Review: MODRL

Norio Kosaka

December 16, 2018

1 Paper Profile

- Title: A Multi-Objective Deep Reinforcement Learning Framework
- Authors: Thanh Thi Nguyen
- Organisation: Deakin University
- Publish Year: March 2018
- URL: https://arxiv.org/ftp/arxiv/papers/1803/1803.02965.pdf

2 Contents

3 Proposal

The present paper proposed a new MODRL framework on DQN coded in Python. It includes the use of linear and non-linear methods to develop the framework which is able to accommodate both single-policy and multi-policy strategies. And the resulting performance on the two previously proposed benchmarks, which are the Deep Sea Treasure and Mountain-Car problems, indicate the convergence to the optimal Pareto solutions effectively.

4 Introduction

The author has summarised the past researches by quoting some good researches as below.

- Q-Learning[6]: Tabular method requiring huge memory so that impractical in practice.
- DRL(e.g., Deep Q network by Mnih et al., 2015[1]) to overcome the problem using experience replay and function approximation techniques
- Mossalam et al. (2016) [2] extended deep Q-network to handle single-policy linear MORL
- Tajmajer(2017) [3] has extended DQN with a non-linear action selection approach based on a subsumption architecture
- Vamplew et al.(2017) [5] developed an MORL framework named MORL_Glue which is based on RL_Glue(Tanner and White, 2009) [4]. Unfortunately, this framework was not compatible with Deep neural networks.

5 Proposal

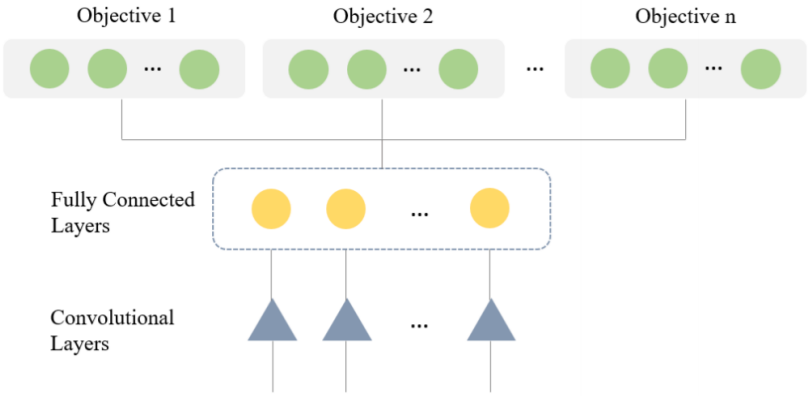

Fig. 1. Neural network structure used in our DQN-based MODRL framework.

Reference

[1] Volodymyr Mnih et al. "Human-level control through deep reinforcement learning". In: *Nature* 518.7540 (2015), p. 529.

[2] Hossam Mossalam et al. "Multi-objective deep reinforcement learning". In: *arXiv preprint arXiv:1610.02707* (2016).

[3] Tomasz Tajmajer. "Multi-Objective Deep Q-Learning with Subsumption Architecture." In: *arXiv preprint arXiv:1704.06676* (2017).

[4] Brian Tanner and Adam White. "RL-Glue: Language-independent software for reinforcement-learning experiments". In: *Journal of Machine Learning Research* 10.Sep (2009), pp. 2133–2136.

[5] Peter Vamplew et al. "Steering approaches to Pareto-optimal multiobjective reinforcement learning". In: *Neurocomputing* 263 (2017), pp. 26–38.

[6] Christopher J.C.H. Watkins and Peter Dayan. "Technical Note: Q-Learning". In: *Machine Learning* 8.3 (May 1992), pp. 279–292. ISSN: 1573-0565. DOI: 10.1023/A:1022676722315. URL: https://doi.org/10.1023/A:1022676722315.