visuAAL

Privacy-Aware and Acceptable Video-Based Technologies and Services for Active and Assisted Living

Trinity College Dublin
Coláiste na Tríonóide, Baile Átha Cliath
The University of Dublin

Improving Home-Based Care Robots' Capabilities using Natural Language Interface

Hassan Zaal

Background

- In the EU, the old-age dependency ratio ($\frac{\geq 65}{15\ to\ 64}$) is projected to grow from 29.6% in 2016 to 51.2% in 2070 [1]
- In ageing societies, the demand for long-term care will increase while there will be shortages in labor to meet this demand [2]
- Recent studies indicate evidence that robotic interventions could support "ageing in place" [3]
- Natural language interaction with robots is challenging, especially when translating high-level abstract instructions to the robot's capabilities
- Previous approaches vary from a set of predetermined instructions to explicit instructions such as Vision-and-Language Navigation. Recently, others have used a set of skills that the robot can perform and use LLMs to select from this set of skills [4]
- LLMs can encode an extensive of semantic knowledge about the world
- A significant **weakness** of LLMs is that they **lack real-world experience**, which makes it difficult to leverage them for decision-making within a given **robot's capabilities**

Aim and Objectives

Aim: To enhance communication and interaction capabilities between the robot and human within the home environment, thereby supporting older adults to better ageing in place. Tho achieve this, the research will address the following objectives:

1. Improve the robot's ability to understand and execute user verbal commands, utilizing extracted information from the environment and the robot's capabilities.
    1) what the robot can/can't do
    2) what the robot can/can't sense
2. Understand the role of LLMs in facilitating the objective
3. To demonstrate the outputs using a simulated environment.

Method

To achieve the aim and objectives the following will be implemented:

- Using pre-trained LLMs to analyze the end-user instructions and convert them into structured low-level commands that the robot can understand and execute
- Evaluate the ability of the robot to adapt to the changes in the environment to perform the user instructions
- Evaluation via NVIDIA Isaac Sim

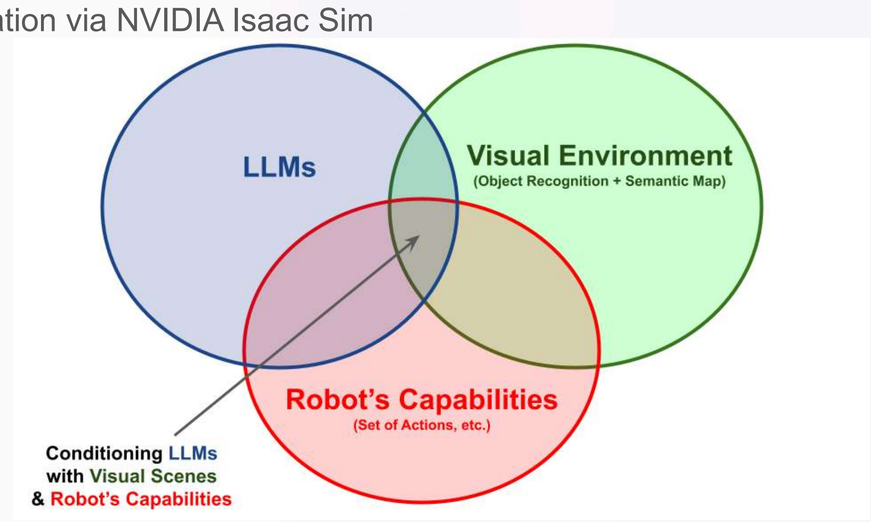

Fig 1. Conditioning LLMs with Visual Scenes & Robot's Capabilities

LLMs to Control Robots

Spatial AI for Scene Description

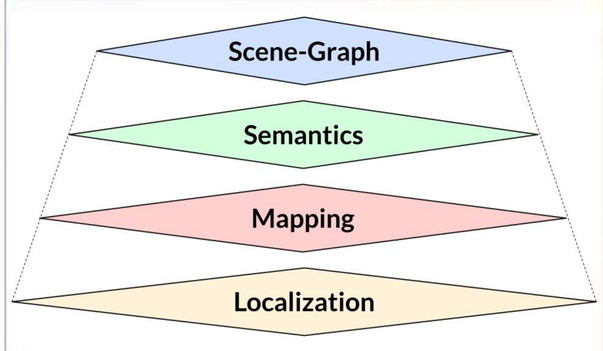

Fig 2. 3D Spatial Perception (source [5])

Level 4: **Scene-Graph**:

How is the world organized?

Level 3: **Semantics**:

What are the objects around me?

Level 2: **Mapping**:

Where are the objects around me?

Level 1: **Localization**:

Where am I?

Prompt Elements for Robotics

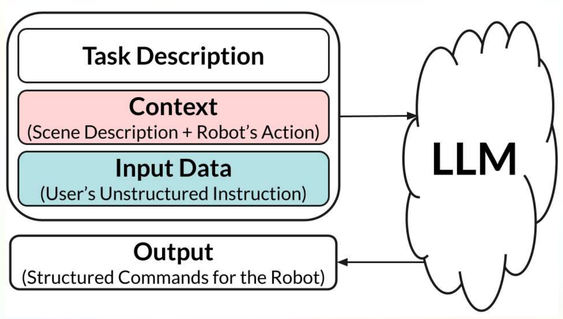

Fig 2. Elements of Prompt Engineering for Robotics

Progress

LLMs can effectively be used to convert abstract, unstructured user instructions into sub-tasks, which are ultimately converted into structured commands that the robot can execute

The next steps are:

- To evaluate the robot in a hierarchically organized environment
- To test on different robot embodiments (has manipulator)

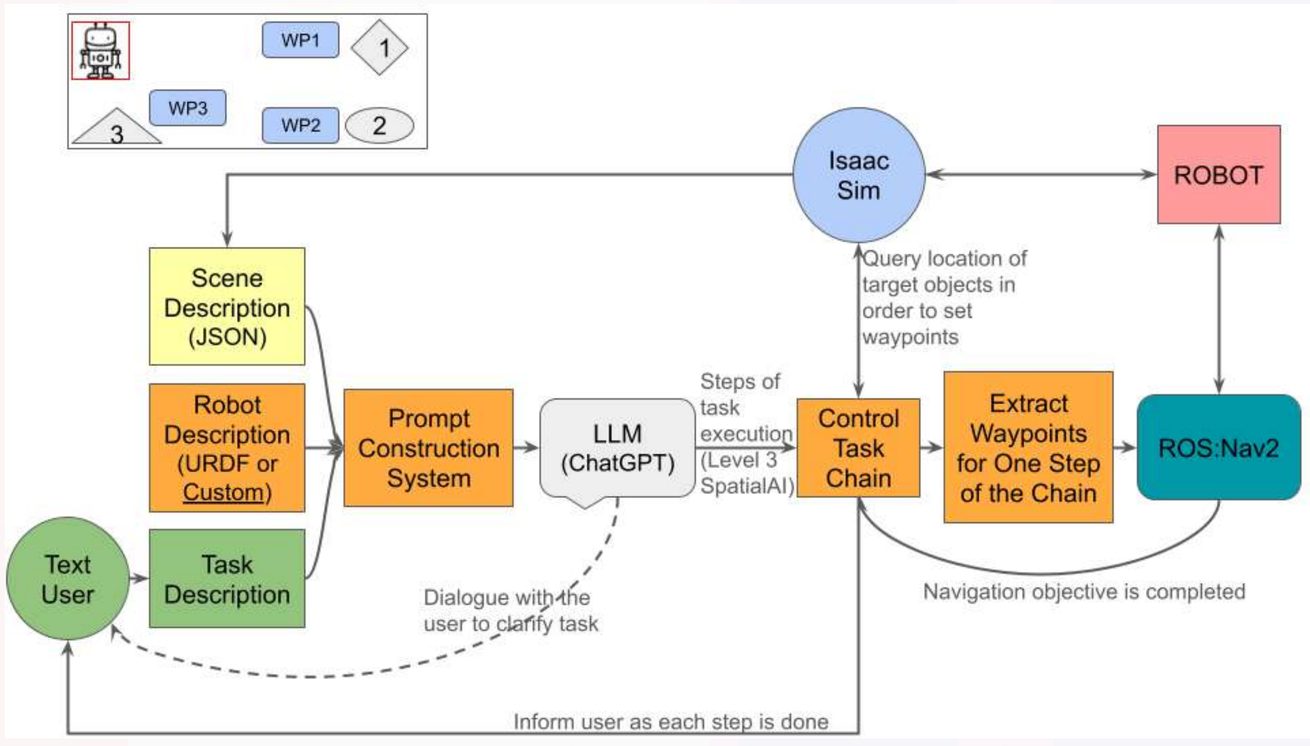

Fig 4. The Proposed Architecture for Using LLMs to Control the Robot in Isaac Sim

Impact

- Parse abstract language instructions provided by the end-user with objects and locations and convert them into navigation commands within constraints of the robot's capabilities
- Robots hold promise in enhancing the quality of life for older adults and caregivers by assisting with ADL, IADL, and medical needs

About Me

- Master by Research in Computer Science
- Postgraduate Certificate in Innovation & Entrepreneurship

  Trinity College Dublin, The University of Dublin

  (Currently)
- Postgraduate Diploma in **Vision and Robotics (VIBOT)**

  Heriot-Watt University, UK
- BSc in **Computer and Automation Engineering**

  Damascus University, Syria
- Completion of research estimated Summer 2024
- Disseminating outputs of the work until the end of the fellowship/ exploring new opportunities

References

1. European Commission (2018), The 2018 Ageing Report, Economic and budgetary projections for the 28 EU Member States (2016-2070), Directorate-General for Economic and Financial Affairs
2. New Job Opportunities in an Ageing Society : Paper Prepared for the 1st Meeting of the G20 Employment Working Group, 25-27 February 2019, Tokyo, Japan. 2019. Geneva: ILO.
3. Mois, G., & Beer, J. M. (2020). Robotics to support aging in place. In *Living with robots* (pp. 49-74). Academic Press.
4. Brohan, A., Chebotar, Y., Finn, C., Hausman, K., Herzog, A., Ho, D., ... & Fu, C. K. (2023, March). Do as i can, not as i say: Grounding language in robotic affordances. In *Conference on Robot Learning* (pp. 287-318). PMLR
5. Rosinol, Antoni. *3D Spatial Perception with Real-Time Dense Metric-Semantic SLAM*. Diss. Massachusetts Institute of Technology, 2023

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 861091".

Universitat d'Alacant
Universidad de Alicante
Project Coordinator

RWTH AACHEN UNIVERSITY

Stockholm University

Trinity College Dublin
Coláiste na Tríonóide, Baile Átha Cliath
The University of Dublin

TU WIEN