

Zizhao Wang

Contact zizhao.wang@utexas.edu, 734-747-4206 **Website** <https://wangzizhao.github.io/>

Google scholar <https://tinyurl.com/zizhaowangscholar>

Research focus

LLM Post-Training, World Models, Reinforcement Learning (RL), Causal Reasoning

Skills

- large language model (LLM), generative AI (genAI), RL post-training (PPO, GRPO), genAI agents, reasoning, safety, tool use
- decision making: model-based RL, imitation learning, planning
- development: Python, machine learning frameworks (PyTorch, TensorFlow, Transformers), distributed training (deepspeed), efficient training (PEFT, LoRA), deployment (vLLM), simulation (Mujoco)
- deep learning: representation learning, generalization

Education

2020 - 25 **PhD**, Electrical and Computer Engineering, GPA: 4.00/4.00
Expected graduation: 2025/12

University of Texas at Austin

2018 - 19 **MS**, Computer Science, GPA: 4.00/4.00

Columbia University

2016 - 18 **BS**, Computer Engineering, GPA: 3.96/4.00

University of Michigan

2014 - 18 **BS**, Electrical and Computer Engineering, GPA: 3.72/4.00

Shanghai Jiao Tong University

Work Experience

2025/03 **Student Researcher**

Google

- Designed an adversarial reinforcement learning post-training framework to enhance LLM tool-use agent security against prompt injections (**LLM Agents, GenAI, RL post-training, Safety**).
- Built the data collection pipeline with vLLM and parallel simulation environments, speeding up LLM agent rollout collection by 8x (**LLM inference, vLLM, parallel environments**).
- Trained the LLM model with the GRPO algorithm in a fast and memory-efficient way (**distributed training, Transformers, deepspeed, LoRA, python, pyTorch**), reducing the attack success rate by 21% and improving task success rate by 18% compared to the untrained model.

2024/06 **Research Intern**

Microsoft Research

- Designed a **generative world model** to synthesize experience of novel scenarios, with object-centric representations and disentangled representations (**world model, genAI, representation learning**).
- Implemented a novel model architecture, training and evaluation pipeline and sped up training with distributed training (**transformer models, state-space models, distributed data parallel, pyTorch**).
- Enhanced the generalization of RL policies by 30%, when trained with generated out-of-distribution data (**model-based RL, python, pyTorch**).

2024/01 **Research Intern**

Honda Research Institute

- Developed a motion prediction algorithm that reduced prediction error by 48%, by applying causal reasoning to vehicle interactions (**world model, autonomous driving, causality**).
- Sped up model training with distributed training and efficient CUDA implementations for sparse attention (**transformer models, distributed data parallel, CUDA, python, pyTorch**).

Research Experience

2021 - 25 **World Model (ICML *oral*, NeurIPS, AAAI *oral*)**

University of Texas at Austin

- Led a team of 3 phd students to build a world model that analyzes causal relationships between state factors, increasing the generalization performance by 46% on long-horizon robot manipulation tasks in simulation (**causality, motion planning, robotics, simulation**).
- Led a team of 4 students to develop a novel intrinsic reward algorithm based on world models, increasing long-horizon robot manipulation task by 3x in simulation (**model-based RL, robotics**).
- Led a team of 4 students to scale latent action world models to multi-agent scenarios, increasing prediction performance by 34% (**genAI, imitation learning, robotics**).

2021 - 22 **Robot Navigation (ICRA, IROS)**

University of Texas at Austin

- Developed a novel framework to dynamically adjust motion planners using interventions and evaluative feedback from humans, reducing navigation time by 45% (**robotics, motion planning, human in the loop, navigation**).
- Led a team of 5 students to develop a novel learning algorithm to generate cheap demonstration data for navigation, enabling robots to navigate in challenging environments where classical motion planners fail (**genAI, imitation learning, robotics**).

Selected Publications

See google scholar (<https://tinyurl.com/zizhaowangscholar>) for a complete list of publications (Machine Learning: **NeurIPS, ICML, AAAI**; Robotics: **CoRL, ICRA, IROS**).

- Adversarial Reinforcement Learning for LLM Agent Safety, *In submission*
Z Wang, D Li, V Keshava, P Wallis, A Balashankar, P Stone, L Rutishauser.
- Dyn-O: Building Structured World Models with Object-Centric Representations, *NeurIPS 2025*
Z Wang, K Wang, L Zhao, P Stone, J Bian.
- SkiLD: Unsupervised Skill Discovery Guided by Local Dependencies, *NeurIPS 2024*
Z Wang*, J Hu*, C Chuck*, S Chen, R Martín-Martín, A Zhang, S Niekum, P Stone.
- Building Minimal and Reusable Causal State Abstractions for Reinforcement Learning, *AAAI 2024 (*oral*)*
Z Wang*, C Wang, X Xiao, Y Zhu, and P Stone.
- ELDEN: Exploration via Local Dependencies, *NeurIPS 2023*
Z Wang*, J Hu*, R Martín-Martín, and P Stone.
- Causal Dynamics Learning for Task-Independent State Abstraction (**Oral**), *ICML 2022 (*oral*)*
Z Wang, X Xiao, Z Xu, Y Zhu, and P Stone.
- Learning to Correct Mistakes: Backjumping in Long-horizon Task and Motion Planning, *CoRL 2022*
Y Sung*, **Z Wang***, and P Stone.
- From Agile Ground to Aerial Navigation: Learning from Learned Hallucination, *IROS 2021*
Z Wang, X Xiao, A Nettekoven, K Umasankar, A Singh, S Bommakanti, U Topcu, and P Stone.
- APPLE: Adaptive Planner Parameter Learning from Evaluative Feedback, *RAL 2021*
Z Wang, X Xiao, G Warnell, and P Stone.
- Maximizing BCI Human Feedback using Active Learning, *IROS 2020*
Z Wang*, J Shi*, I Akinola*, and P Allen.
- Accelerated Robot Learning via Human Brain Signals, *ICRA 2020*.
I Akinola*, **Z Wang***, J Shi, X He, P Lapborisuth, J Xu, D Watkins-Valls, P Sajda, and P Allen.