SC42050 Literature Assignment
Reinforcement Learning for legged robots

Gabriel Lopes

Legged robots require completely different and far more complex modes of actuation then their wheeled counterparts. However, many advantages can be gained from the extra degrees of freedom legs provide. (Kohl and Stone, 2004) take a reinforcement learning approach to optimize the forward trot gait of the commercial robot Aibo. Please read (Kohl and Stone, 2004) carefully and and answer the following questions:

1. Do a small literature review and present the basic concepts of Reinforcement Learning, including the update rules for the Temporal Difference TD(0) algorithm and its gradient descent implementation for parameterized value functions V^{π}.

2. Explain what the authors mean by the phrase at the end of page 2:   *"(the algorithm presented here) can be considered a degenerate form of standard policy gradient reinforcement learning techniques"*. In what way? Please explain by illustrating the differences or similarities in the update rules.

3. Explain the differences or similarities between the algorithm proposed in the paper and the Nelder-Mead algorithm. Please illustrate graphically by drawing the update after some number of samples are collected.

4. Try to enumerate the problems one might face when addressing the gait improvement problem using the standard reinforcement learning techniques for continuous space states.

References

Kohl, N. and Stone, P. (2004). Policy gradient reinforcement learning for fast quadrupedal locomotion. In *Proceedings of the IEEE International Conference on Robotics and Automation*, volume 3, pages 2619–2624.