

Hongji Pu

Grainger College of Engineering
University of Illinois Urbana-Champaign
DCL, 201 N. Goodwin Avenue, MC-258
Champaign, IL, USA

[+1 217-255-3602] 
[hongjip2@illinois.edu] 
[\[Personal Webpage\]](#) 

RESEARCH FOCUS

My research focuses on **language agents**, **continual learning and self-improvement**, and **world models for evaluation and decision-making**. I study how language models and agents use memory and external evidence, learn from research findings and feedback, and acquire new capabilities while retaining what they have learned. I develop methods to verify proposed updates and determine when past experience remains useful as a system changes. I also investigate how world models guide rubric optimization for automated grading and algorithm design on complex networks. Across these directions, I ask when improvements in a model translate into better decisions.

EDUCATION

University of Illinois, Urbana-Champaign <i>M.S. in Financial Engineering</i> <i>GPA: 3.54/4.00</i>	Aug. 2025 - Dec. 2026(Expected) <i>Champaign, IL, USA</i>
Central University of Finance and Economics <i>B.S. in Finance</i> <i>GPA: 3.84/4.00 (Class Rank 1/39)</i>	Aug. 2021 - Jun. 2025 <i>Beijing, CN</i>

PUBLICATIONS

Manuscripts and Preprints

* denotes equal contribution

- | | |
|------|---|
| 2026 | COUNTERMEM: World-Model-Verified Counterfactual Memory for Language Agents
Hongji Pu , Ruixiang Tong, Yongfeng Zhang
Paper Under review at ICLR 2027 |
| 2026 | Network World Models as Environments for Algorithm Design on Complex Networks
Hongji Pu* , Rishab Alagharu*, Zeeshan Memon, Xinyuan Song, Yuntong Hu, Liang Zhao
Paper Under review at ICLR 2027 |
| 2026 | PAPER2LLM++: Continual Self-Evolution of Language Models From Research Papers
Hongji Pu , Yilun Zhao, Wenpeng Yin
Paper Under review at TACL 2026.10 |
| 2026 | Launder-Bench: Can Repository-Level Agentic Code Laundering Be Detected?
Yuntong Hu, Hongji Pu , Zimo Li, Ruomeng Li, Xiaotong Liu, Ziyuan Qin, Qiyue Li, Liang Zhao
Paper Code Under review at ICLR 2027 |
| 2026 | RubricWorld: World Model-Guided Rubric Optimization for LLM Graders
Hongji Pu , Hang Li, Yucheng Chu, Jiliang Tang
Paper Under review at ARR 2026.10 |
| 2026 | SkillOps: Managing LLM Agent Skill Libraries as Self-Maintaining Software Ecosystems
Hongji Pu* , Xinyuan Song*, Liang Zhao
Paper Code Under review at ICLR 2027 |

- 2026 **BeliefRAG: Making Adaptive RAG State-Aware Under Evolving Evidence**
 Hongji Pu
[Paper](#) [Code](#) Under review at TACL
- 2026 **Beyond Forecasting: Recasting Volatility Control as a Routing Problem**
 Hongji Pu, Leyang Zhou
[Paper](#) [Code](#) Under review at ACM ICAIF 2026
- 2026 **RouteProfile: Graph-Based Profiling for cold-start LLM Routing**
 Hongji Pu*, Jingjun Xu*, Tao Feng*, Haozhen Zhang, Jiaxuan You, Ge Liu
[Paper](#) [Code](#) Preprint

RESEARCH EXPERIENCE

CounterMem: World-Model-Verified Counterfactual Memory for Language Agents
 Rutgers University (Remote)

Apr. 2026 – Present
 Prof. Yongfeng Zhang, NJ

- ◆ Developed **CounterMem**, an RL-based framework for constructing and reusing verified counterfactual memory in language agents. The method generates local alternatives to failed actions, verifies their outcomes with executable world models, stores action contrasts with evidence and reuse conditions, and trains a DQN selector to decide when retrieved memory should guide a new task.
- ◆ Evaluated on **12 benchmarks across six domains**. CounterMem improved ReAct and Reflexion in **24/24 benchmark-agent comparisons** with gpt-oss-120b, averaging **+12.6 pp**; across two backbones, four-domain gains reached **8.2–20.4 pp** while task-run token usage decreased by **7.7–42.0%**.

RubricWorld: World Model- Guided Rubric Optimization for LLM Graders.
 Michigan State University (Remote)

Mar. 2026 – Present
 Prof. Jiliang Tang, MI

- ◆ Independently Developed **RubricWorld**, a world model-guided framework for refining LLM grading rubrics without retraining the grader. The framework learns state-dependent edit effects from paired grading decisions, predicts corrections and regressions across score boundaries, and certifies updates using held-out evidence and finite-sample confidence bounds.
- ◆ Evaluated on ASAP 2.0, EIR, and ASAP-SAS across three LLM backbones, achieving the best performance in **20 of 27 metric–dataset–backbone settings**. Improved gpt-5-mini accuracy on ASAP 2.0 from **18.20% to 46.70%**; removing the learned edit-effect model reduced average accuracy from **40.18% to 27.31%**.

Paper2LLM++: Continual Self-Evolution of LLMs From Research Papers
 Pennsylvania State University (Remote)

Mar. 2026 – Present
 Prof. Wenpeng Yin, PA

- ◆ Independently Developed **PAPER2LLM++**, a framework for continual LLM self-evolution from research papers. The framework extracts evidence-grounded findings, verifies model-specific limitations, and converts applicable findings into synthetic supervision, using a try–evaluate–commit procedure to integrate LoRA updates while limiting forgetting and general-capability degradation.
- ◆ Evaluated on **30 research papers spanning seven failure families across two LLM backbones**, achieving positive held-out gains on **24 of 30 papers**. Mixed-paper streams improved targeted performance by **7.1–7.5 points** while retaining **83.7–85.0% of acquired gains**, outperforming five baselines in retention on both backbones.

WORK EXPERIENCE (Additional Experiences Detailed in [Personal Webpage](#))

JPMorgan Chase & Co. — Volatility-Targeted Strategies Design
 Quantitative Research Intern

Jan. 2026 – Mar. 2026
 Abhishek Namballa, New York, NY

- ◆ Co-developed an end-to-end volatility-targeting research pipeline for S&P 500 futures, covering data preprocessing, volatility forecasting, exposure allocation, and backtesting. Compared statistical and machine-learning estimators, including EWMA, GJR-GARCH, HAR-RV, and Random Forest, alongside regime-aware controllers and ensemble strategies.
- ◆ Evaluated estimator–controller combinations using forecasting accuracy, risk-adjusted returns, turnover, and drawdown. In the project backtests, HAR-RV with regime scaling achieved a Sharpe ratio of 0.602 versus 0.575 with naive scaling, while reducing maximum drawdown from 18.71% to 17.58% and increasing the Calmar ratio from 0.355 to 0.364.

SKILLS

- ◆ **Programming:** Python, SQL, Bash, MATLAB, Lean
- ◆ **Tools & Platforms:** PyTorch, Hugging Face Transformers, verl, PyMARL, EPyMARL, MemOS, AgentDojo, NLIP, Qdrant, Git
- ◆ **Soft Skills:** Strong self-motivation, independent research, interdisciplinary collaboration, analytical thinking, and effective multitasking.