SC42050 Literature Assignment

Learning Robust Policies for Control

Ivan Koryakovskiy

Mechanically and electronically, robotics have advanced to the point where cognitive abilities have become the main limiting factor. While robots can flawlessly execute a set of commands to achieve a task, these commands are mostly encoded or tuned by hand. Reinforcement Learning allows finding an optimal sequence of commands without any prior assumption about the world. However, the application of pure learning to real systems is very limited due to intrinsically damaging exploratory policies. Rajeswaran et al. (2017) proposed an algorithm which allows learning a robust policy using an ensemble of simulated systems. The policy can be directly applied to a real system. Please, read the paper and answer the following questions.

1. What does generalization mean in the machine learning context?
2. Explain one of the batch policy optimization algorithms used in Algorithm 1.
3. How does it happen that learned policy is also robust to the effects not modeled in the source domain?
4. List and explain the drawbacks of the approach. Can you name more drawbacks then those mentioned in the conclusion of the paper?

References

Rajeswaran, A., Ghotra, S., Ravindran, B., and Levine, S. (2017). EPOpt: Learning robust neural network policies using model ensembles. In *International Conference on Learning Representations*.