

OPTIMAL POLICY PROBLEM OF BIPEDALWALKER ENVIRONMENT IN OPENAI GYMNASIUM USING DEEP REINFORCEMENT LEARNING

TONE-YAU HUANG
FENG-CHIA UNIVERSITY IN TAIWAN

In this talk, we will focus on deep reinforcement learning algorithms, introducing the Actor-Critic Method, Proximal Policy Optimization (PPO), and other related methods. Furthermore, we will investigate optimal policy generation using the BipedalWalker environment from OpenAI Gymnasium as our primary case study.

REFERENCES

- [1] Richard S.S. and Andrew G.B. (2018). **Reinforcement Learning, an introduction**. The MIT Press.
- [2] John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, Oleg Klimov. **Proximal Policy Optimization Algorithms**. OpenAI (2017)
- [3] Mnih, V. et al. **Playing Atari with deep reinforcement learning**. DeepMind Technologies (2013)
- [4] Schott, L., Hajri, H. and Lamprier, S. **Improving Robustness of Deep Reinforcement Learning Agents: Environment Attack based on the Critic Network**. 2022 International Joint Conference on Neural Networks (IJCNN).
DOI: 10.1109/IJCNN55064.2022.9892901 (2022)
- [5] Lillicrap, T.P., Hunt, J.J., Pritzel, A., Heess, N., Erez, T., Tassa, Y., Silver, D. and Wierstra, D. **Continuous Control With Deep Reinforcement Learning**. Google Deepmind (2016)
- [6] Chen, X.C., Wei, J.Q., Ren, X.Q., Johansson, K.H. and Wang, X.F. **Automatic Overtaking on Two-way Roads with Vehicle Interactions Based on Proximal Policy Optimization**. 2021 IEEE, Intelligent Vehicles Symposium (IV)
DOI: 10.1109/IV48863.2021.9575954 (2021)

(T.-Y. Huang) DEPARTMENT OF APPLIED MATHEMATICS, FENG CHIA UNIVERSITY, TAICHUNG 407102, TAIWAN.

Email address: `huangty@o365.fcu.edu.tw`, `tyhuang2020@mail.fcu.edu.tw`