Development of an Embodied AI Lifelike Social Companion Robot

Mahdi Bamdad*, Cha Seung, Zhijin Meng, Francisco Cruz

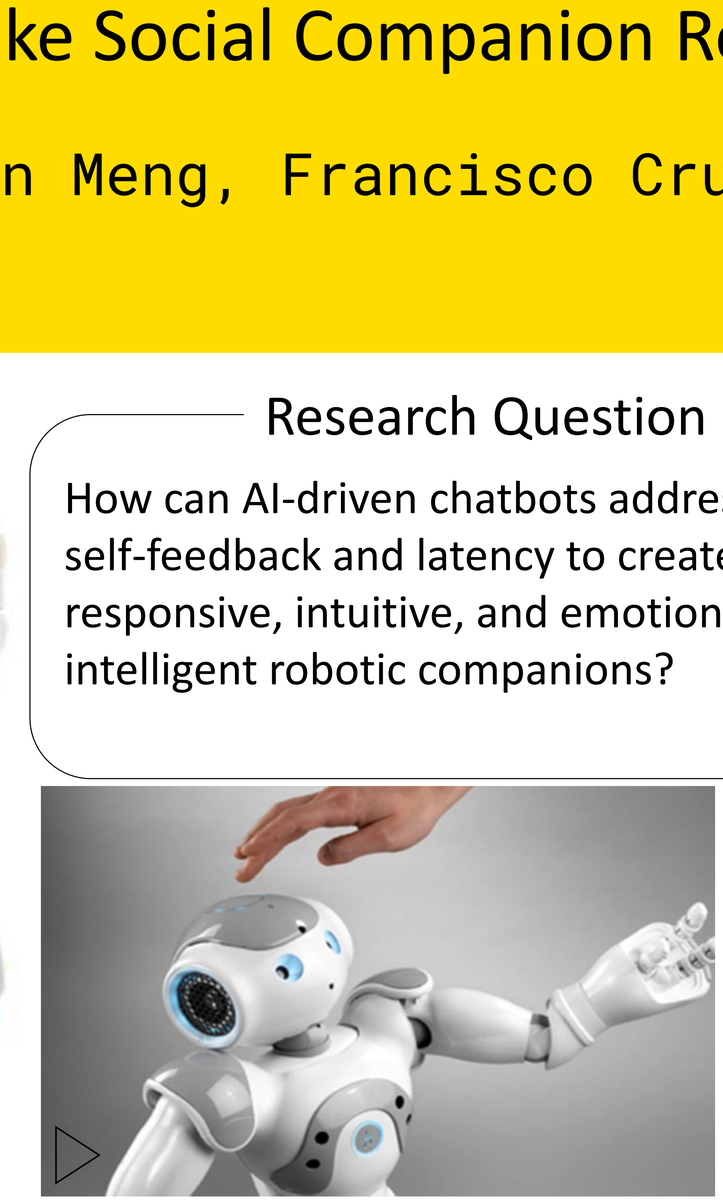
School of Computer Science and Engineering

Motivation

- ***A***chieving seamless and immersive conversation.
- ***B***ridging the gap between companionship and functional utility.
- ***C***reating robots that engage users in a way that feels emotionally intelligent, dynamic, and responsive.
- ***D***esigning an open-source Natural Language Processing.
- ***E***nhancing immersive conversations through ARI and NAO.

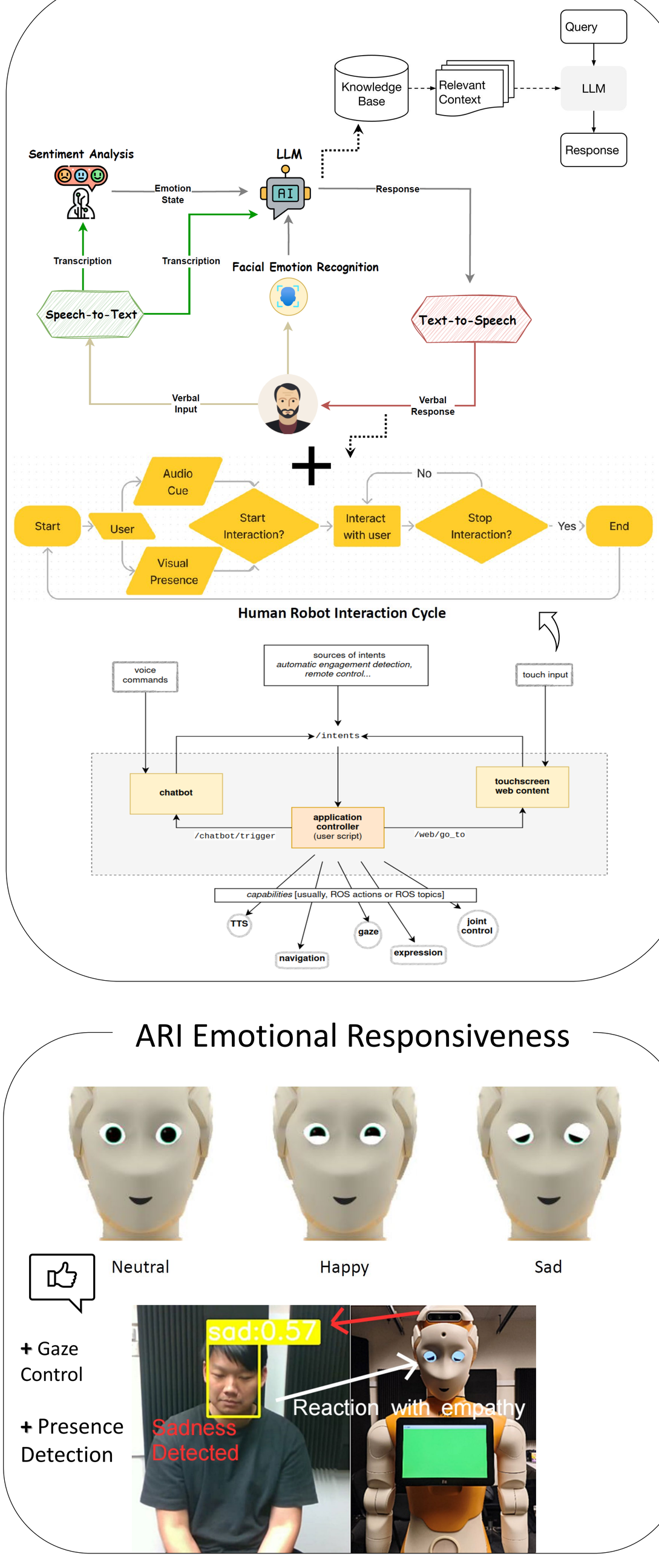
Research Question

How can AI-driven chatbots address self-feedback and latency to create responsive, intuitive, and emotionally intelligent robotic companions?



Video Demonstration

Framework and Models

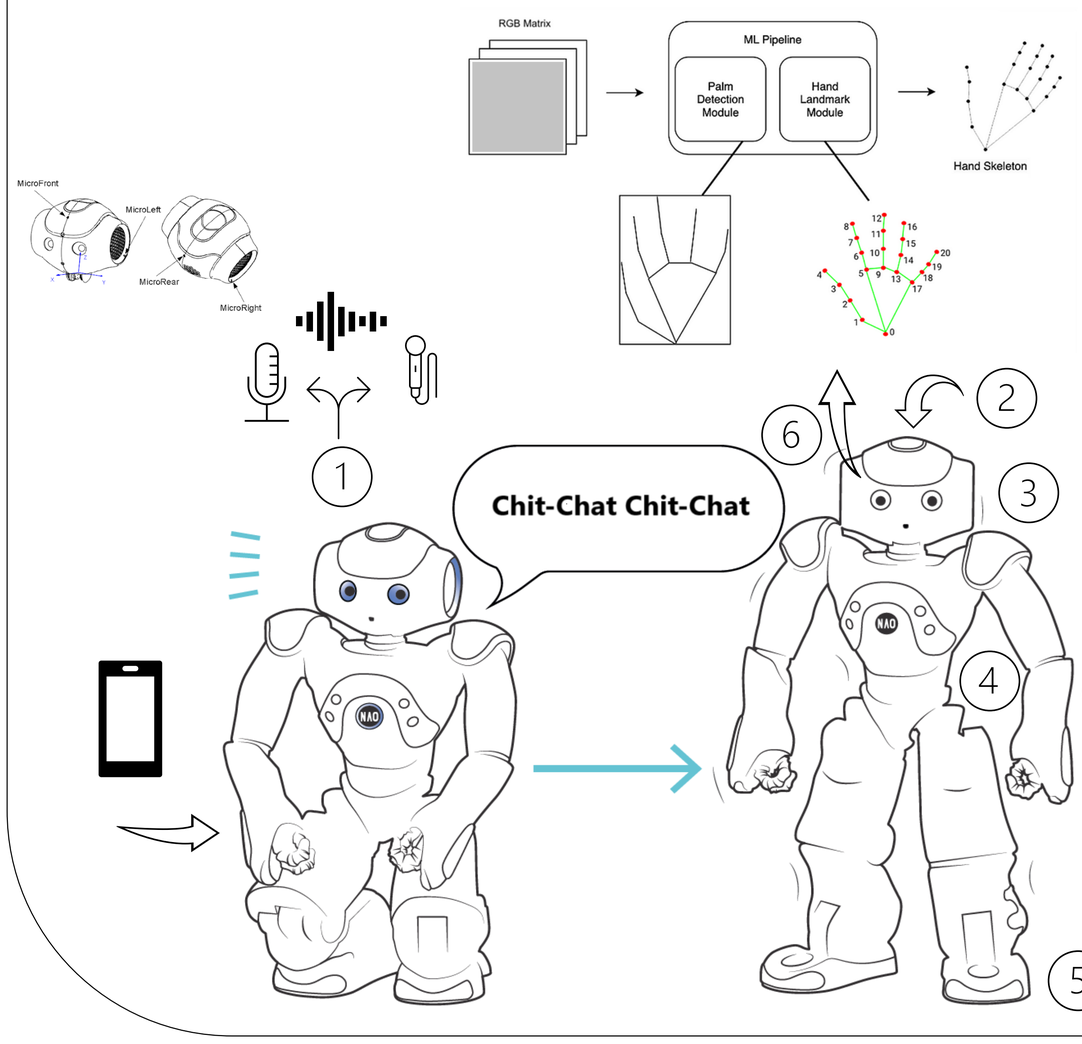


Problems with Robot Chatbots

- **Intuitiveness**: Robot's responses may feel unnatural/disjointed because it needs to wait until it finishes speaking (TTS output) before it can recognize and respond to the user's next input.
- **Emotional Intelligence**: Limited ability to detect/respond appropriately.
- **Monotony**: Robotic voices and repetitive dialogue patterns reduce engagement.
- **Echo Interference**: Robot mistakenly processes its own voice as input, causing conversational errors. Microphone Crosstalk, Self-Voice Recognition Error, Speech-To-Self Error, Audio Loopback.
- **Integration**: Difficulty syncing speech with gestures/non-verbal communication.
- **Hardware**: Limited local processing power (complex AI models).
- **Latency**: Delays in generating responses, especially when dependent on cloud-based LLMs.
- **Data Privacy Concerns**: Risk of user data leakage when relying on cloud services for processing.

NAO Conversational Intuition

- **Sensor-Based Trigger Mechanism**: Utilizing bumper, sonar, camera, and tactile sensors to stop or activate robot functionality.
- Leveraging light advanced AI and minimizing delays or awkward pauses
- Allowing the robot to listen while speaking
- Lightweight GUI web application for displaying chatbot dialogue
- Detecting an open palm gesture to signal the robot to stop its action.



ARI Emotional Responsiveness

Neutral Happy Sad

+ Gaze Control

+ Presence Detection