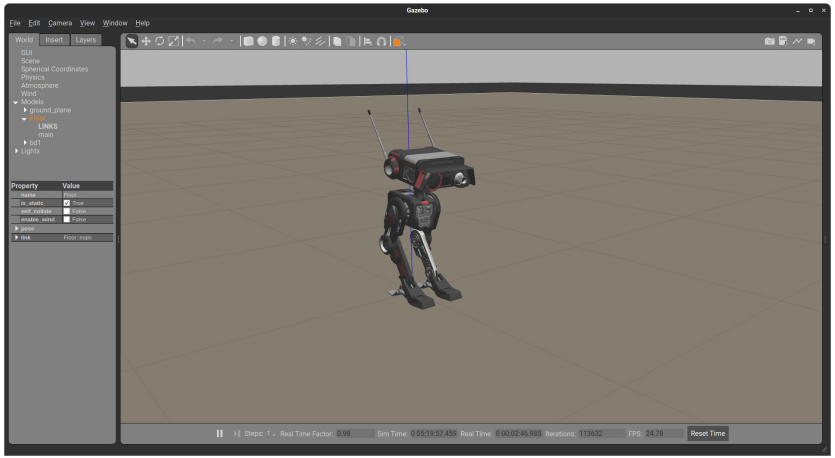
# Bringing Star Wars BD1 Droid to Life

## Trax Bagley

- Computer Science Undergraduate
- Software Engineer at Lightningkite

## Goal

- Move forward on level ground



# What

- Simulate a bipedal robot learning to moving across flat ground
- Using Temporal difference learning algorithm

# Why

- Better understanding of how to apply and implement temporal learning reinforcement learning algorithm
- Better able to move through environment as a human or other bipedal creatures
- Fun, Interesting and cool

# How

- Use ROS Gazebo environment
- Forked from https://github.com/petr-sorokoumov/BD1
- Use the temporal difference learning with a neural network
- Used Pytorch fully connected network
- Take random actions compare that to the result of the neural network

## Observation State

- Float[] of size 22
- Robot position in environment
  - Position [x:float,y:float,z:float
  - Velocity [x:float,y:float,z:float]
- Fall status Boolean
  - Robot position in env Z gets to low
- Joint name such as leg_mid_left- all available joints
  - Right_leg positions:[x:float,y:float,z:float] velocity:[x:float,y:float,z:float]
  - Left_leg positions:[x:float,y:float,z:float] velocity:[x:float,y:float]
  - Head positions:[x:float,y:float] velocity:[x:float,y:float]

## Action Space

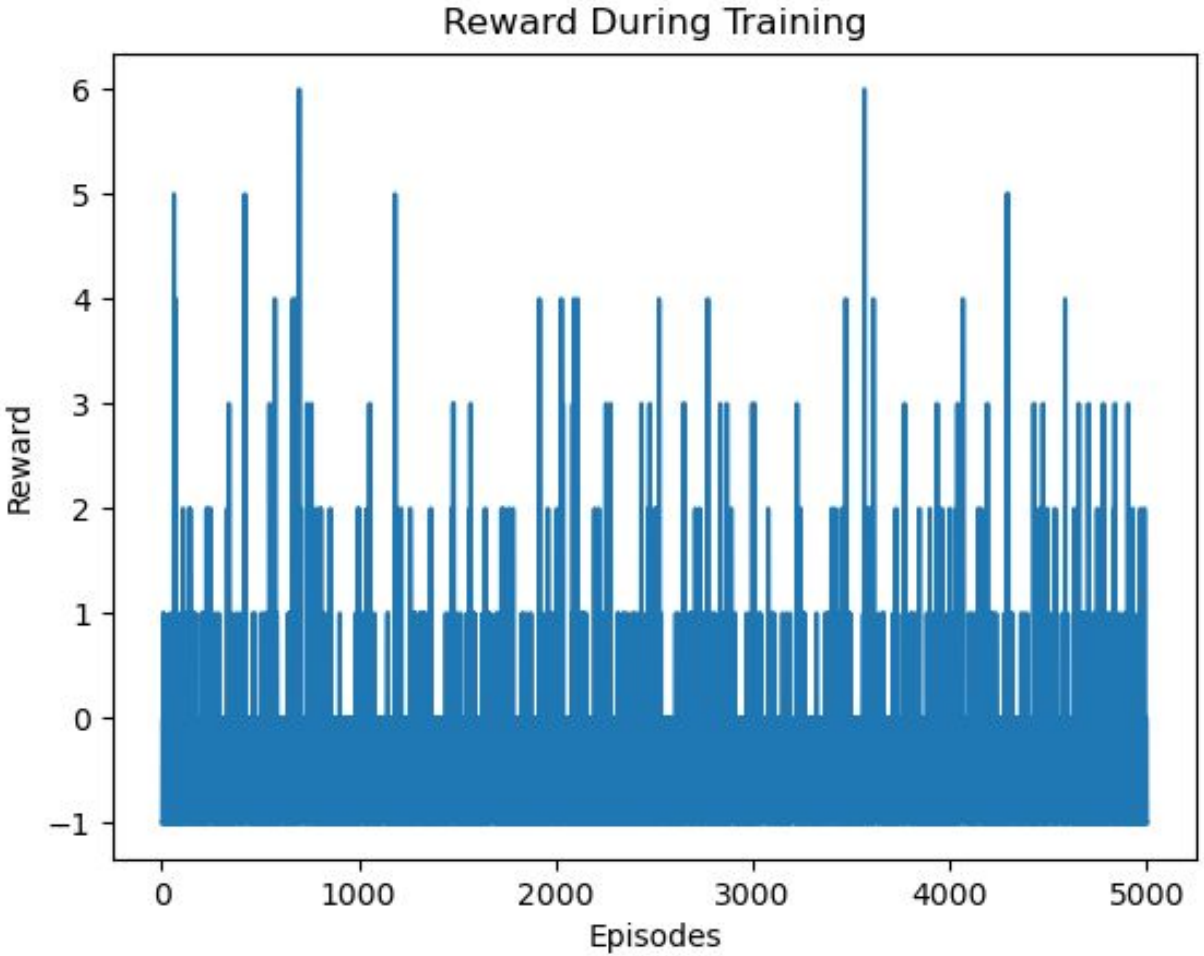
- Float[] of size 16
- Joint Trajectory object For all the joints
  - Leg_mid_left
    - [pos,vel]
  - Leg_mid_right
    - [pos,vel]
  - Leg_up_left
    - [pos,vel]
  - Leg_up_right
    - [pos,vel]
  - Leg_feet_left
    - [pos,vel]
  - Leg_feet_right
    - [pos,vel]
  - Neck_j
    - [pos,vel]
  - Head_j
    - [pos,vel]

## Reward Structure

- Distance on the x axis traveled
  - Positive +1 for every hyperparameter distance
  - Hyperparameter for minum standing height
- Fall
  - Negative reward for having too low of z model height -1

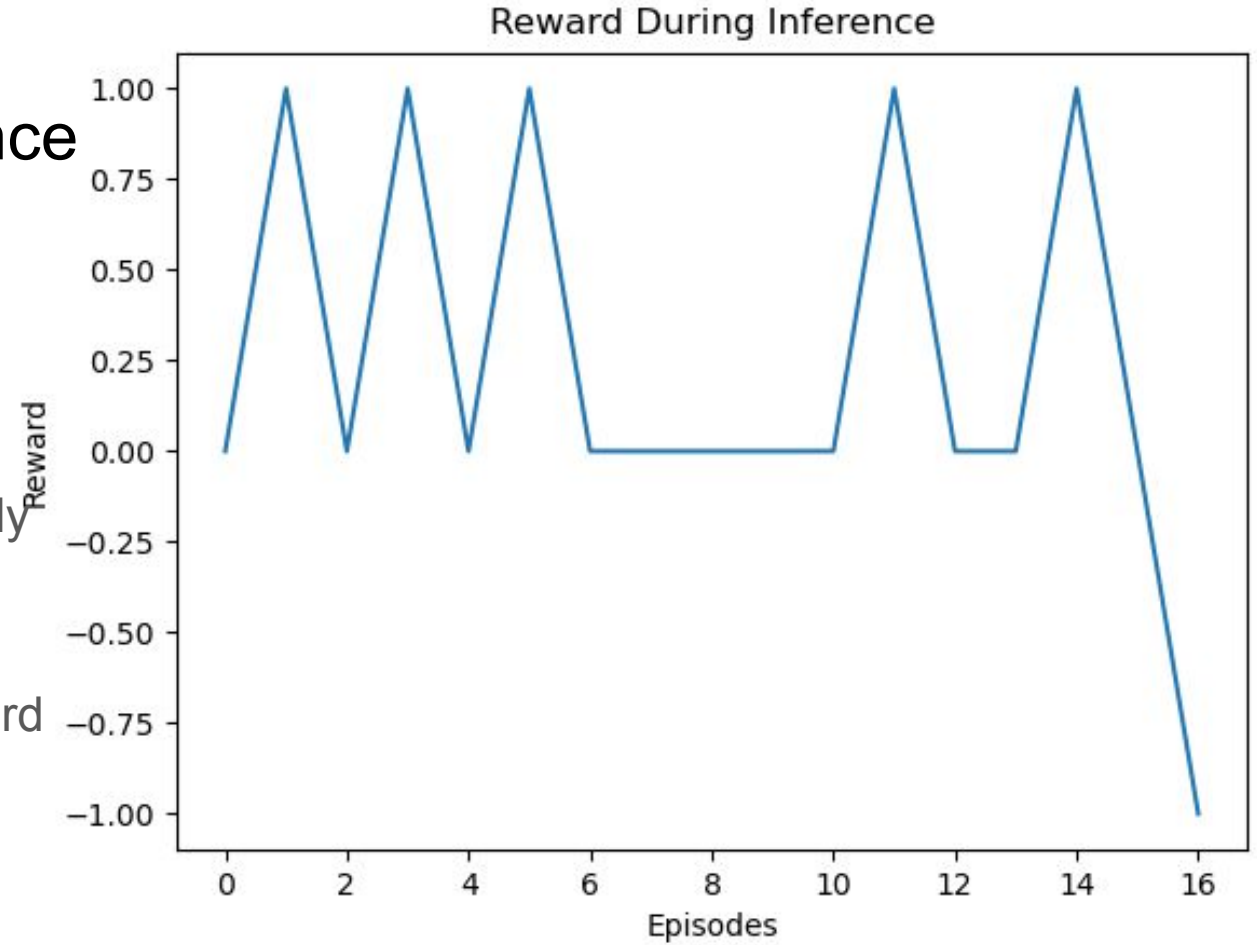
# Graph of the Cumulative Reward for each training episode results

- During training the rewards bounced up and down due to the randomness rate
- Certain times it would get high rewards
- The rewards get a little higher overtime as the rate of randomness goes down



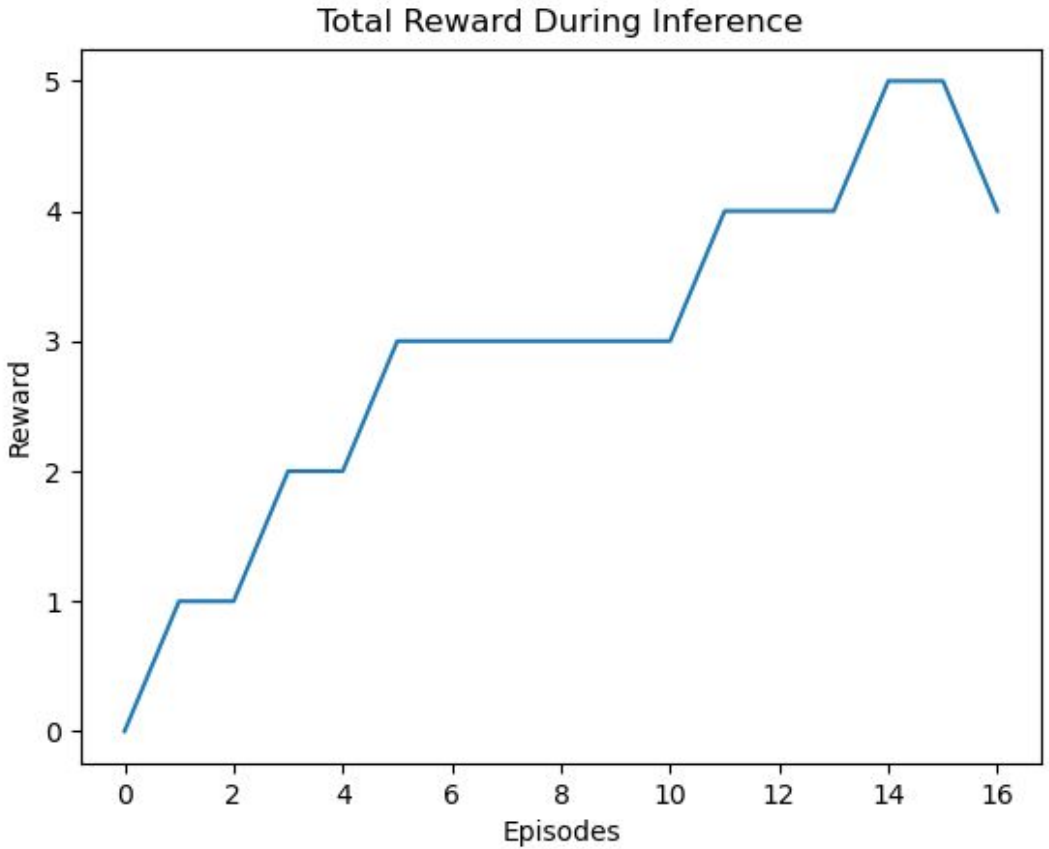
# Graph of the reward received at each episode During inference

- It consistently got 1 reward every two episodes
- Started off really good at the beginning
- Struggled toward the end though



# Cumulative reward during inference

- Got a cumulative of 5 rewards
- It did great at the beginning with receiving rewards
- Struggled in the middle



# Video Results and summary

- Learned to move by using the head to pull itself forward

Future work

- At hip joints to allow better balancing
- Add gyroscope to head and body to help with balance