Enhancing Assistive Home Environments through IoRT Multi-Agent Interaction: Model Selection and Robotic Operations

E. A. B. Huayhua[1], E. Rohmer[2], C. B. da Silva[3], F. A. O. Mota[4], B. M. Portela[5]
[1]Dept. Computer Engineering and Industrial Automation, FEEC, UNICAMP

Introduction: The ongoing improvement in the research and development of LLM (Large Language Models) aligns fascinatingly with Mobile Robotics, presenting a promising trend in artificial intelligence. This convergence holds the potential to revolutionize assistance, especially in the care of adults or individuals with disabilities. Human-machine interaction emerges as a solution, aiming to facilitate daily activities for this group. These models, akin to chatbots and virtual assistants, offer informative assistance. Additionally, there is the possibility of simultaneous assistance from multiple robots for a user or patient, requiring exploration in "Multiple Agent Assistance." This project focuses on a chatbot with a dual purpose: providing voice assistance based on user needs and maintaining an updated database about the environment. These tasks are fundamental to creating an intelligent environment, relying on the project's architecture, which aims to establish a robust connection between robotic devices through a REST server for shared and updated context among agents.

Materials and Methods: To carry out the robot simulations, a personally manufactured humanoid robot developed by B. M. Portela was employed. This robot is utilized for interaction in a specific environment, using specific joints that are controlled from a Python code to Naoqui. Additionally, the simulations included the interaction of the P3DX robot, representing the locomotion part of the robot in the same environment. We also took into account the presence of other IoT agents, represented by AXIS PTZ 214 cameras and Google Assistant, coexisting in the intelligent environment to provide enhanced assistance to the patient. The objects captured by the cameras serve the purpose of contextualizing communication with the Large Language Model (LLMs). In this regard, the LLMs establishes communication with a REST server, responsible for storing all relevant information generated during the simulations.

Results: In the analysis using OpenAI's LLM, ChatGPT version 4.0 demonstrated its ability to generate desired responses during interaction, surpassing other models. Although the variability in prompt usage among models with equal parameter sizes is acknowledged.
ChatGPT 4.0 stands out in providing information on multiple tasks within a conversation. The evaluation was divided into three fields: direct interaction with a user, interaction with the ROS control system, and, finally, supplying context to the REST server. This approach allowed assessing its performance in different contexts, highlighting its versatility in addressing various tasks in the robotic assistant application. The ability to provide real-time contextual information through ROS and maintain coherence in user interaction underscores the effectiveness of ChatGPT 4.0 in this specific environment.

Discussion/Conclusion:
In this project, the initial step involved interacting with multiple models to partially identify the most suitable Large Language Model (LLM) for real robot interaction. The successful selection of OpenAI's model was achieved in part. Additionally, diverse information was gathered to address the Multiagent field, diversifying communication with a server. This allows other agents to interpret actions and multitasking functions.
In the next phase, the aim is to enhance the model's activity by enabling it to choose actions within the Robot Operating System (ROS) environment. The model is expected to maintain a sequence of activities for each agent and pose follow-up questions to facilitate processing within the ROS architecture and cameras. This approach seeks to empower the model's decision-making capability and enable more active participation in system operations.