Seungyong Moon

Seoul, South Korea symoon11@mllab.snu.ac.kr Google scholar Homepage GitHub

Research Interests

My research aims to develop autonomous agents with strong robustness, generalization, and reasoning capabilities. Currently, I am working on enhancing the reasoning capabilities of large language models by leveraging synthetic data generation and designing novel reinforcement learning algorithms tailored for long-horizon reasoning tasks.

Education

Seoul National University Feb 2027
MS/PhD in Computer Science
- Advisor: Hyun Oh Song

Seoul National University Feb 2019
BS in Mathematical Science, BA in Economics, Minor in Computer Science
- Honors: Summa Cum Laude (GPA: 4.10/4.30, Rank: 1/34)

Work Experience

Research Intern Sep 2024–Jan 2025
Qualcomm AI Research Amsterdam, Netherlands
- Developed supervised fine-tuning and reinforcement learning algorithms to improve the algorithmic reasoning capabilities of large language models for code execution.

Research Intern June 2023–Sep 2023
KRAFTON Seoul, South Korea
- Designed a self-reflection algorithm to improve the spatial reasoning capabilities of proprietary large language models in interactive game environments.

Research Intern June 2022–Sep 2022
DeepMetrics Seoul, South Korea
- Built a data preprocessing pipeline and structured SQL database for ventilator waveform data and developed imitation and reinforcement learning algorithms for autonomous ventilator control.

Research Intern July 2018–Aug 2018
NAVER Corp. Seongnam-si, South Korea
- Developed a data augmentation algorithm using sequence-to-sequence models for paraphrase identification.

Publications

[C7] **Seungyong Moon**, Bumsoo Park, Hyun Oh Song
Learning to Better Search with Language Models via Guided Reinforced Self-Training
Neural Information Processing Systems (NeurIPS), 2025

[C6] **Seungyong Moon**, Junyoung Yeom, Bumsoo Park, Hyun Oh Song
Discovering Hierarchical Achievements in Reinforcement Learning via Contrastive Learning
Neural Information Processing Systems (NeurIPS), 2023

[C5] **Seungyong Moon**, JunYeong Lee, Hyun Oh Song
Rethinking Value Function Learning for Generalization in Reinforcement Learning
Neural Information Processing Systems (NeurIPS), 2022

[C4] Deokjae Lee, **Seungyong Moon**, Junhyeok Lee, Hyun Oh Song
Query-Efficient and Scalable Black-Box Adversarial Attacks on Discrete Sequential Data via Bayesian Optimization
International Conference on Machine Learning (ICML), 2022

[C3] **Seungyong Moon***, Gaon An*, Hyun Oh Song
Preemptive Image Robustification for Protecting Users against Man-in-the-Middle Adversarial Attacks
AAAI Conference on Artificial Intelligence (AAAI), 2022

[C2] Gaon An*, **Seungyong Moon***, Jang-Hyun Kim, Hyun Oh Song
Uncertainty-Based Offline Reinforcement Learning with Diversified Q-Ensemble
Neural Information Processing Systems (NeurIPS), 2021

[C1] **Seungyong Moon***, Gaon An*, Hyun Oh Song
Parsimonious Black-Box Adversarial Attacks via Efficient Combinatorial Optimization
International Conference on Machine Learning (ICML), 2019 (long talk)

Awards and Scholarships

NeurIPS Top Reviewers 2022, 2025

NeurIPS Scholar Award 2023

NAVER PhD Fellowship Award 2022

Yulchon AI Star Scholarship 2022

Qualcomm Innovation Fellowship Finalists 2020, 2022

KFAS Computer Science Graduate Student Scholarship 2019–2024

The National Scholarship for Science and Engineering 2015–2016

Gwanak Association Scholarship 2012

Teaching Experience

Teaching Assistant Fall 2020, Fall 2022
Machine Learning

Teaching Assistant Spring 2019
Introduction to Deep Learning

Undergraduate Student Instructor Fall 2017
Basic Calculus 2

Undergraduate Student Instructor Spring 2017
Basic Calculus 1

Academic Services

Conference Reviewer

- NeurIPS (2021–2025), ICML (2022–2025), AAAI (2022–2026), ICLR (2024–2026), AISTATS (2025–2026)

Journal Reviewer

- Neurocomputing (2021), Machine Learning (2023), Transactions on Intelligent Vehicles (2023)

Skills

Programming Languages

- Advanced: Python, PyTorch, JAX, LaTeX
- Intermediate: C++, MATLAB, TensorFlow

Languages

- Korean (native)
- English (fluent)