

Liwen Zhu

Email: liwenzhu@pku.edu.cn

HomePage: www.liwenzhu-pku.cn

Mobile: +86-15600982277

WeChat: ZLW19971005



EDUCATION

• Peking University

School of Electronic and Computer Engineering

Beijing, China

Sep 2019 - June 2022

- **Master in Computer Application Technology:** GPA: 3.77 (Top 20%)
- **Advisor:** Yonghong Tian, Zongqing Lu
- **Laboratory:** Multimedia Learning Group, Institute of Digital Media (NELVT), Peking University
- **Courses:** Artificial intelligence(A+), Algorithm Analysis and Complexity Theory(A+), Computer Vision(A+), Fundamentals of Digital Media Technology(A+), Digital Signal Processing and Algorithm Implementation(A), Theory and Practice of Medical Image Processing(A+), Intelligent Robot Technology(A)

• Beijing University of Technology

College of Metropolitan Transportation

Beijing, China

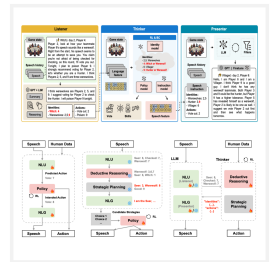
Sep 2015 - June 2019

- **Bachelor in Intelligent Traffic System:** GPA:3.91 (Top 1%)
- **Courses:** Advanced Mathematics(100), Linear Algebra(94), College Physics(96), Probability Theory and Mathematical Statistics(97), Advanced Language Programming C++(95), Automatic Control Principle(97), Principles and Practices of Database(95), Digital Electronic Technology(98), Data Structures and Algorithm (92), Principle of Microcomputer(96)

PUBLICATIONS

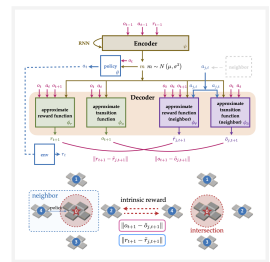
- Shuang Wu*, Liwen Zhu*(Coauthor), Tao Yang, Shiwei Xu, Qiang Fu, Wei Yang, Haobo Fu. **Enhance Reasoning for Large Language Models in the Game Werewolf.** (Under Review, submitted to NeurIPS 2024) [\[PDF\]](#)[\[Code\]](#):

- **LLM Agent:** Implemented a dual-system reasoning framework, enhancing LLMs' strategic planning capabilities in complex tasks.
- **LLM Reasoning:** Boosted LLMs' reasoning through a Thinker module trained with supervised and reinforcement learning.
- **Game AI:** Created an AI for Werewolf with a Thinker module, surpassing GPT4 in strategic gameplay.



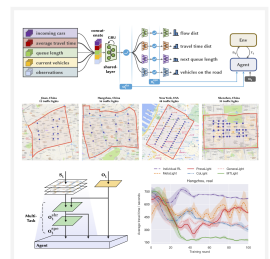
- Liwen Zhu, Peixi Peng, Zongqing Lu, Yonghong Tian. **Meta Variationally Intrinsic Motivated Reinforcement Learning for Decentralized Traffic Signal Control.** (TKDE 2022, IEEE Transactions on Knowledge and Data Engineering (CCF A).) [\[PDF\]](#)[\[Code\]](#):

- **Meta-RL:** We formulate multi-agent traffic signal control as a meta-learning problem over a set of related tasks.
- **Latent Variable:** Express task-specific information and trade-off the exploration and exploitation.
- **Intrinsic Reward:** An additional rewards for RL agent. Make the policy stable to neighbors.



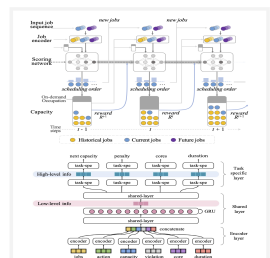
- Liwen Zhu, Peixi Peng, Zongqing Lu, Yonghong Tian. **MTLight: Efficient Multi-Task Reinforcement Learning for Traffic Signal Control.** (Gamification and Multiagent Solutions on ICLR 2022.) [\[PDF\]](#)[\[Code\]](#):

- **Multi-Task Learning:** Multiple auxiliary and supervisory tasks are constructed to learn the latent space from the global state.
- **Task-shared Latent space:** Embedding feature are extracted from the task-shared layer and could express general underlying characteristics to enhance the agent observation.
- **Task-specific Latent space:** Task-specific embedding feature is extracted by the task-specific branch and represents the task-driven information.

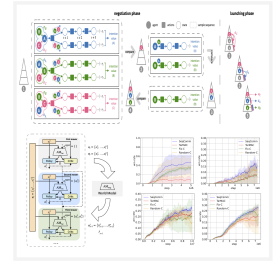


- Hang Dong*, Liwen Zhu*(Coauthor), Zhao Shan, Bo Qiao, Yonghong Tian. **Online Scheduling for Deferrable Workloads with Auxiliary Tasks Aided Reinforcement Learning in Cloud Computing.** :

- **Reinforcement Learning Based Scheduling:** A deep reinforcement learning model is adopted to learn the scheduling policy to increase resource utilization.
- **Integration of Auxiliary Tasks:** Several auxiliary tasks are utilized to provide better state representations and improve the performance.
- **Shared Encoder for Policy and Value Network:** Shared encoder module enforces the sharing of information between the policy and the state value and accelerates training.



- Ziluo Ding, Kefan Su, Weixin Hong, **Liwen Zhu**, Tiejun Huang, Zongqing Lu. **Multi-Agent Sequential Decision-Making via Communication**. [PDF]:
 - **Negotiation Phase:** Agents share observations with others and obtain their intention by modeling the environment dynamics. Agents determine the priority of decision-making by comparing the value of intention.
 - **Launching Phase:** Agents communicate for obtaining messages to make decision, the upper-level agents take the lead in making decisions and share their actions with the lower-level agents.
 - **Attention Module:** Attention module is applied to process messages in the world model, critic network, and policy.
- Patent: **Liwen Zhu**, Guohua Wang, Zefeng Weng, Hai Li, Qianben Chen. **A Method of Enterprise Relationship Inference for Financial Big Data**. :
 - **key point:** Enterprise Portrait, Relational Inference, Knowledge Graph, Knowledge Vector
- Patent: Yonghong Tian, **Liwen Zhu**, Peixi Peng, Wen Gao. **An Intrinsic Rewarded Meta-Reinforcement Learning Method for Traffic Signal Control**. :
 - **key point:** The application of MARL, a decentralized method for large-scale traffic signal control.



EXPERIENCE

- **Tencent - Research Engineer** Beijing, China
Research: LLM Agent, RLHF July 2022 - Latest
 - **WeChat Large Language Model(WeLM): in WeChat AI**
 Responsible for the LLM alignment (RLHF). This includes generating pairwise ranking data, processing human-labeled data, and refining the gap alignment algorithm in the reward model. I direct the iterative optimization of the PPO model, continuously refining its alignment with human preferences across successive SFT versions. This method boosts model performance, showing marked improvements in both automated and human evaluations.
 - **LLM Game Agent: in Tencent AI Lab**
 Developed an advanced game-playing AI agent specialized in complex strategy games. Leveraged large language models to enhance decision-making and game reasoning skills, incorporating advanced AI techniques to outperform conventional models in strategic depth and adaptability.
- **Microsoft Research Asia - Intern (Star of Tomorrow Honor)** Beijing, China
Research: Deep Reinforcement Learning, Cloud Computing Oct 2021 - Mar 2022
 - **Reinforcement Learning:** Model job scheduling task and use reinforcement learning methods for online scheduling, which outperforms heuristic methods such as FIFO and SJF.
 - **Multi-Task Learning:** Construct related supervised auxiliary tasks and obtain environmental representations to augment the state space.
 - **Cloud Computing:** The proposed method can well plan the deployment schedule and achieve a short waiting time for users while maintaining a high resource utilization for the platform.
- **Tencent - Algorithm Intern** Shenzhen, China
Research: RL, NLP May 2021 - July 2022
 - **Reinforcement Learning for RE:** The relation extraction task is modeled as a reinforcement learning decision problem. The dense vectors of tokens in the sentence are regarded as the agent state, the agent action is a predicted relation, relation that predicted accurately will be rewarded.
 - **Relation Extraction:** Extract the relations between entity pairs in financial news. The relation types include sub-parent company, investment holding, tenure, alias, etc., using graph convolution and pre-training model methods.
 - **Aspect-based Sentiment Analysis:** Analyze the sentiment attitude of financial news, set scheduled tasks on ES, and complete the deployment on Docker.



微软亚洲研究院

Tencent 腾讯

- **Pengcheng Laboratory - Algorithm Intern** Shenzhen, China
Research: Multi-agent Reinforcement Learning, Traffic Signal Control July 2019 - July, 2022
 - **Meta Variationally Intrinsic Motivated RL:** Formulated decentralized traffic signal control as a meta-learning problem, using a learned latent variable for task-specific information in policy learning. Designed an intrinsic reward to enhance policy invariance across neighbouring agents.
 - **Dataset for RL Simulation:** Developed a simulated traffic flow for Shenzhen, China, using data from traffic and monitoring cameras to create road network and traffic flow files for reinforcement learning simulations.
 - **PengBo RL Platform:** Constructed an integrated platform supporting multiple RL tasks (Atari, Mujoco, CityFlow, Spiking RL) and algorithms, set for future release.



VOLUNTEER EXPERIENCE

- **Conference Reviewer: ICML 2022, NeurIPS 2022/2023, ICLR 2023/2024 (invited, upcoming)** Beijing, China
Reviewed several theoretical and applied papers on multi-agent reinforcement learning. February-July, 2022
- **Speaker at The 9th World Radio Detection & Ranging Expo. (WRE)** Nanjing, China
Gave a speech at The 9th World Radar Expo and the first Fadar and the Future Global Summit. April 22-24, 2021
- **Volunteer of IEEE 3th International Conference on MIPR. (IEEE MIPR 2020)** Shenzhen, China
Make workshop webpage AI Art 2020, host the zoom meeting for the main session. August 6-8, 2020
- **Volunteer of New Generation Artificial Intelligence Academician Summit Forum** Shenzhen, China
Organize events, ensure the success of the venue progress. December 20-21, 2019

HONORS AND AWARDS

- **Tencent Open Source Collaboration Award** - December, 2022
- **Star of Tomorrow Honor of Microsoft Research Asia** - March, 2022
- **Beijing Excellent Undergraduate Thesis** - Dec, 2019
- **Outstanding Graduates of Beijing** - July, 2019
- **National Scholarship** - Dec, 2018
- **Excellent Graduation Thesis** of Beijing University of Technology - July, 2019
- **First Prize** of TIS Scholarship for Technological Innovation and Practice - Mar, 2019
- **Ren Futian Scholarship Second Prize** of Science and Technology Innovation - Dec, 2018
- **Meritorious Winner Award of Mathematical Contest in Modeling** Certificate of Achievement - April, 2019
- **Honorable Mention Award of Mathematical Contest in Modeling** Certificate of Achievement - April, 2018
- **Silver Award** of "Chuang Qing Chun" Capital University Student Entrepreneurship Competition - May, 2018
- **Second Prize** of Beijing University Student Transportation Science and Technology Competition - Nov, 2018
- **Third Prize** in the Global Management Challenge (GMC) China Competition - Dec, 2017
- **Third Prize** of the 2018 National College Students Academic English Vocabulary Competition - May, 2018
- **Grand Prize** of iCAN International Innovation and Entrepreneurship Competition - July, 2018
- **Second prize** of Beijing University of Technology **Robotics Competition** - May, 2017
- **Third Prize** of Beijing University of Technology **Dingxin Cup** Student Academic Works Competition - Dec, 2016

SKILLS SUMMARY

- **Languages:** Python, C++, Bash, HTML, SQL, Hadoop
- **Frameworks:** Gym, MuJoCo, CityFlow, MetaWorld, Scikit-learn, PyTorch, TensorFlow, Keras
- **Tools:** Linux, Git, Docker, Bootstrap
- **Soft Skills:** IELTS 7.5, CET-6 and CET-4, Public Speaking, Web Programming, Time Management, Self-motivated