# Hao Bai

MS in Computer Science, UIUC

⋈ haob2@illinois.edu

Interests: Reinforcement Learning, Representation Learning

# **Education**

MS in Computer Science	UIUC, USA	Aug 2023 - May 2025
BS in Computer Engineering (Dual)	UIUC, USA	Aug 2019 - May 2023
BE in Computer Engineering (Dual)	Zhejiang University, China	Sep 2019 - Jul 2023
BE in Computer Engineering (Dual)	Zhejiang University, China	Sep 2019 - Jul 202

# Professional Experience

# **UC Berkeley**

Visiting Scholar, Advisor: Sergey Levine, Yi Ma

- Reinforcement learning algorithms and environments for visual language agents.
- Mathematically principled language transformer architectures with better neuron-level interpretability.

### **Microsoft Research**

Research Intern, Advisor: Shilin He • Language-model-based large-scale outage interpretation and prediction.

# **Selected Papers**

#### DigiRL: Training Real-World GUI Agents with Scaled Autonomous RL [PDF] NeurIPS'24

Hao Bai\*, Yifei Zhou\*, Mert Cemri, Jiayi Pan, Alane Suhr, Sergey Levine, Aviral Kumar UC Berkelev Proposed DigiRL, the first algorithmic framework for task-solving Android agents using RL, and introduced a hierarchical reinforcement learning algorithm that achieves 300% better performance than current state-of-the-art methods. Results show that the introduction of reinforcement learning significantly improves sample efficiency.

# RL4VLM: Fine-Tuning Large VLMs as Decision-Making Agents via RL [PDF]

- Y. Zhai, H. Bai, J. Pan, S. Tong, Y. Zhou, A. Suhr, S. Xie, Y. LeCun, Y. Ma, S. Levine
- Proposed an algorithmic framework to fine-tune VLMs with RL, which provides a task description and then prompts it to generate chain-of-thought (CoT) reasoning to enable the VLM to efficiently explore intermediate reasoning steps that lead to the final text-based action. I proposed and implemented format-oriented auto-regressive fine-tuning for better policy initialization, and managed most scaling-up and speed optimization.

White-Box Transformers via Sparse Rate Reduction: Compression Is All There Is? [PDF]

Y. Yu, S. Buchanan, D. Pai, T. Chu, Z. Wu, S. Tong, H. Bai, Y. Zhai, B. Haeffele, Y. Ma UC Berkeley • As part of the research, I designed and pre-trained two mathematically principled language transformers, CRATE-BERT and CRATE-GPT, and empirically show that the architecture is scalable to the GPT-2 level with a comparable performance when halving the parameter size.

### Progressive Responses with Real-Time Internet Search for Conversations [PDF] Revanth Reddy, Sharath Suresh, Hao Bai, ..., Chengxiang Zhai

• As a participant of the Alexa SocialBot challenge, I implemented the progressive response generation to blend search results into the bot's responses while ensuring low response latency, which cuts down user waiting time by 50%.

# Social Conversational Commonsense-Guided Search Query Generation [PDF] Revanth Reddy, Hao Bai, Wentao Yao, Sharath Suresh, Heng Ji, Chengxiang Zhai

• I was in charge of most of the implementation in this work. We proposed to integrate commonsense knowledge to the query generator by generating initial responses from a commonsense response generator and followed by distilling knowledge from LLM. Our model outperforms zero-shot T5 on the quality of the generated query and also final response.

Dec 2023 - Present Berkeley, CA

Nov 2022 - May 2023 Beijing, CN

> NeurIPS'24 UC Berkeley

### WSDM'24

EMNLP'23

UIUC

JMLR

UIUC