Incentivizing exploration in curiosity-driven deep reinforcement learning

Patrik Reizinger[1], Márton Szemenyei[1]

[1] Department of Control Engineering and Informatics, Budapest University of Technology and Economics, Budapest, Hungary

In deep reinforcement learning, the problem of credit assignment is a topic of active research, especially when the reward provided by the environment (called the external reward) is sparse – which is the case for the most environments modelling real tasks. Several model-based methods exist to mitigate the problem, among which curiosity-based ones utilize a so-called intrinsic reward (which is given by the agent to itself) to bridge the gap between external rewards. This approach usually rewards the agent if it is able to model the dynamics of the environment. A recent work of Pathak et al. does this by combining two factors: they define a loss for forward and one for backward dynamics, where the former penalizes if the next state cannot be predicted, while the latter does so for the action between two consecutive states. Their results are quite promising, as the agent trained with this definition of curiosity is able to learn general skills or to be less sensitive for noise injected into the states.

In this work, we investigate the problem of exploration for curiosity-based reinforcement learning agents. Our aim is to improve the exploration of the agent even if rewards are very sparse. Although the methodology referenced above implies that the agent will explore the state space, we argue that using an explicit incentive could help the agent to better explore. To achieve this, we add a new factor to the loss function of the A2C network of the model, which aims to provide a reward if states cannot be predicted well. This additional loss ensures that the agent will visit states never seen before (or which are hard to predict). Our approach is evaluated on standard tasks included in the Gym benchmark suit of OpenAI.

Contact address: rpatrik96@sch.bme.hu