Welcome to ECE-590 Socially Cognizant Robotics!

The course integrates the STEM disciplines of robotics (vision, manipulation, navigation, control) with machine learning (visual learning, language processing) to cognitive modeling, behavioral research and public policy. Collaborative teams will be arranged comprising both STEM and social science students.

Primary Instructor: Prof. Kristin Dana

Additional instructors from the following team of professors: Clint Andrews, Kostas Bekris, Jacob Feldman, Jingang Yi, Aaron Mazzeo, Pernille Hemmer, Hal Salzman, Matthew Stone

Assessment:

- *Assignments (25%) (Python, Pytorch, Mujoco, OpenAIGym)*
- *Quizzes (25%)*
- *Class Presentation (25%)*
- *Final Project (20%)*
- *Class Participation (5%)*

Topics:

1. **Introduction to Socially Cognizant Robotics and AI:** Motivation and tour of open issues in the following application domains of Socially Cognizant Robotics: (1) mobility and strength assistance, (2) food preparation, (3) trash and recycling pickup, (4) smart buildings. Brief Introduction to Socially Cognizant Robotics integrating robotics, cognitive science and social science. Considerations of impact of technology on the individual and society, ethics and unintended consequences. What does a roboticist need to know about social science? What does a social scientist need to know about robotics?
2. **Interdisciplinary Synergies of Vision, Language, Robotics:** Recent Developments in core disciplines: embodiment, control, vision, language, cognitive science. Modern computational frameworks and foundation models in AI that are unifying these subfields. Example synergies among the disciplines: Robot embodiment for human trust and safety; Robot embodiment for human/robot cooperation; Motion planning that accounts for humans in the environment; Motion planning, grasp planning, that accounts for the human intent/theory of mind; Robot vison for semantic state representations supports planning, control and evaluations of cog-sci models (human policy or behavior); language for a natural means of human/robot collaboration
3. **Theory of Mind:** Decision-making by human agents; limitations of human decision-making; Bayesian decision-making. Human navigation; social wayfinding. How can cognitive scientists improve robot simulators? How can robotics be incorporated in behavioral studies in real-world scenarios? Human interpretation of intentional behavior, aka "mindreading." The intentional stance; biological motion; interpreting animacy and intentionality from motion; theory of mind. How can human-style intention interpretation be incorporated into robot design?
4. **Robot Vision (aka Computer Vision):** Deep Learning and Convolutional Neural Nets for interpreting the environment to enable interaction. Image Classification and Object Detection, 3D reconstruction, recognition and semantic segmentation. Representation learning for vision and beyond perception. Human-guided, vision-based robot skill acquisition and learn-by-example paradigms such as learning visuomotor control. Off-the-shelf vision algorithms as a tool for social science studies.

5. **Language Processing:** Language Processing: information retrieval dialogue vs. collaboration through language. Robot-human dialog that uses physical affordances to communicate in multimodal ways (e.g., words, gestures, gaze, positioning and posture). Research on how to use contextual knowledge and behavioral models about users' goals during interaction.
6. **Transformers in Foundation Models of AI and Robotics:** Transformers mechanism of self-attention to weigh the importance of different parts of the input data. Relation of transformers to traditional recurrent neural networks (RNNs) or convolutional neural networks (CNNs). Architectural details of the encoder, decoder, attention mechanism. Impact on AI and Robotics.
7. **Reinforcement Learning (RL), From Classic RL to DQN:** Evolution of RL from Classic Markov decision processes, Q-learning and policy gradient methods to more advanced techniques such as Deep Q-Networks (DQN). The impact of this progression for RL applications requiring sophisticated decision-making, including human-robot interaction. Deep Reinforcement Learning for Learning Action Policies
8. **Unifying RL and Transformers with Decision Transformers:** Framing the RL problem as a sequence modeling task. Using the transformer architecture to process trajectories of state-action-reward sequences, enabling the model to predict optimal actions by attending to relevant past experiences (using the transformer's attention mechanism).
9. **Socially Compliant Planning and Control:** Robot Planning and Control: navigation, kinematics, grasping, motion planning. Data-driven control to improve interactions with humans for interactive social robots. Hybrid methods that use data and experience for evaluation. How can cognitive models predict human desires during interactive processes?
10. **Socially Cognizant Visual Navigation:** Navigation as an interdisciplinary case study: Human navigation from a cognitive science perspective, computational algorithms for automated agent navigation (deep reinforcement learning, hierarchical RL), societal issues in self-navigating autonomous agents, human navigation and wayfinding, navigation and spatial maps, social wayfinding.
11. **Social Interaction: Computational Challenges and Benchmarks** A discussion of new datasets and challenges relevant to socially cognizant robotics including "Agent: A benchmark for core psychological reasoning.", "Watch-and-help: A challenge for social perception and human-ai collaboration.", "Ego4D: Around the World in 3,000 Hours of Video", "Habitat 3.0: A co-habitat for humans, avatars and robots."
12. **Societal Impact: Robots in the City** Frameworks for social science research: collective decision making (social, economic, political), innovation processes, social dimensions of technology development and design, unintended consequences of technical change, interrelationships between technologies and people. Social Science research methods for Robotics: Quantifying how robots and society interact, ethnographic observations, sensing technology (e.g., cameras, counters), survey research and experiments.
13. **Inverse RL (IRL) and RL with Human Feedback (RLHF)** IRL methods estimate the reward function given an agent's observed behavior, allowing an RL agent to mimic these behaviors even without explicit reward signals. The relevance of IRL in computational theory of mind models to infer human intent and state of mind. RLHF for incorporates human feedback directly into the learning process Instead of learning purely from environmental rewards. RLHF is key to socially cognizant robotics since agents receive guidance through human preferences or corrections, ensuring that the learned policies align more closely with human values and expectations. Together, these methods have the potential bridge the gap between human intuition and machine learning, leading to more robust and human-aligned AI systems.
14. **Representation Learning for Robotics: Does computer vision matter for robotics?** Representation learning from computer vision to get higher performance in robot sensorimotor

tasks can be achieved. Techniques to explore include contrastive learning, self-supervised learning, semi-supervised learning.

15. **Robot Embodiment**: Grippers, manipulators, soft-robotics, exoskeletons. Soft robots and manipulators improve safety, empathy, competence in everyday tasks and social settings. Challenge and promise of human-robot exoskeletons in the workplace.
16. **Lectures for Student Presentations of Research Papers** Presentations format, content and delivery are emphasized in this class. The rubric for evaluation will be based on the Create and Deliver Technical Presentation Video (see Linkedin Learning via Rutgers U)