

IN2349 ADLR: Project Ideas

April 29, 2020

Here you can find a number of ideas for projects we collected. Take this as an inspiration for your own project. Some of the ideas are rather "big", meaning they could result in more than one project. After you have registered your team (including a draft proposal) you will discuss the extent of your final proposal with your assigned tutor.

1 Closing the Sim-to-Real Loop: Adapting Simulation Randomization with Real World Experience

- In [Chebotar et al., 2018] an efficient method for solving the sim2real by iteratively adapting the simulation parameters to the real system is proposed.
- Change the experimental setup to a sim2sim setting because no real robot is available (i.e., try to adapt one simulation to a given one with unknown parameters).
- Investigate the convergence of estimated parameters.
- Investigate the influence of a broader starting distribution on the final performance.
- Investigate the influence of fully training the policy in every iteration vs. iterative training (the paper only covers full retraining).

2 Comparing robustness of learned (quadruped walking) policies from different learning algorithms

- Investigate how stochastic policies (SAC, PPO) perform in novel environments when compared to deterministic algorithms like DDPG. SAC and PPO claim to be far more robust.
- Investigate the influence of different distributions used for SAC (first SAC paper used GMMs) on the performance in novel environments.
- Investigate the methods for a quadruped with a more complex powertrain dynamics than the one used in the SAC paper. This can be combined with the Combinable with 3, 4b

3 Comparing the different methods for uncertainty estimation

Interesting methods include (but are limited to) MC-Dropout [Gal and Ghahramani, 2016] and Normalizing Flows [Louizos and Welling, 2017]. These methods could be compared in vastly different settings.

- Investigate how the uncertainty estimation changes during the training process (relevant to RL since we generally don't update the networks until convergence before collecting more data).
- Investigate which methods are best suited for active learning in the framework proposed by [Gal et al., 2017].
- Investigate which methods perform best for DQNs in simple environments similar to the work by [Touati et al., 2018].
- Come up with your own ideas.

4 Curiosity-Driven Learning

A good starting point for projects in this area is the paper by *Large-Scale Study of Curiosity-Driven Learning* [Burda et al., 2018]. Individual projects could first reproduce some of the experiments and then:

- Compare with other curiosity "measures" (count based, predict next state, predict past action, random network distillation, reachability, goal based, and/or your own ideas).
- Extend the authors comparison to other sets of environments. One key question is whether or not there are certain characteristics that a task needs to fulfill in order to profit from curiosity-driven learning.

5 Geometric Representations in Reinforcement Learning

Note: Requires previous experience with GNNs [Kipf and Welling, 2016].

- Modify PyBullet environments (Hopper, Walker, HalfCheetah, Ant) such that the observations contain a graph representing the robot. Use message passing network(s) in addition or instead of the MLP for value/Q function and policy in standard algorithms like PPO [Schulman et al., 2017] or SAC [Haarnoja et al., 2018].

References

- [Burda et al., 2018] Burda, Y., Edwards, H., Pathak, D., Storkey, A., Darrell, T., and Efros, A. A. (2018). Large-scale study of curiosity-driven learning.
- [Chebotar et al., 2018] Chebotar, Y., Handa, A., Makoviychuk, V., Macklin, M., Issac, J., Ratliff, N. D., and Fox, D. (2018). Closing the sim-to-real loop: Adapting simulation randomization with real world experience. *CoRR*, abs/1810.05687.

- [Gal and Ghahramani, 2016] Gal, Y. and Ghahramani, Z. (2016). Dropout as a bayesian approximation: Representing model uncertainty in deep learning. In *Proc. Int. Conference on Machine Learning*.
- [Gal et al., 2017] Gal, Y., Islam, R., and Ghahramani, Z. (2017). Deep bayesian active learning with image data.
- [Haarnoja et al., 2018] Haarnoja, T., Zhou, A., Abbeel, P., and Levine, S. (2018). Soft actor-critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor. In *International Conference on Machine Learning*.
- [Kipf and Welling, 2016] Kipf, T. N. and Welling, M. (2016). Semi-supervised classification with graph convolutional networks.
- [Louizos and Welling, 2017] Louizos, C. and Welling, M. (2017). Multiplicative normalizing flows for variational bayesian neural networks.
- [Schulman et al., 2017] Schulman, J., Wolski, F., Dhariwal, P., Radford, A., and Klimov, O. (2017). Proximal policy optimization algorithms. *CoRR*, abs/1707.06347.
- [Touati et al., 2018] Touati, A., Satija, H., Romoff, J., Pineau, J., and Vincent, P. (2018). Randomized value functions via multiplicative normalizing flows.