

Xueru Wen

wenxueru2022@iscas.ac.cn · Google Scholar · GitHub

Education

University of Chinese Academy of Sciences, Beijing

2023 – Present

Ph.D. Student in Computer Application Technology; expected Jun. 2028; advisor: Prof. Ben He
Chinese Information Processing Laboratory, Institute of Software, CAS

Jilin University, Changchun

2019 – 2023

B.S. Tang Aoqing Honors Program in Science, Computer Science Track

Selected Research

CoVRL: Verifier-free RL [1]

First author, ICML 2026

Proposed variational RL coupling prior and posterior sampling to improve exploration and thought–answer consistency. Qwen2.5-7B average accuracy on math and general reasoning rose from 37.8% to 50.2% (+12.4 percentage points), exceeding the strongest compared baseline by 2.3 points.

Reward Model Evaluation [4]

First author, ICLR 2025 Spotlight

Showed that similar reward-model accuracy can yield substantially different downstream RL outcomes. Analyzed evaluation design and reward overoptimization through the Goodhart effect, exposing the limits of accuracy-based model selection.

Cheems: Chinese Reward Models [5]

First author, ACL 2025

Built a Chinese reward-model training and evaluation framework; released CheemsPreference and CheemsBench, evaluated open-source models on Chinese data, and trained Chinese reward models.

RLFH: Self-alignment [6]

First author, ACL 2025

Proposed on-policy self-alignment using atomic facts and external knowledge to construct dense token-level rewards without human annotations; reduced hallucination and improved accuracy on benchmarks including HotpotQA.

Experience

RedNote (Ace Top Intern since Jul. 2025)

Sep. 2023 – Present

Reward Model Iteration: Iterated RLHF reward models with preference feedback, hallucination detection, and process rewards; evaluated downstream effects and explored scalable oversight mechanisms.

Math Verification Agent: Built an agent with Lean, Python, and natural-language verifiers, then strengthened it with SFT and agentic RL to improve ProofBench performance.

SWE Coding Agent: Built Docker/Kubernetes sampling pipelines for large-scale trajectory collection and validation; used SFT to evaluate data effectiveness. Implemented TITO, R3, and partial rollout for agentic RL, improving SWE-bench Verified, Multilingual, and Pro performance.

Automated Posttrain: Built large-scale data synthesis pipelines on cloud GPU clusters, used world models for targeted synthesis, and implemented segment-wise tokenization for training and inference to improve PostTrainBench performance.

Chinese Information Processing Laboratory, ISCAS

Mar. 2023 – Aug. 2023

Large Language Model Pre-training

- Added LLaMA training, checkpoint conversion, and numerical alignment to Megatron–DeepSpeed
- Integrated optimizations such as fused kernels and Flash Attention, improving training throughput

Honors and Awards

- CAST Young Talent Support Program, Doctoral Student Special Track
- UCAS President’s Scholarship, Excellence Award (CAS system)

2025

Aug. 2026

Publications

- [1] **Coupled Variational Reinforcement Learning for Language Model General Reasoning**
Xueru Wen, Jie Lou, Yanjiang Liu, Hongyu Lin, Ben He, Xianpei Han, Le Sun, Yaojie Lu, Debing Zhang. ICML 2026
- [2] **Decoupling Reasoning and Confidence: Resurrecting Calibration in Reinforcement Learning from Verifiable Rewards**
Zhengzhao Ma, Xueru Wen, Boxi Cao, Yaojie Lu, Hongyu Lin, Jinglin Yang, Min He, Xianpei Han, Le Sun. ICML 2026
- [3] **Learning from Failures: Correction-Oriented Policy Optimization with Verifiable Rewards**
Mengjie Ren, Jie Lou, Boxi Cao, Xueru Wen, Hongyu Lin, Xianpei Han, Le Sun, Xing Yu, Yaojie Lu. COLM 2026
- [4] **Rethinking Reward Model Evaluation: Are We Barking up the Wrong Tree?**
Xueru Wen, Jie Lou, Yaojie Lu, Hongyu Lin, Xing Yu, Xinyu Lu, Ben He, Xianpei Han, Debing Zhang, Le Sun. ICLR 2025 Spotlight
- [5] **Cheems: A Practical Guidance for Building and Evaluating Chinese Reward Models from Scratch**
Xueru Wen, Jie Lou, Zichao Li, Yaojie Lu, Xing Yu, Yuqiu Ji, Guohai Xu, Hongyu Lin, Ben He, Xianpei Han, Le Sun, Debing Zhang. ACL 2025
- [6] **On-Policy Self-Alignment with Fine-grained Knowledge Feedback for Hallucination Mitigation**
Xueru Wen, Jie Lou, Xinyu Lu, Yuqiu Ji, Xinyan Guan, Yaojie Lu, Hongyu Lin, Ben He, Xianpei Han, Debing Zhang, Le Sun. ACL 2025
- [7] **Transferable Post-training via Inverse Value Learning**
Xinyu Lu, Xueru Wen, Yaojie Lu, Bowen Yu, Hongyu Lin, Haiyang Yu, Le Sun, Xianpei Han, Yongbin Li. NAACL 2025
- [8] **AutoAlign: Get Your LLM Aligned with Minimal Annotations**
Xinyu Lu, Dong Xu, Chunkang Zhang, Xinyan Guan, Junxiang Wang, Qingyu Zhang, Pengbo Wang, Yingzhi Mao, Hao Xiang, Xueru Wen, Zichao Li, Yaojie Lu, Hongyu Lin, Le Sun, Xianpei Han. ACL 2025
- [9] **Critic-CoT: Boosting the Reasoning Abilities of Large Language Model via Chain-of-Thought Critic**
Xin Zheng, Jie Lou, Boxi Cao, Xueru Wen, Yuqiu Ji, Hongyu Lin, Yaojie Lu, Xianpei Han, Debing Zhang, Le Sun. ACL 2025
- [10] **The Devil Is in the Details: Tackling Unimodal Spurious Correlations for Generalizable Multimodal Reward Models**
Zichao Li, Xueru Wen, Jie Lou, Yuqiu Ji, Yaojie Lu, Xianpei Han, Debing Zhang, Le Sun. ICML 2025
- [11] **Offline Pseudo Relevance Feedback for Efficient and Effective Single-Pass Dense Retrieval**
Xueru Wen, Xiaoyang Chen, Xuanang Chen, Ben He, Le Sun. SIGIR 2023
- [12] **End-to-End Entity Detection with Proposer and Regressor**
Xueru Wen, Changjiang Zhou, Haotian Tang, Luguang Liang, Hong Qi, Yu Jiang. Neural Processing Letters 2023
- [13] **Type-supervised Sequence Labeling Based on the Heterogeneous Star Graph for Named Entity Recognition**
Xueru Wen, Changjiang Zhou, Haotian Tang, Luguang Liang, Yu Jiang, Hong Qi. arXiv 2022

Open-Source Tools

- **Skynet**: A native macOS agent workspace with local/SSH sessions, queued follow-ups, and code review.
- **Python Hot Patch Skill**: Live CPython hot-patching with preserved state, safe-point execution, and verification.
- **Wait Skill**: Event-driven agent waiting and coordination, with monitoring, timed loops, and dependency workflows.
- **JSONLine Viewer**: A VS Code extension for formatted JSONL/NDJSON previews and indexed navigation of large files.