Research Statement

Raghav Arora
raghavaurora@utexas.edu

Efficient representation development lies at the heart of creating Deep Neural Networks (DNNs) that generalize well in the real world. Slight inaccuracies in this process can lead to catastrophic consequences in high-stakes domains like healthcare and autonomous driving. At Learning Agents Research Group, UT Austin, **I develop representations to make robust and Out of Distribution (OOD) generalizable models for Embodied AI**. My future plans include doing Ph.D. in representation learning and Robotics.

I have a strong background in Electrical and Electronics Engineering and have honed my skills in programming and leadership through multiple collaborative projects. My first hands-on robotics experience was during my undergraduate where I learned about Control Systems, Signal Processing, and Neural Networks. This led me to collaborate with Prof. Meetha Shenoy on seamlessly mosaicing aerial images from a swarm of Unmanned Aerial Vehicles (UAVs). CNNs in conjunction with mosaicing outperformed the then SOTA, but we faced a tradeoff concerning scene generalization and the UAV compute power. I faced a similar problem while on another research project with Dr. Prashant Manohar on molecular dynamics simulation of room temperature ionic liquids. Here, we used dense neural networks to approximate NP-hard Schrodinger's equations, but our challenges still remained in developing *efficient molecular representations*. This led to my collaboration with Prof. Alexandre Tkatchenko on the development of efficient molecular representations.

To work further on the representational challenges in molecular modeling, I visited Prof. Tkatchenko's lab as a student researcher. Here, I pursued my thesis on molecular property prediction based on learned representations in a DL framework. I explored a range of representation learning approaches like **Graph Neural Networks(GNNs) and autoencoders** to effectively capture salient chemical features to create effective molecular descriptors. This was then combined with the outputs of semi-empirical approaches like DFTB for generating electronic descriptors with existing geometrical descriptors like SLATM. This model for predicting atomization energy achieved a testing error of 0.35 kCal/mol, outperforming the current state-of-the-art model, SchNet, which had a testing error of 0.51 kCal/mol. **This project highlighted the importance of creating efficient domain representations that lie at the heart of OOD generalization.**

As a researcher at RRC, I focused on advancing AI by formulating embodied agents that can reason about their domains. In my first project with Prof. Madhava Krishna and Dr. Krishna Murthy, we aimed at providing commonsense reasoning to an agent for robotic Scene Rearrangement. We generated an embedding space for the knowledge base of a household environment based on object and room similarity using GNNs. We also created a **novel contrastive loss function to tackle OOD generalization**. Here, we used a weighing factor based on the similarity of the nodes, obtained from Human Annotation Dataset. Our work was published at IEEE RO-MAN 2023 and I presented it at EEML summer school this year. In both venues, I received constructive feedback on the global world knowledge of foundational LLMs, along with the correctness of plans of classical planners which can be used to extend our research.

We incorporated this feedback into the efficient execution of daily household tasks by integrating classical symbolic planning techniques with modern neural methods. We designed an intelligent agent that can anticipate tasks by learning the pattern in previous tasks. While most existing planning systems rely on smaller domains, we created an extensive household domain in PDDL (Planning Domain Definition Language). I encoded expansive commonsense knowledge of home environments into the domain to generate plans comprising sequences of robotic actions tailored to the inhabitants' lifestyles. While learning about the commonsense displayed by LLMs from Dr. Krishna Murthy, we used in-context-learning to anticipate future tasks, which **reduced the time delays by 31%**, demonstrating the value of learned knowledge. This work has been accepted at IEEE-ICRA 2024 and

was presented at Yokohama, Japan.
Currently, at Learning Agents Research Group, I am actively working with Prof. Peter Stone, and Dr. Roberto Martin-Martin on making an LLM-modulo task and motion planning framework. Like my previous projects, we focus on a robot agent in an unknown household environment, with partial observability. This work involves generating initial beliefs of objects' poses and their locations using LLM commonsense and updating the beliefs using a correlational particle filter. The basic intuition of using a correlational model is that similar objects are more likely to be stored close to each other, so we can update any object's belief pose using its similarity with other *observed* objects. This way, an agent will be able to learn and locate objects in any unknown household environment. We simulate the experiments in pybullet using PDDLstream, and realistic Kitchen-Worlds simulator. We also perform real-world experiments using Toyota's HSRB robot in a simulated kitchen environment.

I am also a member of the UT Austin Villa@Home team that will be participating in the 2025 RoboCup@Home competition. This competition aims at the development of robots with personal domestic applications. We focus our system on domains such as Human-Robot Interaction, Object Recognition (under natural and uneven light conditions), Planning, and Ambient Intelligence.

In summary, I am interested in machine learning and its applications in real world, and always seek to learn new concepts from my peers and colleagues. Additional details about my research and publications are available at my website: raraghavarora.github.io

Thank you for considering my application.