

Master/Bachelor Thesis - Semester Project

Reinforcement Learning via Hindsight Experience Replay (HER)

Background

Deep reinforcement learning has been demonstrated as a powerful framework for robotics to design sophisticated and hard-to-engineer behaviors. However, defining a proper and efficient reward is usually challenging for complicated tasks involving multiple factors and engineering labors. Therefore, it is essential to investigate strategies to learn unshaped and usually sparse rewards. Recently, Hindsight Experience Replay (HER) has been proposed to allow off-policy RL algorithms to perform effective learning in solving goal-based tasks with sparse/binary rewards, such as the manipulation of robotic arms [1]. HER creates a dense reward setting from a sparse/binary task by turning each failed episode to a success by composing a new task whose desired virtual goal is the same as the achieved goal by that episode. On the basis of HER, two research directions are mainly studied, namely, design goal generation methods for more complicated tasks [2] or improve the sample efficiency for fast and better learning process [3]. However, there still are a set of research problems that can lead to widely-applicable RL improvements [4].

Your Tasks

In this thesis, your task will be learning state-of-the-art knowledge of reinforcement learning and HER and then develop more advanced algorithms compared vanilla HER. To be specific:

1. You will first learn basic knowledge of reinforcement learning. Online materials are recommended, such as [David Silver course](#) and [Stanford CS 234](#).
2. You will reproduce the [results from HER](#) and other related research results. By doing this, you will have a deep understanding of HER and the state-of-the-art research results.
3. You will choose from one of the possible ideas that we offer, investigate a novel algorithm that makes improvement on the vanilla HER.

Requirement

- High self-motivation;
- Six month working time;
- Experiences or knowledge from related courses
- Python programming experiences.

Supervisor: Prof. Alois Knoll

Advisor: Zhenshan Bing

bing@in.tum.de

Lehrstuhl für Echtzeitsysteme und Robotik,

Fakultät für Informatik, Technische Universität München

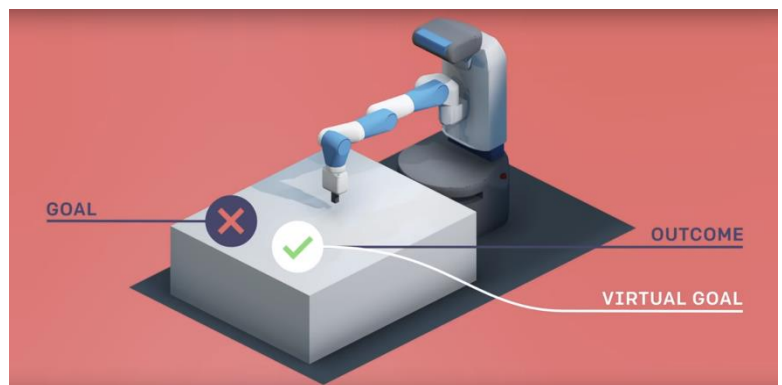


Figure 1 FetchSlide environment

[1] Andrychowicz, Marcin, et al. "Hindsight experience replay." *Advances in Neural Information Processing Systems*. 2017.

[2] Ren, Zhizhou, et al. "Exploration via Hindsight Goal Generation." *NeurIPS* (2019).

[3] Zhao, Rui, Xudong Sun, and Volker Tresp. "Maximum Entropy-Regularized Multi-Goal Reinforcement Learning." *ICML*, 2019

[4] Plappert, et al. "Multi-Goal Reinforcement Learning: Challenging Robotics Environments and Request for Research." 2018