EE239 Reinforcement Learning Project Proposal

Chenyang Wang
Department of ECE
UCLA
105425757

Gongjie Qi
Department of ECE
UCLA
805429380

Hanqing Wu
Department of ECE
UCLA
005429548

Xiao Zeng
Department of ECE
UCLA
505030377

Abstract

In this proposal we explain our interest in reproducing the algorithms proposed in the paper DeepMinic by Peng et al. [1]. We also illustrate the methods that we plan to use to implement the algorithms, how we improve upon the algorithm proposed and how we evaluate our results. In addition, we include the list of relevant related work that we are examining to provide better context for the results later.

1 Scope of work

In this section the essence of the proposed work is described by answering three key questions.

1.1 What is the problem that you will be investigating? Why is it interesting? What data, simulator or real-world RL domain will you be looking at?

In this project, we plan to study the paper DeepMimic: Example-Guided Deep Reinforcement Learning of Physics-Based Character Skills, Peng et al, 2018, and implement the algorithms proposed in this paper and test them on simple environments. In particular, we think this paper is interesting because it tries to use reinforcement learning techniques in the field of modeling the motions of humans and animals. Currently, manually designed controllers show satisfying results in reproducing certain motions, however a main problem here is that these controllers are not generalizable, i.e. they cannot handle new skills and new situations correctly. Another major obstacle of current physics-based human and animal motion simulation approaches is the directability. It is still extremely challenging for users to deliberately educe some desirable behaviors from simulated characters. Therefore, it is really interesting to see how reinforcement learning concepts can be used to solve these kinds of problems. What's more, modeling the motions can be used in many fields from biomechanics to robotics and animation. Thus we think this paper is worth investigating.

1.2 What method, algorithm or theoretical analysis are you proposing?

According to the description of DeepMimic model, we will synthesize a motion controller that takes a character model, a corresponding set of kinematic reference motions and a reward function representing a specific task as inputs, and computes a control policy that enables the character to imitate the reference motions and achieve the task goal using neural networks with proximal policy optimization (PPO) algorithm. To modify the implementation, we hope to apply the framework to a 3D simulated character other than the humanoid model widely demonstrated in the paper. The

Preprint. Under review.

hyper-parameters may also be re-optimized for better results based on our experiments on the new character.

1.3 How will you evaluate your results?

For the evaluation part, we plan to use the normalized return method together with the learning curve plots used in the paper. We will also investigate on more novel evaluation methods as we progress along this project.

2 Related literature

The related literatures cited in the Deepmimic paper can help us to build a fundamental understanding of the background. For instance, Benchmarking Deep Reinforcement Learning for Continuous Control by Duan et al. [2]. introduces many existing reinforcement learning algorithms for various continuous control tasks, and Proximal Policy Optimization Algorithms by Schulman, et al. [3] proposed the PPO algorithm used by DeepMimic. We plan to examine these essential works to obtain enough knowledge for our project.

References

[1] Peng, X.B., Abbeel, P.,Levine,S. & Panne, M. (2018) DeepMimic: Example-Guided Deep Reinforcement Learning of Physics-Based Character Skills. *ACM Trans*.

[2] Duan, Y., Chen, X.,Houthooft,R.,Schulman,J. & Abbeel, P. (2016) Benchmarking Deep Reinforcement Learning for Continuous Control. *OpenAI*.

[3] Schulman, J., Wolski, F.,Dhariwal,P.,Radford,A. & Klimov, O. (2017) Proximal Policy Optimization Algorithms. *OpenAI*.