MARCER: A Multimodal Augmented Reality System for Composing and Executing Robot Action Plans

Bryce Ikeda* bikeda@unc.edu

Maitrey Gramopadhye* maitrey@email.unc.edu

LillyAnn Nekervis lnekervis@unc.edu

Daniel Szafir dszafir@cs.unc.edu

THE UNIVERSITY of NORTH CAROLINA at CHAPEL HILL

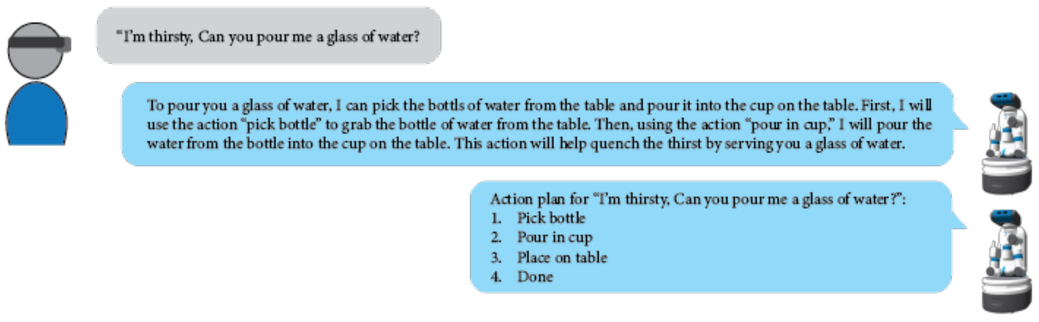
IRON LABORATORY

Introduction

MARCER is a multimodal robot programming system that combines natural language and augmented reality for an intuitive user experience. Users program robots by articulating commands like 'Move the groceries that are on the table to the shelves.' A Large Language Model (LLM) interprets these commands and integrates environmental context to create editable triggers and Action Plans. An augmented reality (AR) interface allows users to set virtual constraints, visualize and edit trigger-action pairs and plans, and use a digital twin to preview and adjust robot activities in real time.

System Design

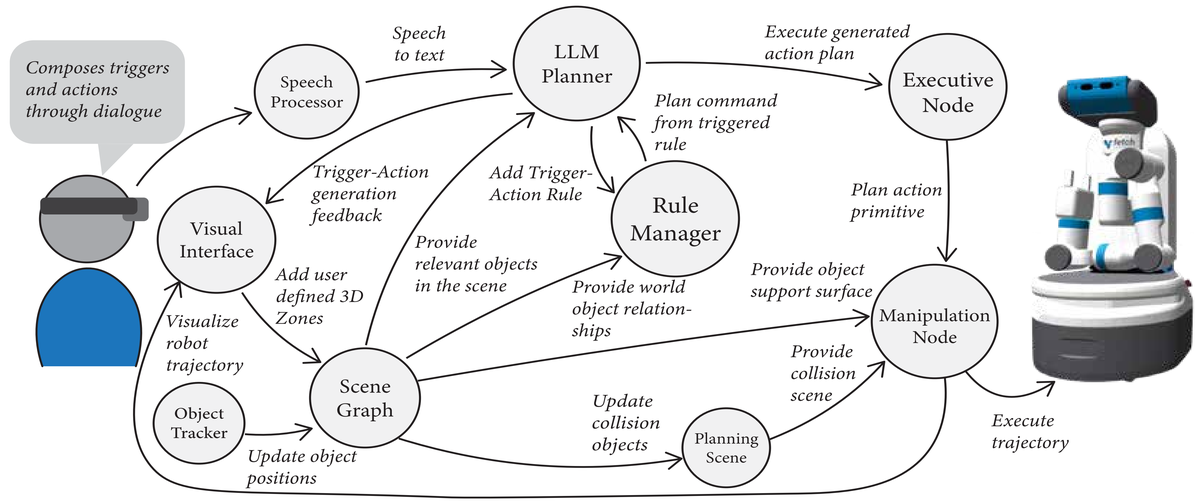
1. LLM Planner: Generates action plan from natural language input



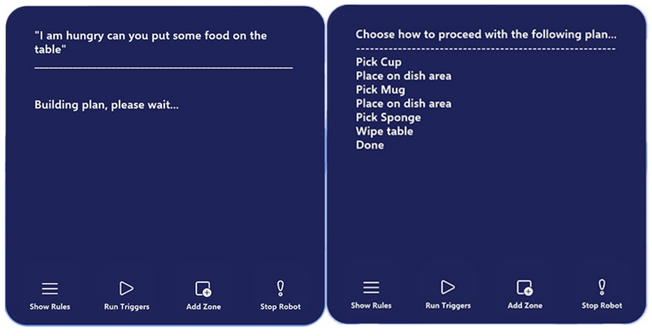
2. Translate detailed reasoning into actionable steps for robot execution.

3. Natural Language Input -> LLM generated Action plan -> Execution with digital twin preview

System Implementation



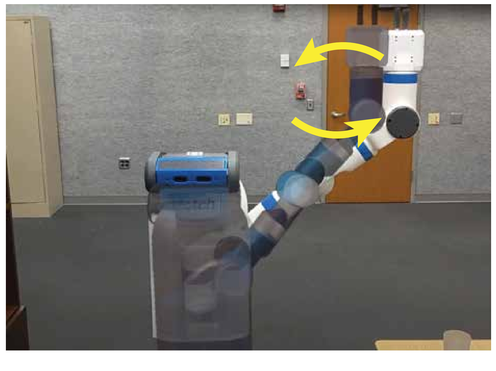
Applications



Direct Verbal Commands

Remote Interaction via World in Miniature (WiM)

Reactive Tasks and Zone Triggers



Social Interaction