

SC42050 Literature Assignment

Evolution strategies or reinforcement learning?

Wouter Wolfslag

Learning robots have the potential to perform tasks that are not fully known or understood beforehand. The (reinforcement) learning algorithms discussed in the lecture build upon Markov Decision Processes (MDPs). Within such a process they learn what actions to take as to maximize a reward. The challenge is to learn quickly and reliably. The paper (Salimans et al., 2017) discusses an alternative to MDP-based reinforcement learning, named evolution strategies (ES). After reading the paper (and references therein), answer the following questions.

1. ES is one of many so called black box optimization techniques. Explain why ES is considered a good option for reinforcement learning? Also explain the choice for the specific ES used.
2. The paper uses two important performance measures: sample-efficiency and wall-clock time. Explain what they are, how they are similar or different and how they depend on each other. Which measure is most important in a robotics setting?
3. For the robot-simulator results, the paper shows that the sample-efficiency of the tested Reinforcement Learning algorithm (TRPO) is often higher than that of ES. Compare the ES-approach with both MDP-based algorithms in general, and TRPO in particular, and explain why this increase in efficiency is either expected or surprising.

References

Salimans, T., Ho, J., Chen, X., and Sutskever, I. (2017). Evolution strategies as a scalable alternative to reinforcement learning. *arXiv preprint arXiv:1703.03864*.