

# Shanqi Liu 刘善琪

No.38 Zheda Road, Yuquan Campus, Zhejiang University  
Hangzhou, Zhejiang Province, China

(+86)18867150519 (Tel)  
shanqiliu@zju.edu.cn (Email)  
github.com/Locke637 (Github)

## Education Background

- **Zhejiang University** *College of Control Science and Engineering* 2019.09 – 2024.06 PhD
- **Zhejiang University** *College of Control Science and Engineering* 2015.09 – 2019.09 Bachelor
- **The University of Western Australia** *Summer Session* 2018.07 – 2018.09 Exchange Student

## Academic Research

**APRIL Laboratory** *Institute of Cyber Systems and Control, ZJU* 2019.09 – Present

- Research interests: Multi-agent reinforcement learning, imitation learning, hierarchical reinforcement learning
- Propose a decentralized, locally observed reinforcement learning algorithm to solve multi-agent in formation (MAiF) tasks. The algorithm uses a hierarchical reinforcement learning structure to decompose the multi-objective task into mutually decoupled tasks. Experimental results show that our algorithm has good mobility for map size variations.
- In order to improve the training process of reinforcement learning under incomplete information self-play games, a comprehensive global information critic is added to the policy gradient method to form a self-play actor-critic (SPAC) method for training agents to play computer games. The results under a variety of tasks such as competition and cooperation show that this method is better than the baseline algorithm.
- Propose a multi-agent reinforcement learning algorithm with centralized training and distributed execution that can adapt to dynamic number of agents. The attention mechanism is used to select several team leaders and establish a dynamic number of teams. The characteristics of the members are integrated within the team and the transfer of characteristics between teams can be realized. The experimental results show that the overall method has achieved better results than the baseline algorithm in several environments with variable number of agents.
- Propose a value function decomposition method for multi-agent reinforcement learning, which not only considers the rewards of independent actions of the agent, but also considers the rewards of cooperation with other nearby agents to solve the monotonic problem of QMIX-like methods. In addition, we propose a greedy action searching method, which can improve the exploration and make the agent's policy not affected by changes in the number of nearby agents or changes in the sequence of nearby agents' actions. Experimental results show that the overall method achieves a significant performance improvement in three challenging MAgent tasks, and can handle unseen cooperation scenarios.
- Propose a method of imitating learning from observations in a non-time-aligned environment, which adopts a hierarchical reinforcement learning structure to dynamically select feasible sub-goals from the observations of the expert trajectory. At the same time, it is possible to learn corresponding types of policy for different types of tasks by using the designed reward structure. In addition, three different methods are proposed to improve the sample efficiency in the hierarchical structure. The experimental results show that the performance and learning efficiency of the overall method are improved.

**The University of Western Australia**

2018.07 – 2018.09

- Participated in the project of the robotic unmanned driving system of the University of Western Australia. The architecture of the system consists of lane detection, traffic sign recognition, parking, communication, and man-machine interface. Several scenarios were considered in the project, including normal patrol mode and parking mode. In the normal patrol mode, the driverless car can automatically drive along the lane and detect various traffic signs in front of it in real time and react accordingly. For example, once a stop sign is detected, it will enter the parking mode and search for parking in the parking lot.

## Papers

- **S. Liu**, L. Wen, J. Cui, X. Yang and Y. Liu, "Moving Forward in Formation: A Decentralized Hierarchical Learning Approach to Multi-Agent Moving Together". (IROS 2021 Accepted)
- **S. Liu**, J. Cao, Y. Wang, W. Chen and Y. Liu, "Self-Play Reinforcement Learning with Comprehensive Critic in Computer Games." (Neurocomputing Accepted)
- W. Liu, **S. Liu**, J. Cao, Q. Wang, X. Lang and Y. Liu, "Learning Communication for Cooperation in Dynamic Agent-Number Environment." (IEEE/ASME Transactions on Mechatronics Accepted)
- S. Sun, J. Zheng, Z. Qiao, **S. Liu**, Z. Lin and T. Bräunl, "The Architecture of a Driverless Robot Car Based on EyeBot System." (Journal of Physics: Conference Series Accepted)
- **S. Liu**, W. Liu, W. Chen, J. Cao and Y. Liu, "Learning Multi-Agent Cooperation via Considering Actions of Teammates." (NeurIPS 2021, Under Review)
- **S. Liu**, J. Cao, W. Chen, L. Wen and Y. Liu, "HILONet: Hierarchical Imitation Learning from Non-Aligned Observations". (IEEE Transactions on SMC: Systems, Under Review)

## Skills

- **Programming Languages:** Python, MATLAB,  $\text{\LaTeX}$ , C Sharp
- **Familiar Libraries:** PyTorch, TensorFlow, Numpy, Matplotlib, Unity, ROS
- **Professional Knowledge:** Reinforcement learning, multi-agent system, imitation learning, hierarchical reinforcement learning
- **English Proficiency:** CET-4, CET-6, IELTS (7.0)
- **Undergraduate grade point:** 3.75/4.0

## Honors & Awards

- **MCM (The Mathematical Contest in Modeling)** *Honorable Mention* 2018.02
- **Outstanding PhD student of Zhejiang University** 2020.09
- **Zhejiang University Academic Scholarship** 2017.10
- **Top level rank in Zhejiang University Student Research Training Program** 2018.02
- **Zhejiang Student Physics Competition** *Third Prize* 2016.06
- **Champion of Zhejiang University CC98 Football Match** *Three times* 2019-2021
- **National Youth Physics Olympiad** *Provincial Third Prize* 2013/2014