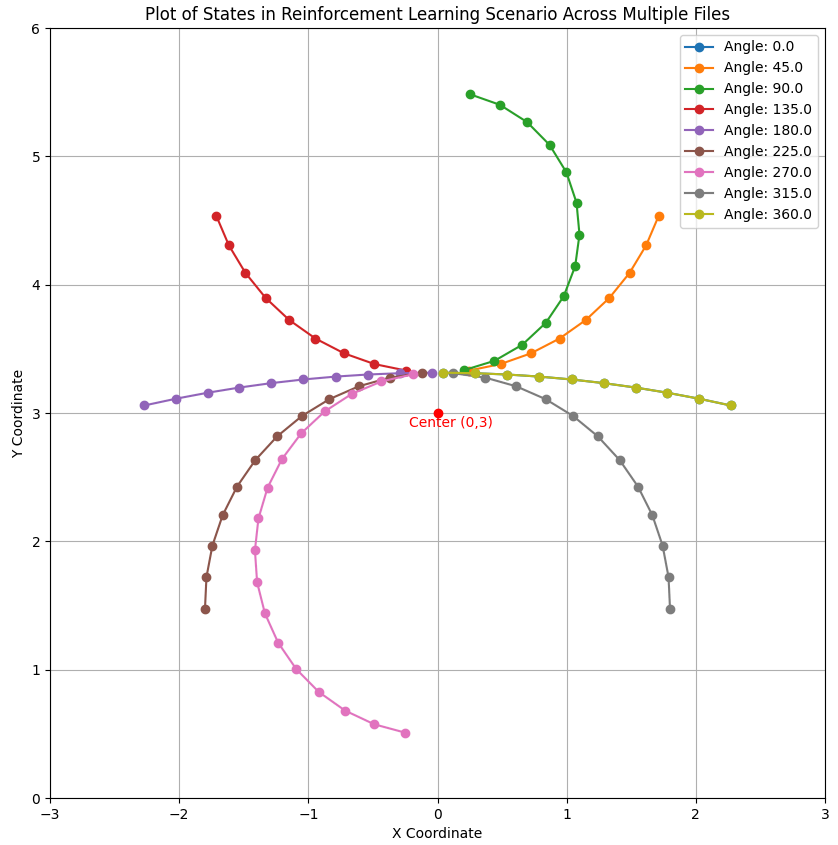
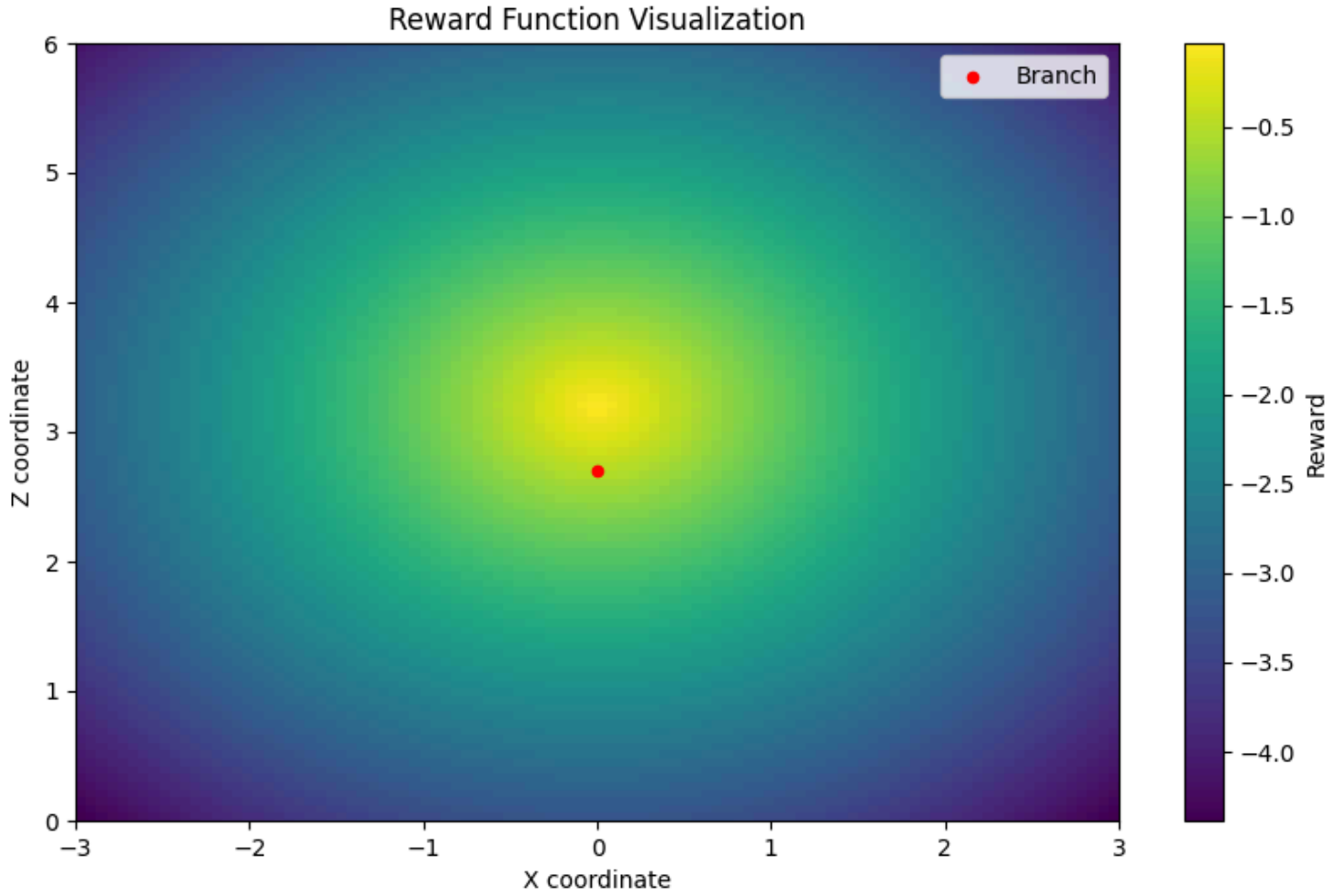
Agile Trajectory Generation for Tensile Perching with Aerial Robots

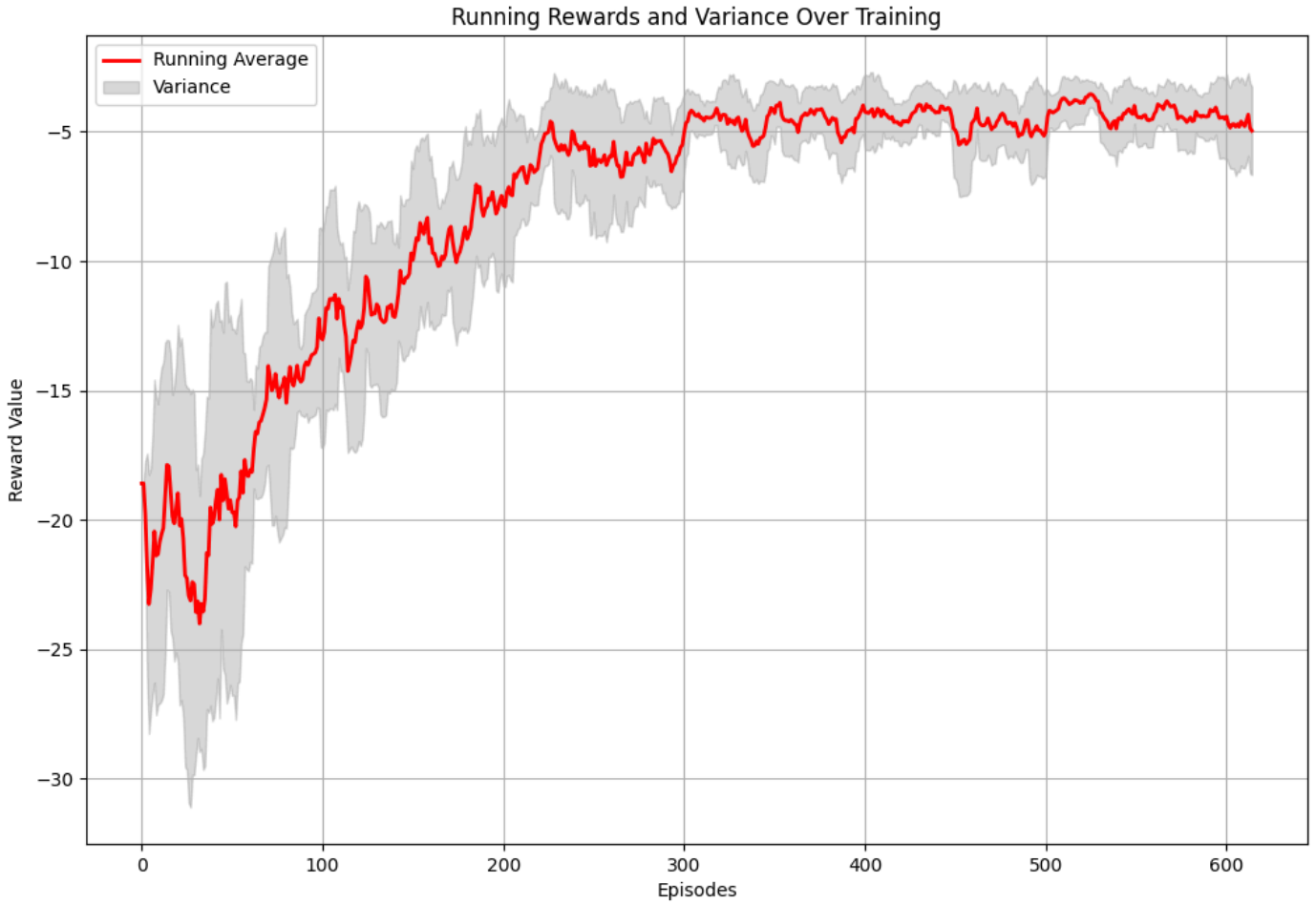
Progress Update

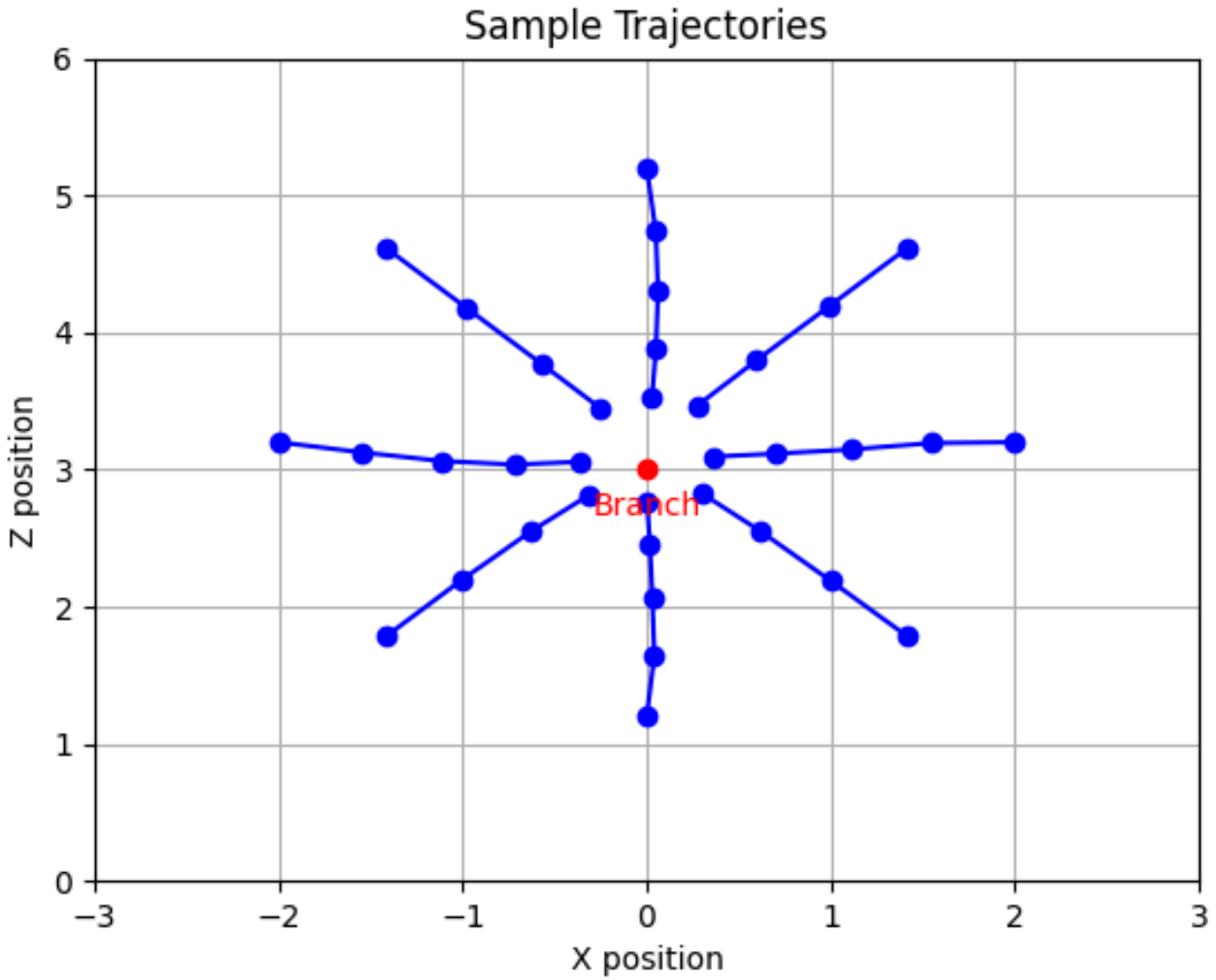
- Demonstrations
  - Demonstrations from both optimised trajectory (Hann) and Previous Work (Fabian).
  - Using these demonstrations in training.

Last Week


Plot of States in Reinforcement Learning Scenario Across Multiple Files
Angle: 0.0
Angle: 45.0
Angle: 90.0
Angle: 135.0
Angle: 180.0
Angle: 225.0
Angle: 270.0
Angle: 315.0
Angle: 360.0
Center (0,3)
X Coordinate
Y Coordinate
6
5
4
3
2
1
0
−3
−2
−1
0
1
2
3


Reward Function Visualization
Branch
Z coordinate
X coordinate
Reward
6
5
4
3
2
1
0
−3
−2
−1
0
1
2
3
−0.5
−1.0
−1.5
−2.0
−2.5
−3.0
−3.5
−4.0

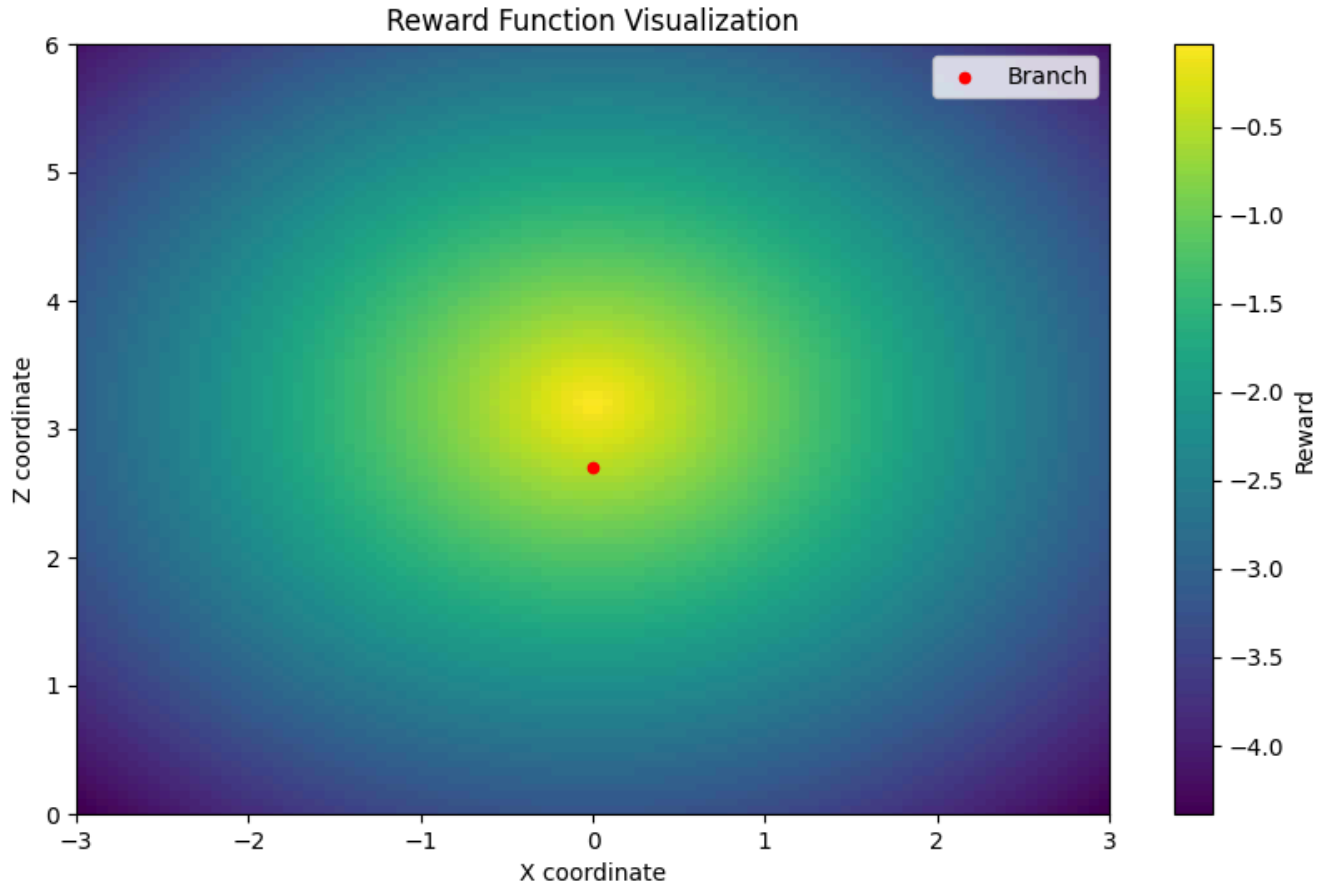

Running Rewards and Variance Over Training
Running Average
Variance
Reward Value
−5
−10
−15
−20
−25
−30
0
100
200
300
400
500
600
Episodes

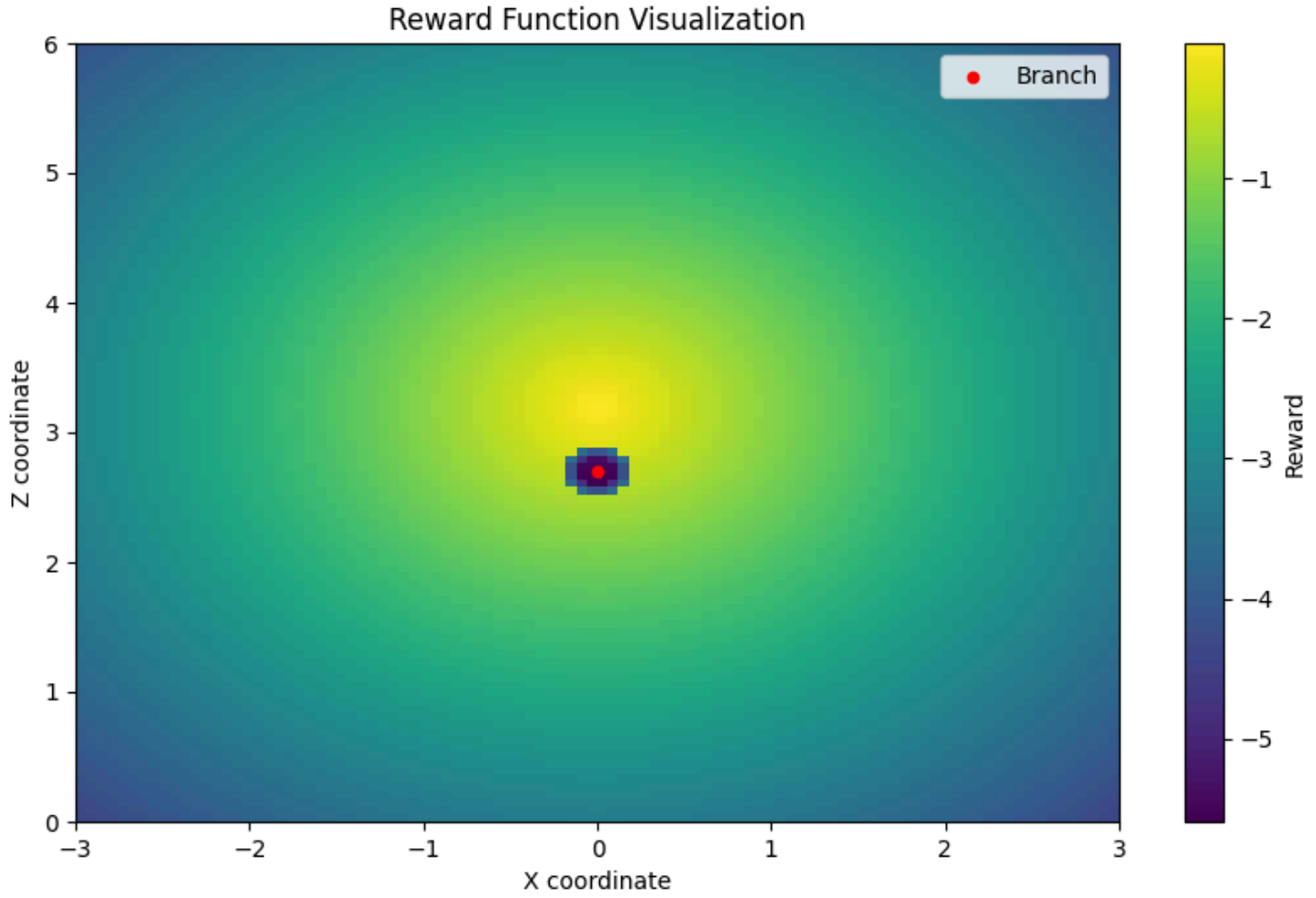

Sample Trajectories
Z position
X position
Branch
6
5
4
3
2
1
0
−3
−2
−1
0
1
2
3

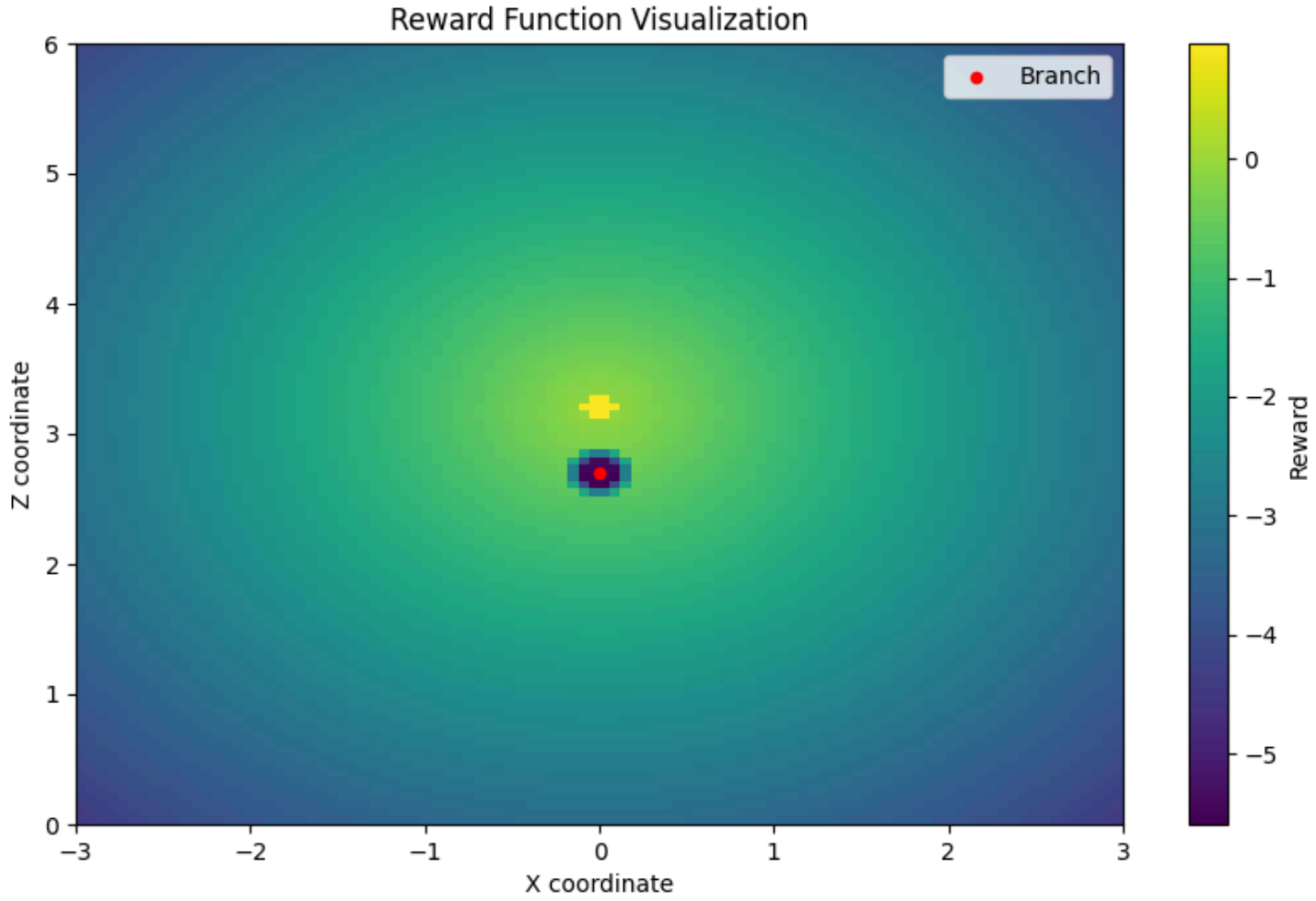
From Previous Week

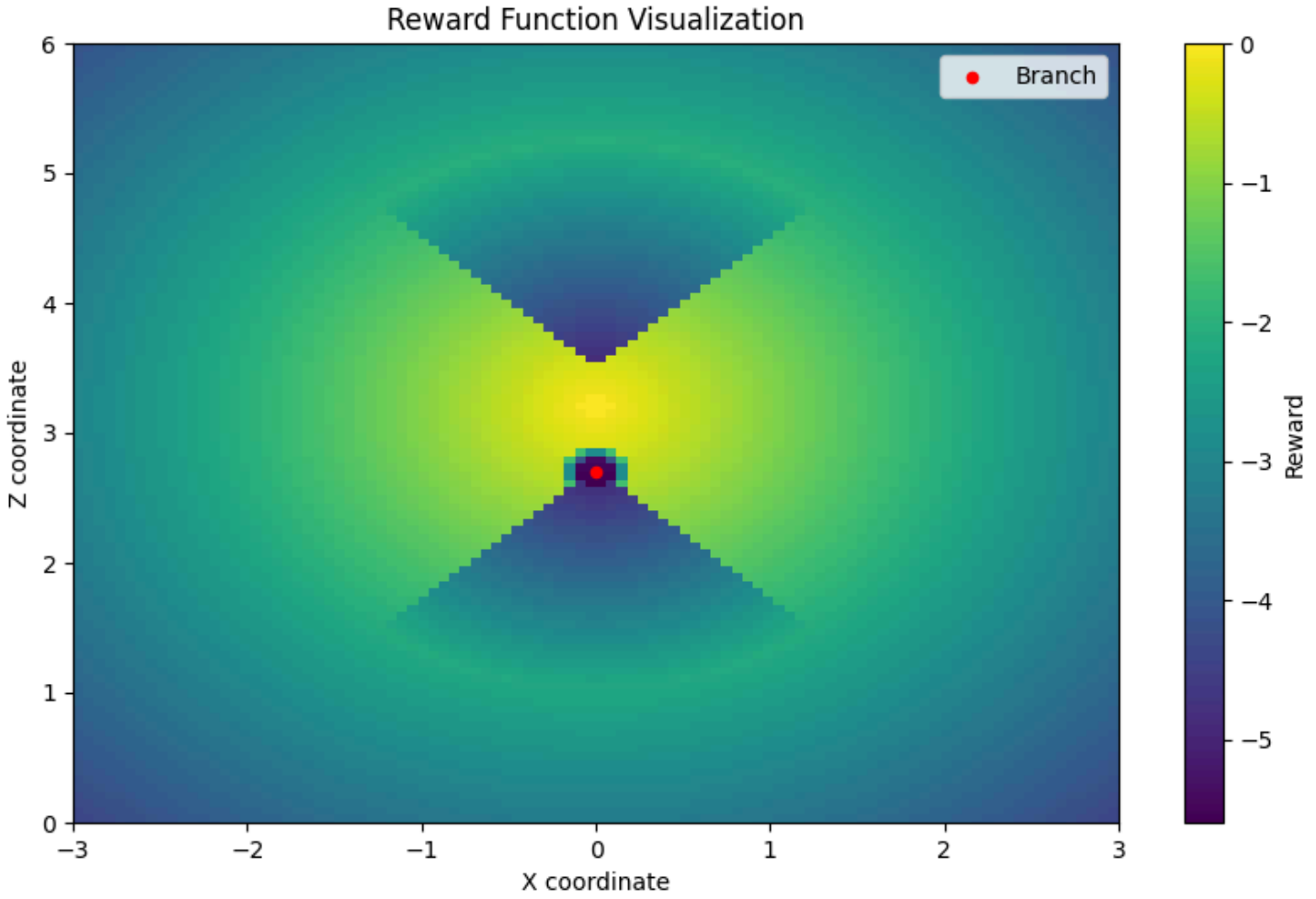
- Issues faced
  - Resetting - fixed an issue an issue with gravity in resetting.
  - Reward Function
    - done based on hitting tether to branch - stop condition is when there is a collision between the centre portion of the tether and the branch.
    - collision - penalise contact between drone and branch - large collision penalty.
    - Trajectory Smoothness Term
- Learning from Demonstrations System
- By next week: Approaching stage finished with comparison results between SAC and SACfD.

Reward

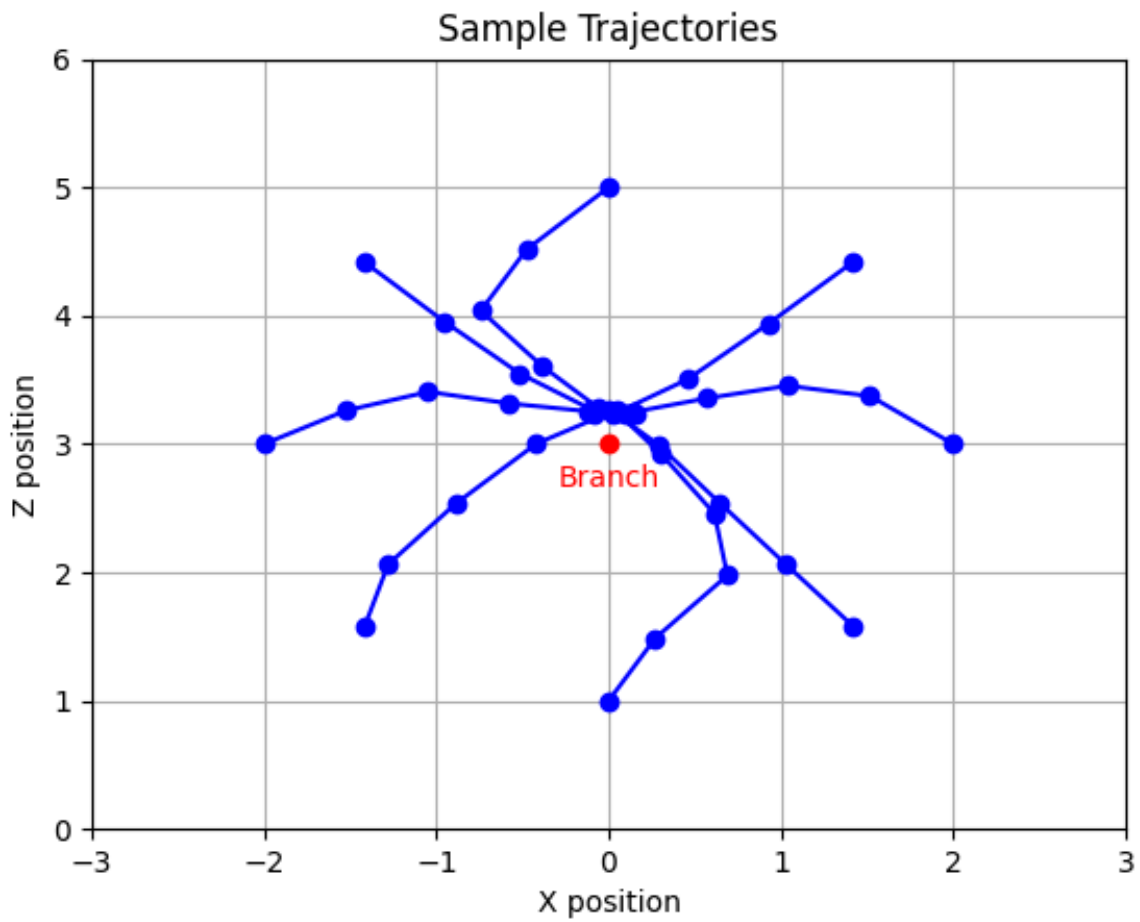

Reward Function Visualization
Branch
Z coordinate
X coordinate
Reward
6
5
4
3
2
1
0
−3
−2
−1
0
1
2
3
−0.5
−1.0
−1.5
−2.0
−2.5
−3.0
−3.5
−4.0


Reward Function Visualization
Branch
Z coordinate
X coordinate
Reward
6
5
4
3
2
1
0
−3
−2
−1
0
1
2
3
−1
−2
−3
−4
−5

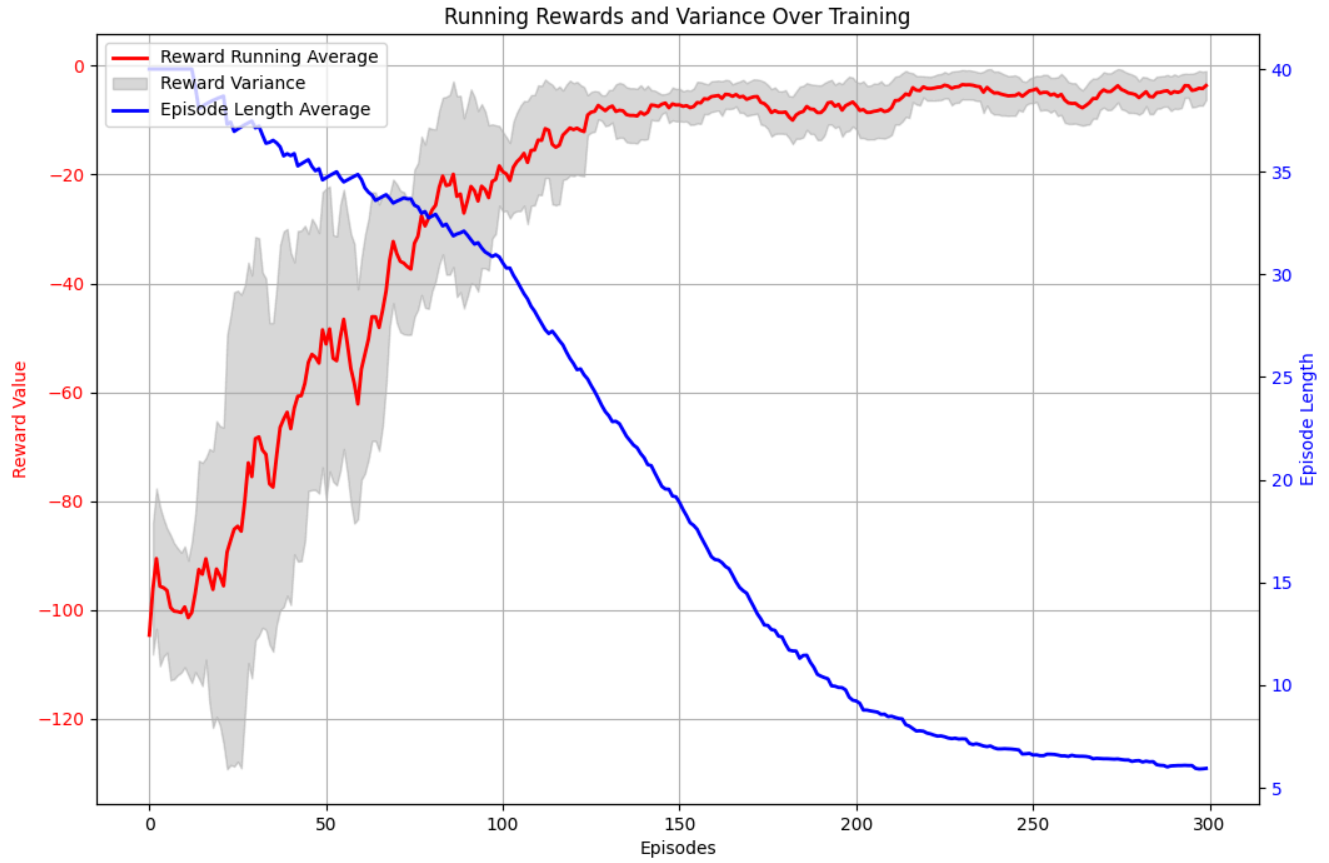

Reward Function Visualization
Branch
Z coordinate
X coordinate
Reward
6
5
4
3
2
1
0
−3
−2
−1
0
1
2
3
0
−1
−2
−3
−4
−5

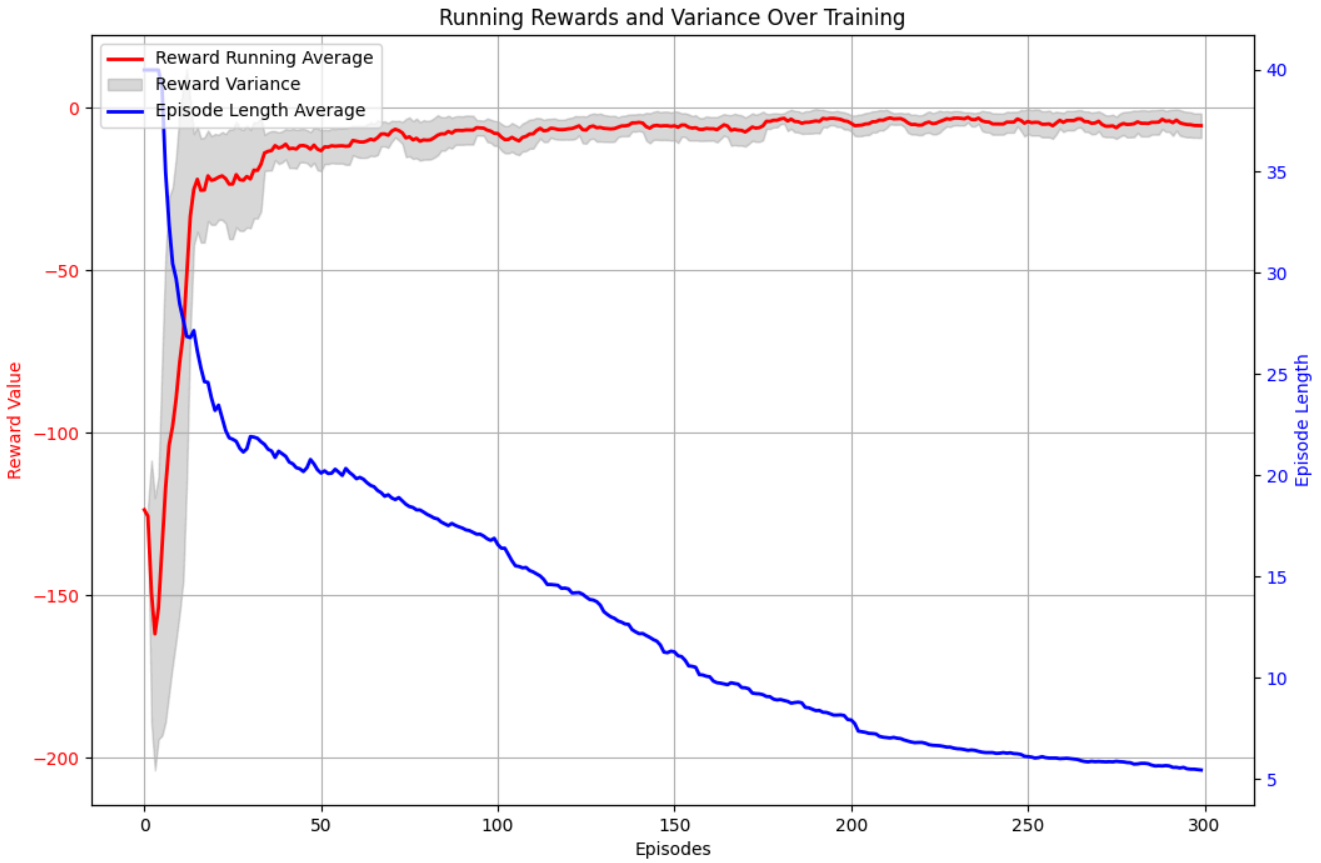

Reward Function Visualization
Branch
Z coordinate
X coordinate
Reward
6
5
4
3
2
1
0
−3
−2
−1
0
1
2
3
0
−1
−2
−3
−4
−5

Trajectories



Comparison between SAC and SACfD




Running Rewards and Variance Over Training
Reward Running Average
Reward Variance
Episode Length Average
Reward Value
Episode Length
Episodes
0
−50
−100
−150
−200
40
35
30
25
20
15
10
5
0
50
100
150
200
250
300

Next Steps

- Further Statistics on the training
  - Crashes.
- Smoothness
  - Already a smoothness term which can be seen in easy tasks.
  - Introduce Prioritised Experience Replay to help combat the harder learning portions.
  - Sampling learnable parameter.
- Move onto next stage:
  - Wrapping

Questions

- For the paper:
  - How would you suggest going about cutting down the background/related work section?