Extending the Hierarchical Deep Reinforcement Learning framework

Stefano Gatto, Master of Science in Computer Science
University of Dublin, Trinity College, 2018

Supervisor: Mads Haahr

Deep Reinforcement Learning (DRL) allows to train an agent to perform a task in an environment. The generality of this concept makes DRL a very powerful tool, applicable to a multitude of fields. Self-driving cars, robot locomotion, drone control and video games AI are only some of the fields which currently use and research DRL. To train an agent to perform a task, a DRL algorithm learns a policy that maximizes the future expected reward, thus ensuring that the agent will act optimally with respect to a reward signal. When this reward is very sparse or the task to learn is very complex, it is hard for a DRL algorithm to learn to perform it. When faced with a complex task to perform, the human brain decomposes the actions to take into simpler ones, developing a hierarchical understanding of the problem. Hierarchical Deep Reinforcement Learning (HDRL) brings this type of understanding to DRL models, allowing them to tackle also very complex and sparse reward tasks.
SAC-LSP is a HDRL algorithm that allows to grow hierarchies of policies in bottom-up layer-wise fashion. Each layer of the hierarchy is trained to perform the main task, and even if the lower layers are not able to completely solve it, they “make the job easier” for the higher layers. SAC-LSP yielded state-of-the-art results on a series of simulated locomotion and control environments. In this research, we propose four different optimizations to SAC-LSP, and evaluate their performance. The main contribution of this work is Distributed SAC-LSP, an optimization that outperforms the baseline by 40%.