Learning to Generate Goals via Curriculum Self Play

August 2021

Project Status

We are interested in training general-purpose reinforcement learning agents that can solve a wide variety of goals. Training such agents without excessive interactions requires automatic generation of a *goal curriculum*. This is challenging as it requires (a) exploring goals of increasing difficulty while ensuring that the agent (b) is exposed to a diverse set of goals in a sample efficient manner and (c) does not catastrophically forget previously solved goals.

In this project, we propose *Curriculum Self Play* (CuSP), an automated goal generation framework that seeks to satisfy these desiderata by virtue of a multi-player game with 4 agents.

Our method builds on the two-player method introduced in Asymmetric Self-Play (ASP) and carefully balances cooperation and competition between two off-policy student learners and two regret-maximizing teachers that generate goals which are challenging yet feasible. The teachers also account for goal coverage and the non-stationary nature of the students, allowing us to automatically induce a curriculum that balances progressive exploration with anti-catastrophic exploitation. Our experiments demonstrate that our method succeeds at generating effective curricula of goals for a range of complex robotics tasks and significantly outperforms ASP and other competing methods at zero-shot test-time generalization to novel out-of-distribution goals. We evaluate across a range of control tasks, from locomotion, to manipulation, to navigation, and find that in more challenging in-distribution goals, our method is able to succeed where the baselines do not.

This project is currently under review and a preprint will be released soon.

Future Work

As future directions, we are currently exploring scaling up CuSP to more complex goal domains and tasks (eg. manipulating a variety of objects), where more simple curriculum generation methods are unlikely to succeed. We are also interested in extending and combining the self play framework with other methods for curriculum generation, such as learning from demonstrations.