Research Statement

Ying Fan

My research advances the frontier of artificial intelligence by developing principled algorithms that solve complex problems through sequential decision making. This approach has led to significant breakthroughs across multiple domains, including reinforcement learning (RL) [1, 2], outlier detection with sequential decision making [3], RL and RLHF for diffusion models [4, 5, 6], etc. My work spans several fundamental areas with a unified framework for AI problem-solving: systematically breaking down complex challenges into sequential decisions.

- **Reinforcement Learning Algorithms:** In our ICML 2021 oral paper [1], we developed a novel model-based RL framework using posterior sampling for efficient exploration with improved theoretical guarantee; In our NeurIPS 2024 paper [2], we develop the first offline RL algorithm utilizing context and goal datasets to solve goal-oriented problems.
- **Outlier Detection with sequential decision making:** Our ICML 2022 oral paper [3] revolutionized outlier detection training by recasting it as a sequential decision process. This novel perspective led to performance improvement while significantly reducing the sample complexity, inspiring a rich body of follow-up work.
- **RLHF for Diffusion Models:** Through our ICML 2023 [4] and NeurIPS 2023 [5] papers, we broke new ground by developing foundational methods for reinforcement learning-based adaptation of diffusion models and pioneered in RLHF for diffusion models, which has been extensively cited and widely adopted, greatly advancing generative AI research.
- **Looped Language Models for Reasoning:** Our recent work on looped Transformers [7] demonstrates architectures inspired by sequential reasoning can substantially improve length generalization for arithmetic tasks without complex and tailor-made positional embeddings, which sparked significant interest from the research community.
- **Improving Reasoning Abilities of LLMs with RL:** Our ongoing work focuses on improving the reasoning ability of LLMs with process supervision and reinforcement learning.

Research Vision: Towards General Problem-Solving AI. My research vision centers on developing AI systems capable of solving complex, real-world problems through decomposition and sequential reasoning. This vision encompasses several interconnected directions:

1. **Improved Reasoning: Compositional and Sequential Learning.** How can AI learn to decompose complex tasks into simpler, reusable components or sequential steps? Developing such algorithms can advance more efficient learning and better generalization.
2. **Alignment for Foundation Models: Model Adaptation and Control.** As foundation models become increasingly powerful, I am interested in efficiently adapting them to specific tasks and aligning them with human preferences, like reinforcement learning from human feedback (**RLHF**) and AI feedback (**RLAIF**).
3. **Applications:** There are various domains where complex decision-making is crucial, including **scientific discovery**, **robotics and automation**, **AI safety and alignment**, etc.

I aim to advance both the theoretical foundations and practical capabilities of AI systems, contributing to the development of more capable, reliable, and beneficial AI systems that can effectively tackle real-world challenges.

References

[1] Y. Fan and Y. Ming, "Model-based reinforcement learning for continuous control with posterior sampling," in *International Conference on Machine Learning*, PMLR, 2021.

[2] Y. Fan, J. Li, A. Swaminathan, A. Modi, and C.-A. Cheng, "How to solve contextual goal-oriented problems with offline datasets?," *Advances in Neural Information Processing Systems*, 2024.

[3] Y. Ming, Y. Fan, and Y. Li, "Poem: Out-of-distribution detection with posterior sampling," in *International Conference on Machine Learning*, pp. 15650–15665, PMLR, 2022.

[4] Y. Fan and K. Lee, "Optimizing ddpm sampling with shortcut fine-tuning," in *International Conference on Machine Learning*, PMLR, 2023.

[5] Y. Fan, O. Watkins, Y. Du, H. Liu, M. Ryu, C. Boutilier, P. Abbeel, M. Ghavamzadeh, K. Lee, and K. Lee, "Dpok: Reinforcement learning for fine-tuning text-to-image diffusion models," in *Thirty-seventh Conference on Neural Information Processing Systems*, 2023.

[6] K. D. Dvijotham, S. Omidshafiei, K. Lee, K. M. Collins, D. Ramachandran, A. Weller, M. Ghavamzadeh, M. Nasr, Y. Fan, and J. Z. Liu, "Algorithms for optimal adaptation ofdiffusion models to reward functions," in *ICML Workshop on New Frontiers in Learning, Control, and Dynamical Systems*, 2023.

[7] Y. Fan, Y. Du, K. Ramchandran, and K. Lee, "Looped transformers for length generalization," in *The 4th Workshop on Mathematical Reasoning and AI at NeurIPS'24*, 2024.