

Internship offer: Integrating Human Demonstrations in Hierarchical Reinforcement Learning

Laboratory: U2IS, ENSTA Paris & LIX, Ecole Polytechnique

Duration: 6 months

Context: Fully autonomous robots have the potential to impact real-life applications, like assisting elderly people. Autonomous robots must deal with uncertain and continuously changing environments, where it is not possible to program the robot tasks. Instead, the robot must continuously learn new tasks and how to perform more complex tasks combining simpler ones (i.e., a task hierarchy). This problem is called lifelong learning of hierarchical tasks.

Summary: Hierarchical Reinforcement Learning (HRL) is a recent approach for learning to solve long and complex tasks by decomposing them into simpler subtasks. HRL could be regarded as an extension of the standard Reinforcement Learning (RL) setting as it features high-level agents selecting subtasks to perform and low-level agents learning actions or policies to achieve them. We recently proposed a HRL algorithm, GARA (Goal Abstraction via Reachability Analysis), that aims to learn an abstract model of the subgoals of the hierarchical task.

However, HRL can still be limited when faced with the states with high dimension and the real-world open-ended environment. Introducing a human teacher to Reinforcement Learning algorithms has been shown to bootstrap the learning performance. Moreover, active imitation learners such as in [1] have shown that they can strategically choose the most useful questions to ask to a human teacher : they can choose, who, when, what and whom to ask for demonstrations [2,3].

This internship's goal is to explore how active imitation can improve the algorithm GARA. The intuition in this context is that human demonstrations can be used to determine the structure of the task (ie. which subtasks need to be achieved) as well as determining a planning strategy to solve it (ie. the order of achieving subtasks).

During this internship we will :

- Study the relevant state-of-art and make a research hypothesis about the usefulness of introducing human demonstrations into the considered HRL algorithm.
- Design and implement a component to learn from human demonstrations in GARA.
- Conduct an experimental evaluation to assess the research hypothesis.

The intern is expected to also collaborate with a PhD student whose work is closely related to this topic.

References:

- [1] Cakmak, M., DePalma, N., Thomaz, A. L., and Arriaga, R. (2009). Effects of Social Exploration Mechanisms on Robot Learning. {IEEE} International Symposium on Robot and Human Interactive Communication(128--134).
- [2] Duminy, N., Nguyen, S. M., and Duhaut, D. (2019). Learning a Set of Interrelated Tasks by Using a Succession of Motor Policies for a Socially Guided Intrinsically Motivated Learner. *Frontiers in Neurobotics*, 12(87).
- [3] Nguyen, S. M. and Oudeyer, P.-Y. (2012). Active choice of teachers, learning strategies and goals for a socially guided intrinsic motivation learner. *Paladyn Journal of Behavioural Robotics*, 3(3)(136-146). SP Versita.