Authors: Andrew Albright, Joshua Vaughan

Date: 03/26/2021

Due Date: 04/09/2021

Try to distill this to 1-2 sentences

This is just a summary of what's in the paper

Main Takeaway / Abstract

Legged locomotive systems have many advantages over their wheeled counterparts such as their ability to navigate rough terrain. They have the ability to deploy many navigating techniques to overcome obstacles, one of which is jumping. Still, there are disadvantages to overcome when using legged systems; an example is their lack of energy efficiency. To combat this apparent lack of efficiency, flexible links can be used to conserve energy that would otherwise be wasted during locomotion. Using Reinforcement Learning to create controllers for flexible legged jumping systems can lead to designs that outperform traditional optimization methods for jumping. Further, using a power-conservative reward function to train the agent results in a control strategy that seeks to maximize jump height while also conserving power.

Introduction

- Discussion on the pros and cons of using legged systems for tasks that require locomotion
    - Legged locomotive robots have many advantages over their wheeled counterparts, for example their ability to more easily navigate harshly uneven terrain. However, there are disadvantages as well, one of which is power consumption. \cite{Altendorfer2001}, \cite{Galloway2011}
    -

  what's the main point to make about these?

- Discussion on the potential of using flexible links to replace rigid ones to increase performance in terms of accomplishing the goal and in conserving power.
    - In an effort to alleviate the power consumption issues seen when using legged systems to accomplish locomotive tasks such as walking, running and jumping, research has been conducted which replaces rigid aspects of said systems with flexible ones. It has been shown that this not only leads to higher performance but also higher efficiency. \cite{Sugiyama2004a}

      Is higher efficiency a type of higher performance?

- Discussion on using RL for controller design for creating higher jumping robots that also conserve power.
    - In this work, RL is used to train an agent (controller) which seeks to jump a simplified jumping robot modeled as a pogo stick. The RL agent is tasked with maximizing jump height while conserving power. It is shown that when tasked as such, the agent finds unique control strategies to maximize the jumping potential of the system, as well as balancing power usage.
- Road map for the rest of the paper
    - discuss related work in the field
    - detail the method used to test the idea
        - system
        - reward
        - system used to compare (DV's)
    - discuss results
    - conclude the results and discuss future work

Related Work

- Discussion on using RL to create controller to control legged locomotive systems \cite{Fankhauser2013}
    - Research has shown that using RL for defining control strategies for legged systems is a viable path for simple and even complex legged locomotive systems. \cite{Yang2019}, \cite{Ha2018a}, \cite{Reda2020c}, \cite{Da2020c}
- Discussion on using RL to create controllers for flexible systems both legged locomotive and non locomotive type \cite{Thuruthelb}, \cite{Ghorbel1990}, \cite{Dwiel2019d}, \cite{He2020f}, \cite{Bhagat2019e}, \cite{Pradhan2012e}, \cite{Cui2019e}

  conducted

    - Research has been conduction which shows the potential of using an RL approach to define controllers for flexible systems which are to have high accuracy. Additionally, the use of RL for defining controllers for flexible legged locomotion is also an emerging study, showing the possibilty and viability of the direction.
- Discussion on flexible systems, legged locomotive or not, being used to be more power conservative \cite{Harper2019}, \cite{Seok2015}, \cite{Seok2013}, \cite{Sugiyama2004}, \cite{Folkertsma2012}
    - Using flexible components within robotic systems has shown great potential for conserving power. Using both flexible links and joins has shown that a system which in less rigid has the ability to more energy efficient. Further more using flexible components in legged locomotive systems has shown that higher performance can be achieved when looking at metrics like running speed and jumping height.
- Discussion on traditional control methods which seek to be more power efficient \cite{Pace2017}, \cite{Harper2019}, \cite{Harper2017}
    - Research has been completed showing the effects of more traditional methods for creating more energy efficient control methods.

Too generic

~~Method~~ Pogo-stick Environment

A. Model and Constraints

A simplified, pogo-stick model is used...

subsection

- Defining pogo stick environment

B. Reward Func.

- Define the reward function used for different training sessions
    - Reward strictly based on jump height
        - this will lead to maximum
    - Reward based on jump height and power usage
- Using RL to develop a controller for the pogo stick env
- Defining different types of reward functions
    - Maximizing jump height
    - Maximizing jump height punishing power conservation
    - Maximizing jump height punishing power conservation past 50% of max jump height

Separate sub-section

Results

- Analyze jump height results form all three agents
- Analyze power consumption results from all three agent
- Make comparisons of jump height reached to power consumed
- Compare results to DV's paper

Conclusions

- Using RL leads higher jumping robot control designs
- Using RL also leads to more energy efficient controller designs

- Mention future work giving RL access to the design parameters to further optimize jump height and power consumption

List of Needed Figures

- Table describing the pogo stick's attributes (masses, spring k, etc.)
- Jumping height from DV's paper
- Jumping height from agent which is maximizing jump height
  - Compared to DV's paper
- Jumping height from power saving agent
  - Compared to other agent and DV's paper
- Power consumption from power saving agent
  - Compared to other agent and DV's paper

- Sketch of model
- reward as function of training time for each?