Navigating a UAV With Deep Reinforcement Learning

John Lacalamita and Nagham Sabbour, Mentored by Ping Wang and Saeed Khodaparast

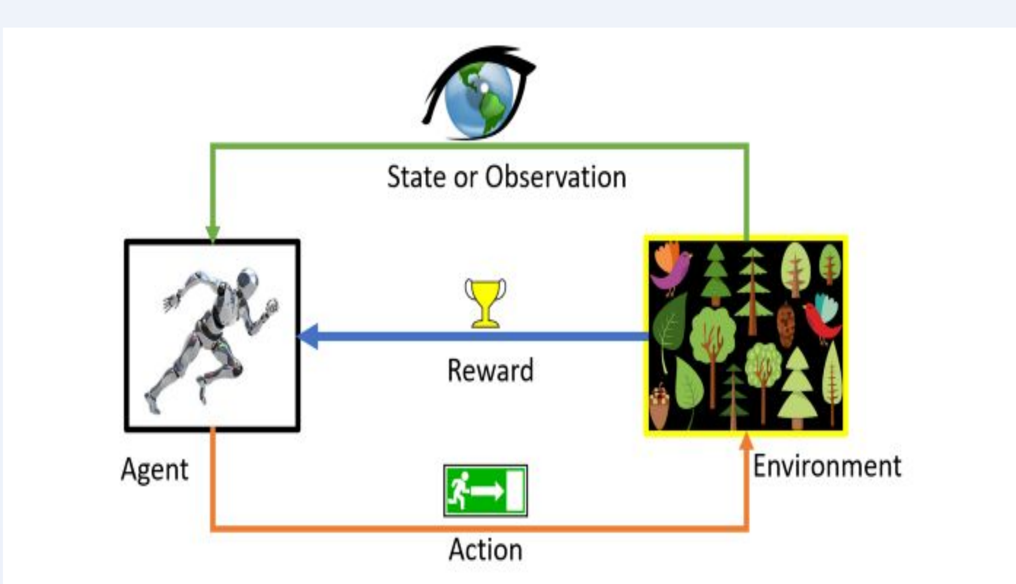
Lassonde School of Engineering

Abstract

To train an unmanned aerial vehicle (UAV) to traverse through a 3D environment in search of a target using deep reinforcement learning. The purpose of this project is meant as an introduction to programming a machine learning model, and the conclusions of the project demonstrate the effectiveness of machine learning and potential it has in many facets of the world.

Background Information

Reinforcement learning is a subset of machine learning in which an agent predicts the best possible decision for an environment to make at any given state. If the decision made by the agent produces a positive result, the environment sends the agent a positive reward. If the agent produces a negative result, the environment will send the agent a negative reward. After many trials, the agent will develop a policy telling the environment the best decision at any given time.



Source:
https://towardsdatascience.com/crystal-clear-reinforcement-learning-7e6c1541365e

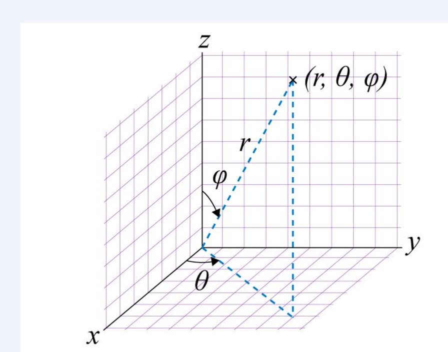
Methods

The project is written in Python 3 and makes use of the machine learning library TensorFlow 2

TensorFlow

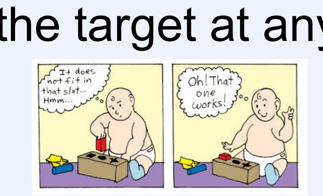
At the beginning of every trial, a random starting spot is given to the UAV, and a target is generated as well as 5 obstacles. The Environment is built around the spherical coordinate system. The X, Y, and Z axis of the UAV is determined by the following formula:
x = x + (speed * sin(theta) * cos(phi))
y = y + (speed * sin(theta) * sin(phi))
z = z + (speed * cos(theta))

Spherical Coordinate System



source: https://byjus.com/maths/spherical-coordinates/

The speed, theta, and phi values are determined by the machine learning agent after every step taken by the UAV. The UAV takes 50 steps per trial and there are 1000 trials in total. If the project works as planned, the agent should create a policy allowing the UAV to reach the target at any given starting position



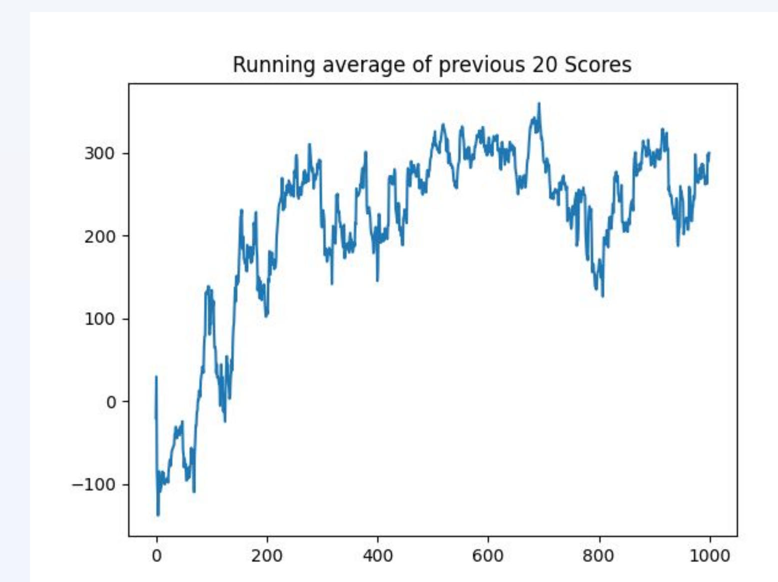
Source:
https://sites.google.com/site/acrader4th2014/home/part-3-unity-and-diversity-of-life/trialanderrorlearning

Reward

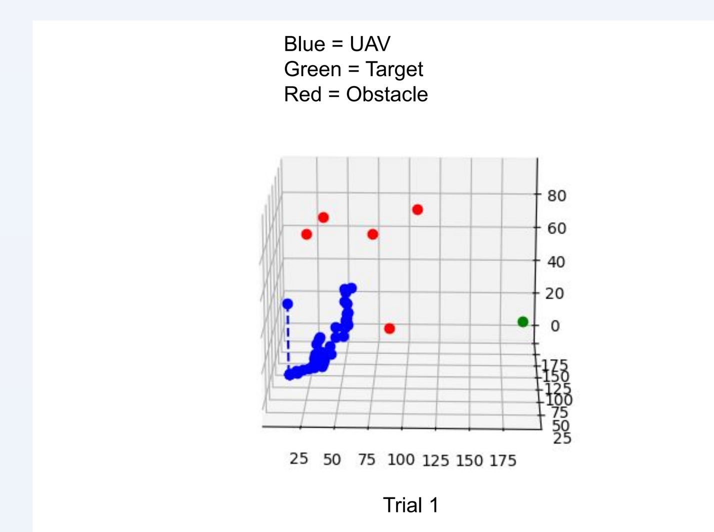
The reward is defined by the following formula:
reward = trans_reward + obs_penalty + finish_reward
Trans_reward calculates the distance after the new step subtracted from the old step in relation to the sensor and multiplies the value (positive or negative) by 2.5. Obs_penalty penalizes the agent by a value of 20 if the UAV hits an obstacle. Finish_reward rewards the agent with a value of 50 if it reaches the target.

Results

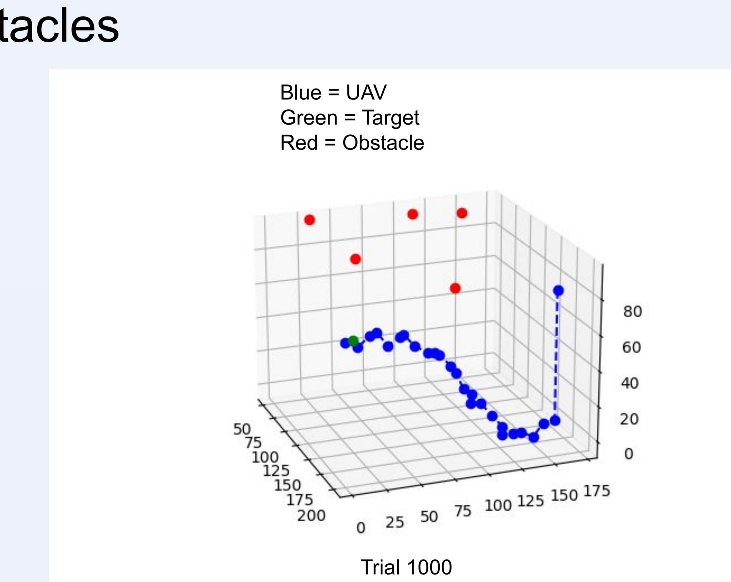
In the beginning, the UAV takes imprecise steps and is not able to reach the target. After about 250 trials, the UAV learns to reach the target and avoid obstacles majority of the time. After this point, the progress plateaus and the UAV no longer makes any positive learning progress.



In the first trial, the UAV makes poor decisions and does not come anywhere close to reaching the target



By the 1000th trial, the UAV can now seamlessly reach the target and avoid any obstacles



Testing the project more times generates very similar results.

Conclusion

- the UAV was successfully able to teach itself to reach the target with deep reinforcement learning.
- Not 100% success, but UAV reached the target majority of the time
- Plateau at around trial 250 demonstrates that eventually the agent will stop learning
- The lack of 100% success even after 1000 trials demonstrates that there is a margin of error in this project
- Whether or not it is possible to optimize the UAV to reach the target 100% of the time remains unanswered
- This project does not consider the power consumption of the UAV, which may alter the conclusion if implemented

Future Directions

The UAV teaching itself to reach the target is just one example of the effectiveness of machine learning. Self-driving cars are an example that utilizes similar concepts to this project. The strongest chess engine in the world, AlphaZero, was created using reinforcement learning and is significantly stronger than any human player. Machine learning is one of the most important fields of computer science and has potential to revolutionize the world.

source:
https://www.greenbiz.com/article/7-companies-steering-self-driving-car-craze

Source:
https://www.rand.org/blog/2020/01/what-chess-can-teach-us-about-the-future-of-ai-and.htm
l

References

https://towardsdatascience.com/crystal-clear-reinforcement-learning-7e6c1541365e

https://tutorial.math.lamar.edu/classes/calciii/SphericalCoords.aspx

http://videolectures.net/rldm2015_silver_reinforcement_learning/

https://www.tensorflow.org/tutorials/quickstart/beginner

https://numpy.org/

https://matplotlib.org/

YORK UNIVERSITÉ UNIVERSITY

LASSONDE SCHOOL OF ENGINEERING

STEM Fellowship