaker/mic_array

- **Person Detection and Tracking**:
  The robot uses an **ESP32-CAM module** with an external antenna to capture video and check if a person is in front. The camera sends the images to the Raspberry Pi for processing. The robot does not need to recognize specific people; it only checks if any person is there.

- **Motor and Movement Control**:
  The robot uses two **N20 DC motors** to drive its movement and a **TB6612FNG motor driver** to control the motors safely. It also uses a **SG90 servo motor** to adjust the camera or other parts if needed.

- **Distance-Based Following with TFMini LiDAR**
  To maintain a consistent following distance of about 1 meter, the robot uses a **TFMini LiDAR** sensor. It helps the robot follow the person smoothly by constantly checking how far ahead the person is. (This is considering the setup that there are no multiple people close to each other)

- **Sound Feedback**:
  The robot responds with a short phrase using an **Adafruit Mini Metal Speaker** and a small **PAM8302 audio amplifier** to make sure the sound is loud enough.

- **Power Supply**:
  The robot is powered by a **2S 7.4V LiPo battery** and uses a **DC-DC buck converter** to safely supply power to the components.

2.4. Milestones

To build the Hexapolo Bot step by step, the work is divided into smaller goals, called milestones. Each milestone focuses on one part of the robot's behavior.

Milestone 1 – Sound Detection & Turning
In a quiet room without obstacles, the robot listens for the word “Polo.” It uses the microphone array and a keyword detection system to recognize the word. Then, it uses the TDoA method to find the direction of the sound. After that, the robot turns toward the sound and uses its camera to look for a person.

Milestone 2 – Walking to the Person
Once the robot detects a person, it moves toward them using its motors. It stops when it is about one meter away. Distance is checked using either the ultrasonic sensor or LiDAR.

Milestone 3 – Following the Person
The robot keeps following the person as they move around. It uses the lidar to keep track of the person’s location and updates its movement.

Milestone 4 (Optional) – Working in Noisy Rooms
The robot will be tested in a noisy environment. It will use sound-processing libraries from the ReSpeaker project to help it detect the “Polo” keyword even when there is background noise.

Milestone 5 (Optional) – Obstacle Avoidance
The robot will use a **TFmini Plus LiDAR scanner** to detect obstacles while moving. If something is in its path, it will stop and change direction to avoid collisions.

3. Complete Technical Drawings

3.1. CAD Drawings

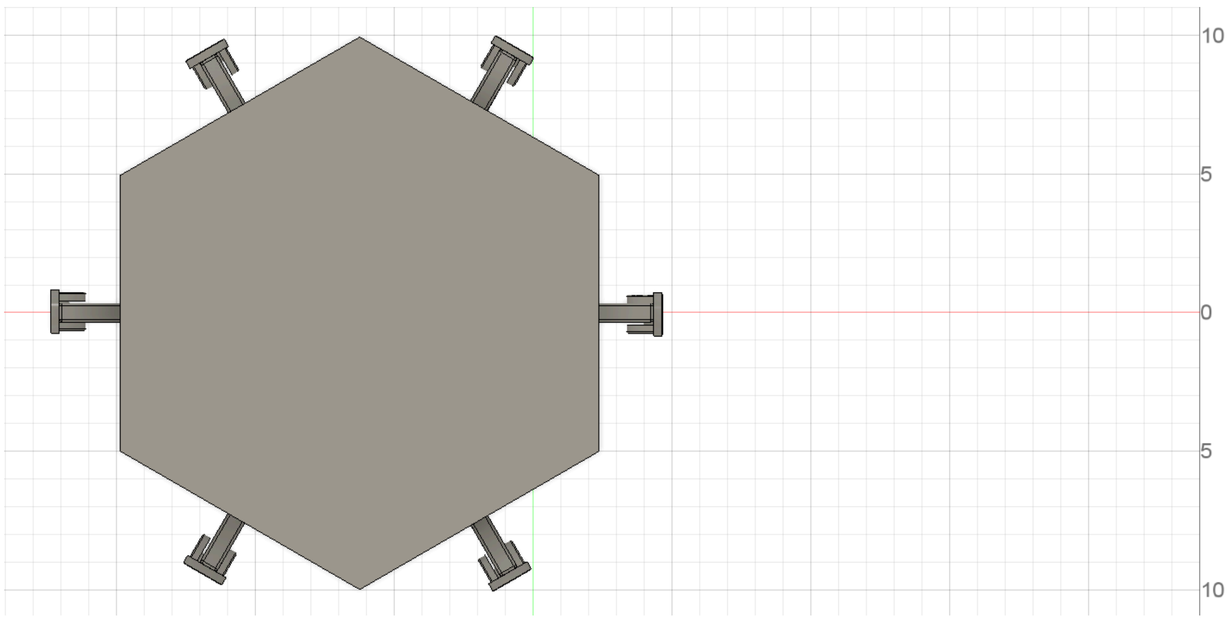


Figure 1: The prototype looking from above (the numbers on the vertical axis are in cm)

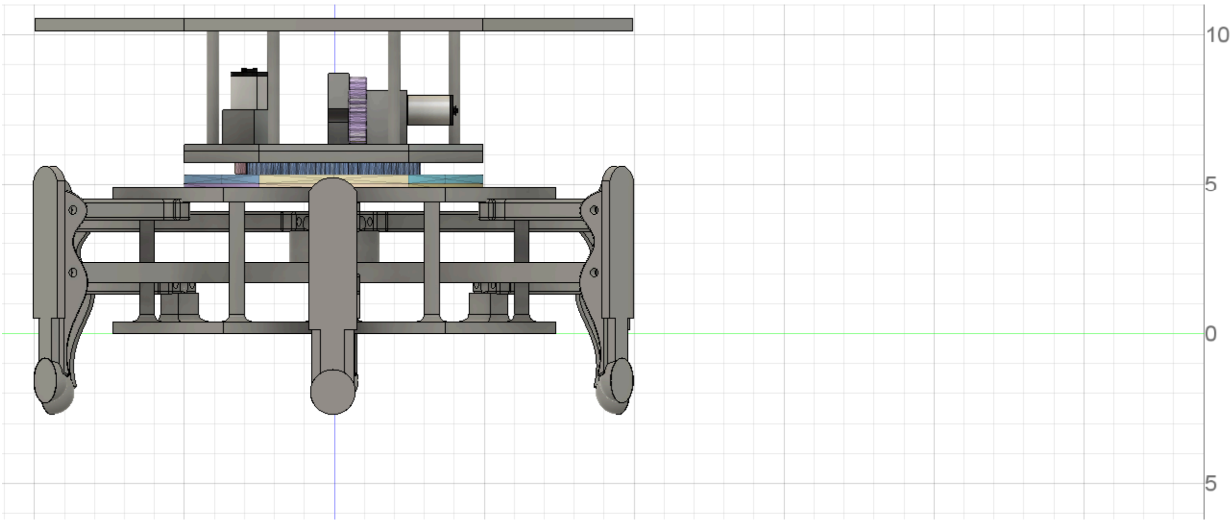


Figure 2: The prototyping looking from the right

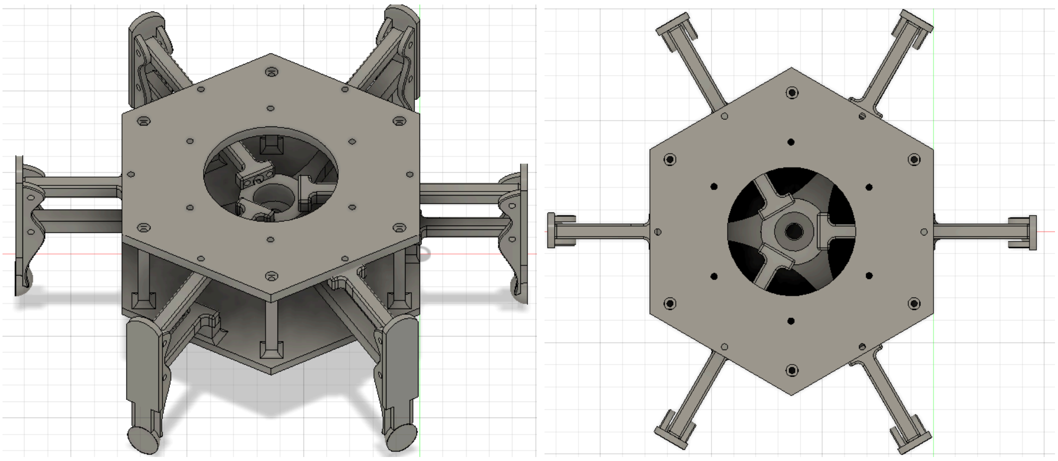


Figure 3: Part 1, the Legs and mechanical links to control the legs

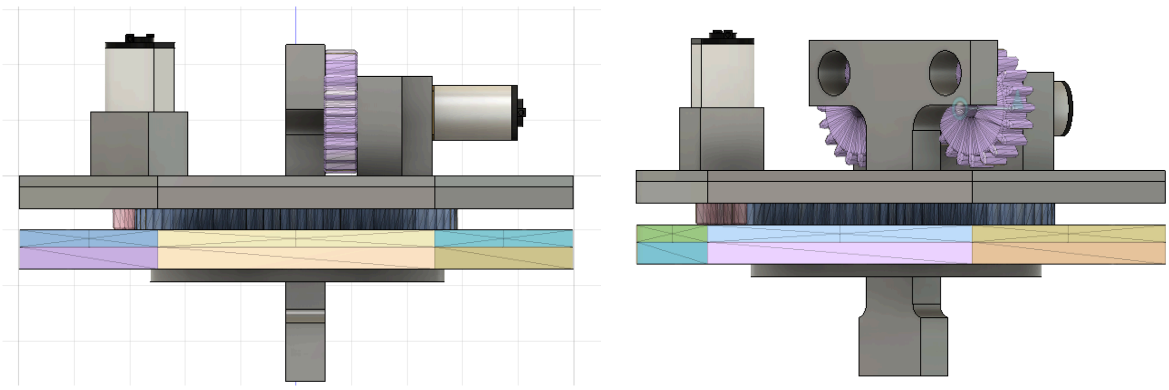


Figure 4: Part 2, T-bar, gears and motors to rotate heads and control the movement of the legs

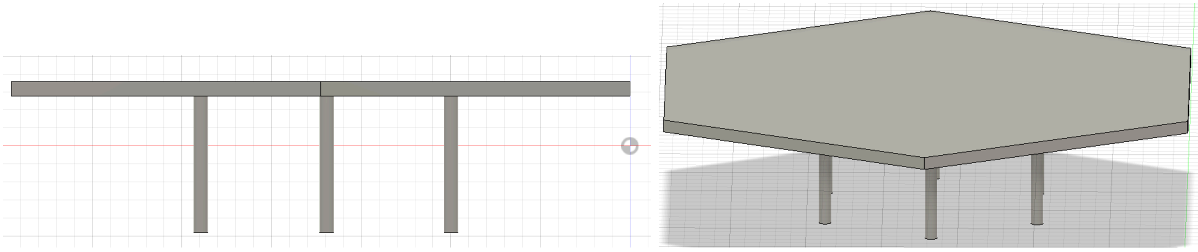
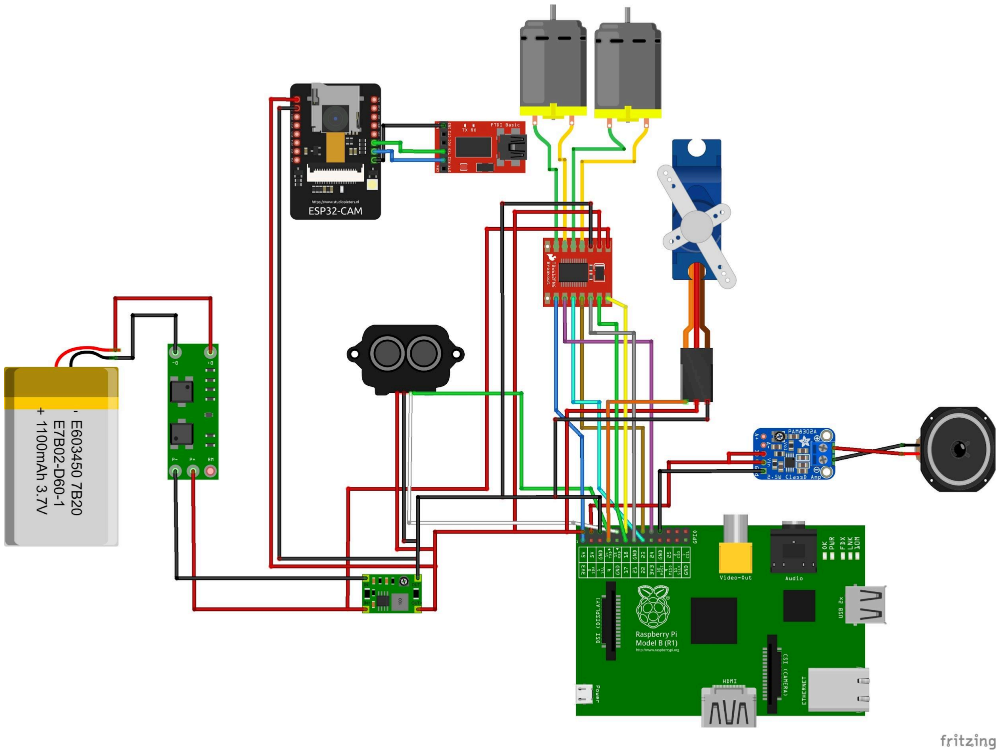


Figure 5: Part 3, the platform that holds all the electronics

3.2. Electrical or Circuit Schematics

Please note that the following diagram does not include the USB connections between the FTDI board and the Raspberry Pi as well as that between the microphone array (also not included) and the Raspberry Pi. Furthermore, note that the speaker amplifier will also be wired to a 3.5mm headphone jack, which will plug into the microphone array, through its A+ and A- ports.



4. Bill of Materials (BOM)

- AI thinker esp32-cam with antenna - 13. - cs358 stock
    - FTDI USB-to-serial - 2. - cs358 stock
    - sg90 micro servo - 3. - cs358 stock
    - raspberry pi 1 B - in stock
    - Adafruit Mini Metal Speaker - 8.9 - https://www.digitec.ch/en/s1/product/adafruit-mini-metal-speaker-loudspeaker-electronics-modules-5998730

- DFRobot Fermion TB6612FNG 2x1.2A DC Motor Driver - 14.9 - https://www.digitec.ch/en/s1/product/dfrobot-fermion-tb6612fng-2x12a-dc-motor-driver-development-boards-kits-24935800
- Tfmini Plus Lidar - 89 - https://www.fruugoschweiz.com/tfmini-plus-lidar-reichweitensensormodul-12-m-reichweite-hohe-bildrate-kleiner-blindbereich-hohe-genauigkeit/p-213832614-455740347?language=de&ac=bing
- 300rpm 6V n20 motor - in stock
- 200rpm 6V n20 motor - in stock
- 2S 7.4V lipo 1100mAh 15C (with xt60 connector) - 12.95 - https://www.conrad.ch/fr/p/reely-pack-de-batterie-lipo-7-4-v-1100-mah-nombre-de-cellules-2-15-c-softcase-xt60-2582340.html
- WaveShare 6-36V to 5V/3.3V 4A DC-DC buck step down converter - 10.9 - https://www.digitec.ch/en/s1/product/waveshare-6-36v-to-5v33v-4a-dc-dc-buck-step-down-converter-development-boards-kits-53542221
- OV5640 camera - 9.95 - https://www.fruugoschweiz.com/ov5640-module-de-camera-pour-carte-logicielle-esp32-cam-5-millions-de-pixels-c/p-331210929-730594189
- Goobay Jack plug 3.5 mm mono - 2.7 - https://www.digitec.ch/en/s1/product/goobay-jack-plug-35-mm-mono-audio-connectors-20604603
- Purecrea Piezo buzzer active - 5.9 - https://www.digitec.ch/en/s1/product/purecrea-piezo-buzzer-active-development-boards-kits-38633745
- ReSpeaker Mic Array v2.0 - 58.51 - https://www.reichelt.com/ch/en/shop/product/respeaker_mic_array_v2_0-248721?country=ch&CCTYPE=private&LANGUAGE=en
- PAM8302 2.5W single channel Class D Audio power amplifier - 7.95 - https://www.fruugoschweiz.com/pam8302a-cjmcu-832-pam8302-25-w-einkanal-class-d-audio-leistungsverstarkermodul-entwicklungsplatine/p-263124514-578035388?language=de&ac=bing

Total cost: 239.66

5. Work Breakdown Structure

5.1. Task Breakdown

The chart below outlines our planned task breakdown for the project. While it gives a clear structure for how we aim to progress week by week, we are approaching the plan with flexibility and will adapt it as needed based on testing outcomes and unexpected challenges.

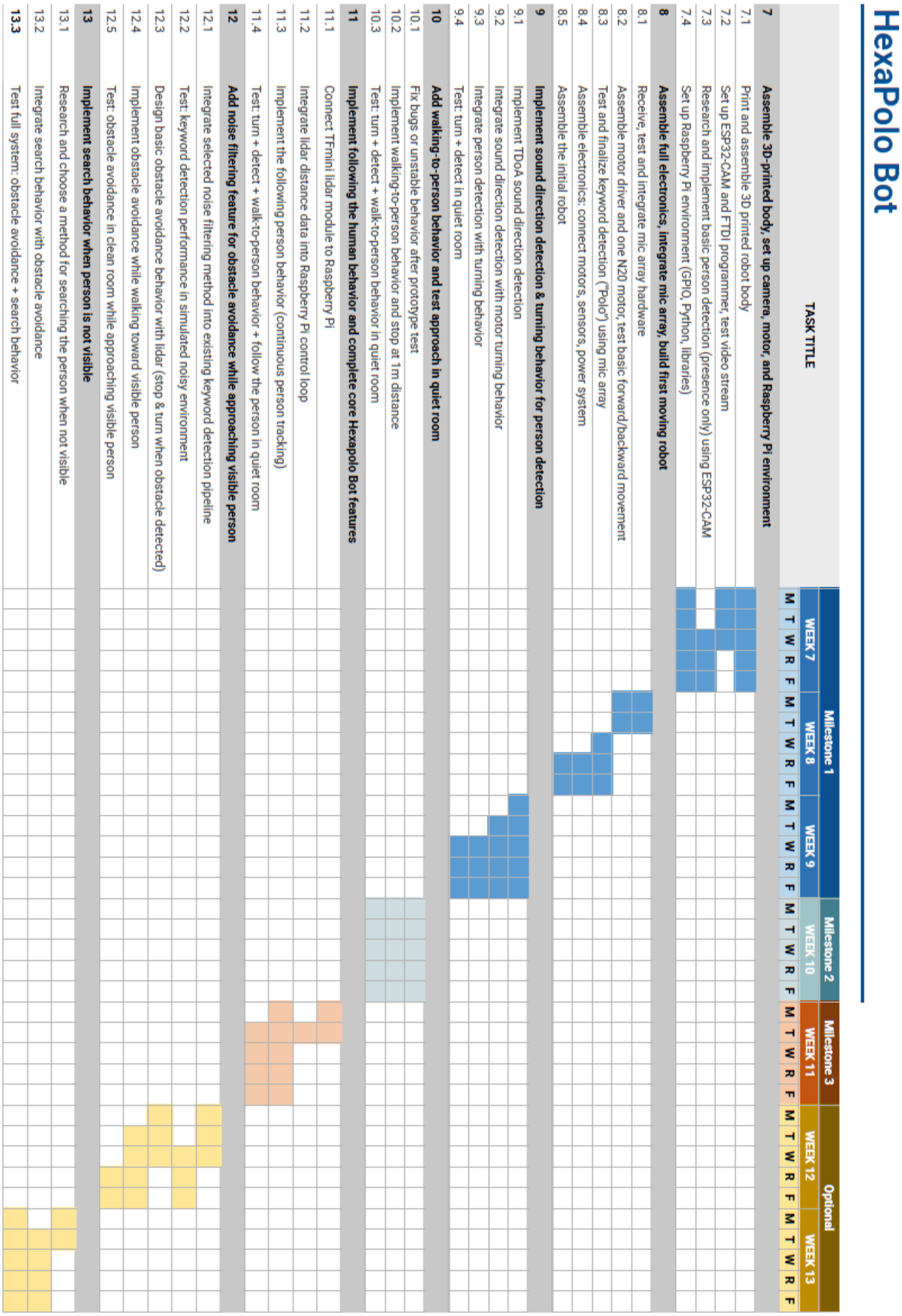
The tasks marked as optional represent features we would like to implement if time allows, but they will only be pursued after all core features are completed and working reliably.

The timeline for the core features is intentionally planned with more buffer time to handle possible errors and unexpected issues, while the optional features are scheduled more tightly so that, if we finish the core tasks earlier than expected, we can immediately move forward without losing time.

Week	Task	Dependencies / Note
Week 7: Assemble 3D-printed body, set up camera, motor, and Raspberry Pi environment		
7.1	Print and assemble 3D printed robot body	Independent
7.2	Set up ESP32-CAM and FTDI programmer, test video stream	In-stock components
7.3	Research and implement basic person detection (presence only) using ESP32-CAM	In-stock components
7.4	Set up Raspberry Pi environment (GPIO, Python, libraries)	In-stock components
Week 8: Assemble full electronics, integrate mic array, build first moving robot		
8.1	Receive, test and integrate mic array hardware	Needs delivery
8.2	Assemble motor driver and one N20 motor, test basic forward/backward movement	Need delivery
8.3	Test and finalize keyword detection (“Polo”) using mic array	
8.4	Assemble electronics: connect motors, sensors, power system	Needs motor driver, battery
8.5	Assemble the initial robot	This is the very first moving robot we will have
Week 9 (Milestone 1): Implement sound direction detection & turning behavior for person detection		
9.1	Implement TDoA sound direction detection	Needs working mic array
9.2	Integrate sound direction detection with motor turning behavior	Needs sound direction + motor setup
9.3	Integrate person detection with turning behavior	Needs vision + turning working
9.4	Test: turn + detect in quiet room	First prototype (Milestone 1)
Week 10 (Milestone 2): Add walking-to-person behavior and test approach in quiet room		
10.1	Fix bugs or unstable behavior after prototype test	System improvement
10.2	Implement walking-to-person behavior and stop at 1m distance	Needs motor + person detection
10.3	Test: turn + detect + walk-to-person behavior in quiet	Second prototype

	room	(Milestone 2)
Week 11 (Milestone 3, full prototype): Implement following the human behavior and complete core Hexapolo Bot features		
11.1	Connect TFmini lidar module to Raspberry Pi	
11.2	Integrate lidar distance data into Raspberry Pi control loop	
11.3	Implement the following person behavior (continuous person tracking)	Builds on turn-detect-walk behavior
11.4	Test: turn + detect + walk-to-person behavior + follow the person in quiet room	The full Hexapolo bot prototype with all of the essential features (Milestone 3)
Week 12 (Milestone 4): *(Optional)* Add noise filtering feature and connect lidar module for obstacle avoidance while approaching visible person		
12.1	[Optional] Integrate selected noise filtering method into existing keyword detection pipeline	Needs working keyword detection
12.2	[Optional] Test: keyword detection performance in simulated noisy environment	Test until which level of noise it can handle, and document it (Milestone 4)
12.3	[Optional] Design basic obstacle avoidance behavior with lidar (stop & turn when obstacle detected)	
12.4	[Optional] Implement obstacle avoidance while walking toward visible person	
12.5	[Optional] Test: obstacle avoidance in clean room while approaching visible person	
Week 13 (Milestone 5): *(Optional)* Implement search behavior when person is not visible		
13.1	[Optional] Research and choose a method for searching the person when not visible (strategy decision in addition to what we already have, ex. Saying Marco again after 4 trial)	
13.2	[Optional] Integrate search behavior with obstacle avoidance	Requires both search logic + avoidance working
13.3	[Optional] Test full system: obstacle avoidance + search behavior	Final integrated test (Milestone 5)

5.2. Gantt Chart



https://docs.google.com/spreadsheets/d/1xnIqV8nKbMYVABpta7GscwPPvCl3ueHyudp6yGAwllQ/edit?usp=sharing