Statement of Purpose

For application for summer research internship in CoRe Lab.
April 15^{th}, 2023

Yaowen Ye
elwin@connect.hku.hk
University of Hong Kong

Recent years have witnessed the widespread use of Artificial Intelligence (AI) systems in various fields. Large Language Models (LLMs) like ChatGPT have profoundly amazed the general public with many tasks that even humans cannot perform well [3]. However, such amazement does not cover the fact that we are still far from an AI system that can learn, reason, and generalize as humans do. The quest for human-like machine intelligence still goes on, and that is the very quest where my passion for research lies, as well as the motivation for my application to your lab.

I started self-studying machine learning in senior high school, creating a project about generating music with RNNs by myself. Later, I worked on another project about building smart glasses for the visually impaired with CNNs for perception and vibration motors plus earphones for human-machine interaction. Although these projects were application-oriented, my passion for machine intelligence remained steadfast. Upon joining the University of Hong Kong as a computer science student, I immediately applied to become a student research assistant. In December 2021, I joined the Data Intelligence Lab@HKU under the supervision of Dr. Chao Huang. Although the lab's focus may not be on learning, reasoning, or generalization, I remained committed to the work due to my attraction to graph machine learning and self-supervised learning. Moreover, I deemed it as an opportune jumping-off point for my research career. In the lab, I worked on enhancing sequential recommender systems with self-supervised learning consistently for over a year, reproducing and studying state-of-the-art (SOTA) works, and assisting senior students in conducting research [1]. I came up with a novel idea of graph transition path masking, which semantically aligns with the sequential pattern of recommendation datasets. Following this, I implemented my idea and conducted comprehensive experiments on it over a six-month period and published my findings in SIGIR 2023 [5]. After the long journey, I finally realized that I may now be ready to pursue my true passion in the field of machine intelligence.

I spent weeks going through papers related to reasoning, with a particular focus on those published in NeurIPS, trying to find a suitable and elegant problem to start with. Subsequently, I encountered your work, *On the Learning Mechanisms in Physical Reasoning* [2]. It has such a clear logic flow and comprehensive experiments with in-depth analysis. Moreover, when I browsed CoRe Lab's website after accessing the project link, I was thrilled to find that you also study topics related to abstract reasoning, and most astonishingly, the entire lab focuses on reasoning, which is exactly what I was looking for! Furthermore, I appreciate your emphasis on integrating cognitive science and AI in your work. Maybe neuroscience is worth adding here since you recommended students watch the Brains, Minds, and Machines (BMM) Summer Course. Coincidently, I also went through its recordings in my first year of university. Upon recovering from the excitement of discovering your lab, I delved further into your published works, with a particular emphasis on those exploring abstract reasoning, especially by your research scientist, Chi Zhang. He contributed so much in this direction with his consistent line of work, and I was impressed by how he did this step by step, from the RAVEN dataset [6] to various models attempting this challenging task [7, 8, 10]. Overall, I appreciate the reasoning-focused and cognitive science-inspired mission of CoRe Lab, and I am convinced that CoRe Lab is the ideal next chapter in my academic journey.

My ultimate goal is to achieve human-like machine intelligence by designing AI systems capable of learning, reasoning, and generalizing like humans. Presently, I plan to concentrate my research efforts on reasoning, which is well aligned with CoRe Lab's research direction. My (incomplete) literature review brought me the following ideas for further exploration:

- How can one design benchmarks for evaluating a model's reasoning ability comprehen-

sively? For example, in the context of physical reasoning, how to develop benchmarks that are sufficiently complex in the sense of dynamics while elegant and focused enough to facilitate the study of reasoning without excessive irrelevant noise, such as those come with some real-world datasets.

- Similar to the sentiment neurons discovered in LLMs [4], is there an analogous phenomenon in abstract reasoning tasks? Is there a specific part of the neural network structure associated with a particular reasoning process (e.g., basic logical operations, syllogism, etc.) akin to the human cortex?
- Would pre-training on abstract reasoning tasks boost models' reasoning ability in real-world tasks and enable more robust generalization against distribution shifts? Many works have shown that deep neural networks tend to learn by memorization instead of reasoning [9], resulting in unsatisfying generalizability. From my intuition, pre-training on datasets that enforce the model to reason might be able to mitigate this issue, as the learned reasoning procedure may be possibly transferable to other tasks.

The ideas presented above are merely in their nascent stage, but investigating them will represent my first attempt towards achieving my research goal.

I am a highly self-motivated student with a strong passion for machine intelligence and reasoning, but I am also an inexperienced researcher who eagerly needs guidance from a professional. Joining CoRe Lab would provide me with the opportunity to work with top-notch scientists and gain valuable knowledge, advice, as well as supervision. My long-term goal is to pursue a Ph.D. in this field, eventually become a professor, and establish my own lab, if possible. I believe that CoRe Lab is the ideal place for me to start this journey. If given the opportunity, I am confident that I would make significant strides toward achieving my goals. Thank you once again for considering my application for the summer internship!

References

[1] Li, C., Huang, C., Xia, L., Ren, X., Ye, Y., and Xu, Y. Masked graph transformer for recommendation.

[2] Li, S., Wu, K., Zhang, C., and Zhu, Y. On the learning mechanisms in physical reasoning. *arXiv preprint arXiv:2210.02075* (2022).

[3] OpenAI(2023). Gpt-4 technical report. *arXiv preprint arXiv:2303.08774* (2023).

[4] Radford, A., Jozefowicz, R., and Sutskever, I. Learning to generate reviews and discovering sentiment. *arXiv preprint arXiv:1704.01444* (2017).

[5] Ye, Y., Huang, C., and Xia, L. Graph masked autoencoder for sequential recommendation.

[6] Zhang, C., Gao, F., Jia, B., Zhu, Y., and Zhu, S.-C. Raven: A dataset for relational and analogical visual reasoning. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (2019), pp. 5317–5327.

[7] Zhang, C., Jia, B., Gao, F., Zhu, Y., Lu, H., and Zhu, S.-C. Learning perceptual inference by contrasting. *Advances in neural information processing systems 32* (2019).

[8] Zhang, C., Jia, B., Zhu, S.-C., and Zhu, Y. Abstract spatial-temporal reasoning via probabilistic abduction and execution. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (2021), pp. 9736–9746.

[9] Zhang, C., Raghu, M., Kleinberg, J., and Bengio, S. Pointer value retrieval: A new benchmark for understanding the limits of neural network generalization.

[10] Zhang, C., Xie, S., Jia, B., Wu, Y. N., Zhu, S.-C., and Zhu, Y. Learning algebraic representation for systematic generalization in abstract reasoning. In *Computer Vision–ECCV 2022: 17th European Conference, Tel Aviv, Israel, October 23–27, 2022, Proceedings, Part XXXIX* (2022), Springer, pp. 692–709.