C.R.A.W.LAB
http://www.ucs.louisiana.edu/~jev9637/

Learning Control Strategies and Design Parameters for Flexible-legged Locomotive Systems

Andrew Albright, Joshua Vaughan

andrew.albright1@louisiana.edu

UNIVERSITY OF LOUISIANA Lafayette. Mechanical Engineering

If you include a figure, you need to explain and reference it in the text

Flexible Robotics

Figure 1: Boston Dynamics Atlas Robot [1]

Legged systems have many advantages ~~when~~ compared to their wheeled counterparts. For example, if designed properly, they can more easily navigate uneven and unstable terrain. However, there are disadvantages as well, including dramatically less energy efficient locomotion.

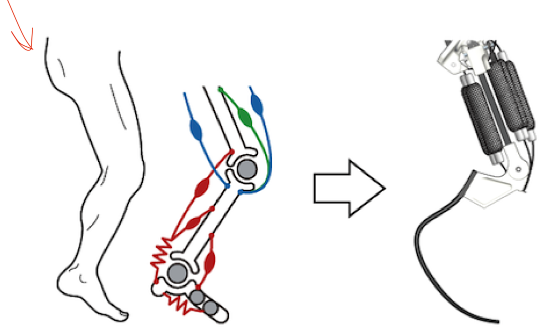

Figure 2: Athlete Robot [2]

In an effort to mitigate this disadvantage, research has been conducted that shows using flexible components in legged locomotive systems not only increases their efficiency but also their performance [3]. However, nonlinear models and controllers for flexible systems are difficult to develop using traditional methods.

Reinforcement Learning

A solutions to combat the nonlinear difficulties is Reinforcement Learning (RL) which uses neural networks to represent the nonlinear aspects of the control architecture. The RL approach does not require closed-form term models of the system dynamics to develop a controller. Rather, the approach creates a machine learning-based agent to interact with an environment collecting samples of rewards and actions that get mapped to system states. With this map, the agent defines inputs for the system based in its state. Previous work has shown that this method can be successful at defining both mechanical parameters and control strategies for rigid systems [4].

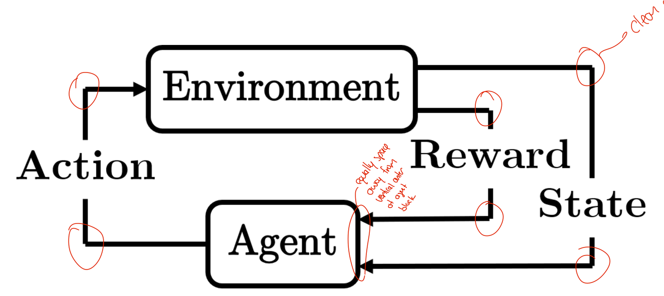

Figure 3: RL Block Diagram

Reinforcement Learning for Flexible Robotics

In this work, Reinforcement Learning is used to develop optimal control strategies leading to higher performance both in desired movement and energy efficiency for a flexible system. The pogo stick, shown in Figure 4 is used to train an agent to determine the acceleration of the mass, m_a, along the rod, m_l, causing system to jump. The agent is tasked to learn a control strategy that jumps as high as possible in a single jump while also limiting power consumption. Two training strategies are deployed to accomplish this task. In the first strategy, the agent is rewarded based on the height it achieves receiving a score directly correlated to the height ~~it reaches~~. For the second strategy, the agent is rewarded using a normalized positive score correlated to jump height, and punished a normalized negative score proportional to power use ~~for using power~~. ~~The reward function is tuned for training, allowing the agent to accomplish the specified tasks.~~ The agent is able to learn both optimal jumping height strategies as well as power conserving strategies.

Doesn't add anything to reader's understanding.

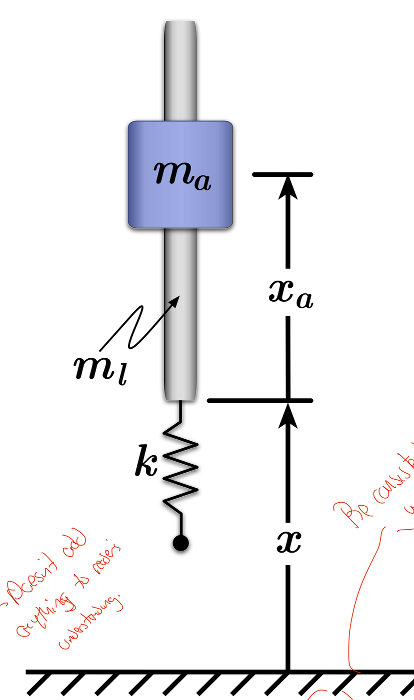

Figure 4: Pogostick Model

Be consistent with spacing, etc

Jumping Analysis

We need some metric for energy efficiency. There are some for walking. I but some exist for jumping too. Energy use per m jump height or similar. Not sure we need for this poster, but will eventually

~~Time series jumping results and power usage data are collected during an evaluation study after the agent is trained.~~ Figure 5 shows the jumping profiles for the individually trained agents. It is apparent that when training the agent to jump high regardless of power usage, the maximum jump height reached is higher than when training the agent to conserve power. However, the power consumption data shown in Figure 6 shows that the power conserving agent is more energy efficient, using 64% less power than the agent which disregards power usage. ~~Because the difference in jump height is a lesser 44%, it is concluded that the~~ The agent has learned a more efficient control strategy.

Defining The Reward

Two reward functions are defined for training the two different agents; the first rewards the agent based on the height it reaches; the second seeks to balance reward height with power usage. The reward function ~~is expressed as the equation~~:

$$R = \frac{\frac{x_t - x_{min}}{x_{max} - x_{min}} \omega_x + m_a \frac{a_t v_t - a_{max} v_{max}}{a_{min} v_{min} - a_{max} v_{max}} \omega_p - \omega_p}{\omega_x + \omega_p} \quad (1)$$

Constants/weights usually go before the term that they apply to

where x_t, a_t and v_t are position, velocity and acceleration of mass m_a respectively, ω_x and ω_p are weights used to tune ~~how much to conserve power~~.

Evaluation Data

Units are most usually italicized... change all

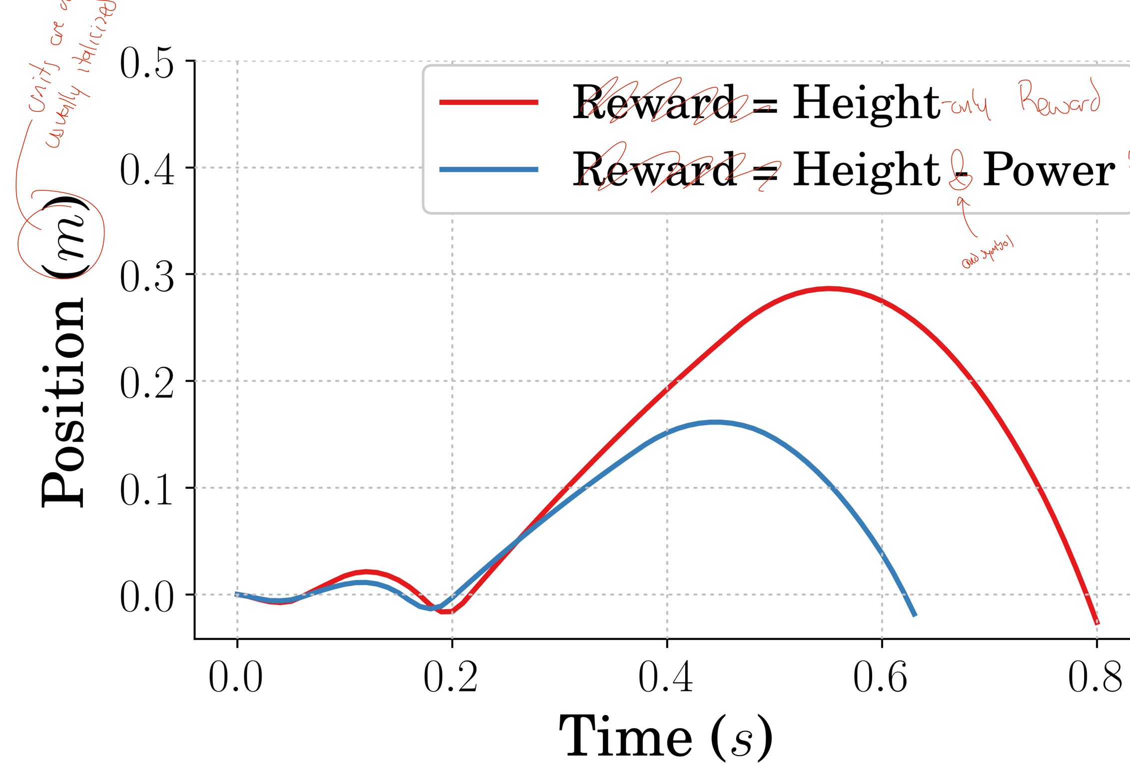

Figure 5: Jumping Height Comparison

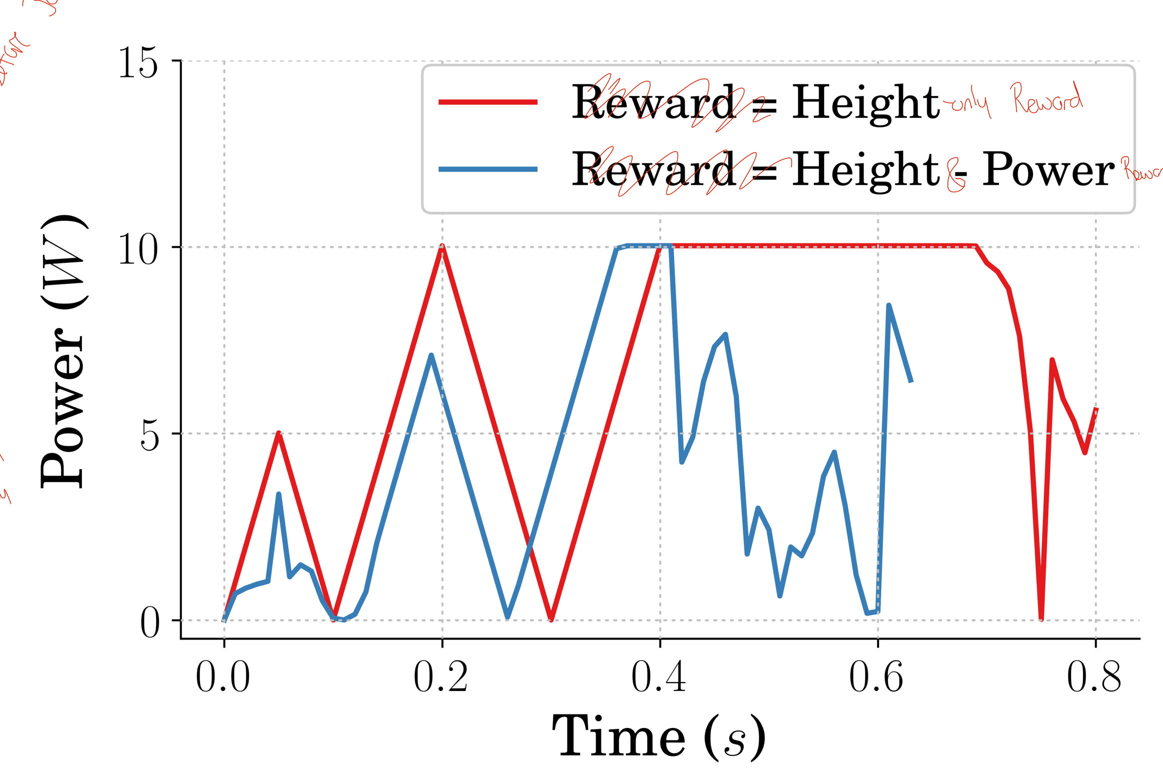

Figure 6: Power Usage Comparison

Implications of Results

The jumping profiles presented in Figure 5 show that the agent is controlling the pogo stick to jump with a profile matching what would be called a stutter jump. ~~This can be visualized~~ in Figure 7. ~~It has been studied that~~ this kind of jump is an optimal jump for reaching maximum heights [5]. Therefore, it is concluded that a reinforcement learning agent can learn optimal jumping control strategies while minimizing power consumption.

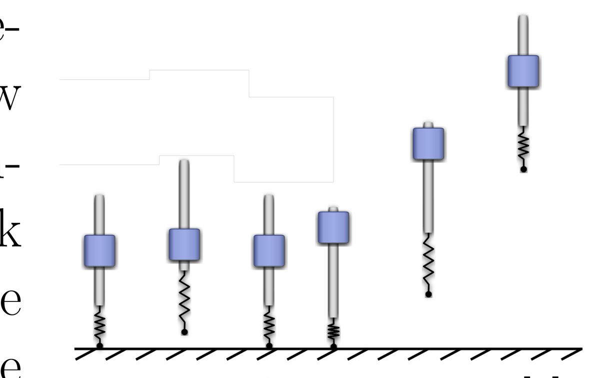

Figure 7: Stutter Jump [5]

Conclusion

Using a reinforcement learning approach to define control strategies for flexible jumping systems leads to optimal performance. Using a correctly tuned reward function can further improve performance, leading to control strategies that require less power.

References

[1] Boston Dynamics.
Atlas® boston dynamics, 2021.

[2] designboom.
ryuma niiyama athlete robot, 2010.

[3] Yuuta Sugiyama and Shinichi Hirai.
Crawling and jumping of deformable soft robot.
2004 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 4(c):3276–3281, 2004.

[4] David Ha.
Reinforcement learning for improving agent design.
Artificial Life, 25(4):352–365, 2019.

[5] Joshua Vaughan.
Jumping Commands For Flexible-Legged Robots. 2013.

Acknowledgements

Special thanks to the Louisiana Crawfish Promotion and Research Board for their support.