

Kai-Chieh Hsu

F-310 Engineering Quadrangle, 41 Olden Street, Princeton, New Jersey, U.S.A. | ☎ (+1) 508-345-3157

✉ kaichieh@princeton.edu | 🏠 kaichiehhsu.github.io/ | 📧 kaichiehhsu | 🌐 kai-chieh-hsu | 📷 Kai-Chieh Hsu

I work on combining game-theoretic counterfactual reasoning and machine learning techniques for safe human-centered robotic systems.

Research Interests

Machine Learning	Safe reinforcement learning (RL), adversarial RL and safe Sim2Real transfer
Human-Robot Interaction	Generative models and inverse RL for strategy and intent inference
Multi-Agent Planning	Game-theoretic counterfactual reasoning and iterative linear quadratic game

Education

Princeton University (PU)

Ph.D. Candidate in Electrical and Computer Engineering (ECE)

M.A. in Electrical and Computer Engineering

Princeton, NJ, USA

Sept. 2021 - June 2024 (EXPECTED)

Sept. 2019 - May 2021

- Concentration: Machine learning and Robotics
- Achieved 4.0/4.0 GPA
- Thesis Advisor: Prof. Jaime Fernández Fisac

National Taiwan University (NTU)

B.S. in Electrical Engineering (EE)

Taipei, Taiwan

Sept. 2014 - Jan. 2019

- Concentration: Signal processing and Digital IC design
- Achieved 4.19/4.30 overall GPA and ranked in **top 5%**
- Research Advisors: Prof. An-Yeu (Andy) Wu and Prof. Jean-Fu Kiang

Work Experiences

Research Scientist Intern [P1]

Remote

NVIDIA Corporation (Manager: **Prof. Marco Pavone**, Mentor: **Prof. Karen Leung**, Yuxiao Chen)

May 2022 - Dec. 2022

- Formalize **responsibility** by safety margin decrease and policy shift with **counterfactual reasoning**
- Estimate the responsibility level online with **hidden Markov model**
- Incorporate the estimated responsibility into the **trajectory prediction** models

Research Projects

Adversarial Safety Game [P2]

PU, NJ, USA

Safe Robotics Laboratory, Prof. Jaime Fernández Fisac, Duy Phuong Nguyen

Feb. 2022 - PRESENT

- Robustify the **reachability-based RL** by jointly training an adversarial agent under the self-play spirit
- Apply a novel shielding scheme to combine both the **model-based counterfactual rollout** and model-free safety value function

Inverse Specification

PU, NJ, USA

Safe Robotics Laboratory, Prof. Jaime Fernández Fisac

July 2020 - PRESENT

- Use Bayesian optimization to infer constraints **interactively with humans** by asking for ranking feedback
- **Select queries actively** to speed up the convergence to the true preference based on information-theoretic metrics
- Survey in **inverse optimal control** and **imitation learning**

Safe Sim2Real Transfer (Sim-to-Lab-to-Real) [J1]

PU, NJ, USA

Safe Robotics Laboratory, Prof. Jaime Fernández Fisac, Duy Phuong Nguyen

Intelligent Robot Motion Lab, Prof. Anirudha Majumdar, Allen Z. Ren

May 2021 - Jan. 2022

- Use **Reachability-Based RL** and a **supervisory control** scheme to allow the least-restrictive safe exploration
- Combine with **PAC-Bayes control** to provide a tight performance lower bound to unseen environments

Reach-Avoid Reinforcement Learning [C2] [J1]

PU, NJ, USA

Safe Robotics Laboratory, Prof. Jaime Fernández Fisac

July 2020 - Mar. 2021

Hybrid Systems Laboratory, Prof. Claire J. Tomlin, Vicenç Rubies-Royo

- Derive a time-discounted formulation of the reach-avoid optimal control problem that lends itself to **(deep) RL**
- Deploy our reach-avoid Q-Learning in a range of nonlinear systems, including an **attack-defense game**
- Reach-avoid reinforcement learning allows learning from near defeat and fits in **safe reinforcement learning**

ECG Real-Time Telemonitoring [J2] [C3]

NTU, Taiwan

Access IC Lab, Prof. An-Yeu (Andy) Wu

Aug. 2017 - Mar. 2019

- **Edge Classification:** Incorporate **compressed sensing**, task-driven dictionary learning (predictive sparse coding) and PCA to render light-weighted classifier and overcome limited labeled data challenge
- **On-Demand Recovery:** Design a two-stage algorithm that classifies and reconstructs only problematic signals. This algorithm utilizes the information from classification stage to speed up the reconstruction algorithm
- **Hardware Design and Chip Implementation:** Propose a hardware architecture for on-demand recovery to allow hardware sharing between classification and reconstruction algorithms

Publications

Preprint

- [P1] **K.-C. Hsu**, Karen Leung, Yuxiao Chen, Jaime F. Fisac, Marco Pavone, Interpretable Trajectory Prediction for Autonomous Vehicles via Counterfactual Responsibility, submitted to *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Feb 2023.
- [P2] **K.-C. Hsu***, Duy P. Nguyen*, and Jaime F. Fisac, *ISAACS: Iterative Soft Adversarial Actor-Critic for Safety*, submitted to *Learning for Dynamics and Control (L4DC)*, Nov 2022.

Journal Papers

- [J1] **K.-C. Hsu***, Allen Z. Ren*, Duy P. Nguyen, Anirudha Majumdar⁺, and Jaime F. Fisac⁺, *Sim-to-Lab-to-Real: Safe Reinforcement Learning with Shielding and Generalization Guarantees*, in *Artificial Intelligence Journal*, Oct 2022. | Spotlight in *ICLR Workshop* and *NeurIPS Workshop*
- [J2] C.-Y. Chou, **K.-C. Hsu**, B.-H. Cho, K.-C. Chen and A.-Y. (Andy) Wu, *Low-Complexity On-demand Reconstruction for Compressively Sensed Problematic Signals*, in *IEEE Trans. Signal Process.*, vol. 68, pp. 4094-4107, July 2020.
- [J3] **K.-C. Hsu** and J.-F. Kiang, *Joint Estimation of DOA and Frequency From A Mixture of Frequency Known and Unknown Sources with Orthogonal Coprime Arrays*, in *Sensors*, 19(2), 335, Jan. 2019.

Conference Papers

- [C1] H. Chen, **K.-C. Hsu**, W. Turner, P.-H. Wei, K. Zhu, D. Pan and H. Ren, Reinforcement Learning Guided Detailed Routing for FinFET Custom Circuits, in *International Symposium on Physical Design (ISPD)*, Held Virtually, Mar 2023.
- [C2] **K.-C. Hsu***, V. Rubies-Royo*, C. J. Tomlin and J. F. Fisac, *Safety and Liveness Guarantees through Reach-Avoid Reinforcement Learning*, in *Proceedings of Robotics: Science and Systems (RSS)*, Held Virtually, July 2021.
- [C3] **K.-C. Hsu***, B.-H. Cho*, C.-Y. Chou and A.-Y. (Andy) Wu, *Low-Complexity Compressed Analysis in Eigenspace with Limited Labeled Data for Real-Time Electrocardiography Telemonitoring*, in *IEEE GlobalSIP*, Anaheim, USA, Nov 2018.

Honors & Awards

SEAS Travel Grant

SEAS, PU, NJ, USA

Nov. 2022

Teaching Assistant Award

Dept. of ECE, PU, NJ, USA

- For the new *Intelligent Robotic Systems* course


Sept. 2022

3rd Prize in Integrated Circuit Design Contest

Ministry of Education, Taiwan

- Out of about 300 teams

July 2018

- 2nd Prize** in Taiwan Creative Electromagnetic Implementation Competition
 • Under the supervision of Prof. Tzong-Lin Wu |  High-speed RF and mm-Wave Tech. Center, Taiwan
 Aug. 2017
- 8th place** in Data Structure and Programming Contest
 • Out of about 250 students
 Cadence, Taiwan
 Mar. 2017
- Graduate Representative** in NTUEE graduate ceremony
 • Given to top ten students of four years
 Dept. of EE, NTU, Taiwan
 June 2018
- Professor Chun-Hsiung Chen Scholarship**
 • Rewarded outstanding performances in electromagnetic fields
 Electromagnetic Industry-Academia Consortium, Taiwan
 Jan. 2018
- Presidential Awards**
 • Given to top ten students of that semester
 Dept. of EE, NTU, Taiwan
 second semester of 2014 and 2016

Research & Teaching Experiences

Teaching Assistant

ECE346/566: Intelligent Robotic Systems, Prof. Jaime Fernández Fisac
 ELE364: Machine Learning for Predictive Data Analytics, Prof. Niraj Jha

PU, NJ, USA
 Jan. 2022 - May 2022
 Sept. 2020 - Dec. 2020

Research Assistant

Access IC Lab, Prof. An-Yeu (Andy) Wu
 Group of Electromagnetic Applications, Prof. Jean-Fu Kiang

NTU, Taiwan
 Feb. 2018 - Mar. 2019
 Feb. 2017 - Mar. 2019

Teaching Assistant

Digital System Design

NTU, Taiwan
 Feb. 2018 - June 2018

Professional Activities

Reviewer

Artificial Intelligence, IEEE Open Journal of Control Systems, IEEE Trans. on Vehicular Technology, IETE Technical Review, IEEE Trans. on Signal Processing, ICRA, L4DC

Program Committee

NeurIPS Workshop on **Human in the Loop Learning** and **Trustworthy Embodied AI**

Skills

- **Program Languages** Python, MATLAB, Verilog, C++
- **Others** PyTorch, Jax, Git, SLURM, NumPyro, CVX, \LaTeX