Learning to Walk

Akhil Nagariya, Vishala Arya, Sameer

October 2019

1 Introduction

2 Related Work

[?] model the statics of tensegrity robots, [?] used this model to come up with a kinematic controller [?] uses Monte Carlo simulations to learn goal directed behaviors requiring thousands of trials making this approach ineffective for real systems. [?] transfers policies learned in simulation to real systems but requires hand tuning. [?] exploited symmetry of a simulated super ball to learn an efficient rolling controller. Classical approaches[] for bipedal locomotion uses analytical model for dynamics and hand engineered control policies and difficult to generalize for tensegrity robots which are highly underactuated with coupled dynamics. Direct policy search, searches in the space of policies and can effectively scale to complex high dimensional systems but require large number of samples and often get stuck in poor local minima. [?] Guided policy search[1] uses differential dynamic programming to generate suitable guiding samples and uses importance sampling to incorporate these samples directly into the policy. Learning preodic gaits are a key problem in locomotion tasks, [?] solves this problem by extending MDGPS[?] to learn a controller for the SUPERball tensegrity structure by sequentially training policies for a perodic behaviour starting from different initial states and then combining them to perform continous locomotion for wide range of condions. [?] introduces noise in simulation to train policies that are successful in a wide range of simulated terrains.

3 Proposed Approach

- planning to use [?]

[1] [2] [3] [4] [5] [6]

References

[1] S. Levine and V. Koltun, "Guided policy search," in *Proceedings of the 30th International Conference on International Conference on Machine Learning - Volume 28*, ser. ICML'13. JMLR.org, 2013, pp. III–1–III–9. [Online]. Available: http://dl.acm.org/citation.cfm?id=3042817.3042937

[2] ——, "Variational policy search via trajectory optimization," in *Advances in Neural Information Processing Systems 26*, C. J. C. Burges, L. Bottou, M. Welling, Z. Ghahramani, and K. Q. Weinberger, Eds. Curran Associates, Inc., 2013, pp. 207–215. [Online]. Available: http://papers.nips.cc/paper/5178-variational-policy-search-via-trajectory-optimization.pdf

[3] ——, "Learning complex neural network policies with trajectory optimization," in *ICML '14: Proceedings of the 31st International Conference on Machine Learning*, 2014.

[4] S. Levine, C. Finn, T. Darrell, and P. Abbeel, "End-to-end training of deep visuomotor policies," *CoRR*, vol. abs/1504.00702, 2015. [Online]. Available: http://arxiv.org/abs/1504.00702

[5] W. Montgomery and S. Levine, "Guided policy search as approximate mirror descent," *CoRR*, vol. abs/1607.04614, 2016. [Online]. Available: http://arxiv.org/abs/1607.04614

[6] X. Geng, M. Zhang, J. Bruce, K. Caluwaerts, M. Vespignani, V. SunSpiral, P. Abbeel, and S. Levine, "Deep reinforcement learning for tensegrity robot locomotion," *CoRR*, vol. abs/1609.09049, 2016. [Online]. Available: http://arxiv.org/abs/1609.09049