SIDDHANT BHAMBRI

699 S Mill Ave, Tempe, AZ 85281 - United States of America

+1 (480)-227-4177 ◇ siddhantbhambri@asu.edu

Siddhant Bhambri ◇ sbhambr1 ◇ Siddhant Bhambri ◇ sbhambr1 ◇ sbhambr1

Research Objective: The goal of my research is to advance the field of Human-Aware Artificial Intelligence (HAAI). I aim to understand the interactions of robotic agents in human-AI collaborative settings. My primary research interests lie in the fields of *Reinforcement & Preference-based Learning*, using *LLMs for human-AI teaming*, and *Multi-Agent RL*.

EDUCATION

Ira A. Fulton School of Engineering, Arizona State University *2021 - Present*
PhD Student in Computer Science GPA: 4.0/4.0
Advised by *Dr. Subbarao Kambhampati*

Delhi Technological University, India *2016-2020*
B.Tech in Computer Science CGPA: 8.7/10.0

RESEARCH & PROFESSIONAL EXPERIENCE

Graduate Research Associate: ASU *Presently*
Mentored by *Dr. Subbarao Kambhampati*
Working on Collaborative Human-Aware AI problem settings to formulate robust and seamless interaction between humans and AI agent/robot, particularly focusing on modeling the AI agent for compatibility with human behaviors in teaming scenarios.

Research Intern: Nokia Bell Labs, NJ, USA *Summer 2022*
Data & Devices Group
Developed real-world experiment test-bed for testing the transfer of Reinforcement Learning algorithms to conduct Sim-to-Real simulations on real robots, integrating software and robotic hardware for seamless transfer.

Research Intern: IIIT-Delhi & IIT-Madras, India *2018-2020*
Mentored by *Dr. Arun Balaji Buduru (IIIT) & Dr. Chester Rebeiro (IIT)*
Learned and utilized human preferences for smart-home domain, with specific emphasis on power consumption patterns for IoT devices (2018-2019).
Reviewed black-box adversarial attack techniques on face-recognition and object-tracking systems (2019-20).

PUBLICATIONS & MANUSCRIPTS

Preference Proxies: Evaluating Large Language Models in capturing Human Preferences in Human-AI Tasks
Mudit Verma*, **Siddhant Bhambri***, Subbarao Kambhampati
(ICML) International Conference on Machine Learning 2023 - Workshop on Theory of Mind in Communicating Agents & Workshop on The Many Facets of Preference-based Learning

Exploiting Action Distances for Reward Learning from Human Preferences
Mudit Verma, **Siddhant Bhambri**, Subbarao Kambhampati
(ICML) International Conference on Machine Learning 2023 - Workshop on The Many Facets of Preference-based Learning & Submitted to NeurIPS 2023

Exploiting Unlabeled Data for Feedback Efficient Human Preference-based Reinforcement Learning
Mudit Verma, **Siddhant Bhambri**, Subbarao Kambhampati
AAAI 2023 - Workshop on Representation Learning for Responsible Human-Centric AI *(Pre-print)*

Reinforcement Learning Methods for Wordle: A POMDP/Adaptive Control Approach
Siddhant Bhambri, Amrita Bhattacharjee, Dimitri Bertsekas

IEEE Conference on Games (CoG) 2023 *(Pre-print)*

Using Deception in Markov Game to Understand Adversarial Behaviors through a Capture-The-Flag Environment
Siddhant Bhambri, Purv Chauhan, Frederico Araujo, Adam Doupé, Subbarao Kambhampati
Conference on Decision and Game Theory for Security (GameSec 2022) & AAAI-2023 Workshop on Artificial Intelligence for Cyber Security (AICS) *(Link)*

Contrastively Learning Visual Attention as Affordance Cues from Demonstrations for Robotic Grasping
Yantian Zha, **Siddhant Bhambri**, Lin Guan
IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) 2021 *(Link)*

Multi-objective Reinforcement Learning based approach for User-Centric Power Optimization in Smart Home Environments
Saurabh Gupta, **Siddhant Bhambri**, Karan Dhingra, Arun Balaji Buduru, Ponnurangam Kumaraguru
2020 IEEE World Congress on Services - Smart Data Service (SMDS) *(Link)*

A Survey of Black-Box Adversarial Attacks on Computer Vision Models
Siddhant Bhambri, Sumanyu Muku, Arun Balaji Buduru
arXiv 2019 *(Link)*

Multiple Resource Management and Burst Time Prediction using Deep Reinforcement Learning
Vaibhav Kumar, **Siddhant Bhambri**, Prashant Giridhar Shambharkar
International Journal of Advances in Computer Science and its Applications 2019 *(Link)*

TEACHING & SERVICE

- **Teaching:**
  - Teaching Assistant: CSE 471-Intro to Artificial Intelligence (Fall '21)
  - Teaching Assistant: CSE 574-Planning & Learning in AI (Fall '22)
- **Reviewing:**
  - (ICML) International Conference on Machine Learning 2023 - Workshop on Theory of Mind in Communicating Agents
  - (GameSec) Conference on Decision and Game Theory for Security 2023
  - (ICAPS) International Conference on Automated Planning and Scheduling 2023 - Human Aware and Explainable Planning Workshop
  - (SBP-BRiMS) International Conference on Social Computing, Behavioral-Cultural Modeling & Prediction and Behavior Representation in Modeling and Simulation 2023 (*Sub-reviewer*)
  - (RA-L) IEEE Robotics and Automation Letters 2022
  - (IROS) IEEE International Conference on Intelligent Robots And Systems 2021, 2022
  - (TDSC) IEEE Transactions on Dependable and Secure Computing 2021
- **Other Services:**
  - Technical Program Committee (PC) Member: GameSec - Conference on Decision and Game Theory for Security 2023

TECHNICAL STRENGTHS

- **Programming Languages:** Python, PDDL, C/C++, JAVA.
- **Tools & Technologies:** PyTorch, Sklearn, Pandas, Numpy, Jupyter, ROS, Gazebo.

NOTABLE AWARDS

- **Doctoral Fellowship**: Awarded by the School of Computing, Informatics, and Decision Systems Engineering (CIDSE), Arizona State University
- Rank : **In top 10 percent** in JEE Advance 2016 among **150,000** candidates.
- Secured **99.97** percentile in JEE Main 2016 among 1.2 million students.