UF | Herbert Wertheim College of Engineering
UNIVERSITY of FLORIDA

Coordination of Distributed Agents through Stochastic Policies in a Cooperative Jamming Scenario

May 30, 2024

Robert Thomas Lattus
PhD Student, WING Lab

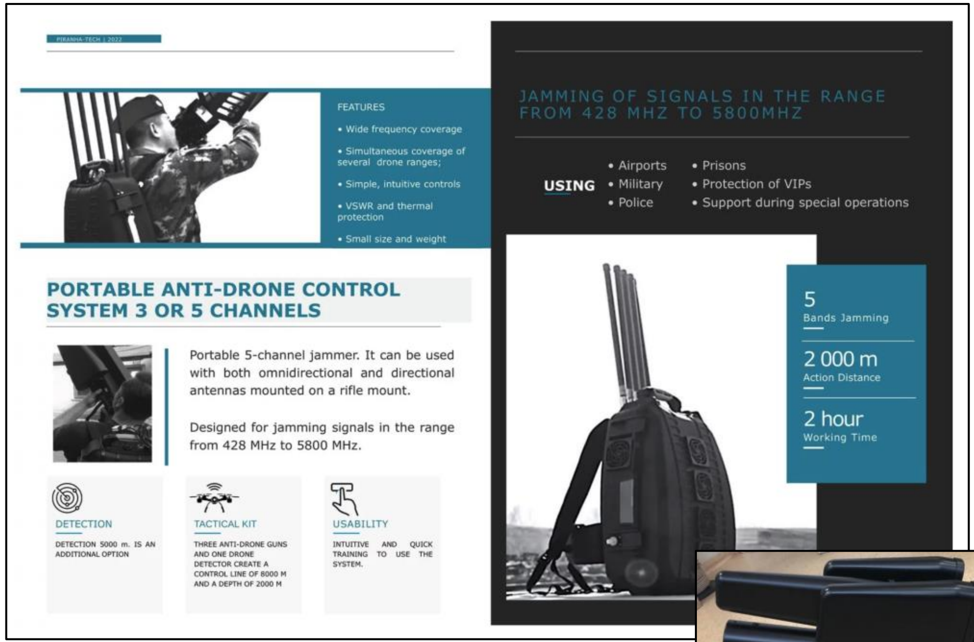
MOTIVATION

How Russia Is Trying To Complicate Ukraine's Drone Defenses

Published Mar 07, 2024 at 11:33 AM EST | Updated Mar 07, 2024 at 5:21 PM EST

The Invisible War in Ukraine Being Fought Over Radio Waves

How drone combat in Ukraine is changing warfare

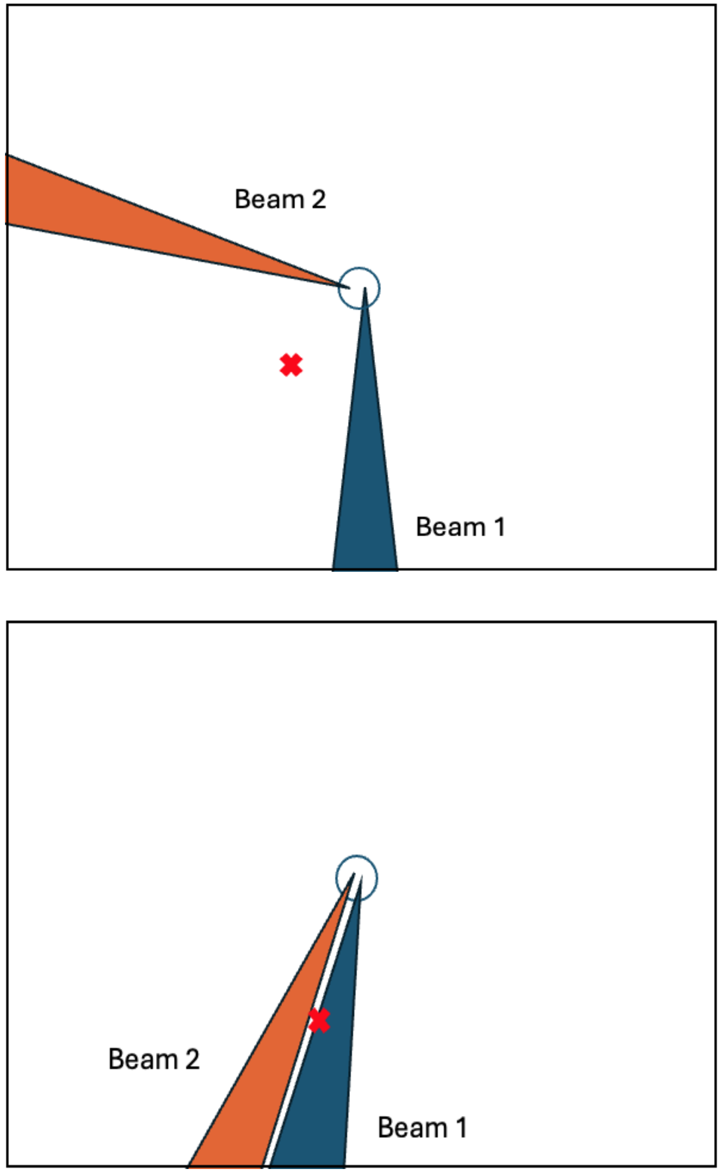


https://piranha-tech.net

How can agents coordinate their actions without direct communication in a cooperative jamming scenario?

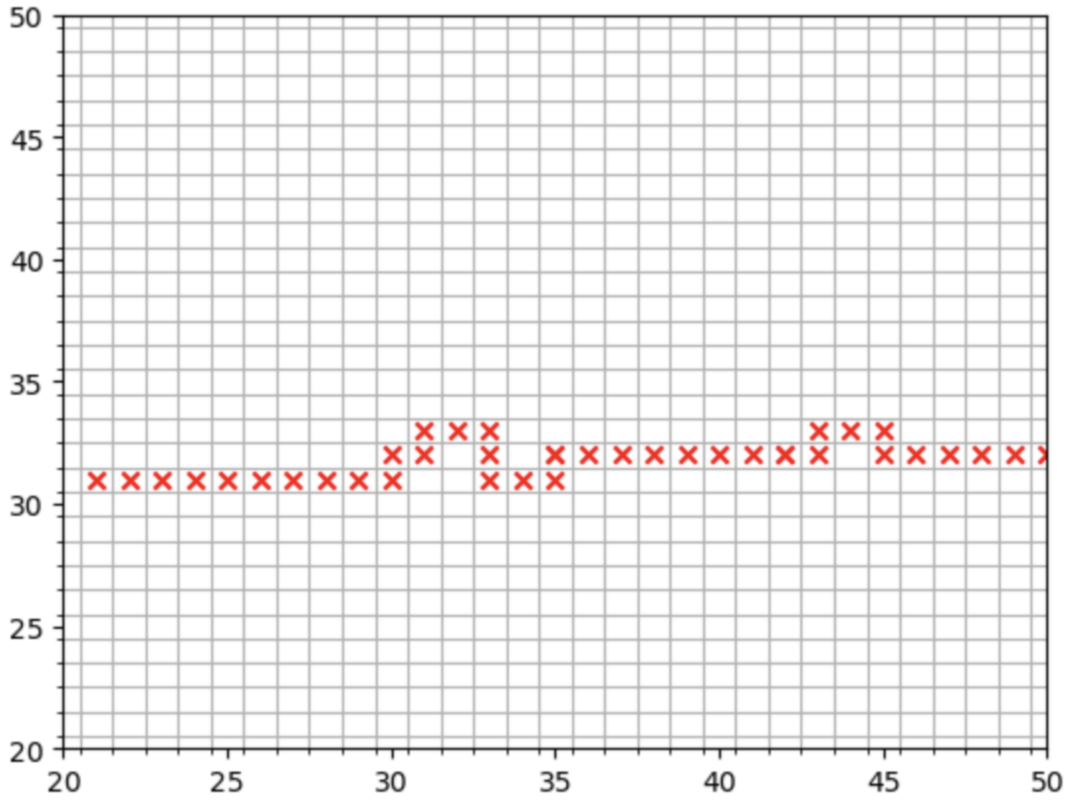
SYSTEM MODEL — JAMMING AGENTS

- Multiple agents at fixed locations in a region they are tasked to protect
- Each agent has a fixed-beam directional antenna modelled as a single lobe of width $B°$
- Fixed beamwidth with main lobe
- If adversary is in main lobe from any agent, it is jammed
- Consider discrete angular positions at each antenna dividing covered area into $\frac{360°}{B°}$ subsections
- Each agent learns resulting adversary position and beam position of fellow agents after choosing their own action



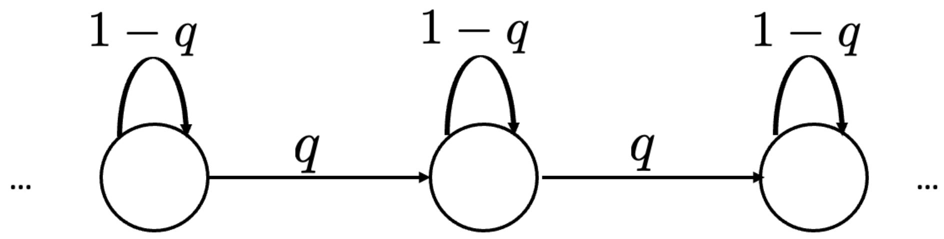
SYSTEM MODEL — ADVERSARY

- Consider movement on gridded space with discrete actions: up, down, left, right, stationary
- Momentum parameter, μ, equal to probability agent repeats previous action
- Smaller momentum probability corresponds to more evasive maneuvers/stochasticity
- General pattern of left-to-right and then right-to-left

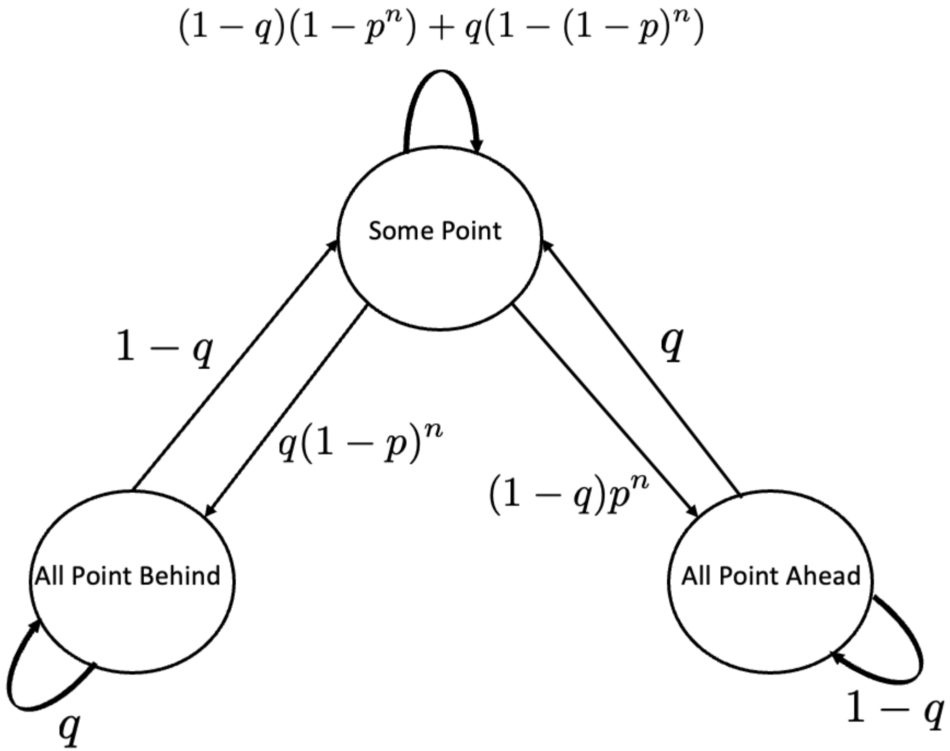


MOTIVATING ANALYSIS — ONE-D INFINITE MARKOV MOTION MODEL

- Consider simplified 1-D, unidirectional motion model on infinite quantized line
- Adversary moves with probability q, remains with probability 1-q
- Assume agents begin by pointing at adversary
- Let agents point at the next state with probability p, or continue pointing at current state with $1 - p$
- Agents have knowledge of q
- If there are N-1 agents, choose p to **maximize probability of at least one agent pointing at adversary**



MOTIVATING ANALYSIS — MARKOV CHAIN MODEL FOR AGENTS' STATE

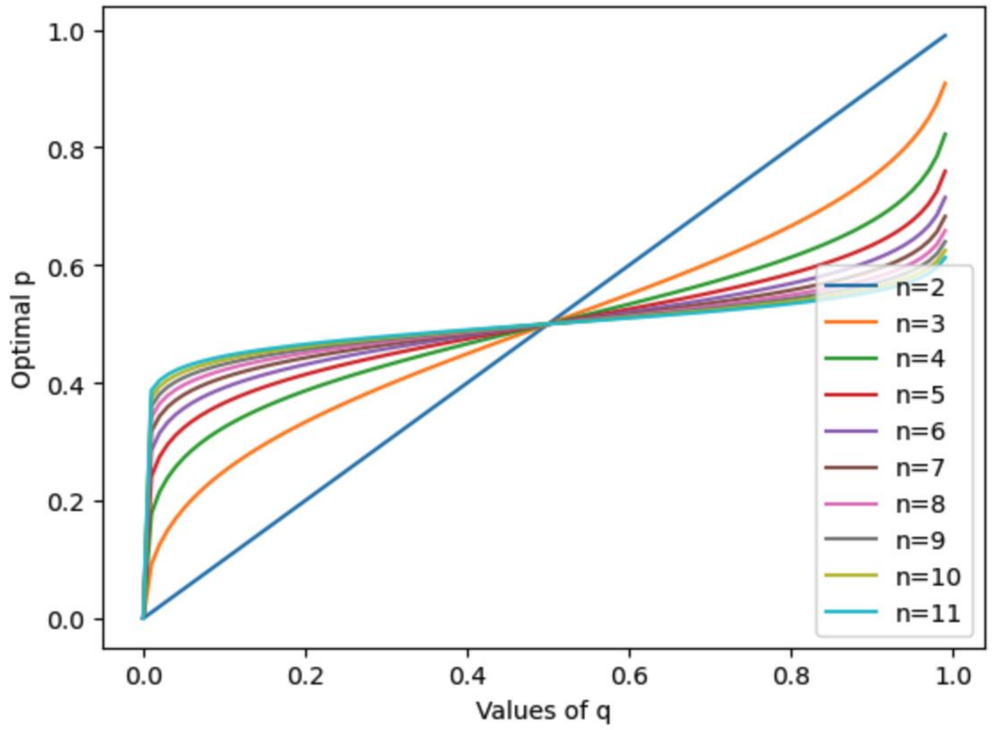


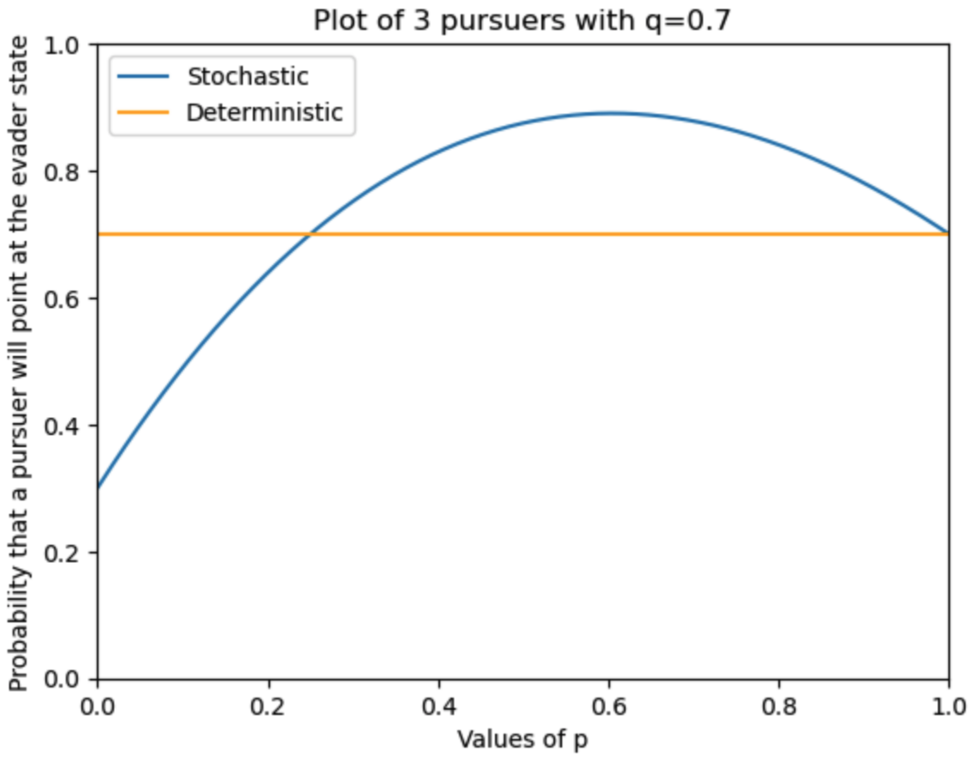
- Can model state of agents' jamming as simple finite-state Markov chain
- If an agent points behind the adversary, it deterministically moves to the next position to "catch up"
- If an agent points ahead of the adversary, it deterministically remains stationary to wait for the adversary to "catch up"

- Solve for the optimal p in terms of q and n

$$p_{opt} = \frac{\frac{q}{1-q}^{\frac{1}{n-1}}}{1 - \frac{q}{1-q}^{\frac{1}{n-1}}}$$

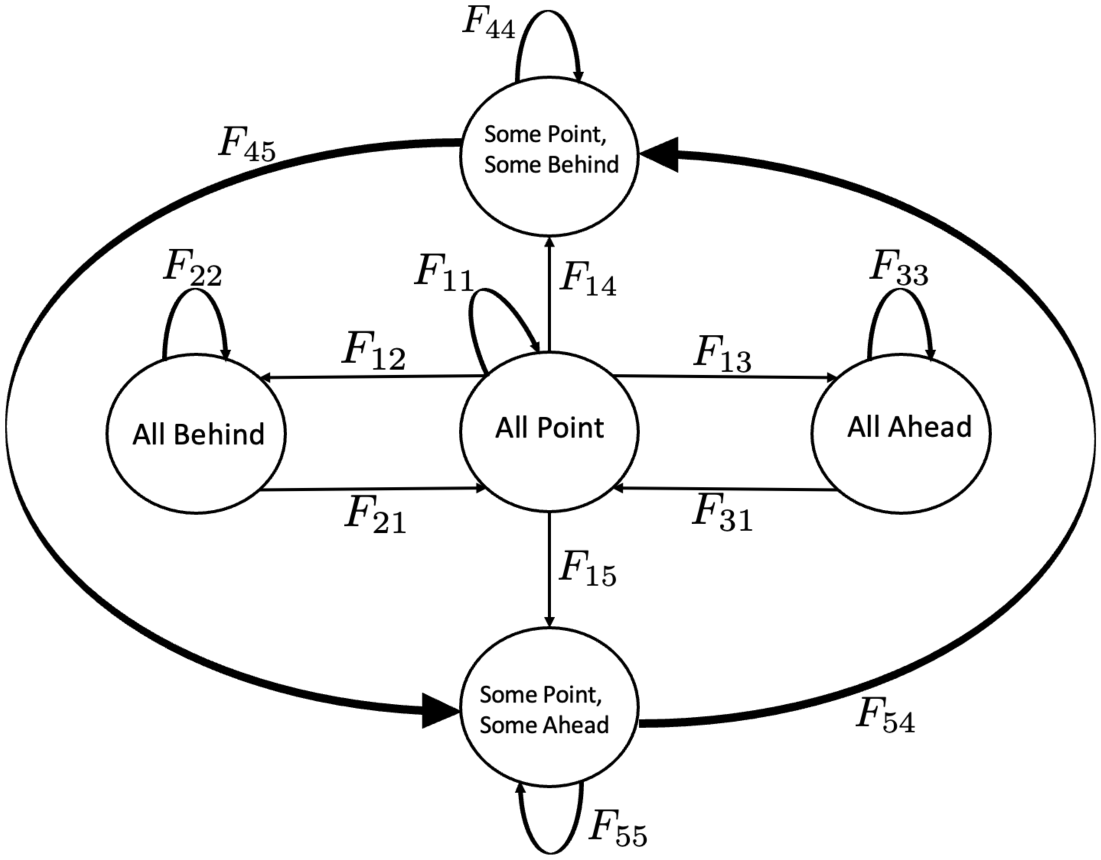
- Note that as the number of agents rises, p becomes more random to eliminate redundancy





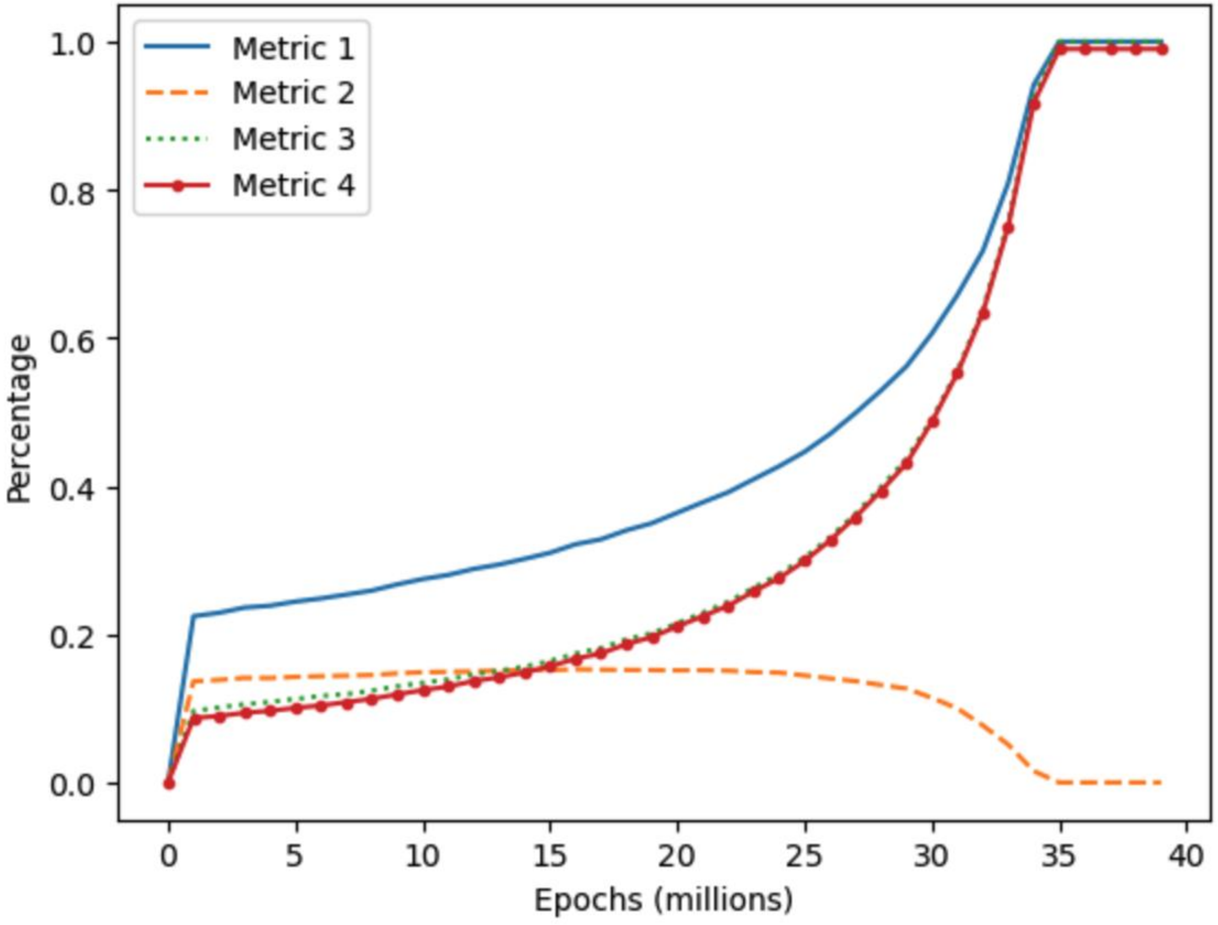
MOTIVATING ANALYSIS — EXPANSION OF MARKOV CHAIN

- Previous Markov chain can be expanded to reveal recurrent subchain (outer states)
  - In this subchain, at least one agent is always pointing at the adversary, resulting in optimum performance
- Reaching recurrent states requires stochastic policy
- Deterministic agents make same decision & thus can never reach this optimal configuration
- Note that in this recurrent subchain, agents may still switch their roles (which ones advance where they point and which ones are stationary)

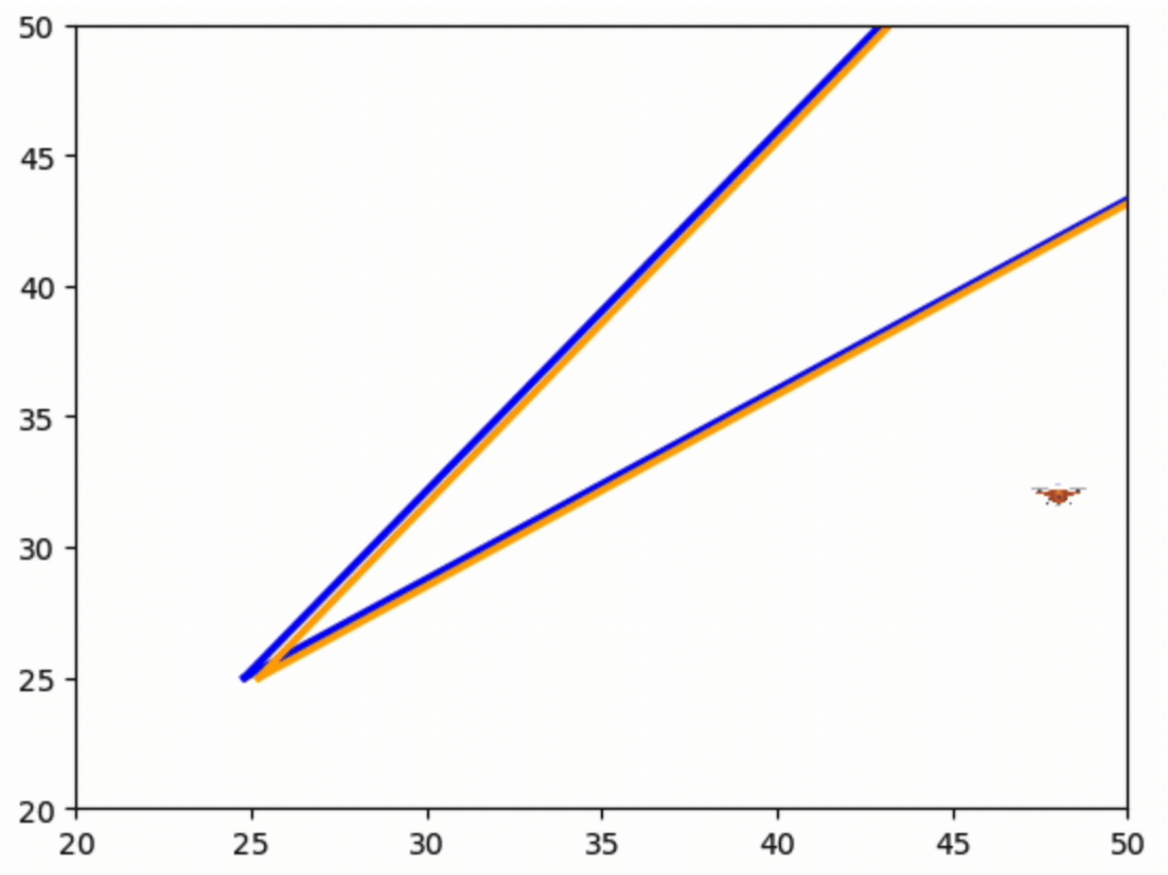


SIMULATION—FINITE MARKOV CHAIN

- Simulation is completed of the expanded Markov chain
- Reinforcement learning to see if the agents could learn optimal policy
- Four metrics (%):
  1) An agent points at the adversary
  2) The agents are in the same state
  3) A deterministic policy is used
  4) The agents are in the absorbing states

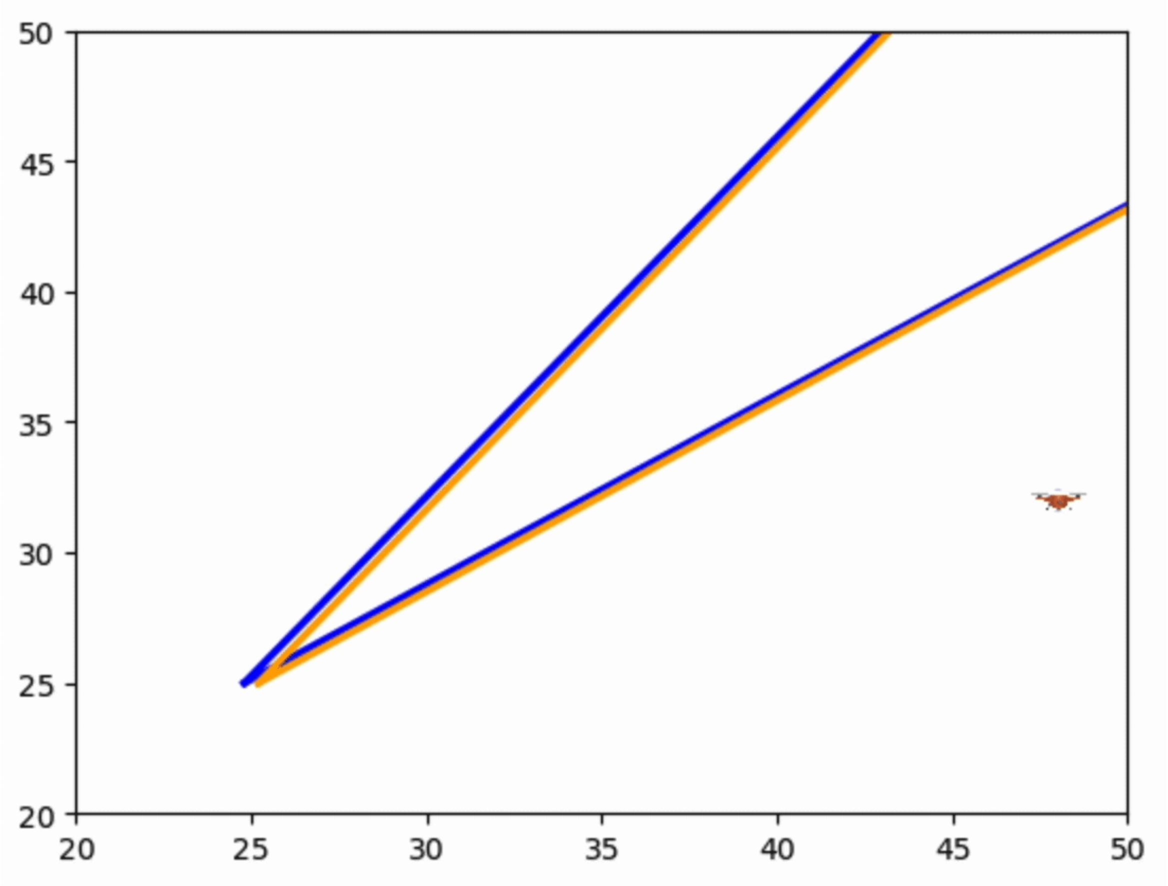


TRAINED Q-TABLE EXAMPLE



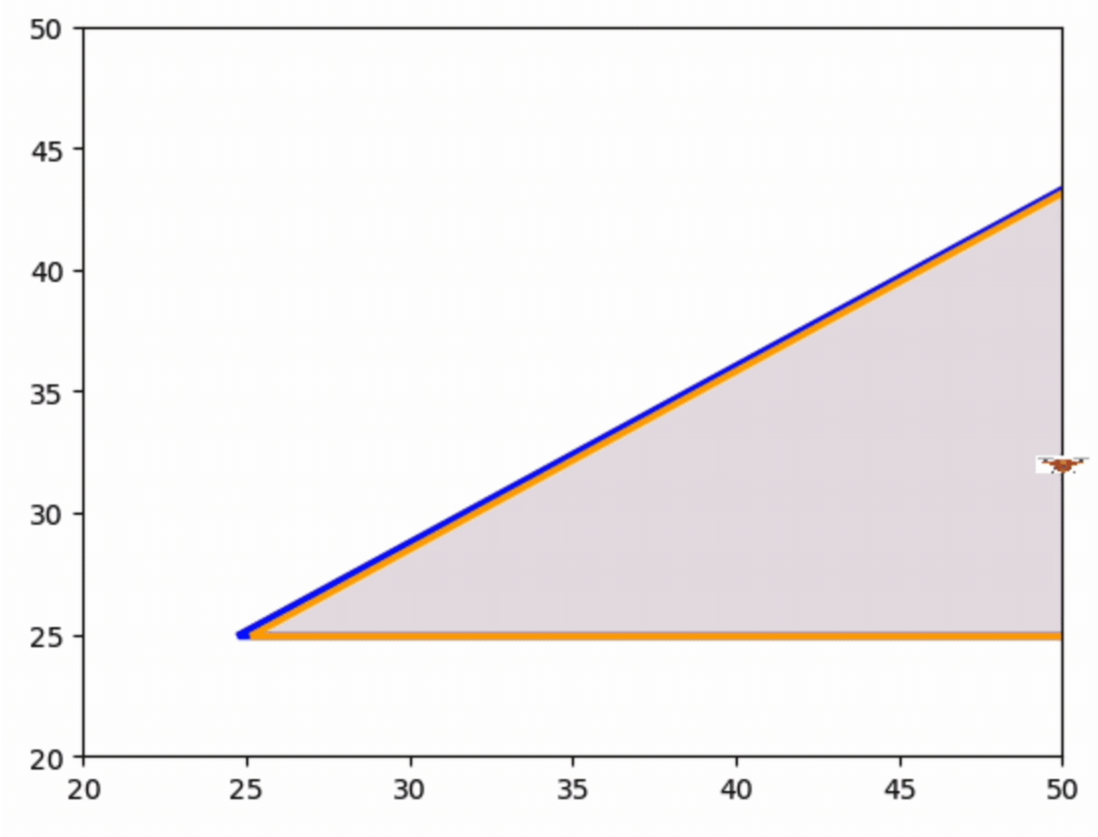
Deterministic Cooperation

TRAINED Q-TABLE EXAMPLE



Stochastic Cooperation

TRAINED Q-TABLE EXAMPLE



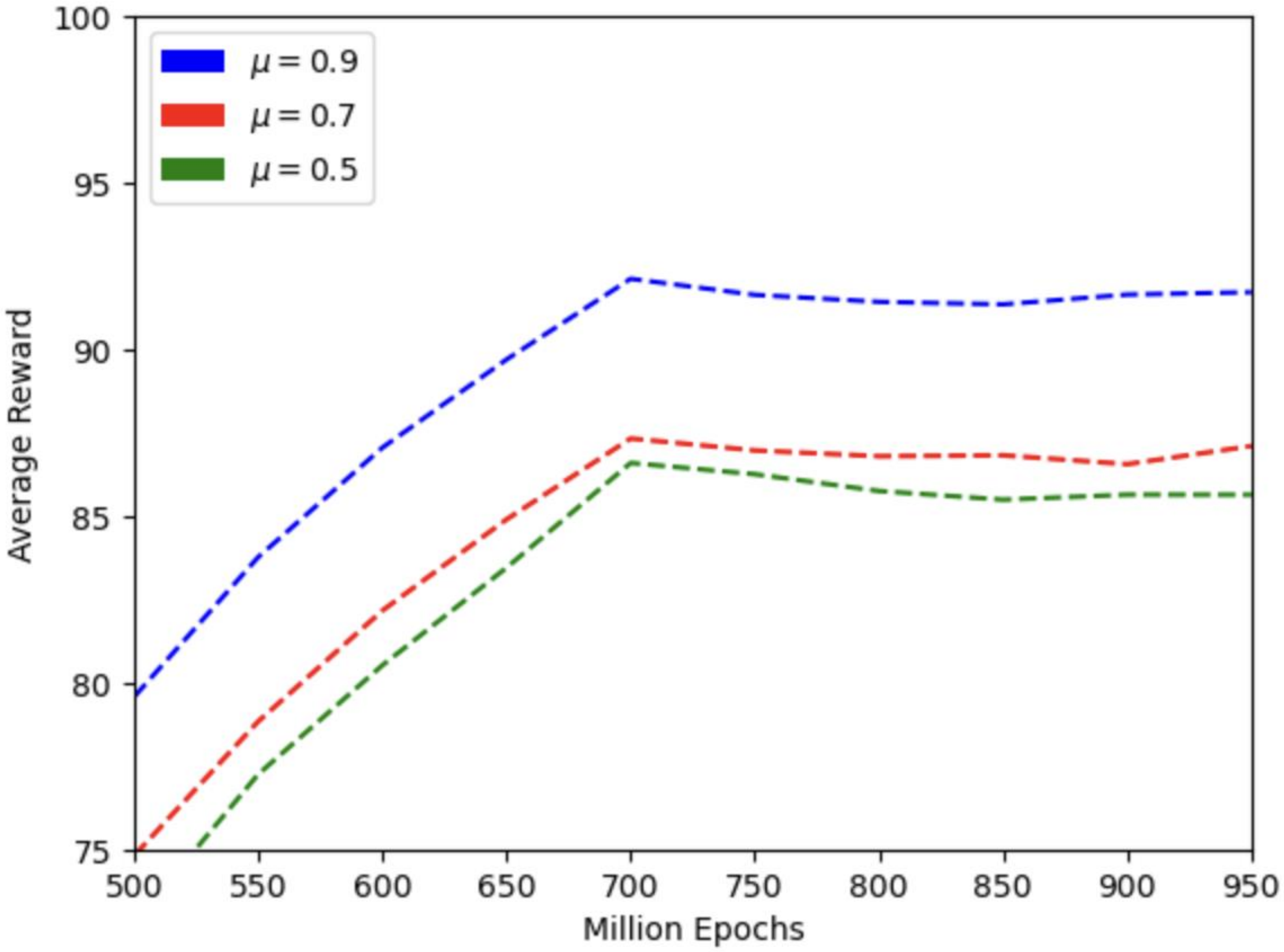
Stochastic Cooperation with Two Antenna Beamwidths

SIMULATION—SYSTEM MODEL

- Two agents adjacently placed near the center of the environment
- Beamwidths of $36°$ with 10 distinct radial positions
- Two action space strategies:
  - Three deterministic-only pmfs (e.g [1.0,0.0,0.0])
  - Thirteen stochastic and deterministic pmfs (e.g. [0.7,0.3,0.0])
- Simulation run for three values of adversary momentum:
  $\mu = 0.9, \mu = 0.7, \mu = 0.5$

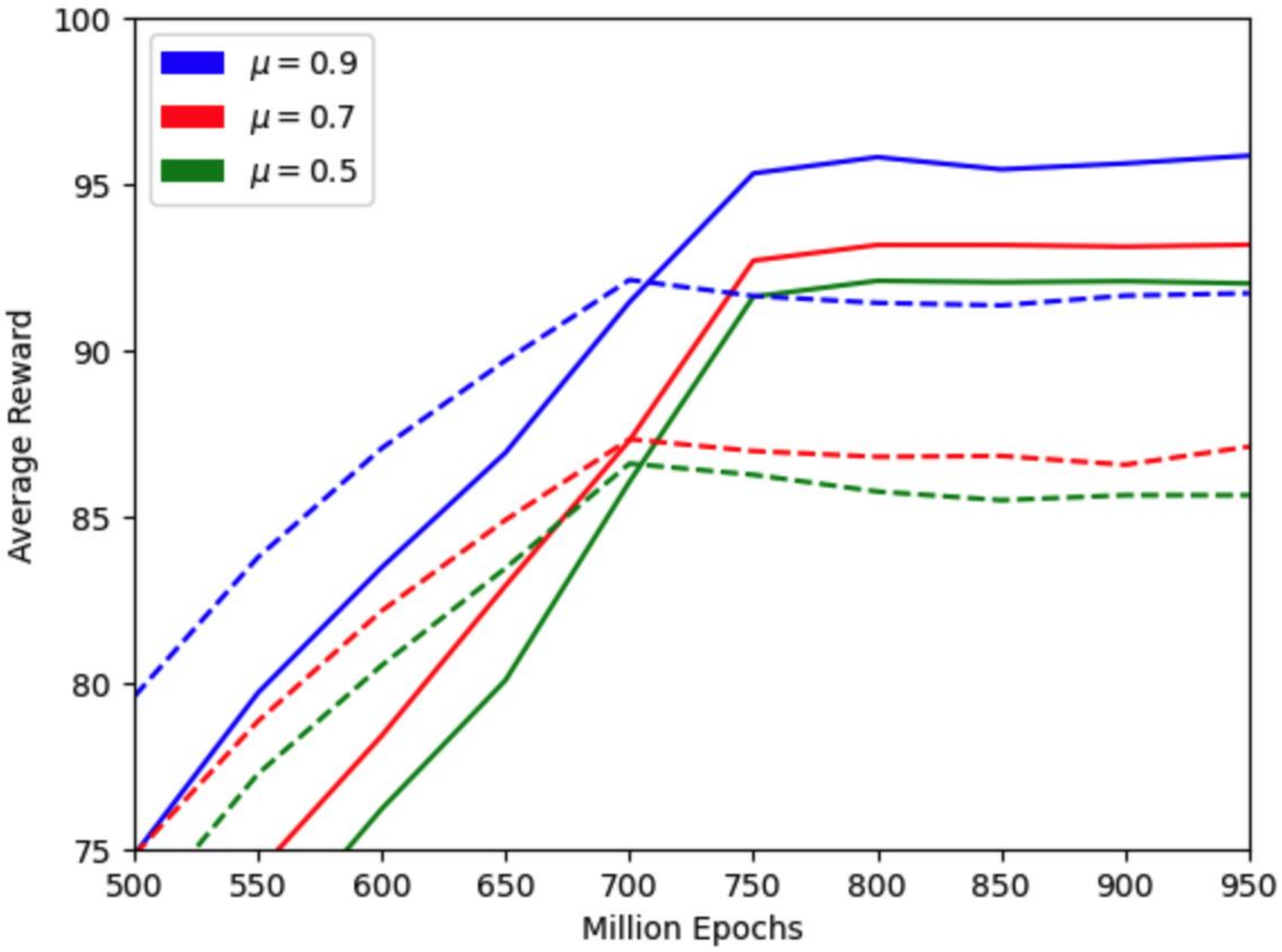
SIMULATION—RESULTS

- Consider deterministic policies first



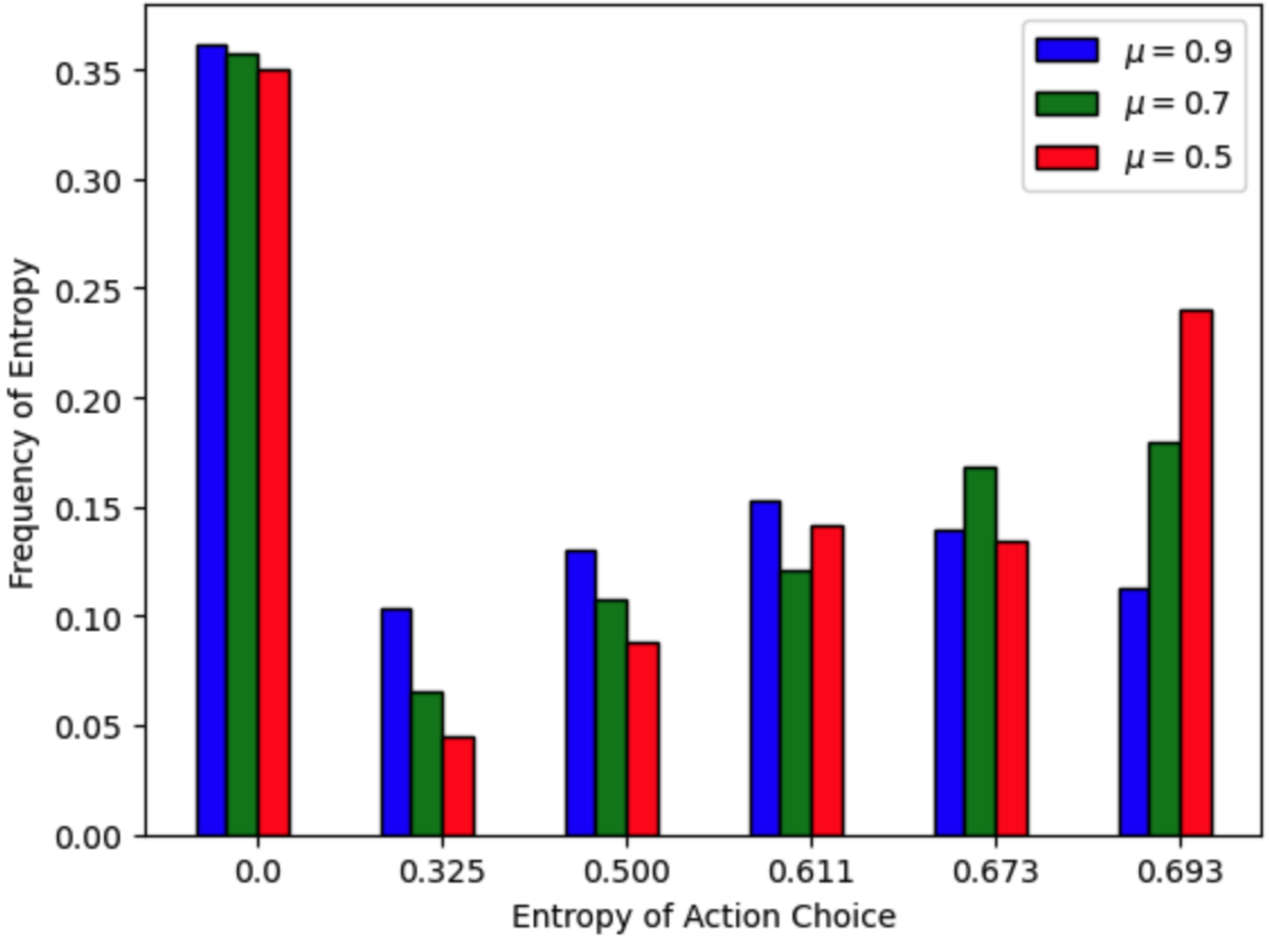
SIMULATION—RESULTS

- Now allow stochastic policies (solid lines):

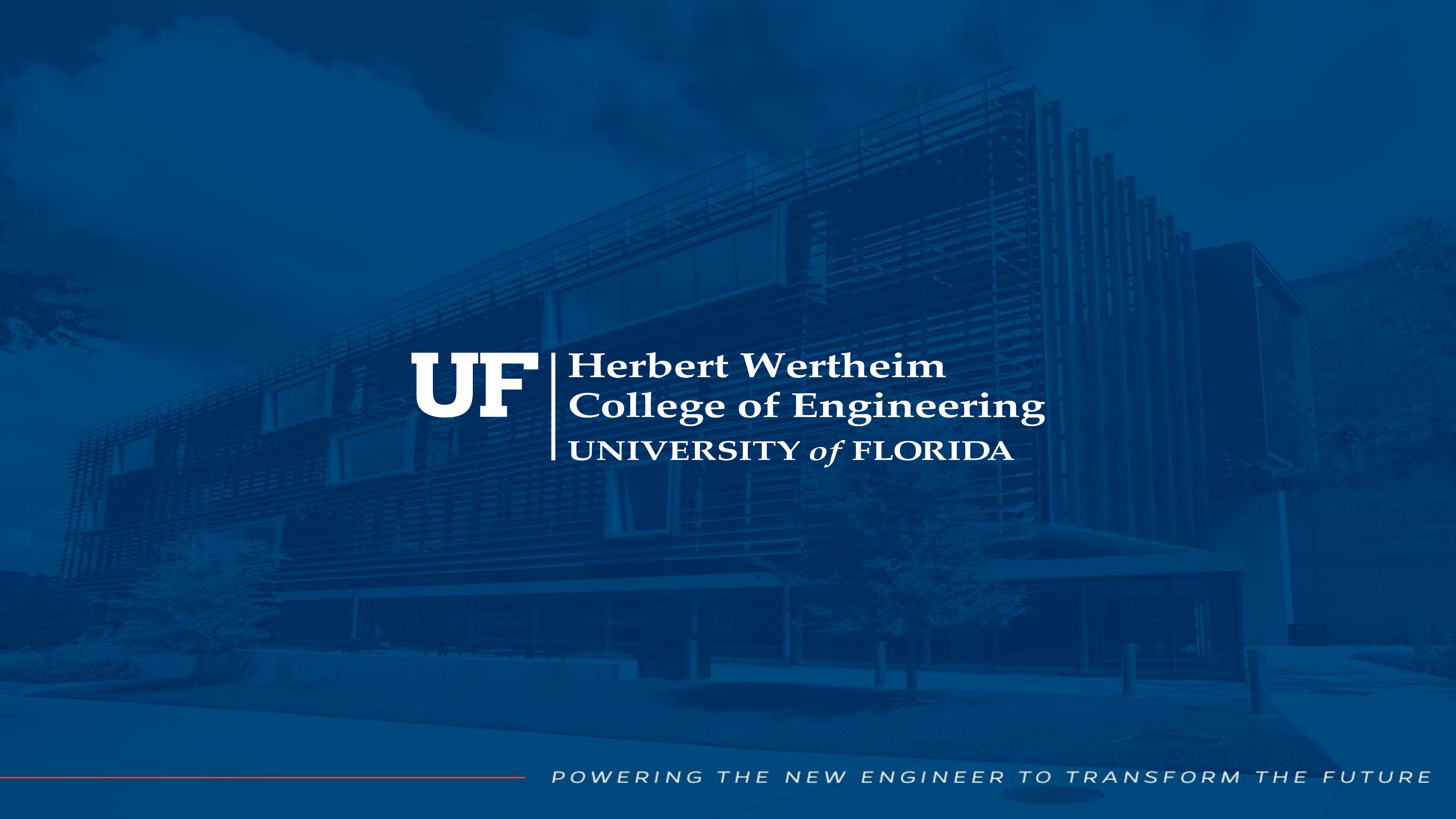


SIMULATION - ENTROPY

- States where beams are in overlap and adversary momentum was leaving were collected to analyze trained decisions



Momentum	Entropy (Nats)
$\mu = 0.9$	0.363820
$\mu = 0.7$	0.387328
$\mu = 0.5$	0.402358



UF Herbert Wertheim College of Engineering
UNIVERSITY *of* FLORIDA

POWERING THE NEW ENGINEER TO TRANSFORM THE FUTURE