



JIWON JEON (Ph.D. Student)

Large-Language-Models & Reinforcement Learning

🌐 Homepage ✉ email [in](#) Linked-In
☎ +82-10-7206-8566

ABOUT ME

I received my B.S. and M.S. degrees in Electrical Engineering from KAIST in 2020 and 2022. I am currently a Ph.D. student working with Prof. Youngchul Sung. My research focuses on LLM post-training and multi-agent reinforcement learning, and my recent interest is in LLM harnessing.

EDUCATION

Ph.D. Candidate (Electrical Engineering) KAIST (Daejeon, Korea) Advisor: Prof. Youngchul Sung	2022 - ongoing
M.S. (Electrical Engineering) KAIST (Daejeon, Korea) Advisor: Prof. Youngchul Sung	2020 - 2022
B.S. (Electrical Engineering) KAIST (Daejeon, Korea)	2015 - 2020
Exchange Student (Electrical Engineering) Technical University of Denmark (Lyngby, Denmark)	08.2018-02.2019

PUBLICATIONS (LLMS)

Rebellious Student: Reversing Teacher Signals for Reasoning Exploration with Self-Distilled RLVR

Jeonghye Kim*, **Jiwon Jeon***, Dongsheng Li, Yuqing Yang (* Equal contribution)
NeurIPS 2026 (Poster), **COLM 2026 Workshop** (Oral Presentation) [paper]

Inject a Turn, Not a Token: Learning Where to Redirect LLM Reasoning for Self-Correction

Jiwon Jeon*, Myungsik Cho*, Jeonghye Kim, Youngchul Sung (*Equal contribution)
ICLR 2027 (Submitted), **COLM 2026 Workshop** (Spotlight) [paper]

Why Does Self-Distillation (Sometimes) Degrade the Reasoning Capability of LLMs?

Jeonghye Kim, Xufang Luo, Minbeom Kim, Sangmook Lee, Dohyung Kim, **Jiwon Jeon**, Dongsheng Li, Yuqing Yang
COLM 2026 (poster) [paper]

Be My Tutor: On-Policy Co-Distillation for Mutual LLM Improvement via Peer Feedback

Woohyeon Byeon, **Jiwon Jeon**, Jeonghye Kim, Youngchul Sung
NAACL 2027 (submitted) [paper]

PUBLICATIONS (RL)

MASTARS: Multi-Agent Sequential Trajectory Augmentation with Return-Conditioned Subgoals

Jiwon Jeon, Myungsik Cho, Woojun Kim, Seongmin Kim, Woohyeon Byeon, Seungyul Han, Youngchul Sung
NeurIPS 2026 (poster)

AGiR: Mitigating Gift Over-Reliance in Mixed-Motive Games

Woohyeon Byeon, Seongmin Kim, **Jiwon Jeon**, Woojun Kim, Youngchul Sung
NeurIPS 2026 (poster)

Addressing Exogeneous Variability in Cooperative Multi-Agent Reinforcement Learning

Seongmin Kim, Woohyeon Byeon, **Jiwon Jeon**, Seungyul Han, Youngchul Sung
NeurIPS 2026 (Spotlight)

STAIRS-Former: Spatio-Temporal Attention with Interleaved Recursive Structure Transformer for Offline Multi-task Multi-agent Reinforcement Learning

Jiwon Jeon*, Myungsik Cho*, Youngchul Sung (* Equal contribution)
ICLR 2026 (poster) [paper]

Generalized Per-Agent Advantage Estimation for Multi-Agent Policy Optimization

Seongmin Kim, Giseung Park, Woojun Kim, **Jiwon Jeon**, Seungyul Han, Youngchul Sung
AAMAS 2026 (Best Paper Award Nominee) [paper]

MASER: Multi-Agent Reinforcement Learning with Subgoals Generated from Experience Replay Buffer

Jiwon Jeon, Woojun Kim, Whiyoung Jung, Youngchul Sung
ICML 2022 (poster) [paper]

WORK EXPERIENCE

Research Intern LG AI Research, (Seoul, Korea)	07.2023-03.2024
Undergraduate Researcher NICA Lab in KAIST, (Daejeon, Korea)	02.2019 - 12.2019
Student Intern Future Design System, (Daejeon, Korea)	12.2017-02.2018

TEACHING EXPERIENCE

Mathematical Foundations of Reinforcement Learning (EE619, KAIST) Role: Teaching Assistant	2025 Fall
Introduction to Electronics Design Lab (EE305, KAIST) Role: Teaching Assistant	2025 Spring
Introduction to Electronics Design Lab (EE305, KAIST) Role: Teaching Assistant	2024 Fall
Mathematical Foundations of Reinforcement Learning (EE619, KAIST) Role: Teaching Assistant	2023 Spring
Probability and Introductory Random Processes (EE210, KAIST) Role: Teaching Assistant	2022 Fall
Mathematical Foundations of Reinforcement Learning (EE619, KAIST) Role: Teaching Assistant	2022 Spring
Special Lecture on Leadership Entrepreneurship (CC010, KAIST) Role: Teaching Assistant	2022 Spring

ACADEMIC SERVICE

- Silver Reviewer, 2026 ICML