GOODBROTHER / VISUAAL CONFERENCE



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

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

ESR 9. HASSAN ZAAL





Overview

- Background
- Healthcare Robots for Old Adults
- Work Aim and Objectives
- Background Research on Large Language Models (LLMs) for Robotics
- Proposed Architecture for using LLM to Control Robot
- Conclusion





Background

- In the EU, the old-age dependency ratio $\left(\frac{\geq 65}{15 \text{ to } 64}\right)$ 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]
 - European Commission (2018), The 2018 Ageing Report, Economic and budgetary projections for the 28 EU Member 1. States (2016-2070), Directorate-General for Economic and Financial Affairs
 - 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.
 - Mois, G., & Beer, J. M. (2020). Robotics to support aging in place. In Living with Robots (pp. 49-74). Academic Press.













Use Case Scenario



Cara, 87 years old, Ireland

"I am looking for technologies to help me to be more independent, and they can understand me."

- Retired, Ex-kindergarten teacher
- Widowed, 1 son died 3 years ago, 3 grandchildren
- Suffers from osteoarthritis.
- Suffer from lung cancer.
- Suffer from visual impairment.

Cara, 87 years old, used to work in kindergarten as a teacher. Her husband died 15 years ago and her son died 3 years ago. Her grandchildren work overseas, so she does not see them that much.

Cara has a lung cancer and she is under medications.

Cara wears glasses to correct her visual impairment. Unfortunately, she always forget them after bath, waking-up, etc. She usually spend times looking for her glasses, key, etc. **Cara likes gardening, but she** cannot carry on the gardening tools.

What she would like is a personal assistant robot that can interact with similar to the interactions between humans. She would like the robot to understand her spoken language and allocate to her missing glasses or a missing key, etc. Also, **She wants the robot to carry for her the tools**. **She wants the robot to remind her about her medicine to fetch them**. She would like the robot to help her with her self-management needs.







Healthcare Robots for Old Adults



ASTRO



Jaco



Care-O-bot 4



ReWalk







Astro (Amazon)











5

Evaluation of Healthcare Robots for Users' Needs

Users' Needs: ADLs/IADLs/Social	ASTRO	Jaco	Care-O-bot 4	ReWalk	PARO	AIBO	Astro (Amazon)*	Bomy
Dressing								
Personal Hygiene								
Feeding		Х						
Toileting								
Transferring/ Walking	X			X				
Transportation/ Shopping								
Managing Medications			Х					Х
HouseWork								
Social					X	Х	Х	Х
Indoor Objects Transferring			×					





Work Aim and Objective

• Aim: To enhance communication and interaction capabilities between the robot and human within the home environment

- Objectives:
 - 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
 - Understand the role of LLMs in facilitating the objective





Grounding Language Visual Scenes

- By the time we reach the opposite <u>bank</u>, the <u>boat</u> was sinking fast.
- I had to take out a **bank** loan to start my own business.

Captioning: a cat staring out the window at a group of birds.

FOIL: a dog cat staring out the window at a group of birds.

Referring Expression, Recognition: Bird to the left of the feeder.

Visual Question Answering: Q: How many birds are there? A: four





Q: Left image has twice as many cats as the right image, and at least two cats are black. **A:** True

Visual Dialog:

A-Bot: Image shows a cat staring out the window at a group of birds.

Q-Bot: How many cats are there ? **A-Bot:** 1

Q-Bot: Can you see its face? [*it* = *cat*; *visual coreference*] **A-Bot:** no

Q-Bot: I think we were talking about **Image 2**.







Background Research on Vision-and-Language Navigation







Background Research on LLMs for Robotics





CONFIDENTIAL - Do not disclose this information to any third party without the prior written consent of the Disclosing Party

Background Research on LLMs for Robotics

 Large Language Models (LLMs) can encode a wealth of semantic knowledge about the world



 A significant weakness of language models is that they lack real-world experience, which makes it difficult to leverage them for decision-making within a given robot's capabilities.





Background Research on LLMs for Robotics - SayCan







LLMs for Robotics - Limitations in SayCan

- The robot is equipped with a repertoire of learned skills
 => specific responses.
 - It does not adapt to new environments, or different robot's capabilities.
- Go to: It is assumed the location of objects are known => these are specified as "go to location" and "find object".
- The planning does not take into account Spatial AI and semantic maps in real world (different rooms, etc).

CONFIDENT the prior writ







How to Describe an Environment Using Spatial AI

- 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?



14



Method

 To analyze the end-user instructions using pre-trained LLMs 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







Method – Prompt Engineering for Robotics

- LLMs are N-gram models that learn to predict N-th word given the preceding N-1 words.
- LLMs learn compact representations of conditional distributions given the data.
- Prompting is increasing the contents of the context window.







Standard Ways to Improve LLM Response

- Prompting ("in-context learning") [doesn't change LLM parameters]
 - If you don't like what an LLM is giving as an answer to your prompt, you can add additional prompts
 - The LLM will then take the new context window (including what it said and what you said) to predict the next sequence of words/tokens.
 - Every word in the context window including the ones LLM added is changing the conditional distribution with which the next token being predicted.
 - □ Note that all these conditional distributions have been precomputed!
 - o Nothing inside LLM is changing because of your prompts
 - The undeniable attraction of "prompting" is that it is natural for us! It is sort of how we interact with each other!
 - □ If you give k example of good answers as part of the prompt, it is called "k-shot in-context learning"

Fine tuning [Changes LLM parameters]

- Fine tune the parameters of a pre-trained LLM by making it look specifically at the data of interest to you
- Because fine tuning is changing the parameters of the LLM, while its performance on specific task (be a better planner, be less offensive) may be improve, it also changes its performance on other task in unpredictable ways.



- Do not disclose this information to any third party without

consent

CONFIDENTIAL the prior written co



Prompt Engineering for Robotics

Expected the robot

has an arm

You are a guide for a domestic robot that can only navigate inside a room. This room has a sofa, a table, a cup of coffee on the table, and another person next to the table, write in order steps how the robot will bring me a coffee. In case some steps you cannot perform ask the user in the room to do that step and when he has done it to write DONE to continue.

- Certainly, here are the steps for the domestic robot to bring you a cup of coffee:
- 1. Start in a ready position: Ensure that the robot is in a ready state to receive commands.

000

- Locate the coffee cup: Use your built-in sensors and cameras to identify the coffee cup on the table.
- Calculate the distance: Determine the distance from the robot's current position to the table where the coffee cup is located.
- Plan a path to the table: Based on your map of the room and the distance calculation, plan a
 path to reach the table. Ensure that there are no obstacles in the way.
- Navigate to the table: Start moving towards the table using your wheels or tracks. Be cautious and slow to avoid any collisions.
- Stop at the table: When you reach the table, stop a safe distance away to prevent knocking anything over.
- Extend your arm: If you have a robotic arm, carefully extend it toward the coffee cup on the table.
- 8. Grip the coffee cup: Use your gripper or arm to grasp the coffee cup securely.
- 9. Lift the coffee cup: If your robot has the capability, lift the coffee cup off the table.
- Plan a path back to the user: Calculate a path to return to the person sitting next to the table.
- Navigate back to the user: Start moving back to the user, making sure to avoid any
 obstacles or collisions.
- Stop in front of the user: When you reach the user, stop in front of them, ensuring you're at a comfortable distance for them to take the coffee.
- 13. Lower the coffee cup: If you have lifted the coffee cup, carefully lower it to the table i G Regenerate





Prompt Engineering for Robotics

You are a guide for a domestic robot that can only navigate inside a room. You cannot pick up or place any objects, but you have a tray on top of you that a user can place objects on it. This room has a sofa, a table, a cup of coffee on the table, and another person his name Ali and he is next to the table, write in order steps how the robot will bring me a coffee. In case some steps you cannot perform ask Ali who is in the room to do that step and when he has done it to write DONE to continue.

- Certainly, here are the steps for the domestic robot with a tray to bring you a cup of coffee, involving Ali's assistance where necessary:
- 1. Start in a ready position: Ensure that the robot is in a ready state to receive commands.
- Locate the coffee cup: Use your built-in sensors and cameras to identify the coffee cup on the table.
- Calculate the distance: Determine the distance from the robot's current position to the table where the coffee cup is located.
- Plan a path to the table: Based on your map of the room and the distance calculation, plan a
 path to reach the table. Ensure that there are no obstacles in the way.
- Navigate to the table: Start moving towards the table using your wheels or tracks. Be cautious and slow to avoid any collisions.
- Stop at the table: When you reach the table, stop a safe distance away to prevent knocking anything over.
- 7. Notify All: Politely inform Ali that you're ready to assist with bringing coffee.
- Request Ali to place the coffee on your tray: Ask Ali to place the coffee cup on the tray on top of you.
- 9. Wait for Ali to place the coffee: Allow Ali to place the coffee cup on your tray.
- 10. Verify the placement: Use your sensors to ensure that the coffee cup is securely placed on

your tray. If it's not, politely request Ali to adjust it.

- 11. Request Ali to command you to "serve": Ask Ali to give a command, such as "Serve coffee."
- 12. Wait for Ali's command: Wait for Ali to give the "serve" command.
- 13. Once Ali issues the "serve" command, navigate back to you: Based on Ali's comma





Method– Connecting ChatGPT with Robotics Operating System







The Proposed Architecture for using LLMs to Control the Robot







Control the Robot in Isaac Sim using ChatGPT









Navigation To Goal







What is Next?

- Test with different robots' embodiments.
- Study the issues that LLMs have, but for Spatial AI
 - Hallucination
 - Memory
 - Safety, etc.







Conclusion

 Robots hold promise in enhancing the quality of life for older adults and caregivers by assisting with ADL, IADL, and medical needs

• LLMs can encode an extensive of semantic knowledge about the world

Using pre-trained LLMs can enhance communication and interaction capabilities between the robot and human within the home environment
 Expected impact: supporting older adults to better ageing in place.









Thank You for Listening

zaalh@tcd.ie