Abstract

Reinforcement learning can solve decision-making problems and train an agent to behave in an environment according to a predesigned reward function. However, such an approach becomes very problematic if the reward is too sparse and the agent does not come across the reward during the environmental exploration. The solution to such a problem may be in equipping the agent with an intrinsic motivation, which will provide informed exploration, during which the agent is likely to also encounter external reward. Novelty detection and state prediction are the promising branches of intrinsic motivation research. We present **Self-supervised Network Distillation** (SND), a class of intrinsic motivation algorithms based on the distillation error as a novelty indicator, where the target model is trained using self-supervised learning. We adapted three existing self-supervised methods for this purpose and experimentally tested them on a set of ten environments that are considered difficult to explore. We also applied self-supervised learning in the training of the forward model, denoted **Self-supervised Predictor** (SP) and showed that even for this approach to intrinsic motivation, it leads to an improvement in the agent's performance. The results show that our approach achieves faster growth and higher external reward for the same training time compared to the baseline models, which implies improved exploration in a very sparse reward environment.

Keywords: reinforcement learning, self-supervised learning, intrinsic motivation, novelty detection, state prediction, knowledge distillation, sparse reward