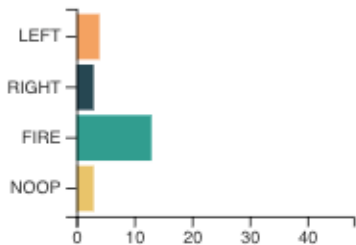


Episode

Selected Episode

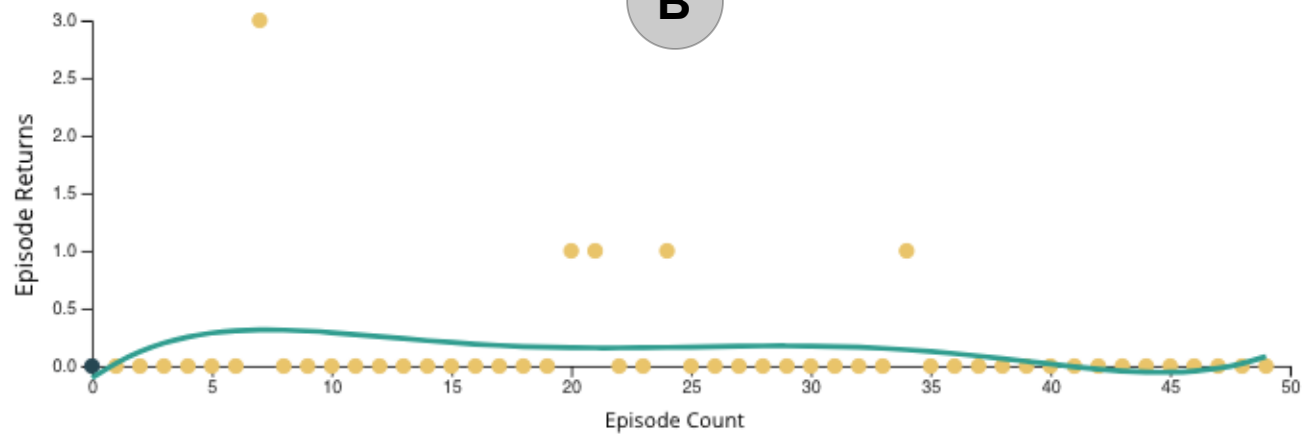
0

Action Distribution

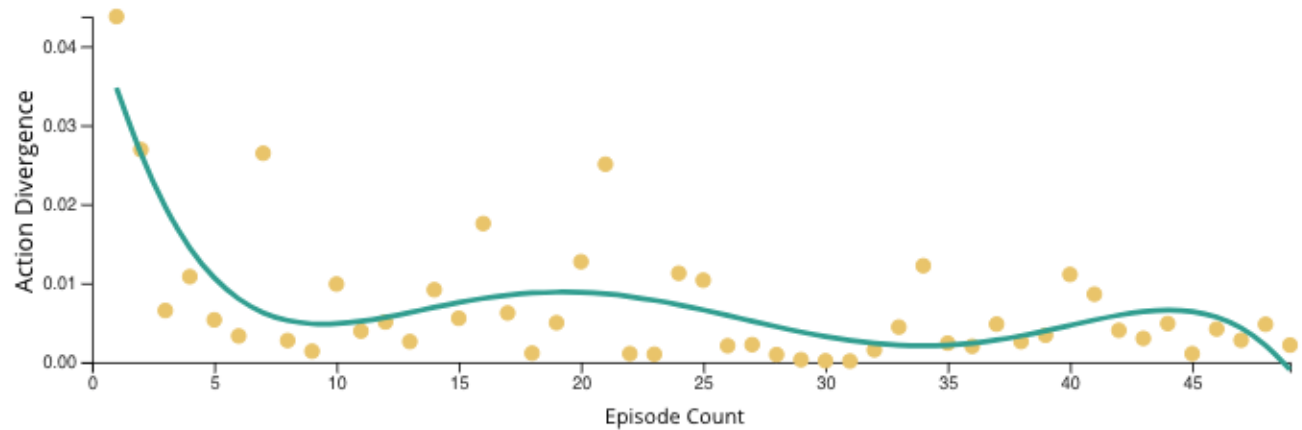


+

Episode Returns



Action Divergence



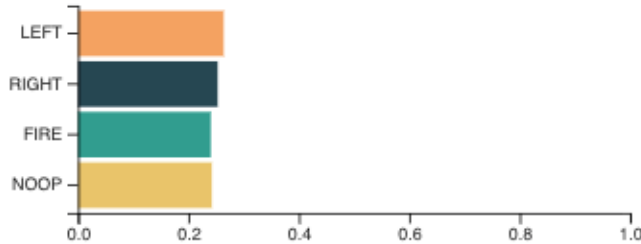
+

Timestep

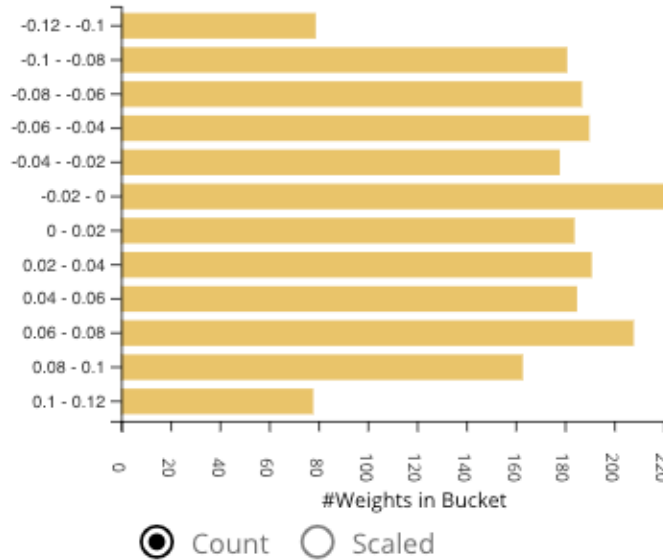
Selected Timestep

0

Action Probabilities



Weight Histogram



Reward

]

D



Timestep

PLAY

X1