Yuan Xia

yuanxia@usc.edu | (949) 374-6628 | 900 S Figueroa St, Los Angeles, CA 90015

https://www.linkedin.com/in/yuan-xia-0227/

Education

University of Southern California **Los Angeles, CA**

Ph.D., Computer Science, GPA: 3.95/4.0 *08/2020-12/2025*

University of California, Irvine **Irvine, CA**

B.S., Computer Science with Honors degree, GPA: 3.91/4.0 *09/2017- 12/2019(with earlier graduation)*

Cousework: The Science of Large Language Models, Deep Learning, Natural Language processing, Computer Networks, Probabilistic and Deterministic Graphical Models, Computer Vision, Algorithms, Machine Learning, Deep Learning, Python, Java, C++.

Research Projects

Learning-based property mining on Distributed Systems with LLM-aided

- Developed a data-driven, template-free invariant generation tool while monitoring distributed systems online or offline, adaptable to any programming language.
- Designed and integrated a framework with Large Language Models (**LLMs**) to enhance automated invariant **synthesis**, outperforming traditional LLMs and dynamic invariant generators.

LLM Reasoning

- Built a novel framework incorporating multiple techniques that significantly boost **LLM reasoning** capabilities.
- Achieved an **80%** increase in accuracy over the baseline model, Llama 3.1.

Natural Language to Temporal Logic Property Translation

- Created an innovative synthetic data generator addressing data scarcity in the Temporal Logic (TL) domain, producing a significant synthetic dataset for Natural Language to TL (**NL-TL**) **translation**.
- Trained a state-of-the-art LLM for NL-TL translation, capable of converting context-insensitive language to context-sensitive language effectively.

Probabilistic Guarantees for Autonomous Systems using Gaussian Process Model with Conformal Inference

- Implemented conformal inference with Gaussian Process, Neural Network, and Perceptron models to ensure probabilistic guarantees for cyber-physical systems.
- Validated these methods in real-world scenarios including **Mountain Car, Lane Keep Assist via Reinforcement Learning, F-16 Flight Control Systems, and Artificial Pancreas models.**

Work Experience

Machine Learning Engineer & Researcher **May 2024 - Present**

Bell Labs, NOKIA *Muray Hill, NJ*

- Lead initiatives in Large Language Models (**LLMs**), such as Llama 3 & Llama 3.1, and transformer **reinforcement learning,** driving forward innovations in machine learning.
- Innovated and implemented **multiple groundbreaking techniques**, establishing a new framework that markedly improved the reasoning accuracy of LLMs.
- Filed a **patent** and prepared a scholarly **paper** for submission, reflecting pioneering contributions to the field of LLMs.

Software Engineering Research Assistant **May 2019 - Present**

CPS-VIDA LAB, USC & UCI *Los Angeles & Irvine, CA*

- Advanced the development of property translation and reasoning techniques using Large Language Models (**LLMs**), which further aids the monitoring system for **autonomous driving.**
- Research in mining logical **properties** and performing statistical verifications for **distributed systems,** enhancing system reliability and performance.
- **Event detection** with **LLMs** and **motion prediction** with **perception** data within autonomous driving systems.

Publications

The Reasoning of Large Language Models, ICLR 2025 review.

Discovering Likely Invariants and Anomalies in Distributed Systems through Runtime Monitoring and Learning, VMCAI 2025 review.

Data-Driven Template-Free Invariant Generation, ICSE 2024 Review**.**

Systematic Translation from Natural Language Robot Task Descriptions to STL, AISOLA 2024.

Statistical Verification and Risk Estimation of Cyber-Physical Systems using Surrogate Models and Conformal Inference, TCPS 2023.

Statistical Verification of Autonomous Systems using Surrogate Models and Conformal Inference, ICCPS, 2022.

A Comprehensive Study of Autonomous Vehicle Bugs, ICSE, 2020.