Jiajun **Fan**

(+86) 17822018001 | fanjj21@mails.tsinghua.edu.cn | www.jiajunfan.com

Education

Tsinghua University   *Beijing, China*

MASTER OF ELECTRONIC INFORMATION (COMPUTER SCIENCE)   *Sep 2021 - Jun 2024 (Expected)*

Nankai University   *Tianjin, China*

BACHELOR OF INTELLIGENT SCIENCE AND TECHNOLOGY   *Sep 2017 - Jun 2021*
- GPA: 3.93/4.0
- Courses: Reinforcement Learning, Computer Vision, Computer Composition Principle, Probability Theory and Mathematical Statistics

Academic Experience

Research on High Sample-efficiency Reinforcement Learning Paradigm   *Beijing, China*

$\mathcal{R\&D}$ INTERN   *Sep 2020 - Jun 2021*
- Extended the basic paradigm of RL called the **Generalized Policy Iteration (GPI)** into a more generalized version, which is called the **Generalized Data Distribution Iteration (GDI)**.
- **Unified massive RL algorithms into the GDI paradigm**, which can be considered as one of the special cases of GDI.
- Provided theoretical proof of why GDI is better than GPI and how it works.
- My work has achieved SOTA in Atari57 and broken through **massive human world records**, which also **outperformed Agent57**.
- Published a paper called **"GDI: Rethinking What Makes Reinforcement Learning Different From Supervised Learning"**. In addition, I have invented more than **9 national patents**.

Research on Model-free Reinforcement Learning Algorithm   *Beijing, China*

$\mathcal{R\&D}$ INTERN   *Sep 2020 - Jun 2021*
- Proposed **a unified framework of model-free reinforcement learning** called **CASA**.
- Proposed an entropy regularization free mechanism for policy-based reinforcement learning based on bandits.
- Published two papers called **"CASA: A Bridge Between Gradient of Policy Improvement and Policy Evaluation"** and **"An Entropy Regularization Free Mechanism for Policy-based Reinforcement Learning"**. Additionally, I have published several patents.

Research on Model-based Reinforcement Learning Algorithm   *Tianjin, China*

$\mathcal{R\&D}$ INTERN   *Apr 2020 - Sep 2020*
- Proposed a novel method to improve the algorithm learning performance and test the algorithm performance on the **MuJoCo simulation environment**.
- Published a paper called **"Critic PI2: Master Continuous Planning via Policy Improvement with Path Integrals and Deep Actor-Critic Reinforcement Learning"**.

Multifunctional Home Service Robot   *Tianjin, China*

$\mathcal{R\&D}$ INTERN   *Sep 2018 - Jul 2019*
- Proposed a solution for the entire project and a finite state machine diagram.
- Implemented autonomous navigation and RRT path planning related algorithms based on ROS.
- Project results have been presented in the **2019 ROBOCUP Sydney World Finals**.

Publications

Paper
- **Fan, J.**, Xiao, C., & Huang, Y. (2021). GDI: Rethinking What Makes Reinforcement Learning Different From Supervised Learning. arXiv preprint arXiv:2106.06232.
- **Fan, J.**, Ba, H., Guo, X., & Hao, J. (2020). Critic PI2: Master Continuous Planning via Policy Improvement with Path Integrals and Deep Actor-Critic Reinforcement Learning. arXiv preprint arXiv:2011.06752.
- Xiao, C., Shi, H., **Fan, J.**, & Deng, S. (2021). An Entropy Regularization Free Mechanism for Policy-based Reinforcement Learning. arXiv preprint arXiv:2106.00707.
- Xiao, C., Shi, H., **Fan, J.**, & Deng, S. (2021). CASA: A Bridge Between Gradient of Policy Improvement and Policy Evaluation. arXiv preprint arXiv:2105.03923.

Patent

- **Fan, J.** "Real-time multi-hyperparameter controller." CN113052252A. 2021.06.29.
- **Fan, J.** "Hyperspace multi-coupling parameter optimizer based on multi-arm gambling machine combined with democratic voting."CN113052253A. 2021.06.29.
- **Fan, J.** "Fast and generalizable hyperspace coupling multi-parameter nonlinear optimizer." CN113052248A. 2021.06.29.
- **Fan, J.** "Reinforcement learning algorithm based on generalized combination strategy space." CN113052312A. 2021.06.29.
- **Fan, J.** "Asynchronous multi-arm gambling machine hyperparameter optimizer based on electoral college voting mechanism." CN112949850A. 2021.06.11.
- **Fan, J.** "Hyperparameter tuning algorithm based on multi-arm gambling machine optimizer." CN112926629A. 2021.06.08.
- **Fan, J.**, Xiao, C. "An unbiased estimation algorithm of behavior value function." CN112926628A. 2021.06.08.
- **Fan, J.**, Xiao, C. "Fan, J., Xiao, C. Policy gradient algorithm based on double robust qualification trace." CN112926735A. 2021.06.08.
- **Fan, J.**, Xiao, C. "Unified framework for model-free reinforcement learning algorithms." CN112766497A. 2021.05.07.

Awards

2021	**Outstanding Graduates**, Nankai University	*Tianjin, China*
2021	**Excellent Graduation Thesis**, Nankai University	*Tianjin, China*
2021	**Tang Lixin Scholarship**, Nankai University	*Tianjin, China*
2020	**National Scholarship**, Nankai University	*Tianjin, China*
2020	**Nomination for Zhou Enlai Scholarship**, Nankai University	*Tianjin, China*
2019	**National Scholarship**, Nankai University	*Tianjin, China*
2019	**3st Prize**, Robocup@HOME Education World Final	*Sydney, Australia*
2019	**Bronze Medal**, ACM / ICPC Asia Regional Contest	*Xuzhou, China*
2018	**National 2st Prize**, National College Students Mathematical Contest in Modeling	*Tianjin, China*
2018	**The First Prize Scholarship**, Nankai University	*Tianjin, China*