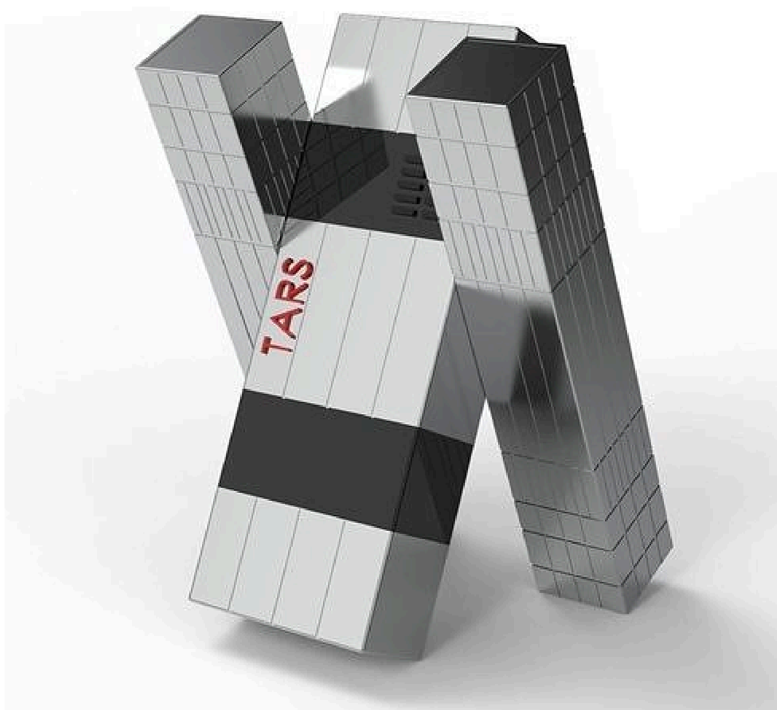
Senior Design Proposal - TARS Robot

Jack McGarrity, Garrett Young, Jack Corrao, Matthew Sims, Xander Steele



1 Introduction

Our project aims to design and build a robot inspired by TARS from Interstellar, combining the advanced functionality with a simple, humanlike design. Our TARS robot will incorporate the agility mechanics found in the movie, voice-communication, and artificial intelligence. By the end, we plan on having a companion-like robot that is able to converse, maneuver, and interact with its environment semi-autonomously.

2 Problem Description

In today's society, there is a growing demand for interactive and intelligent companions that extend beyond virtual assistants like Siri or Alexa. While these AI systems have made significant strides in voice recognition and response, they lack a physical presence that can engage users on a more personal and tangible level. Existing robotic companions in the market are often limited by high costs, simplistic interactions, or lack of adaptability, leaving a gap for a more accessible and emotionally engaging solution. This underscores the need for a small, versatile robot that can serve as a personable AI companion, offering both functional assistance and meaningful interactions.

Our project aims to address this need by designing and creating a compact robot inspired by TARS from the movie *Interstellar*. This robot will function as an AI companion, capable of understanding and responding to user inputs through speech and visuals, thereby providing a more immersive interactive experience. By integrating advanced robotics with AI technologies, we intend to develop a companion that is not only capable of assisting with everyday tasks but also of forming a personal connection with the user. This endeavor will contribute to the field of personal robotics by making sophisticated AI companionship more accessible and by pushing the boundaries of human-robot interaction.

3 Proposed Solution

To develop a compact AI companion robot inspired by TARS from *Interstellar*, the proposed solution integrates mobility, artificial intelligence, voice recognition, display capabilities, motor control, speech synthesis, music streaming connectivity, computer vision, and remote control functionality. The robot aims to interact with users through speech, visual displays, and physical movements, providing a multifaceted user experience. The solution is organized into several logical functions, each responsible for specific features of the robot.

<u>Logical Functions:</u>

1. Mobility and Motor Control (TARS-like Mobility - see Walk Modulation 2):
    - <u>Function</u>: Enables the robot to move and navigate its environment in a manner similar to TARS.
    - <u>Description</u>: The robot will have segments that can rotate or pivot to facilitate movement. Motor control systems will manage the movement of these segments.
    - <u>Key Technologies</u>: Servo motors; sensors (gyro, IMU, ultrasonic); motor controllers.
2. Artificial Intelligence Integration (AI API):
    - <u>Function</u>: Provides advanced AI capabilities for conversation, decision-making, and learning.
    - <u>Description</u>: Integration with AI APIs will enable the robot to process complex user inputs, engage in natural conversations, and learn from interactions. This could involve cloud-based AI services for robust processing.

    - <u>Key Technologies</u>: Utilization of AI APIs such as OpenAI's Realtime GPT-4 for conversational AI; cloud connectivity via Wi-Fi;
3. Voice Recognition Module:
    - <u>Function</u>: Allows the robot to understand and process voice commands from the user.
    - <u>Description</u>: Implementing speech-to-text technology to convert spoken words into text for processing. Supports natural language understanding for diverse user inputs. Features wake command to “wake-up” TARS.
    - <u>Key Technologies</u>: Microphone; OpenAI’s Realtime API; Espressif ESP-Skainet WakeNet.
4. Display Screen:
    - <u>Function</u>: Provides visual feedback and interface for the user.
    - <u>Description</u>: A built-in display screen will show information such as text responses, images, or graphical interfaces. It enhances interaction by providing visual cues and data.
    - <u>Key Technologies</u>: Integration of LCD or OLED screen with compatible library.
5. Speech Synthesis Module (Speaking Capabilities):
    - <u>Function</u>: Enables the robot to communicate verbally with the user.
    - <u>Description</u>: Text-to-speech technology will allow the robot to produce spoken responses, making interactions more natural and engaging.
    - <u>Key Technologies</u>: Speaker with amplifier; OpenAI’s Realtime API.
6. Spotify API Connectivity:
    - <u>Function</u>: Allows the robot to play music via Spotify upon user request.
    - <u>Description</u>: Integration with Spotify's API to control music playback, create playlists, and respond to music-related commands.
    - <u>Key Technologies</u>: Spotify Developer API; authentication protocols (OAuth); audio playback libraries; internet connectivity modules.
7. Computer Vision Module:
    - <u>Function</u>: Enables the robot to perceive and interpret visual information from its environment.
    - <u>Description</u>: Using cameras and image processing algorithms, the robot can recognize objects, faces, or gestures, enhancing its ability to interact with the surroundings.
    - <u>Key Technologies</u>: Cameras (e.g., USB or Raspberry Pi Camera Module); OpenCV for image processing; OpenAI’s Realtime API for object identification.
8. Remote Controller Compatibility (PS5/Xbox Controller):
    - <u>Function</u>: Allows manual control of the robot using a gaming controller.

- Description: Users can control the robot's movements and functions remotely using a PS5 or Xbox controller, providing an alternative interaction method.
- Key Technologies: Bluetooth connectivity; controller input libraries (e.g., Pygame for Python); mapping of controller inputs to robot functions.

9. Power Management System:
    - Function: Supplies power to all components and manages energy consumption.
    - Description: A rechargeable battery system with power regulation circuits will ensure efficient operation. Power monitoring will prevent over-discharge and extend battery life.
    - Key Technologies: Lithium-polymer or lithium-ion batteries; power management ICs; battery charging modules; voltage regulators.
10. Connectivity and Networking:
    - Function: Provides internet access and network communication for API integration and updates.
    - Description: Wi-Fi and Bluetooth modules will enable internet connectivity and communication with external devices like smartphones or controllers.
    - Key Technologies: Wi-Fi modules (e.g., built-in Wi-Fi, bluetooth on ESP32).

4 Demonstrated Features

- TARS-like mobility
- Conversational AI
- Voice Recognition
- Display Screen
- Motor Control
- Speaking Capabilities
- Spotify-API Connectivity
- Computer Vision
- Remote Controller (PS5/XBox)

5 Available Technologies

1) Mechanical Systems
    a) Actuators and Servos (~ $25)
    b) Frame Materials (3D-Printed)
2) Microcontroller and Processing Unit
    a) ESP32-C6 (Wifi/Bluetooth) (~ $50)
3) Wireless Communications

a) RC (PS or XBox controller) (~ Common)

4) Sensors
    a) Camera for Computer Vision (~ $20)
    b) UltraSonic Sensor for Proximity and Collision Avoidance (~ $5)
    c) Accelerometer for Balance (~ $5)
    d) Microphone for Voice Recognition (~ $5)
5) Interface
    a) OLED Display (~ $30)
    b) Speaker (~ $10)
6) Power
    a) Li-ion Batteries (~ $5)
    b) TP4056 for battery charging (~ $10)
7) Software (Open-source)
    a) PlatformIO
    b) Control Algorithms
    c) Speech and Command Processing using Speech APIs
8) Fabrication
    a) 3D-Printing
    b) Custom PCB

Current cost-analysis has us at ~ $170. This is a gross-underestimate as we anticipate the actual cost of manufacturing will exceed this amount. We did not include the cost of 3D-printed materials into our analysis.

6 Engineering Content

1. Mechanical and Mobility Systems

We need to design the body frame and implement the unique movement mechanism of TARS. This involves designing the frame to ensure structural stability and house all the other electrical components while keeping in mind weight distribution. We also need to integrate servomotors and develop control algorithms to allow precise control and coordinated movement.

2. Power

We need to be able to power all the motors and electronics within TARS by battery. This means we need to design circuits to allow us to safely and efficiently charge the battery and make sure it is not overcharged or over-discharged. All of the different electronic components will have different power requirements so we need to design circuits to manage the battery power so that every device gets what it requires.

3. Sensors

We need to design circuits capable of interfacing with all the proposed sensors integrated into TARS, including the camera, ultrasonic sensor, accelerometer, and microphone. Additionally, we will develop software to process the sensor data and enable the robot to interpret and respond to its environment effectively.

4. Wireless communications (WiFi/bluetooth)

We need to integrate WiFi and bluetooth capabilities to enable connectivity for cloud-based APIs and external utilities such as the remote controller. The antennae must ensure strong connections and protocols must be secure.

5. Hardware Organization

We need to decide on how to organize our required tasks at the hardware level. This includes the number of boards to be designed, microcontrollers to be used, and the tasks to be assigned to each board (i.e motor control/power, main comms/I/O).

6. Speech to speech conversion

We need to develop a solution that accepts audio voice input, convert the speech to text, run the text through an LLM, get the output of the LLM, and convert the text output to speech. OpenAI's Realtime API simplifies these tasks into one API call which would be optimal for our low-power MCU design.

6. Output Interfaces

We need to seamlessly integrate output interfaces, including speaker and display modules. The display's text output should sync up with the audio output to enhance the user experience.

7 Conclusion

Our project aims to create an innovative robotic companion inspired by TARS from *Interstellar*, bridging the gap between functional AI systems and personable, interactive robots. Our design focuses on creating a robot capable of TARS-like movement, human interaction, and conversational task execution. Utilizing a lightweight frame with servo-motor joints for mobility, sensors for environmental awareness, and a listen-learn-speak framework for communication, the robot will provide an engaging and dynamic user experience. Powered by an ESP32-C6 microcontroller, it will support wireless operation, real-time processing, and user-friendly interaction through a display and speaker. By integrating advanced technologies such as AI, voice recognition, and computer vision into a compact and cost-effective design, our TARS-inspired robot will serve as a fun and helpful personal assistant, bringing a beloved movie character to life in a meaningful way.