

# Cover Letter

My name is Changhun Lee, a Postdoctoral Associate in the Graduate School of Artificial Intelligence at UNIST (Ulsan National Institute of Science and Technology). I recently received my Ph.D. degree with the thesis titled *“Tackling Three Problems in Controlled Sequence Generations: Bridging Reinforcement Learning with Language Models.”* Based on my achievement and deep interest in the interplay area between natural language generations and reinforcement learning (RL), I would love to expand the theoretical horizons of machine learning and contribute to the advancement of our knowledge from natural to social sciences.

**During the last two years of my Ph.D. course, I worked on developing controllable language models (CLMs).** For example, I regarded language models (LMs) as agents and fine-tuned them using policy-based RL algorithms (e.g., REINFORCE, PPO, etc.) to control LMs’ behavior as intended. **More recently, my research orientation has pivoted towards establishing the theoretical foundations of controllable artificial intelligence,** primarily in the context of language, i.e., sequential, models. With the increasing reliance on artificial intelligence (AI) in various fields, I believe that unprincipled, or rule-of-thumb, use of AI will greatly increase the risk of out-of-control situations, and improving theoretical understanding of the controllable AI will help counter the potential risks associated with unprincipled practices. In this regard, **I am interested in research that analyzes the nature of controllability and uses it to improve the controllability of AI models.**

Specifically, **I have three academic keywords that are of personal interest as post-doc research topics.** 1) The first is **“Pareto optimality.”** To build controllable AI, a model must be optimized for the likelihood and control objectives simultaneously (jointly). Given this multi-objective nature, studying controllability from the perspective of Pareto optimality can provide hints for the theoretical foundations of controllable AI. 2) The second is **“maximum mutual information.”** In terms of Pareto optimality, the more the Pareto improvement, the higher the controllability. Since mutual information represents the joint probability of two random variables, mapping both likelihood and control objectives onto the probability space together and maximizing their mutual information will derive the Pareto improvement. 3) The last keyword is **“approximation methods.”** Computing mutual information is expensive, so it is necessary to approximate the maximum mutual information without estimating the exact amount of mutual information. Therefore, variational inference or perturbation techniques are the research topic in itself.

As a post-doc researcher, I hope to further delve into **identifying theoretical foundations** to guarantee the controllability of AI models, **building controllable AIs** that are reliable from theoretical perspectives, and **applying them to solve real-world applications** (e.g., drug discovery). I am confident that my research experience would bring value to your team. If you find my profile aligning with your research interest, I would greatly appreciate being informed of an opportunity to further discuss my potential contributions to your team.

I dream of becoming a researcher who is willing to stand up against prejudice and inequality in the world with logical thinking and a scientific attitude. I am looking forward to a heart-pounding journey of research that will create a better world and promote the future with great colleagues. I will end this article by introducing a quote from one of my favorite scholars, Alfred Marshall:

*“Cool head, but warm heart!”*

For a more detailed overview of my work and accomplishments, I invite you to visit my personal website at <https://chlee-leo.github.io/>.