



Exploration Strategies for Deep Reinforcement Learning

Nat Dilokthanakul

Imperial College London, London, UK
E-mail: n.dilokthanakul14@imperial.ac.uk

Autonomous vehicles, data center cooling systems and warehouse operation planners are some examples of engineering systems that operate by acquiring data and using it to make automated decisions. In order to create such systems, classic solutions use expert-knowledge and hand-engineer rules to specify appropriate responses for all possible scenarios. Reinforcement learning is a machine-learning paradigm that could be used to replace expensive hand-engineering tasks with systems that can learn through trial-and-error. However, the current state of reinforcement learning requires large sets of data which becomes impractical in some applications. In this work, we explore several techniques that can improve data inefficiency in reinforcement learning. We focus on an area called **deep reinforcement learning** (DRL) that aims to solve high dimensional problems via the use of deep neural networks. DRL commonly uses shortsighted exploration strategies, which results in the need for large amounts of interactions with the environment. This problem is especially severe in the problems where a series of precise coordinated actions are needed to achieve some learning-feedback signals. In order to alleviate this problem, we propose two exploration strategies; (i) an exploration heuristic based on risk measurement which encourages risk-seeking behaviours and (ii) a hierarchical architecture that endows our algorithm with an ability to explore in temporally abstracted manner. We evaluated our methods on a set of Atari 2600 games, which are classics games in Atari console designed to be difficult even for human players. We highlight the abilities of our methods in one of the hardest game for DRL -- Montezuma's Revenge --where we show a large improvement from the baselines in terms of final scores and data efficiency.

Keywords: Deep Reinforcement Learning, Exploration-Exploitation, Data Efficiency

Explain the impacts of your work especially to Thailand.

This work is a basic research, which aims to improve our knowledge in machine learning and artificial intelligence. The ideas and techniques studied can then be adopted by practitioners to create automated systems of various kinds. For example,

- (i) An automated energy control systems can be used to reduce energy consumptions in industrial factories, data centers or power-plants.
- (ii) An autonomous robot that can performs complicated tasks to assist in any demand from professionals to household use.
- (iii) An autonomous planner which can suggests best courses of actions for business managers, doctors, law-enforcements, etc.

Explain how your work fit with the ATPER conference theme

This work falls into the categories of autonomous robot and resource management for smart society.