HW2

Alejandro Escontrela

## Experiment 1

Chart

Description automatically generated

Chart

Description automatically generated

* Reward to go performs significantly better
* Normalizing advantages leads to more stable learning
* Batch size reduced the amount of training steps required to achieve the max reward

## Experiment 2

Graphical user interface, chart, histogram

Description automatically generated

## Experiment 3

Chart, line chart

Description automatically generated

## Experiment 4

Chart

Description automatically generated

A screenshot of a computer

Description automatically generated with medium confidence

## Experiment 5

Chart, histogram

Description automatically generated