# Personalized and Interpretable Artificial Intelligence for Positive Human-Robot Interaction

Rohan Paleja, Technical Staff Researcher
Artificial Intelligence Group, MIT Lincoln Laboratory

Abstract: Collaborative robots and machine-learning-based virtual agents are increasingly entering the human workspace with the aim of increasing productivity, enhancing safety, and improving the quality of our lives. These agents will dynamically interact with a wide variety of people in dynamic and novel contexts, increasing the prevalence of human-machine teams in applications spanning from healthcare and manufacturing to household assistance. My research aims to create transparent systems that can support users and interact with humans, pushing the frontier of real-world robotics systems towards those that understand human behavior, maintain interpretability, communicate efficiently, and coordinate with high performance. In this talk, I will cover a set of works that enable robots to 1) coordinate well with other agents utilizing simulated experience and communication, 2) understand and learn from diverse human users, 3) learn interpretable, human-readable tree-based control policies directly via reinforcement learning, and 4) provide users with information online to improve situational awareness and facilitate effective human-robot collaboration.
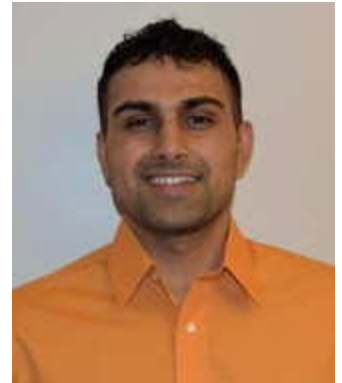
About the Speaker: Dr. Rohan Paleja is a Technical Staff Researcher in the Artificial Intelligence Group at MIT Lincoln Laboratory, where he is researching new approaches to help operators understand, program, utilize, and team with language and logic-based systems. His aim is to develop new computational methods to support robot learning and human-machine collaboration in the diverse and unstructured environments that will be encountered by these agents. Prior, Dr. Paleja completed his Ph.D. in Robotics at the Georgia Institute of Technology in the Cognitive Optimization and RElational (CORE) Robotics Lab, advised by Dr. Matthew Gombolay. His research interests cover a broad of topics, namely Explainable AI, Human-Machine Teaming, Multi-Agent Systems, and Robotics. During his Ph.D., he received a Best Paper Finalist Award at the Conference of Robot Learning (CoRL) and a Best Workshop Paper Award at the ICCV Multi-Agent Relational Reasoning Workshop.

SEP 22, 2023
Friday
12 pm – 1 pm
ENGR 201A

Lunch will be served!