Project - CS747

Hoang-Dung Bui
NetID: G01301478

1 Literature Search

One of the challenge in robotic path-finding field is the robust of path-finder to deal with the changing of environment or robot's starting and ending points. The traditional methods will face a problem of long computation time, and it usually does not meet the requirement for real-time applications. There are number of research applying machine learning to handle the problem of robotic path-finding. In [1, 2, 3], the authors developed motion planning networks whose inputs include new environments or new starting and ending points. The networks were designed to handle static environments, however, it cannot work for dynamical ones.

That is a gap which reinforcement learning (RL) can be applied. In [4, 5, 6, 7], the authors used RL algorithms to to determine pathfinding for multi-agent. However, the proposed frameworks were not handled well for variant environments, or in others words, they need to be retrained for new environments.

In [8, 9], the authors contributed frameworks for on robotic path-finding for 8-connected (8 actions) gripmap for single or multi-agent. The propose algorithm could work on dynamical environments with partial observation for the agents. The simulated environments were gridmaps with dimension varying from 10x10 to 30x30. They are quite small as comparing to the real industrial application.

In this research, we will develop a deep RL algorithm which can help a single agent to find a path from a start- to end- points on a gripmap with the size of 50 or bigger.

2 Methods

To achieve this objective, we construct an gripmap environment with any random obstacles. The rewards for the agent is determined by a potential function, where states closes to obstacles will receive a low negative score. It ensures that the agent will avoid the obstacles as far as possible.

Based on the state-of-the-art RL algorithms such as PPO [10], Double DQN [11], Dueling DQN [12], we will combine and improve the algorithms to apply into our scenarios. In the end of the projects, we expect our algorithms can meet so requirements as following:

- Work well on any start- and end- points
- Avoid the obstacles
- Work well on several environments

References

[1] Ahmed H Qureshi, Anthony Simeonov, Mayur J Bency, and Michael C Yip. Motion planning networks. In *2019 International Conference on Robotics and Automation (ICRA)*, pages 2118–2124. IEEE, 2019.

[2] Simon X Yang and Max Meng. An efficient neural network method for real-time motion planning with safety consideration. *Robotics and Autonomous Systems*, 32(2-3):115–128, 2000.

[3] Jiankun Wang, Wenzheng Chi, Chenming Li, Chaoqun Wang, and Max Q-H Meng. Neural rrt*: Learning-based optimal path planning. *IEEE Transactions on Automation Science and Engineering*, 17(4):1748–1758, 2020.

[4] Hongda Qiu. Multi-agent navigation based on deep reinforcement learning and traditional pathfinding algorithm. *arXiv preprint arXiv:2012.09134*, 2020.

[5] Robbert Reijnen, Yingqian Zhang, Wim Nuijten, Caglar Senaras, and Mariana Goldak-Altgassen. Combining deep reinforcement learning with search heuristics for solving multi-agent path finding in segment-based layouts. In *2020 IEEE Symposium Series on Computational Intelligence (SSCI)*, pages 2647–2654. IEEE, 2020.

[6] Guillaume Sartoretti, Justin Kerr, Yunfei Shi, Glenn Wagner, TK Satish Kumar, Sven Koenig, and Howie Choset. Primal: Pathfinding via reinforcement and imitation multi-agent learning. *IEEE Robotics and Automation Letters*, 4(3):2378–2385, 2019.

[7] Jinglun Yu, Yuancheng Su, and Yifan Liao. The path planning of mobile robot by neural networks and hierarchical reinforcement learning. *Frontiers in Neurorobotics*, 14, 2020.

[8] Aleksandr I Panov, Konstantin S Yakovlev, and Roman Suvorov. Grid path planning with deep reinforcement learning: Preliminary results. *Procedia computer science*, 123:347–353, 2018.

[9] Vasilii Davydov, Alexey Skrynnik, Konstantin Yakovlev, and Aleksandr I Panov. Q-mixing network for multi-agent pathfinding in partially observable grid environments. *arXiv preprint arXiv:2108.06148*, 2021.

[10] John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov. Proximal policy optimization algorithms. *arXiv preprint arXiv:1707.06347*, 2017.

[11] Hado Van Hasselt, Arthur Guez, and David Silver. Deep reinforcement learning with double q-learning. In *Proceedings of the AAAI conference on artificial intelligence*, volume 30, 2016.

[12] Ziyu Wang, Tom Schaul, Matteo Hessel, Hado Hasselt, Marc Lanctot, and Nando Freitas. Dueling network architectures for deep reinforcement learning. In *International conference on machine learning*, pages 1995–2003. PMLR, 2016.