

強化学習による四足歩行ロボットの歩行運動の学習

～ PPO による方策勾配法～

東京都立大泉高等学校 データサイエンス機 鈴木權人

Abstract

The objective of this research is to obtain efficient walking motions for quadruped robots through reinforcement learning using computer simulations. The significance of this research is that it will enable the development of inexpensive quadruped robots, which will lead to their widespread use. In this study, PPO (Proximal Policy Optimization) was used as the algorithm, but parallelization could not be implemented. Future prospects include implementation in other languages.

研究背景・目的・意義

現状：昨今の無人口ボットの
現場における活躍



目的：強化学習による構築の容易化



意義：さらなる現場への普及

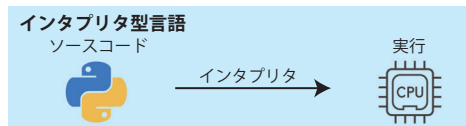
結果・考察

PPO 実装に失敗

原因：Trajectory 収集時の
並列化エージェント処理でのエラー



Python はインタプリタ言語であり、
並列処理が苦手である



研究方法

ロボットモデル：Spot micro (図1)

実装言語：Python

強化学習ライブラリ：TensorFlow

物理シミュレーションエンジン：PyBullet

強化学習アルゴリズム：PPO

$$L^{CPI}(\theta) = \hat{\mathbb{E}}_t \left[\frac{\pi_{\theta}(a_t|s_t)}{\pi_{\theta_{old}}(a_t|s_t)} \hat{A}_t \right] = \hat{\mathbb{E}}_t [r_t(\theta) \hat{A}_t]$$

$$L^{CLIP}(\theta) = \hat{\mathbb{E}}_t \left[\min(r_t(\theta) \hat{A}_t, \text{clip}(r_t(\theta), 1 - \epsilon, 1 + \epsilon) \hat{A}_t) \right]$$

PPO モデル式

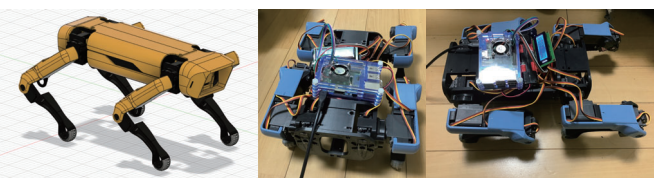


図1 Spot micro

実際に作成した Spot micro のモデル (試作)

結論・展望

Python 以外の言語での実装→エラー回避

ex) C, Java, Go



様々なシーンへのロボットの普及



引用・参考文献

- [1] John Schulman, Sergey Levine, Philipp Moritz, Michael I. Jordan, Pieter Abbeel, Trust Region Policy Optimization, 2017b.
- [2] Horia Mania, Aurelia Guy, Benjamin Recht, Simple random search provides a competitive approach to reinforcement learning, 2018.
- [3] nicrusso, 2021 rex-gym, [Source code], <https://github.com/nicrusso7/rex-gym>.
- [4] OpenQuadruped, 2020, spot mini mini, [Source code], <https://github.com/OpenQuadruped/spot mini mini>.
- [5] めんだこ, "ハムスターでもわかる Proximal Policy Optimization (PPO) ①基本編", どこから見てもメンダコ, 2020-10-18, <https://horomary.hatenablog.com/entry/2020/10/18/225833>, (2022-12-18).
- [6] Tensorflow, "Actor-LearnerAPI を使用した SAC ミニチュアロボット", Agents, 2021, https://www.tensorflow.org/agents/tutorials/7_SAC_miniaur_tutorial?hl=ja, (2022-12-18).
- [7] 齊藤康毅, ゼロから作る Deep Learning, オライリー・ジャパン, 2016
- [8] 齊藤康毅, ゼロから作る Deep Learning ④ — 強化学習, オライリー・ジャパン, 2022