Research Statement

Xinrun Wang
School of Computing and Information Systems, Singapore Management University
Tel: (65) 91659032; Email: xrwang@smu.edu.sg

December 15, 2024

1 Background

My research is primarily on the decision making, which is pervasive in the modern society. Specifically, my research can be divided into three categories:

- **Fundamental Research on Decision Making**. Reinforcement learning (RL) plays the core role in complex decision making, including both single-agent, e.g., Atari, and multi-agent problems, e.g., StarCraft II. My research in this category focus on improving the scalability and generalizability of the RL methods, which is important for the large-scale real-world deployment. My ultimate goal is to build the decision foundation model which can generalize across any decision making tasks.
- **Applications of Decision Making Research.**. FinTech, especially quantitative trading, is an ideal testbed for the RL methods. In this category, we focus on developing high-performance trading algorithms for portfolio management and high frequency trading and building an open-sourced platform. Our goal is building the generalist RL agent for FinTech which can be deploy to the real trading settings.
- **Foundation Agent**. Foundation agent is an emerging research area. Leveraging the (multimodal) language model, foundation agent can potentially tackles the decision making problems with extremely long-term horizon and complex semantic instructions. My goal in this category is to build the foundation agent which can efficiently explore in the new complex environments.

To summarize, the goal of my research is **advancing the understanding of decision making methods, regarding the scalability, genealizability and solution quality**.

2 Research Areas

In this section, I will introduce a representative work and the future directions in the three categories.

2.1 Fundamental Research on Decision Making

Configurable Mirror Descent [ICML 2024]. Decision-making problems, categorized as single-agent, e.g., Atari, cooperative multi-agent, e.g., Hanabi, competitive multi-agent, e.g., Hold'em poker, and mixed cooperative and competitive, e.g., football, are ubiquitous in the real world. Various methods are proposed to address the specific decision-making problems. Despite the successes in specific categories, these methods typically evolve independently and cannot generalize to other categories. Therefore, a fundamental question for decision-making is: *Can we develop **a single algorithm** to tackle **ALL** categories of decision-making problems?* There are several main challenges: i) different decision-making categories

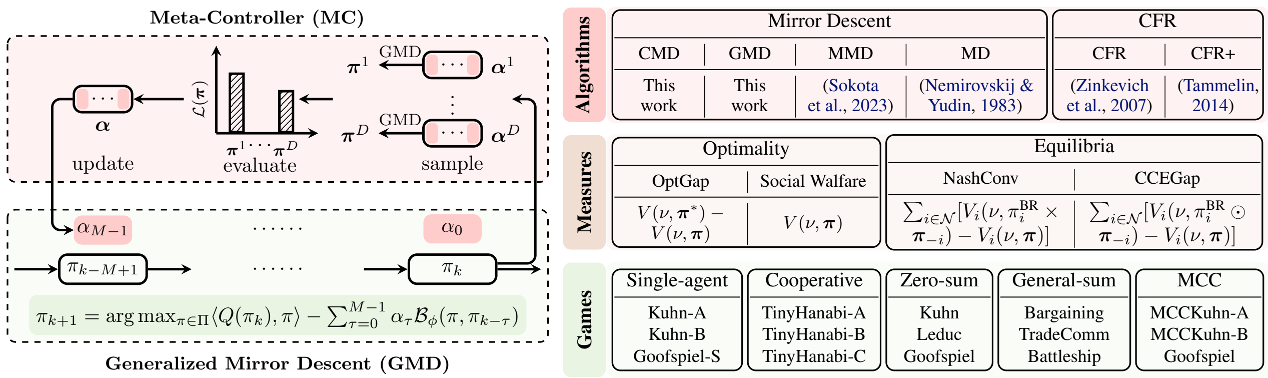

Figure 1: Configurable Mirror Descent

involve different numbers of agents and different relationships between agents, ii) different categories have different solution concepts and evaluation measures, and iii) there lacks a comprehensive benchmark covering all the categories. This work presents a preliminary attempt to address the question with three main contributions. i) We propose the generalized mirror descent (GMD), a generalization of MD variants, which considers multiple historical policies and works with a broader class of Bregman divergences. ii) We propose the configurable mirror descent (CMD) where a meta-controller is introduced to dynamically adjust the hyper-parameters in GMD conditional on the evaluation measures. iii) We construct the GAMEBENCH with 15 academic-friendly games across different decision-making categories. Extensive experiments demonstrate that CMD achieves empirically competitive or better outcomes compared to baselines while providing the capability of exploring diverse dimensions of decision making.

Future Directions. We will continue to explore the unification of decision-making processes. Building on the CMD framework, our focus will be on the following directions:

- **Parameterizing and Optimizing Bregman Divergence**: Developing methods to effectively parameterize and optimize Bregman divergence for improved decision-making performance.
- **Efficient Learning of the Meta-Controller**: Designing techniques to efficiently learn the meta-controller, enabling adaptive and robust decision-making across various contexts.
- **Addressing Complex Decision-Making with Neural Network-Based Policies**: Investigating the use of neural network-based policies to handle complex decision-making problems.

Our goal is to develop a plug-and-play framework for complex decision-making, which is critical for real-world deployment with scalability, generalizability, and efficiency.

2.2 Applications of Decision Making Research

MacroHFT [KDD 2024]. High-frequency trading (HFT) that executes algorithmic trading in short time scales, has recently occupied the majority of cryptocurrency market. Besides traditional quantitative trading methods, reinforcement learning (RL) has become another appealing approach for HFT due to its terrific ability of handling high-dimensional financial data and solving sophisticated sequential decision-making problems, *e.g.,* hierarchical reinforcement learning (HRL) has shown its promising performance on second-level HFT by training a router to select only one sub-agent from the agent pool to execute the current transaction. However, existing RL methods for HFT still have some defects: 1) standard RL-based trading agents suffer from the overfitting issue, preventing them from making effective policy adjustments based on financial context; 2) due to the rapid changes in market conditions, investment decisions made by an individual agent are usually one-sided and highly biased, which might lead to significant loss in extreme markets. To tackle these problems, we propose a novel Memory Augmented Context-aware Reinforcement learning method On HFT, *a.k.a.* MacroHFT, which consists of two training phases: 1) we first train multiple types of sub-agents with the market data decomposed according to various financial indicators, specifically market trend and volatility, where each agent owns a conditional adapter to adjust its trading policy according to market conditions; 2) then we train a hyper-agent to mix the decisions from these sub-agents and output a consistently profitable meta-policy to handle rapid market fluctuations, equipped with a memory mechanism to enhance the capability of decision-making. Extensive experiments on various cryptocurrency markets demonstrate that MacroHFT can achieve state-of-the-art performance on minute-level trading tasks.

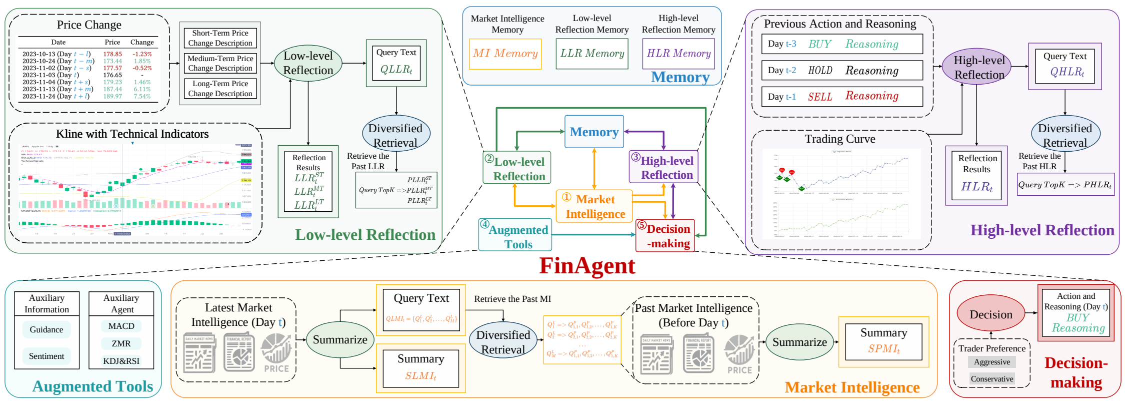

Figure 2: Overview of FinAgent

FinAgent [KDD 2024]. Financial trading is a crucial component of the markets, informed by a multimodal information landscape encompassing news, prices, and Kline charts, and encompasses diverse tasks such as quantitative trading and high-frequency trading with various assets. While advanced AI techniques like deep learning and reinforcement learning are extensively utilized in finance, their application in financial trading tasks often faces challenges due to inadequate handling of multimodal data and limited generalizability across various tasks. To address these challenges, we present FinAgent (as showed in Figure 2), a multimodal foundational agent with tool augmentation for financial trading. FinAgent's market intelligence module processes a diverse range of data-numerical, textual, and visual-to accurately analyze the financial market. Its unique dual-level reflection module not only enables rapid adaptation to market dynamics but also incorporates a diversified memory retrieval system, enhancing the agent's ability to learn from historical data and improve decision-making processes. The agent's emphasis on reasoning for actions fosters trust in its financial decisions. Moreover, FinAgent integrates established trading strategies and expert insights, ensuring that its trading approaches are both data-driven and rooted in sound financial principles. With comprehensive experiments on 6 financial datasets, including stocks and Crypto, FinAgent significantly outperforms 9 state-of-the-art baselines in terms of 6 financial metrics with over 36% average improvement on profit. Specifically, a 92.27% return (a 84.39% relative improvement) is achieved on one dataset. Notably, FinAgent is the first advanced multimodal foundation agent designed for financial trading tasks.

Future Directions. We will continue to explore the AI techniques for FinTech. There are several topics to explore:

- **Developing Mini-FinAgent.**: We will develop the minimal FinAgent for fast development, test and deployment by non-AI people. We will seek the collaboration with companies to deploy the FinAgent into the real world.

- **Developing FinAgent2**: We will develop FinAgent2 which can handle the transactions with different time-scales, e.g., seconds, minutes and days, the different types of transactions, e.g., stocks, cryptocurrencies and securities, with the capabilities to extract the data from Internet.

- **Extending the Platform to Financial Management**: Instead of trading, we will extend the work to financial management such as accounting and auditing for individuals, companies and governments. We will build the comprehensive platform for AI to empower the financial services.

2.3 Foundation Agent

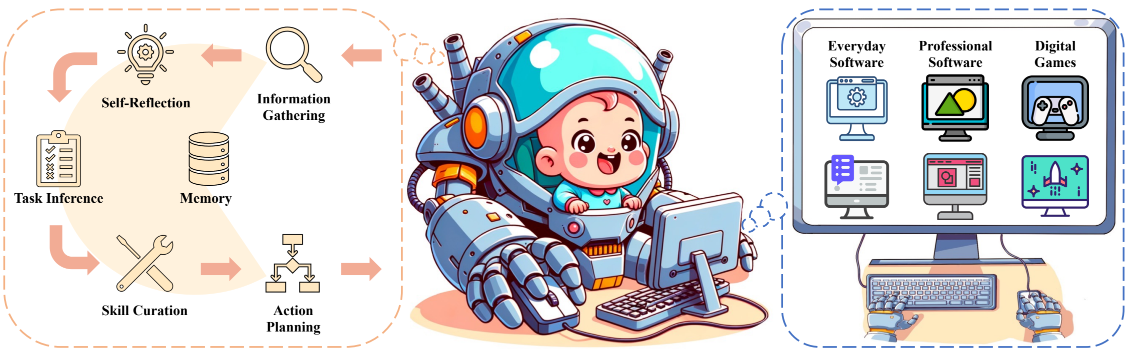

Figure 3: Cradle

Cradle [ICLR 2024, LLM Agent Workshop]. Despite the success in specific tasks and scenarios, existing foundation agents, empowered by large models (LMs) and advanced tools, still cannot generalize to different scenarios, mainly due to dramatic differences in the observations and actions across scenarios. In this work, we propose the General Computer Control (GCC) setting: building foundation agents that can master any computer task by taking only screen images (and possibly audio) of the computer as input, and producing keyboard and mouse operations as output, similar to human-computer interaction. The main challenges of achieving GCC are: 1) the multimodal observations for decision-making, 2) the requirements of accurate control of keyboard and mouse, 3) the need for long-term memory and reasoning, and 4) the abilities of efficient exploration and self-improvement. To target GCC, we introduce Cradle, an agent framework with six main modules, including: 1) information gathering to extract multimodality information, 2) self-reflection to rethink past experiences, 3) task inference to choose the best next task, 4) skill curation for generating and updating relevant skills for given tasks, 5) action planning to generate specific operations for keyboard and mouse control, and 6) memory for storage and retrieval of past experiences and known skills. To demonstrate the capabilities of generalization of Cradle, we deploy it in the complex AAA game Red Dead Redemption II, serving as a preliminary attempt towards GCC. To our best knowledge, our work is the first to enable LMM-based agents to follow the main storyline and finish real missions in complex AAA games, with minimal reliance on prior knowledge.

Future Directions. The future directions we will explore are:

- **Developing Agents Capable of Efficiently Exploring New and Complex Games**: This involves designing algorithms that enhance the agent's ability to navigate, learn, and adapt to new and intricate scenarios, thereby improving their overall performance and decision-making capabilities.

- **Constructing Multi-agent Systems Comprising Foundational Agents that can Both Collaborate and Compete with Each Other**: This includes designing interaction protocols, communication strategies, and coordination mechanisms to enable seamless cooperation and competition, ultimately leading to more robust and versatile AI systems.

- **Developing Distributed Foundation Agents**: Our goal is to design distributed foundation agents capable of integrating and leveraging both closed-source and open-source models to enhance the overall capability and flexibility of our AI systems, i.e., cloud platform over foundation models

Our ultimate goal is to build Artificial General Intelligence (AGI) — creating intelligent systems that can understand, learn, and perform any intellectual task in a human-competitive or super-human level.

3 Selected Publications and Outputs

[1]. Wentao Zhang, Lingxuan Zhao, Haochong Xia, Shuo Sun, Jiaze Sun, Molei Qin, Xinyi Li, Yuqing Zhao, Yilei Zhao, Xinyu Cai, Longtao Zheng, **Xinrun Wang**, Bo An. A multimodal foundation agent for financial trading: Tool-augmented, diversified, and generalist. Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data (**KDD**), 2024.

[2]. Pengdeng Li, Shuxin Li, Chang Yang, **Xinrun Wang**†, Shuyue Hu, Xiao Huang, Hau Chan, Bo An. Configurable mirror descent: Towards a unification of decision making. *Proceedings of the 41st International Conference on Machine Learning* (**ICML**), 2024.

[3]. **Xinrun Wang***, Chang Yang*, Shuxin Li, Pengdeng Li, Xiao Huang, Hau Chan and Bo An. Reinforcement Nash equilibrium solver. *Proceedings of the 33rd International Joint Conference on Artificial Intelligence* (**IJCAI**), 2024.

[4]. Pengdeng Li, Shuxin Li, Chang Yang, **Xinrun Wang**†, Xiao Huang, Hau Chan, Bo An. Self-adaptive PSRO: Towards an automatic population-based game solver. *Proceedings of the 33rd International Joint Conference on Artificial Intelligence* (**IJCAI**), 2024.

[5]. Longtao Zheng, Rundong Wang, **Xinrun Wang**†, Bo An†. Synapse: Trajectory-as-exemplar prompting with memory for computer control. *Proceedings of the 2024 International Conference on Learning Representations* (**ICLR**), 2024.

[6]. Weihao Tan, Wentao Zhang, Shanqi Liu, Longtao Zheng, **Xinrun Wang**†, Bo An†. True knowledge comes from practice: Aligning large language models with embodied environments via reinforcement learning. *Proceedings of the 2024 International Conference on Learning Representations* (**ICLR**), 2024.

[7]. Pengdeng Li*, Shuxin Li*, **Xinrun Wang**†, Jakub Cerny, Youzhi Zhang, Stephen Marcus McAleer, Hau Chan, Bo An. Grasper: A generalist pursuer for pursuit-evasion problems. *Proceeding of the 23rd International Conference on Autonomous Agents and Multi-Agent Systems* (**AAMAS**), 2024.

[8]. Molei Qin, Shuo Sun, Wentao Zhang, Haochong Xia, **Xinrun Wang**†, Bo An†. EarnHFT: Efficient hierarchical reinforcement learning for high frequency trading. *Proceedings of the 38th AAAI Conference on Artificial Intelligence* (**AAAI**), 2024.

[9]. Pengdeng Li, Runsheng Yu, **Xinrun Wang**†, Bo An. Transition-informed reinforcement learning for large-scale Stackelberg mean-field games. *Proceedings of the 38th AAAI Conference on Artificial Intelligence* (**AAAI**), 2024.

[10]. Shuo Sun*, Molei Qin*, Wentao Zhang, Haochong Xia, Chuqiao Zong, Jie Ying, Yonggang Xie, Lingxuan Zhao, **Xinrun Wang**†, Bo An†. TradeMaster: A holistic quantitative trading platform empowered by reinforcement learning. *Thirty-seventh Conference on Neural Information Processing Systems* (**NeurIPS**) Dataset and Benchmark Track, 2023.

[11]. Pengdeng Li, **Xinrun Wang**[†], Shuxin Li, Hau Chan, Bo An. Population-size-Aware Policy Optimization for Mean-Field Games. *The 11th International Conference on Learning Representations* (**ICLR**), 2023.

[12]. Shuxin Li, **Xinrun Wang**[†], Youzhi Zhang[†], Wanqi Xue, Jakub Cerny, Bo An. Solving large-scale pursuit-evasion games using pre-trained strategies. *Proceedings of the 37th AAAI Conference on Artificial Intelligence* (**AAAI**), 11586-11594, 2023.

[13]. Aye Phyu Phyu Aung[*], **Xinrun Wang**[*†], Runsheng Yu[*], Bo An[†], Senthilnath Jayavelu, Xiaoli Li[†]. DO-GAN: A double oracle framework for generative adversarial networks. *Proceedings of the 2022 IEEE Conference on Computer Vision and Pattern Recognition* (**CVPR**), pp. 11275-11284, 2022.

[14]. Wei Qiu, **Xinrun Wang**, Runsheng Yu, Rundong Wang, Xu He, Bo An, Svetlana Obraztsova, Zinovi Rabinovich. RMIX: Learning risk-sensitive policies for cooperative reinforcement learning agents. *Proceedings of the 35th Annual Conference on Neural Information Processing Systems* (**NeurIPS**), pp. 23049-23062, 2021.

[15]. Shuxin Li, Youzhi Zhang[*], **Xinrun Wang**[*], Wanqi Xue, Bo An. CFR-MIX: Solving imperfect information extensive-form games with combinatorial action space. *Proceedings of the 30th International Joint Conference on Artificial Intelligence* (**IJCAI**), pp. 3663-3669, 2021.

[16]. Zhenyu Shi[*], Runsheng Yu[*], **Xinrun Wang**[*], Rundong Wang, Youzhi Zhang, Hanjiang Lai, Bo An. Learning expensive coordination: An event-based deep RL approach. *Proceedings of the 2020 International Conference on Learning Representations* (**ICLR**), 2020.